



HAL
open science

Développement d'une approche SIG pour l'intégration de données Terre/Mer

Ariane Mascret

► **To cite this version:**

Ariane Mascret. Développement d'une approche SIG pour l'intégration de données Terre/Mer. Informatique [cs]. Ecole nationale supérieure d'arts et métiers - ENSAM, 2010. Français. NNT : . tel-00492953

HAL Id: tel-00492953

<https://theses.hal.science/tel-00492953>

Submitted on 17 Jun 2010

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.

École doctorale n°432 : Sciences des Métiers de l'Ingénieur

Doctorat ParisTech

T H È S E

pour obtenir le grade de docteur délivré par

l'École Nationale Supérieure d'Arts et Métiers

Spécialité " Informatique "

présentée et soutenue publiquement par

Ariane MASCRET

15 Mars 2010

**Développement d'une approche SIG pour l'intégration
de données Terre/Mer**

Directeur de thèse : **Christophe CLARAMUNT**
Co-encadrement de la thèse : **Thomas DEVOGELE**

Jury

Mme. Christiane WEBER, Professeur, Image, Ville et Environnement, Strasbourg
Mme. Anne RUAS, Ingénieur HDR, COGIT, IGN Paris
M. Christophe DELACOURT, Professeur, UBO/IUEM, Brest
M. Marc ROBIN, Professeur, LETG, Nantes
M. Christophe CLARAMUNT, Professeur, IRENav, Ecole Navale, Brest
M. Thomas DEVOGELE, Professeur, IRENav, Ecole Navale, Brest

Rapporteur
Rapporteur
Examineur
Examineur
Examineur
Examineur

**T
H
È
S
E**

REMERCIEMENTS

Au crépuscule de la rédaction de ce manuscrit, j'aimerais remercier les personnes qui ont participé à ce projet de longue haleine et qui ont permis sa réalisation.

Merci à Christophe Claramunt tout d'abord pour m'avoir proposé ce sujet de thèse alors que je venais postuler sur une tout autre thématique, ainsi qu'à Thomas Devogele pour avoir accepté de m'encadrer alors qu'il ne me connaissait pas. Je leur suis reconnaissante d'avoir osé se lancer dans l'aventure en n'ayant de moi que l'image d'une « survoltée » capable de faire une présentation de soixante slides en dix minutes chrono.

Merci également au jury de thèse d'avoir accepté d'évaluer ce travail. Certains me connaissaient pour avoir longuement parlé avec moi de mon sujet de recherche et m'avaient mise en garde contre certains points très délicats de l'approche proposée. D'autres n'avaient jamais entendu parlé de mon travail ni des algorithmes sur lesquels il se base mais ont décidé de participer quand même à son évaluation : merci pour leurs retours et la pertinence de leurs commentaires.

Merci à tous les collaborateurs et intervenants qui ont participé directement ou indirectement à ce travail : la société Seazone qui a travaillé sur la fusion de données et m'a gracieusement fait part de ses résultats, le SHOM et l'IGN qui entreprenaient le programme Litto3D et nous ont permis l'accès aux données, le laboratoire Géomer pour avoir collaboré avec nous et participé à la relecture d'un de mes chapitres et enfin Jacques Populus pour nous avoir fourni des données lidar haute définition.

Par ailleurs une thèse est toujours supportée par ce que j'appelle affectueusement le « staff technique » : informaticiens, professeurs d'anglais/de math, secrétaires et tous les gens de l'école (officiers militaires en particulier) qui ont pris le temps de s'intéresser à mon travail ou m'ont aidée dans mes démarches. Je tiens d'eux certaines ficelles administratives, la recette d'un pain de Noël aux épices (merci Mr Marion) et la conception de fléchettes en papier pour sarbacanes.

Ensuite... tout ne se passe pas toujours comme prévu, et malgré tout le meilleur optimisme du monde, parfois trop c'est décidément trop : c'est là que les proches prennent le relais. J'aurais donc des pensées reconnaissantes envers tous ceux qui ne m'ont pas laissée baisser les bras quand j'en avais vraiment raz le bol. Hé oui, c'est souvent à la fin où il faudrait garder le plus de détermination, que Motivation se fait joyeusement la malle avec Persévérance et Sérieux. Le moral en prenant un bon coup dans l'aile, la tentation est grande de laisser en plan le boulot quasi fini.

Que nenni donc, mes proches m'ont affrontée moi et ma mauvaise volonté :

- Mon Cher et Tendre en relisant et corrigeant un boulot qui n'était pas de lui m'a remis sur les rails désertés de la rédaction.

- Mathieu qui par son flegme et sa détermination sans faille à écrire ses deux pages par jour a réussi à me redonner envie de relever le pari.

- Marie Laure, qui supportant mes jérémiades et pistant mes nombreux moments d'inaction, de flemme, et de vagabondage m'a « forcée » à terminer malgré moi.

- Mériam pour son côté réaliste : la rédac c'est long, pénible et loin d'être grisant (saoulant à la rigueur et c'est pourtant un synonyme) !! Ceux qui vous soutiennent le contraire sont aux choix des génies ou des cabotins. C'est par elle aussi que j'ai appris tout le sens du mot « procrastination » : fait de découvrir que tout ce que vous aviez laissé en plan depuis des siècles est décidément beaucoup plus passionnant que rédiger. C'est ainsi que j'ai exhumé le cycle complet de « Fondation » d'Isaac Asimov, les chroniques d'Alvin le Faiseur suivies du cycle d'Ender d'Orson Scott Card, l'intégrale de la compagnie noire de Glen Cook, Terry

Pratchett et ses Annales du disque-monde et j'en passe (merci au passage à Thierry et Antoine pour m'avoir fait découvrir ces auteurs). L'apport culturel a beau être indéniable, les vingt malheureuses pages de thèse restantes à rédiger n'ont pas avancé d'un chouia...

- Eric, compagnon d'infortune de rédaction qui en bavait en même temps que moi pour terminer au plus vite.

- Calixte dans son île lointaine qui m'a soutenue et répété sans cesse « TERMINE TA THESE, on verra ça APRES !! ».

- Marie pour avoir relu et corrigé ma thèse de ses fautes d'orthographe.

Il y a encore bien d'autres personnes parmi mes amis et ma famille mais je ne puis malheureusement citer tout le monde, sous peine d'avoir des remerciements aussi éléphantesques que le reste de la thèse (déjà deux pages, ça commence à faire). Je pense néanmoins à eux et les remercie d'avoir été là tout au long du périple.

Bien, mais et après ? Beaucoup m'ont parlé d'un hypothétique vide ressenti après la thèse tout ça... Bon, soyons honnête, pas vraiment ! Faisons le point : il y a de bons films à voir, d'excellents livres à lire (même si je n'en dévorerais probablement pas autant que ce que j'ai déjà fait), des tas de projets plus fous les uns que les autres restés en suspens et un post-doc (suivi temporel des dunes sous-marines pour les curieux) à faire courir jusqu'en décembre 2010 : la routine reprend son cours...

TABLE DES MATIERES

<i>Rémerciements</i>	1
<i>Table des Matières</i>	3
<i>Résumé</i>	7
<i>Contexte de la thèse</i>	9
1 Acteurs et données géographiques littorales	9
2 Continuité et cohérence au sein des données géographiques	10
3 Contexte de recherche	11
4 Démarche et problématique	14
5 Plan de thèse	16
I Littoral et intégration : présentation, état de l'art et discussion	17
I.1 Présentation du milieu littoral	18
I.1.1 Agents morphogéniques des reliefs littoraux	19
I.1.1.1 Actions mécaniques	19
I.1.1.2 Actions physico-chimiques et biologiques	20
I.1.2 Zonalité des formes littorales selon les grands domaines climatiques	21
I.1.2.1 Zones tempérées	21
I.1.2.2 Zones sub-tropicales	21
I.1.2.3 Zones polaires	21
I.2 Intégration d'éléments 2D	22
I.2.1 Définitions	22
I.2.1.1 Dimensions géographiques des objets	22
I.2.1.2 Comparaison entre deux objets ou un ensemble d'objets	22
I.2.1.3 Appariement de deux jeux de données	25
I.2.1.4 Fusion de données.....	26
I.2.2 Intégration de points	27
I.2.2.1 Outils d'appariement de points	27
I.2.2.2 Fusion de points	28
I.2.3 Intégration de lignes	28
I.2.3.1 Outils de comparaison de lignes	28
Distance moyenne	28
Distance de Hausdorff.....	28
Distance de Fréchet et dérivées.....	30
Conclusion sur les distances linéaires	32
Outils de comparaison de formes pour des objets linéaires	32
Outils de comparaison par zone définie	33
I.2.3.2 Outils d'appariement géométrique de lignes	34
I.2.3.3 Fusion et déformation élastique de lignes.....	35
I.2.4 Intégration de polygones	38
I.2.4.1 Comparaison de surfaces	39
Distance entre surfaces.....	39
Comparaison de ressemblance de forme entre surfaces	39

I.2.4.2	Appariement de surfaces.....	39
I.2.4.3	Fusion de surfaces.....	40
I.3	Intégration de Données 2.5D (MNT)	40
I.3.1	Caractéristiques des MNT	40
I.3.1.1	Acquisition et génération d'un MNT.....	40
MNT terrestres : acquisitions et traitements	41	
Modes d'acquisition de MNT maritimes	43	
Synthèse sur les modes de production des MNT	45	
I.3.1.2	Formats des MNT.....	46
Les nuages et grilles de points	46	
Les profils	46	
Les courbes de niveau	47	
Triangular Irregular Network (TIN).....	47	
Grille raster	47	
Méthodes d'interpolation de l'altitude	48	
I.3.1.3	Qualité des MNT	49
Types d'erreurs présentes dans les données	50	
Précision et résolution.....	50	
I.3.2	Fusion de MNT : travaux en cours ou déjà réalisés.....	51
I.3.2.1	Intégration de données Terre/ Mer.....	52
ICZMap.....	52	
Litto 3D.....	54	
MNT intégré sur la Baie du Mont-Saint-Michel	55	
Projet Américain	56	
I.3.2.2	Fusion de MNT terrestres avec recouvrement total	57
I.3.2.3	Bilan des travaux présentés.....	57
I.4	Bilan sur l'état de l'art.....	59
II	Extensions de la distance linéaire de Fréchet et algorithmes d'appariement	61
II.1	Extension 1 : distance moyenne de Fréchet et couples de points homologues..	63
II.2	Extension 2 : distance de Fréchet sur des lignes d'emprises ou de natures différentes.....	67
II.2.1	Distance de Fréchet discrète partielle.....	67
II.2.2	Distance de Fréchet discrète entre deux lignes fermées	68
II.2.3	Distance de Fréchet discrète partielle entre une ligne ouverte et une ligne fermée.....	72
II.2.4	Distance de Fréchet discrète partielle / partielle entre deux lignes ouvertes	73
II.3	Algorithme d'appariement	78
II.3.1	Recherche de la nature des lignes comparées.....	79
II.3.2	Relations d'emprises entre les lignes.....	79
II.3.3	Types d'appariement entre lignes.....	81
II.3.3.1	Appariement simple « 1 – 1 » de deux lignes :.....	81
II.3.3.2	Appariement total/partiel « 1 – P » entre deux lignes	82
II.3.3.3	Appariement « partiel/partiel » (PP) et « deux partiel/partiel » (2PP) entre deux lignes	82
II.3.4	Méthode globale : appariement par fichier.....	86
II.4	Exemple d'applications de la méthode sur la comparaison de traits de côte ...	89
II.4.1	Présentation des données et de la zone d'étude	89
II.4.2	Méthode de comparaison des traits de côte	89
II.4.3	Résultats et perspectives	90
II.5	Bilan sur les apports des extensions.....	93
III	Méthodologie d'intégration.....	95
III.1	Présentation des données	97
III.1.1	Données géographiques acquises	97
III.1.1.1	Description de la zone d'étude.....	97

III.1.1.2	Description des jeux de données.....	100
III.1.2	Evaluation de la cohérence des données géographiques	101
III.1.2.1	Etude des valeurs de différences entre données altimétriques	101
III.1.2.2	Comparaison de la densité d'échantillonnage des données altimétriques.....	104
III.1.3	Données Simulées	107
III.1.3.1	Mécanisme de production des données simulées.....	107
III.1.3.2	Modèles de base.....	108
III.1.3.3	MNT bruités	109
III.2	Méthode d'intégration globale	114
III.2.1	Comparaison de MNT	115
III.2.1.1	Conversion dans un même référentiel cartographique.....	116
III.2.1.2	Partition en domaine de recouvrement	116
III.2.1.3	Analyse de corrélation	117
III.2.2	Segmentation et extraction de lignes caractéristiques.....	117
III.2.2.1	Segmentation	117
III.2.2.2	Extraction de lignes caractéristiques.....	118
III.2.3	Intégration des modèles de terrain.....	119
III.2.3.1	Interpolation des trous	119
III.2.3.2	Appariement	120
III.2.3.3	Déformation élastique et Fusion de MNT	120
III.3	Validation de la méthodologie sur les données simulées.....	122
III.3.1	Modèle du chenal	122
III.3.2	Modèle Numérique de Terrain représentant un littoral	130
III.3.2.1	Fusion d'un modèle bruité et d'un modèle d'origine.....	130
III.3.2.2	Fusion du modèle topographique et du modèle bathymétrique	136
III.4	Perspectives d'application de la méthodologie sur des données littorales	153
III.4.1	Conversion géographique.....	154
III.4.2	Partition en domaines de recouvrement	154
III.4.3	Analyse de corrélation.....	154
III.4.4	Analyse paysagère.....	154
III.4.4.1	Extraction et pré-enrichissement	155
III.4.4.2	Segmentation	156
III.4.4.3	Enrichissement.....	156
III.4.5	Appariement	157
III.4.6	Corrections élastiques	157
III.5	Bilan sur la méthodologie d'intégration.....	157
	Conclusion générale.....	159
1	De la reconnaissance d'éléments caractéristiques à leur appariement.	160
2	Méthode globale d'intégration des données topo-bathymétriques	161
3	Perspectives.....	162
	Références bibliographiques.....	167
	Table des Figures	175
	Table des Tableaux.....	180

RESUME

Le domaine maritime littoral constitue un milieu fragile, placé à l'interface Terre/Mer où s'exercent de nombreuses pressions issues des milieux terrestres, maritimes et atmosphériques. Les approches globales de Gestion Intégrée des Zones Côtières (GIZC) ont pour but de pallier les lacunes des politiques précédentes au niveau des interactions simultanées des différents acteurs du littoral. Cette approche requiert un Modèle Numérique de Terrain (MNT) littoral continu ainsi que des données maritimes et terrestres cohérentes entre elles.

En raison d'acquisitions ou de productions différentes pour les MNT terrestres et maritimes, l'ensemble du relief n'est pas traité de manière ni homogène ni continue. De plus les techniques d'intégration des MNT actuelles entraînent des lissages des reliefs, voire la perte des éléments caractéristiques.

Des outils et une nouvelle méthodologie d'intégration de MNT ont donc été définis. Les outils utilisent la distance de Fréchet pour calculer une distance maximale d'écartement entre les points homologues de deux lignes. La valeur de cette distance donne une mesure de ressemblance de forme qui sert de base à leur appariement ultérieur.

La méthodologie de fusion proposée dans cette thèse se veut générique tout en s'appuyant sur des lignes caractéristiques spécifiques du paysage. Elle effectue un enchaînement de traitements répartis en trois étapes majeures : 1) transformations cartographiques et analyse des zones de recouvrement ; 2) Segmentation et extraction de lignes caractéristiques ; 3) appariement et fusion à l'aide de déformations élastiques.

Ces outils d'intégration ont été testés sur des données réelles en 2D pour la comparaison du trait de côte, puis validés sur des MNT simulés 3D.

Mots-clés:

Modèle numérique de terrain (MNT), intégration terre/mer de MNT, géomorphologie littorale, déformation élastique, Système d'Information Géographique (SIG), distance de Fréchet, appariement et fusion de données.

CONTEXTE DE LA THESE

Sommaire

1	Acteurs et données géographiques littorales.....	9
2	Continuité et cohérence au sein des données géographiques	10
3	Contexte de recherche.....	11
4	Démarche et problématique	14
5	Plan de thèse	16

Les récentes catastrophes survenues en domaine littoral et leur impact sur la population humaine (marées noires, tsunamis) ont amené les acteurs du littoral à prendre conscience de l'importance de la frange littorale et à repenser son mode de gestion. En effet, le littoral constitue une zone géographique aux limites floues où la confrontation Terre/Mer génère des espaces complexes où interfèrent de multiples composantes, physiques, biologiques, ou anthropiques, et ce à différentes échelles spatio-temporelles.

Ce système riche et diversifié se trouve être le siège d'enjeux à la fois environnementaux, patrimoniaux, économiques et nationaux, où les conflits pour les ressources, les paysages ou l'espace sont les plus exacerbés. Dans ce cadre, la compréhension des interactions entre les activités anthropiques et le milieu littoral est un objectif majeur de la recherche pour l'aide au développement durable des sociétés humaines. C'est pourquoi il est nécessaire de se focaliser sur les modalités d'utilisation et d'exploitation de l'environnement par l'homme. L'information géographique, par sa capacité à représenter une information localisée dans le temps et l'espace, s'impose comme un support privilégié pour avoir une vision globale et interdisciplinaire du territoire et de ses acteurs.

1 Acteurs et données géographiques littorales

Jusqu'à présent les interactions simultanées des politiques envisagées, des enjeux de durabilité et des conditions écologiques n'étaient pas respectées par les approches sectorielles classiques. En effet, de nombreux intervenants se partagent le « droit » au littoral, créant des tensions en cas de mauvaises concertations ou d'intérêts contradictoires [RUOA, 2006]. Plusieurs pôles majeurs centralisent ces enjeux parfois divergents.

Le premier pôle concerne la **problématique environnementale** au travers des **programmes de recherche scientifique** sur l'étude et la protection des écosystèmes côtiers [RUOA, 2006], ainsi que la prévention des risques naturels présents (érosion côtière¹, tsunamis, inondations).

¹ <http://www.eurosion.org/>

A ces problématiques s'ajoutent de nombreuses **applications cartographiques** en géologie (IFREMER, BRGM) et des thématiques plus géographiques centrées sur la remise à jour du trait de côte² ([Allain et al., 2000], [Le Berre et al., 2004]) ou l'établissement des cartes marines (SHOM). De façon plus marginale, le besoin de cartographie continue du littoral intervient dans beaucoup d'applications militaires sur la frange littorale. Elles concernent les problèmes de débarquement des troupes ou de véhicules amphibies sur des rivages.

A cette préservation du milieu s'opposent les **activités d'aménagement du littoral** d'une part et les activités **économico-industrielles** d'autre part. Les premières concernent les aménagements urbains, touristiques et de loisirs, tandis que les secondes représentent les constructions industrielles : installations portuaires, usines à énergie marémotrice. Les activités économiques ont trait à l'exploitation des ressources maritimes et de l'estran (pêche, aquaculture).

L'ensemble de ces acteurs ont besoin d'une connaissance précise du littoral ainsi que d'une représentation suffisamment fine pour obtenir une bonne compréhension des phénomènes étudiés et orienter au mieux les décisions envisagées. Ceci passe par l'intermédiaire d'informations acquises soit par des organismes publics, soit par des sociétés privées dans le cadre de programmes de recherche spécifiques.

Parmi toutes les données collectées, le **modèle numérique de terrain** (MNT) tient une place importante en raison de sa capacité à modéliser l'environnement d'étude au plus proche de la réalité ([Li et al., 2005]). Il permet de localiser dans un espace géographique un phénomène (répartition biologique, construction, courant marin, etc.), tout en permettant un déplacement dans une dimension supplémentaire. En effet, il représente le relief à la manière d'une carte avec d'autres informations majeures : l'altitude et ses dérivées. Même s'il n'intervient pas directement dans les études, il sert de support pour la majorité des informations collectées grâce à son impact visuel fort :

- l'impact des mécanismes d'érosion, et donc de l'évolution du trait de côte, est fortement dépendant de la force des vagues et de la « rudesse » du paysage concerné (altitudes hautes ou basses, reliefs émoussés ou à crêtes saillantes) ;
- un plan de prévention des risques utilise un MNT pour prédire les zones potentiellement inondables suite à une montée des eaux, que ce soit en raison de grandes marées ou de tsunamis (raz de marée). L'étude des raz de marée utilise le calcul des pentes et de leur angle de stabilité : la formation de ces vagues meurtrières est induite par un glissement de terrain provoqué lors d'un tremblement de terre ;
- l'aménagement du littoral s'appuie également sur une vision du milieu pour étudier le rendu d'un ouvrage important (pont, barrage), ou tout simplement l'impact, aussi bien visuel qu'écologique, d'une construction touristique ;
- en biologie marine, la finesse d'un MNT permet de prédire les lieux privilégiés d'habitats côtiers (dans les infractuosités par exemple) ;
- toute étude multi-échelle ou temporelle de la géodynamique d'un milieu en géomorphologie repose sur un modèle de relief.

Dans ce cadre, un logiciel capable d'importer et de gérer un grand nombre de données de sources variées impose la reconsidération de leur cohérence et de leur complémentarité.

2 Continuité et cohérence au sein des données géographiques

La vulgarisation des Systèmes d'Informations Géographiques (SIG) a considérablement favorisé l'échange d'informations ainsi que la concertation entre ces divers intervenants au

² Notamment : <http://www.logicacmg.com/coastchart/default.htm>

travers d'une vision d'ensemble du littoral. Un SIG est un logiciel informatique capable d'organiser et de présenter des données numériques spatialement référencées en vue de les analyser ou de les cartographier.

En parallèle, ce mouvement est favorisé par le constat de la nécessité réelle de mettre en place une politique globale de gestion intégrée des zones côtières (GIZC) qui ne soit plus dépendante des secteurs économiques ou des communautés urbaines. Il y a donc un besoin croissant de données suffisamment **homogènes** pour croiser les informations, et **continues** dans le temps (pour les visions à long terme) et dans l'espace littoral (pour le suivi des phénomènes naturels qui ignorent les frontières terrestre/marin). Ces requis impliquent une reconsidération des méthodes de production des données et de leurs traitements.

En effet, en France (comme dans certains autres pays, européens ou non), la majorité des données littorales disponibles est acquise par des organismes publics différents selon le milieu considéré : le service hydrographique de la marine (SHOM) pour la bathymétrie et l'institut géographique national (IGN) pour la topographie. Certains problèmes surviennent alors : référentiels, systèmes de coordonnées, formats de données, standards utilisés, méthodes d'acquisition différentes et surtout continuité de l'information non assurée.

Tout ceci entraîne la présence simultanée de plusieurs bases de données géographiques qui fournissent des représentations du milieu qui varient selon les spécifications, l'échelle de saisie, la qualité, l'actualité. Ces différences sont susceptibles d'engendrer des réponses discordantes sur une requête faite via un SIG ou lors d'un travail de synthèse d'informations. Il est donc impératif de mettre les données en cohérence, ou de disposer d'un standard (comme par exemple le format S57 du SHOM), avant toute analyse. Par ailleurs la majorité des activités ignore les frontières (phénomènes de pollution par exemple), ce qui implique également de disposer de données ininterrompues.

3 *Contexte de recherche*

Le manque de cohérence et de continuité des données a entraîné la mise en place depuis une trentaine d'années, aux niveaux **international** et **national**, de programmes relatifs à la gestion intégrée des zones côtières et de recherches scientifiques dans ce domaine.

EuroSDR³ (European Spatial Data Research) [EuroSDR, 2007] est un organisme européen établi à Paris en 1953 qui comprend dix-sept pays au travers d'organisations géographiques nationales, d'instituts de recherche et d'entreprises du secteur privé. Il a pour finalité d'identifier les besoins de la recherche au niveau du financement de la gestion et de la distribution de données spatiales. Ceci se réalise notamment grâce à des collaborations internationales sur des projets de recherche ou des conférences. L'objectif est de déterminer et de documenter l'état de l'art et les progrès effectués en Europe au niveau de l'intégration de bases de données marines et terrestres. Les rapports issus des différentes conférences ont montré la nécessité :

- de définir un standard de données spatiales (nommé S-100 en référence au format actuel des données marines S-57) afin de faciliter l'échange de données hydrographiques ;
- de promouvoir ce format auprès des organismes géographiques ;
- et de planifier au niveau des organismes participants des campagnes d'acquisition continue du littoral.

³ http://www.eurocdr.net/workshops/landsea_2007/organisers.htm

L'association **Eurocoast**⁴ (**European Coastal Association for Science and Technology**), fondée en 1989, représente un regroupement de scientifiques, ingénieurs et acteurs de la prise de décision en milieu littoral. Les objectifs de cette association sont, en premier lieu, de créer un réseau d'échange et de coopération technique sur des sujets relatifs à la protection, au développement et à la gestion des zones côtières, et ainsi améliorer les connaissances existantes dans ce domaine. En second lieu, elle vise à promouvoir les recherches inter disciplinaires ainsi que la synthèse de thèmes communs entre utilisateurs de domaines différents.

Son objectif global est la conception d'une base de données comportant tous les aspects de la zone côtière et sa large communication au sein de la communauté scientifique. Par exemple, en 2002, cette association a permis des échanges internationaux pour la mise en place d'un projet de gestion intégrée des zones côtières (**ICZM**⁵) en Ukraine⁶.

Parmi les multiples symposiums centrés sur la thématique de la gestion côtière, **CoastGIS**⁷ se démarque par ses orientations SIG et cartographie assistée par ordinateur appliquée aux domaines côtiers et marins. La première rencontre CoastGIS, qui a eu lieu à Cork (Irlande) en 1995, n'avait pas une thématique bien définie mais a servi « d'état de l'art – bilan » aux groupes d'intérêt variés alors participants. Depuis, tous les deux ans, ce symposium constitue un portail d'échanges complémentaires et interdisciplinaires entre experts du SIG et spécialistes côtiers. Le but est d'initier un partage d'expérience et/ou de connaissances entre chercheurs et ingénieurs tout en favorisant des collaborations mutuelles. Il vise également à utiliser l'expertise scientifique aux travers de SIG afin de résoudre les problèmes de gestion côtière.

En parallèle à ces organismes et congrès, de nombreux programmes européens ont été lancés avec des thématiques assez diverses mais concernant toute la recherche et le développement en milieu littoral : les programmes successifs **ROCC** et **PROTECT** visent à étudier l'érosion des falaises de craie du littoral de la manche. **EUROSION** et son successeur **RESPONSE** étudient l'érosion côtière en général, tout en prenant en compte les problèmes de variations côtières et l'apparition de risques liés aux changements climatiques.

Au niveau **national**, le projet **Litto 3D** [Louvart & Gateau, 2005] a été initié pour répondre à une directive européenne datant de 2002. En effet, suite aux catastrophes pétrolières successives, l'Europe a décidé de renforcer sa politique environnementale côtière en recommandant aux états membres d'effectuer un inventaire détaillé du littoral. En France, les deux organismes de cartographie, l'IGN et le SHOM, se sont alors associés pour mettre en commun leurs compétences. Une gestion optimale du littoral nécessite de prendre en compte plusieurs thématiques :

- environnementales, au travers du suivi du trait de côte sous l'effet de l'érosion, de la protection du domaine côtier et de la protection de la faune et de la flore ;
- humaines, avec la prévention des risques d'inondation, de pollution, ou de catastrophes naturelles ;
- d'équipement, avec les aménagements de loisirs ou industriels ;
- économiques, avec l'exploitation des ressources maritimes ;
- de recherche scientifique, avec l'étude des écosystèmes spécifiques à l'estran, ou de l'impact des marées noires sur la faune et la flore par exemple ;

⁴ <http://www.eurocoast.org/index.html>

⁵ <http://ec.europa.eu/environment/iczm/home.htm>

⁶ <http://www.eurocoast.org/ukraine%20project.pdf>

⁷ <http://www.coastgis.org/>

- militaires, dans le cadre de débarquements de troupes et d'applications amphibies.

Suite au bilan de ces besoins, le constat a été fait que les données existantes dans le domaine étaient insuffisantes au niveau de leur précision, de leur validité et de leur continuité. En effet, les données bathymétriques du SHOM sont de grande précision mais discontinues dans le temps et l'espace. A l'inverse, les données de l'IGN sont denses avec une grande fréquence d'acquisition mais insuffisamment précises pour produire des cartes de navigation ou d'estimation de risque lié à l'érosion ou autres phénomènes côtiers. Dans ce cadre, il a été décidé de procéder à la réacquisition complète du littoral à l'aide de nouveaux moyens de mesure. Dans l'attente de la production de cette cartographie exhaustive du littoral, une base de données historique, **Histolitt** [Louvart & Grateau, 2005], a été constituée à partir des bases de données du SHOM et de l'IGN. Au final, cette base de données comporte une densité d'information bien supérieure à celle figurant sur les cartes, car elle rassemble toutes les données historiques collectées depuis la création des deux organismes.

De même, au niveau de la recherche française, la problématique plus large de l'intégration de représentations multiples est traitée par l'**axe de représentations multiples de l'information géographique**, antenne du Groupe de Recherche SIGMA⁸ (Systèmes d'Information Géographique, Méthodologies et Applications) dit CASSINI. L'objectif de ce groupe de travail est de proposer des méthodologies facilitant la multi représentation en fonction des besoins et permettant la gestion conjointe de données disponibles à différents niveaux de détail. La multi représentation naît de différentes traductions de phénomènes naturels ou anthropiques complexes, dans des bases de données géographiques ou des cartes, selon les spécifications, le modèle attendu, les besoins des utilisateurs et/ou les objectifs de l'application. Elle se traduit par des niveaux de détail, des représentations de simulation, des représentations temporelles ou des points de vue différents, nécessaires à une bonne prise de décision.

Les données possédant de multiples représentations du même phénomène entraînent des difficultés lors du processus d'intégration et de fusion. C'est pourquoi le but final de ce groupe de travail est de proposer un formalisme et/ou des méthodologies qui permettent d'intégrer la représentation multiple dans la gestion de l'information géographique sous toutes ses formes. De nombreux intervenants travaillent sur l'intégration de bases de données vectorielles, l'appariement automatique [Mustière & Devogele, 2008], l'évaluation de la cohérence entre les représentations [Sheeren, 2005] ou entre MNT dans un contexte de risques [Rousseaux, 2005]. Leurs approches offrent une vision complémentaire à cette thèse au niveau de la gestion des incohérences entre représentations, de la problématique d'appariement⁹ ou de l'intégration.

En parallèle, le groupe de recherche **SIG** de l'institut de recherche de l'école navale (**IRENav**) a donné à cette thèse son aspect interdisciplinaire entre les divers domaines de la géographie, de l'informatique, et de la géomatique. En effet, les axes de recherche de ce groupe se focalisent sur la modélisation et l'analyse de l'information spatiale et temporelle dans les SIG à deux et trois dimensions appliquées au contexte maritime. Ils s'organisent selon deux axes majeurs :

- le développement de modèles spatio-temporels d'intégration et d'analyse de l'information géographique ;

⁸ <http://recherche.ign.fr/labos/cogit/sigma/index.htm>

⁹ Processus qui vise à identifier entre deux bases de données les représentations du même phénomène du monde réel.

- le développement de modèles de représentation et de traitement des données maritimes dans les domaines du transport et de la sécurité maritime.

Ces thématiques comprennent le raisonnement spatial et temporel, la géométrie algorithmique et géomorphométrique et la conception de SIG. Le sujet de cette thèse se place donc transversalement entre ces domaines d'études : premièrement parce qu'elle tient compte des changements morphologiques temporels, sources d'incohérences lors du processus d'intégration [Sriti et al., 2005]. Deuxièmement parce qu'elle se base également sur la caractérisation de relief via ses éléments topographiques [Saux et al., 2004]. Et finalement parce qu'elle reprend les travaux développés par [Devogele, 2002] sur l'appariement de bases de données à l'aide de l'algorithme de Fréchet. Toutes ces notions sont utilisées afin de réaliser un outil SIG d'intégration de MNT basé sur l'analyse du paysage littoral afin de proposer une représentation continue terre/mer. En effet de nombreuses organisations ou organismes gouvernementaux ont établi l'importance et la nécessité de posséder une cartographie complète et précise du littoral.

4 Démarche et problématique

Pour faire face à cette demande, de nombreux programmes de recherche et de campagnes de ré acquisition ont été lancés. Cependant il n'existe aucune indication sur la manière de traiter par la suite ces données au mieux afin d'obtenir une représentation continue ni comment les intégrer avec celles déjà existantes. Cette étape a longtemps été considérée comme secondaire. Néanmoins elle entraîne de nombreuses erreurs et imprécisions liées aux traitements effectués parfois de manière non appropriée. En effet, les processus d'intégration employés sont définis selon un jeu de données spécifique ayant une qualité qui lui est propre, et la connaissance qu'a l'utilisateur du milieu qu'il étudie. Il en résulte la plupart du temps, la mise en place de méthodes ponctuelles non réutilisables dans un autre contexte. L'idéal serait de pouvoir disposer d'outils permettant dans un premier temps un pré traitement générique des données et dans un deuxième temps, en cas de difficultés, de permettre à l'utilisateur de décider des traitements les meilleurs. Ceci se rapproche de ce qui est réalisé pour l'intégration de bases de donnée géographiques.

La démarche ici est donc d'établir un lien entre les domaines de la géométrie algorithmique, de la géomorphologie et enfin des systèmes d'information géographique au travers de l'intégration de données Terre/Mer. Les processus d'intégration actuellement décrits font appel à des notions et des outils développés par la géométrie algorithmique et appliqués aux bases de données géographiques multi représentations (deux dimensions). Les données Terre/Mer utilisées dans ce travail sont des modèles numériques de terrain (trois dimensions) pour lesquels très peu d'outils SIG ont été développés.

L'objectif de cette thèse est de proposer une méthodologie d'intégration de MNT de manière à obtenir une représentation continue la plus exacte possible en fonction des données disponibles. Pour cela divers outils SIG sont nécessaires. Certains des outils identifiés ont donc été implémentés afin de répondre à cette attente. Cette thèse se place ainsi à la frontière entre la géographie, la géométrie algorithmique, et les SIG (voir figure 1).

La géographie, par l'intermédiaire de la géomorphologie littorale apporte la compréhension nécessaire du paysage pour réaliser une intégration pertinente des MNT. La géométrie algorithmique met à disposition des outils de comparaison et d'appariement d'objets linéaires, ponctuels ou surfaciques. La programmation de ces outils permet d'effectuer une fusion de ces objets. Les SIG, par la visualisation de ces données nouvellement intégrées, offrent une multitude d'applications possibles.

Le point fort de la méthodologie présentée est qu'elle est générique à tous types de MNT (issus du programme Litto3D ou antérieurs) mais s'adapte aux divers paysages traités dans les modèles numériques. Ainsi, les éléments qui font la particularité d'un relief sont conservés. L'originalité de cette préservation de l'information clé du relief est d'éviter une perte de l'information pertinente pour des modélisations ultérieures (présence ou non d' « accidents » dans le paysage).

Les différentes étapes d'intégration impliquent premièrement de pouvoir mesurer des similitudes entre objets issus de deux bases de données. A partir de ces mesures il s'agit de déterminer si les deux objets étudiés représentent bien le même phénomène du monde réel ou non (appariement). Les outils de comparaison et d'appariement qui permettent de réaliser cette étape sont issus de ceux déjà existants pour les données 2D.

Ensuite les différents types de relief sont identifiés au sein de chaque modèle numérique de terrain. Une typologie basée sur des connaissances en géomorphologie du littoral recense de manière exhaustive les diverses classes de paysages côtiers.

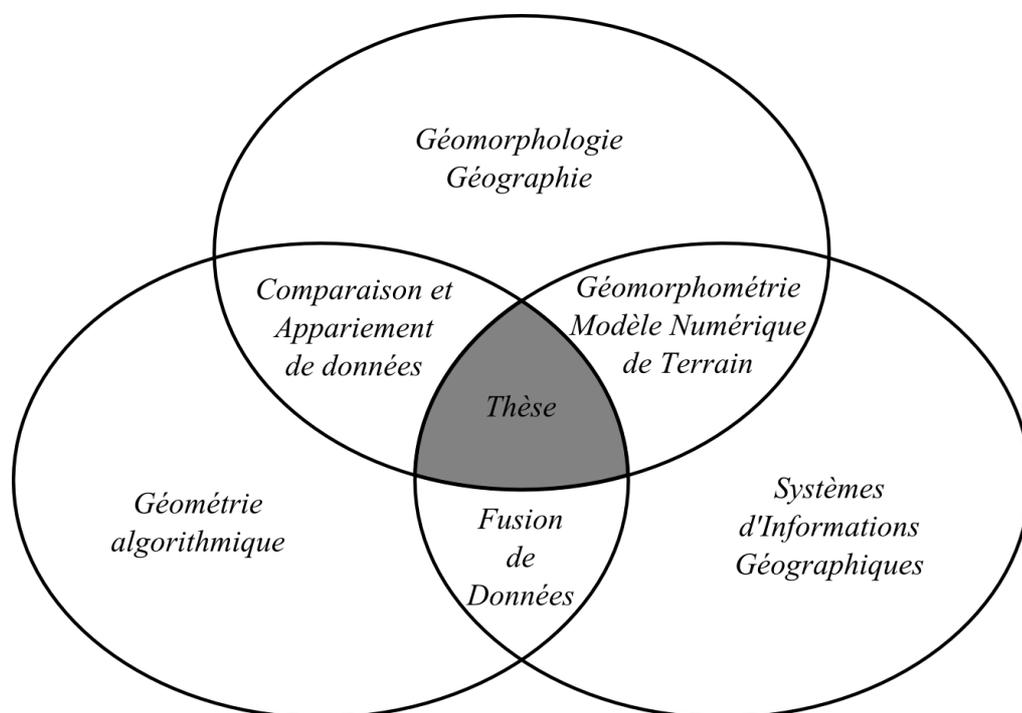


Figure 1 : Positionnement de la thèse.

Cette typologie est alors complétée à l'aide de valeurs numériques données par la géomorphométrie. Ces valeurs permettent l'extraction des reliefs identifiés par la typologie dans n'importe quel MNT. Par ailleurs, des algorithmes de traitement ont été suggérés pour chaque type paysager en fonction de ses caractéristiques. De cette manière il est possible d'adapter les traitements effectués sur le modèle et ainsi préserver les informations importantes lors de l'intégration.

La **méthodologie d'intégration** mise en œuvre a pour but de produire au final un modèle continu et cohérent entre les divers modèles. Elle effectue pour cela un enchaînement de traitements répartis en trois phases principales qui visent à préparer au mieux les données au processus d'intégration (conversion dans un système cartographique, enrichissement et segmentation des modèles, appariement, etc...). L'intérêt de cette méthode est qu'elle travaille de manière relative sur les données elles-mêmes. Elle s'affranchit donc des problèmes de positionnement, de variabilité temporelle ou de différence de qualités. En conséquence, elle permet d'intégrer des données déjà existantes comme des données

nouvellement produites avec les mêmes considérations paysagères. En ce sens, elle se veut adaptative tout en étant générique.

5 Plan de thèse

Les différentes phases de la méthodologie d'intégration sont directement dépendantes des outils sur lesquels elle se base. Par exemple la fusion proprement dite ne pourra s'effectuer qu'après une étape préalable de recalage des données et de la détection des éléments représentant le même phénomène du monde réel.

La première partie de cette thèse réalise tout d'abord une description de l'ensemble du milieu littoral, des agents édificateurs du relief présents et des classes de paysages rencontrés sur les côtes. Elle définit ensuite les termes utilisés lors du processus d'intégration et pose ainsi les fondements des processus d'intégration pour des données 2D. En complément, elle détaille la structure des données 3D et ce qui a été établi dans le domaine de la fusion de MNT.

Les algorithmes de fusion choisis suite à l'état de l'art et les types de relief à traiter étant connus, il est nécessaire dans une seconde partie d'expliquer quelles sont les extensions imposées pour réaliser cette intégration, conformément aux objectifs définis.

Une troisième partie effectue la synthèse de tout ce qui a été introduit précédemment en décrivant en détail la méthodologie d'intégration proposée et son application sur des données simulées simples. L'application à ce type de données permet de vérifier la robustesse des outils implémentés, et ainsi les valider, mais également d'en mesurer les limites et de proposer des améliorations futures.

La conclusion dresse un bilan des apports de cette thèse vis-à-vis du processus d'intégration. En perspectives, elle propose également des améliorations des outils développés en vue d'une chaîne de traitement la plus automatique et complète possible.

I LITTORAL ET INTEGRATION : PRESENTATION, ETAT DE L'ART ET DISCUSSION

Sommaire

I.1	Présentation du milieu littoral	18
I.1.1	Agents morphogéniques des reliefs littoraux	19
I.1.2	Zonalité des formes littorales selon les grands domaines climatiques	21
I.2	Intégration d'éléments 2D	22
I.2.1	Définitions	22
I.2.2	Intégration de points	27
I.2.3	Intégration de lignes	28
I.2.4	Intégration de polygones	38
I.3	Intégration de Données 2.5D (MNT)	40
I.3.1	Caractéristiques des MNT	40
I.3.2	Fusion de MNT : travaux en cours ou déjà réalisés.....	51
I.4	Bilan sur l'état de l'art.....	59

Un processus d'intégration de plusieurs jeux de données, qu'elles soient planes ou de relief, vise la compilation de l'ensemble des informations disponibles de manière cohérente et non redondante sur une même zone. Selon la nature et l'organisation des données, la méthode de fusion ne peut pas être la même. Pour les données planes, de nombreuses méthodes et processus de fusion ont été amplement développés et décrits dans la littérature ([Laurini, 1996], [Devogele, 2000], [Beerli et al., 2004]). Ils se basent en général sur les identifiants communs contenus dans les tables attributaires des données vectorielles. Si aucune information analogue n'est disponible, un processus de mise en correspondance des données homologues entre bases (appariement) est mis en jeu.

Pour des données de relief en revanche, ces méthodes ont été très peu ou pas développées. En effet la fusion d'un modèle numérique de terrain (MNT), qui ne possède pas de table attributaire, se borne le plus souvent à moyenniser les altitudes « Z » selon le positionnement géographique X et Y des points (ou des cellules pour une grille raster). Si ces derniers sont décalés l'un par rapport à l'autre, des difficultés apparaissent. En effet il existe peu de moyens pour vérifier la bonne mise en correspondance des modèles entre eux.

Un processus de fusion complet se doit, pour garantir la qualité du résultat final, tout d'abord de comparer les données, puis d'apparier celles qui sont homologues et peuvent être mises en correspondance et enfin de les fusionner entre elles.

C'est pourquoi l'objectif de ce chapitre est de faire la synthèse des travaux faits précédemment en matière d'intégration, tout d'abord sur des données planes (2D), puis sur des modèles de relief numériques.

Afin de bien comprendre le processus d'appariement et comment il s'effectue au travers d'algorithmes divers, il convient tout d'abord de détailler la représentation numérique des données 2D, puis de la même façon de présenter les données 3D et leurs différents formats.

La première partie de ce chapitre esquisse le milieu littoral dans son ensemble, en décrivant les agents intervenant dans la morphogenèse (formation et évolution du relief) ainsi que le zonage lié aux grands domaines climatiques.

La seconde partie définit les notions d'élément plan, de comparaison, d'appariement, et de fusion. Un état de l'art énumère ensuite les différents types d'éléments 2D. Enfin il expose les méthodes de comparaison, d'appariement et de fusion déjà implémentées pour chacune de ces catégories.

La troisième partie présente de la même manière les MNT, à savoir leur mode de production, leurs formats, l'estimation de leur qualité ainsi qu'une synthèse bibliographique des précédents travaux en cours ou déjà réalisés en matière de fusion de MNT.

La quatrième partie se focalise sur les problèmes d'appariement de données de relief en retenant notamment les difficultés liées à la nécessité de surmonter les décalages planimétriques et altimétriques. Le processus d'appariement de modèle numérique est introduit au travers de l'énumération des éléments qui vont permettre sa réalisation.

1.1 Présentation du milieu littoral

Le littoral représente l'interface entre la terre, la mer et l'atmosphère, mais plus qu'un simple trait de côte, c'est avant tout un espace géographique vivant qui possède sa dynamique propre issue des trois domaines qu'il côtoie. Il comporte l'estran, superficie comprise entre le niveau

de plus haute mer et celui de plus basse mer (zone intertidale), la zone infratidale et l'arrière côte (zone supratidale).

La zone infratidale est constituée par les fonds océaniques toujours immergés qui débutent au niveau des plus basses mers. L'arrière côte désigne les terres susceptibles d'être atteintes par l'eau salée (lors d'une tempête ou d'une inondation), ou directement influencées par l'action des eaux marines et le vent (dans les deltas, les estuaires ou les marais maritimes). De part sa position, l'élaboration du relief de la zone littorale reflète les effets de multiples processus dynamiques, qu'ils soient maritimes, continentaux et/ou atmosphériques. Pour cette raison, certains milieux possèdent une influence continentale plus marquée (marais maritimes) tandis que d'autres voient l'expression majoritaire des agents marins. En effet, le modelé paysager est dépendant de l'action d'agents morphogéniques et de ses environnements géographique et climatique.

Indépendamment du contexte lithologique qui, en lui-même, entraîne des variations morphologiques importantes, cette section décrit les agents morphogéniques (actions mécaniques et actions physico-chimiques et biologiques). Elle présente également les principaux domaines climatiques dans lesquels ces agents s'expriment plus ou moins.

I.1.1 Agents morphogéniques des reliefs littoraux

Les agents morphogéniques sont des processus élémentaires qui se combinent de manière à façonner le relief [Coque, 1993]. Selon le type d'altération engendrée sur le milieu, il est possible d'établir un classement des divers processus : actions à dominantes mécanique, physico-chimique et biologique.

I.1.1.1 Actions mécaniques

Les agents à dominante mécanique agissent majoritairement par l'intermédiaire de l'eau (via les vagues, les courants marins, les marées, le gel et le sel). Leur action combinée au vent entraîne le plus souvent l'érosion (fragmentation et désagrégation) de la roche en relation avec sa porosité et sa fracturation. Sur d'autres types de substrats comme le sable l'action se manifeste également par un déplacement de matière puis une sédimentation.

Vagues

Les vagues peuvent avoir sur le paysage une action à dominante érosive (phénomène d'ablation), de dépôt (jet de rive) ou de transport (dérive littorale), modifiée ou non par les marées. Ces actions sont gouvernées par la bathymétrie qui entraîne les phénomènes de turbulence, ou de déferlement en bord de plage. Ce dernier a un impact plus fort sur l'élaboration des formes (morphogénèse) en raison de l'importante dissipation d'énergie contenue dans la vague. L'effet de cette action mécanique est plus ou moins marqué selon la topographie, l'exposition et le substrat rocheux rencontrés. Ainsi, sur des côtes à falaises, les vagues peuvent détacher des blocs importants, créer des encoches susceptibles d'entraîner des éboulements, et engendrer l'abrasion des affleurements rocheux grâce au sable et aux blocs véhiculés dans l'eau [Valadas, 2005].

Courants marins

Les courants marins sont des déplacements à l'échelle planétaire des masses d'eau des mers et des océans. Les facteurs qui déterminent ces mouvements sont d'une part les vents, et d'autre part la différence de températures, de salinités et donc les différences de densités des eaux. Ces facteurs distinguent deux types de courants : les courants de surface et les courants profonds [Saint-Guily, 2006]. Concernant leur influence sur le modelé littoral, seuls les

courants de surface ont un réel impact sur le modelé littoral avec les courants de débris (ou de dérives littorales), les courants de décharge, et les courants de marée [Derruau, 1988].

Les courants de marée

La marée est le mouvement montant (flux) ou descendant (reflux) des eaux des mers ou des océans causé par l'effet conjugué des forces de gravitation de la lune et du soleil. Selon la force de son courant (influencé par la pente générale), elle peut soit avoir une action érosive, soit au contraire favoriser la sédimentation et édifier des cordons littoraux par exemple [Tricart, 1995].

Gel

Le gel agit en tant que processus climatique majoritairement dans les régions froides, et à une échelle moindre aux latitudes tempérées, de trois manières :

- Par l'intermédiaire de la glace d'hiver qui constitue un *pied de glace* adhérent aux falaises au niveau des hautes mers. Lorsque celle-ci descend, le poids entraîne alors le détachement de blocs entiers du support [Derruau, 1988]. Par ailleurs la succession de phases de gel et de dégel favorise l'éclatement de la roche par agrandissement de ses fissures (phénomènes de cryoclastie).
- Par l'intermédiaire des icebergs qui emprisonnent des blocs qui peuvent soit être délestés soit raclés sur les fonds marins de faible profondeur [Derruau, 1988].
- Lors de la fonte de coins de glace en été, ce qui a pour conséquences d'évacuer par la mer des morceaux de substrat présents entre les coins de glace et d'entraîner le recul de la côte de plusieurs mètres par an [Derruau, 1988].

Vents

En plus d'être à l'origine de la houle, les vents ont un impact fort sur des paysages de sable dépourvus de végétation. En effet, selon sa force, sa régularité et la granulométrie des sables, le vent mobilise les éléments sableux, le plus souvent vers l'intérieur des terres, pour constituer des dunes littorales [Valadas, 2005].

I.1.1.2 Actions physico-chimiques et biologiques

Ce type d'actions privilégie en majorité les processus d'altérations : corrosion et dissolution. La composition du substrat va donc jouer un rôle beaucoup plus important que pour les actions mécaniques.

Dissolution

Le processus de dissolution, produit essentiellement par les organismes vivants et l'eau de mer, exploite les fractures et lignes de faiblesse de la roche, majoritairement les roches carbonatées. Elle s'effectue par les embruns au dessus du niveau de marée haute (appelée dissolution karstique en raison de sa présence majoritaire dans les paysages de type karst) et sur l'estran grâce à la présence de flaques [Derruau, 1988]. Les formes de dissolution des calcaires se traduisent par :

- Les lapiès littoraux (forme la plus courante des karsts, paysages façonnés dans des roches solubles carbonatées) qui se manifestent par une surface déchiquetée par de nombreuses rainures, fissures et crevasses de tailles variables plus ou moins profondes issues de la dissolution et regroupées en ensembles denses.
- Les vasques de quelques mètres de diamètre et dont le fond plat permet la persistance de mares à marée basse.
- Les encoches de corrosion, plus fréquentes en milieu tropical, et favorisées par les sécrétions des êtres vivants.

- La désagrégation granulaire, engendrée par les variations de volume dues aux cristaux salins formés par évaporation de l'eau [Tricart, 1995] (processus d'haloclastie).

Actions biologiques

La faune et la flore présentes dans l'écosystème littoral ont divers types d'actions sur leur environnement. Les débris des organismes à coquilles et tests fournissent des matériaux meubles qui augmentent la quantité de matériel sédimentaire des plages et des cordons littoraux [Tricart, 1995]. Certains autres organismes comme les algues encroûtantes ou les récifs coralliens sont des organismes constructeurs qui permettent l'édification de modelés tout en les protégeant de l'attaque érosive des vagues. D'autres au contraire participent au démembrement des roches, soit en les perforant pour y vivre, soit en les corrodant comme certaines algues cyanophycées sur le calcaire [Tricart, 1995].

I.1.2 Zonalité des formes littorales selon les grands domaines climatiques

Les modelés d'une région sont marqués par la présence de formes élémentaires spécifiques dues à des combinaisons de processus morphogéniques particuliers dépendant du contexte géographique et climatique. Les travaux de [Valadas, 2005] que nous utilisons ici permettent de différencier trois grands domaines à l'échelle planétaire.

I.1.2.1 Zones tempérées

Ces types de littoraux comportent surtout des côtes rocheuses, des côtes basses (sableuses ou en sédimentation) des vasières et accumulation de galets. Ils sont caractérisés par une forte action mécanique des vagues (vents, tempêtes fréquentes) entraînant une érosion importante et un recul du trait de côte [documentation française, 1997]. Le dépôt du matériel érodé engendre une superficie assez conséquente des côtes construites à partir d'accumulation de sédiments (marins, éoliens ou terrestres) : plages, dunes et marais [Valadas, 2005].

I.1.2.2 Zones sub-tropicales

Ces latitudes favorisent les actions physico-chimiques et biochimiques comme agents morphogéniques du relief. Il est également fréquent qu'une forêt littorale pénètre jusqu'à la mer par le biais de mangroves et joue un rôle fixateur de matériaux fins, favorisant le remplissage de baies et de marais [Valadas, 2005]. Les côtes sableuses, les vasières littorales et les deltas sont de tailles très importantes en raison de l'abondance de sédiments apportés par les fleuves.

I.1.2.3 Zones polaires

Ces milieux sont marqués par une saisonnalité forte dans les zones où il existe un dégel saisonnier. L'action mécanique est dominante ainsi que des conditions climatiques extrêmes. Tous les processus morphogéniques sont majoritairement commandés par le gel, ce qui entraîne la formation d'immenses calottes glaciaires (inlandsis) qui, lorsqu'elles débordent sur la mer, forment des plates-formes épaisses de plusieurs dizaines ou centaines de mètres à l'origine des icebergs. Les plates-formes d'érosion littorales, les deltas, et les plages sont également très développés [Valadas, 2005].

1.2 Intégration d'éléments 2D

Cette section présente les éléments plans pour lesquels les processus d'intégration de données ont tout d'abord été développés vis à vis des données 3D. Les notions de dimensions géographiques d'un objet, de fusion, de comparaison et d'appariement sont donc définies en premier lieu afin de permettre au lecteur d'avoir une base bien définie pour considérer la suite du rapport. En effet, dans le domaine de l'intégration de données, un même terme a souvent des significations multiples ou est utilisé de différentes manières. Ensuite les multiples processus d'appariement et leur mode de comparaison des données sont détaillés.

1.2.1 Définitions

1.2.1.1 Dimensions géographiques des objets

La notion de dimension peut avoir plusieurs sens : on parle soit de dimension d'un objet relativement à sa taille (longueur, largeur, profondeur, diamètre), soit de sa dimension comme étant le nombre de variables qui définissent l'espace vectoriel dans lequel est décrit cet objet (X, Y, Z, mais il peut y avoir d'autres grandeurs comme le temps par exemple).

Dans cette thèse, un objet plan avec deux coordonnées (X et Y) est de dimension « 2 ». La dimension 2.5 s'utilise pour les modèles qui à un X et un Y n'associent qu'une seule et unique hauteur Z. S'il s'agit d'informations de relief, ces représentations sont appelées des Modèles Numériques de Terrain.

La 2.75D [Cambray, 1994] ajoute un niveau de détail en permettant d'associer deux altitudes Z à une même position tandis que la 3D autorise les trous et donc plusieurs Z pour un même X et Y.



Figure 2 : Exemple de représentation d'un pont selon les différents types de modèles disponibles (2D, 2.5D, 2.75D et 3D, source : [Devogele, 1997])

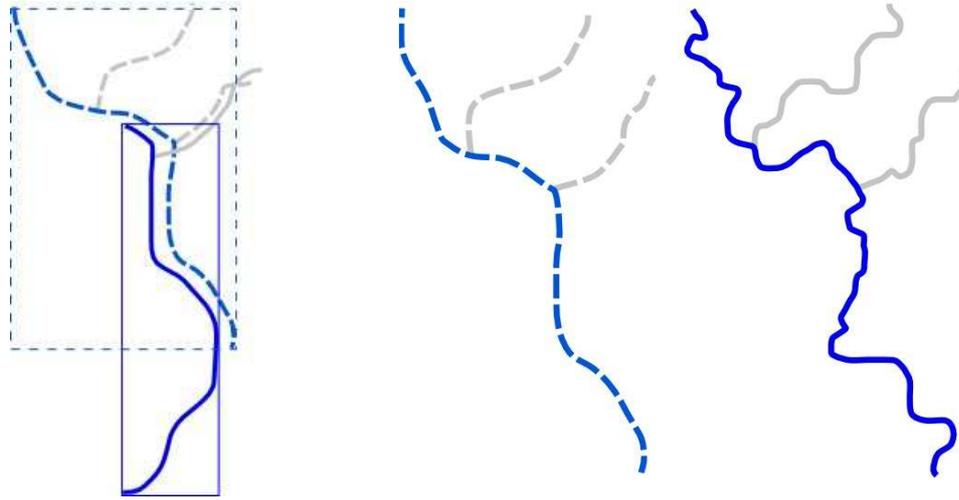
A titre d'exemple, la figure 2 illustre différentes représentations d'un pont dans les diverses dimensions : 2D, 2.5D, 2.75D et 3D.

1.2.1.2 Comparaison entre deux objets ou un ensemble d'objets

Les sections qui suivent se focalisent uniquement sur la définition générale de la comparaison et l'inventaire des disparités les plus communes. En effet, les outils propres au processus de comparaison sont différents selon le type de géométrie considéré. Ils sont donc détaillés plus amplement par la suite dans les parties consacrées aux objets 2D et 3D.

Une **comparaison** entre objets met en jeu toute fonction qui évalue des différences ou des similitudes. Ces mesures s'effectuent à l'aide de divers outils disponibles comme l'emprise de l'objet (figure 3(a)), sa forme (figure 3(b)), sa géométrie (figure 4), ses attributs thématiques (figure 5), et ses relations topologiques (figure 6) avec un ensemble d'objets voisins. La comparaison entre éléments issus de deux jeux de données divers n'est pas un processus simple car il existe un grand nombre de disparités entre les objets 2D.

La figure 3 illustre un exemple de comparaison entre deux objets issus de deux jeux de données différents, chacun représentant des parties distinctes d'un même fleuve. La figure 3(a) montre *les emprises* décalées avec néanmoins un recouvrement entre elles. L'emprise d'un objet géographique est souvent représentée par le plus petit rectangle englobant sa géométrie. Ce rectangle est généralement orienté selon les axes X et Y.



(a) (b)
Figure 3 : Représentation d'un même fleuve dans deux bases géographiques avec a) des emprises différentes, b) des niveaux de détail différents d'une base à l'autre.

La figure 3(b) dessine des niveaux de granularité variables entre les objets des deux bases de données. Le tracé en pointillé a été simplifié (par un processus de généralisation cartographique [Ruas, 2002a] par exemple) alors que pour le deuxième tracé, un niveau de détail important est présent. Pour cet exemple figure 3(b), si des outils de comparaison des formes sont employés, les deux objets ne sont pas reconnus comme identiques car l'écart entre les niveaux de détail est trop grand. Il en va de même pour des objets variables dans le temps comme le trait de côte ou le lit des fleuves.

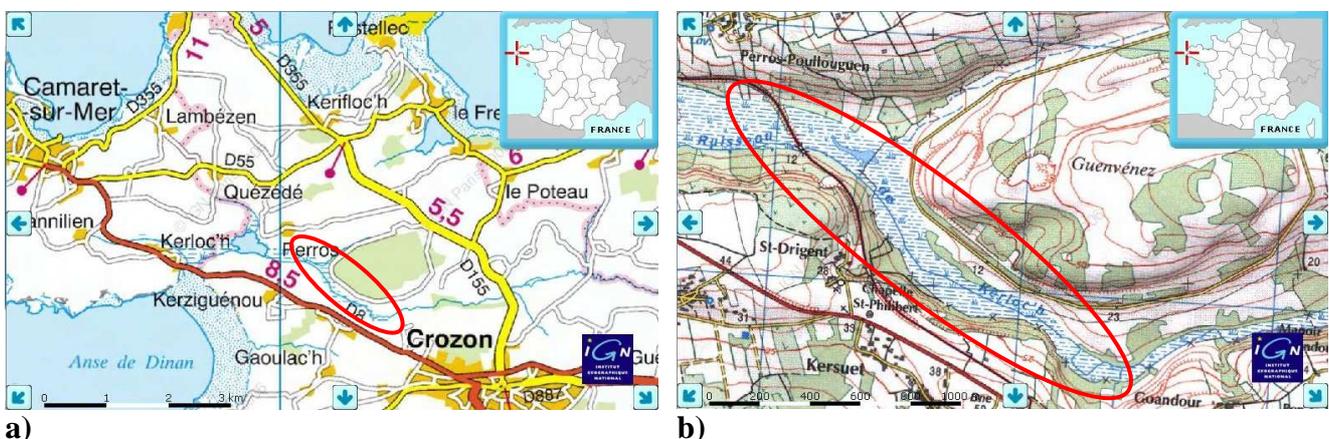


Figure 4 : Exemple de différences de géométries pour un fleuve selon l'échelle : a) une ligne (au Nord de Crozon) et b) un polygone (avec un zoom plus fort sur le fleuve) (Source : <http://www.geoportail.fr/>)

L'exemple de la figure 4 représente un même fleuve dont la géométrie diffère (voir ellipse rouge) selon l'échelle et le type de carte considérés : une ligne au 1/ 50 000 et une surface au 1/25 000 (Scan 25 de l'IGN). La comparaison de ces deux objets n'est alors possible qu'après avoir effectué auparavant une squelettisation de la surface [Mustière, 1995]]. Par ailleurs, il est à noter qu'au sein d'une même carte un objet peut avoir une sémiologie graphique évolutive. Un fleuve par exemple est dessiné par une ligne à sa source, là où il est le plus petit. A l'inverse dans un estuaire, où son étendue est plus importante, il est représenté par une surface.

Au niveau des attributs thématiques, il existe aussi des disparités illustrées dans la figure 5 : les trois objets dans la base de données 1, ayant un attribut « type de cours d'eau » de valeur fleuve, sont décrits comme un « cours d'eau important » ou un « cours d'eau moyen » dans la seconde base. Les dissemblances entre les attributs proches et l'absence d'identifiant commun obligent l'utilisateur à favoriser les outils de comparaison géométrique pour les données géographiques.

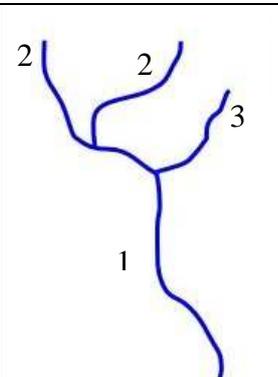
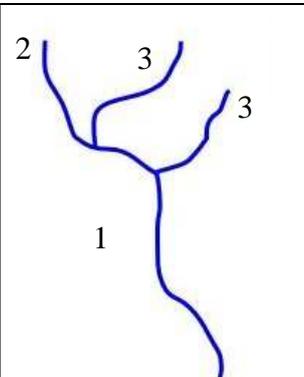
Attributs thématiques BD1 :		Attributs thématiques BD2 :	
1 : Fleuve 2 : Rivière 3 : Torrent		1 : Cours d'eau important 2 : Cours d'eau moyen 3 : Petit cours d'eau	

Figure 5 : Exemple de tables attributaires thématiques différentes d'une base de données à l'autre.

Un phénomène du monde réel peut être représenté par un objet, un ensemble d'objets ou une partie d'un objet plus général. Cette diversité oblige à compléter les outils de comparaison totale par des outils de comparaison partielle et des outils de comparaison entre un objet et un ensemble d'objets.

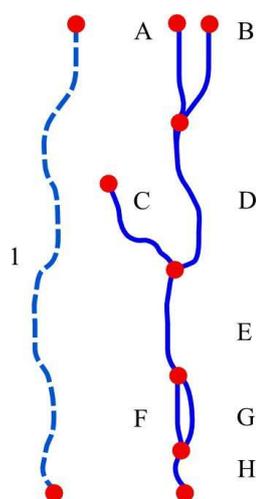


Figure 6 : Représentation d'un même fleuve dans deux bases géographiques avec des relations topologiques variables (absences de nœuds)

Pour ces derniers outils les relations topologiques sont utiles pour regrouper les objets. Le terme relation topologique s'emploie pour évoquer l'organisation spatiale et la connectivité des points (nœuds) et arcs (entre deux nœuds) des lignes qui constituent la base de données. Dans le cas d'un objet linéaire, le processus de comparaison mesure les différences ou les similarités entre cet objet et un objet composite formé d'une suite d'arcs et de nœuds de l'autre base. La figure 6 représente deux fleuves, l'un sous la forme d'un unique objet, l'autre sous la forme de plusieurs arcs.

Dans cet exemple, l'objet 1 peut être comparé avec l'un des objets composites suivants : ADEGH, ADEFH, BDEGH et BDEFH.

Pour recenser l'ensemble de ces disparités, il existe des fonctions de comparaison qui renvoient soit une distance (d'autant plus importante que la différence entre les objets comparés est grande) soit une similitude. Cette dernière se présente sous forme d'une valeur comprise le plus souvent entre 0 (aucune similitude) ou 1 (objets complètement identiques).

I.2.1.3 Appariement de deux jeux de données

La notion d'appariement est liée à celle d'homologie. Des objets, issus de divers jeux de données, sont homologues s'ils représentent le même phénomène du monde réel. **L'appariement**, aussi appelé « conflation » ou « data matching » en anglais, est le processus qui consiste à établir les correspondances entre ces objets homologues [Devoegele, 1997]. Cette mise en relation s'appuie sur les outils de comparaison et elle peut être soit totale entre les deux objets, soit partielle.

Ce processus comprend de nombreuses applications et s'adapte aussi bien à la reconnaissance automatique d'empreinte digitale, qu'au principe du morphing (passage continu d'une image à une autre, [Sederberg & Greenwood, 1992]) ou à la mise à jour de bases de données géographiques [Mustière & Devoegele, 2008].

L'appariement géométrique s'applique sur la comparaison de la forme et de la position des objets. **L'appariement thématique** consiste à rechercher les objets homologues grâce aux valeurs des informations thématiques portées par leurs attributs [Bel Hadj Ali, 2001].

L'appariement topologique utilise les relations topologiques et est souvent utilisé pour mettre en correspondance les données de type réseau ([Gabay & Doytcher, 1994], [Dimitrijévic, 2000]).

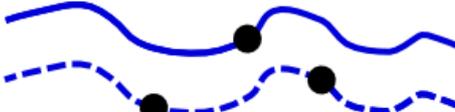
Type de relation		
1-1		BD 1
		BD 2
1-n		
n-m		

Figure 7 : Illustration des différents types de relations entre bases de données.

Le propos de cette thèse étant l'appariement et la fusion finale de deux types de MNT différents, nous ne détaillerons pas l'appariement thématique ni topologique, tous deux peu adaptés à un modèle numérique de terrain à l'heure actuelle.

L'absence fréquente d'identifiant commun entre les bases, les différences de représentation, les imprécisions de positionnement géographique ou bien de description des phénomènes font qu'aucune solution simple et générique n'existe et qu'un seul critère de comparaison ne suffit pas. Le type d'appariement entre objets peut être : 1-1, 1-n, n-m.

Un appariement 1-1, signifie qu'à un objet du jeu de données 1 correspond un et un seul objet du jeu de données 2. La relation 1-n signifie qu'à un objet du jeu de données 1 équivalent plusieurs objets du jeu de données 2 tandis que n-m indique qu'avec plusieurs objets du jeu de données 1 coïncident plusieurs objets du jeu de données 2 (voir figure 7).

Les processus d'appariement se basent rarement sur une seule méthode de comparaison pour déterminer si les objets sont homologues ou non. En effet une seule mesure ne suffit pas à lever les indéterminations dues aux disparités présentées dans la partie comparaison.

Par ailleurs, il existe également le pré-appariement : ce processus consiste en une pré-sélection des objets potentiellement homologues parmi les deux jeux de données. Ceci est souvent réalisé à l'aide de mesures de zones (comme l'emprise abordée précédemment), qui permet de sélectionner parmi un jeu de données les candidats possibles, puis l'appariement est finalisé grâce à des mesures de distance et/ou de forme (détaillées ci-après).

I.2.1.4 Fusion de données

Le terme fusion désigne un ensemble de processus qui consiste à partir de jeux de données multi sources à n'en obtenir plus qu'un seul : le jeu fusionné, possédant une information plus riche que les deux jeux d'origine.

Le résultat du processus de fusion est de :

- supprimer la redondance ;
- résoudre les problèmes d'hétérogénéité ;
- regrouper les informations complémentaires apportées par chacune des données afin de produire une description plus exhaustive du phénomène représenté.

Il existe essentiellement trois applications de la fusion :

- soit la propagation de mises à jour d'une base de données à l'autre à l'aide des outils d'appariement [Bel Hadj Ali, 2001],
- soit la fusion de la zone de recouvrement de deux jeux afin d'obtenir un modèle continu (base de donnée, modèle numérique, etc.),
- soit l'enrichissement d'un jeu de données à l'aide d'un autre qui dispose d'informations complémentaires ([Rousseaux, 2005], [Podobnikar, 2005]).

Une fusion de jeux donnés de recouvrement total a pour but de produire une base de données multi représentations homogène au niveau des attributs recopiés et sans objets redondants. En géographie, une *représentation* est une interprétation et une décomposition du monde réel en bases de données exploitables par des SIG [Denègre & Salgé, 1996].

Si le recouvrement n'est pas total il s'agit alors de faire une fusion dans la zone commune et de conserver les données d'origine dans les zones propres à chaque jeu en assurant une transition continue entre les trois zones. Prenons l'exemple de deux bases de données, la première contenant les thalwegs issus d'une base de données géographique terrestre (en rouge sur la figure 8) et la seconde leur prolongement dans le domaine marin proche (au niveau de leur embouchure, figure 8). Les deux bases ont des emprises différentes qui se recouvrent au niveau de l'estran. Le résultat de la fusion permet d'avoir un jeu final d'une emprise plus large et continu d'un domaine à l'autre.

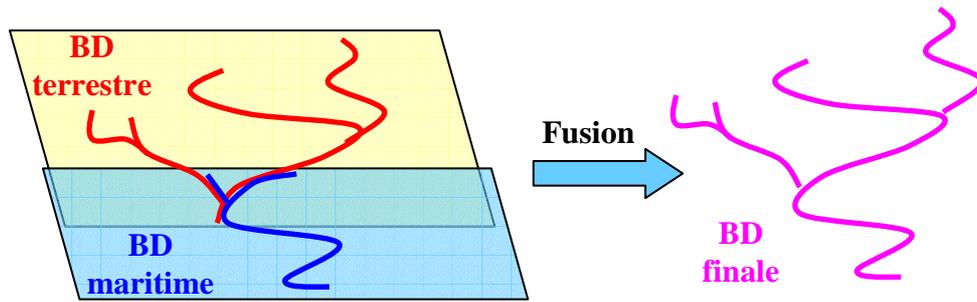


Figure 8 : Exemple de fusion entre deux bases de données terrestre et maritime.

En conclusion, le processus d'intégration va permettre d'apparier entre eux les éléments de différentes bases de données grâce aux résultats de leur comparaison afin de les fusionner en dernier lieu. Les différentes étapes sont les suivantes : premièrement des outils de comparaison, basés sur des mesures de distances et/ou des valeurs d'emprise, sont utilisés. Ensuite vient l'appariement de tout ou d'une partie des géométries issues des deux bases qui représentent les mêmes phénomènes du monde réel. Enfin le modèle final est obtenu en fusionnant les objets redondants au niveau de la zone de recouvrement des deux bases. Finalement les parties non appariées sont recalées pour que l'ensemble reste cohérent.

I.2.2 Intégration de points

I.2.2.1 Outils d'appariement de points

Les méthodes d'appariement actuelles sont pour la majeure partie fondées soit sur l'appariement des objets les plus proches (appariement géométrique) soit sur l'appariement thématique. Seuls les appariements géométriques basés sur le positionnement d'un objet seront détaillés ici. Cet état de l'art s'appuie majoritairement sur l'article publié par [Beeri *et al.*, 04] sur le sujet.

L'appariement géographique entre deux ensembles ponctuels se base sur la distance euclidienne afin de déterminer le voisin le plus proche. La principale méthode utilisée pour l'appariement géométrique est celle du « **one-sided nearest neighbor join** ». Celle-ci fusionne un objet d'une base de données avec l'objet dans l'autre base le plus proche géographiquement.

Ce type d'appariement pose un premier problème dans le cas où la localisation dans une des deux bases est entachée d'un biais et un second si l'appariement entre objets est de type « 1-n ». Par ailleurs, l'auteur démontre la non symétrie des résultats obtenus : un appariement A-B ne donnera pas les mêmes résultats qu'un appariement B-A.

Trois autres méthodes d'appariement ont été définies par [Beeri *et al.*, 2004] de manière à présenter de meilleurs résultats au niveau de la symétrie quel que soit le taux de recouvrement tout en autorisant des relations de type 1-2. En complément, deux autres notions sont définies :

- un facteur de choix (« choice factor ») qui est le nombre d'objets dans un cercle d'investigation dont le rayon d'action est égal à l'intervalle d'erreur.
- le degré de recouvrement qui représente la mesure de la fraction d'objets qui possèdent un homologue dans l'autre jeu.

La première méthode utilisée, appelée « **mutually nearest method** », est basée sur l'idée que deux objets sont homologues s'ils sont mutuellement les plus proches. L'avantage principal de cette méthode est une faible sensibilité au degré de recouvrement entre données, en

revanche dans le cas d'un facteur de choix important, les résultats se montrent décevants puisque seulement deux voisins proches au maximum sont pris en compte. Par ailleurs, les tests ont montré qu'elle donnait de meilleurs résultats pour un pourcentage d'objets communs (ou degré de recouvrement) faible.

La méthode probabiliste (« **probabilistic method** »), troisième méthode décrite pour la fusion de jeux de points, s'adapte beaucoup mieux lorsque le facteur de choix est large. Par contre elle nécessite un recouvrement important des données.

La dernière méthode introduite, « **Normalized-Weights-Method** » voit son efficacité optimale pour des recouvrements moyens à petits. Cette approche probabiliste permet de proposer une méthode d'appariement qui après améliorations donne les meilleurs résultats quels que soient le type, la densité et le degré de recouvrement des données. Néanmoins, les appariements « 1-n » et « n-m » entre bases ne peuvent pas être gérés.

I.2.2.2 Fusion de points

La fusion entre des éléments ponctuels est une fonction qui calcule une somme : soit une moyenne des points homologues des deux jeux, soit une conservation des points les meilleurs (indice de qualité ou précision plus grande). Dans les deux cas une fonction supprime les redondances entre attributs des points modifiés puis recalcule l'ensemble des points non appariés de façon à ce que le jeu fusionné reste cohérent avec les jeux de points de départ.

I.2.3 Intégration de lignes

Après avoir détaillé les méthodes existantes pour l'intégration de points, des outils d'intégration de lignes vont maintenant être présentés. Ceux-ci comprennent essentiellement la comparaison, basée sur des mesures de distance, l'appariement et la fusion qui inclut une étape de déformation élastique.

I.2.3.1 Outils de comparaison de lignes

Les distances représentent les mesures les plus adéquates pour comparer des objets et donc évaluer leur similitude car elles permettent d'estimer les écarts moyen, minimal et maximal entre eux. En ce qui concerne le processus d'appariement, qui s'appuie sur ces valeurs, l'écart minimal n'a que peu d'intérêt. Nous détaillerons donc les notions de distance moyenne et de distance maximum.

Distance moyenne

Il existe plusieurs distances moyennes, dont une définie dans cette thèse au chapitre IV. Un exemple de distance moyenne proposé par [McMaster, 1986] est la surface entre deux lignes homologues, divisée par la longueur de référence. La surface est définie à partir des deux lignes à comparer, du segment reliant les nœuds initiaux et du segment reliant les nœuds finaux. Cette surface est divisée par la longueur de la ligne en référence pour rendre cette mesure indépendante de cette dernière.

Cette mesure peut être rendue symétrique en divisant la surface par la moyenne des longueurs des arcs. La distance moyenne entre les deux lignes est ainsi obtenue.

La distance moyenne nous donne une idée de l'écart moyen entre deux lignes et donc de la précision globale de l'une par rapport à l'autre. Cependant, elle doit toujours être couplée avec une distance maximale (distance de Hausdorff ou distance de Fréchet), sinon elle risque de donner des résultats erronés [McMaster, 1986].

Distance de Hausdorff

Par définition, la distance de Hausdorff entre deux lignes (L_1 , L_2) est l'écart maximum entre celles-ci [Hausdorff, 19]. Elle représente la plus grande des deux composantes :

- d_1 qui est la plus grande distance parmi les distances minimales entre les points de L_1 et le point le plus proche de L_2 ,
- d_2 qui est la plus grande distance parmi les distances minimales entre les points de L_2 et le point le plus proche de L_1 .

$$d_1 = \max_{p_1 \in L_1} \left[\min_{p_2 \in L_2} [dist(p_1, p_2)] \right]$$

$$d_2 = \max_{p_2 \in L_2} \left[\min_{p_1 \in L_1} [dist(p_2, p_1)] \right]$$

$$d_H = \max(d_1, d_2)$$

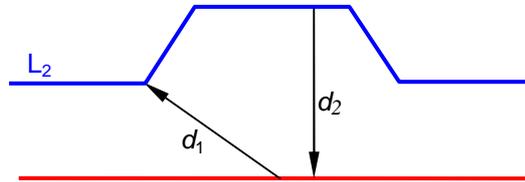


Figure 9 : Définition et illustration de la distance de Hausdorff.

Ces deux distances, d_1 et d_2 , ne sont pas forcément identiques ($d_1 > d_2$ pour la figure 9), surtout si les lignes n'ont pas la même emprise.

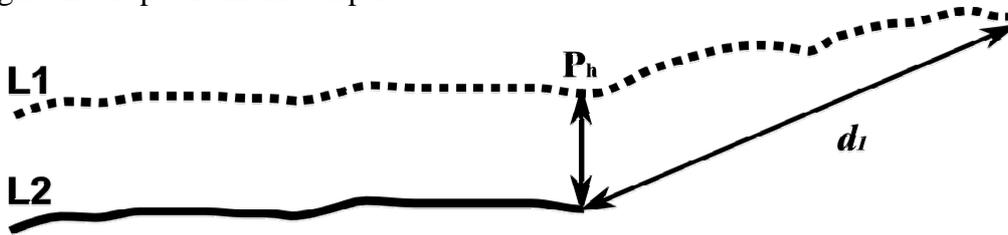


Figure 10 : Distance de Hausdorff entre deux lignes d'emprises différentes.

La figure 10 illustre le cas où les deux lignes possèdent des tailles différentes : une des deux composantes de la distance de Hausdorff (d_1) calculées n'est révélatrice que de l'écartement maximal des extrémités des lignes et non de l'écartement maximal de points homologues des lignes. Pour cet exemple, si la valeur d_1 est la seule prise en compte, les lignes sont considérées comme dissemblables. Si la composante d_1 de la distance de Hausdorff est calculée à partir d'une partie de L_1 homologue à l'extrémité de L_2 (partie à gauche du point « P_h »), le résultat sera bien meilleur.

De la même manière, la non prise en compte de l'orientation des lignes fait que la distance de Hausdorff n'est pas très appropriée pour la comparaison de lignes complètement différentes mais voisines géographiquement.

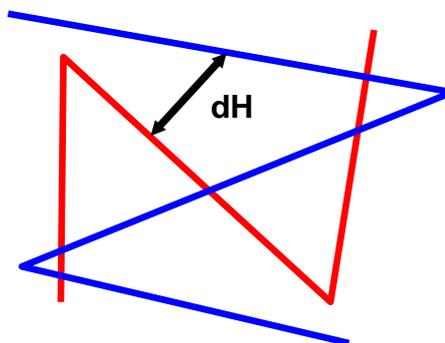


Figure 11 : Problème de la distance de Hausdorff avec des lignes sinueuses.

Les deux lignes présentées en figure 11 possèdent une rotation de l'une par rapport à l'autre de 90° environ, pourtant la mesure de la distance entre les deux est faible. Au vu uniquement de cette valeur dH , ces lignes peuvent être considérées comme identiques alors qu'elles ne se

ressemblent pas du tout. d_H ici n'a pas de signification réelle sur l'état de similitude des lignes car elle n'est pas calculée sur les points homologues mais les points les plus proches. Dans des cas similaires de lignes sinueuses, cette distance ne paraît pas appropriée.

Distance de Fréchet et dérivées

La distance de Fréchet ainsi que les autres distances qui en ont été dérivées vont maintenant être définies.

La distance de Fréchet :

Elle se base sur l'assimilation de toute ligne comme une suite de points orientés équivalente à une fonction continue $f : [a, b] \rightarrow V$ où $a, b \in \mathfrak{R}$, $a < b$ et V est l'espace vectoriel. La distance de Fréchet (d_F) [Fréchet, 1906] est la suivante :

Soient $f : [a, a'] \rightarrow V$ et $g : [b, b'] \rightarrow V$ deux polygones et $\| \cdot \|$ la norme usuelle,

$$d_F(f, g) = \inf_{\substack{\alpha: [0,1] \rightarrow [a, a'] \\ \beta: [0,1] \rightarrow [b, b']}} \max_{t \in [0,1]} \|f(\alpha(t)) - g(\beta(t))\| \quad (1)$$

Une illustration intuitive de cette distance peut être donnée par l'exemple d'un maître se promenant avec son chien en laisse. Chacun suit son chemin propre, s'arrête et avance à volonté mais ne peut pas revenir en arrière. La distance de Fréchet entre ces deux chemins est la plus petite longueur de laisse qui permet aux deux protagonistes d'avancer en même temps en se pliant aux impératifs précédemment énoncés.

Une difficulté majeure à son application vient de sa complexité à programmer (algorithme d'ordre $O(N M \log^2(N M))$ avec N et M les nombres de segments des polygones [Alt & Gadau, 1995]). Le développement de la distance de Fréchet discrète a résolu ce problème et a aussi permis la naissance d'extensions pour cette distance.

La distance de Fréchet discrète :

Eiter et Mannila [Eiter & Mannila, 1994] donnent une approximation de la distance de Fréchet réduite à un algorithme d'ordre $O(N M)$: la distance de Fréchet discrète (d_{dF}).

Les lignes L_1 et L_2 sont discrétisées à l'aide d'une suite de points : $\langle L_{1,1} \dots L_{1,n} \rangle$ et $\langle L_{2,1} \dots L_{2,m} \rangle$.

Pour reprendre notre illustration précédente, d_{dF} représenterait la longueur de la laisse minimale telle qu'un chemin depuis le couple de points de départ $(L_{1,1}, L_{2,1})$ jusqu'au couple de points de fin $(L_{1,n}, L_{2,m})$ existe. Le chemin donne une suite ordonnée de couples $(L_{1,i}, L_{2,j})$ telle que le couple suivant celui-ci soit l'un de ces trois couples :

- $(L_{1,i+1}, L_{2,j+1})$ le maître et son chien avancent
- $(L_{1,i+1}, L_{2,j})$ seul le maître se déplace
- $(L_{1,i}, L_{2,j+1})$ seul le chien avance.

d_{dF} est une bonne estimation de d_F car l'approximation est limitée par la distance maximale entre deux points consécutifs (LongMaxSeg) [Eiter & Mannila, 1994] :

$$d_F(L_1, L_2) \leq d_{dF}(L_1, L_2) \leq d_F(L_1, L_2) + \text{LongMaxSeg} \quad (2)$$

Comme la valeur de la distance de Fréchet partielle calculée dépend de la longueur du segment entre deux points (LongMaxSeg), il est possible de limiter l'erreur d'approximation

à un epsilon (ϵ) en ajoutant des points aux segments de ligne par ré-échantillonnage des jeux de données. Ainsi la longueur du plus grand segment est inférieure ou égale à ϵ .

La distance de Fréchet discrète entre L_1 et L_2 est calculée récursivement comme suit :

$$d_{Fd}(L_1, L_2) = \max \left(\begin{array}{l} d_E(L_{1,n}, L_{2,m}) \\ \min \left(\begin{array}{l} d_{Fd}(\langle L_{1,1} \dots L_{1,n-1} \rangle, \langle L_{2,1} \dots L_{2,m} \rangle) \forall n \neq 1 \\ d_{Fd}(\langle L_{1,1} \dots L_{1,n} \rangle, \langle L_{2,1} \dots L_{2,m-1} \rangle) \forall m \neq 1 \\ d_{Fd}(\langle L_{1,1} \dots L_{1,n-1} \rangle, \langle L_{2,1} \dots L_{2,m-1} \rangle) \forall n \neq 1, m \neq 1 \end{array} \right) \end{array} \right) \quad (3)$$

$\langle L_{1,1} \dots L_{1,n-1} \rangle$ et $\langle L_{2,1} \dots L_{2,m-1} \rangle$ représentent les lignes. Le calcul prend fin lorsque les deux lignes sont réduites à deux points singuliers $\langle L_{1,1} \rangle$ et $\langle L_{2,2} \rangle$.

L'exemple illustré figure 12 présente un calcul de la distance de Fréchet discrète entre deux lignes. La première ligne L_1 est composée de 8 points (et $N = 8$, de $L_{1,1}$ à $L_{1,8}$) et la deuxième ligne L_2 de 7 (et $M = 7$ de $L_{2,1}$ à $L_{2,7}$).

Le calcul des deux matrices, de dimension $N \times M$ (voir), remplace le traitement récursif. Ces matrices représentent :

- La matrice de distances euclidiennes (MD). Chaque cellule $MD_{i,j}$ contient la valeur de la distance entre $L_{1,i}$ et $L_{2,j}$
- La matrice de Fréchet (MF) qui permet de calculer itérativement la distance de Fréchet [Eiter et Mannila, 94]. La formule de calcul se présente ainsi :

$$MF_{i,j} = \max (d_E(L_{1,i}, L_{2,j}), \min(MF_{i-1,j}, MF_{i,j-1}, MF_{i-1,j-1})) \quad (4)$$

La distance de Fréchet discrète (d_{dF}) est la valeur de la cellule $MF_{n,m}$.

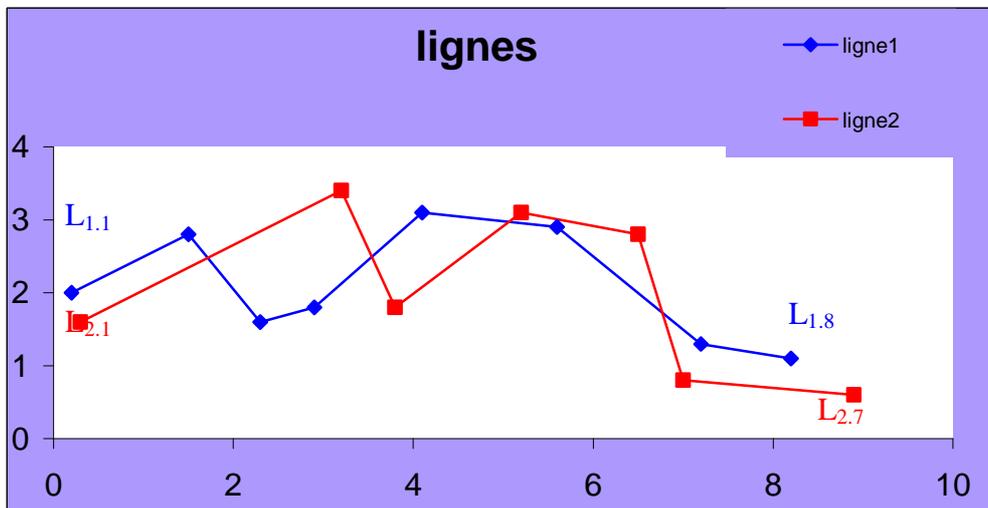


Figure 12 : Exemple de couple de lignes.

Le tableau 1 montre les Matrices de Distances (MD) et de Fréchet (MF) calculées à partir des lignes de la figure 12. La d_{dF} entre ces deux lignes est de : 1,80. Les cellules en jaune dans la matrice de Fréchet sont les valeurs minimales de distance calculées entre deux points des lignes L_1 et L_2 .

La distance de Fréchet discrète est égale à la distance euclidienne entre les points homologues $L_{1,2}$ et $L_{2,2}$. Pour ces deux lignes, la distance de Hausdorff est certes plus petite (<1), mais pas

significative en terme d'appariement de points homologues. En effet celle-ci est calculée entre les points $L_{1,6}$ et $L_{2,4}$ (voir figure 12) qui sont les plus proches car les deux lignes sont décalées l'une par rapport à l'autre.

Tableau 1 : Matrice de distance (euclidiennes) et matrice de Fréchet des lignes de la Figure 12.

x1i	0.2	1.5	2.3	2.9	4.1	5.6	7.2	8.2	
y1j	2	2.8	1.6	1.8	3.1	2.9	1.3	1.1	
Matrice de Distance (MD)									
x2i	y2j	1	2	3	4	5	6	7	8
0.3	1.6	0.41	1.70	2.00	2.61	4.09	5.46	6.91	7.92
3.2	3.4	3.31	1.80	2.01	1.63	0.95	2.45	4.52	5.50
3.8	1.8	3.61	2.51	1.51	0.90	1.33	2.11	3.44	4.46
5.2	3.1	5.12	3.71	3.26	2.64	1.10	0.45	2.69	3.61
6.5	2.8	6.35	5.00	4.37	3.74	2.42	0.91	1.66	2.40
7	0.8	6.91	5.85	4.77	4.22	3.70	2.52	0.54	1.24
8.9	0.6	8.81	7.72	6.68	6.12	5.41	4.02	1.84	0.86
		1	2	3	4	5	6	7	8
1	0.41	1.70	2.00	2.61	4.09	5.46	6.91	7.92	
2	3.31	1.80	2.01	2.00	2.00	2.45	4.52	5.50	
3	3.61	2.51	1.80	1.80	1.80	2.11	3.44	4.46	
4	5.12	3.71	3.26	2.64	1.80	1.80	2.69	3.61	
5	6.35	5.00	4.37	3.74	2.42	1.80	1.80	2.40	
6	6.91	5.85	4.77	4.22	3.70	2.52	1.80	1.80	
7	8.81	7.72	6.68	6.12	5.41	4.02	1.84	1.80	

Conclusion sur les distances linéaires

La distance moyenne est très utile pour produire une estimation de l'écart moyen global de deux lignes, cependant aucune n'a été définie pour les points homologues. Par ailleurs elle doit toujours être couplée à une distance maximale afin de savoir si un écart fort l'est en un seul endroit (erreur de mesure ponctuelle) ou s'il est moyen sur l'ensemble des lignes (biais de positionnement par exemple).

En ce qui concerne les distances maximales (Hausdorff et Fréchet), la distance de Fréchet a l'avantage d'être calculée à partir des points homologues. La distance de Hausdorff, elle, ne travaille qu'avec les points les plus proches (qui ne sont pas toujours homologues). Pour cette dernière des difficultés apparaissent lors de la comparaison de lignes sinueuses. En effet la mesure ne reflète pas toujours la similitude des lignes. Ceci pose problème dans le cas de cette thèse qui utilise comme base des données géographiques naturelles. Ces données sont marquées par une complexité et un aspect de type fractal fort (le trait de côté Finistérien par exemple).

Outils de comparaison de formes pour des objets linéaires

En complément de l'utilisation des distances, d'autres critères ont été employés afin de déterminer la similitude entre deux objets linéaires. Ceux-ci se basent sur la forme géométrique même des lignes et non plus sur la mesure de leur écartement moyen ou maximal. [McMaster, 1986] et [Mustière, 1995] définissent comme critères :

- le rapport entre les longueurs des arcs des deux lignes ;
- le rapport des longueurs moyennes des virages ;

- la différence entre les directions de chaque segment ;
- le rapport du nombre de points intermédiaires ;
- le rapport des sommes des angles entre les segments (angles exprimés entre 0 et π) ;
- le rapport du nombre de virages (ce dernier est défini comme le sous-arc compris entre deux points d'inflexion).

Plusieurs autres mesures additionnelles ont été présentées dans [Plazanet, 1996].

Outils de comparaison par zone définie

Une troisième voie pour la comparaison consiste à employer des zones représentant les emprises des objets pour chaque ligne et à les comparer entre elles. Ces zones sont calculées à partir des objets eux-mêmes de différentes manières : il peut s'agir par exemple d'un rectangle incluant l'objet, d'une zone définie par l'ensemble convexe de taille minimale qui contient la géométrie (enveloppe convexe), etc. Les rectangles englobants sont les zones les plus utilisées. Seules les zones qui sont employées dans cette thèse sont présentées dans les paragraphes suivants.

Rectangle englobant

Le rectangle englobant se définit comme le rectangle incluant l'objet en entier. Il peut être calculé de deux manières :

- soit ses axes sont parallèles aux axes des abscisses et des ordonnées (voir figure 13, cadre noir en pointillés), plus courant car plus simple à calculer,
- soit ses axes sont parallèles à la direction principale de l'objet (voir figure 13, cadre noir en trait continu) et dans ce cas il constitue le rectangle englobant minimum [Devoegele, 1997].

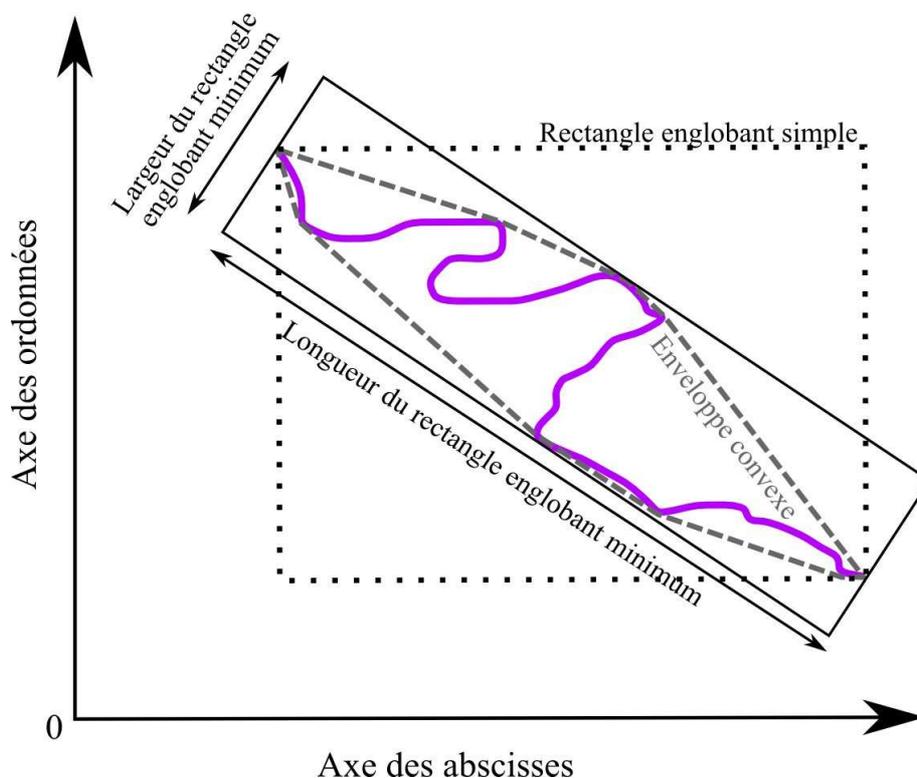


Figure 13 : Rectangle englobant simple et rectangle englobant minimum calculés à partir d'un objet. En comparaison l'enveloppe convexe a également été représentée.

Le rectangle englobant minimum permet de savoir rapidement si deux objets ont des emprises équivalentes.

Bande epsilon

A chaque point ou segment constitutif de la ligne, est assignée une zone de tolérance circulaire dont le rayon varie selon la nature du point représenté. Les cercles sont ensuite associés par leur tangentes communes afin de former une bande de tolérance [Perkal, 1956] : la bande epsilon. Cette technique a l'avantage de permettre un appariement partiel des lignes au niveau des points ou des segments et non uniquement en totalité.

La zone tampon est un cas particulier de cette bande [Perkal, 1956]. Elle est déterminée par une ligne et une distance d à celle-ci. L'objet linéaire à comparer devra donc appartenir à cette zone tampon pour être considéré comme homologue. Cette zone est assez adaptée pour modéliser finement l'imprécision des jeux de données.

L'association rectangle + zone tampon permet de mieux tenir compte lors d'un processus de comparaison du degré d'incertitude et d'imprécision des bases de données entre elles.

La comparaison à l'aide de zones définies s'effectue au travers de deux mesures : une relation d'inclusion ou un pourcentage de recouvrement. La première consiste à vérifier s'il existe bien une relation entre la zone d'un objet et la zone élargie de l'autre. Si celle-ci est réciproque, alors les objets ont même extension géographique, sinon l'objet contenu a une taille plus petite. Pour le pourcentage de recouvrement [Bel Hadj Ali, 1997] a proposé différentes méthodes de calcul. Lorsqu'il s'agit d'un recouvrement total (très proche de 100%), les objets comparés sont potentiellement similaires. Pour une valeur inférieure, il faut arrêter une valeur seuil au-delà de laquelle l'écart entre les objets est trop grand pour qu'ils puissent être considérés comme identiques (voir figure 3(a), cas des emprises). Par ailleurs, dans le cas d'objets partiellement similaires, il faut utiliser des indicateurs capables de distinguer comparaisons globale et locale entre deux géométries.

En conclusion, le rectangle englobant est rapide à calculer, mais donne un résultat assez peu précis : deux lignes de même taille formant un « X » ont le même rectangle englobant mais n'en sont pas pour autant identiques. L'enveloppe convexe permet elle d'être plus proche de la forme de l'objet et d'éviter ainsi ce type de confusion, seulement elle présente également un temps de calcul plus long. De ces remarques nous en déduisons que le rectangle englobant (simple ou non) sert de première approximation rapide et simple à mettre en place. Il constitue en effet un très bon outil de pré-appariement qui permet de sélectionner les candidats potentiellement homologues. Cependant il doit être complété par une autre mesure (distance maximale par exemple) pour pallier certaines configurations ambiguës entre les objets.

I.2.3.2 Outils d'appariement géométrique de lignes

Afin de déterminer si deux objets linéaires géographiques représentent le même phénomène du monde réel, le processus doit tenir compte des imprécisions de positionnement et de forme. Pour cela, quatre catégories d'appariement se complètent :

- appariement par zone définie
- appariement via des mesures de distances entre objets (basé sur les mesures issues du processus de comparaison de lignes)
- appariement par l'utilisation de la forme géométrique propre des objets linéaires (grâce aux outils de comparaison de formes détaillés plus haut)
- appariement par méthodes probabilistes.

Appariement par zone définie :

Ce type d'appariement utilise les mesures issues de la comparaison des zones définies à partir des objets contenus dans les bases de données géographiques. Les plus utilisées sont celles calculées à partir des rectangles englobants (minimaux et élargis d'une bande epsilon) ou des enveloppes convexes. Il en existe également d'autres comme celles basées sur les diagrammes de Voronoï [Arnaud & Emery, 2000]. Deux objets sont candidats à l'appariement si :

- le pourcentage de recouvrement entre eux est suffisant
- un des rectangles englobants simples de l'objet 1 est inclus dans le rectangle englobant élargi d'une bande epsilon de l'objet 2.

Le type d'appariement peut être défini en fonction de la relation d'inclusion. Si l'inclusion est vérifiée alors l'objet inclus est de taille inférieure. Par contre si l'inclusion est réciproque alors les objets sont de taille similaire ou de type 1-n si un élément d'une base en contient plusieurs de l'autre base.

Appariement à l'aide de distances :

Ce processus utilise les distances calculées pour la comparaison des lignes à savoir :

- distance de Fréchet
- distance de Hausdorff

Le principe est d'utiliser les mesures fournies par ces diverses distances afin de déterminer si une ligne est candidates ou, dans des cas d'appariement partiel, si une partie de la ligne est candidate à l'appariement avec tout ou une partie d'une autre ligne.

Appariement basé sur la ressemblance de forme :

Ce type d'appariement s'appuie sur les mesures obtenues grâce aux outils de comparaison de formes décrits précédemment. Tous ces critères (rapport de longueurs, de sommes, différence entre directions des lignes) aident à l'appariement d'objets selon leurs ressemblances, mais seulement pour des relations de type 1-1, ce qui interdit les cas d'appariements partiels. Pour cette raison des mesures de distances et/ou de zones leur sont souvent associées afin de prendre en compte les cas d'appariement 1-n.

Combinaison des différentes techniques d'appariement :

En général une catégorie d'appariement seul ne suffit pas pour prendre en compte les différents cas, 1-1 (total), 1-n, n-m (partiels) et il est nécessaire de combiner plusieurs mesures.

[Devogele, 1997], [Mustière & Devogele, 2008] proposent une pré-sélection des candidats potentiels à l'aide d'appariements par zone définie. Des distances maximales sont alors calculées entre les candidats et l'objet à apparier. Finalement, des critères topologiques sont employés pour définir les appariements 1-1 et 1-n.

Des approches **probabilistes** [Servigne, 1993] [Servigne, 1994] [Salmeron & Milgram, 1986] [Jamet & Phalakarn, 1989, Phalakarn, 1991] [Le Men & Jamet, 1990] sont aussi suggérées pour combiner les techniques d'appariement. Ces méthodes calculent une probabilité d'association pour chaque candidat sélectionné selon des critères de recouvrement, d'angle relatif ou autre mesure de comparaison.

I.2.3.3 Fusion et déformation élastique de lignes

Les processus de fusion sont souvent simplifiés à la somme (pondérée ou non) des objets les plus proches parmi ceux des deux bases comparées. Le postulat qui sous tend cette opération est que si la localisation géographique est correcte et que les bases possèdent le même recouvrement alors l'objet de la BD1 homologue de l'objet de la BD2 est à la même position

géographique ou presque. Le processus de fusion prend alors soit un jeu de données en référence et calque le deuxième dessus, soit effectue une somme pondérée des deux.

Ces deux opérations entraînent des déformations des géométries des données d'origine. Afin de garder une représentation cohérente au sein du modèle final, il est nécessaire d'effectuer des déplacements des objets non appariés.

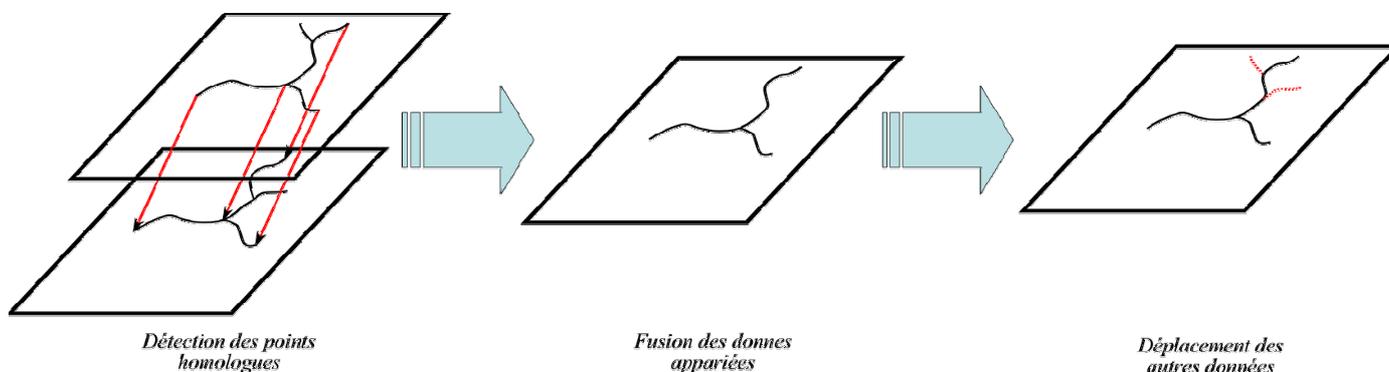


Figure 14 : Etapes de fusion et de déformation élastique de lignes.

Le processus de déformation élastique pallie ce problème en se basant sur des points corrélés spatialement d'une base à l'autre et désignés soit par l'utilisateur soit automatiquement (voir figure 14, détection des points homologues). Cette opération est réalisée par une distorsion non linéaire de manière à forcer l'ajustement d'éléments géographiques (points, lignes) à des positions données. La figure 14 montre les étapes de déformation : les points sont tout d'abord désignés par l'utilisateur puis les données appariées sont fusionnées. De manière à maintenir la cohérence au sein du jeu de données final, les données qui n'ont pas pu être appariées sont déformées de manière à respecter leur géométrie d'origine avec l'ensemble de la base [Airault, 1996]. Si le déplacement global est inférieur à une distance x par exemple, alors la déformation élastique permettra une correction via un lissage, au moins partiel, sinon complet, grâce à des points de contrôle espacés d'une distance inférieure à x . Ces mêmes raisons impliquent que les formes des traits seront d'autant plus déformées que le biais de déplacement est grand par rapport à x .

[TCI, 1999] propose par l'intermédiaire de son application « Adjust » dans AutoCAD de réaliser une déformation semi-automatique. L'utilisateur doit sélectionner manuellement deux jeux de points de calage et les corrélés en utilisant la relation « From-To ». Le processus calcule alors automatiquement un nombre de points homologues intermédiaires sur les deux lignes grâce à deux modes disponibles au choix. Dans cette chaîne de traitement, les points des deux jeux, de départ (= From) et d'arrivée (= To) sont fusionnés en un seul : celui d'arrivée. Le type de relation entre les deux ne peut être que de 1-1 avec cette méthode. De plus, comme le point gardé est celui d'arrivée, les données prises en référence doivent être de meilleure précision que celles de départ.

Une intégration complète de base de données géographique avec emprises limitrophes a été proposée par [Laurini, 1996]. Cette intégration étend la procédure définie dans [Spaccapietra et al., 1992] aux bases de données géographiques. Le principe repris est de mettre en correspondance deux éléments de bases (classes, attributs, objets) s'ils décrivent les mêmes phénomènes du monde réel, ceci indépendamment de leur représentation. Cette identification des éléments homologues des différentes bases est appelée « déclaration des correspondances ». Cette étape d'appariement est complétée par :

- des pré-traitements pour éliminer les problèmes liés aux bases de données,
- un mécanisme de transformation élastique pour le raccordement géométrique aux frontières.

Cette intégration n'a pas été appliquée à des bases de données qui se superposent. Le processus de déformation peut être amélioré en ajoutant une bande de correction élastique. Cette bande applique des modifications plus importantes aux données au voisinage de la zone de recouvrement et nulles à l'autre extrémité de la bande. Il faut utiliser pour cela une fonction de déformation plus fine qui permet d'obtenir des déplacements en fonction de la distance relative à la frontière [Laurini, 1996].

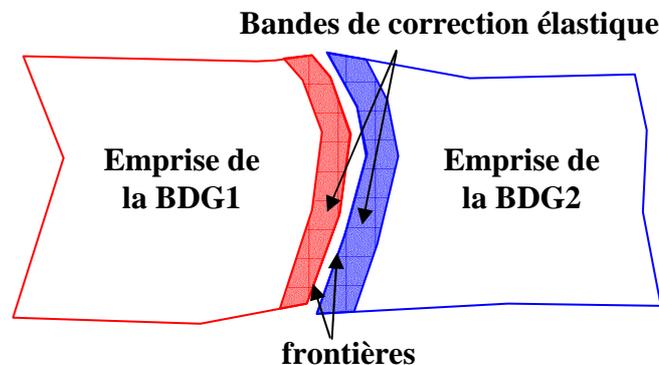


Figure 15 : Exemple de bande de correction élastique (Source : [Devogele, 2002])

Les méthodes de déformation possèdent les inconvénients d'obliger l'utilisateur à sélectionner les points homologues, de modifier tous les points, même ceux non sélectionnés, et de ne pas respecter les géométries de départ (parallélisme, angles).

Dans l'optique de remédier à cela, effectuer une déformation élastique à partir de la distance de Fréchet partielle permet de disposer directement des points homologues calculés lors du processus et ceci de façon automatique [Devogele, 2002]. En revanche les transformations élastiques constituent des fonctions non linéaires : le parallélisme entre les géométries et la rectangularité de certains angles ne sont donc pas conservés.

Une solution est de compléter la déformation par un processus d'équarrissage [Airault, 1996] qui contraint la déformation de manière à respecter les éléments géométriques d'origine comme les angles droits ou le parallélisme.

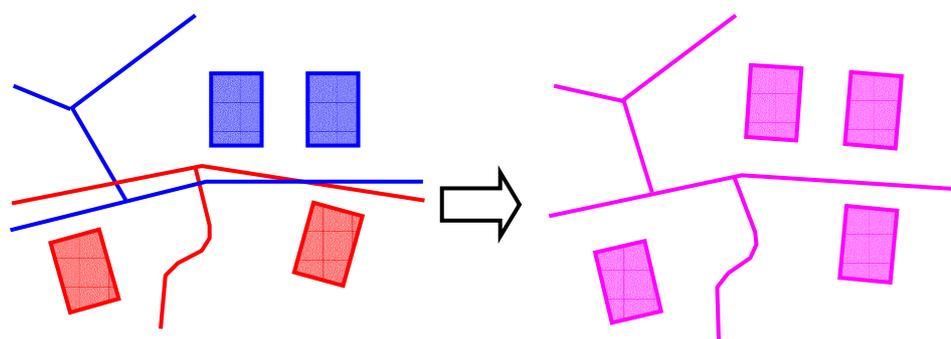


Figure 16 : Exemple de fusion + déformations élastiques (Source : [Devogele, 2000])

Concernant la fusion, la société Geomod a également réalisé un logiciel de mise en cohérence des données du cadastre avec les bases de données de l'IGN (BD Topo®, BD ortho®) dans le cadre du projet de production de la BD Parcellaire® Vecteur. Les données du cadastre sont au format EDIGÉO (norme française permettant de modéliser n'importe quel lot de données). Pour mener à bien cette conversion, il faut mettre les données en cohérence. Pour cela,

l'utilisateur doit en premier lieu effectuer une analyse du paysage afin de maîtriser les déplacements futurs des objets.

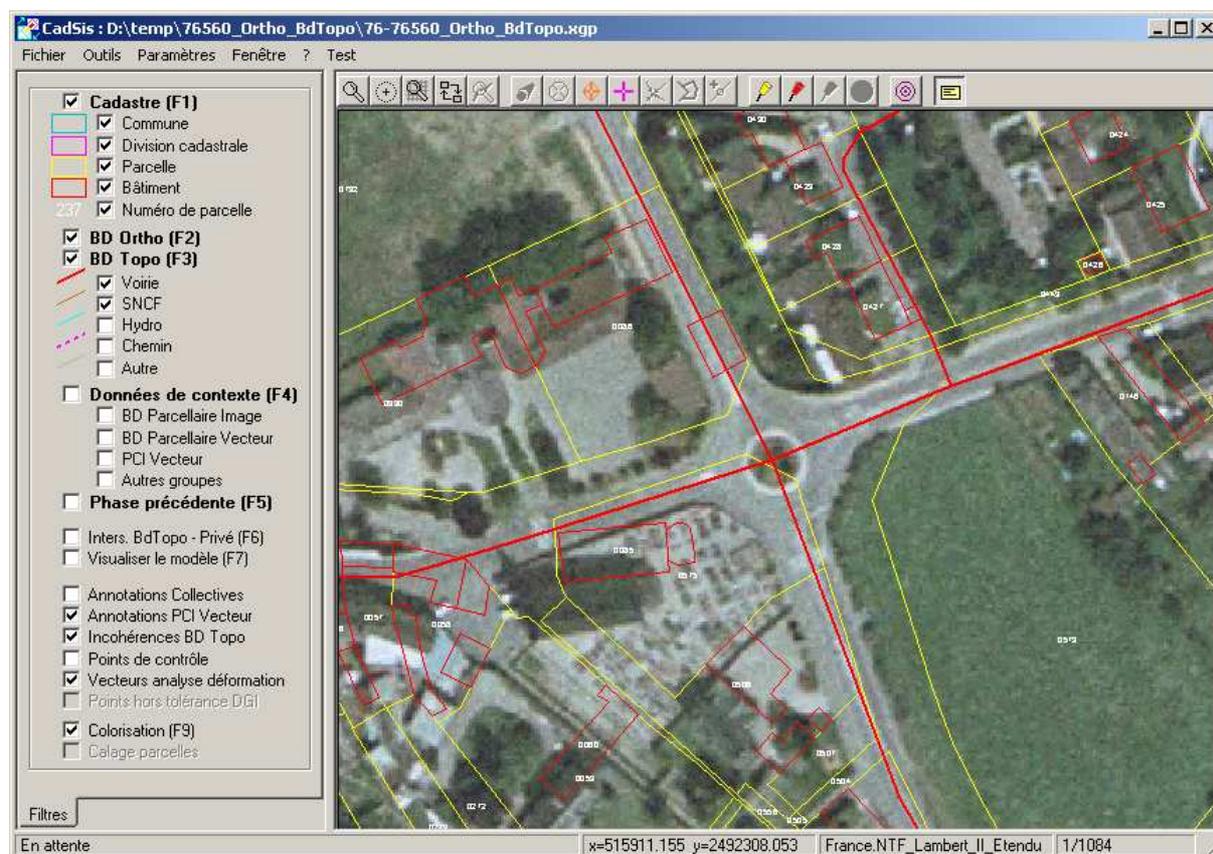


Figure 17 : Superposition des données du cadastre et de l'IGN avant tout processus de fusion et de déformation élastique.

Le traitement se fait en plusieurs étapes : une étape de déformation de chaque section cadastrale, puis des étapes de raccordement des limites de sections, puis de communes, etc. La figure 17 illustre les décalages entre les données du cadastre et les données de l'IGN (BD Ortho® et BD Topo®). Pour y remédier, il faut tout d'abord définir des vecteurs de déplacement. Une fois cette étape réalisée, une grille de déformation élastique « affine » est appliquée. De cette manière un vecteur de déplacement basé sur le calcul des vecteurs de déplacement définis peut être obtenu pour tout point de l'espace. A la fin de cette étape la section cadastrale est repositionnée dans le référentiel IGN.

Au niveau des jonctions intra communales, l'utilisateur doit définir des points raccord pour relier les sections entre elles. Ensuite le logiciel calcule les limites grâce à un algorithme qui tient compte de l'échelle et de la qualité de chaque planche cadastrale. Le processus d'assemblage proprement dit est basé sur une déformation des sections avec les nouvelles limites. L'assemblage final entre communes est géré de la même manière que les raccords intra communales.

I.2.4 Intégration de polygones

Cette partie n'est que brièvement détaillée car elle ne comporte pas d'informations indispensables aux outils de fusion de MNT développés dans cette thèse. Néanmoins, elle a été abordée pour permettre un aperçu exhaustif des outils existants pour la comparaison, l'appariement et la fusion d'éléments 2D.

I.2.4.1 Comparaison de surfaces

Distance entre surfaces

Pour évaluer la distance des positions entre surfaces [Vauglin, 1997] propose la **distance surfacique** qui mesure le rapport des surfaces communes par rapport à l'union de ces surfaces. Il existe également la probabilité d'association ([Phalakarn, 1991], [Lemarié, 1996], [Bel Hadj Ali, 2001]) ainsi que la fonction de ressemblance [Bel Hadj Ali, 1997].

Comparaison de ressemblance de forme entre surfaces

Pour les objets surfaciques, les critères de formes sont présentés essentiellement par [Kidner, 96]. Il définit entre autres comme mesures pertinentes de similitude :

- Le rapport entre les aires
- Le rapport entre les périmètres
- Le rapport entre les compacités (périmètre/aire)
- Le rapport entre les élongations
- Le rapport entre les excentricités
- Le rapport entre les allongements (longueur/largeur)
- Le rapport entre les rectangularités
- Le rapport entre les distances maxima au centroïde
- Le rapport entre les distances radiales au centroïde
- La comparaison entre les transformations de Fourier
- La comparaison entre les dimensions fractales.

Le test de ces mesures sur des surfaces simples représentant le même phénomène du monde réel mais à des échelles différentes, montre que l'aire, la rectangularité et l'excentricité sont des mesures fiables tandis le périmètre donne de mauvais résultats. [Bel Hadj Ali, 2001] évoque également la distance entre fonctions angulaires, mais celle-ci présente le désavantage d'être inadaptée à des objets de complexité élevée comme les objets géographiques. De nombreuses autres distances sont également présentées dans cette thèse, mais elles sont plus liées aux propriétés mathématiques intrinsèques aux surfaces comme la distance associée à l'espace des moments. Cette dernière permet, via le calcul de moments mathématiques propres aux surfaces, de détecter les différences de formes entre les deux entités mesurées.

Pour les surfaces complexes, comportant des trous par exemple, il est possible d'utiliser en complément le nombre de trous, le nombre de composants connexes, l'enveloppe convexe et le nombre d'Euler.

I.2.4.2 Appariement de surfaces

L'appariement de surfaces peut se faire soit à l'aide des valeurs fournies par les outils de distance soit celles obtenues avec les outils de ressemblance de forme [Bel Hadj Ali, 2001]. De même que pour les points et les lignes, une seule valeur ne se suffit pas. Tout processus d'appariement utilise donc un éventail de mesures de comparaison, précédées d'une sélection de candidats potentiels et complétées par une information de localisation. La combinaison de plusieurs mesures détecte les différentes configurations des écarts en tenant compte des écarts de forme et de position [Bel Hadj Ali, 2001]. Le positionnement géographique quant à lui évite que deux objets de même forme mais très éloignés géographiquement soient détectés comme homologues.

I.2.4.3 Fusion de surfaces

Les quelques outils développés pour la fusion de surfaces à proprement parler se basent sur les frontières des objets, comme par exemple le projet de conversion du cadastre au format EDIGéo. Aucun algorithme ne travaille avec les surfaces directement.

1.3 Intégration de Données 2.5D (MNT)

Le propos de cette thèse est l'intégration de données contenant l'information de relief (altitude), aussi appelées Modèle Numérique de Terrain (MNT). Cette partie introduit en premier lieu les caractéristiques des MNT qui les démarquent des données 2D. Celle-ci détaille tout d'abord les différents modes d'acquisition et de production existants pour ce type de données (terrestre et maritimes), bien plus diversifié que pour les données 2D, ce qui implique de nombreuses difficultés pour les processus d'intégration ultérieurs. Ensuite les différents formats de MNT sont énumérés de façon à disposer d'un aperçu du mode de stockage de l'information. Les avantages et inconvénients de chaque format pour la formalisation du relief sont ensuite énumérés. Enfin le dernier paragraphe est consacré à la définition et l'évaluation de la qualité d'un MNT.

En second lieu sont décrits les précédents travaux effectués en matière d'intégration de MNT. L'approche présentée dans cette thèse en section II.3 peut être ainsi positionnée vis-à-vis des méthodes existantes.

I.3.1 Caractéristiques des MNT

Il existe différentes dénominations pour le terme générique de MNT suivant la technique de production utilisée pour les obtenir. En domaine terrestre, certains ne modélisent que le sol à proprement parler (modèle numérique de terrain), tandis que d'autres comprennent également le sursol, comme la végétation ou le bâti (modèle numérique de surface, MNS, ou modèle numérique d'élévation, MNE). Ces derniers s'obtiennent généralement par des jeux de données brutes acquises grâce à des techniques telles que le Lidar, la photogrammétrie, ou la radargrammétrie. Les MNT constituent le plus souvent un dérivé de ces modèles, « filtré » de tous les éléments constitutifs de la couverture topographique pour ne garder que le sol nu. Le terme « couverture topographique » désigne tous les éléments situés au dessus de la surface terrestre continentale : bâtis et végétation haute comme les forêts.

En domaine maritime, quel que soit le mode d'obtention des modèles (par altimétrie Lidar, échosondeurs ou autres) tous sont nommés modèles bathymétriques. Par la suite, aucune distinction ne sera faite et seul le terme de MNT est utilisé indistinctement pour tous les modèles. Par ailleurs, indépendamment de ce qu'il modélise (surface du sol avec sa « couverture » ou simplement le terrain nu) chaque modèle se distingue par son mode d'acquisition, son format de stockage ou sa qualité. Les sections suivantes présentent ces différentes caractéristiques.

I.3.1.1 Acquisition et génération d'un MNT

L'acquisition de MNT peut se faire par des capteurs de mesure de signal embarqués sur plusieurs types de vecteurs (avions, bateaux), ou avec des levés de terrain. Ils sont également générés à partir de couples d'images (satellites ou aériennes), ou encore des digitalisations de cartes. Ces différents modes d'acquisition entraînent par la suite un des traitements qui leur sont spécifiques. Ces traitements peuvent comprendre : le calcul du relief à l'aide d'un couple d'images stéréoscopiques (dans le cas de photos aériennes), une étape de traitement du signal

(pour des acquisitions laser par exemple) ou une phase d'interpolation (lors de levés de terrain). Cette dernière permet d'obtenir une représentation continue à l'aide d'un nombre fini de points.

La figure 18 présente la chaîne de traitement simplifiée nécessaire à la production d'un MNT. L'acquisition des données, comme par exemple une mesure de signal renvoyé par le sol ou des images de télédétection, sert en premier lieu de base à la génération du modèle. Ensuite vient une étape de traitement de ces données qui fournissent rarement une information d'altitude directement et leur stockage dans un format propre. Le modèle peut rester sous ce format brut ou être à nouveau transformé en un modèle continu grâce à des algorithmes d'interpolation.

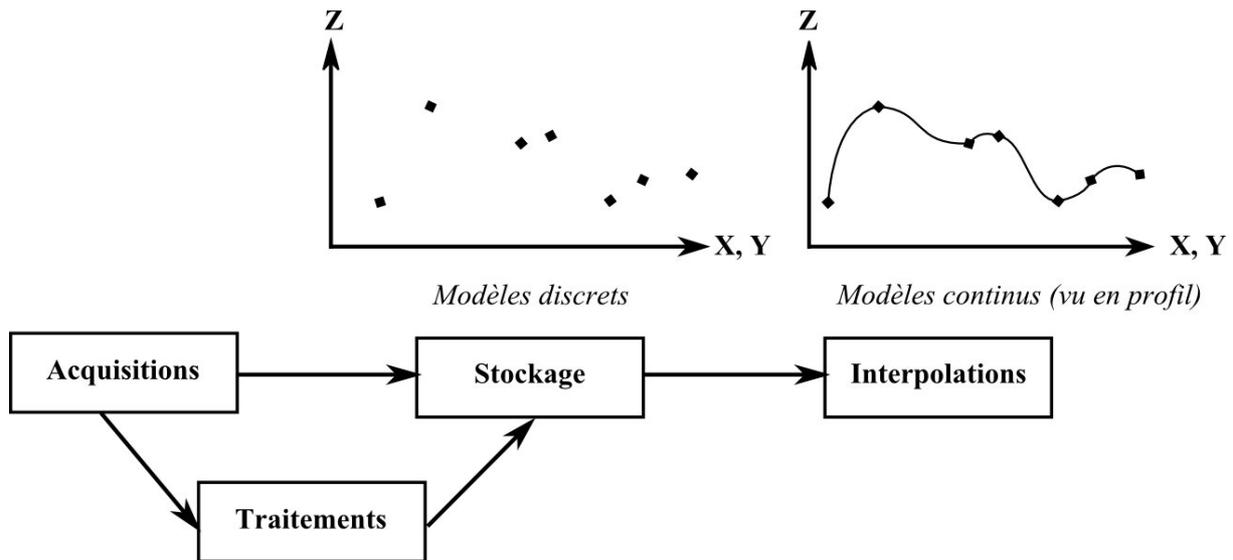


Figure 18 : Etapes de production d'un MNT.

Il est à noter qu'à chacune de ces étapes des erreurs sont ajoutées : lors de l'acquisition les erreurs physiques dues aux capteurs et aux méthodes d'acquisition sont insérées. Par ailleurs des erreurs sont introduites lors des traitements par les méthodes appliquées sur les données (approximations). De plus, le stockage dans un modèle de même que leur interpolation entraîne des erreurs de conversion d'un format à l'autre [Li, 1990].

MNT terrestres : acquisitions et traitements

En ce qui concerne les MNT terrestres, plusieurs vecteurs sont à l'origine des données qui vont permettre leur génération. Les satellites avec leurs capteurs radar ou optiques vont tout d'abord fournir l'imagerie grâce à laquelle on obtient l'information d'altitude via des traitements appropriés (présentés ci-après). Ensuite il existe également les systèmes embarqués sur des avions (laser, radar, ou photographie aérienne). Ceux-ci produisent soit une information d'altitude via des processus automatiques (laser) soit à nouveau des images qui sont retraitées par la suite. Enfin il y a également les levés de terrain acquis sous forme de points régulièrement répartis sur la surface à modéliser.

Les satellites d'observation (type SPOT, ERS, ENVISAT) mesurent l'information de surface selon différents principes : les capteurs optiques mesurent le rayonnement solaire rétrodiffusé par la surface terrestre (même principe qu'une photographie) tandis que les capteurs radar émettent eux-mêmes un signal (indépendant des conditions extérieures de luminosité et atmosphériques) et enregistrent l'onde réfléchi par le sol (information de rugosité).

Les capteurs aéroportés (avions) utilisent le même principe que les satellites mais le degré de détail des images produites est supérieur en raison d'une altitude d'acquisition moindre.

Le Lidar (LIght Detection And Ranging) mesure la distance sol/capteur via une onde électromagnétique cohérente (laser) émise dans une longueur d'onde donnée (le plus souvent l'infrarouge) et enregistrée à l'aide d'un capteur embarqué dans un vecteur aéroporté (avion, hélicoptère). Il se présente comme un système actif émetteur d'ondes laser réfléchies par les points d'impact. L'utilisation de miroirs réflecteurs permet d'orienter et de balayer le signal au sol avec une fréquence modulable (principe du laser scanning). L'altimètre laser déduit du temps de parcours du signal rétrodiffusé la distance entre la plate-forme et le point d'impact du laser. Ces mesures recueillies ne sont exploitables que si elles sont correctement géoréférencées dans un référentiel géodésique (le WGS84 le plus souvent car il a l'avantage d'être global). Pour cela, il faut combiner un système GPS et un système de navigation inertielle, afin d'en déduire les coordonnées tridimensionnelles du capteur puis de la cible. La densité d'échantillonnage est de l'ordre du point par m² en moyenne et peut augmenter jusqu'à six points par m² en zone de recouvrement [Oliveira et al., 2006] tandis que la précision altimétrique obtenue est inférieure au décimètre. L'enregistrement se compose de multiples échos (first pulse, last pulse) pour chaque impulsion émise en fonction des objets du sursol.

Les levés topographiques sont effectués à l'aide d'un tachéomètre le plus souvent et ont pour objectif la création d'un plan ou d'une carte à partir d'informations obtenues sur le terrain. Une fois le tachéomètre¹⁰ positionné de manière à ce que l'axe vertical de l'appareil soit perpendiculaire au plan horizontal de la station, les points caractéristiques du terrain sont relevés. A partir de ces points de référence, si le terrain est très étendu, la série de mesures est réalisée par rayonnement via un cheminement précis au cours duquel l'appareil de mesure est déplacé pour couvrir toute la surface à carter.

En dehors de la technique Lidar qui permet de fournir directement des altitudes par la mesure de la distance entre la cible et le capteur via un faisceau laser, l'information altimétrique n'est pas mesurée directement par les autres méthodes d'acquisition satellitaires (images optique ou radar) ou aériennes (photographie). Il est alors nécessaire d'appliquer des traitements spécifiques selon le type d'image afin d'obtenir l'information d'altitude : photogrammétrie, radargrammétrie, interférométrie ou clinométrie.

La photogrammétrie ainsi que **la radargrammétrie** exploitent la vision stéréoscopique d'un couple d'images optiques (numériques ou non) ou radar pour reconstituer le relief. L'information topographique (relief du terrain) est déterminée à partir du calcul de la parallaxe d'un couple de points homologues. La parallaxe représente la variation apparente de la position d'un objet résultant d'un changement du point de vue de l'observation. L'appariement des points est réalisé avec des stéréo-restituteurs analytiques ou numériques ou par corrélation automatique ([Charleux-demargne, 2001], [Dupont et al., 1998]).

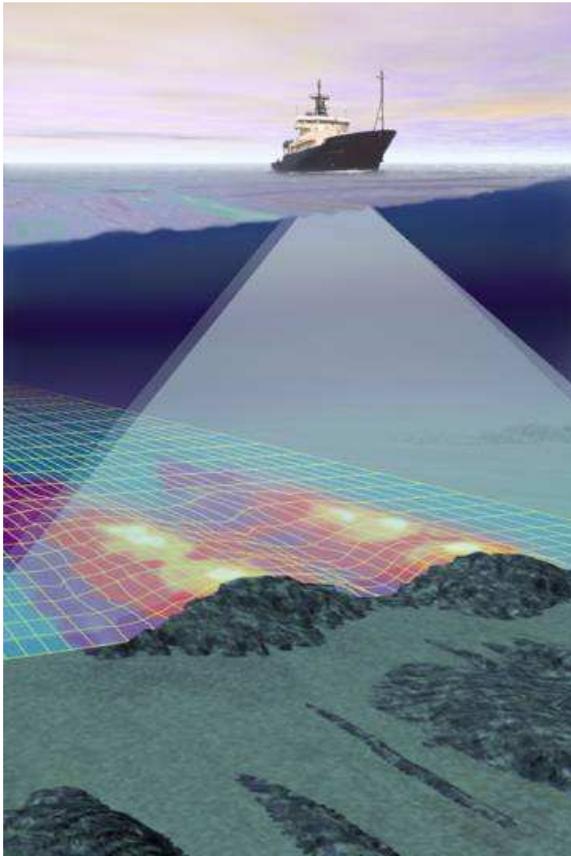
L'interférométrie exploite la différence de phases entre deux images radar préalablement recalées et prises dans les mêmes conditions géométriques. Celle-ci informe seulement sur la distance entre le capteur et la cible aux deux instants de prise de vue. C'est pourquoi la construction du MNT requiert de connaître précisément le positionnement du capteur lors des prises de vues ([Adragna, 1997]).

La clinométrie ou **radar-clinométrie** est une technique qui reconstruit le relief à partir de la radiométrie d'une image, en se basant sur une reconstruction par ombrage ([Tannous et al., 1997], [Nocéra, 1993], [Paquerault & Maître, 1997]).

Il est à noter que des modèles d'altitude terrestres et maritimes peuvent également être obtenus par l'intermédiaire de processus de **numérisation** de lignes de niveau et de points cotés à partir de cartes topographiques ou marines.

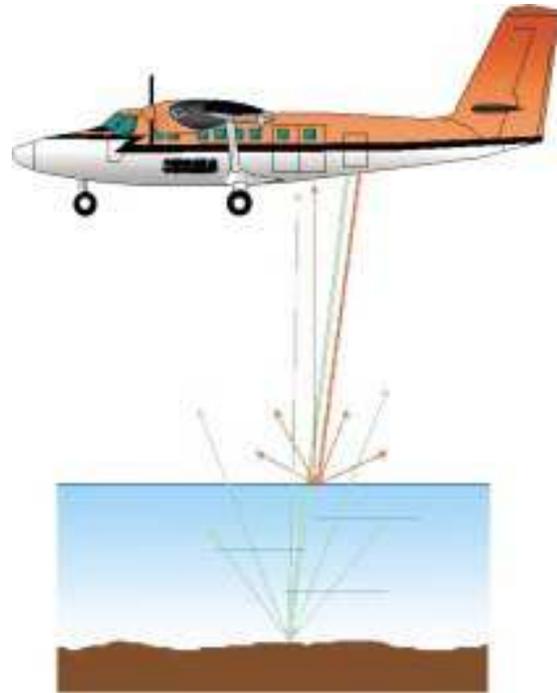
¹⁰ Un tachéomètre est un appareil utilisé en topographie pour des relevés à grande et à moyenne échelle pour mesurer des distances et des dénivellations.

Modes d'acquisition de MNT maritimes



(a) (Source :

<http://www.lr.tudelft.nl/live/pagina.jsp?id=643ece3f-6e04-4fc9-8903-4bd3a3e54297&lang=en>)



(b) Source : Lettre du SHOM N°21.

Figure 19 : Principe d'acquisition d'un échosondeur monofaisceau et du Lidar bathymétrique.

Les levés bathymétriques sont obtenus principalement à l'aide de trois méthodes d'acquisition différentes : le sondeur acoustique monofaisceau (figure 19(a)), le sondeur acoustique multifaisceaux et le lidar topo bathymétrique (figure 19(b)). Notons qu'il a également existé le plomb de sonde, mais cette technique n'est plus usitée. Les trois modes d'acquisition principaux sont donc présentés ci-après.

Le principe de fonctionnement d'un **échosondeur** (mono ou multi faisceaux) est basé sur l'émission et la réception d'un signal acoustique émis par une antenne fixée sur la coque d'un navire. Le capteur mesure le temps aller/retour de l'onde et en déduit la profondeur p avec la formule :

$$p = c \cdot dt / 2$$

où « c » représente la célérité du son dans l'eau, et « dt » la durée aller/retour de l'onde du fond marin au navire [SHOM, 2009].

Par rapport au sondeur monofaisceau, le sondeur multifaisceaux dispose en plus d'un grand nombre de faisceaux (figure 20) qui lui permettent d'explorer une large bande. La précision de mesure est d'autant plus grande que les faisceaux sont étroits.

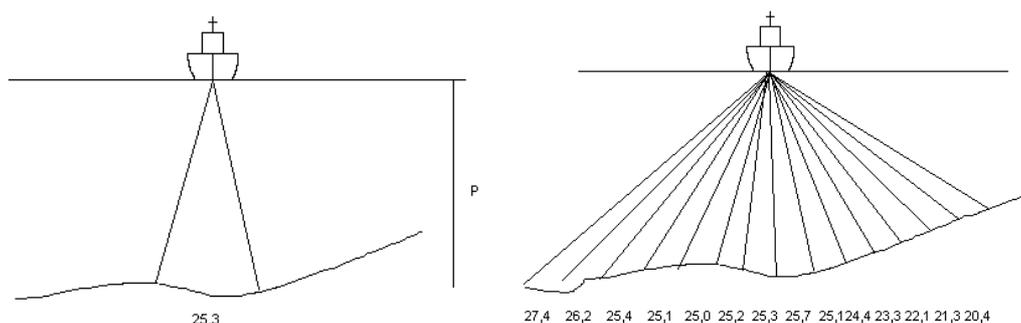


Figure 20 : Illustration des techniques d'acquisition d'échosondeur. (Source : http://www.shom.fr/fr_page/fr_act_acquisition/smf/principe.htm)

La profondeur calculée tient compte si possible des variations de la célérité du son dans la colonne d'eau traversée par le faisceau, de l'inclinaison du faisceau et de l'attitude du bateau (roulis, tangage). Une étape de traitement du signal permet d'optimiser l'estimation de ces paramètres. La précision des mesures est ainsi augmentée tandis que le rapport signal/bruit est diminué. Cette diminution améliore le pouvoir séparateur entre deux objets (résolution spatiale et temporelle) [Lurton, 2002].

La profondeur des sondes dans un système de référence (et non la hauteur du bateau vis à vis du fond marin) s'obtient avec un modèle de marée suffisamment précis pour repositionner la valeur des sondes à une mesure absolue. Cette valeur de référence est le zéro hydrographique qui représente l'origine du système altimétrique marin à partir duquel sont comptées les hauteurs de marées. C'est le niveau théorique estimé des plus basses mers astronomiques (coefficient 120), parfois différent du niveau réellement observé. En effet il arrive très exceptionnellement que sous l'influence de conditions météorologiques extrêmes (dépressions, etc.) le niveau de la mer descende en deçà de ce coefficient ([Wöppelmann et al., 1999], [Marguerit, 2004]).

En réalité les notions de « zéro hydrographique » et de « plus basses mers astronomiques » ne sont pas tout à fait identiques. Pour qu'elles le soient, il faudrait que les positions des astres lors de fortes marées soient les mêmes en tous points, de façon à ce que l'écart entre le zéro hydrographique choisi et le niveau des plus basses mers astronomiques soit nul, ce qui n'est jamais vérifié. Il existe donc entre les deux un écart difficilement évaluable [Simon, 2001].

Le calcul du zéro de référence des cartes marines s'est effectué à une période donnée dans chaque port de référence grâce aux mesures marégraphiques établies sur une durée d'observation d'au moins un an. Seulement une fois fixé, il est en général conservé même si des calculs ultérieurs montrent qu'il n'était pas le niveau des plus basses mers. Il arrive, si un écart est trop important, qu'il y ait un réajustement, comme à Brest en 1996 où il a été rehaussé de 0,5 mètre [SHOM, 1998].

A ceci s'ajoutent les problèmes de la concordance des zéros définis à des dates différentes au sein des ports ainsi que son évaluation entre deux ports de référence. Elle s'effectue le plus souvent par une interpolation affine entre les deux valeurs en supposant linéaire la géométrie moyenne du plan d'eau. Ce qui peut s'avérer inexact pour des mesures réalisées en domaine côtier proche comme des zones estuariennes (effets de pente hydraulique), en présence de courant ou de rupture topographique sous-marine importante (jusqu'à 40 cm de différence en baie du Mont Saint Michel par exemple pour des données de précision centimétrique [Kaveh & Deffontaines, 2005]).

De même, l'utilisation de bateaux pour réaliser les levés hydrographiques limite la possibilité de recueil des données à des profondeurs inférieures ou égales au mètre. Afin de pallier cette

lacune, il est possible d'utiliser par petits fonds non ou peu turbides d'autres techniques comme le Lidar bathymétrique [SHOM & Allain, 2004].

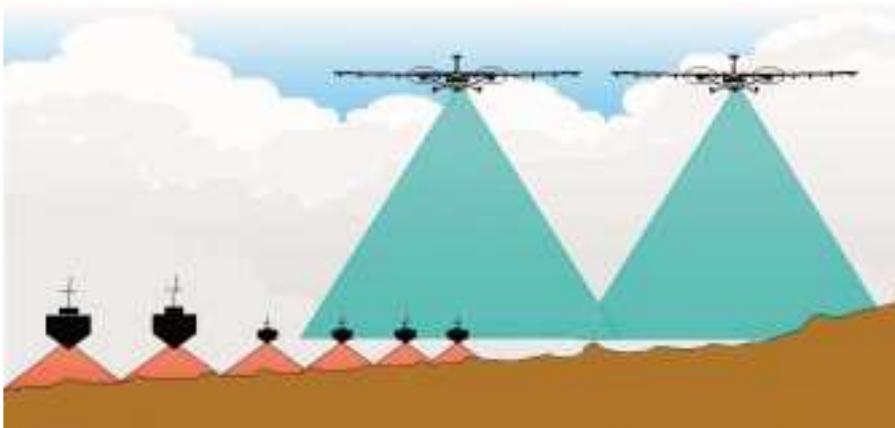


Figure 21 : Systèmes de mesure du fond marin selon la profondeur. (Source : lettre du SHOM N°21)

Le principe du **Lidar topo-bathymétrique** est le même que celui du Lidar topographique. Un rayon bleu vert est utilisé en complément du signal infrarouge car ce dernier se trouve réverbéré par la surface de l'eau (figure 19). La hauteur d'eau est déduite de la différence de temps entre les deux signaux [SHOM & Debes, 2006]. Cette technique n'est utilisable que si la turbidité de l'eau est faible (densité de particules en suspension) car celle-ci restreint l'intensité du signal laser propagé dans l'eau [SHOM & Allain, 2004].

La figure 21 montre le recouvrement qu'il est possible d'obtenir avec l'acquisition conjointe réalisée par toutes ces techniques. Le recouvrement entre sonar petit fond et Lidar bathymétrique est suffisante pour assurer la génération d'un modèle continu.

Synthèse sur les modes de production des MNT

Le suivant récapitule les différents modes présentés pour l'acquisition et la production de MNT, qu'ils soient terrestres ou maritimes. Chacune de ces techniques présente des avantages mais également des limites. Les mesures de terrain sont très précises mais nécessitent un algorithme d'interpolation afin de générer un modèle continu, ce qui ajoute des approximations et des artéfacts d'interpolation.

Tableau 2 : Sources et techniques de calcul du positionnement d'un jeu de points de référence.

Type de MNT	Sources	Techniques
Terrestre	Points côtés	Mesures topométriques
	Images radar : capteurs actifs aéroportés ou satellitaires	radargrammétrie, radar-clinométrie, interférométrie
	Images optiques : capteurs passifs aéroportés ou satellitaires	photogrammétrie, clinométrie
	Mesures laser topographiques	Distance optique
Bathymétrique	Points de sondes : mono ou multi faisceaux	Echosondage acoustique
	Mesures laser bathymétriques	Distance optique

La photogrammétrie ainsi que la radargrammétrie sont pour leur part très dépendantes des images qui ont servi à produire le MNT : celles-ci doivent être assez différentes pour avoir un effet stéréoscopique important et donc une meilleure estimation de l'altitude. Cependant, si les changements entre les deux prises de vue sont trop grands, le bruit sur la parallaxe est supérieur et la hauteur estimée est mauvaise ([Leberl, 1990], [Tannous et al., 1997]).

Pour l'interférométrie, si la pente est trop forte, il y a une perte de cohérence entre les deux images, il est donc plus difficile de reconstituer le relief. De même si l'incidence locale des images est trop importante ([Massonet & Rabaute, 1993], [Adragna, 1997], [Dupont, 1997], [Pichon et al., 1997]). La clinométrie est plus adaptée pour une restitution du relief local car elle nécessite l'homogénéité de la réflectance sur l'ensemble de la scène (de sorte que les variations radiométriques soient uniquement dues à la pente).

I.3.1.2 Formats des MNT

Une fois l'acquisition effectuée et les traitements appliqués afin d'extraire l'information de relief, les modèles numériques doivent être stockés dans un format facilement réutilisable par la suite au sein d'un SIG. Les formats de MNT les plus couramment employés sont : nuage ou grille de points, courbes de niveau, profils, TIN (Triangular Irregular Network), et grille régulière ou non (figure 22).

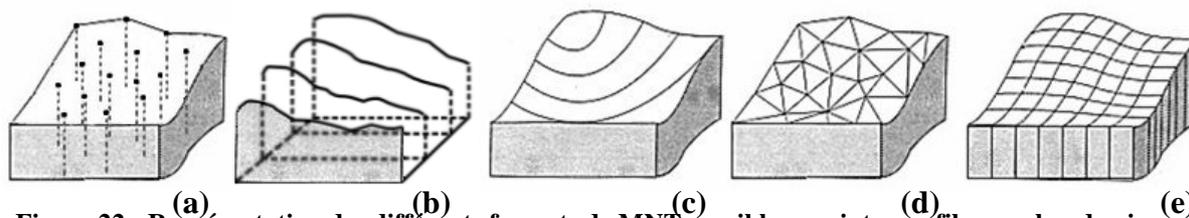


Figure 22 : Représentation des différents formats de MNT possibles : points, profils, courbes de niveau, TIN (Triangular Irregular Network), Grille.

Les nuages et grilles de points

Les nuages de points (figure 22(a)) sont la forme la plus brute d'un MNT : chaque point, disposé de façon non régulière, représente une position géographique donnée dotée d'une valeur d'altitude. Ils sont issus en général de levés topographiques ou de sondes bathymétriques.

La grille de points est obtenue grâce à un ré-échantillonnage à un pas régulier d'un autre modèle (nuage de points par exemple). Il s'agit d'appliquer une redistribution des valeurs régulièrement au sein du modèle en interpolant celles qui n'existent pas à partir des données d'origine. Au final, au lieu d'avoir des points avec une disposition aléatoire, chaque point sera disposé au centre d'une cellule à un intervalle constant sur tout le modèle. Ces deux types de formats discrets souffrent de la même difficulté à appréhender le paysage représenté.

Les profils

Les modèles de profils (figure 22(b)) représentent l'altitude à l'aide de points disposés le long de tracés horizontaux. L'intervalle entre deux points côtés d'un même profil est variable et peut s'adapter à la morphologie locale. La précision intrinsèque de ces modèles peut être très bonne, néanmoins dans le cas de profils successifs très distants entre eux, le relief peut être décrit de façon non homogène dans toutes ses directions. C'est le cas par exemple de certaines acquisitions avec un sondeur mono faisceau, qui donne un espacement entre profils d'environ 200m [Oliveira et al., 2006].

Les courbes de niveau

Les courbes de niveaux (figure 22(c)) sont des isolignes avec une altitude constante, chacune représentée par un ensemble de points qui constitue une polyligne fermée ou ouverte. On parle d'isohypse pour des courbes topographiques et d'isobathes pour des cartes marines. Ces courbes sont généralement tracées de façon à avoir le même pas d'altitude entre deux courbes successives. L'information d'altitude est présente uniquement au niveau des polylignes. En conséquence cette représentation ne révèle que partiellement la morphologie du terrain. Les points caractéristiques sont rarement renseignés de même que les lignes de rupture de pente. Par ailleurs, les zones de faible pente sont très peu décrites.

Triangular Irregular Network (TIN)

Dans ce type de modèle (figure 22(d)), la surface topographique est constituée d'une multitude de facettes triangulaires adjacentes : les facettes relient les points de référence. Elles sont de formes et de tailles variables. Cette représentation retransmet globalement une bonne idée du relief car le TIN peut être construit à partir des points caractéristiques du terrain représenté (sommets, cols, cuvettes). [Dakowicz & Gold, 2003] proposent une amélioration conséquente de ce format en réutilisant la méthode présentée par [Amenta et al., 1998] sur la squelettisation des courbes de niveau. Concernant les TIN, il s'agit de générer la transformation de l'axe médian (ou squelette) et de se baser dessus pour générer le TIN et ainsi éviter les problèmes de triangles plats (fréquent dans les modèles obtenus à partir d'isolignes). Les modèles se retrouvent ainsi enrichis de lignes de force comme les lignes de crête, et les thalwegs. En complément, il est possible de les compléter par d'autres lignes caractéristiques comme les ruptures de pente, contours ou autres limites (trait de côte, frontières, réseau hydrographique ou routier par exemple). De part leur nature, ces modèles possèdent un rapide temps de calcul ainsi qu'une incorporation simple des discontinuités et des caractéristiques structurales du relief.

Grille raster

Le format raster (figure 22 (e)) est un maillage (ou une grille) qui représente l'espace continu de manière discrète à l'aide d'unités de base finies, régulières, dont la longueur est souvent égale à la largeur : les cellules. Pour les grilles régulières (coordonnées X et Y à un delta constant entre X/X+1 et Y/Y+1), l'altitude est donnée par la valeur de chaque pixel considéré. Seules les coordonnées X et Y de la première cellule sont nécessaires. Les autres cellules peuvent être calculées à l'aide des delta X et delta Y. Le croisement ou l'extraction d'informations sont très simples avec ce type de format car les méthodes employées (additions, soustractions, intersections, etc.) se rapprochent beaucoup de celles du traitement d'image. En contrepartie, aucune information géomorphologique pertinente (points/lignes caractéristiques) n'est stockée. De plus l'ajout de ces informations est difficile, contrairement aux TIN par exemple. En effet, la régularité de la grille implique une description constante du relief quel que soit le type de paysage figuré en zone plate ou en montagne. De même, l'extraction de toute information (réseau hydrographique, lignes de rupture de pente, de crêtes et de thalweg, chemin d'écoulement) va être très dépendante du pas de la maille ainsi que des huit directions préférentielles d'une grille carrée. Des études comme celles poursuivies par [Kienzle, 1996], [Lagacherie et al., 1996] ou [Zhang & Montgomery, 1994] décrivent l'influence de la taille de la grille sur les résultats obtenus.

Une réponse à certains de ces problèmes peut être l'utilisation de grille irrégulière en arbre quaternaire (Quad Tree en anglais). Le principe de ces grilles est de segmenter chaque zone qui nécessite une description plus fine, sachant qu'une maille est subdivisée à nouveau en quatre nouvelles mailles carrées. Par exemple une région de pente constante sera représentée par une même zone de l'image peu détaillée tandis qu'un brusque accident de la topographie

nécessitera une segmentation en blocs de plus en plus fins selon le degré de précision requis. Cette irrégularité du pas de la grille permet ainsi de s'adapter au mieux à des MNT avec des zones hétérogènes.

Méthodes d'interpolation de l'altitude

Dans le cas de MNT avec des hauteurs portées par des géométries (points, isolignes, profils), une interpolation permet de connaître l'altitude de façon continue dans des zones où aucune information n'est présente. Ces fonctions peuvent également servir à convertir un modèle d'un format à l'autre ou effectuer un ré-échantillonnage des valeurs.

Ces techniques d'interpolation ont amplement été décrites dans la littérature ([Weibel & Heller, 1991], [Longley et al., 1999], [Arnaud & Emery, 2000]). Notre but n'est pas d'en dresser un inventaire exhaustif mais de rappeler celles les plus couramment utilisées dans les SIG.

Ainsi, l'interpolation des altitudes à partir d'un nuage de points doit respecter des contraintes précises de manière à ce que le relief modélisé représente le plus fidèlement possible le terrain réel.

Deux catégories de méthodes se distinguent :

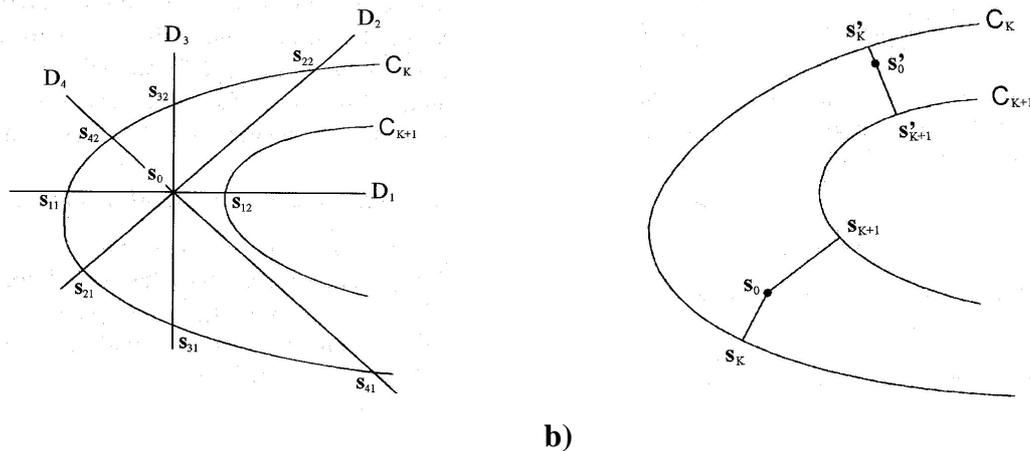
- la première considère uniquement les points disposés de manière irrégulière dans l'espace. Ces méthodes sont appelées **méthodes d'interpolation spatiale** ;
- la seconde considère également **l'information morphologique** des lignes caractéristiques du relief (crêtes, thalwegs, réseaux hydrographiques, courbes de niveau, etc., [Arnaud & Emery, 2000]).

Parmi la première catégorie de méthodes d'interpolation, il existe plusieurs sortes d'interpolations : celles basées sur le voisinage local d'un point ou celles basées sur des statistiques (Krigage essentiellement). Parmi les méthodes de voisinage local, les fonctions de l'inverse de la distance, bilinéaires, voisins naturels ou des fonctions polynomiales diverses sont les plus implémentées. Elles permettent le calcul de pente, le ré-échantillonnage ou de réseaux hydrographiques, de façon simple et plus rapide. Et même si elles sont à l'origine de certains artéfact d'interpolation, elles suffisent en général pour la plupart des applications.

Ces méthodes font assez peu cas de la nature des données traitées. Pourtant, pour le cas spécifique du MNT, des algorithmes uniquement mathématiques, qui ne tiennent pas compte du relief, ne permettront pas d'obtenir des estimations suffisamment réalistes des altitudes interpolées [Arnaud & Emery, 2000]. En parallèle à ces méthodes, il existe aussi la triangulation calculée à partir d'interpolation. Celle de Delaunay est la plus usuelle mais il en existe d'autres. Cette triangulation est obtenue de manière à ce qu'un cercle passant par les trois points d'un triangle ne contienne aucun autre point. Les altitudes à l'intérieur du triangle sont calculées le plus souvent à l'aide d'une interpolation barycentrique de l'inverse de la distance [Arnaud & Emery, 2000]. La triangulation présente l'avantage de se baser sur les points d'origine : si ceux-ci sont des marqueurs précis, le modèle construit respecte la « logique » du relief représenté.

Pour pallier les visions trop « mathématiques » des méthodes d'interpolation, une seconde catégorie de fonctions s'appuie plus particulièrement sur les lignes structurantes du relief : crêtes, thalwegs ou lignes de plus grande pente. A la différence des méthodes d'interpolation spatiales, la position des points dans l'espace prime sur leur nombre. Par exemple, le calcul des points hauts ne nécessite pas les altitudes des bas, qui ne seront donc pas utilisées. Il existe dans cette catégorie deux fonctions présentées en figure 23 :

- celle employant des axes (voir figure 23(a)),
- celle utilisant la ligne de plus grande pente [Arnaud & Emery, 2000] (voir figure 23(b)).



a) b)
Figure 23 : Illustration des deux méthodes d'interpolation utilisant les lignes structurantes. En a) interpolation à partir des lignes de niveau C_k et C_{k+1} , et b) interpolation basée sur la ligne de plus grande pente. Source [Arnaud & Emery, 2000].

La première méthode consiste à déplacer un système d'axes (entre un et quatre, notés D_1 à 4 sur la figure 23(a)) centré sur chaque nœud de la grille. Chaque point à l'intersection de ces axes avec les courbes de niveau (points $S_{22}, S_{32}, S_{42}, S_{11}, S_{21}, S_{31}, S_{41}$ sur la figure 23(a)) est utilisé pour calculer le point résultant (S_0).

La méthode utilisant la ligne de plus grande pente consiste à rechercher pour chaque point à interpolier (S_0 sur la figure 23(b)) entre deux points S_K et S_{K+1} , les segments de droite allant de S_0 à S_K et de S_0 à S_{K+1} qui s'approchent au mieux de la ligne de plus grande pente.

Une interpolation intermédiaire entre ces deux catégories est représentée par la triangulation : elle combine une interpolation linéaire entre deux points et un enrichissement morphologique en se basant sur les lignes de force du relief [Dakowicz & Gold, 2003].

L'estimation de la qualité de ces méthodes d'interpolation n'est pas aisée. Elles sont pour la plupart des interpolations « exactes ». Ce terme signifie que la surface interpolée passe par tous les points de l'échantillon de départ ou au moins au plus proche comme dans le cas de splines. Il n'est donc pas possible d'évaluer l'écart entre les points d'origine et ceux du résultat final : il n'y en a pas.

Une réponse à ce problème est d'utiliser la technique de « l'échantillon de base et de l'échantillon test » [Arnaud & Emery, 2000]. Les points sont partagés en deux sous-ensembles : échantillon de base et échantillon test. Une surface est interpolée à partir de l'échantillon de base. L'échantillon test est utilisé pour mesurer la qualité en confrontant le « z » des points de test et le « z » de la surface interpolée [Arnaud & Emery, 2000].

I.3.1.3 Qualité des MNT

Le terme de « qualité » de MNT la notion d'exhaustivité et de cohérence des données [Longley et al., 1999]. Toute qualité intrinsèque de MNT peut être évaluée par différents éléments qui lui sont propres : en premier lieu ses valeurs altimétriques et en second par tout autre paramètre calculé à partir de ces altitudes. Ces paramètres sont des dérivées premières ou secondes de la surface topographique, et sont nommés *paramètres géomorphologiques*. Ils comprennent par exemple la pente, l'azimut, les lignes caractéristiques. **La pente** est l'inclinaison de la surface par rapport à l'horizontale, **l'azimut** est l'angle horizontal entre la direction d'un objet et une direction de référence. **Les lignes caractéristiques** représentent toutes lignes informant sur la structure du relief (lignes de points hauts, de points bas, lignes de changement brutal de pente, voir chapitre III).

Il est également courant d'estimer la qualité d'un modèle soit par comparaison de celui-ci avec des données de référence (**validation externe**) soit sans données mais en utilisant un savoir à priori sur les propriétés du relief [Polidori, 1995] (**validation interne**).

L'évaluation de la qualité des valeurs altimétriques passe tout d'abord par l'identification des différents types d'erreurs présentes dans les données (liées aux capteurs ou aux données elles-mêmes). Celle-ci est systématiquement complétée par une comparaison du modèle avec les données de référence disponibles [Charleux-demargne, 2001].

Types d'erreurs présentes dans les données

Les données sont biaisées par deux sources principales : les erreurs commises lors de l'acquisition et celles dues aux traitements ultérieurs des données (insertion d'artéfacts d'interpolation, etc.).

Parmi les erreurs affectant le capteur, il existe ([Zhilin et al., 2005]) :

- **Les erreurs ponctuelles** (« gross errors » ou « mistakes » en anglais), dues à des dysfonctionnements accidentels du capteur, ou par exemple à la présence d'obstacles dans la couche d'eau traversée [SHOM & Debese, 2006]. Certaines peuvent être détectées et corrigées lors de la phase de validation des données. Pour les sondeurs multifaisceaux, cette étape est fastidieuse en raison d'un volume important de données ainsi que des contraintes de sûreté de navigation imposées par la navigation petit fond [SHOM & Debese, 2006]. Pour le Lidar bathymétrique, la détection de ces erreurs s'appuie sur l'exploitation de photographies numériques acquises en même temps que les levés laser [SHOM & Debese, 2006]. La probabilité de ces erreurs est très faible, mais se retrouve souvent lors des étapes de corrélation automatique d'images, lors de la production de MNT, en raison de discordances entre points de recalage.
- **Les erreurs systématiques** (« systematic errors ») liées au fait que bon nombre de capteurs sont des systèmes multi composants : association par exemple d'un capteur optique, radar ou acoustique avec une centrale GPS et parfois une centrale d'attitude [SHOM, 06] pour les mesures bathymétriques. Ces erreurs se caractérisent par leur présence sur l'ensemble de la fauchée et leur détection passe par l'étude des discontinuités sur les zones de recouvrement entre fauchées [SHOM & Debese, 2006].
- **Les bruits du capteur** (« random errors ») directement dépendants du principe physique mis en œuvre par les capteurs lors de l'acquisition de l'information. Il n'est pas possible de l'éliminer comme les deux autres types d'erreurs mais il peut être évalué et intervient donc par la suite dans l'estimation de la qualité des données [SHOM & Debese, 2006]. Ce type d'erreur peut aussi se référer au « bruit aléatoire » dans une image et au « bruit blanc » en statistiques. Il possède une distribution normale et l'application d'un filtrage réduit ses effets [Zhilin et al., 2005].

Par ailleurs, à cela s'ajoutent les erreurs dues aux traitements appliqués sur les données brutes, à leur stockage dans un modèle ou aux algorithmes d'interpolation utilisés pour convertir un modèle à l'autre ou le ré-échantillonner.

Précision et résolution

La notion de précision peut présenter certaines ambiguïtés car elle comporte plusieurs significations (degré de détail, degré d'exactitude des mesures, ... etc.) et est souvent confondue avec la précision de positionnement : une mesure précise au centimètre près (20.25 m) peut comporter un erreur de positionnement de l'ordre du mètre (+/- 1m).

[Ruas, 2002b] définit trois termes pour exprimer la qualité des données :

- la précision qui est le degré de détail dans les mesures (précision des données),

- l'exactitude, qui représente l'écart entre la mesure stockée dans le modèle et une mesure de référence qui ne serait entachée d'aucune erreur (précision de positionnement),
- la résolution géométrique, qui est la taille du plus petit objet identifiable,

Pour cette dernière il conviendrait mieux de parler d'écartement entre cellules, qui représente l'aire couverte par un pixel. En effet le terme de résolution implique la notion de pouvoir séparateur, propre à une image et non à un MNT. Il existe deux résolutions différentes : la résolution planimétrique ou spatiale, relative à la **position planimétrique** des points, et la **résolution altimétrique** relative à l'unité de mesure de l'altitude des points.

Tableau 3 : Caractéristiques des MNT selon les données source et les techniques utilisées, d'après [Charleux-demargne, 2001].

	Résolution planimétrique	Estimation de la précision altimétrique
Cartes topographiques + digitalisation	10-100 m	5-50m
Photos aériennes + Photogrammétrie	2-30 m	<4m
Images radar ERS + interférométrie	30-40 m	5-25m
Altimétrie laser (Lidar)	+/- 1.5 m	+ /- 10cm
Sonde bathymétrique multifaisceaux (+ GPS cinématique)	< 5cm	+/- 15 cm [rapport SHOM]

Le récapitule les résolutions planimétriques et altimétriques rencontrées en fonction des types de données : cartes papier scannées, photos aériennes, images satellites, Lidar, sondes bathymétriques. L'utilisation du GPS cinématique¹¹ avec les dernières technologies Lidar et échosondeur permet d'obtenir une précision planimétrique supérieure à la précision altimétrique.

I.3.2 Fusion de MNT : travaux en cours ou déjà réalisés

La fusion de MNT en domaine littoral couple l'ensemble des problèmes :

- relatifs à la précision de positionnement planimétrique/altimétrique,
- liés à l'absence de données,
- liés à une divergence de correspondances entre les modèles à fusionner.

Les données terrestres et maritimes sont acquises par des organismes distincts qui ne travaillent pas avec les mêmes référentiels ni avec les mêmes outils d'acquisition. Par ailleurs les exigences de production des données sont également variables : la sûreté de navigation impose par exemple le choix d'un zéro hydrographique qui assure une hauteur d'eau minimale en toutes circonstances. Ce problème ne se pose pas en domaine terrestre où un niveau moyen des mers est retenu comme niveau d'origine. La résolution des conflits d'usages en domaines maritimes nécessite également des outils de gestion qui exploitent au mieux la précision des données existantes ([Tillier & Robin, 2008], [Gourmelon et al., 2007]). La fusion de MNT terre/mer est fondamentale pour la Gestion Intégrée des Zones Côtières (GIZC) ainsi que la prévention des risques en domaine côtier comme les marées noires, tsunamis, glissements de terrain ([Delacourt et al., 2007]). Plusieurs programmes d'études ont donc été engagés depuis quelques années ainsi que des congrès concernant cette thématique

¹¹ Le GPS cinématique est un instrument de mesure qui offre une précision centimétrique. Il utilise les satellites en orbite autour de la Terre afin de positionner précisément le lieu de réalisation des mesures. Cinq satellites au minimum sont nécessaires afin d'assurer un calcul de positionnement précis.

([EuroSDR, 2007], [Seatechweek, 2006]) ou ses applications ([Saferseas, 2007]). L'objectif des travaux de recherche amorcés est de proposer un modèle continu relativement standard et/ou des méthodes d'intégration optimales. Ils sont présentés avec leur contexte de mise en place dans les paragraphes suivants.

I.3.2.1 Intégration de données Terre/ Mer

ICZMap

Depuis 2001 en Angleterre, l'OS (Ordnance Survey) et l'UKHO (United Kingdom Hydrographic Office) collaborent à un projet de recherche sur la cartographie des zones côtières : ICZMap® (Integrated Coastal Zone Mapping - Data research project, 2001)¹².

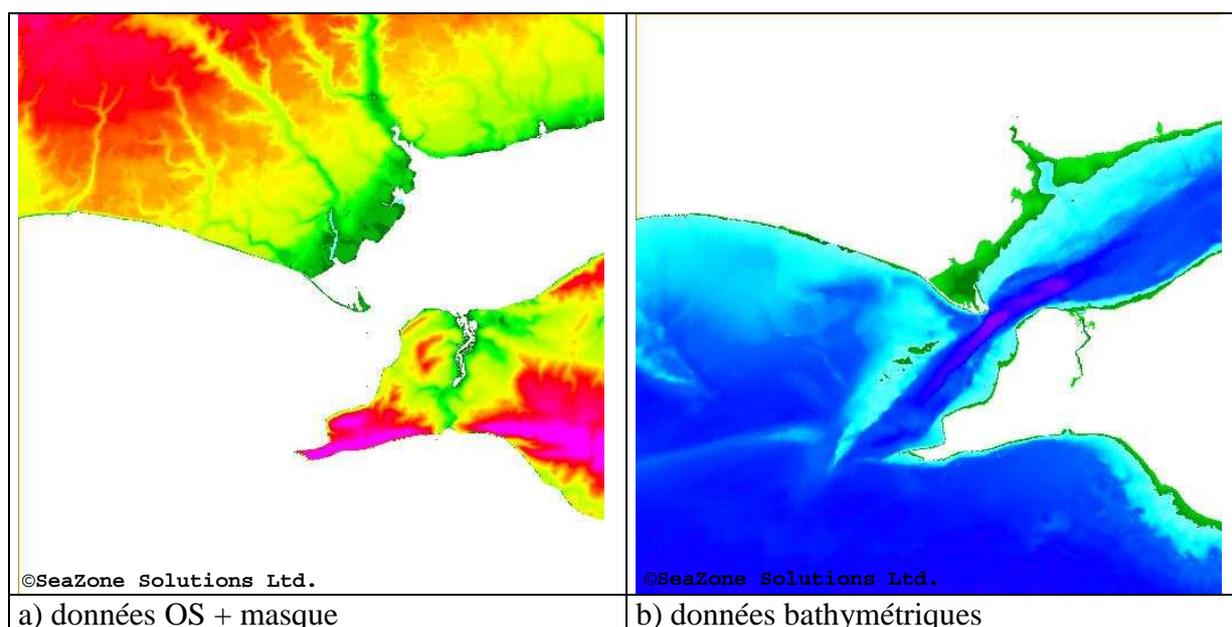


Figure 24 : Présentation du masque appliqué sur les données topographiques et bathymétriques successivement sur une zone test (Sud de l'Angleterre : Solent, près de Portsmouth).

Le but est de créer une base de données intégrées accessible à tout utilisateur, à partir des données existantes sur le littoral (terrestres, géologiques, marines) de façon à produire au final un MNT continu. Il s'agit également de produire une méthodologie de transformation des multiples zéros de référence. Ainsi l'ensemble des altitudes peut être exprimé de façon cohérente quel que soit le zéro de référence. Ce projet décrit les incohérences entre données : projections, systèmes géodésiques, et zéros de référence différents, échelles d'acquisition variables et spécifications non homologues. Il propose en réponse des techniques de conversion de données entre systèmes de référence. Une fois qu'un système commun est défini, les données 2.5D de l'OS et de l'UKHO [ICZMap, 2004] sont découpées selon un masque puis fusionnées.

Dans le cadre de la collaboration avec la société Seazone chargée de réaliser cette intégration, celle-ci nous a communiqué certains de ses résultats et fait participer à son projet, notamment pour la quantification des écarts entre données OS et UKHO [Mascret et al., 2006]. La figure 24 montre successivement les données terrestres (produites par l'OS) et les données bathymétriques découpées à l'aide du masque sur la zone. La ligne de séparation du masque est établie à partir du trait de côte fourni par les données vecteur OS Mastermap qui

¹² http://www.thsoa.org/hy03/4b_2.pdf et <http://www.seazone.com/files/finalreport.pdf>

présentent l'avantage d'avoir une précision bien meilleure que celles fournies par la norme S-57. Le format S-57 est le format normalisé de l'Organisation Hydrographique Internationale (OHI) dans lequel les données marines sont distribuées. Le but de cette norme est entre autres de fournir des exigences minimales en termes de qualité et de précision concernant la représentation de l'information maritime.

Des profils de chaque type de MNT (marin et terrestre) positionnés dans le même système cartographique ont été comparés visuellement. L'objectif est de vérifier si le recalage planimétrique est suffisamment satisfaisant pour permettre une fusion.

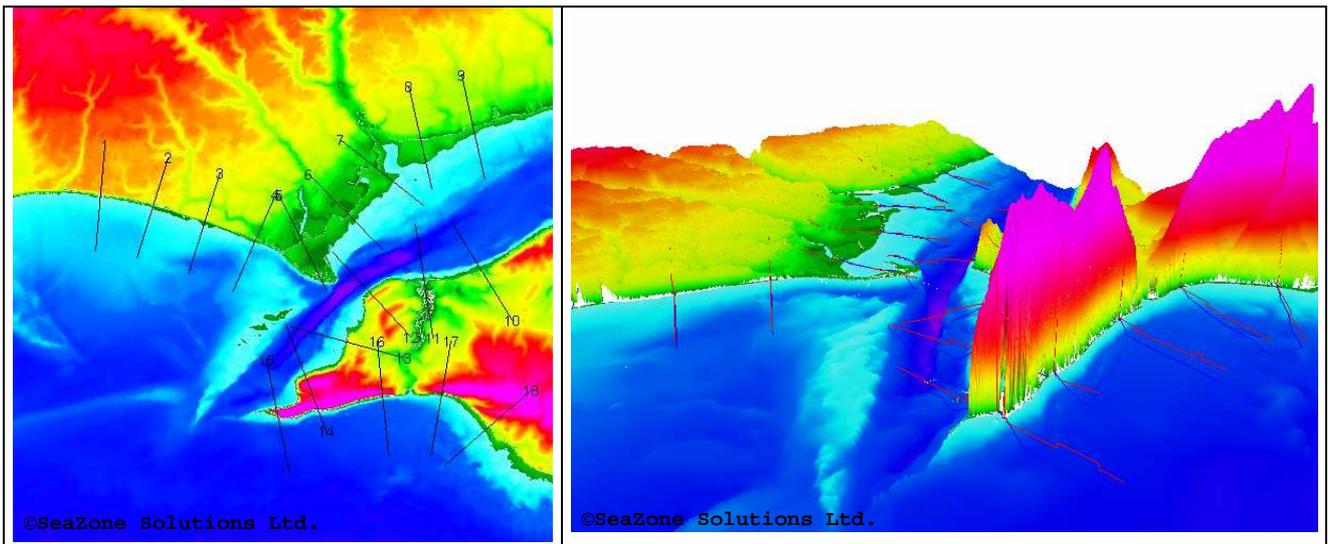
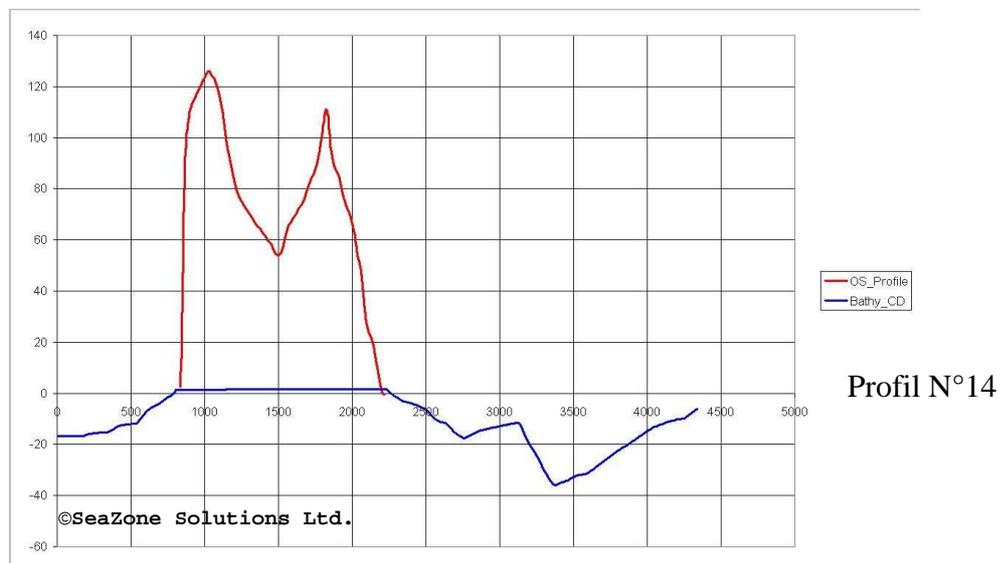


Figure 25 : Placement des profils topo-bathymétriques réalisés sur les modèles superposés et visualisation en relief avec une exagération verticale de 10.



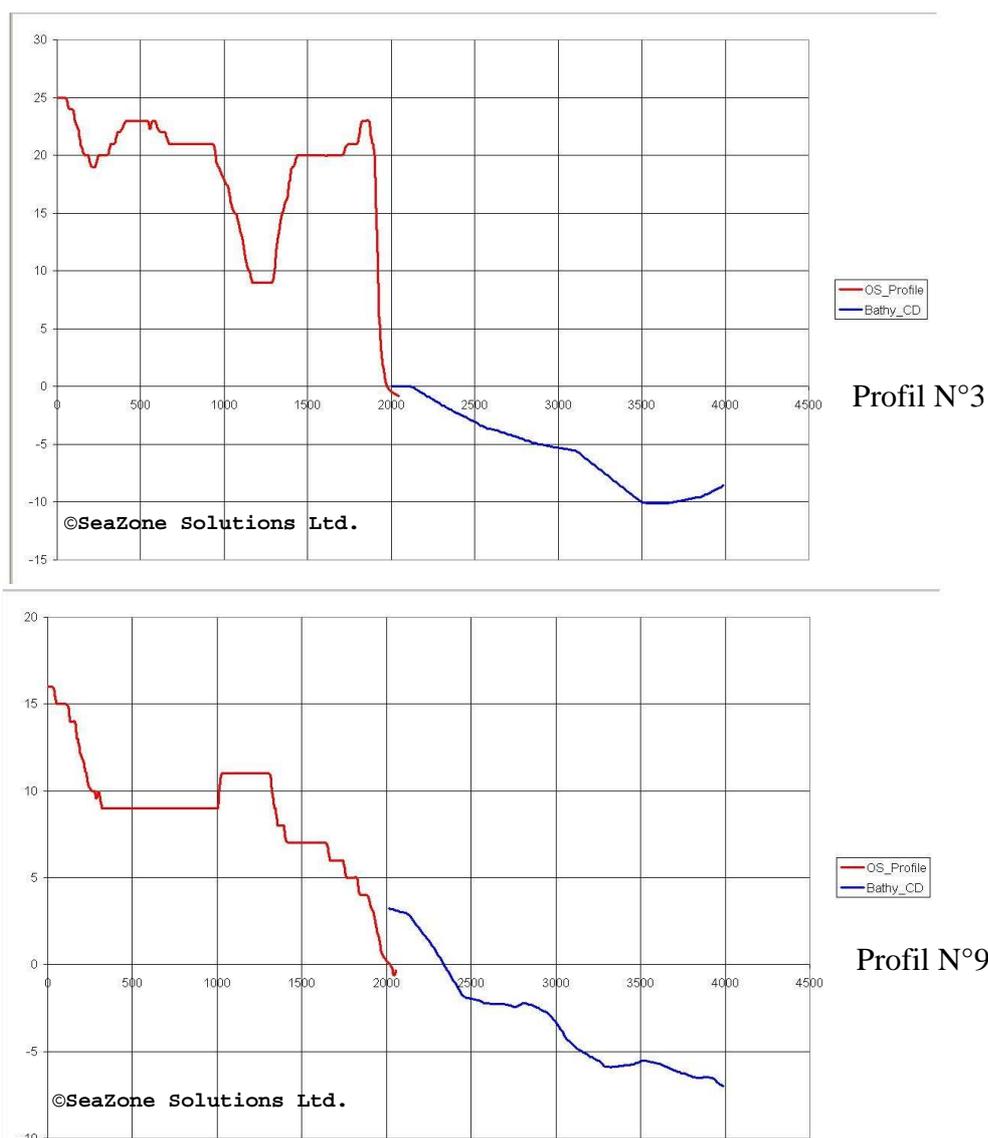


Figure 26 : Dessins des profils numéro 14, 3 et 9 présentés en Figure 25.

La figure 25 montre la position de la série de profils effectués sur la zone test. Trois des profils les plus pertinents sont présentés figure 26 : le premier cas (profil N°14) montre la superposition des deux MNT ainsi que l'interpolation réalisée sur le domaine terrestre de la partie bathymétrie. Le profil N° 3 présente un décalage moyen sur la zone d'étude tandis que le profil N°9 expose l'écart maximal entre données, soit près de quatre mètres, après correction géographique.

Cette étude est encore en phase d'analyse. Les algorithmes de fusion sont en cours de développement et à l'heure actuelle aucun modèle n'a été présenté ni validé qualitativement.

Litto 3D

En France, plusieurs travaux ont été entrepris avec pour objectif de dresser l'inventaire qualitatif des données existantes, en établir un catalogue, et définir celles qui serviront de référence nécessaire à la création et au fonctionnement d'une banque d'information géographique littorale (CNIG GT littoral, 03). Une des réponses apportées fut l'élaboration du RGL (référentiel géographique littoral) basé sur le RGE (référentiel à grande échelle) de l'IGN. Le RGL a pour objectif de fournir à la communauté d'utilisateurs une information géographique cohérente et fiable sous la forme d'un ensemble minimal de données de

référence. Il comporte notamment certaines composantes du RGE (limites administratives, parcellaires) et d'autres spécifiques : orthophotographie littorale, altimétrie continue terre-mer. Sa future élaboration est issue de la collaboration du SHOM et de l'IGN pour aboutir à la mise en place d'un programme d'observation et de cartographie continues du littoral : Litto3D. Ce projet a pour finalité la production d'une cartographie continue Terre/Mer de la frange côtière, de l'estran, et d'une partie de la mer territoriale. Considérant que les levés actuels SHOM/IGN sont trop incohérents, trop imprécis pour certains et de recouvrement insuffisant pour modéliser les phénomènes côtiers, de nouvelles acquisitions doivent servir de base à la production de nouveaux modèles. Les techniques d'acquisition mises en œuvre comportent : pour les grands fonds des échosondeurs multifaisceaux, des sondeurs petit fond pour des profondeurs moins importantes, et du lidar topo-bathymétrique pour compléter les zones de zéro à un mètre, non relevées avec les sondeurs. Ce type de Lidar permet selon la turbidité de l'eau d'acquérir des données à des altitudes allant de quelques mètres jusqu'à moins dix mètres.

Le résultat final doit se présenter sous la forme d'un MNT localisé dans une frange d'une dizaine de kilomètres de part et d'autre du trait de côte, sur la base duquel pourront se superposer différentes couches de données géographiques thématiques. Pour le moment aucun modèle continu n'a été produit via une méthode de fusion. L'accent a été mis sur l'évaluation de la qualité des données source et de leurs techniques d'acquisition.

Les premiers prototypes de résultats ont été acquis dans le golfe du Morbihan, zone d'expérimentation assez délicate en raison de son modèle de marée complexe. Les tests préliminaires ont permis de calibrer les capteurs afin de produire les données les plus précises possibles avec un recouvrement maximal. En effet, la turbidité de l'eau empêche dans certains cas d'avoir une couverture continue. De plus certaines questions techniques sur la mise au point de la méthodologie d'acquisition ont été soulevées (levés Lidar avant ou après le levé SMF ?, à marée haute ou basse ?, de jour ou de nuit ?, etc.) sans pour autant avoir de réponses pleinement satisfaisantes. En attendant que la couverture complète soit réalisée, une base de données « historique » a été constituée à partir des données existantes du SHOM et de l'IGN : Histolitt [Louvart & Grateau, 2005]. Celle-ci rassemble des informations topographiques, bathymétriques, des modèles de marée et le trait de côté. En domaine terrestre, les données s'étendent soit jusqu'à l'altitude dix mètres, soit de trois à dix kilomètres en planimétrie dans les terres, et en domaine marin, jusqu'à six miles marins du trait de côte.

MNT intégré sur la Baie du Mont-Saint-Michel

Un MNT intégré de la baie du Mont Saint Michel combinant des données terrestres et maritimes d'origines différentes a été réalisé par [Juc & Deffontaines, 2003]. Les données initiales se composent de la BD Alti® de l'IGN, de levés topographiques, de films négatifs noirs et blancs issus de la mission IPLI (Inventaire Permanent du Littoral) et de cartes scannées de sondes bathymétriques. Cette hétérogénéité entraîne de nombreuses difficultés lors du recalage des différents référentiels (terrestres et maritimes) entre eux, de la conservation de l'information et des calculs de corrélation automatique. Toutes ces données, hormis les levés de terrain, ont tout d'abord été convertis dans le même référentiel. Une pondération a ensuite été définie en fonction de l'altitude du point pour enfin produire le MNT intégré en utilisant un modèle quadratique de poids. La précision des données d'origine étant assez inégale, le MNT final ne présente pas non plus une qualité constante. C'est pourquoi les levés de terrain ont servi de points de référence pour évaluer la qualité du modèle final obtenu. Cette intégration présente l'avantage de mélanger plusieurs types de données et de bien inventorier les problèmes rencontrés. Cependant, le traitement décrit ici n'est développé que pour ce type de données précises et dans ce contexte particulier. Il n'est donc

pas généralisable à d'autres zones. Par ailleurs il s'agit de traitements manuels des données adaptés selon les problèmes rencontrés et appliqués en zone de terrain de très faible pente.

Projet Américain

Les deux organismes nationaux américains collecteurs de données terrestres (USGS) et maritimes (NOAA) ont initié une collaboration afin de développer des techniques et des outils facilitant la fusion de leurs données disponibles [Gesch & Wilson, 2001]. Le site choisi fut la Baie de Tampa pour des raisons de qualité des données.

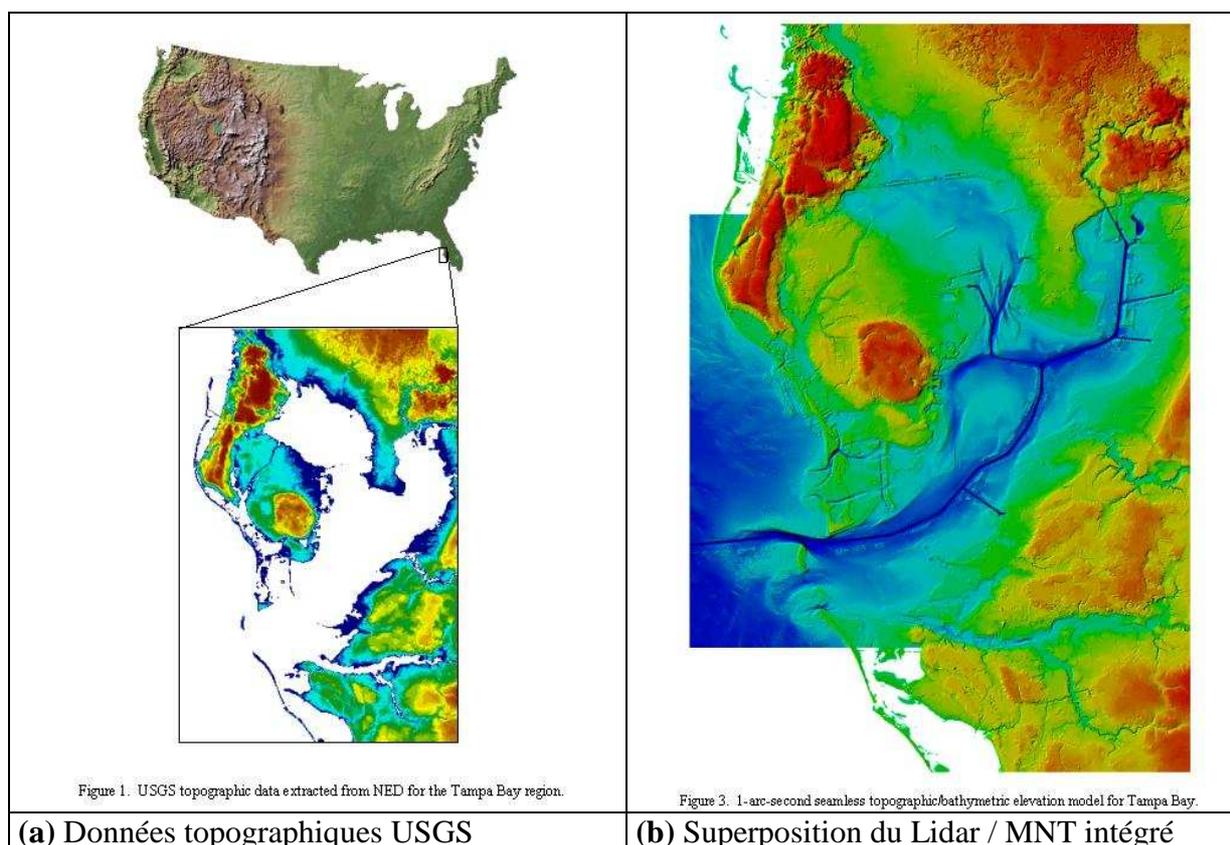


Figure 27 : Données topographiques seules et superposition avec un MNT Lidar de meilleure résolution.

La figure 27(a) représente l'étendue et la localisation des données topographiques USGS National Elevation Dataset (NED) de résolution dix mètres utilisées pour la fusion. La partie bathymétrie est extraite de la base de données des levés hydrographiques NOAA. Le recouvrement entre les données étant assuré partout, les points clés ont été la sélection des données les meilleures parmi celles disponibles, et la résolution des problèmes liés aux référentiels différents. Une fois les sondes bathymétriques transformées en grille et converties dans le même système cartographique que les autres données, une ligne de référence est choisie : le trait de côte équivalent au zéro. Celui-ci permet d'établir une sélection des points bathymétriques et topographiques (obtenus à partir des grilles) compris dans la zone de recouvrement définie par un buffer de 600m à partir du trait. Les points sélectionnés sont interpolés en une grille d'un arc seconde qui représente la zone de recouvrement. Afin d'éviter les effets de bords dus à l'interpolation, la grille obtenue est découpée à l'aide d'un buffer de 300m à partir du trait de côte sur les 600m initiaux. L'étape finale de traitement réalise une mosaïque entre les points NOAA grillés et transformés géographiquement, les grilles USGS et la grille de la zone de recouvrement. La fusion est effectuée avec la fonction MOSAIC d'ArcGrid qui accomplit une moyenne pondérée cellules à cellules selon leur

proximité au bord de la zone de recouvrement. L'étude est validée par l'incorporation dans le référentiel spatial créé d'autres modèles numériques comme du Lidar (présenté figure 27**(b)**). Cependant il faut noter qu'aucun contrôle qualité n'est présenté. La cohérence du paysage décrit est bien conservée malgré la différence de précisions (un mètre pour le Lidar contre dix pour le modèle continu), mais celui-ci ne présente pas de grosses difficultés en raison de sa très faible pente. Le résultat est bien un modèle continu Terre/Mer mais le problème du non recouvrement des données et du contrôle qualité du modèle obtenu n'est pas abordé.

I.3.2.2 Fusion de MNT terrestres avec recouvrement total

En matière de fusion de MNT avec contraintes géomorphologiques, un travail novateur a été proposé par [Podobnikar, 2005]. L'article présente une méthode de fusion basée sur l'intégration de MNT de résolutions variables pondérées par des coefficients. La valeur de ces coefficients est calculée en fonction de la qualité du modèle, ce qui permet d'améliorer la qualité du MNT final. Comme le modèle obtenu est néanmoins plus lisse que ceux de départ, un enrichissement géomorphologique est effectué en ajoutant des détails extraits du MNT de meilleure qualité. Cette méthode souffre de deux déficiences : une en raison de l'absence de méthodologie d'intégration, et l'autre au niveau de la non prise en compte des décalages planimétriques éventuels des MNT.

I.3.2.3 Bilan des travaux présentés

Toutes ces études décrivent les différents travaux développés pour fusionner au mieux des données hétérogènes issues de capteurs différents destinés à l'acquisition dans des milieux physiques opposés. Les trois premiers projets (ICZMap, [Juc & Deffontaines, 2003] et [Gesch & Wilson, 2001]) présentent l'avantage de se baser sur des données déjà existantes. Ils inventorient donc les problèmes rencontrés au niveau de la conversion d'un référentiel à l'autre et les traitements utilisés pour la fusion : découpage selon une ligne choisie puis collage ou somme pondérée. Parallèlement à ceci, ICZMap et [Juc & Deffontaines, 2003] établissent également une estimation de la qualité du modèle obtenu. Elle s'effectue soit à l'aide de profils réalisés perpendiculairement à la ligne de découpe (évaluation visuelle dans ICZMap), soit avec une comparaison utilisant des levés de terrain [Juc & Deffontaines, 2003]. Le processus de validation dans le projet américain s'est restreint à incorporer dans le modèle créé des données de bien meilleure précision comme du Lidar [Gesch & Wilson, 2001]. Pour le projet Litto3D l'accent a été mis sur le recensement de tous les problèmes amont possibles lors de la phase d'acquisition ainsi que sur la haute qualité des données obtenues. En revanche, comme aucun MNT n'a été produit, les données comportent entre elles des anomalies : présence de marches d'escalier au passage sonar/Lidar, de même que certains trous en raison de la turbidité de l'eau ou de la présence de navires pendant les mesures.

Toutes ces études proposent des approches complémentaires :

- homogénéisation du système de coordonnées pour les données existantes,
- transformation des données en un modèle commun,
- évaluation de la qualité du produit final pour certains,
- enrichissement des modèles de manière à atténuer l'effet de perte d'informations à chaque traitement
- prise en compte du caractère « morphologique » des MNT lors de la fusion.

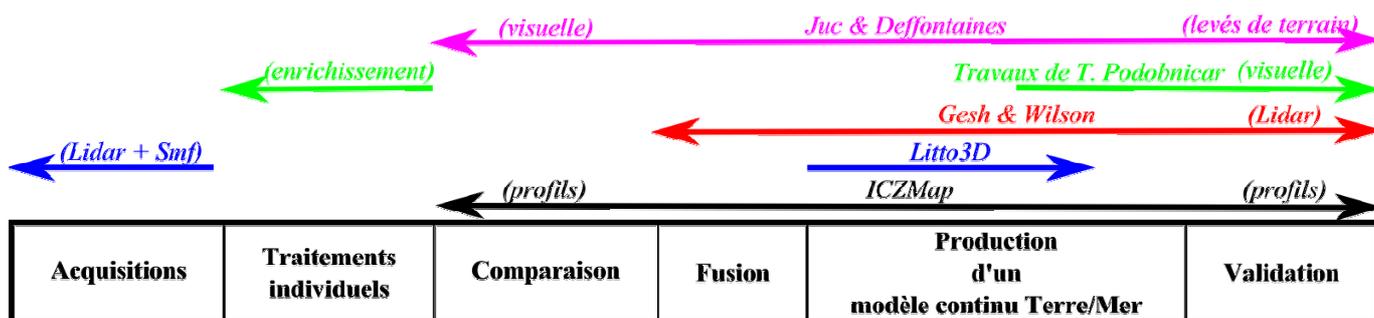


Figure 28 : Bilan des apports de chaque méthode dans le processus d'intégration de données.

Tableau 4 : résumé des points forts et faiblesses des différentes études présentées dans ce chapitre.

	Avantages	Inconvénients
ICZMap	<ul style="list-style-type: none"> - utilisation de données existantes - homogénéisation des données : techniques de conversion entre systèmes de référence - évaluation de la qualité du modèle produit 	<ul style="list-style-type: none"> - découpe selon une ligne de référence « artificielle » - prise en compte d'un biais global uniquement - validation visuelle
Gesh & Wilson	<ul style="list-style-type: none"> - utilisation de données existantes - validation par incorporation d'un autre modèle plus précis 	<ul style="list-style-type: none"> - problème de non recouvrement des données non traité - technique valable uniquement en zone très plate
Juc & Defontaines	<ul style="list-style-type: none"> - utilisation de données existantes - pondération selon la qualité des données - estimation de la qualité du modèle grâce à des levés de terrain 	<ul style="list-style-type: none"> - pas de généralisation possible de la chaîne de traitements
Podobnikar	<ul style="list-style-type: none"> - utilisation de données existantes de précisions différentes - enrichissement - prise en compte du caractère morphologique des MNT 	<ul style="list-style-type: none"> - problème de non recouvrement des données non traité - pas de méthodologie d'intégration - pas de prise en compte des décalages planimétriques entre MNT - validation uniquement visuelle
Litto3D	<ul style="list-style-type: none"> - acquisition de données 	<ul style="list-style-type: none"> - pas de lien avec les données existantes - obligation de réacquérir toutes les zones littorales

La figure 28 résume les apports majeurs de chaque méthode lors du processus d'intégration. Ces contributions se placent soit en amont lors de l'acquisition des données (projet Litto3D), soit en aval lors de la mise en cohérence des données existantes entre elles. A ce niveau, elles se situent :

- lors de la phase de traitement avec l'enrichissement géomorphologique qui permet de compenser la perte de détails (lissage) liée à la fusion elle-même (travaux de T.Podobnicar) ;
- lors de la comparaison des modèles entre eux, visuellement ou à l'aide de profils ;
- au niveau de la validation qui est soit qualitative, par comparaison avec des données de référence (travaux en Baie du Mont St Michel), soit visuelle (ICZMap), soit par intégration d'un modèle plus précis avec celui déjà produit (projet Américain).

Les points fort présentés par ces études (résumés pour chacune dans le tableau 4) sont : (1) une bonne description des problèmes liés à la fusion de données déjà existantes, et plus particulièrement de données terre/mer ; (2) une évaluation de la qualité finale du modèle pour certaines ou (3) une incorporation de données de plus grande précision dans le modèle continu réalisé et (4) un enrichissement des données après intégration. En revanche, aucune méthodologie générique n'est développée, les décalages planimétriques persistants après transformation géographique ne sont pas traités et aucune méthode ne préserve les détails du modèle (éléments principaux du relief) avant la fusion. Il est donc nécessaire de développer des méthodes d'intégration de MNT plus complètes que la fusion d'altitudes repositionnées en X et Y avec une interpolation des trous présents entre les modèles.

1.4 Bilan sur l'état de l'art

Le début de ce chapitre présente le milieu littoral d'un point de vue géographique : plus qu'une simple ligne de démarcation entre trois domaines, c'est avant tout une interface ambivalente sous influence des dynamiques continentale et maritime nettement différentes. Les actions initiées par les différents agents morphogéniques sont donc décrites, ainsi que leur impact sur le relief. Certaines actions possèdent un caractère marin dominant, comme les processus mécaniques engendrés par les vagues ou les courants marins, tandis que d'autres naissent de la conjonction des deux influences, comme les modifications dues aux organismes biologiques. Indépendamment de la lithologie, leur influence sur le littoral dépend principalement du contexte climatique dans lequel ils apparaissent.

Dans la seconde section sont définies les notions de comparaison, d'appariement et de fusion d'objets 2D, étapes indispensables pour la réalisation d'une intégration de données. Le processus de comparaison donne une valeur de similitude (mesure de distance ou autre) entre les objets afin de déterminer s'ils sont potentiellement homologues. L'appariement établit les correspondances, totales ou partielles, entre ces objets. La fusion produit un seul jeu de données finales cohérentes de même qualité, mais de contenu thématique ou d'emprise plus étendue que les données de départ considérées individuellement. Ces trois processus présentés se basent sur le positionnement géographique des entités (points, lignes ou polygones) lorsque des identifiants communs n'existent pas.

La troisième section de ce chapitre est consacrée à la présentation et la description des MNT, données plus complexes que les précédentes car possédant une dimension de plus. Une première partie présente leurs caractéristiques, modes d'acquisition, formats, qualités, tandis qu'une seconde partie décrit le processus d'intégration de ces données au travers des travaux déjà réalisés. Nous y découvrons que les étapes comparaison/appariement précédant la fusion sont absentes mais remplacées par une étape de mise en cohérence des données au niveau du référencement géographique ou du format. Le processus de fusion consiste en une moyenne, pondérée ou non et pixel à pixel. De plus aucune de ces études ne propose de réelle méthodologie globale d'intégration ni ne développe des outils destinés à l'intégration

automatique de données d'origines diverses. Par ailleurs, la question de zones de non recouvrement n'est jamais abordée.

L'apport de cette thèse dans le domaine est donc de définir une méthodologie générique d'intégration de MNT hétérogènes en zone côtière. Pour cela il faut également développer des outils destinés à l'intégration de modèles existants. Ces outils doivent par ailleurs être assez robustes pour pouvoir mettre en cohérence les MNT continus obtenus avec d'autres données à haute précision issues de programmes comme Litto3D. Pour ce faire, ces outils doivent détecter les éléments structurants du relief et réaliser l'intégration des modèles en se basant sur ces éléments. Cet impératif permet tout d'abord de s'affranchir des problèmes de recalage de données et préserver les informations clés (points, lignes, ou surfaces remarquables) du paysage en adaptant les pré et post traitements des MNT à intégrer. Ainsi les éléments homologues utilisés ont une signification géographique : l'appariement est donc réalisé sur des éléments réellement pertinents. Cette approche amène à l'extraction d'un grand nombre de lignes dont les propriétés sont très différentes (ouverte, fermée, partielle). Le chapitre suivant explique donc comment l'algorithme de Fréchet a été étendu afin de gérer la complexité des lignes extraites.

II EXTENSIONS DE LA DISTANCE LINEAIRE DE FRECHET ET ALGORITHMES D'APPARIEMENT

Sommaire

II.1	Extension 1 : distance moyenne de Fréchet et couples de points homologues..	63
II.2	Extension 2 : distance de Fréchet sur des lignes d'emprises ou de natures différentes.....	67
II.2.1	Distance de Fréchet discrète partielle	67
II.2.2	Distance de Fréchet discrète entre deux lignes fermées	68
II.2.3	Distance de Fréchet discrète partielle entre une ligne ouverte et une ligne fermée.....	72
II.2.4	Distance de Fréchet discrète partielle / partielle entre deux lignes ouvertes	73
II.3	Algorithme d'appariement	78
II.3.1	Recherche de la nature des lignes comparées	79
II.3.2	Relations d'emprises entre les lignes.....	79
II.3.3	Types d'appariement entre lignes.....	81
II.3.3.1	Appariement simple « 1 – 1 » de deux lignes :.....	81
II.3.3.2	Appariement total/partiel « 1 – P » entre deux lignes.....	82
II.3.3.3	Appariement « partiel/partiel » (PP) et « deux partiel/partiel » (2PP) entre deux lignes	82
II.3.4	Méthode globale : appariement par fichier.....	86
II.4	Exemple d'applications de la méthode sur la comparaison de traits de côte ...	89
II.4.1	Présentation des données et de la zone d'étude	89
II.4.2	Méthode de comparaison des traits de côte	89
II.4.3	Résultats et perspectives	90
II.5	Bilan sur les apports des extensions.....	93

Le chapitre Précédent a montré que les éléments caractéristiques importants pour l'intégration de MNT étaient essentiellement linéaires. Il faut donc disposer de distances permettant de comparer, appairer puis intégrer ces éléments. L'algorithme retenu pour réaliser l'intégration de données terre/mer à l'issue de cet inventaire est celui basé sur la distance de Fréchet car elle est la mesure la plus appropriée pour comparer des éléments linéaires complexes (thalwegs crêtes, etc.) Celle-ci, introduite par [Fréchet, 1906], se base sur l'assimilation de toute ligne à une suite de points orientés équivalente à une fonction continue. Le résultat est une distance maximale entre les deux lignes comparées. La difficulté à utiliser ce calcul vient de sa complexité à programmer. [Eiter & Mannila, 1994] proposent une solution en développant une estimation de cette distance : la distance de Fréchet discrète.

Cette distance a pour intérêt de réaliser un appariement basé sur les points homologues des lignes comparées et non les points les plus proches. Cependant, telle qu'elle a été définie, elle est insuffisante pour répondre à toutes les attentes et contraintes d'un appariement de données géographiques. En effet l'algorithme de départ ne s'applique qu'à des lignes ouvertes, sensiblement de même taille et d'orientations identiques. Dans le cadre de l'appariement, il est nécessaire de l'adapter afin de pouvoir comparer des lignes d'emprises, de tailles, de natures et d'orientations différentes.

Dans le cas de données réelles, les lignes numérisées (1) peuvent présenter l'aspect de lignes « fermées » (polygones dont les points sont ordonnés) et (2) ne sont pas obligatoirement saisies dans le même sens. Il convient donc de déterminer avant toute chose la « nature » d'une ligne : si le premier point est identique au dernier, alors la ligne représente une ligne fermée. Ensuite il faut également s'assurer des orientations respectives des lignes entre elles. Il s'agit de vérifier que les suites ordonnées de points de chaque ligne ouverte sont bien dans le même sens et que les points de début et de fin sont homologues et non diamétralement opposés.

Les deux extensions développées dans ce chapitre se proposent donc de répondre aux différents problèmes posés lors de la comparaison de lignes en inventoriant les différents cas de figure rencontrés.

Les algorithmes détaillés par la suite expliquent comment utiliser au mieux ces distances lors des processus d'appariement de lignes. La première partie présente succinctement les lignes et les attributs qui sont utilisés dans les processus de comparaison et d'appariement.

La seconde partie détaille toutes les distances introduites pour la comparaison de lignes :

- la distance moyenne tout d'abord, qui permet d'avoir une estimation de l'écart moyen entre les lignes. Elle donne une information complémentaire à la distance maximale définie par Fréchet ;
- la distance de Fréchet partielle ensuite, définie par [Devogele, 2002] pour la comparaison de lignes de tailles variables ;
- la distance de Fréchet discrète entre deux lignes fermées ;
- la distance de Fréchet discrète partielle destinée à comparer des lignes de tailles variables et de natures différentes (ligne ouverte ou fermée) ;
- et enfin la distance de Fréchet discrète partielle calculée entre deux segments de lignes.

La troisième partie se base sur ces distances pour détailler les algorithmes d'appariement. La première section présente les étapes d'identification de la nature des lignes comparées. La seconde présente les différents types d'inclusion qu'il existe entre les emprises des lignes étudiées. Et la dernière section explique quel appariement est effectué selon les différents cas de figure identifiés et quelle distance est adaptée. Une dernière section réalise une synthèse en intégrant ces appariements simples dans un algorithme de traitement global de deux jeux de polygones.

Finalement, la dernière partie sert de validation des algorithmes appliqués au cas de comparaison de lignes en présentant le résultat d'une étude sur le suivi du trait de côte. Cette étude s'est réalisée au moyen d'indicateurs visuels extraits d'images satellites et de la distance de Fréchet pour la quantification des déplacements entre une ligne de référence et une ligne numérisée à une date donnée.

II.1 Extension 1 : distance moyenne de Fréchet et couples de points homologues

La distance de Fréchet discrète (d_{dF}) représente l'écartement maximal entre deux points homologues des lignes comparées. Afin de compléter l'information fournie par cette mesure, d'autres distances, toutes dérivées de la d_{dF} (voir Chapitre I), ont été introduites. La première distance exposée dans cette section, la **distance de Fréchet moyenne** (d_{aF}), permet l'évaluation de l'écartement moyen global entre les points homologues des lignes. Elle est la moyenne des distances euclidiennes entre couples de points issus du **chemin minimum**.

Le « chemin minimum » est composé d'une suite de couples de points $\{(L_{1.1}, L_{2.1}), \dots, (L_{1.n}, L_{2.m})\}$ pour lesquels la distance d'écartement entre les points du couple est inférieure ou égale à d_{dF} . Les points ainsi retenus sont appelés **points homologues**.

Pour reprendre l'exemple du maître et de son chien introduit au chapitre II, plusieurs chemins possibles respectent la distance de Fréchet. Parmi ces chemins candidats, le « chemin minimum » constitue celui pour lequel le promeneur et son chien évoluent le plus proche possible l'un de l'autre à chacun de leur déplacement.

Pour calculer ce chemin, les deux matrices de distance et de Fréchet sont nécessaires ainsi que l'opérateur inférieur ou égal (\leq) entre deux couples de réels. Cet opérateur se définit ainsi :

$$\begin{aligned} (a,b) \text{ et } (c,d) &\in \mathfrak{R}^2 & (1) \\ (a,b) \leq (c,d) &\text{ si } a < c \text{ ou si } a = c \text{ et } b \leq d \end{aligned}$$

où « a » et « c » représentent les valeurs de cellules de la matrice de Fréchet (MF). De même « b » et « d » sont les valeurs des cellules de la matrice de distance (MD). L'opérateur inférieur ou égal (\leq) va permettre de choisir le couple de points homologues parmi ceux possibles. Pour les lignes de la figure 29, les matrices de distance et de Fréchet ont été calculées dans le **Erreur ! Source du renvoi introuvable.** A partir de ces matrices sont extraits les couples MF(i, j) et MD(i, j) replacés dans la cellule (i, j) du tableau 6.

Le chemin minimum (C_M) de la figure 30 est construit par « retour arrière » (« backtracking ») à travers les matrices. Il débute donc par le couple de points final : $(L_{1.n}, L_{2.m})$. Le couple suivant, $(L_{1.i}, L_{2.j})$, est recherché parmi trois couples candidats, $(L_{1.i-1}, L_{2.j-1})$, $(L_{1.i-1}, L_{2.j})$ et $(L_{1.i}, L_{2.j-1})$, à l'aide de l'opérateur \leq défini précédemment. La valeur de la MF associée au couple de points considéré signifie que pour aller du couple de départ au couple considéré, la distance de Fréchet vaut MF. Le couple retenu parmi les trois possibles est celui qui permet d'utiliser le sous-chemin le plus petit possible.

Par exemple entre les trois couples $(L_{1.7}, L_{2.7})$, $(L_{1.7}, L_{2.6})$, $(L_{1.8}, L_{2.6})$ de la figure 29, le couple $(L_{1.7}, L_{2.7})$ n'est pas retenu car $(MF_{7.7}) > (MF_{6.7})$ et $(MF_{6.8})$. Entre les deux autres couples, la valeur de la distance de MF est égale ce qui signifie, si l'analogie du maître et de son chien est

reprise, que les deux chemins passant par ces couples utilisent une laisse de même longueur. Dans ce cas, le couple choisi est celui dont les points sont les plus proches (dans la matrice de distance MD) : $(MD_{6,7}) < (MD_{6,8})$.

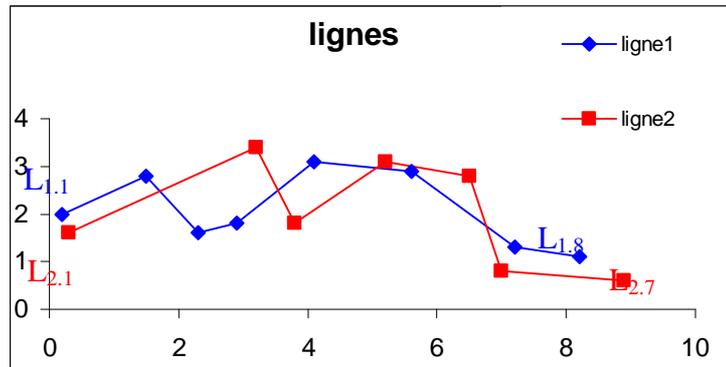


Figure 29 : Couple de lignes homologues.

Tableau 5 : Matrice de distance (euclidienne) et matrice de Fréchet des lignes de la figure 29

x1i	0.2	1.5	2.3	2.9	4.1	5.6	7.2	8.2
y1j	2	2.8	1.6	1.8	3.1	2.9	1.3	1.1

Matrice de Distance (MD)

x2i	y2j		1	2	3	4	5	6	7	8
0.3	1.6	1	0.41	1.70	2.00	2.61	4.09	5.46	6.91	7.92
3.2	3.4	2	3.31	1.80	2.01	1.63	0.95	2.45	4.52	5.50
3.8	1.8	3	3.61	2.51	1.51	0.90	1.33	2.11	3.44	4.46
5.2	3.1	4	5.12	3.71	3.26	2.64	1.10	0.45	2.69	3.61
6.5	2.8	5	6.35	5.00	4.37	3.74	2.42	0.91	1.66	2.40
7	0.8	6	6.91	5.85	4.77	4.22	3.70	2.52	0.54	1.24
8.9	0.6	7	8.81	7.72	6.68	6.12	5.41	4.02	1.84	0.86

Matrice de Fréchet (MF)

			1	2	3	4	5	6	7	8
1			0.41	1.70	2.00	2.61	4.09	5.46	6.91	7.92
2			3.31	1.80	2.01	2.00	2.00	2.45	4.52	5.50
3			3.61	2.51	1.80	1.80	1.80	2.11	3.44	4.46
4			5.12	3.71	3.26	2.64	1.80	1.80	2.69	3.61
5			6.35	5.00	4.37	3.74	2.42	1.80	1.80	2.40
6			6.91	5.85	4.77	4.22	3.70	2.52	1.80	1.80
7			8.81	7.72	6.68	6.12	5.41	4.02	1.84	1.80

Le processus est itéré selon l’algorithme décrit figure 30 jusqu’au couple de départ $C_{1,1}$. Le chemin minimum obtenu est représenté dans le tableau 6 sous forme de cellules grisées contenant les couples de réels. C’est une suite ordonnée de neuf couples (voir

tableau 6) : $(C_M) = \{C_{1,1} (L_{1,1}, L_{2,1}), C_{2,2} (L_{1,2}, L_{2,2}), C_{3,3} (L_{1,3}, L_{2,3}), C_{3,4} (L_{1,4}, L_{2,3}), C_{4,5} (L_{1,5}, L_{2,4}), C_{4,6} (L_{1,6}, L_{2,4}), C_{5,6} (L_{1,6}, L_{2,5}), C_{6,7} (L_{1,7}, L_{2,6}), C_{7,8} (L_{1,8}, L_{2,7})\}$.

Cette programmation dynamique avec l’opérateur \leq donne des résultats plus rapidement que dans celle proposée par [Devoegele, 2002]. En effet dans [Devoegele, 2002], l’ensemble des chemins possibles sont testés avant de choisir celui qui est le plus optimal, ce qui peut être très long pour des lignes comportant énormément de points. Avec la méthode par backtracking présentée ici, seules $(N+M)$ comparaisons au maximum sont effectuées.

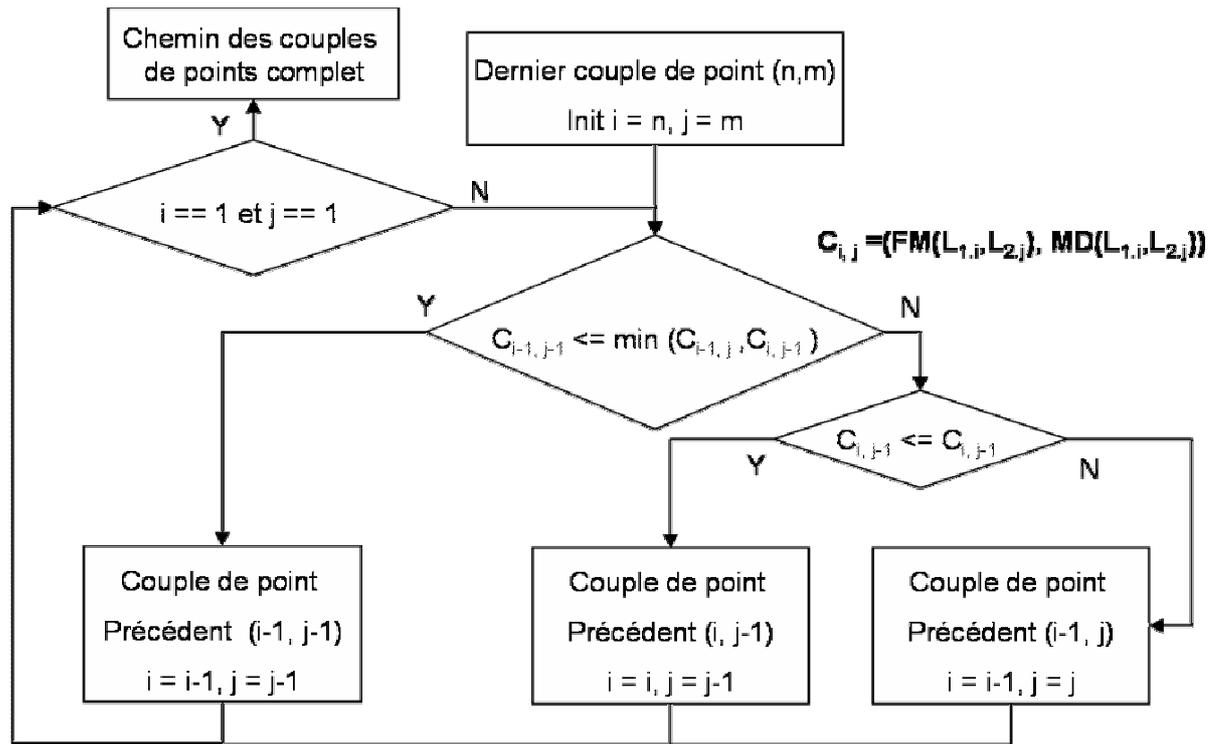


Figure 30 : Algorithme pour le calcul du chemin minimum.

Tableau 6 : Le « chemin minimum » des lignes $(L_{1,i}, L_{2,i})$ et $(L_{1,m}, L_{2,m})$ est défini en sélectionnant les couples de réels en grisés.

	1	2	3	4	5	6	7	8
1	0.41 0.41	1.70 1.70	2.00 2.00	2.61 2.61	4.09 4.09	5.46 5.46	6.91 6.91	7.92 7.92
2	3.31 3.31	1.80 1.80	2.01 2.01	2.00 1.63	2.00 0.95	2.45 2.45	4.52 4.52	5.50 5.50
3	3.61 3.61	2.51 2.51	1.80 1.51	1.80 0.90	1.80 1.33	2.11 2.11	3.44 3.44	4.46 4.46
4	5.12 5.12	3.71 3.71	3.26 3.26	2.64 2.64	1.80 1.10	1.80 0.45	2.69 2.69	3.61 3.61
5	6.35 6.35	5.00 5.00	4.37 4.37	3.74 3.74	2.42 2.42	1.80 0.91	1.80 1.66	2.40 2.40
6	6.91 6.91	5.85 5.85	4.77 4.77	4.22 4.22	3.70 3.70	2.52 2.52	1.80 0.54	1.80 1.24
7	8.81 8.81	7.72 7.72	6.68 6.68	6.12 6.12	5.41 5.41	4.02 4.02	1.84 1.84	1.80 0.86

La figure 31 montre les couples de $(L_{1,i}, L_{2,j})$ associés au chemin minimum. Les points issus des couples sont considérés comme homologues. Cette figure montre qu'un point peut être associé à plusieurs points de l'autre ligne.

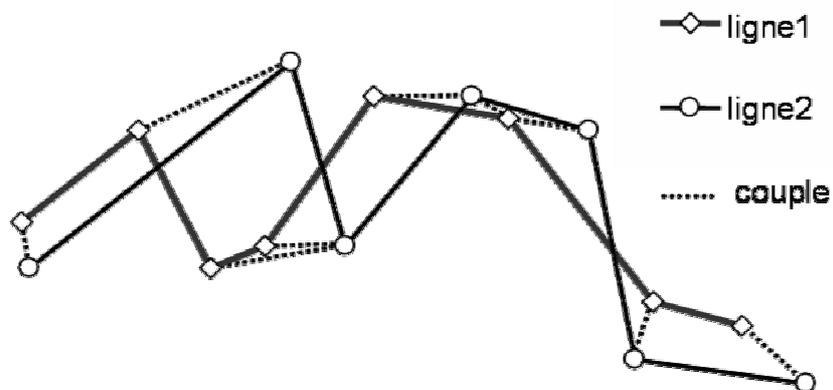


Figure 31 : Couples de points $(L_{1,i}, L_{2,j})$ du chemin minimum représentés par les lignes pointillées.

La distance de Fréchet moyenne est la moyenne des distances entre couples de points homologues issus du chemin minimum. Elle donne un indice de l'écartement global moyen des lignes alors que la distance de Fréchet fournit un écart maximal.

Soit k couples de points homologues $(L_{1,i}, L_{2,j}) \in C_M$:

$$d_{aF} = \frac{1}{k} \sum_k d(L_{1,i}, L_{2,j})$$

avec $(L_{1,i}, L_{2,j}) \in C_M$.

Pour l'exemple (d_{aF}) vaut 0,9524. Une valeur forte de la distance de Fréchet (d_{dF}) peut être due :

- soit à un « accident » local important des points d'une des deux lignes (cas (a) figure 32). Dans ce cas la valeur de la (d_{aF}) est bien meilleure ;
- soit à un mauvais calage général d'une ligne sur l'autre (cas (b) figure 32). La valeur de la (d_{aF}) est alors très proche de celle de la (d_{dF}) .

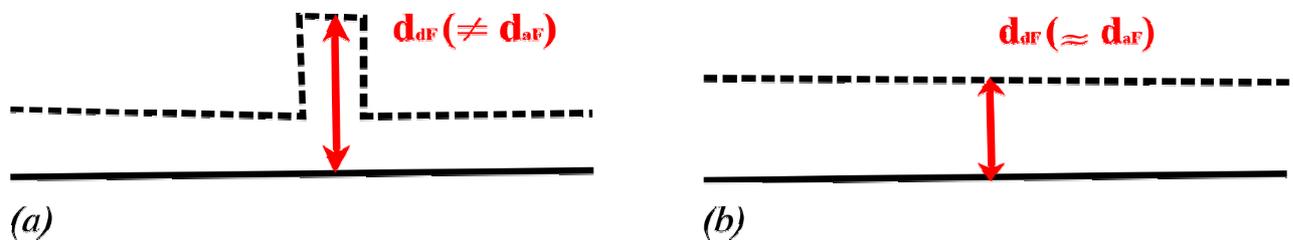


Figure 32 : Cas pour lesquels à distances moyennes égales, les distances de Fréchet maximales correspondantes sont très différentes.

Il est à noter qu'un ré échantillonnage des lignes réduit l'imprécision de la discrétisation sur la valeur de cette distance.

Tableau 7 : Evolution des valeurs de d_{dF} et d_{aF} selon le taux de ré échantillonnage des lignes de la figure 29

	Sans	1	0.1	0.01	0.001
d_{dF}	1.8028	1.2260	1.2015	1.2012	1.2012
d_{aF}	0.9524	0.5843	0.5116	0.5030	0.4997

Le tableau 7 montre la variation des valeurs selon l'échantillonnage effectué. L'architecture des jeux de données de départ ne gère pas le positionnement des points ni leur espacement : l'appariement est alors calculé entre les points constitutifs des lignes mais la longueur des segments comparés est très approximative. Sans ré échantillonnage, la mesure obtenue ne donne pas une valeur « fiable » de la distance maximale réelle entre les segments mais seulement une estimation de l'écartement entre les extrémités des segments.

Pour obtenir une valeur d_{aF} de précision centimétrique, il faut effectuer un échantillonnage avec un pas de l'ordre de quelques centimètres. Ceci s'explique par la corrélation de la distance de Fréchet discrète et de la longueur des segments (« LongMaxSeg », [Eiter & Mannila, 1994]) à partir desquels elle est calculée (voir Chapitre II).

II.2 Extension 2 : distance de Fréchet sur des lignes d'emprises ou de natures différentes

Les distances de Fréchet définies dans cette section ont pour objectif d'apparier des lignes d'emprises et de natures différentes. Elles s'appliquent à différents cas de figure illustrés figure 33 :

- comparaison de lignes ouvertes d'emprises différentes (cas (a)) ;
- comparaison entre deux lignes fermées (cas (b)) ;
- comparaison entre une ligne ouverte et une ligne fermée (cas (c)) ;
- comparaison entre parties de lignes ouvertes (cas (d)).

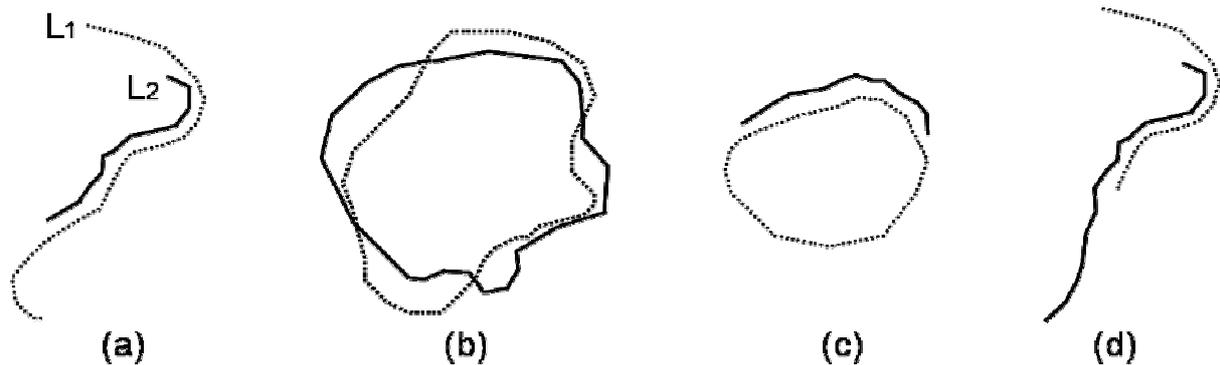


Figure 33 : Exemples de couples de géométries pertinentes pour le calcul des mesures dérivées de la distance de Fréchet.

II.2.1 Distance de Fréchet discrète partielle

Une mesure dérivée de la distance de Fréchet a été introduite par [Devoegele, 2002] afin d'apparier partiellement des lignes qui ne sont pas de même emprise (figure 33 (a)) : la distance de Fréchet discrète partielle (d_{pdF}).

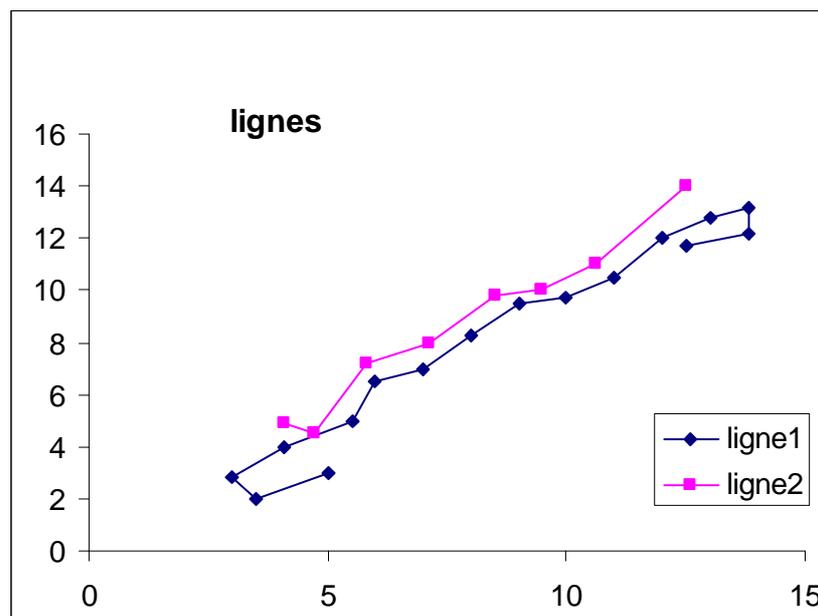


Figure 34 : Lignes partielles homologues représentées avec leur points.

La figure 34 illustre un couple de lignes qui peuvent être partiellement appariées : L_2 avec la partie homologue de L_1 comprise entre les points $\langle L_{1.4} \dots L_{1.13} \rangle$. Une distance de Fréchet discrète classique ne peut pas être utilisée pour ce cas de figure en terme d'appariement car elle n'a pas de sens. Elle donnerait un résultat exprimé entre l'un des deux couples extrémités des lignes, par conséquent moins bon car non calculé entre points homologues.

Pour calculer la d_{pdF} et déterminer la ligne partiellement homologue, il est nécessaire de détecter automatiquement dans la ligne de plus grande emprise les points de départ ($L_{1.début}$) et de fin ($L_{1.fin}$) de la ligne partiellement homologue. Les points $L_{1.début} \dots L_{1.fin}$ sont testés par ordre de distance aux points de L_2 . Cet ordre est utilisé pour optimiser la recherche de ($\langle L_{1.début} \dots L_{1.fin} \rangle$) : il évite ainsi de tester systématiquement tous les couples possibles du tableau 8. Le mécanisme de sélection des points homologues parmi les candidats pour $L_{1.début}$ et $L_{1.fin}$ est décrit par [Devogele, 2002]. Ces points sont des points de L_1 tels que la distance de Fréchet discrète entre L_2 et la ligne ($\langle L_{1.début} \dots L_{1.fin} \rangle$) soit minimale : $d_{pdF} = d_{dF}(L_2, \langle L_{1.début} \dots L_{1.fin} \rangle)$ où début = 4 et fin = 13 dans l'exemple de la figure 34.

Tableau 8 : Matrice de Distance des lignes homologues partielles représentées figure 34.

		Matrice de Distance															
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1		2.1	2.96	2.37	0.9	1.4	2.48	3.58	5.14	6.72	7.64	8.89	10.6	11.9	12.8	12.1	10.8
2		1.53	2.77	2.4	0.78	0.94	2.39	3.4	5	6.59	7.46	8.7	10.5	11.7	12.6	11.9	10.6
3		4.28	5.69	5.22	3.62	2.22	0.73	1.22	2.44	3.94	4.91	6.16	7.84	9.12	10	9.43	8.07
4		5.42	7	6.62	5	3.4	1.86	1	0.93	2.42	3.39	4.63	6.33	7.61	8.48	7.91	6.55
5		7.65	9.26	8.9	7.28	5.66	4.14	3.18	1.63	0.58	1.5	2.6	4.13	5.41	6.3	5.82	4.43
6		8.32	10	9.7	8.07	6.4	4.95	3.91	2.3	0.71	0.56	1.58	3.2	4.48	5.36	4.83	3.45
7		9.77	11.5	11.2	9.55	7.87	6.44	5.38	3.78	2.19	1.39	0.64	1.72	3	3.88	3.42	2.02
8		13.3	15	14.7	13.1	11.4	9.92	8.9	7.3	5.7	4.93	3.81	2.06	1.3	1.53	2.22	2.3

Le chemin minimal est déterminé de la même manière que pour la distance de Fréchet discrète.

II.2.2 Distance de Fréchet discrète entre deux lignes fermées

La distance de Fréchet discrète entre deux lignes fermées représente la distance de Fréchet discrète entre deux limites de polygones. La figure 33 (b) illustre un exemple pour lequel la distance de Fréchet entre deux lignes fermées doit être utilisée.

La procédure décrite dans [Devogele, 2002] utilise une fonction qui convertit les limites de polygones P_1 et P_2 en suites ordonnées de points, les lignes L_1 et L_2 . La difficulté vient de l'absence de point de début et de fin pour un polygone, contrairement à une ligne ouverte. Chaque ligne issue de la transformation du polygone va donc comporter un point doublé. Ce point marque le caractère « fermé » et symbolise à la fois le point de début et de fin de la ligne ainsi obtenue.

L'objectif est alors d'obtenir des lignes de même sens avec des points de début relativement proches.

L'algorithme présenté figure 35 explique le mécanisme de recherche d'un point de début d'une ligne L_2 en vue de son appariement avec une ligne L_1 . Dans un premier temps, un point de L_1 est choisi de manière à servir de point de début. Ce point dénommé $L_{1.1}$ est arbitrairement sélectionné pour son éloignement maximal vis-à-vis de tous les points de L_2 . Par la suite l'algorithme recherche parmi les points de L_2 le point $L_{2.1}$ le plus proche de $L_{1.1}$.

Ces deux points $L_{1,1}$ et $L_{2,1}$ deviennent alors temporairement les points de début respectifs de L_1 et L_2 localisés à proximité l'un de l'autre : il est alors possible de tester le sens des deux lignes.

Une ligne L_2 est dite **dans le même sens** qu'une ligne L_1 si l'indice du point de L_2 le plus proche du point $L_{1,1}$ est **avant** l'indice du point de L_2 le plus proche de $L_{1,n}$. Si les lignes ne sont pas dans le même sens, la ligne L_2 est inversée. Cette étape permet le calcul de la d_{Fd} entre les deux lignes de façon à ce que la valeur soit la plus petite possible et le chemin minimal continu [Devogele, 2002].

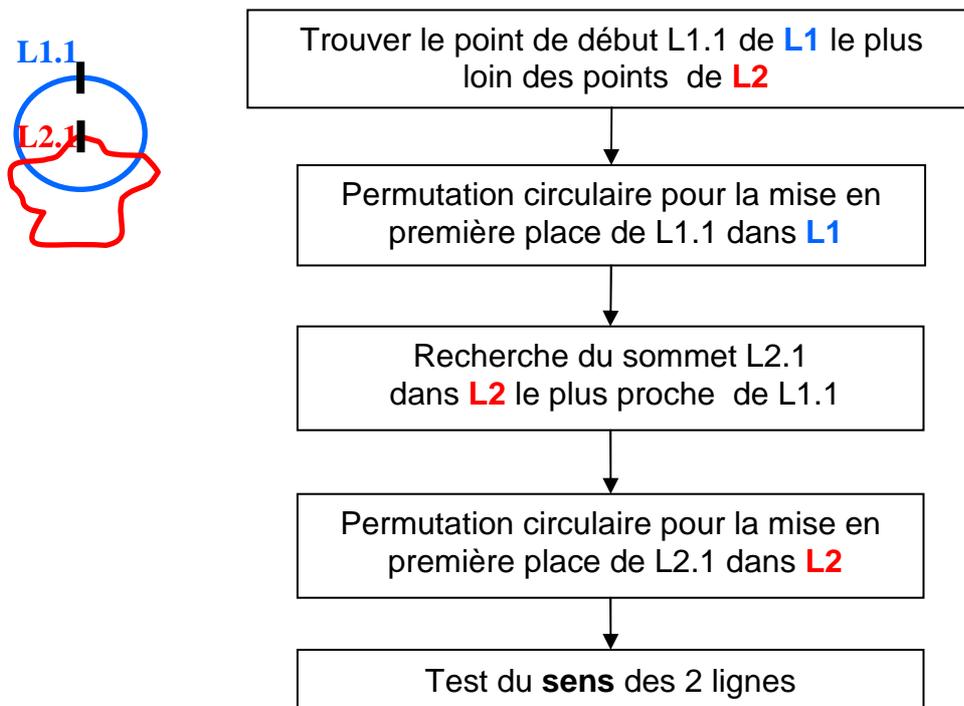


Figure 35 : Algorithme de recherche des points de début de L_1 et L_2 en vue d'un appariement des deux lignes fermées.

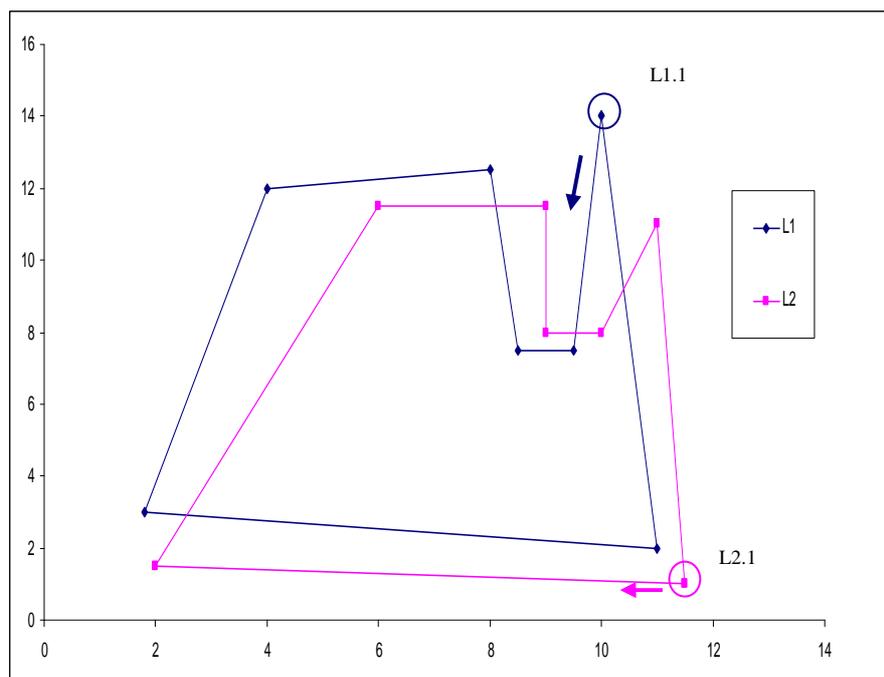


Figure 36 : Lignes fermées représentées avec des sens opposés.

Tableau 9 : Matrices de Distance et de Fréchet des lignes fermées de la figure 36.

Matrice de Distance

	1	2	3	4	5	6	7	8
1	13.086	14.841	4.717	2.6926	6.0828	6	3.1623	13.086
2	6.8007	9.6047	5.3151	4.0311	0.7071	0.7071	3.8079	6.8007
3	7.1589	8.8459	4.717	4.0311	0.7071	1.5811	4.3012	7.1589
4	12.021	12.53	2.2361	1.4142	4.6098	4.9244	3.3541	12.021
5	13.314	10.689	2.0616	5.0249	6.4031	7.2111	7.0711	13.314
6	9.904	1.5133	9.481	11.14	8.7658	9.6042	12.192	9.904
7	1.118	9.0139	10.735	9.7082	6.3246	6.0828	9	1.118
8	13.086	14.841	4.717	2.6926	6.0828	6	3.1623	13.086

Calcul de Fréchet

	1	2	3	4	5	6	7	8
1	13.086	14.841	14.841	14.841	14.841	14.841	14.841	14.841
2	13.086	13.086	13.086	13.086	13.086	13.086	13.086	13.086
3	13.086	13.086	13.086	13.086	13.086	13.086	13.086	13.086
4	13.086	13.086	13.086	13.086	13.086	13.086	13.086	13.086
5	13.314	13.086	13.086	13.086	13.086	13.086	13.086	13.314
6	13.314	13.086	13.086	13.086	13.086	13.086	13.086	13.086
7	13.314	13.086	13.086	13.086	13.086	13.086	13.086	13.086
8	13.314	14.841	13.086	13.086	13.086	13.086	13.086	13.086

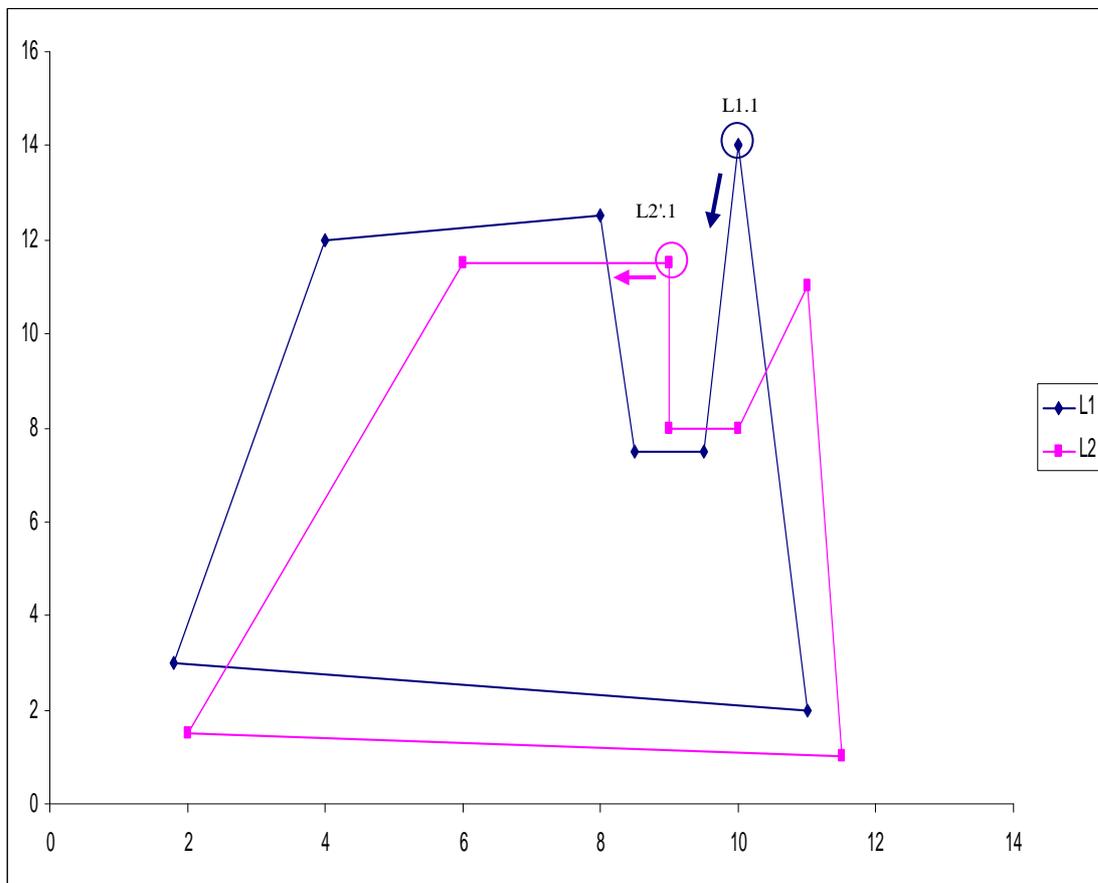


Figure 37 : Etape d'inversion du sens de la ligne fermée L_2 en L_2' .

Tableau 10 : Matrice de distance résultant de l'inversion de la ligne L_2 , présentée figure 37. Le cheminement optimal est représenté en gris. Pour cet exemple $d_{Fd} = 6,08$.

Matrice de Distance								
	1	2	3	4	5	6	7	8
1	2.6926	4.717	14.841	13.086	3.1623	6	6.0828	2.6926
2	4.0311	5.3151	9.6047	6.8007	3.8079	0.7071	0.7071	4.0311
3	4.0311	4.717	8.8459	7.1589	4.3012	1.5811	0.7071	4.0311
4	1.4142	2.2361	12.53	12.021	3.3541	4.9244	4.6098	1.4142
5	5.0249	2.0616	10.689	13.314	7.0711	7.2111	6.4031	5.0249
6	11.14	9.481	1.5133	9.904	12.192	9.6042	8.7658	11.14
7	9.7082	10.735	9.0139	1.118	9	6.0828	6.3246	9.7082
8	2.6926	4.717	14.841	13.086	3.1623	6	6.0828	2.6926

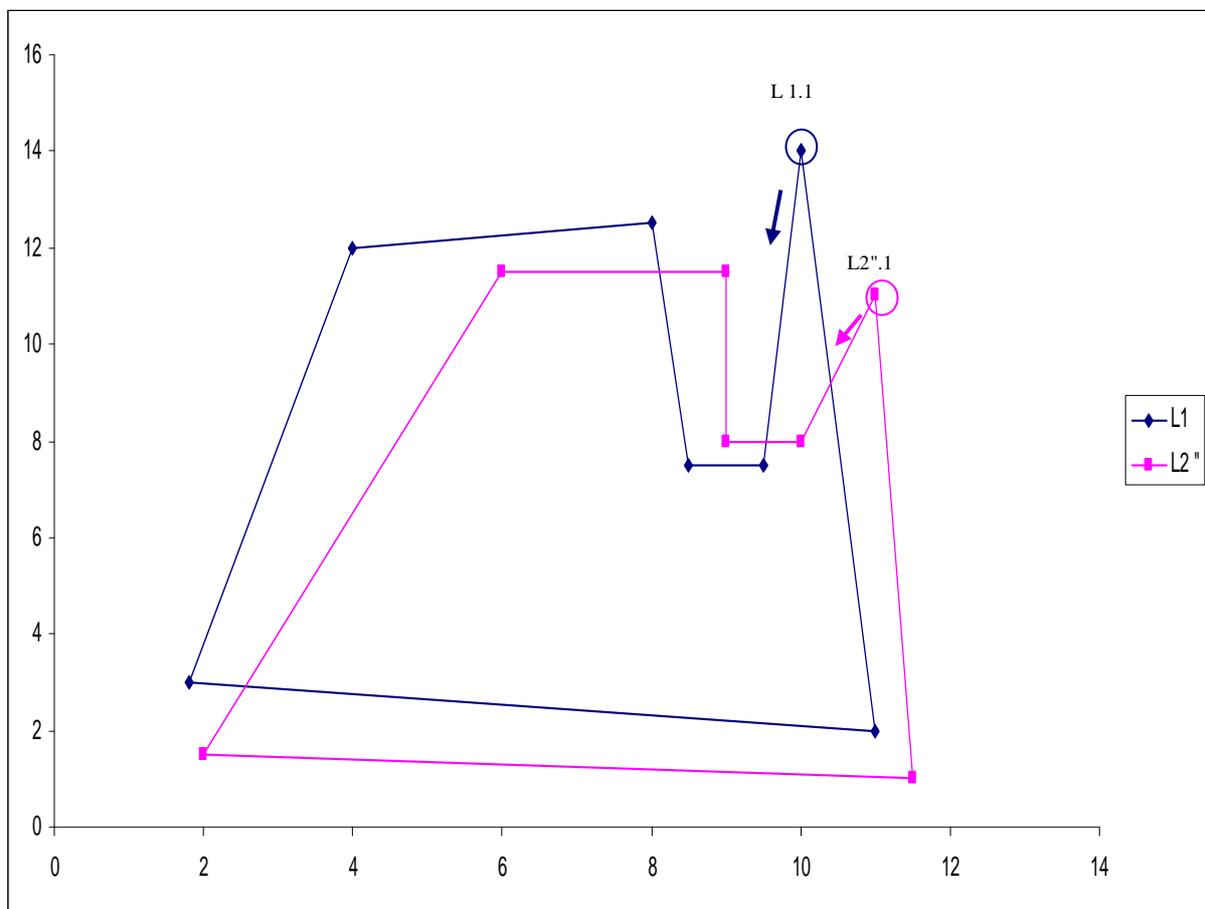


Figure 38 : Recherche du point homologue de $L_{1,1}$ parmi les points de L_2 , et permutation du point trouvé de manière à ce qu'il devienne $L_{2',1}$.

Les figure 36 et suivantes illustrent le traitement de deux lignes fermées de sens opposés. Le tableau 9 présente les matrices de distance et de Fréchet calculées à partir des points des lignes : aucune recherche ni modification de sens n'a été réalisée. La distance de Fréchet est de : 13,09 ; ce qui est une valeur trop forte pour permettre un appariement. Un nouveau point de début pour L_2 est alors recherché : $L_{2,4}$ devient $L_{2',1}$ après une permutation circulaire. Le test du sens des lignes confirme que celles-ci sont de sens opposés : L_2 est alors inversée.

Tableau 11 : Matrice de Distance résultante de la dernière permutation, distance de Fréchet.

Matrice de Distance								
	1	2	3	4	5	6	7	8
1	3.1623	6	6.0828	2.6926	4.717	14.841	13.086	3.1623
2	3.8079	0.7071	0.7071	4.0311	5.3151	9.6047	6.8007	3.8079
3	4.3012	1.5811	0.7071	4.0311	4.717	8.8459	7.1589	4.3012
4	3.3541	4.9244	4.6098	1.4142	2.2361	12.53	12.021	3.3541
5	7.0711	7.2111	6.4031	5.0249	2.0616	10.689	13.314	7.0711
6	12.192	9.6042	8.7658	11.14	9.481	1.5133	9.904	12.192
7	9	6.0828	6.3246	9.7082	10.735	9.0139	1.118	9
8	3.1623	6	6.0828	2.6926	4.717	14.841	13.086	3.1623

Ordre des points de L2 avant modification ddF = 13,086

L2.j.x	11.5	2	6	9	9	10	11	11.5
L2.j.y	1	1.5	11.5	11.5	8	8	11	1

Point de L2 inversée au plus proche (L2.4 est le plus proche de L1.1) ddF = 6,083

L2.j.x	9	6	2	11.5	11	10	9	9
L2.j.y	11.5	11.5	1.5	1	11	8	8	11.5

Point de L2 permutation circulaire (L2.7) ddF = 3,162

L2.j.x	11	10	9	9	6	2	11.5	11
L2.j.y	11	8	8	11.5	11.5	1.5	1	11

Pas d'autres points à tester car les autres distances (L1.1, L2.j) sont supérieures à 3,162

La figure 37 et ses matrices (tableau 10) montrent l'évolution de la distance de Fréchet après ces changements : la distance obtenue est de 6,08.

Afin d'apparier les points homologues entre eux et non uniquement les plus proches géographiquement, une permutation circulaire est à nouveau réalisée. La figure 38 et le tableau 11 présentent la valeur de la distance obtenue : $d_{Fd} = 3,16$. Des permutations suivantes ne sont pas entreprises car les autres distances dans la matrice de Fréchet sont supérieures à 3,162.

La comparaison de lignes fermées se fait impérativement sur leur totalité car il faudrait sinon rechercher quatre inconnues : deux points de début et deux points de fin. Comme ceci n'est pas réalisable, il n'existe pas de distance de Fréchet discrète partielle pour deux lignes fermées.

II.2.3 Distance de Fréchet discrète partielle entre une ligne ouverte et une ligne fermée

La définition d'une distance de Fréchet discrète partielle (d_{pdF}) entre une ligne L_1 et une partie de frontière de polygone P_2 représente un apport de cette thèse vis-à-vis de ce qui a déjà été présenté. Elle se calcule de façon intermédiaire vis-à-vis des deux mesures décrites précédemment.

Le principe est de découper la ligne fermée pour obtenir la partie homologue de la ligne ouverte et ensuite de calculer une distance de Fréchet discrète. La figure 33 (c) montre un exemple entre une ligne ouverte et une ligne fermée qui impose le calcul d'une d_{pdF} .

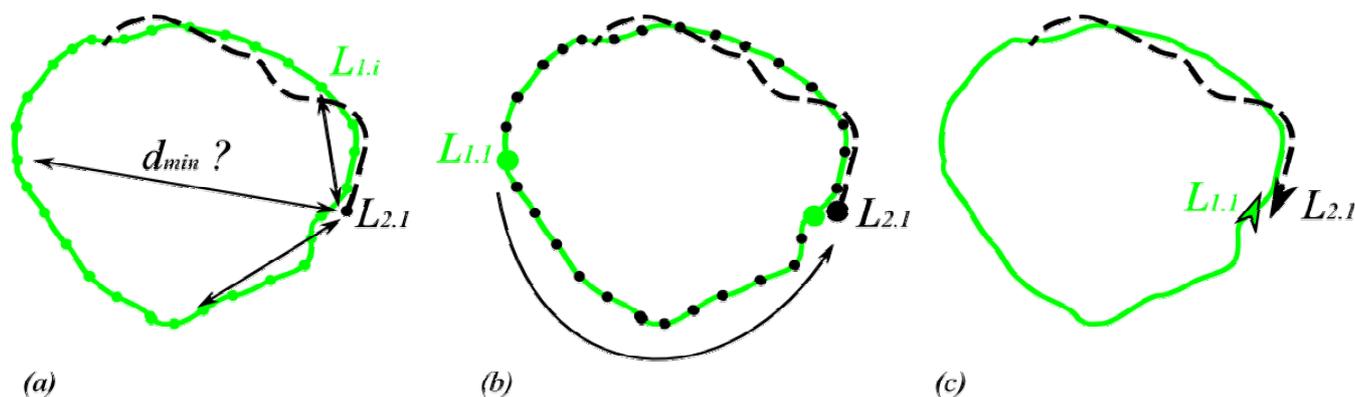


Figure 39 : détail de l'appariement d'une ligne fermée avec une ligne ouverte.

La figure 39 reprend le détail du traitement. Premièrement les distances entre le point de départ de la ligne ouverte $L_{2,1}$ et tous les points de L_1 sont calculées. La distance minimale est conservée (figure 39 (a)). A partir de cette distance, une permutation circulaire de L_1 est effectuée de manière à avoir $L_{1,\text{début}}$ situé à l'emplacement du point où la distance est minimale (figure 39 (b)). Cette étape de rapprochement géographique, à l'instar de l'exemple précédent entre deux lignes fermées, facilite la détermination du sens de L_1 par rapport à L_2 . S'il s'avère que les lignes ont des sens opposés (figure 39 (c)), une des deux lignes est inversée. La ligne fermée L_1 est ensuite comparée à la ligne ouverte L_2 puis le point homologue à $L_{2,n}$ est recherché. La distance de Fréchet discrète partielle est ensuite calculée de la même manière qu'elle le serait entre deux lignes de même emprise en construisant le chemin minimal et en recherchant les points homologue et non ceux les plus proches.

II.2.4 Distance de Fréchet discrète partielle / partielle entre deux lignes ouvertes

La distance de Fréchet discrète partielle entre deux tronçons de lignes ouvertes diffère de celles présentées précédemment car la relation entre les lignes n'est plus la même. Jusqu'à présent à une ligne entière correspondait la totalité d'une autre ligne (distance de Fréchet discrète) ou au moins un des ses tronçons (distance de Fréchet discrète partielle).

Ici, à une partie d'une ligne ne correspond uniquement qu'une autre partie de la ligne comparée, ce qui amène à définir également une nouvelle relation entre les objets (cf chapitre II) : la relation d'appariement *partiel* « p-p », très liée à la relation n-m définie en partie II.1.1.3. Si à n objets de la BD 1 correspondent m objets de la BD2, cela signifie que pour deux objets comparés, une partie d'un objet géographique du jeu de données 1 équivaut à une partie d'un objet du jeu de données 2 (voir figure 40). D'où la relation partielle/partielle « PP » entre objets. Ce type de relation exclut automatiquement les relations de type 1-1.

Type de relation		
n-m		BD 1 BD 2
p-p		

Figure 40 : Mise en parallèle des relations n-m et p-p entre deux objets.

En pratique cela induit la comparaison de la fin d'une ligne avec le début d'une autre. En effet, contrairement aux mesures précédentes, les **deux** lignes sont comparées partiellement de manière à trouver dans certaines configurations deux points de début et de deux points de fin (voir figure 41 cas (a)).

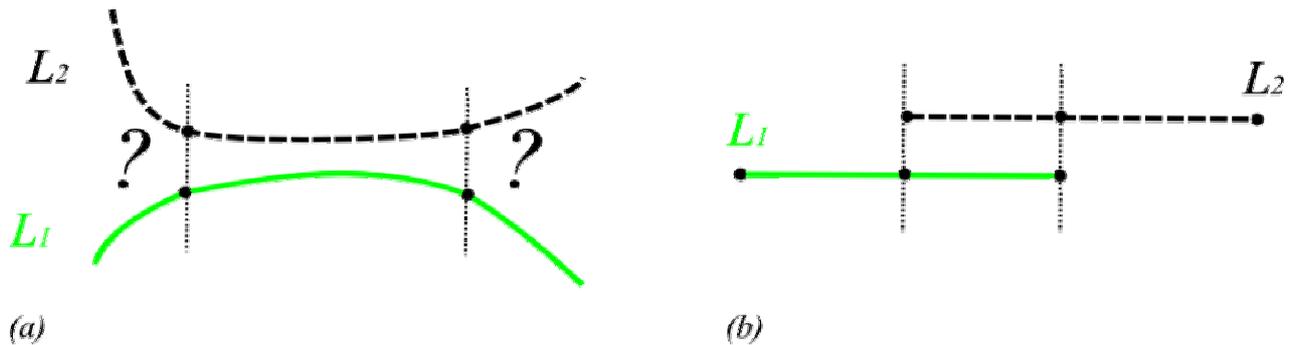


Figure 41 : Cas d'appariement partiel-partiel de lignes ouvertes : a) exemple impossible à résoudre en l'état actuel de nos connaissances ; b) exemple qui possède une solution unique.

Ce type d'appariement vise à identifier un segment de ligne homologue à un autre segment de ligne. Il va alors se poser des **problèmes de contraintes** pour certains cas. L'exemple (a) de la figure 41 est insolvable car le nombre d'inconnues est trop grand : la disposition des lignes impose la recherche d'un point de début et de fin pour L_1 et L_2 , soit **quatre points** à déterminer au total. Aucun indice ne permet de déterminer la partie de ligne L_1 à appairer avec la partie de ligne L_2 , il faut donc retenir :

- soit les segments les plus longs possibles et ainsi obtenir une distance de Fréchet discrète très grande ;
- soit rechercher les segments les plus courts afin d'avoir une distance de Fréchet bien meilleure, au risque de réduire les segments aux deux points les plus proches.

Entre ces deux extrêmes, il existe une infinité de solutions possibles. Pour apporter une solution à ce problème, il est donc nécessaire d'introduire deux autres contraintes :

- un des points de début est le point d'origine de la ligne ;
- un des points de fin est le point de fin de l'autre ligne considérée.

Ces deux contraintes sont illustrées par l'exemple (b) figure 41 qui ne demande la recherche que d'un point de début sur une ligne L_1 et un point de fin sur la ligne L_2 .

D'un point de vu calcul matriciel, ces contraintes équivalent à chercher les points candidats dans une sous-matrice située dans l'angle haut droit de la matrice globale. La résolution mathématique comporte en effet de fortes similitudes avec le calcul d'une distance entre lignes partielle/totale :

- dans un cas partiel/totale, les points de début et de fin sont cherchés dans une sous matrice incluant la première et dernière ligne de la matrice ;
- pour un cas PP les points de débuts et de fin candidats sont recherchés uniquement parmi les points de la **première ligne** et de la **dernière colonne** de la matrice globale (tableau 13).

Sous certaines conditions, ces règles de recherches de points peuvent être étendues à la recherche de quatre points si les lignes possèdent entre elles non plus **une** extrémité homologue (cas « PP ») mais **deux** (exemple b) figure 42) : la relation définie est appelée « **deux partielle-partielle** » (2PP) et réalise deux appariements PP.

A partir de ces considérations, les cas de figure qui peuvent être traités sont illustrés figure 42 :

- (a) : relation PP ;
- (b) : relation 2PP ;
- (c) : relation ambiguë d'un cas PP.

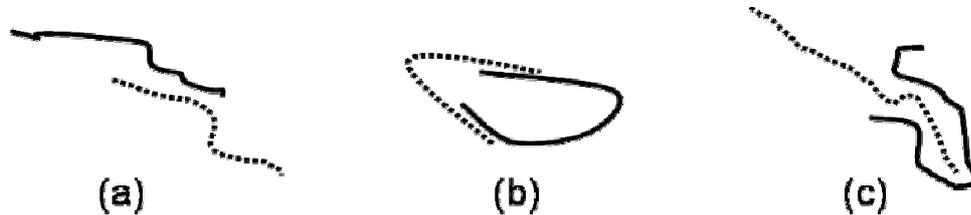


Figure 42 : Illustration des différents positionnements des lignes lors de leur appariement.

Les exemples suivants montrent des cas de lignes pour lesquelles il y a impossibilité à déterminer des points de début et de fin : 1) pour des lignes ouvertes/fermées (figure 43), et 2) pour des lignes fermées/fermées (figure 44).

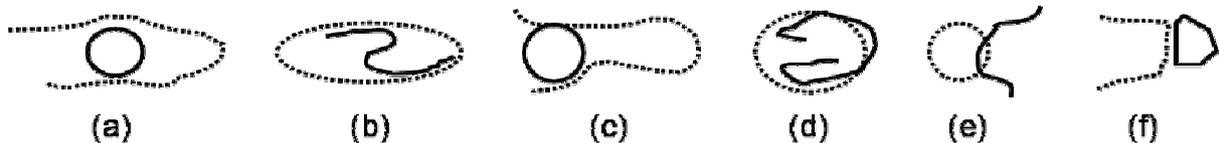


Figure 43 : Inventaire des cas d'appariement impossibles à traiter entre une ligne ouverte et une ligne fermée. En (a), (b) et (c) dans un contexte d'appariement deux partiel/partiel ; et en (d), (e), (f) pour des cas partiel/partiel.

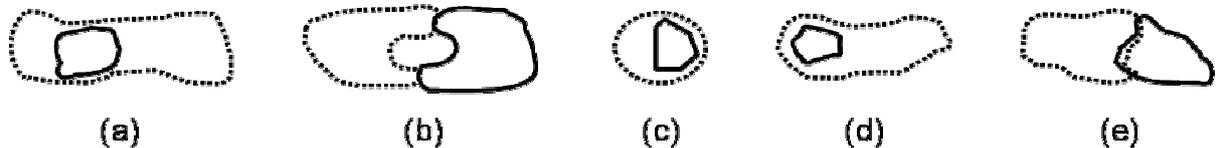


Figure 44 : Inventaire des cas impossibles à traiter entre deux lignes fermées. En (a), (b) et (c) pour un appariement deux partiel/partiel ; et en (d), (e) pour un appariement partiel/partiel.

L'exemple (b) « 2PP » de la figure 42 est présenté ci-après de manière à étayer plus précisément la méthode de traitement.

Visuellement les lignes ne sont homologues que partiellement, mais simultanément sur leur deux extrémités.

L'outil implémenté ne permet pas lors du processus de comparaison de déterminer quelle extrémité de la ligne est homologues à l'autre ni si elle est homologues sur une partie ou sur deux. Quatre distances partielles sont donc calculées :

- L_1 et L_2 ;
- L_{1inv} et L_2 ;
- L_1 et L_{2inv} ;
- L_{1inv} et L_{2inv} .

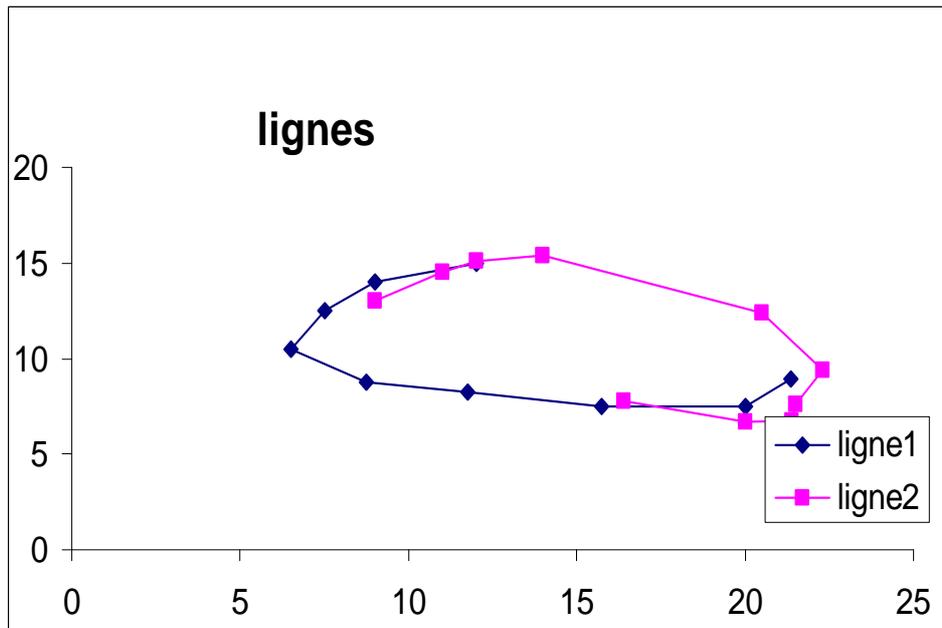


Figure 45 : Exemple de lignes pour lesquelles les deux extrémités sont homologues (cas 2PP).

Tableau 12 : Matrice de Distance et Matrice de Fréchet des lignes L_1 et L_2 présentées figure 45.

Matrice de Distance (MD)									
	1	2	3	4	5	6	7	8	9
1	3.606	1	1.581	3.536	4.257	5.489	8.701	12.3	13.02
2	1.118	2.062	4.031	6.021	6.175	6.295	8.451	11.4	11.77
3	0.1	3.195	5.197	7.17	7.133	6.855	8.466	11.03	11.21
4	2.04	5.192	7.118	8.959	8.473	7.496	8.082	9.92	9.8
5	8.889	11.61	13	14.13	12.3	9.684	6.817	4.925	3.568
6	11.72	14.07	15.12	15.84	13.57	10.61	6.817	2.983	1.038
7	12.04	14.04	14.83	15.28	12.8	9.772	5.751	1.503	1.346
8	12.52	14.37	15.05	15.37	12.81	9.769	5.704	1.598	2.21
9	11.52	13.2	13.78	14.02	11.43	8.393	4.325	0.79	2.617
10	8.438	9.654	10.06	10.26	7.709	4.672	0.712	3.612	5.099

Matrice de Fréchet (MF)									
	1	2	3	4	5	6	7	8	9
1		1	1.581	3.536	4.257	5.489	8.701	12.3	13.02
2		1.118	4.031	6.021	6.175	6.295	8.451	11.4	11.77
3		0.1	5.197	7.17	7.133	6.855	8.466	11.03	11.21
4		2.04	7.118	8.959	8.473	7.496	8.082	9.92	9.92
5		8.889	13	14.13	12.3	9.684	7.496	7.496	7.496
6		11.72	15.12	15.84	13.57	10.61	7.496	7.496	7.496
7		12.04	14.83	15.28	13.57	10.61	7.496	7.496	7.496

Les matrices de distance et de Fréchet des lignes de la figure 45 sont présentées dans le tableau 12.

Le point de départ retenu est $L_{1,2}$ et le point d'arrivée $L_{2,7}$, le résultat de la distance de Fréchet discrète partielle obtenue est de : **7,496**.

La ligne L_1 est ensuite inversée (L_{1inv}) et le même calcul est réitéré avec L_2 (voir tableau 13).

Tableau 13 : Matrice de Distance et de Fréchet des lignes de la figure 45 avec les points de la ligne L_1 inversée.

Matrice de Distance									
	1	2	3	4	5	6	7	8	9
1	13.02	12.3	8.701	5.489	4.257	3.536	1.581	1	3.606
2	11.77	11.4	8.451	6.295	6.175	6.021	4.031	2.062	1.118
3	11.21	11.03	8.466	6.855	7.133	7.17	5.197	3.195	0.1
4	9.8	9.92	8.082	7.496	8.473	8.959	7.118	5.192	2.04
5	3.568	4.925	6.817	9.684	12.3	14.13	13	11.61	8.889
6	1.038	2.983	6.817	10.61	13.57	15.84	15.12	14.07	11.72
7	1.346	1.503	5.751	9.772	12.8	15.28	14.83	14.04	12.04
8	2.21	1.598	5.704	9.769	12.81	15.37	15.05	14.37	12.52
9	2.617	0.79	4.325	8.393	11.43	14.02	13.78	13.2	11.52
10	5.099	3.612	0.712	4.672	7.709	10.26	10.06	9.654	8.438

Matrice de Fréchet									
	1	2	3	4	5	6	7	8	9
1								1	3.606
2								2.062	1.118
3								3.195	1.118

Le point de départ retenu est $L_{1,8}$ et le point d'arrivée $L_{2,3}$, le résultat de la distance de Fréchet discrète partielle obtenue est de : **1,118**.

Les deux autres calculs sont effectués exactement sur le même modèle et les distances de Fréchet obtenues sont :

- pour L_1 et $L_{2inv} = \mathbf{1,598}$;
- pour L_{1inv} et $L_{2inv} = \mathbf{7,495}$.

Dans cet exemple, six appariements sont donc possibles : deux appariements 2PP et quatre appariements PP. Dans le cas d'un appariement 2PP les résultats des couples retenus sont les suivants :

- couple L_1 et L_2 et L_{1inv} et $L_2 = \{\mathbf{7,496} ; \mathbf{7,495}\}$;
- couple L_{1inv} et L_2 et L_1 et $L_{2inv} = \{\mathbf{1,118} ; \mathbf{1,598}\}$.

Pour un appariement PP seules les distances entre les deux lignes sont retenues :

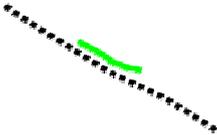
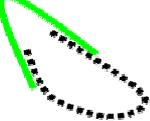
- L_1 et $L_2 = 7,496$;
- L_1 et $L_{2inv} = 1,598$;
- L_{1inv} et $L_2 = 1,118$;
- L_{1inv} et $L_{2inv} = 7,495$;

Se pose alors la question du choix entre un appariement 2PP ou PP. Il a été décidé de préférer un appariement 2PP pour les cas où il existe **un couple** possédant deux valeurs de distances de Fréchet inférieures à ϵ . Dans le cas où une seule distance est inférieure, un appariement PP est choisi.

Pour cet exemple les couples L_{1inv}/L_2 , et L_1/L_{2inv} , présentent des distances de Fréchet très proches en valeurs et assez faibles. L'appariement est donc effectué entre les points homologues des lignes retenues dans les chemins minimaux et de type 2PP. Cependant il ne faut pas prendre en compte les couples de points qui possèdent deux fois la même extrémité (exemple figure 42 (c)). Pour ces types de cas, un appariement PP doit être retenu en fonction des d_{pF} et du contexte (défini par l'utilisateur).

Le tableau 14 synthétise les différents types d'appariement résolus grâce à soit l'utilisation de distances de Fréchet déjà existantes, soit la définition de nouvelles distances.

Tableau 14 : Récapitulatif des distances d'appariements déjà traitées et celles introduites dans cette section.

	Total/total 1 - 1	Partiel 1 - P	Partiel - Partiel	
			PP	2PP
Distances déjà existantes				
Apports de la thèse				

Les différentes catégories d'appariements, basés sur les multiples distances déjà existantes ou créées selon les nouveaux cas de comparaisons effectuées, vont maintenant être présentées.

II.3 Algorithme d'appariement

L'objectif est de proposer le meilleur appariement lors de la comparaison de deux lignes quelconques. Pour ce faire, l'algorithme procède à une séquence de tests réalisés sur les lignes de manière à identifier quelle distance de Fréchet doit être calculée en vue de leur appariement final.

Cette section présente les trois étapes essentielles intervenant lors du processus d'appariement de lignes (voir figure 46). La première est la détermination du type de lignes comparées. Cette étape est effectuée soit manuellement par l'utilisateur (comparaison et appariement simple) soit automatiquement (appariement global par fichier). Dans la seconde étape il s'agit d'identifier le type de relations entre les emprises des lignes, symbolisées par les rectangles englobants. Enfin, la troisième et dernière étape consiste à calculer la distance de Fréchet choisie en fonction du type de lignes identifié en première étape et de la position relative des lignes déterminée lors de l'étape deux. En fonction du résultat de cette distance, un appariement approprié est effectué ou non.

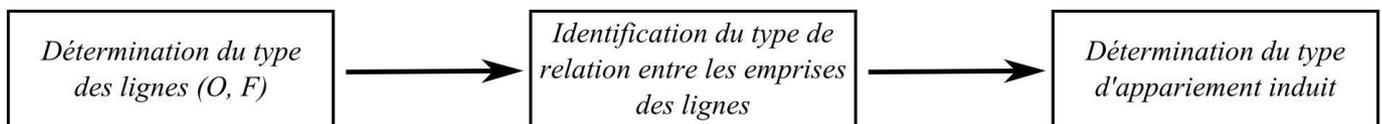


Figure 46 : Description des trois principales étapes définies pour l'appariement.

Cette chaîne de traitement vise à résoudre automatiquement le plus grand nombre de cas possibles lors d'un processus de comparaison. Ceux impossibles à résoudre par cette méthode sont également décrits en fin de section.

II.3.1 Recherche de la nature des lignes comparées

Les algorithmes d'appariement testés ne sont pas les mêmes selon qu'il s'agisse de comparer deux lignes ouvertes de même emprise, deux lignes ouvertes d'emprise différentes, deux lignes fermées ou une ligne ouverte avec une ligne fermée. Par souci d'optimisation du nombre de calculs effectués, l'algorithme va effectuer un premier tri : l'identification de la nature des lignes à comparer. En effet certaines catégories de comparaison de lignes vont obligatoirement exclure certains cas d'appariement qu'il n'est donc pas nécessaire de tester. La première étape consiste donc à vérifier pour une ligne donnée si les points de début et de fin sont les mêmes. Une fois que le type de ligne est connu, plusieurs types de comparaison s'offrent alors :

- ouvert/ouvert total ;
- ouvert/ouvert partiel ;
- ouvert/fermé (toujours partiel) ;
- fermé/fermé (toujours total).

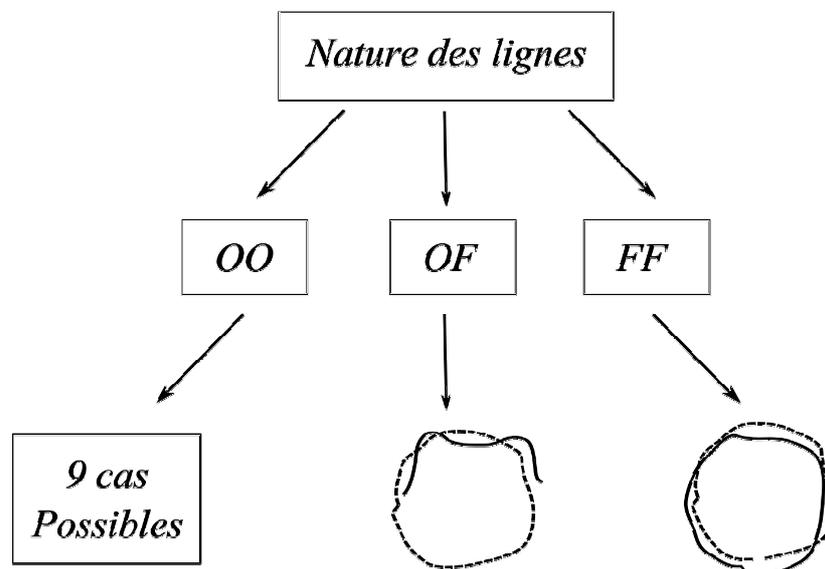


Figure 47 : Inventaire des appariements possibles selon la nature des lignes comparées.

La figure 47 récapitule les différents appariements possibles selon la nature des lignes comparées. Dans le cas de deux lignes ouvertes, tous les types d'appariement sont possibles. Pour déterminer quelle distance de Fréchet calculer, un autre tri doit être réalisé grâce aux relations des emprises des lignes comparées.

La comparaison d'une ligne ouverte avec une ligne fermée se rapporte à calculer une distance de Fréchet partielle et donc induire un appariement partiel : soit 1-P, soit PP soit 2PP.

La comparaison de deux lignes fermées est la plus simple car elle n'entraîne qu'un appariement total.

II.3.2 Relations d'emprises entre les lignes

Les différents appariements possibles sont déduits à partir des relations d'inclusion des emprises des lignes. Ces emprises sont représentées par le rectangle simple d'une ligne (BB : Bounding Box) et le rectangle englobant élargi d'un epsilon (BBE : Bounding Box Extended) de l'autre ligne comparée. Cet élargissement a pour but de créer une zone tampon

d'incertitude liée à la qualité des données. Cette valeur epsilon (ϵ) est choisie par l'utilisateur selon la précision de positionnement. Si les données sont très précises, ϵ peut être plus petit. Si au contraire les données sont peu précises, il va être nécessaire d'utiliser un ϵ plus grand de manière à s'affranchir des problèmes de calage qui peuvent aller jusqu'à plusieurs dizaines de mètres. Dans ce cas, l'appariement est effectué si la valeur de la distance de Fréchet obtenue est inférieure à une dizaine de mètres.

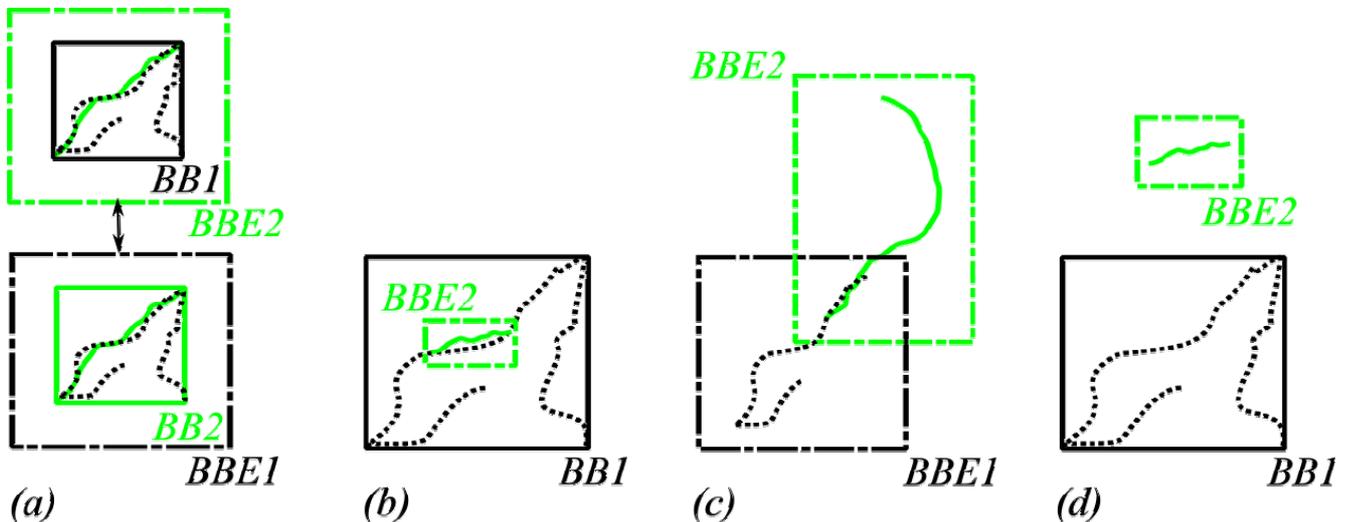


Figure 48 : Inventaire des relations d'inclusion existant entre objets géographiques : (a) emprises équivalentes, (b) emprise incluse dans une autre, (c) emprises possédant une zone de recouvrement, (d) emprises sans intersection.

Quatre types de relations ont été définis :

- l'inclusion réciproque de deux lignes entre elles dans le cas où elles sont d'emprises relativement **équivalentes** (cas (a), figure 48) : $BB1 \subset BBE2$ **et** $BB2 \subset BBE1$. L'appariement est alors soit total (1-1), soit partiel (1-P, PP ou 2PP) ;
- l'**inclusion** simple d'une ligne dans le rectangle englobant d'une autre (cas (b), figure 48) : $BB1 \subset BBE2$. Sa relation complémentaire se nomme **est inclus** : $BB2 \subset BBE1$. L'appariement est toujours partiel (1-P, PP ou 2PP) ;
- le **recouvrement** partiel (cas (c), figure 48) : $BBE1 \cap BBE2 \neq 0$. L'appariement effectué est partiel de type PP ou 2PP ;
- **aucune inclusion** (cas (d), Figure 48) : $BBE1 \cap BBE2 = 0$. Dans ce cas il n'y a pas d'appariement possible.

Ces relations sont appelées relations d'inclusion entre les emprises (RIE). Pour chaque type d'inclusion et selon la nature des lignes comparées, les appariements engagés et les distances de Fréchet discrètes utilisées ne sont pas les mêmes. Par exemple pour la cas (a), le principe est que si les emprises des lignes sont équivalentes, il y a une très forte probabilité pour que les lignes soient au même endroit et à peu près de la même taille. La distance calculée est une distance de Fréchet totale et l'appariement testé est de type 1-1. Si la d_F obtenue est inférieure à ϵ , l'appariement est validé, sinon il est éliminé des appariements possibles.

De la même manière, si la relation entre les emprises est l'inclusion, cela exclut automatiquement un appariement de type total.

S'il n'y a pas d'intersection entre les emprises, aucune distance de Fréchet n'est calculée et aucun appariement n'est effectué.

La section suivante rappelle les types d'appariement existants et quels types de lignes et de relations d'emprise sont concernés.

II.3.3 Types d'appariement entre lignes

II.3.3.1 Appariement simple « 1 – 1 » de deux lignes

Ce type d'appariement concerne les lignes d'emprises équivalentes qui ont une grande probabilité d'avoir une relation 1-1 dans leur comparaison : à une ligne L_1 correspond une et une seule ligne L_2 .

Deux lignes ouvertes

L'équivalence des emprises des lignes est vérifiée par test croisé : $BB1 \subset BBE2$ et inversement. Ensuite la distance de Fréchet est calculée entre L_1/L_2 . Si la distance minimale trouvée est supérieure à ϵ , alors aucun appariement n'est effectué.

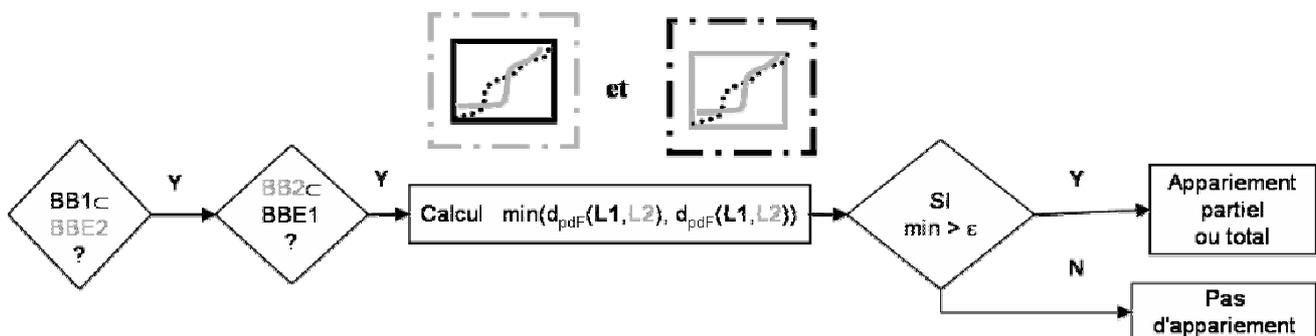


Figure 49 : Algorithme d'appariement total « 1-1 » entre deux lignes ouvertes. Les illustrations indiquent quel cas est alors identifié lors du calcul de la d_{pdf} .

Deux lignes de natures différentes : ouverte/fermée

La comparaison d'une ligne ouverte avec une ligne fermée se réalise en découpant la ligne fermée selon l'emprise de la ligne ouverte. Le calcul de la distance de Fréchet partielle est ensuite lancé. Si la valeur est satisfaisante ($< \epsilon$), un appariement partiel entre les deux tronçons est alors effectué.

Deux lignes fermées

Le cas des lignes fermées est probablement le plus simple car seules les lignes de même emprise sont appariables (figure 50).

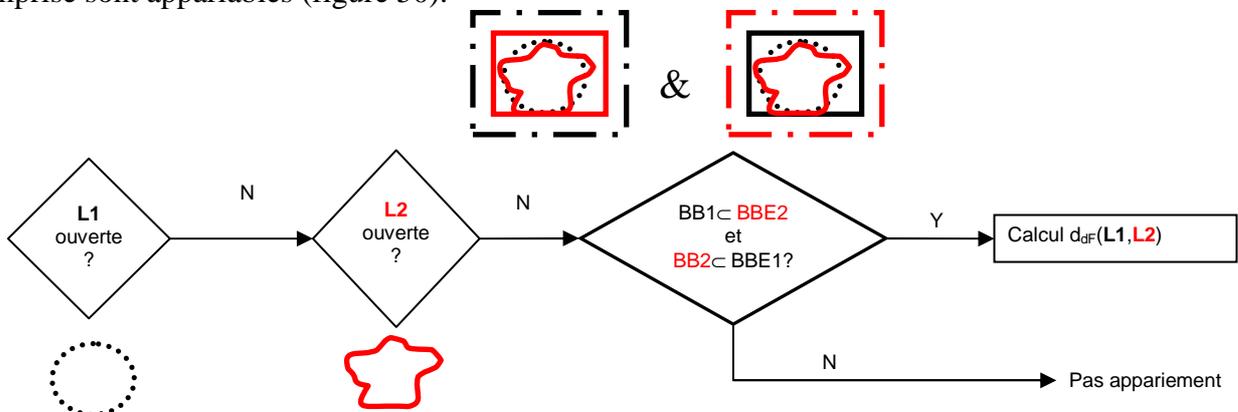


Figure 50 : Algorithme d'appariement de deux lignes fermées.

II.3.3.2 Appariement total/partiel « 1 – P » entre deux lignes

Cette catégorie d'appariement concerne les lignes qui ont une relation 1-n : à une ligne d'une BD1 correspond une, plusieurs lignes ou une partie d'une ligne de la BD2. Un découpage va donc être nécessaire, ce qui implique que cet appariement ne se réalise que sur des lignes ouvertes.

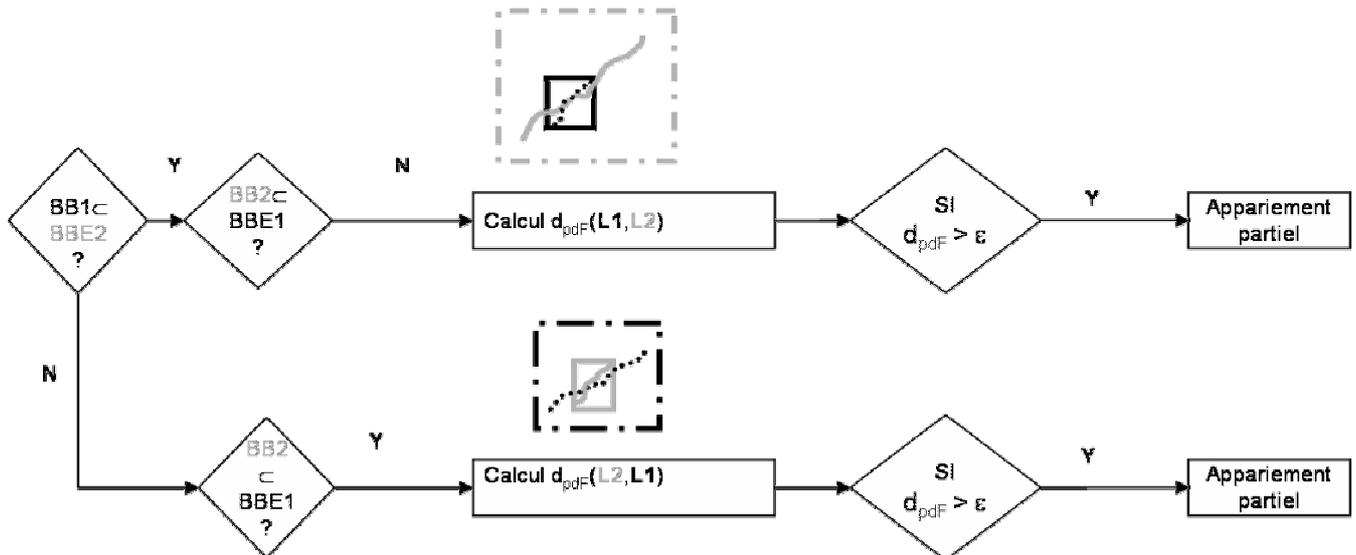


Figure 51 : Algorithme d'appariement total/partiel entre deux lignes ouvertes.

La méthodologie présentée figure 51 détaille les différentes actions effectuées pour le calcul de la distance de Fréchet partielle entre des lignes incluses. La première étape consiste à vérifier la relation d'inclusion entre les rectangles englobants simple de L_1 et élargi de L_2 . Si la relation est vraie, alors la droite la plus grande (L_1 dans notre exemple) est découpée selon le rectangle englobant élargi de la plus petite (L_2) à l'aide de la fonction « sélection rectangle » (voir figure 54, étape 2).

Celle-ci renvoie le segment de L_1 candidat à l'appariement avec L_2 . L'orientation des lignes est par la suite vérifiée. A partir de ces lignes, les deux matrices, de distance euclidienne et de Fréchet, sont alors construites comme dans l'exemple présenté figure 29 et **Erreur ! Source du renvoi introuvable.** Les couples de points homologues de début et de fin entre L_1 et L_2 sont recherchés, et les distances de Fréchet discrète partielle et moyenne sont calculées.

II.3.3.3 Appariement « partiel/partiel » (PP) et « deux partiel/partiel » (2PP) entre deux lignes

L'appariement de type partiel/partiel se calcule à l'aide de la distance de Fréchet partielle et peut s'appliquer à tout type de lignes quelles que soient leur emprise respective (voir figure 55 pour le récapitulatif de tous les cas possibles). Cependant, dans notre problématique d'intégration, il n'a été implémenté que pour les lignes ouvertes possédant des emprises avec un recouvrement partiel.

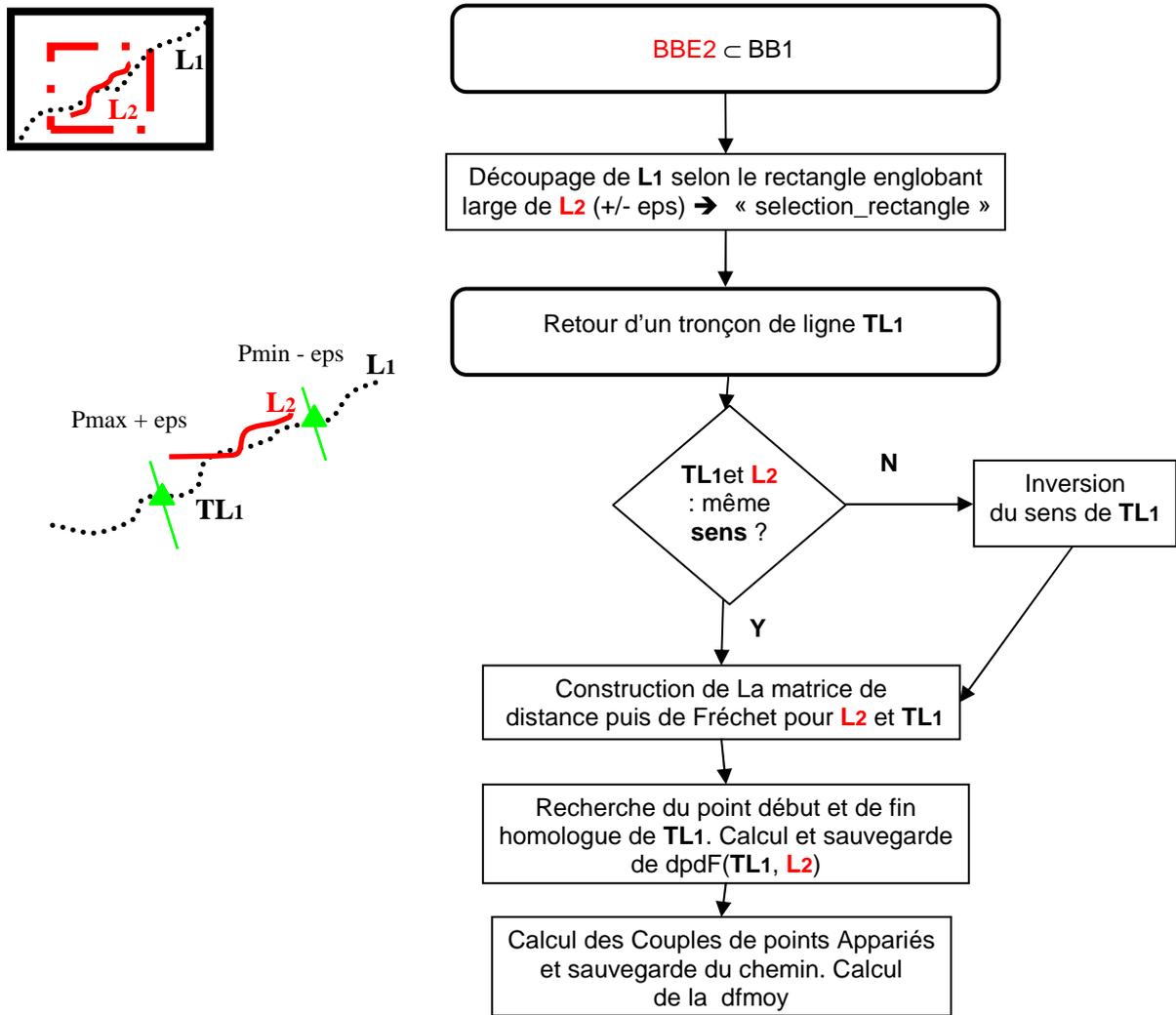


Figure 52 : Méthodologie de découpage d'une ligne L_1 selon une ligne L_2 d'emprise plus petite en vue du calcul de la distance de Fréchet partielle d_{pdF} .

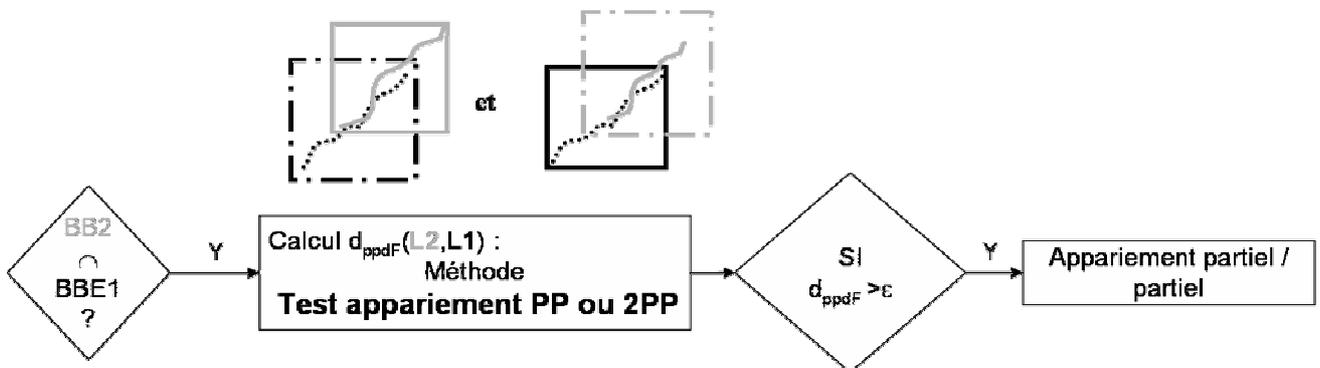


Figure 53 : Algorithme d'appariement partiel/partiel ou deux partiel/partiel de lignes possédant un domaine de recouvrement entre leurs emprises.

Lors du test du type d'appariement entre PP ou 2PP (« **Test appariement PP ou 2PP** » figure 53), neuf catégories sont identifiées (voir détail figure 54) :

- quatre concernent uniquement l'inclusion d'une seule extrémité d'une ligne dans le rectangle élargi de la seconde (illustration figure 42 (a)) : appariement partiel/partiel ;
- une concerne les deux extrémités des lignes (illustration figure 42 (b)) : appariement « deux partiel/partiel » ;
- quatre concernent une extrémité d'une ligne L_1 et deux extrémités d'une ligne L_2 (illustration figure 42 (c)) : appariement partiel/partiel.

Il est à noter que l'algorithme effectue le test des neuf cas en parallèle et que la (ou les) meilleure(s) valeur(s) de d_{pdF} est/sont retenue(s) à la fin après tri.

Le premier test de l'algorithme présenté figure 54 vise tout d'abord à identifier pour deux lignes (L_1 et L_2), dont le recouvrement n'est ni total ni inclus, si au moins une des extrémités (le début par exemple) de L_1 est comprise dans le rectangle englobant élargi de L_2 (fin non incluse). Si oui, le test suivant regarde si l'inverse est vrai, à savoir qu'une des extrémités de L_2 (par exemple le début) est comprise dans le rectangle englobant élargi de L_1 et pas l'autre (fin de L_2). Si c'est le cas, la d_{pdF} est calculée et conservée pour comparaison avec les autres valeurs.

Les trois tests suivants vérifient uniquement l'orientation des lignes entre elles en calculant les d_{pdF} pour chaque sens de ligne.

Le quatrième test concerne le cas où les deux extrémités (début et fin) de chaque ligne (L_1 et L_2) sont proches. Elles sont alors comprises dans les rectangles englobants élargis respectifs. Comme il n'est pas possible ici de déterminer a priori le sens des lignes, les quatre d_{pdF} sont calculées et la meilleure est gardée (se référer pour cela à l'exemple exposé figure 45).

Le test cinq traite le cas illustré figure 42 (c) où il est difficile de déterminer si c'est le début ou la fin de la ligne L_1 qui est homologue à l'extrémité de L_2 .

Les trois autres tests réitèrent le calcul en changeant les extrémités concernées de façon à vérifier l'orientation des lignes entre elles.

Le tableau 15 résume l'algorithme de la figure 54 :

- si un seul des cas A, B, C ou D est vrai, alors seul un appariement PP est effectué ;
- si A et D ou B et C sont vrais, alors un appariement 2PP est testé ;
- si un autre couple est vrai, la décision de l'appariement à réaliser est laissée au choix de l'utilisateur.

Tableau 15 : Récapitulatif des cas possibles d'appariement PP et 2PP.

Segment inclus de L1	Segment inclus de L2	distance	cas
$L1.n \subset BBE2$	$L2.1 \subset BBE1$	$d_{pdF}(L_1, L_2) < \epsilon$	A
$L1.1 \subset BBE2$	$L2.1 \subset BBE1$	$d_{pdF}(L_{1inv}, L_2) < \epsilon$	B
$L1.n \subset BBE2$	$L2.m \subset BBE1$	$d_{pdF}(L_1, L_{2inv}) < \epsilon$	C
$L1.1 \subset BBE2$	$L2.m \subset BBE1$	$d_{pdF}(L_{1inv}, L_{2inv}) < \epsilon$	D

Méthode test appariement PP ou 2PP

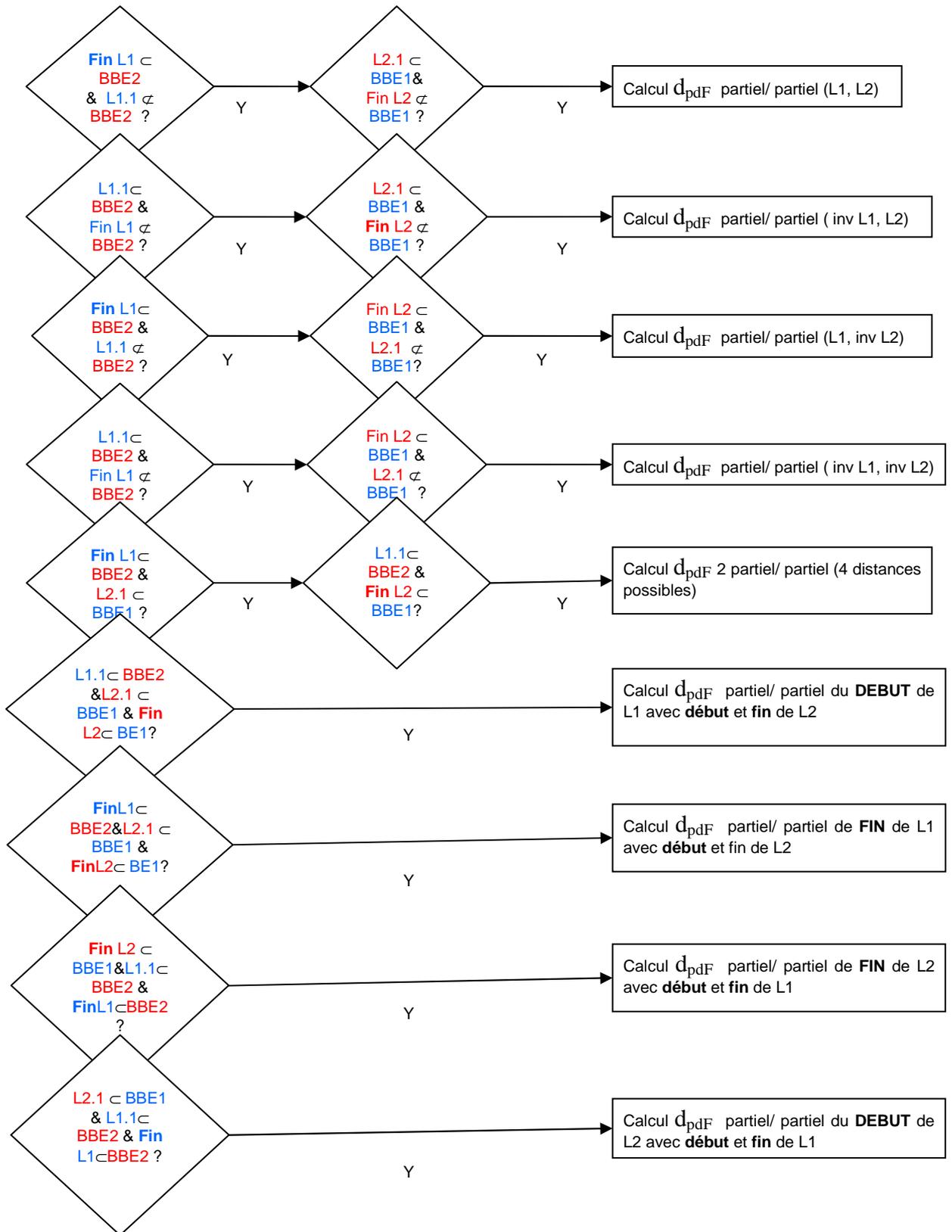


Figure 54 : Algorithme de test d'appariement partiel/partiel (PP) et deux partiel/partiel (2PP).

Pour proposer une synthèse des différents appariements possibles et des distances de Fréchet calculées dans chacun des cas, voici un tableau récapitulatif exprimé en fonction des relations d'inclusion :

Tableau 16 : Résumé des différents appariements présentés et de leurs distances de Fréchet associées.

Type appariement	Type L_1	Type L_2	RIE (Relation d'inclusion entre les emprises) (L_1, L_2)	Distances
1-1	O	O	équivalence	$d_F(L_1, L_2) < \epsilon$
	F	F		
1-P (non 1-1)	O	O	équivalence	$d_{pdF}(L_1, L_2) < \epsilon$ ou $d_{pdF}(L_2, L_1) < \epsilon$. Si les deux sont vrai \rightarrow choix utilisateur
			inclut	$d_{pdF}(L_1, L_2) < \epsilon$
			est inclus	$d_{pdF}(L_2, L_1) < \epsilon$.
	O	F	est inclus	$d_{pdF}(L_2, L_1) < \epsilon$.
PP et 2PP (non 1-P)	O	O	Non disjoint	7 cas possibles
2PP	O	O	recouvrement	$d_{ppdF}(L_1, L_2) < \epsilon$ ou $d_{ppdF}(L_1, L_{2inv}) < \epsilon$ ou $d_{ppdF}(L_{1inv}, L_2) < \epsilon$ ou $d_{ppdF}(L_{1inv}, L_{2inv}) < \epsilon$.

Les différents cas de figure rencontrés lors de la comparaison de lignes et les appariements induits sont repris dans l'algorithme qui décrit la méthode globale d'appariement par fichier. Les lignes ne sont plus comparées individuellement selon le choix de l'utilisateur mais par l'intermédiaire de fichiers regroupant l'ensemble des lignes.

II.3.4 Méthode globale : appariement par fichier

La méthode globale reprend tous les modules de comparaison et appariement décrits précédemment et les enchaîne selon un ordre précis. Les bases de données de lignes à comparer sont regroupées dans deux fichiers. Chaque fichier est importé dans un tableau (ou vecteur) de lignes.

L'algorithme de la figure 55 présente la méthode de traitement des fichiers de lignes dans son ensemble. Une comparaison deux à deux (L_i, L_j) avec L_i dans le fichier 1 et L_j dans le fichier 2 est effectuée. Après importation des lignes, leur nature (fermée ou ouverte) est ensuite déterminée : en effet les comparaisons successives impliquées ne sont pas les mêmes selon la situation considérée. Les trois cas possibles évoqués précédemment (ouvert/ouvert, ouvert/fermé ou fermé/fermé) déterminent les types d'appariement testés : total (ou partiel), total/partiel seul, partiel/partiel ou deux partiel/partiel.

Les distances de Fréchet partielles inférieures à ϵ sont sauvegardées dans un fichier dans un format qui permet une visualisation rapide des résultats via un SIG.

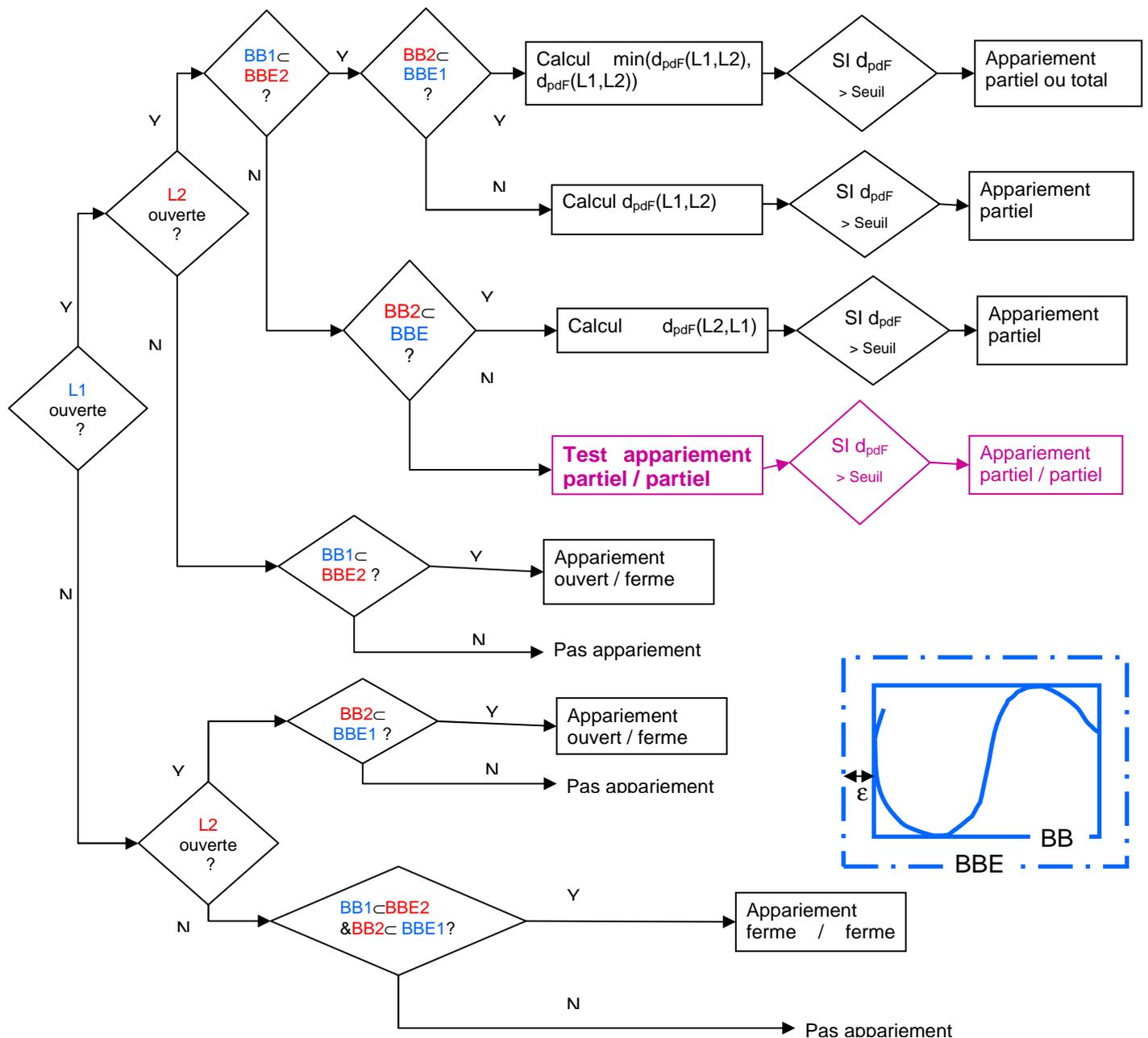


Figure 55 : Algorithme d'appariement global.

Le tableau 17 dresse un inventaire de tous les cas existants : ceux identifiés et traités, et ceux impossibles à résoudre. Les cas sont classés en fonction de la nature des deux lignes comparées (ouverte/fermée), de leur emprise (totale, incluse ou partielle) et du type d'appariement. Les rectangles englobants simples sont en trait plein et ceux étendus en trait pointillé.

Les cases contenant seulement une croix sont celles pour lesquelles il n'existe pas de cas possible. Les dessins grisés barrés d'une croix sont les situations pour lesquelles il n'y a pas de solution en l'état actuel de la recherche. Les dessins en gris sont les cas traités algorithmiquement mais non encore codés dans les outils d'intégration.

Cette méthode globale permet de traiter la majorité des cas parmi ceux identifiés : sont exclus ceux qui ne possèdent pas de solution simple (cas en gris clair dans le tableau).

L'implémentation de toutes ces méthodes de détection de cas, de tests des différents types de lignes, de leurs sens respectifs, et de calculs des appariements possibles, représente environ 5.300 lignes de codes réparties en 11 classes et 209 méthodes.

Afin de vérifier la validité de ce code au travers de la méthodologie globale, il est impératif de la tester sur des données simulées, représentatives de données réelles. De ces données sont extraites les lignes caractéristiques nécessaires à l'intégration, mais le problème de recouvrement est exclu.

Tableau 17 : Récapitulatif des différents cas de figure rencontrés lors de la comparaison de lignes (1) pour chaque type de ligne (ouverte ou fermée) ; (2) pour chaque type de recouvrement : total, inclus ou partiel (rectangles englobants simple et étendu associés).

	Total/total	Partiel/total	Partiel/ partiel	2 partiel /partiel
OO				
OO	X			
OO	X	X		
OF	X			X
OF	X		X	
OF	X	X		
FF			X	X
FF	X	X		
FF	X	X		

Le chapitre suivant développe plus en détail ces deux aspects : l'extraction d'éléments caractéristiques (présentés au chapitre III) de MNT possédant un recouvrement total, puis une intégration établie sur des modèles Terre/Mer avec un recouvrement partiel.

L'étude suivante détaille dans la section suivante une application de cet algorithme à la comparaison de traits de côte. L'objectif est de quantifier les différences entre une ligne de référence et une ligne obtenue avec des données plus récentes mais de moins bonne résolution.

II.4 Exemple d'applications de la méthode sur la comparaison de traits de côte

Cette étude, détaillée dans les travaux de [Le Berre et al., 2005], a été réalisée dans le cadre de l'appel d'offre CNES-IFEN d'août 2002 relatif au suivi du littoral à grande échelle par imagerie SPOT 5. L'étude présente l'expérimentation réalisée pour la délimitation et la caractérisation du « trait de côte géomorphologique » tel que défini par le Conseil National de l'Information Géographique (CNIG). Ce trait est défini comme étant « *la ligne de discontinuité morphologique caractéristique de la limite probable atteinte par la mer sur la côte* » [CNIG., 2003]. Une comparaison des traits de côte issus de la photo-interprétation de la BD ORTHO® et d'une image SPOT 5 avec des données de référence produites lors de levés topographiques sur le terrain est réalisée à l'aide de méthodes basées sur la distance linéaire de Fréchet. Cette comparaison produit un résultat numérique quantitatif qu'il est possible de comparer à une évaluation qualitative du numéricien.

II.4.1 Présentation des données et de la zone d'étude

Les données de référence pour la validation des résultats sont extraites des bases d'information géographique développées au laboratoire GEOMER sur le littoral finistérien ou ont été fournies par le Ministère de l'Équipement¹³ via le CETMEF et la DDE du Finistère, partenaires de ce projet. La comparaison quantitative des différents traits de côte produits et acquis est réalisée par l'institut de recherche de l'école navale (IRENav) à l'aide de méthodes basées sur la distance de Fréchet. Cette comparaison complète ainsi l'évaluation qualitative de la numérisation déjà réalisée.

Le projet a bénéficié de la programmation du satellite SPOT 5 pour l'acquisition d'une image au cours d'une basse mer de vive-eau sur le littoral nord-ouest finistérien. Une image en mode multispectral à 2,5 m de résolution a été acquise le 17 avril 2003, lors de la basse mer, ce qui permet une très bonne visualisation de l'estran.

La zone d'application retenue pour la comparaison des traits de côte se situe dans deux secteurs du nord Finistère. Ils ont été choisis du fait (1) de leur représentativité des différents types de côtes de la région, (2) de la disponibilité des données de référence et (3) de la présence sur le terrain de l'équipe dans la période encadrant l'acquisition de l'image.

II.4.2 Méthode de comparaison des traits de côte

La méthode de comparaison des traits de côte employée s'appuie sur des levés topographiques, sur l'interprétation de l'image SPOT 5 et de la BD ORTHO® IGN, puis sur l'analyse comparative des différents traits de côte ainsi produits.

Des levés topographiques détaillés de sections du trait de côte ont été réalisés sur le terrain dans la perspective de comparer les tracés obtenus à ceux issus de l'interprétation de l'image SPOT 5 et de l'orthophotographie. Ces levés ont été effectués le jour du passage du satellite dans la presqu'île de Sainte Marguerite en Landeda dans des zones de falaises et de plages. Dans les secteurs de falaises, les deux lignes de référence définissant généralement le trait de

¹³ Nouvellement renommé le Ministère de l'écologie, de l'énergie, du développement durable et de la mer.

côte, à savoir le pied de falaise [Bonnot-Courtois & Levasseur, 2003] et le rebord sommital de la partie érodée du versant [Robin, 2002], ont été mesurées. Dans le cas du secteur sableux, seul le contact entre la limite de la végétation dunaire a été pris en considération pour la cartographie du trait de côte [Robin, 2002]. Sur le secteur artificialisé c'est la base des ouvrages qui a été relevée.

II.4.3 Résultats et perspectives

L'image SPOT 5 tout comme la BD ORTHO[®] IGN sont traitées par interprétation assistée par ordinateur afin d'étudier, de manière fine, les possibilités de délimitation et de caractérisation du trait de côte.

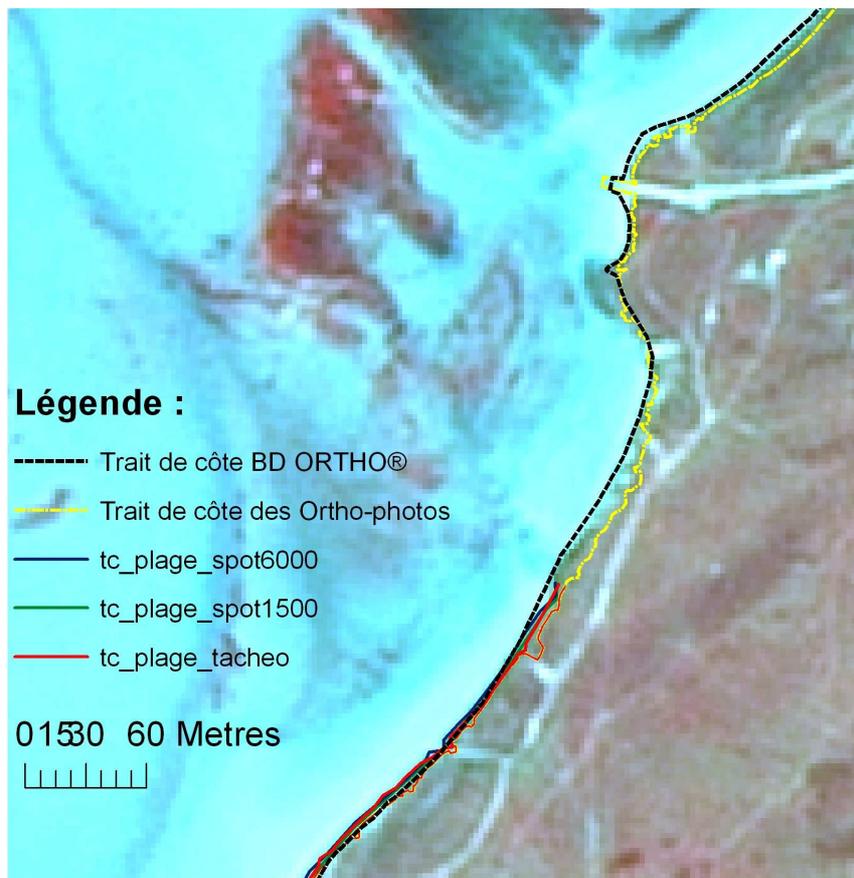


Figure 56 : Représentation des différents traits de côtes issus de la BD ORTHO[®], des levés de terrain et photo interprétés à partir des ortho-photographies, et de l'image SPOT 5.

L'image SPOT 5 a fait l'objet de deux numérisations séparées, la première à 1/6000 qui constitue une échelle confortable pour l'interprétation de l'image, la seconde à 1/1500 qui constitue l'échelle à partir de laquelle les pixels deviennent nettement visibles et commencent à constituer une gêne pour l'interprétation de l'image. La BD ORTHO[®] pour sa part a été numérisée à 1/1500, échelle de confort pour son interprétation.

Chaque numérisation aboutit à la production de couches d'information géographique vectorielles (voir figure 56). Deux traits de côte sont photo-interprétés d'après l'image SPOT 5 (aux échelles de 1/6000 et de 1/1500 : « tc_plage_spot6000 » et « tc_plage_spot1500 », légende Figure 56) et un troisième d'après la BD ORTHO[®] (à l'échelle de 1/1500 : « Trait de côte des Ortho-photos »). Les traits de côte de référence sont acquis par tachéométrie ou par GPS différentiel (nommés « tc_plage_tacheo »).

Les points des extrémités des segments constituant ces traits de côte sont exportés dans un fichier texte à l'aide d'Arc/Info. Ces lignes sont traitées par un exécutable développé en C++. Elles sont suréchantillonnées tous les demi-mètres pour être comparées à l'aide de mesures dérivées de la distance de Fréchet [Devoegele, 2000]. La précision de numérisation entraîne une estimation décimétrique de l'écart entre une ligne de référence et une ligne saisie. Elles qualifient la précision de numérisation en comparant les traits de côte numérisés avec les traits de côte de référence.

Tableau 18 : Distance en mètres entre les lignes saisies et les lignes de référence : distance de Fréchet discrète partielle moyenne suivie de la distance de Fréchet discrète partielle entre parenthèses (précision décimétrique).

Photo- interprétation	Trait de côte de référence (levé tachéomètre ou GPS)					
	Rebord sommital de falaise (Brouennou)	Falaise végétalisée (baie des Angés)	Dune (Ste Marguerite)	Grève (baie des Angés)	Cordon littoral (baie des Angés)	Ouvrages côtiers (baie des Angés)
SPOT 5, 1/6000	3,34 (6,07)	4,54 (11,98)	2,25 (6,08)	0,96 (3,90)	2,01 (4,18)	2,37 (6,26)
SPOT 5, 1/1500	1,62 (3,78)	2,77 (10,84)	1,17 (3,79)	1,22 (5,20)	1,45 (3,39)	3,58 (14,01)
BD ORTHO®, 1/1500	1,03 (3,10)	2,8 (7,25)	2,1 (8,64)	1,41 (5,38)	1,82 (5,08)	1,77 (7,37)

Sur la zone de falaise vive et verticale, les résultats du tableau 18 montrent qu'en moyenne la précision de photo-interprétation réalisée à partir de la BD ORTHO® est meilleure que celle issue de SPOT 5. Pour la zone de falaise végétalisée, les résultats sont relativement proches et montrent que la végétation constitue manifestement une contrainte pour une numérisation précise.

En zone dunaire, les traits de côte issus de SPOT 5 sont légèrement meilleurs que ceux issus de la BD ORTHO®. La simple évolution qu'a connue la ligne de rivage en haut de plage entre le moment de la prise orthophotographique (juin 2000) et des acquisitions SPOT 5 et tachéométrique (avril 2003), explique la meilleure précision obtenue ici à partir de l'image SPOT 5. Elle démontre ainsi l'intérêt de l'imagerie SPOT 5 pour le suivi à court terme de l'évolution des littoraux meubles.

De même, les meilleurs résultats obtenus à l'aide de SPOT 5 pour l'interprétation du cordon de galet et de la zone de grève (figure 57) doivent être interprétés avec prudence et à la lumière d'une possible évolution du trait de côte ou de la végétation qui la recouvre entre les deux prises de vues.

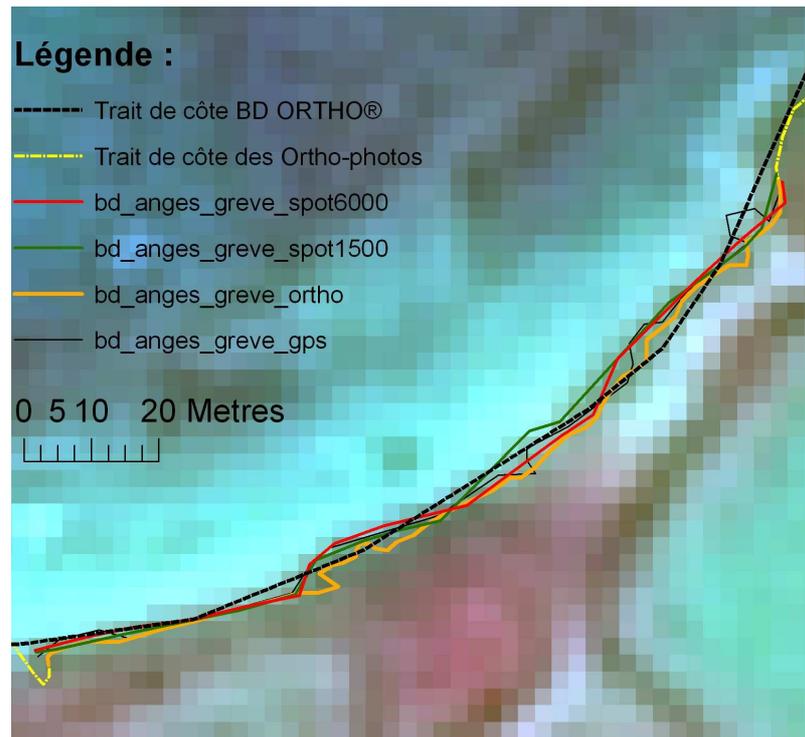


Figure 57 : Illustration des différents traits de côte numérisés au 1/1500 et 1/6000 comparés à la BD ORTHO® et à des levés GPS établis sur une grève.

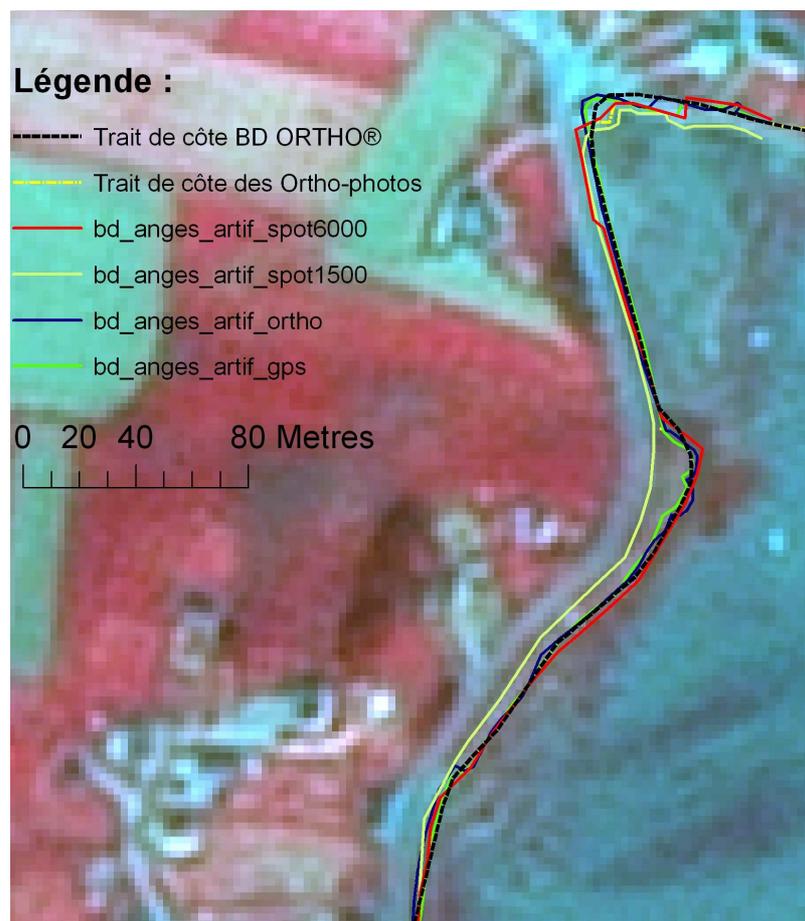


Figure 58 : Comparaison des traits de côte issus de photo interprétation à partir de l'image SPOT 5 avec le trait de côte de la BD ORTHO® sur une côte artificielle.

En zone artificielle l'ortho photographie permet l'obtention de résultats sensiblement meilleurs (25% d'amélioration).

Le faible résultat obtenu pour le trait de côte numérisé à 1/1500 (trait bleu figure 58) d'après SPOT 5 est essentiellement lié à une différence dans l'interprétation et dans le choix de l'objet à numériser, qui est également à l'origine de la forte distance de Fréchet discrète (écart maximal) dans les trois cas. Cette valeur témoigne de la difficulté à identifier précisément les ouvrages côtiers.

Ainsi, globalement, sur cette zone d'étude, si l'on excepte la numérisation effectuée au 6000ème d'après l'image SPOT 5, les photo-interprétations de SPOT 5 ont une précision du même ordre de grandeur que celles de la BD ORTHO®.

Au vu des résultats obtenus, tant en comparaison visuelle qu'à l'aide des valeurs obtenues avec la distance de Fréchet, il apparaît que l'utilisation de SPOT 5 semble appropriée pour la remise à jour du trait de côte. En effet, malgré des incertitudes de délimitation liées notamment aux effets d'ombre, certains secteurs donnent de bons résultats. Il s'agit principalement des zones dunaires sur lesquelles le contraste entre la partie terrestre et le domaine marin est souligné par la végétation, qui permet une bonne détermination de la limite côtière. Compte tenu de la mobilité de ce type de côte, une voie potentiellement intéressante peut être explorée.

D'après les tests de numérisation et la comparaison des mesures effectuées in situ, il apparaît que ces secteurs de côte peuvent être cartographiés avec une précision satisfaisante au regard de celle obtenue à partir de l'ortho-photographie. L'image SPOT 5 offre donc un certain nombre de perspectives en matière de suivi à court terme de l'évolution du trait de côte qui s'additionnent aux potentialités offertes par la répétitivité d'acquisitions de données.

II.5 Bilan sur les apports des extensions

Ce chapitre introduit tout d'abord les différents types de lignes présentes lors d'une comparaison : ligne ouverte, fermée ou d'orientations différentes. Ces notions sont essentielles pour comprendre la nécessité des extensions apportées à la distance de Fréchet. En effet la méthode définie par [Eiter & Mannila, 1994] ne s'applique qu'à des lignes de tailles similaires et de même orientation. Les données géographiques remplissent rarement ces deux impératifs.

Par ailleurs, la comparaison de lignes de forme complexe amène à définir d'autres distances complémentaires nécessaires aux besoins de cette thèse : distance moyenne, distance partielle entre lignes, distance entre lignes fermées, ou encore entre des extrémités de ligne uniquement (distance « deux partiel/partiel »).

La section trois de ce chapitre reprend donc ces distances et les intègre dans les algorithmes d'appariement de lignes en décrivant les trois étapes du processus d'appariement. La première étape détaille tout d'abord le processus de recherche de la nature des lignes comparées et vérifie leur orientation. La seconde étape inventorie les types d'appariement mis en œuvre dans la troisième étape au moyen des relations d'inclusion existant entre les lignes :

- total,
- total/partiel,
- partiel/partiel ou deux partiel/partiel.

La dernière partie synthétise les cas possibles au travers de la description d'une méthode de comparaison et d'appariement globale entre des fichiers de lignes.

L'exemple d'application détaillé en dernière section permet d'illustrer une application possible de la distance de Fréchet tout en validant les divers algorithmes de comparaison de lignes présentés en section 4.1. En revanche cette comparaison s'est faite sur des lignes 2D numérisées choisies par l'utilisateur.

Il s'agit donc maintenant de tester l'algorithme d'appariement global sur des lignes extraites des MNT (thalwegs, crêtes, ruptures de pente) en l'intégrant à une méthodologie d'intégration générique. Cette méthodologie doit permettre d'intégrer tous types de données quel que soit le cas de figure rencontré. Pour chaque élément apparié, des mesures de distances moyenne, maximale ainsi que les couples de points homologues sont obtenues. Cette étape originale est absente des autres méthodes d'intégration présentées en état de l'art. Elle permet d'apparier des éléments morphologiques caractéristiques extraits des modèles numériques de terrain. La méthodologie globale d'intégration des MNT terrestre et maritime va maintenant être détaillée ainsi que son application à des données simulées.

III METHODOLOGIE D'INTEGRATION

Sommaire

III.1	Présentation des données	97
III.1.1	Données géographiques acquises	97
III.1.2	Evaluation de la cohérence des données géographiques	101
III.1.3	Données Simulées	107
III.2	Méthode d'intégration globale	114
III.2.1	Comparaison de MNT	115
III.2.2	Segmentation et extraction de lignes caractéristiques	117
III.2.3	Intégration des modèles de terrain.....	119
III.3	Validation de la méthodologie sur les données simulées.....	122
III.3.1	Modèle du chenal	122
III.3.2	Modèle Numérique de Terrain représentant un littoral	130
III.4	Perspectives d'application de la méthodologie sur des données littorales	153
III.4.1	Conversion géographique.....	154
III.4.2	Partition en domaines de recouvrement	154
III.4.3	Analyse de corrélation.....	154
III.4.4	Analyse paysagère.....	154
III.4.5	Appariement	157
III.4.6	Corrections élastiques	157
III.5	Bilan sur la méthodologie d'intégration.....	157

La comparaison de différents traits de côte en vue de l'évaluation de leur qualité respective constitue une validation pertinente de l'algorithme d'appariement car il permet d'expérimenter le cœur même de la méthode. En effet ce test a eu pour conséquences d'induire de nombreuses améliorations de l'algorithme. Ces modifications ont eu pour but d'optimiser l'efficacité des corrections, d'augmenter le nombre de cas solutionnés et d'automatiser certains traitements longs et fastidieux.

Par ailleurs cette étude a soulevé des problèmes consécutifs liés à la structure des données elle-même :

- sens des lignes différents ;
- données non continues en raison d'une numérisation parfois approximative ;
- cas particuliers difficilement gérables ;
- systèmes de coordonnées non similaires.

Cette étude a également permis d'identifier les cas insolubles en l'état actuel de nos travaux. Toutes les manipulations nécessaires à la préparation des données pour le processus d'appariement ont été réalisées manuellement dans un premier temps.

L'aboutissement de cette thèse est de dépasser le simple appariement 2D de lignes en proposant une chaîne de traitement complète et exhaustive organisée sous forme d'une méthodologie d'intégration de MNTs générique et robuste. L'adaptation de l'algorithme aux modèles numériques amène à baser leur appariement sur des éléments morphologiques structuraux extraits de ces modèles : lignes de forces, points remarquables, etc. Le principe est d'extraire ces éléments sur la zone de recouvrement des deux modèles et d'y rechercher les éléments caractéristiques homologues. Pour satisfaire cette attente, la méthodologie a été ordonnancée en une succession d'étapes qui ont pour but de répondre à l'ensemble des problèmes qui peuvent se présenter lors d'un processus d'intégration.

Ce chapitre s'organise autour de quatre parties qui procèdent les unes des autres par leur organisation. La première partie présente tout d'abord les données utilisées pour tester et valider cette méthodologie. Elles comportent d'une part des données réelles suffisamment précises et redondantes pour constituer des zones d'études bien renseignées. Ce type de données implique une certaine complexité de traitement de systèmes géographiques différents, de la résolution variable, de zone de recouvrement insuffisante, ou de densité de données importante. D'autre part des données simulées sont utilisées pour maîtriser toutes ces imprécisions mais aussi écarter dans un premier temps les problèmes de cohérence entre données.

Puis, la seconde partie détaille la méthode et ses trois étapes fondamentales :

- tout d'abord une phase d'analyse des données en présence. L'objectif de cette phase est de corriger les problèmes de cohérence entre données : vérification des référentiels et étude de la surface de recouvrement entre les données ;
- ensuite une phase de segmentation qui vise à extraire les lignes caractéristiques essentielles à l'étape d'appariement ;
- et enfin la partie intégration à proprement parler : gestion des zones sans informations, appariement, correction élastique, fusion et production d'un MNT continu de la zone.

La troisième partie expose ensuite une validation de cette méthodologie au travers de deux exemples d'intégration appliquée à des données simulées. La première intégration s'effectue sur un modèle de chenal simplifié : les différents types de déformations y sont comparés et leurs tendances de corrections (bonne amélioration de la qualité du modèle, persistance d'artéfacts locaux, etc...) analysées. La seconde intégration concerne un modèle simulé

approchant au mieux un environnement littoral. Les différents types de corrections sont à nouveau appliqués. Les tendances des résultats sont comparées avec celles obtenues sur le chenal afin de vérifier si les méthodes utilisées donnent toujours de bons résultats ou si c'est seulement dans une configuration précise.

La quatrième partie retrace les perspectives d'application à des cas plus concrets et les améliorations possibles et souhaitées pour chaque étape de la méthodologie.

III.1 Présentation des données

Divers types de données ont été utilisés pour tester d'une part la méthodologie d'intégration et d'autre part la valider. Un premier jeu de données réelles issu de la base de données géographique a été extrait sur des zones pertinentes pour cette étude. Un second jeu a été créé artificiellement pour tester la méthodologie premièrement sur un élément géographique pertinent et ensuite sur des modèles de relief approchant la réalité. Ces données, qu'elles soient réelles ou simulées, vont maintenant être présentées.

III.1.1 Données géographiques acquises

III.1.1.1 Description de la zone d'étude

Deux zones préférentielles ont été choisies (voire figure 59 et figure 60) en raison de la diversité des informations présentes (données issues de sonars du SHOM, LIDAR topographique et/ou bathymétrique, de cartes, de levés de terrain) et de la variabilité paysagère existante.

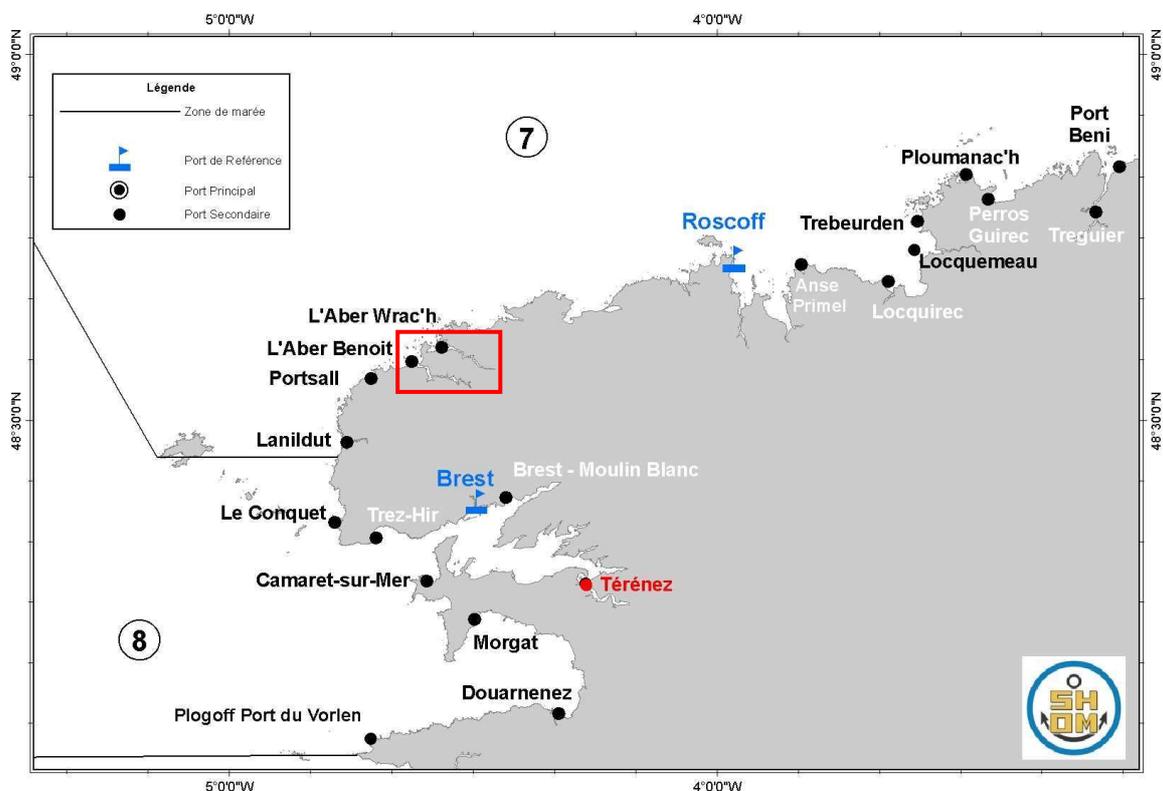


Figure 59 : Localisation de la première zone d'étude (encadrée en rouge sur la carte du SHOM) : zone des Abers.

En effet, ces secteurs (Abers et Golfe du Morbihan) présentent l'intérêt d'avoir soit un bon échantillonnage des différents types de reliefs littoraux, soit une complexité suffisante pour concentrer la majeure partie des problèmes rencontrés lors de l'intégration.

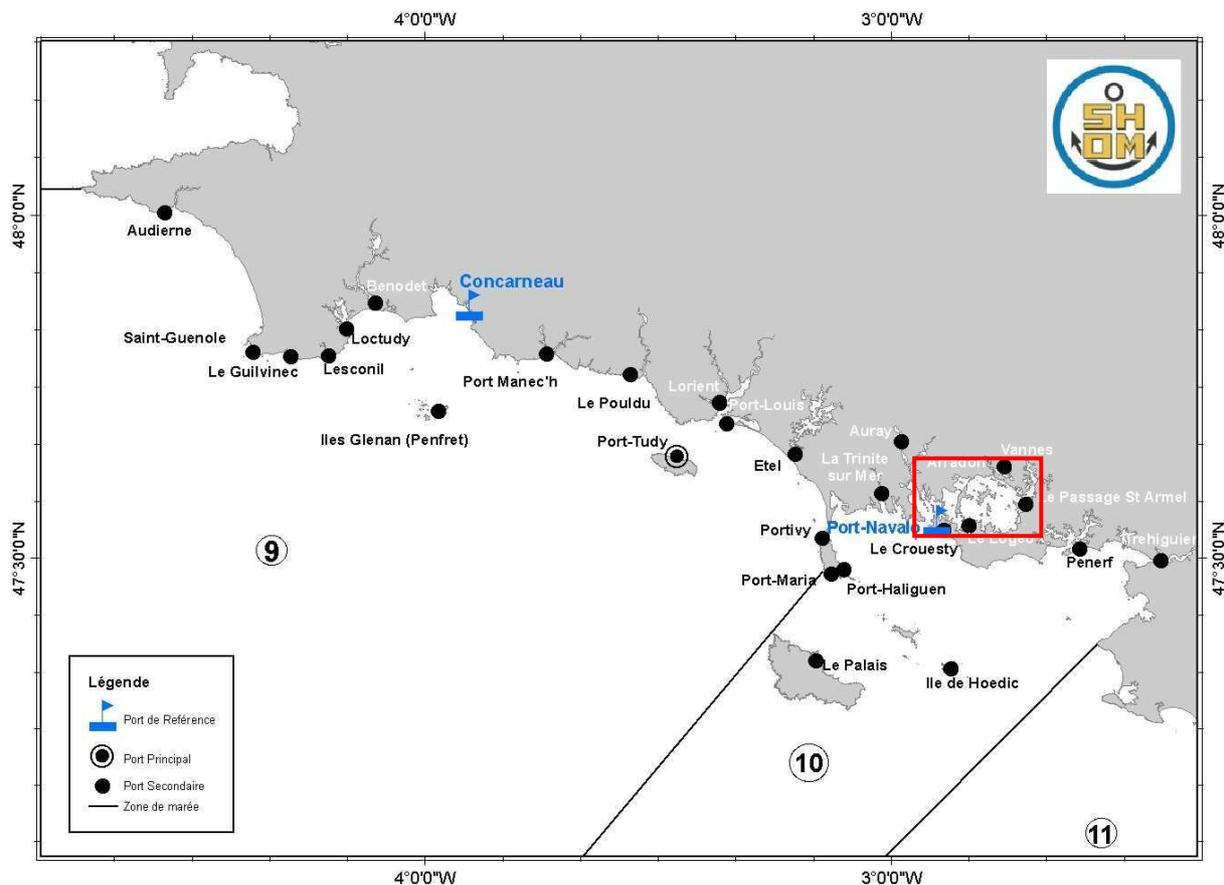


Figure 60 : Localisation de la deuxième zone d'étude (encadrée en rouge sur la carte du SHOM) : golfe de Morbihan.

Zone des Abers :

Le secteur des Abers (figure 59) est compris entre Ploudalmézeau et Plouguerneau au Nord de la Bretagne. Il comporte des portions de falaises meubles, des côtes rocheuses, des ensembles dunaires, de larges estrans et deux estuaires : l'Aber Benoît et l'Aber Wrac'h (voir photo 1 et suivantes).



Photo 1 : Vue d'ensemble de l'embouchure de l'Aber Wrac'h (presqu'île de Sainte Marguerite). Photo prise le 8 Avril 2007 à 16h40 du haut des dunes situées dans le haut estran (orientation de la prise de vue : ouest).

Des données sur cette zone ont été collectées dans le cadre de plusieurs études scientifiques. La première concerne les données topographiques aéroportées Lidar de haute résolution

fournies par l'IFREMER¹⁴. Le but de ces levés était d'avoir une meilleure connaissance de la topographie de la zone côtière. La seconde a trait aux données SPOT et aux levés de terrain effectués en vue de l'évaluation des données SPOT 5 pour la cartographie et la remise à jour du trait de côte [Le Berre et al., 2005].



Photo 2 : Cette photo (encadré 1 de la photo 1) souligne les très faibles ruptures de pente dans le haut estran (schorre). Au premier plan, la limite de la dernière pleine mer de vive eau est marquée par la laisse de haute mer (cordon d'algue). Orientation de la prise de vue : Est.



Photo 3 : Photo panoramique de l'estran à l'embouchure de l'Aber Wrac'h (encadré 2 de la photo 1). Elle souligne le passage de la zone sableuse au platier rocheux dans le bas estran (orientation de la prise de vue : Sud Ouest).

Golfe du Morbihan :

Le golfe du Morbihan (figure 60) s'ouvre sur la baie de Quiberon par un étroit passage entre Locmariaquer et Port-Navalo. Sa géomorphologie de criques, pointes, rochers, îles et îlots ainsi que son régime marégraphique complexe font qu'il a été choisi par le SHOM comme zone test pour la réalisation du démonstrateur du projet litto3D [SHOM & Debese, 2006]. Celui-ci concentre en effet la plupart des difficultés susceptibles d'être rencontrées sur les côtes françaises tant au niveau de la diversité du relief qu'en terme de modélisation marégraphique. Il est constitué d'un régime hydrologique complexe caractérisé par l'existence de multiples zones de marée, ce qui impose d'utiliser plusieurs référentiels hydrographiques. Ces levés conjoints Lidar topo-bathymétrique et échosondeur multifaisceaux ont en outre permis de tester les chaînes de mesures et soulever certains problèmes quant à l'ordre d'acquisition entre mesures Lidar et acoustique.

¹⁴ Avec l'amabilité de Mr Jacques Populus.

III.1.1.2 Description des jeux de données

La liste des données disponibles sur les deux zones d'intérêt des Abers et du golfe du Morbihan avec leurs caractéristiques géographiques et numériques est donnée ci-après :

Lidar topographique (source IFREMER)

- Résolution : pixel de 2 mètres pour la grille raster ;
- Système cartographique : projection Lambert 93 associée au système géodésique (datum) RGF93 ;
- Système altimétrique : zéro IGN 69 (en mètre système altimétrique terrestre);
- Format de départ des données : grille raster interpolée avec la méthode de Krigeage (résolution choisie : 2 mètres), points ré échantillonnés (2 mètres) et points bruts ;
- Poids des données : grille raster (137 Mo) ; points ré échantillonnés (300 Mo) ; points bruts (2,7 Go).

Lidar topo-bathymétrique (source IFREMER)

- Résolution : pixel de 1 mètre ;
- Système cartographique : coordonnées géographiques WGS84 ;
- Système altimétrique : hauteur ellipsoïdale (EGM96) ;
- Format de départ des données : points bruts ;
- Poids des données : 5 Go.

Image SPOT 5 (source UBO)

- Résolution : pixel de 5 mètres ;
- Système cartographique : recalé en Lambert 93 ;
- Format de départ des données : image raster ;
- Poids des données : 140 Mo.

Levés sonar mono faisceau (source SHOM)

- Résolution : semis de points irréguliers ;
- Système cartographique : coordonnées géographiques WGS84 ;
- Système altimétrique : zéro hydrographique des cartes marines ;
- Format de départ des données : fichier texte xyz de points ;
- Poids des données : 1 Mo.

Lidar topo-bathymétrique (source SHOM)

- Résolution : pixel de 1 mètre ;
- Système cartographique : WGS84 ;
- Système altimétrique : hauteur ellipsoïdale (EGM96) ;
- Format de départ des données : points bruts ;
- Poids des données : 200 Go.

Cette description des jeux de données montre qu'il y a des disparités dans les systèmes de projection ainsi que dans les systèmes altimétriques, spécifiques aux méthodes d'acquisition en milieu littoral. Les difficultés majeures qui vont se poser lors de l'intégration sont donc de définir un référentiel commun à toutes les données et avoir une homogénéité temporelle suffisante entre les jeux de données. Nous allons donc étudier la cohérence des données géographiques et produire une estimation de celle-ci.

III.1.2 Evaluation de la cohérence des données géographiques

L'ensemble des SIG sur le marché propose, lors de l'importation de données géographiques, une fonction qui permet de les repositionner dans un système cartographique choisi par l'utilisateur. Ce recalage se fait de manière plus ou moins précise selon les paramètres par défaut utilisés par le logiciel ou ceux précisés par l'utilisateur dans le cas d'un système géographique local. En revanche très peu proposent des outils permettant d'évaluer la cohérence des données sur l'altitude (ou la profondeur selon la zone géographique étudiée). Celle-ci dépend d'une part des changements altimétriques au cours du temps (variations temporelles) et d'autre part des méthodes d'acquisition mises en jeu.

Evaluer la cohérence des données c'est établir les différences entre les modèles numériques et calculer les densités relatives de chacune des méthodes d'acquisition. Cette étude n'est réalisée que sur la zone des Abers.

III.1.2.1 Etude des valeurs de différences entre données altimétriques

Une carte de soustraction pixel à pixel des modèles raster interpolés à partir des points xyz apparaît alors comme l'outil le plus pertinent pour évaluer la cohérence altitudinale des modèles. Ces cartes ont été générées avec le module « Raster calculator » d'ArcGIS 9.0.

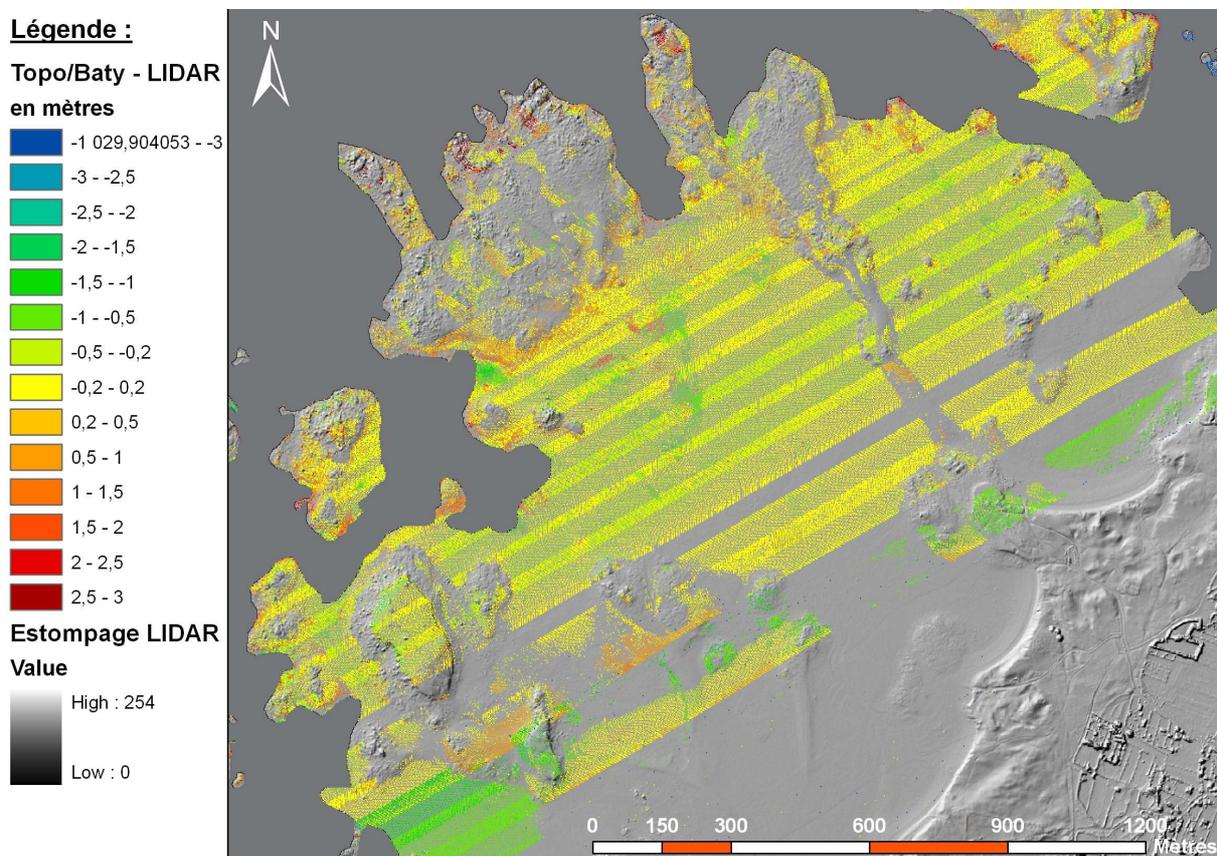


Figure 61 : Carte de différences exprimées en mètres entre les données issues du Lidar topo-bathymétrique et celles acquises par le Lidar topographique après correction de la hauteur ellipsoïdale.

La figure 61 montre un estompage du MNT Lidar topographique et le résultat de la soustraction entre le modèle Lidar topo-bathymétrique et le modèle topographique simple. La gamme de couleurs montre une différence altimétrique globale des deux MNT comprise entre 51 et 55 mètres. L'importance de l'écart est due au choix du référentiel altitudinal pour exprimer la profondeur bathymétrique. En effet les acquisitions en altitude se sont faites parallèlement par rapport à la **hauteur ellipsoïdale**, et non une **altitude**, exprimée par rapport au marégraphe de Marseille. Après conversion du référentiel, les modèles ne présentent plus qu'une différence maximale de moins d'un mètre (voir figure 61).

La figure 62 illustre les différences d'altitudes entre les points acquis par sonar mono faisceau et la topographie Lidar. Les points ont été interpolés en grille avec la méthode de voisins naturels avec un rayon d'investigation choisi de telle sorte que les zones sans information ne soient pas comblées.

La palette colorimétrique montre entre les deux modèles une différence moyenne de 1 mètre (en positif ou en négatif) et une différence maximale de 7 mètres pour les régions de l'estran les moins bien couvertes par le sonar (domaine proche terrestre). Le fort écart pour ces zones de l'estran s'explique par la difficulté du sonar à acquérir en milieu peu profond. Les différences minimales entre les données sur le reste des berges de l'Aber impliquent une bonne cohérence entre les représentations.

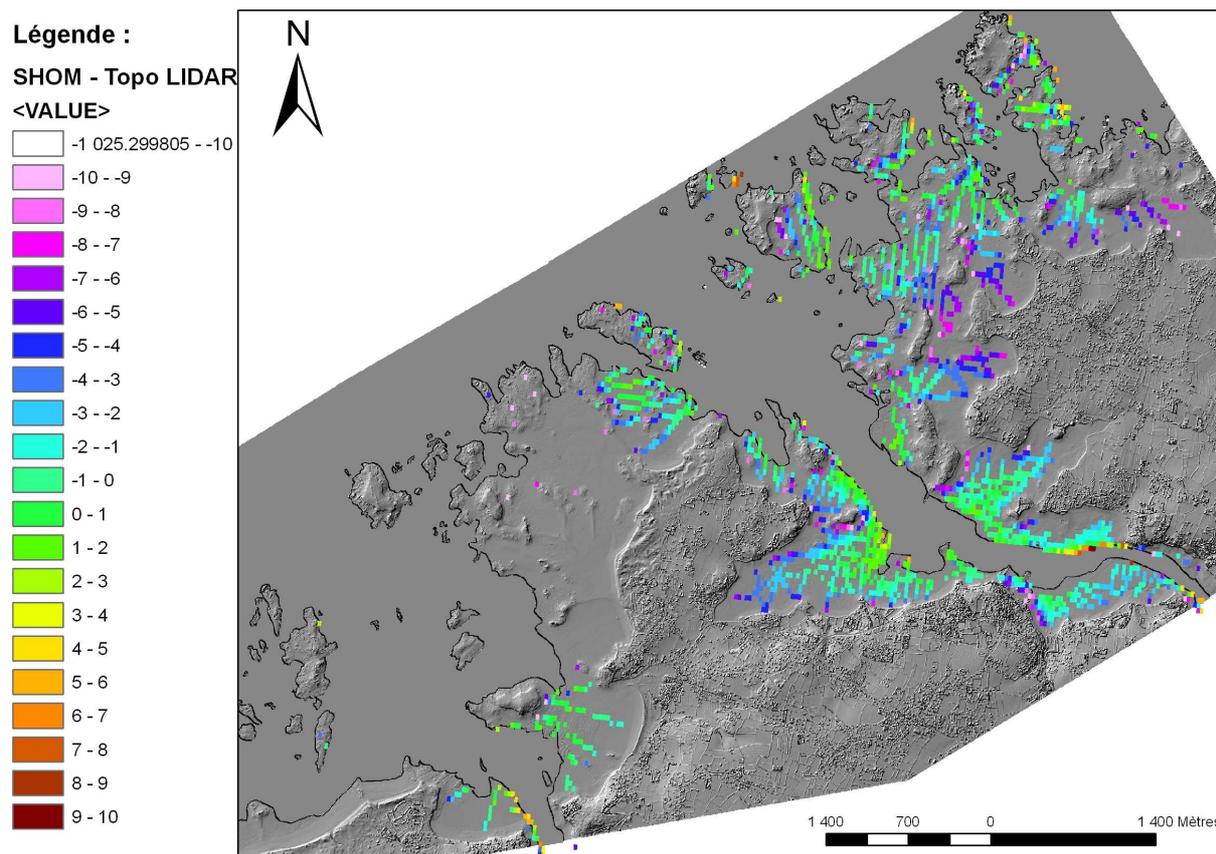


Figure 62 : Carte des valeurs de différences en mètres entre les données d'acquisition sonar mono-faisceau du SHOM et celle acquises par le lidar topographique (Ifremer).

La figure 63 représente la soustraction d'une grille générée à partir des données Lidar topo-bathymétrique et d'une grille produite à partir des données sonar. Là encore, la forte différence entre les deux types de données est due aux différents référentiels utilisés pour les mesures respectives des profondeurs bathymétriques. Néanmoins, après conversion des

données, des écarts importants de l'ordre d'une dizaine de mètres sont encore observables au niveau des chenaux des Abers Benoît et Wrac'h.

Un début d'explication est fourni par le fait que les données sonar sont traitées conformément aux règles de sûreté de navigation. Lors de l'acquisition, le signal enregistré représente en réalité un intervalle réduit de valeurs de profondeurs. La mesure finale gardée est celle qui traduit la profondeur la plus faible de sorte à toujours posséder une distance de sûreté en cas d'erreur de mesure, si minime soit elle.

Par ailleurs, nous ne disposons pas pour les levés Lidar d'estimations de la qualité et de la fiabilité des levés, simplement de l'indication d'une valeur par défaut égale à -100 lorsque le fond n'a pas pu être enregistré. De même, les levés sonar étant plus anciens que ceux établis par le Lidar, une modification de la morphologie du chenal due à l'érosion peut également expliquer les différences trouvées.

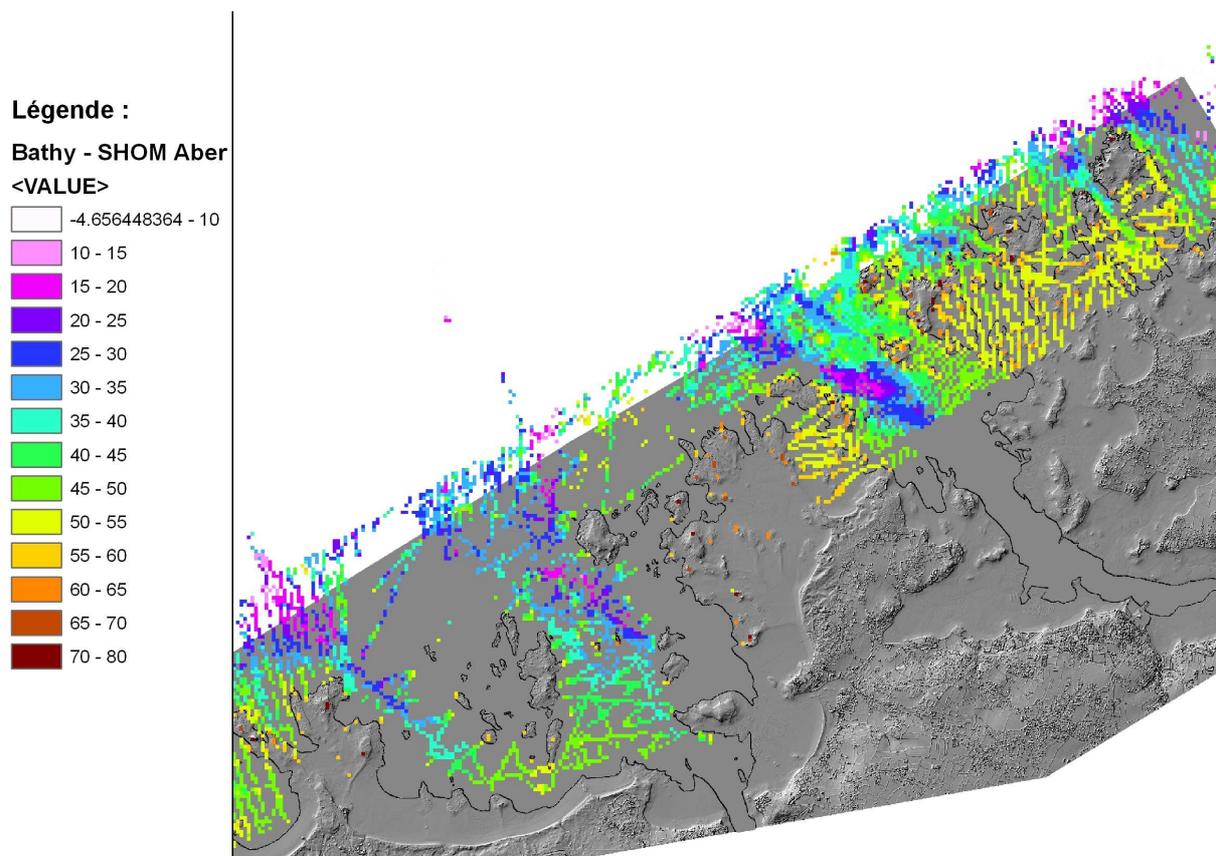


Figure 63 : Carte de différences entre les données bathymétriques acquises par échosondeur mono faisceau du SHOM et par du lidar topo-bathymétrique (source : Ifremer).

Au vu de ces cartes de différences, un autre problème lié à la qualité intrinsèque des données est apparu : il concerne leur homogénéité d'acquisition sur une région donnée.

En effet si les données Lidar possèdent une densité d'échantillonnage importante, ainsi qu'un recouvrement conséquent, elles n'en ont pas moins des limitations concernant les milieux d'acquisition. Ces points ont été détaillés au Chapitre II : le Lidar topographique ne peut acquérir en milieu marin, aussi faible que soit la profondeur, le Lidar topo-bathymétrique est limité par la turbidité de l'eau tandis que le sonar nécessite une profondeur minimale pour ses acquisitions.

Il en résulte des mesures parfois non homogènes sur l'ensemble des régions acquises. Des cartes de densité pour chacune des données ont donc été établies de manière à présenter la répartition des données.

III.1.2.2 Comparaison de la densité d'échantillonnage des données altimétriques

En plus des différences altimétriques présentées précédemment, il existe des disparités d'échantillonnage entre les jeux de données. Ces densités ont été calculées pour chaque lot de données avec le module « spatial analyst » d'ArcGIS et sont exprimées en points/m² ou en points/10m² de manière à lisser les résultats.

Lidar topographique :

Le calcul de la densité des données Lidar topographique est montré figure 64. La grille « LIDAR topographique » donne une densité de 20 points par 10m² (ou 0.2 /m²). Les données desquelles est extraite cette grille ré échantillonnée donne des résultats différents : un extrait est présenté en superposition des données régulières figure 64 au niveau de l'Aber Wrac'h. Les bandes de recouvrement des données lors de l'acquisition sont très nettement visibles. La densité y est alors plus que doublée et peut avoisiner les 7 points /m². Les zones de plan d'eau sont peu voire pas couvertes du tout : seule la partie totalement émergée de l'estran est bien représentée. Le trait de côte en rouge indique la fin du domaine terrestre à proprement parler. La laisse de haute mer (trait pointillé superposé au trait de côte) et la laisse de basse mer (trait bleu sur la figure) délimitent l'estran. Il est à noter que la laisse de basse mer marque également la limite des acquisitions Lidar topographique sur domaine maritime proche.

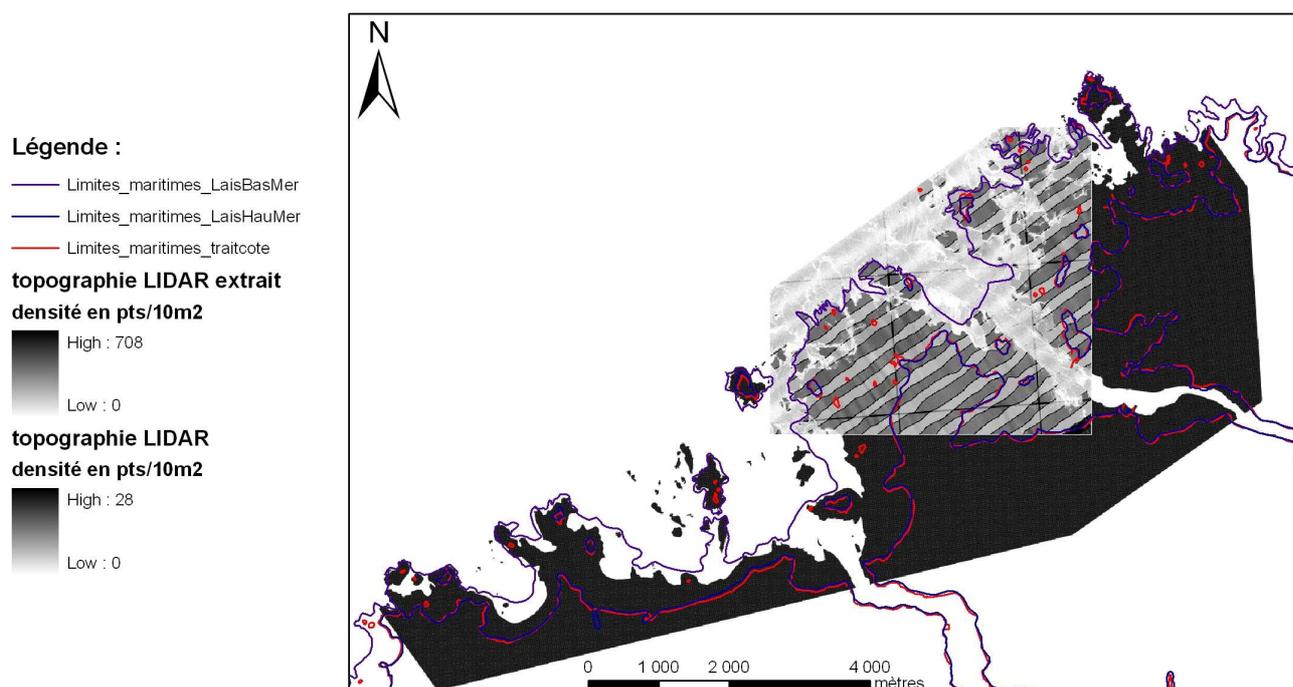


Figure 64 : Carte de densité des données Lidar topographique. En noir homogène, la densité de la grille ré échantillonnée à 2 mètres de résolution. En extrait sur l'Aber Wrac'h, le semis de points sources.

La densité moyenne des points avoisine alors les 3 points/m² (zone de recouvrement et de non recouvrement confondues) mais n'est pas totalement homogène sur l'ensemble de l'acquisition.

Lidar topo bathymétrique :

Les données Lidar topo-bathymétriques, détaillées figure 65, présentent le même aspect de *bandes* de densité dû au recouvrement des fauchées lors des levés aéroportés. Dans

l'ensemble l'estran est très bien couvert avec une densité moyenne de 4 points/m² tandis que la zone maritime voit la densité de points descendre à moins de 1 point/m². Cette chute brutale s'explique par la turbidité importante de l'eau qui a entraîné un non retour du signal conjointement à l'enregistrement d'un bruit non négligeable lors des mesures.

Sondes bathymétriques :

Le calcul de densité des données sonar est illustré figure 66.

Légende :

- Limites_maritimes_LaisBasMer
- Limites_maritimes_LaisHauMer
- Limites_maritimes_traitcote

topo bathymétrie LIDAR

densité en points/10m²

- 0
- 0 - 5
- 5 - 10
- 10 - 15
- 15 - 20
- 20 - 25
- 25 - 30
- 30 - 35
- 35 - 40
- 40 - 45
- 45 - 50
- 50 - 55
- 55 - 60
- 60 - 65
- 65 - 70
- 70 - 75
- 75 - 80

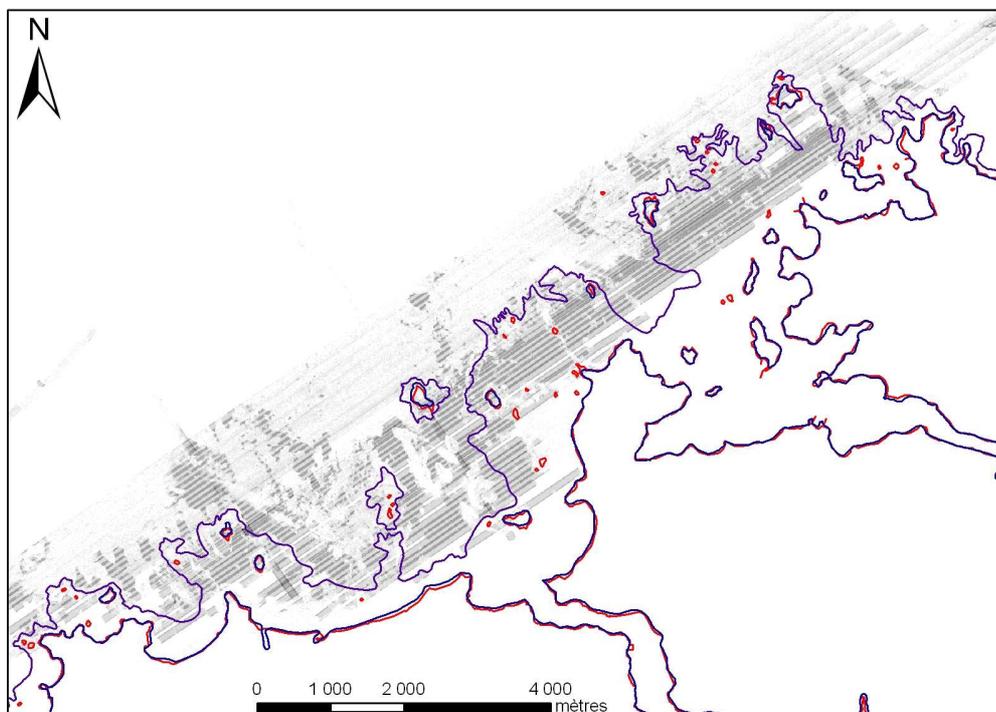


Figure 65 : Carte de densité des points topo-bathymétriques (Source des données : Ifremer).

Légende :

- Limites_maritimes_LaisBasMer
- Limites_maritimes_LaisHauMer
- Limites_maritimes_traitcote

bathymétrie des Abers

VALUE

- 0
- 0 - 5
- 5 - 10
- 10 - 15
- 15 - 20
- 20 - 25
- 25 - 30
- 30 - 35
- 35 - 40
- 40 - 45
- 45 - 50
- 50 - 55
- 55 - 60
- 60 - 65

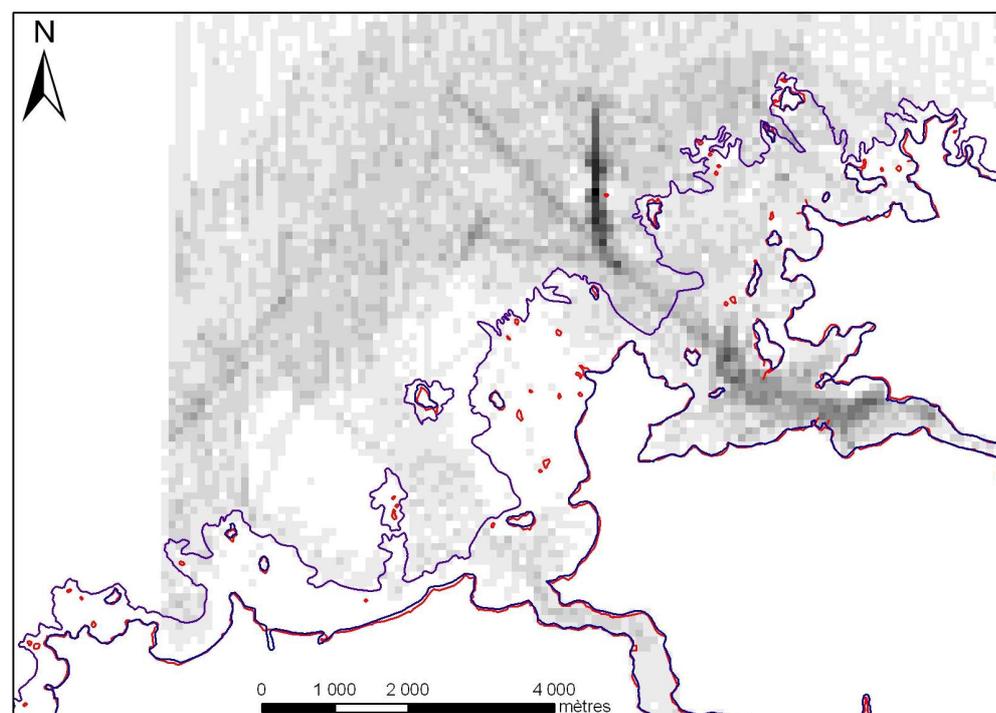


Figure 66 : Carte de densité des points acquis par écho sondeur mono faisceau (Source des données : Service Hydrographique et Océanographique de la Marine).

La mauvaise qualité de la carte de densité est la conséquence d'une part de la grande hétérogénéité de répartition des points et d'autre part du faible échantillonnage du semis de points des côtes sonar. En effet la densité moyenne des données lidar bathymétrique est de l'ordre de 0.4 point/m² contre 0.2 point/m² pour les données sonar. Par ailleurs, les contraintes d'acquisition du sonar mono faisceau oblige à avoir une tranche d'eau minimale. C'est pourquoi l'embouchure du fleuve (en noir figure 66) est très bien détaillée (densité maximale) contrairement à l'estran.

Synthèse des cartes de densités :

Le calcul de la densité de tous les points regroupés dans un même fichier permet d'avoir un bon aperçu des zones délicates à traiter lors de l'intégration : la zone uniquement terrestre, délimitée par le trait de côte n'entraîne pas de difficultés d'acquisition particulières. De même pour les zones maritimes de profondeur suffisante pour permettre le passage d'un échosondeur, complément optimal du Lidar. En revanche l'espace intermédiaire entre la laisse de haute mer et la laisse de basse mer comporte un plan d'eau mobile qui rend certaines zones accessibles uniquement lors des marées hautes.

Dans l'ensemble la figure 67 montre une densité satisfaisante de points (de 0.2 jusqu'à 0.8 point/m²) sur toute la partie terrestre et le proche estran émergé. Par contre toutes les zones d'interface terre/mer présentent une densité moyenne inférieure de moitié (0.1 contre 0.3) ainsi que des valeurs maximales très éparées. Cette hétérogénéité dans la qualité de l'information sur le milieu intertidal explique les difficultés d'intégration entre les données terre/mer.

Légende :

- Limites_maritimes_LaisBasMer
- Limites_maritimes_LaisHauMer
- Limites_maritimes_traitcote

**Toutes données confondues
densité en pts/10m²**

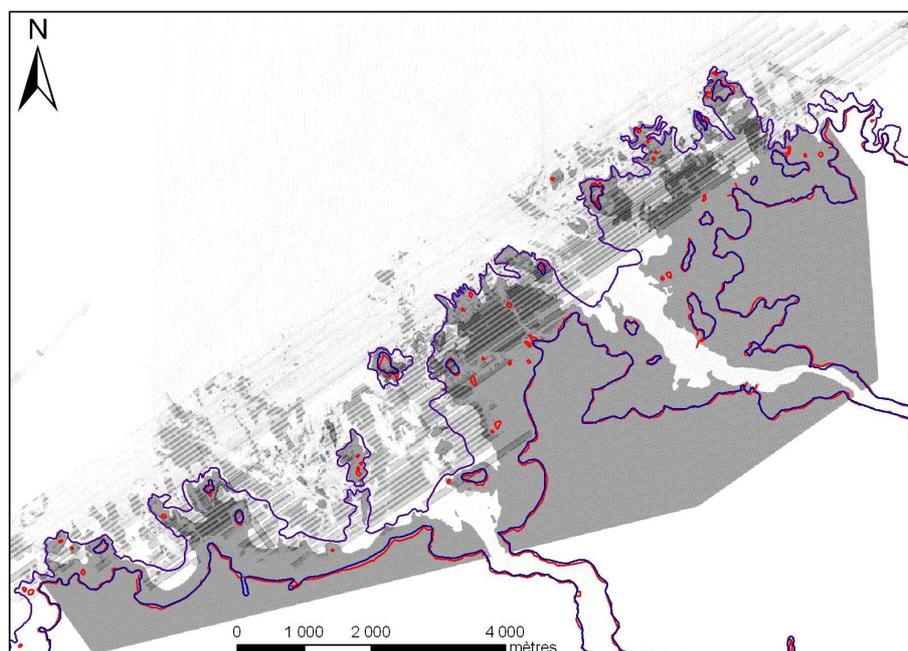
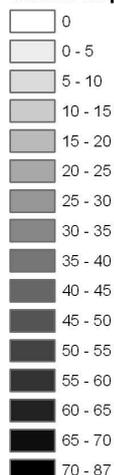


Figure 67 : Carte de densité de l'ensemble des données sur la zone des Abers. Les jeux de données comprennent : la grille LIDAR topographique, le LIDAR topo bathymétrique et les points de sondes.

Les cartes de différences altimétriques soulignent les difficultés liées aux problèmes de conversion entre référentiels altimétriques, aux méthodes d'acquisition utilisées (précision et principes physiques variables selon les milieux) et à l'évolution morphologique au cours du temps.

Les cartes de densité révèlent d'autant mieux les qualités hétérogènes respectives de chaque lot de données selon les milieux considérés.

Ces éléments rajoutent une difficulté supplémentaire au test de la méthodologie définie dans le Chapitre IV ainsi qu'à la nécessité de la valider de façon fiable. En effet en cas de mauvais résultat, il est plus délicat de décider si l'erreur provient d'un mauvais recalage, de l'insuffisance des données ou d'une mauvaise conception de la méthodologie.

Ce constat nous a amenés à envisager l'utilisation de données simulées pour valider la méthodologie et ainsi s'affranchir des problèmes de non recouvrement, de recalage et de densité hétérogène.

Il a fallu pour cela définir les éléments paysagers essentiels à notre méthode d'intégration (Chapitre III) et faire en sorte de les retrouver dans la simulation de paysage littoral.

III.1.3 Données Simulées

La complexité des paysages décrits dans les données réelles, leur hétérogénéité, leurs incohérences dues à des acquisitions temporelles différentes et leur manque de recouvrement amènent à tester les algorithmes en premier lieu sur des données simulées. En effet la simulation de données (terrestres et maritimes) permet de tester les algorithmes implémentés sur des cas plus ou moins simples, où le recouvrement est suffisant, et où le bruit des données est maîtrisé.

Posséder un recouvrement entre les données permet tout d'abord de correctement tester l'appariement entre données homologues en s'assurant de leur présence dans les deux jeux de données. Ensuite il est également possible de s'affranchir du problème des trous et de leur éventuelle interpolation.

A la suite, quantifier et déterminer la nature du bruit introduit dans les données facilite la validation de la méthode pour chaque type d'erreur existante.

III.1.3.1 Mécanisme de production des données simulées

Représentation d'un objet géographique :

Le premier modèle simulé vise à figurer un élément du relief littoral comportant des éléments caractéristiques sur lesquels appuyer l'intégration. Un chenal simplifié (lit d'une rivière ou d'une fleuve toujours présent dans les estuaires) a été choisi et représenté à la fois sur le domaine continental et maritime. Deux grilles régulières de points ont été modélisées (voir figure 68) avec un recouvrement suffisant pour permettre une intégration des MNT.

Simulation de MNT topographique et bathymétrique :

Le second modèle vise à représenter un paysage naturel plus complexe au niveau de ses caractéristiques qu'un simple objet géographique. Un modèle basé sur la génération de relief fractal a donc été produit de manière à être suffisamment représentatif de la variabilité du relief réel. La méthode utilisée pour la génération du MNT est basée sur des algorithmes d'approximation des mouvements browniens [Barnsley et al., 1988]. Selon l'idée que « la forme du relief est indépendante de l'échelle d'observation » (Théorie de Mandelbrot), il est possible de construire des reliefs à l'aide d'un modèle fractal. L'algorithme retenu pour effectuer ces modélisations est la méthode d'approximation d'un mouvement brownien. Cette représentation montre un aspect semblable à celui d'un profil topographique. L'approximation de cette fonction aléatoire continue $X(x, y)$ s'effectue par le calcul d'une transformée de Fourier inverse. Ceci permet de contrôler le niveau de filtrage et donc la complexité du relief [Thibaud, 2002]. Cette méthode permet d'obtenir des reliefs pour lesquels il est possible de paramétrer la profondeur moyenne, l'amplitude des points hauts ainsi que la complexité du paysage représenté (rugosité importante ou non).

A partir de ce relief numérique continu terre/mer, une acquisition de sondeurs multifaisceaux est émulée sur la partie maritime à l'aide d'un logiciel de simulation de données bathymétriques : « SimSmF » [Thibaud, 2002]. La méthode d'échantillonnage du MNT reprend de façon simplifiée les principes de fonctionnement des échosondeurs : précision, densité d'échantillonnage, et influence des mouvements du porteur selon les conditions d'acquisition.

III.1.3.2 Modèles de base

Schématisation d'un chenal :

La figure 68 montre la totalité des points utilisés pour représenter le chenal sur les deux zones terrestre et maritime.

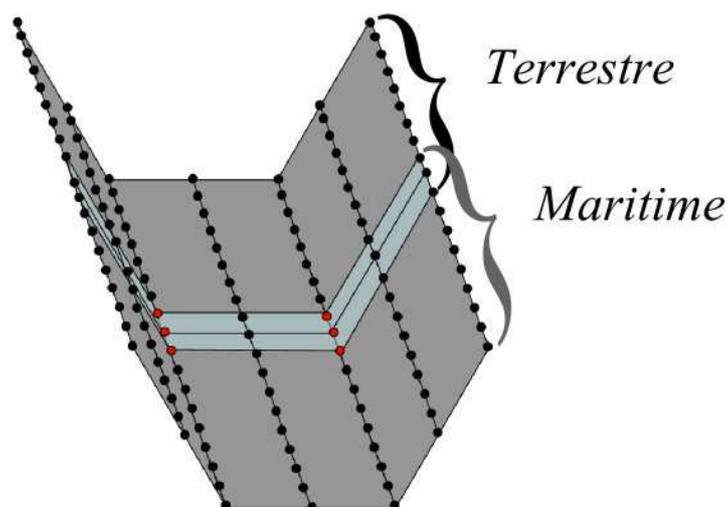


Figure 68 : Illustration d'un chenal en 3D. Les points utilisés pour construire le modèle sont représentés sur le schéma.

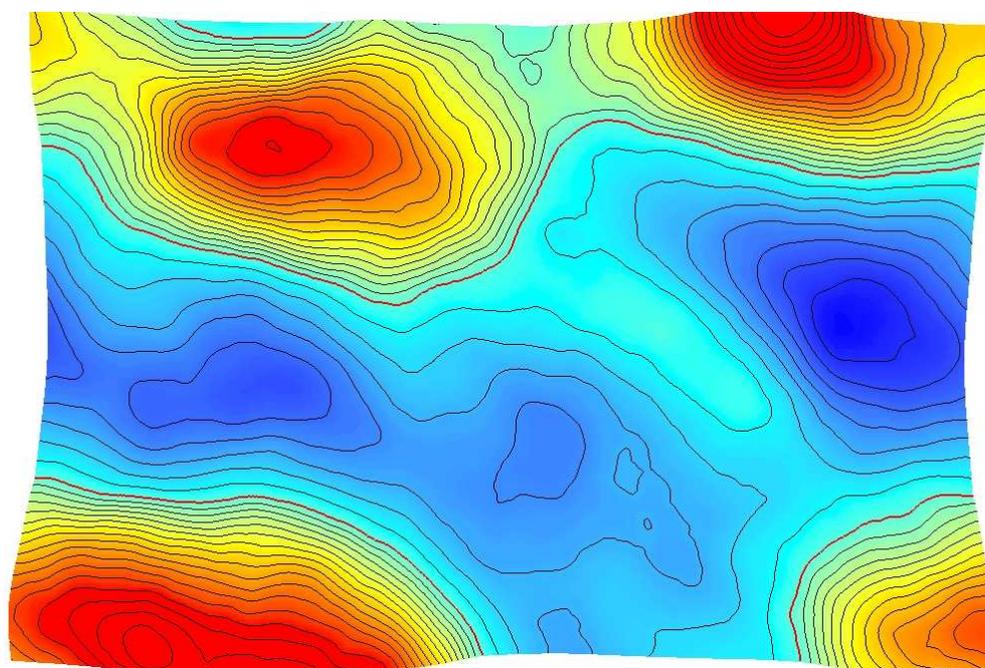


Figure 69 : MNT de référence généré à partir d'un logiciel d'émulation de reliefs fractals. La ligne rouge représente l'isohypse zéro¹⁵ selon laquelle les données seront par la suite subdivisées en modèles terrestre et maritime.

¹⁵ Ligne reliant les points d'égale altitude.

La bande transverse représente l'aire de recouvrement commune aux modèles. Les points de couleur rouge sont les points appartenant aux lignes caractéristiques extraites des MNT (rupture de pente) sur lesquelles les calculs de déformation des modèles sont effectués. Les points ont été interpolés de manière à produire une grille régulière qui sert de modèle de référence. **Les modèles topographiques et bathymétriques** sont des extraits bruités de ce MNT de référence, de manière à produire des lignes géomorphométriques caractéristiques de valeur xyz différentes.

Modèles Numériques de Terrain représentant un paysage littoral :

Plusieurs modèles numériques de terrain ont été créés avec le générateur de reliefs fractals. Le modèle retenu comporte une texture assez simple qui se rapproche de reliefs émoussés d'une plage dunaire par exemple (figure 69).

Les formes élémentaires spécifiques du paysage ont été choisies parallèles entre elles de manière à obtenir par la suite une simulation sonar « réaliste » (qui ne passe pas à travers les sommets hauts par exemple). Par ailleurs une rupture de pente à l'altitude zéro (ligne rouge sur le MNT de référence) a été forcée lors de la simulation du modèle numérique de manière à servir de ligne « élément caractéristique » sur laquelle baser les déformations. A partir de ce modèle de référence un MNT **terrestre** et un MNT **bathymétrique** ont été produits en y ajoutant plusieurs types d'erreurs de mesures présentes dans les données réelles et dues aux mécanismes d'acquisition des MNT.

III.1.3.3 MNT bruités

Le mode de génération de bruit pour les MNT topographiques et bathymétriques ne sont pas les mêmes car il doit reproduire les erreurs présentes dans les données réelles. **Le modèle bathymétrique** a donc utilisé le logiciel SimSmf pour simuler une acquisition réalisée par sondeur multifaisceaux [Thibaud, 2002]. La figure 70 montre une vue 3D d'un des deux modèles obtenus après calcul. Deux acquisitions (nord et sud) sont simulées de manière à rester proche de ce qui aurait été fait dans la réalité pour couvrir une surface similaire.

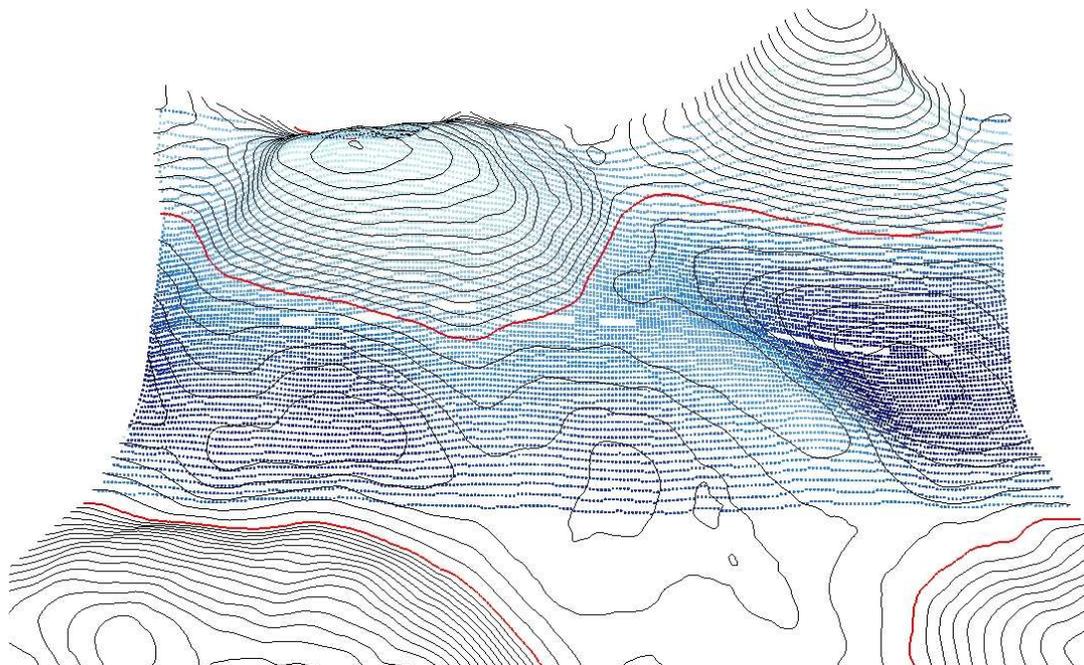


Figure 70 : Levés bathymétriques de la partie Nord obtenus à partir du MNT de référence et du logiciel SimSmf.

Par contre tous les points extraits du modèle de départ sont conservés dans un premier temps, même ceux supérieurs au niveau moyen des mers (ligne rouge sur la figure 70).

Le modèle topographique est déformé à l'aide des trois types de bruit présents dans les données réelles :

- un bruit blanc uniforme ajouté sur l'ensemble des données qui représente les bruits du capteur lié aux principes physiques mis en oeuvre pour l'acquisition ;
- un biais sur l'ensemble du MNT qui représente les erreurs systématiques du capteur ;
- des erreurs locales qui représentent des erreurs grossières accidentelles survenues lors de la mesure.

Le bruit blanc est obtenu à l'aide d'une fonction aléatoire qui ajoute des valeurs comprises entre $[- 0.01 ; + 0.01]$ (intervalle choisi par l'utilisateur), soit une erreur de l'ordre du centimètre pour des données de précision de l'ordre du mètre.

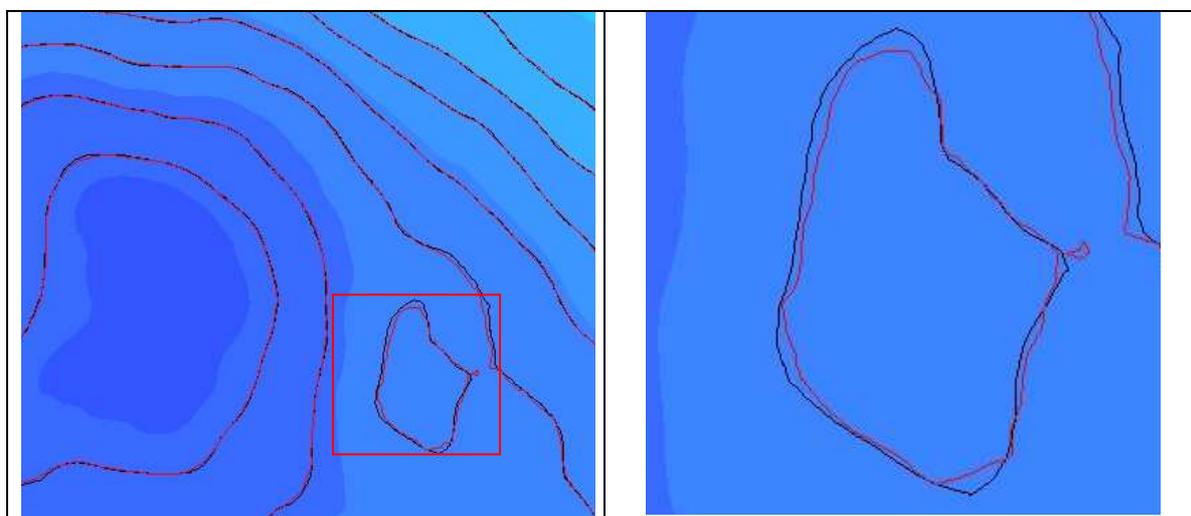


Figure 71 : Visualisation du bruit blanc inséré dans les données par superposition des lignes de niveau de même altitude. Vue de gauche : extrait général d'une zone d'étude ; vue de droite : zoom (cadre rouge) sur une partie des courbes de niveau.

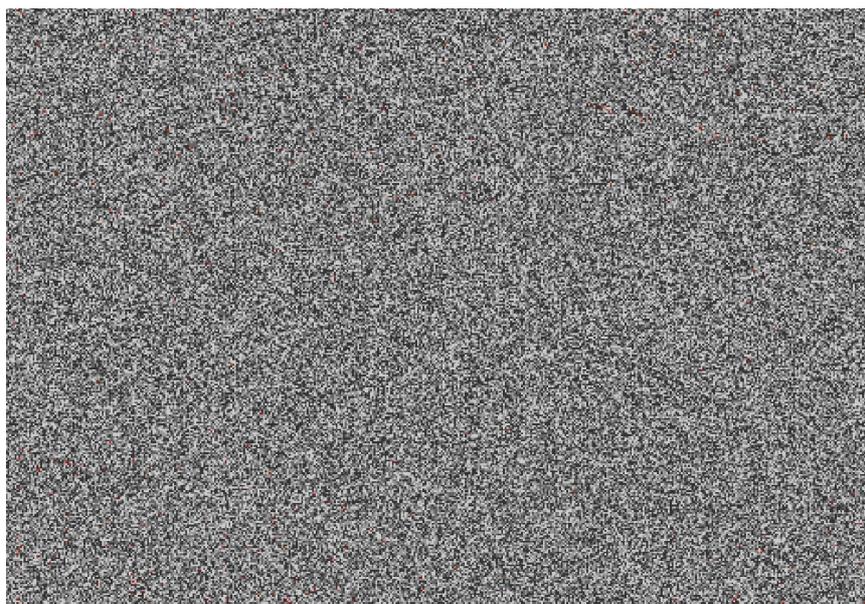


Figure 72 : Résultat de la soustraction effectuée entre le modèle de référence et le modèle bruité avec un bruit blanc.

La figure 71 présente une superposition des courbes de niveau du modèle de référence et du modèle bruité avec un bruit blanc, interpolées tous les 0.5m. La vue de gauche montre peu voire aucune différence entre ces courbes.

Un agrandissement sur la vue de droite permet de distinguer quelques différences minimales entre les tracés. Le résultat de la soustraction pixel à pixel des deux modèles illustré figure 72, montre une grille d'aspect « poivre et sel » caractéristique de ce type de bruit. Les valeurs, comprises entre $[-0.01 ; +0.01]$, se répartissent de manière aléatoire et uniforme.

Une étude parallèle du bruit blanc a été effectuée sur les données LIDAR topographique afin d'évaluer l'intervalle de validité d'un bruit incorporé dans des données simulées. Pour cela un profil topographique est réalisé :

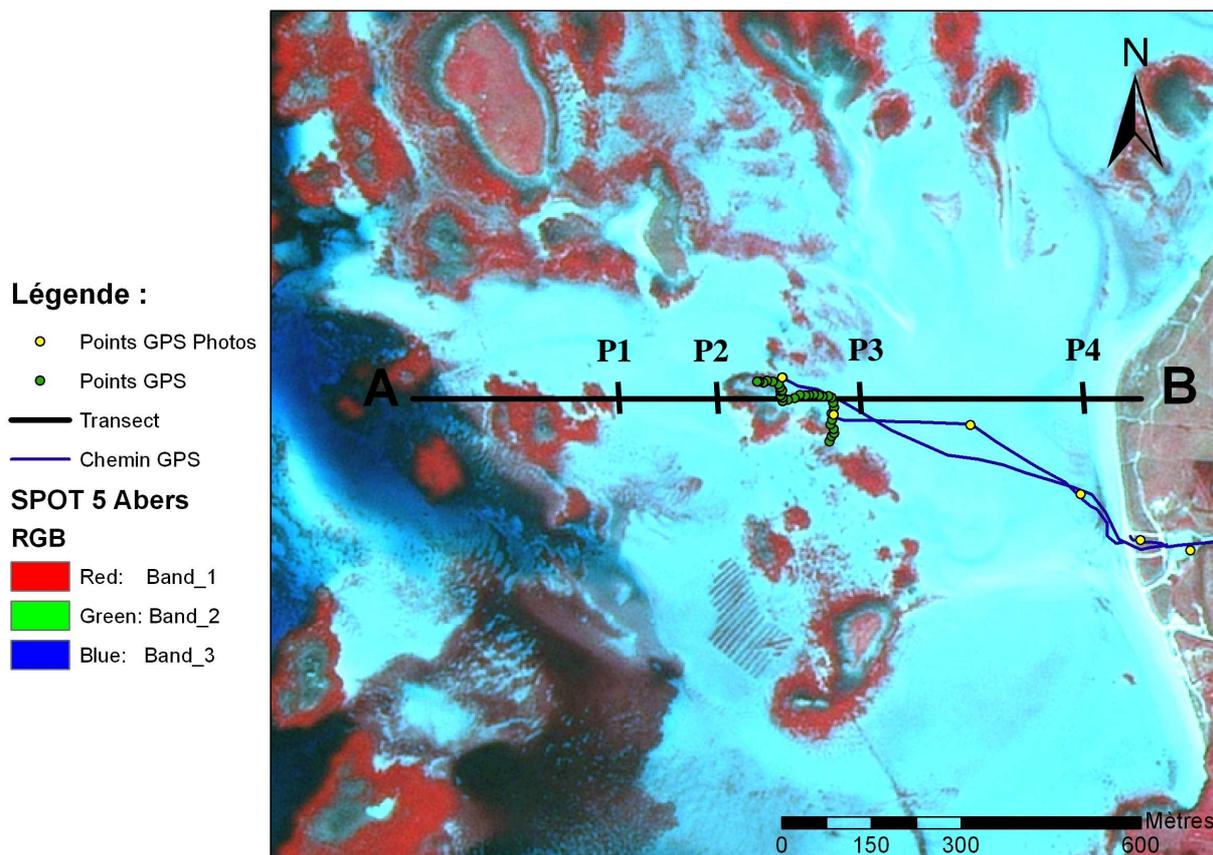


Figure 73 : Localisation du profil topographique effectué sur les données Lidar en vue de l'estimation de l'erreur de bruit blanc.

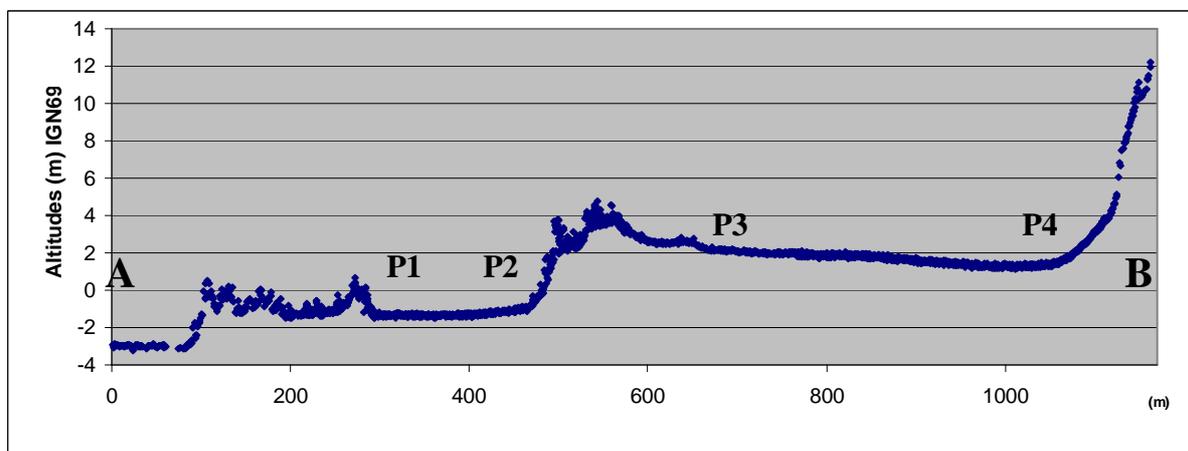


Figure 74 : Profil topographique des points exportés à partir du trait de coupe présenté Figure 73.

Les points Lidar sont sélectionnés selon un buffer de 2m autour de ce trait puis exportées dans Excel (voir figure 74). Deux zones d'étude ont été choisies pour leur faible pente : les points (P1-P2) situés entre 300 et 450m du point A (positionné à 0 sur le profil figure 74) et ceux (P3-P4) situés entre 770 et 920m du profil. La seconde zone comporte une densité double de points car il existe un recouvrement fort entre eux (figure 75).

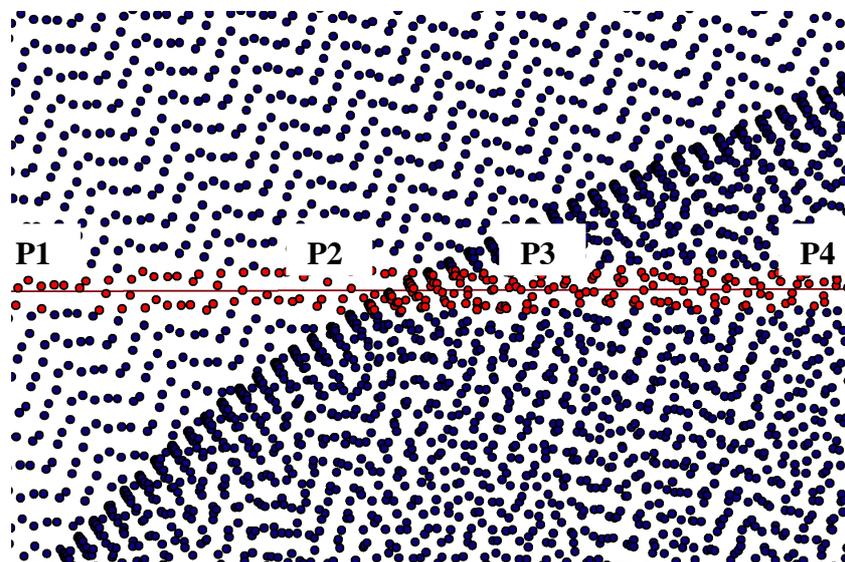


Figure 75 : Visualisation des points Lidar total et ceux situés à 2m du trait de coupe de l'étude.

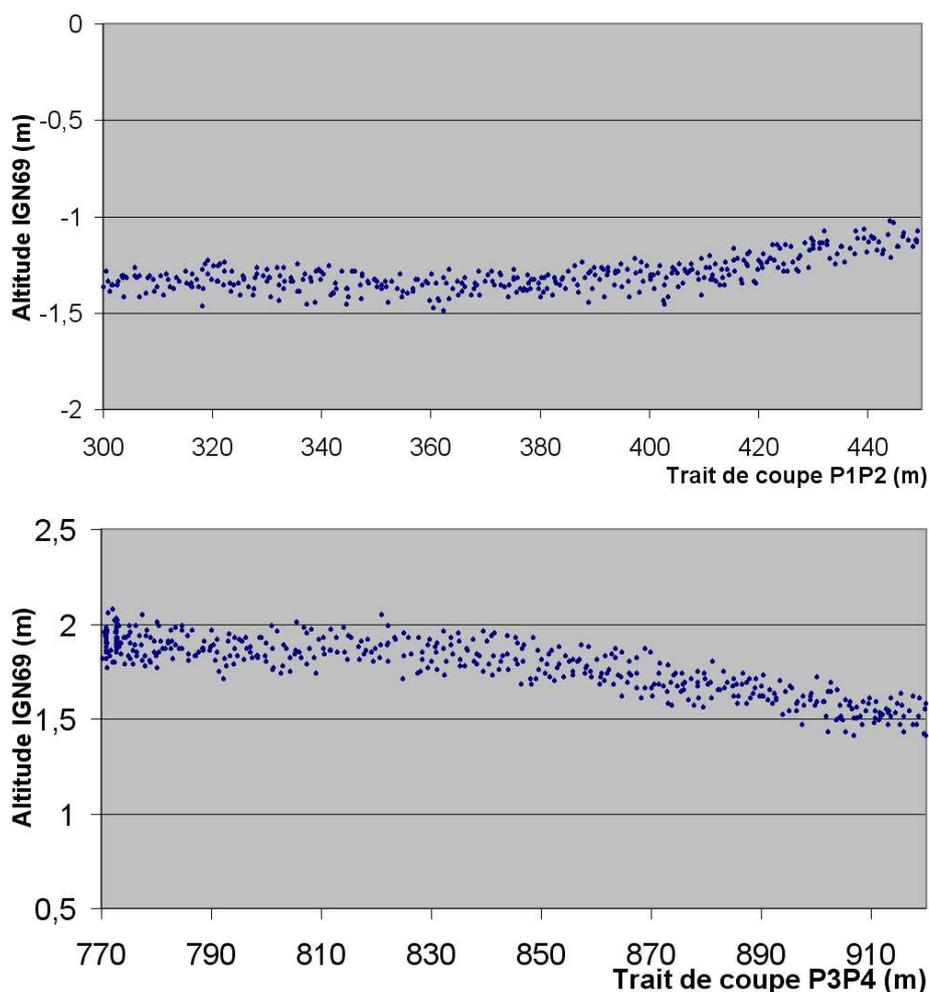


Figure 76 : Répartition des points extraits des traits de coupe P1-P2 et P3-P4.

Les points bleus représentent la donnée totale brute sans nettoyage ni interpolation des points. Les points rouges sont ceux positionnés à 2m du trait de coupe (trait noir sur la figure 73) extraits pour l'étude.

Les données Lidar (de précisions métrique à décimétrique pour ce jeu de données) présente donc une variation altimétrique de l'ordre du décimètre (voir variation des points autour de la valeur moyenne de l'altitude figure 76), que ce soit en domaine de recouvrement ou non des fauchées. La valeur centimétrique du bruit blanc choisi pour les données simulées semble donc sous estimé. Cependant il ne nous est pas paru important d'insister davantage sur un bruit qui n'a que peu d'impact sur la méthode vu son caractère aléatoire et sa répartition uniforme. Les deux autres catégories de bruits, en revanche, ont volontairement été exagérées afin de tester la robustesse de la méthode.

Le second type de bruit incorporé dans les données représentant le domaine terrestre est un **décalage** constant du modèle topographique par rapport au modèle de référence. Ce décalage symbolise les erreurs altimétriques et planimétriques systématiques présentes sur l'ensemble de la fauchée.

Enfin le dernier type d'erreur, **les erreurs grossières**, a été ajouté dans le modèle au moyen de vecteurs localisés. Ces vecteurs sont générés aléatoirement dans le modèle et servent de noyaux de déformation locale des points du MNT de référence.

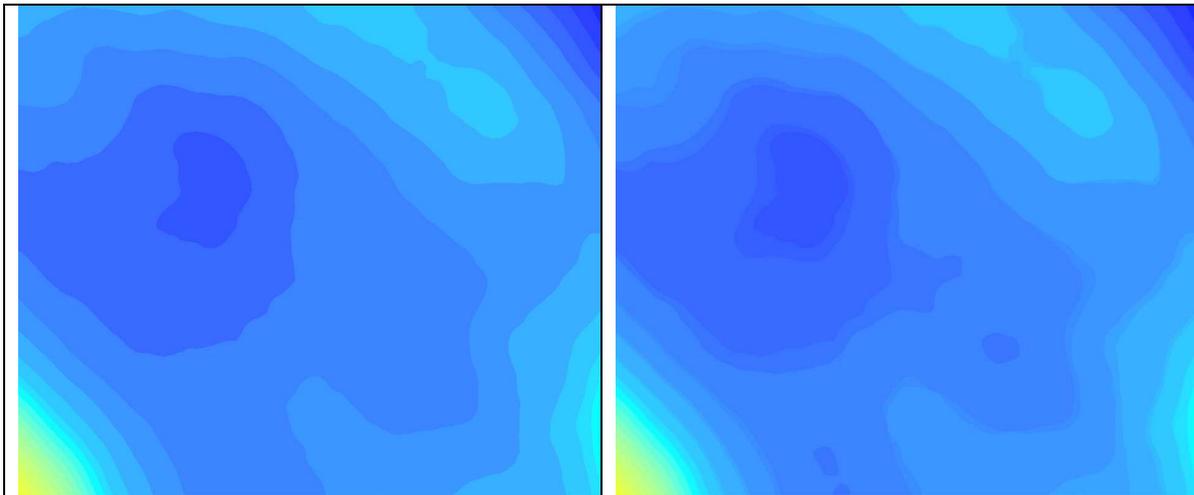


Figure 77 : Vue de gauche : extrait du MNT de référence. Vue de droite : superposition du MNT de référence avec le MNT bruité de manière à faire apparaître la différence de contour d'une même ligne.

La figure 77 montre à gauche un agrandissement du MNT simulé de référence et à droite, le MNT bruité avec les erreurs grossières superposé à ce modèle et mis en transparence. La symbologie choisie est la même de manière à faire correspondre les classes d'altitudes identiques. Sur cette vue, les endroits qui ne comportent pas de vecteur de déformation (en bas à gauche et en haut à droite de la vue de droite), ne présentent aucune différence avec le modèle de départ. Par contre au milieu de la figure, une divergence des lignes de niveau est notable. Les lignes sont déformées de façon circulaire (déformations positive ou négative) et il apparaît également de nouveaux reliefs.

La figure 78 montre la position de ces erreurs dans le modèle de la figure 69, ainsi que leur aspect circulaire, qui est le rayon d'influence du vecteur lors de la déformation. Les zones en dépression sont marquées par des cercles noirs tandis que les zones en surélévation sont en blanc.

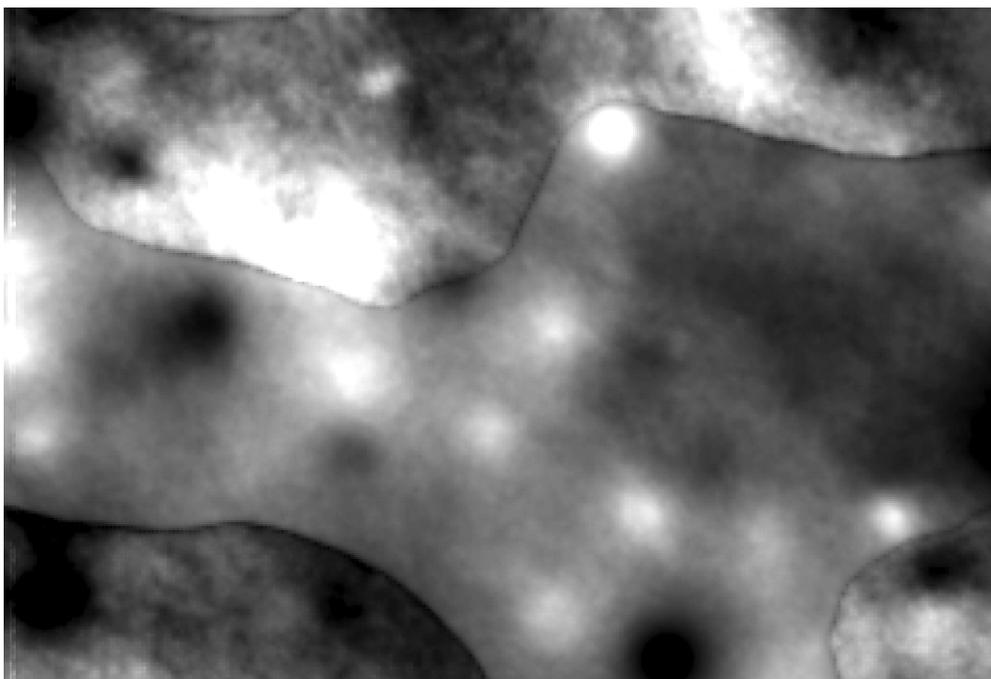


Figure 78 : Visualisation des erreurs grossières ajoutées au modèle de référence.

Nous allons maintenant détailler la méthodologie mise en place pour réaliser l'intégration de ces données ainsi que son partitionnement en phases de manière à résoudre au mieux les problèmes précédemment évoqués.

III.2 Méthode d'intégration globale

La méthodologie d'intégration proposée dans cette thèse (décrite figure 79) se veut générique de manière à traiter n'importe quel type de données, déjà acquises ou futures. Cependant, en s'appuyant sur les lignes caractéristiques contenus dans chaque modèle numérique de terrain, elle vise à conserver des information clé du relief (rupture de pente, ligne de crête, etc...). Elle met en œuvre plusieurs outils dans un ordre précis pour répondre aux besoins posés lors de la fusion (voir Chapitre I) :

- conversion dans un même référentiel,
- correction des biais globaux,
- interpolation des zones sans information,
- détection des objets homologues,
- ou fusion des données appariées.

L'enchaînement de ces traitements s'organise en trois phases principales (voir figure 79) :

- une étape de **comparaison** de l'ensemble **des MNT** qui comprend : une conversion dans un même référentiel géographique, une étape de partition des MNT et une étape de corrélation et correction des biais ;
- une **segmentation** qui effectue une extraction de lignes caractéristiques ;
- et enfin **l'intégration du MNT** avec tout d'abord l'interpolation des trous, ensuite l'appariement, la fusion puis finalement les corrections élastiques.

Les trois parties suivantes s'attachent donc à décrire en détail chacune de ces phases, les traitements réalisés sur les données et leur apport vis-à-vis des méthodes d'intégration déjà existantes.

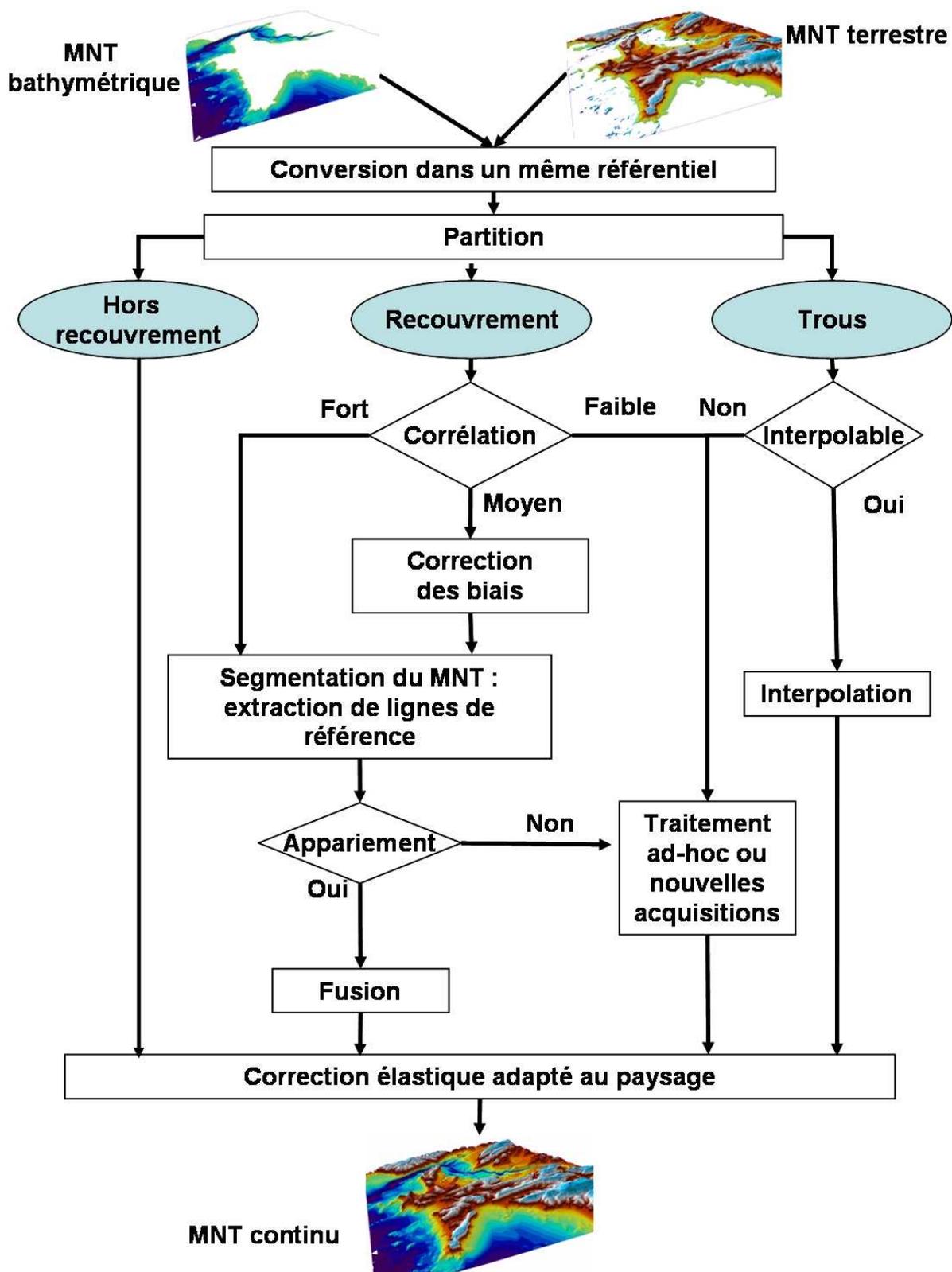


Figure 79 : Représentation de l'ensemble de la méthodologie d'intégration.

III.2.1 Comparaison de MNT

La première phase de comparaison de MNT a pour but d'évaluer la qualité respective des données à intégrer. La partie « Evaluation de la cohérence des données géographiques » a montré qu'il est rare d'avoir une information homogène, cohérente et également répartie sur

toute la zone d'étude. Les données sont, en effet, couramment dans des systèmes géographiques différents selon la précision de positionnement recherchée, de résolutions différentes selon les systèmes d'acquisition, et de recouvrement souvent inégal.

Cette phase vise donc tout d'abord à s'assurer que toutes les données sont bien dans le même référentiel (local ou global) (étape « transformation de référentiels » figure 80). Ensuite, un tri est effectué sur le modèle : un recouvrement total entre données n'existant pas dans ce contexte, toutes les zones des MNT ne sont pas à traiter (étape « partition »). Enfin, la dernière étape s'assure que toutes les transformations effectuées garantissent une cohérence suffisante entre les données pour effectuer la deuxième phase de la méthodologie (étape « corrélation »).

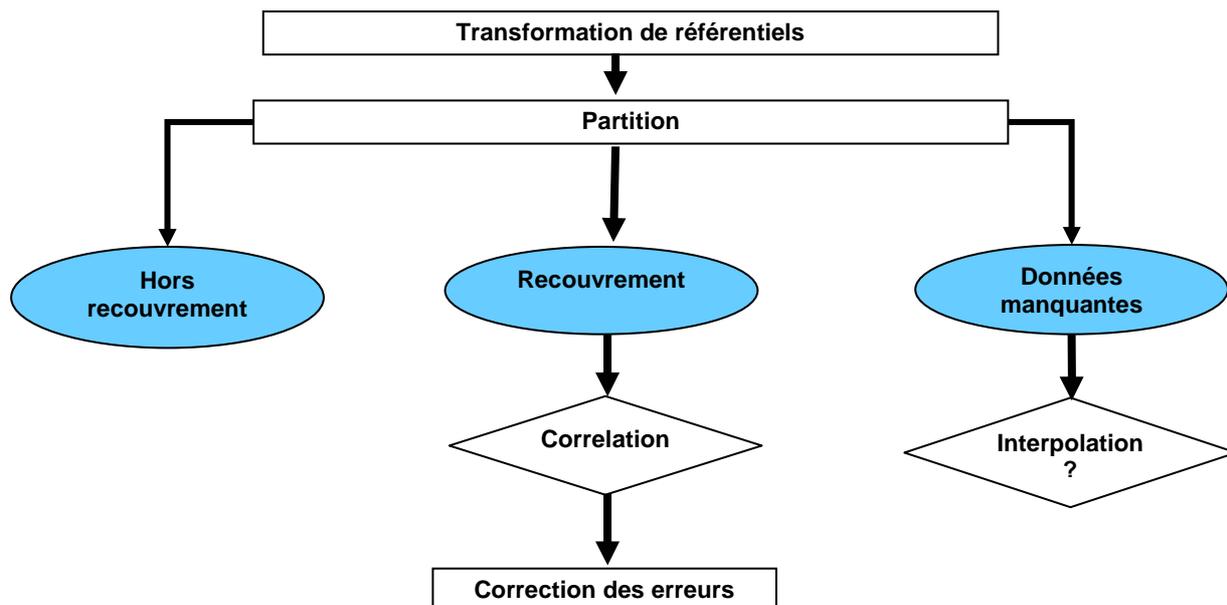


Figure 80 : Détail de la première étape : comparaison des MNT et analyse de corrélation.

Chacune de ces étapes vont maintenant être présentées plus en détail.

III.2.1.1 Conversion dans un même référentiel cartographique

La première étape convertit les deux modèles numériques dans un même référentiel cartographique choisi par l'utilisateur. Un des deux systèmes géodésiques sert de base pour transformer le second MNT dans le même référentiel. Ces systèmes peuvent être soit globaux, soit locaux et cette conversion peut s'effectuer au moyen d'un SIG classique ou d'un logiciel développé spécifiquement. Même si la conversion n'est pas extrêmement précise, les écarts entre les deux modèles sont ainsi minimisés. Ainsi en définissant les paramètres locaux, il est possible de convertir les données historiques du SHOM avec un positionnement suffisant pour permettre une comparaison avec un modèle plus récent. Les deux MNT sont également convertis dans un même format (TIN, grille régulière, nuage de points) de manière à faciliter les traitements ultérieurs.

III.2.1.2 Partition en domaine de recouvrement

La seconde étape concerne la partition des modèles selon trois types de zones :

- les zones géographiques communes aux deux MNT ;
- les zones sans information (trous) ;

- et les zones où il n'y a pas de recouvrement et qui n'ont pas besoin d'être considérées lors du traitement.

Les zones sans information sont pour leur part traitées ultérieurement manuellement (ou automatiquement) par interpolation selon leur étendue et la possibilité ou non d'enrichissement.

III.2.1.3 Analyse de corrélation

Sur les zones de recouvrement, des coefficients de corrélation similaires à ceux définis dans [Guth, 2006] sont employés pour vérifier l'importance des décalages planimétriques et altimétriques entre les deux MNT. Les zones de faibles coefficients sont considérées comme trop incohérentes pour être traitées directement. Elles nécessitent au préalable de nouvelles acquisitions de données ou un traitement plus spécifique. En ce qui concerne les zones de corrélation moyenne, les jeux de données sont corrigés des biais planimétriques et altimétriques. Les zones avec une cohérence élevée sont directement traitées en vue d'une intégration.

III.2.2 Segmentation et extraction de lignes caractéristiques

La seconde phase de la méthodologie concerne une segmentation effectuée sur les zones de recouvrement entre modèles. Ces traitements se réalisent sur les zones des MNT qui ont montré une bonne corrélation suite à la première phase et qui ne comportent pas ou peu de trous.

Cette étape effectue une classification du MNT en zones homogènes (segmentation) et l'extraction pour chaque zone de lignes caractéristiques propres sur lesquelles se base l'appariement (extraction).

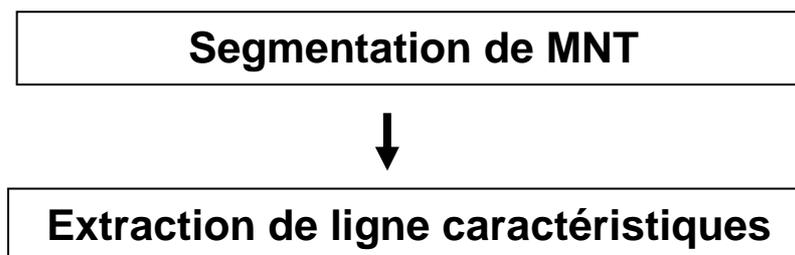


Figure 81 : Détail de la seconde phase : segmentation en vue de l'extraction des lignes nécessaires au processus d'appariement.

III.2.2.1 Segmentation

Le principe de la segmentation est de réaliser une partition d'une surface en régions homogènes en accord avec un ou plusieurs critères [Cocquerez et al., 1995]. Ceci peut se réaliser de plusieurs manières :

- soit en caractérisant les formes paysagères au travers de leur « hétérogénéité » au sein du modèle. Il s'agit alors de détecter des contours ou des éléments précis au moyen de termes manipulables par l'ordinateur comme les mesures géomorphométriques (méthode **descendante**) ;

- soit en regroupant des pixels (ou des groupes de pixels) selon des critères d'homogénéité de façon à les agréger pour ensuite reconnaître des formes paysagères prédéfinies (méthode **ascendante**) ;
- soit en cherchant des **limites de zones** (lieux de ruptures) et ensuite essayer de structurer ces ensembles de limites.

La méthode de segmentation la plus fréquemment utilisée dans la littérature est une approche **ascendante** basée sur la détection de zones homogènes au travers de formes caractéristiques ([MacMillan & Shary, 2008], [Bonin et al., 2006], [Saux et al., 2004], [Pike, 2000], [Monier, 1997], [Wood, 1996], [Dikau, 1989], [Pike, 1988], [Evans, 1980], [Peucker & Douglas, 1975]). Certains auteurs ([Wood, 1996], [Evans, 1980], [Peucker & Douglas, 1975]), proposent de classer ces formes selon six primitives de base : le plan, le chenal, la crête, le col, le sommet et la cuvette (cf III.2.3).

Wood est le premier à mettre au point un logiciel de segmentation de paysage basé sur ces primitives : « Landserf » ([Wood, 2002]). La première version souffre en revanche d'être très sensible à la résolution du MNT. Pour remédier à ce problème, Fischer et Wood ont proposé de se baser sur de la logique floue, ce qui autorise également une analyse multi-échelle [Fisher et al., 2004].

D'autres méthodes ont également été implémentées à partir de primitives différentes. [Saux et al., 2004] décompose une surface donnée selon un attribut qui combine les propriétés de la pente et de la déviation. [Monier, 1997] s'appuie sur une information composée des réseaux de crêtes et de thalwegs extraits du MNT et des modèles surfaciques de pente et de dureté.

En reprenant les travaux de Wood, le COGIT ([Bonin & Poupeau, 2005, Bonin et al., 2006]) a implémenté un logiciel de segmentation sur la plate-forme de développement « Oxygène ». L'application est destinée spécifiquement aux modèles TIN. L'algorithme se base sur les lignes de contour pour détecter cinq formes paysagère élémentaires : sommets, thalwegs, pentes, collines et bassins. Le TIN est alors découpé en aires de tendance mais toutes les zones ne sont pas obligatoirement classées.

D'autres SIG, référencés dans [Hengl & Reuter, 2008], proposent également des méthodes de segmentation de MNT ou d'analyse géomorphométrique à partir de paramètres paysagers différents.

La segmentation basée sur la recherche de limites de zones ([Schneider & Wood, 2004]) décompose le MNT en bassins, interfluves, thalwegs et cols, ce qui exploite quatre des six formes géomorphométriques. L'algorithme le plus souvent utilisé est l'interpolation bilinéaire de Morse-Smale ([Magillo et al., 2007], [Danovaro et al., 2007], [Comic et al., 2005]) qui présente l'avantage d'être déterministe et donc d'obtenir des résultats propres à chaque jeu de données. Néanmoins la segmentation obtenue est dépendante de la précision du modèle de départ, ce qui pose problème dans le cas de structures trop petites pour être détectées.

Dans la méthodologie telle qu'elle est pensée, la méthode de segmentation appliquée permet uniquement d'extraire les lignes caractéristiques des zones homogènes obtenues et de comparer ces jeux de lignes entre eux.

III.2.2.2 Extraction de lignes caractéristiques

La détection des éléments caractéristiques par zones identifiées doit être affinée après la segmentation. Ceci peut s'effectuer de diverses manières. Par manque d'outils SIG adaptés à nos attentes, l'extraction de lignes caractéristiques s'est réalisée pour ce travail par reconnaissance visuelle.

III.2.3 Intégration des modèles de terrain

Cette étape traite du processus d'intégration effectué sur les modèles à fusionner lorsque la segmentation est terminée. Les lignes caractéristiques issues des zones à forte corrélation sont comparées en vue d'un appariement tandis que les zones trop incohérentes nécessitent de nouvelles acquisitions ou des traitements spécifiques. A l'issue de l'appariement, les zones concernées sont déformées selon un processus de correction choisi par l'utilisateur et le modèle intégré final est produit.

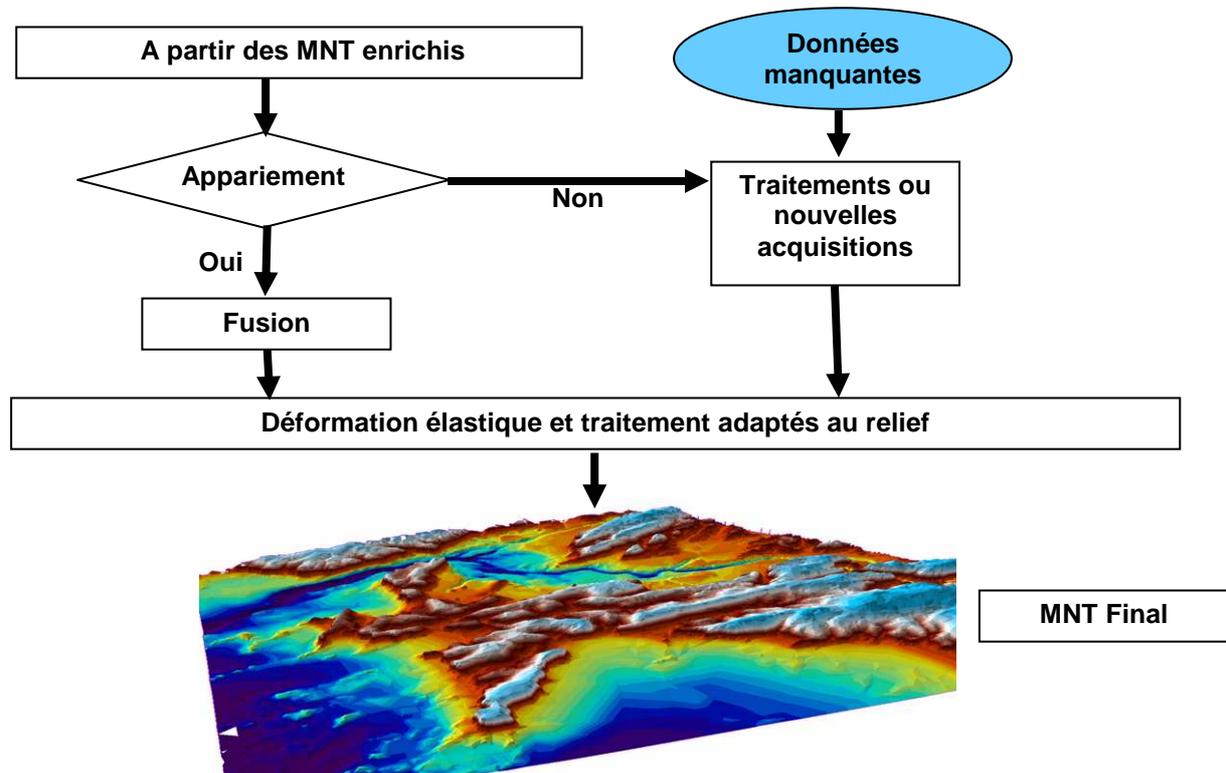


Figure 82 : Détail de la troisième et dernière étape : appariement, fusion et déformation élastique.

III.2.3.1 Interpolation des trous

Une zone sans information peut être considérée comme « interpolable » en fonction des critères suivants :

- si la superficie concernée est de taille *raisonnable* (à définir par l'utilisateur selon la précision des besoins et les spécifications) ;
- si le paysage est homogène : une lacune située au milieu d'une zone de plage peut être interpolée. En revanche l'interpolation d'une lacune située entre une limite de falaise et de platier rocheux présente des risques : le changement de paysage ;
- si des éléments caractéristiques morphométriques sont présents de part et d'autre ;
- si un enrichissement est accessible ;
- si la précision attendue pour l'application le permet.

Dans le cas d'un trou de petite taille par exemple, situé entre deux zones de pente homogène, si la phase d'enrichissement détecte un chenal de part et d'autre, une interpolation de l'information peut être effectuée en tenant compte de cet élément caractéristique.

En complément, des méta-données portant sur la ou les zones interpolées, peuvent être établies de manière à estimer la qualité de l'information en chaque zone du MNT final intégré.

Dans le cadre de l'application de la méthodologie à des données réelles, une étude a été menée par les élèves de l'école navale lors d'un PVA (Projet de Voie d'Approfondissement) pour implémenter un outil de détection automatique de zones de lacunes entre MNT ([Kervazo & Moutoulatchimy, 2006]). L'addon est un module compatible avec le logiciel ArcMap qui permet d'obtenir un raster composé uniquement de deux valeurs :

- 1 si il y a recouvrement
- 0 si il y a lacune.

Suite à de nombreuses difficultés, l'outil n'a pu être exploité sur les données réelles présentes sur les zones d'étude qui se révélaient trop complexes tant au niveau de leur structure que de leur facilité d'exploitation. Par ailleurs ce module ne propose pas d'interpoler ces lacunes en tenant compte des éléments caractéristiques mis en évidence lors de l'enrichissement car de nombreux travaux de recherche sont encore nécessaires.

III.2.3.2 Appariement

Le processus d'appariement consiste à détecter dans la zone de recouvrement des MNT les éléments morphologiques linéaires homologues (crêtes, thalwegs, réseaux hydrographiques) pour lesquels des points homologues ont été identifiés. Dans l'exemple de la figure 83, les réseaux hydrographiques (représentés en (a)) sont reconnus comme homologues si la distance de Fréchet partielle/partielle calculée est inférieure à un seuil défini selon la précision des données [Mascret et al., 2006]. L'appariement de ces deux lignes compare chacun des points et détermine ceux qui sont homologues (trait noirs fin, figure 83 (b)). Pour leur fusion, il est possible soit de garder une moyenne des deux lignes (figure 83 (c) et résultat final en (d)), [Devoegele, 2002], soit d'en privilégier une et de déplacer l'autre.

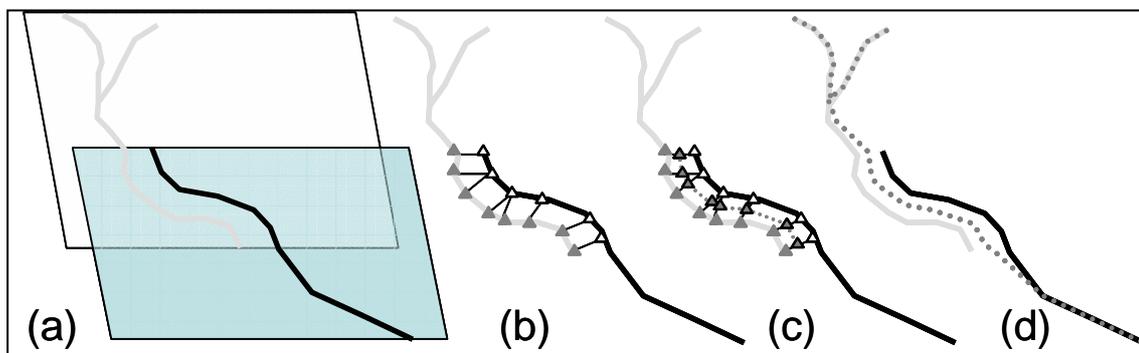


Figure 83 : Schéma d'intégration de deux MNT. (a) : enrichissement, (b) : appariement, (c) : fusion, (d) : déformations élastiques.

III.2.3.3 Déformation élastique et Fusion de MNT

Il existe différentes sortes de processus de déformation de données :

- déformation rigide ;
- déformation élastique.

La déformation « rigide » consiste essentiellement en une déformation selon un vecteur constant plus une rotation : utilisée seule, elle se trouve inadaptée à notre problématique d'intégration car elle sous-entend une déformation constante sur la totalité du modèle. Dans ce cas elle se retrouve également très sensible aux bruits inhérents aux MNT. Par ailleurs elle ne permet pas un appariement partiel de nos données.

La déformation élastique ([Langlois, 1994], [Laurini, 1998]), quant à elle, corrige localement les données via une bande de déformation élastique. Le rayon de déformation de la bande doit être choisi avec soin sous peine de voir apparaître des artefacts induits par la correction.

Une solution optimale est de proposer une combinaison de ces deux types de déformation : un premier algorithme (« TRANS ») estime la présence ou non d'un biais global homogène dans les données, tandis qu'un second algorithme se charge des déformations locales via un rayon d'action.

Le biais global est corrigé uniquement à l'aide d'une fonction de translation et non de rotation. Nous considérons en effet que le recalage des données et la conversion des systèmes géographiques dans un référentiel commun effectué auparavant doit traiter ce problème.

La correction peut s'effectuer de deux manières :

- soit il y a conservation d'une donnée considérée meilleure et qui sert alors de référentiel absolu ;
- soit il y a une fusion pondérée des deux données en considérant qu'elles peuvent avoir une qualité qui varie de l'équivalent (poids de 50-50) au très différent (poids de 90-10 par exemple).

Les éléments morphologiques des modèles sont alors corrigés au moyen d'une déformation élastique [Devoegele, 2000] globale qui n'est autre qu'une fonction inverse de la distance au carré de formule :

$$\overrightarrow{p_i pte_i} = \sum_{k=1 \dots n} \left(\overrightarrow{Cph_{k,i} pf_k} \times \left(\frac{d_E(p_i, Cph_{k,i})^{-2}}{\sum_{l=1 \dots n} d_E(p_i, Cph_{k,i})^{-2}} \right) \right)$$

où k est le nombre de couples de points homologues ($Cph_{k,1}$, $Cph_{k,2}$). $Cph_{k,1}$ est le point de L_1 qui correspond à un point $Cph_{k,2}$ de L_2 .

Le vecteur obtenu est basé sur une moyenne pondérée des vecteurs de déplacement calculés entre les couples de points homologues issus du chemin minimal ($Cph_{k,i}$). Le nouveau point p_i déformé est déplacé selon ce vecteur pondéré de manière inversement proportionnelle à la distance entre p_i et les $Cph_{k,i}$.

En complément, un coefficient d'amortissement est utilisé afin d'adapter la correction effectuée sur les points sans homologue contenu dans la bande élastique. La valeur de ce coefficient est variable en fonction de l'éloignement du point considéré vis-à-vis du couple de points homologues voisin et de son vecteur de déformation. Si le point à corriger est très proche du point p_i la déformation appliquée est presque maximale. Si au contraire il est à la limite du rayon d'influence, la correction s'apparente à zéro.

Ce coefficient est calculé à partir de la formule :

$$coef_i = \left(\cos \frac{d_E(p, p_i) \times \pi}{ri} + 1 \right) / 2 \text{ si } d_E(p, p_i) \leq RI, 0 \text{ sinon.}$$

La figure 84 représente la variation de l'impact de ce coefficient en fonction de la distance entre le point à déformer et p_i sur un rayon d'influence de quatre.

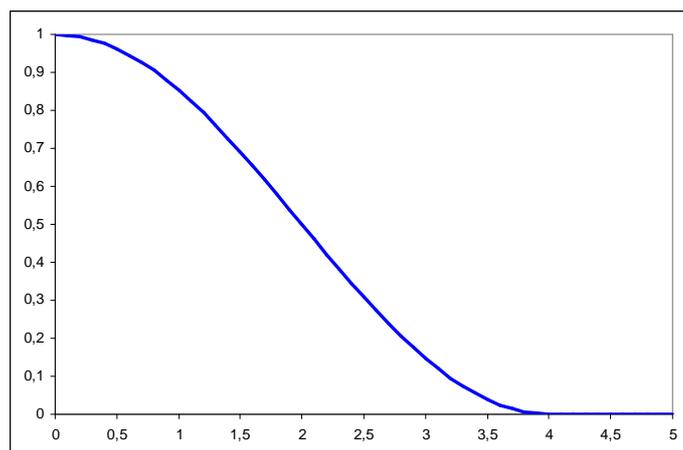


Figure 84 : Courbe de variation du coefficient d'amortissement entre le point p_i et l'origine d'un vecteur de déformation sur un rayon d'influence égal à 4.

Le point considéré est alors déplacé d'un vecteur \vec{v} qui varie en fonction de la position du point d'origine p_i . Ce vecteur \vec{v} est calculé à l'aide d'une somme pondérée des \vec{v}_{li} . Le poids associé à chaque \vec{v}_{li} est fonction du coefficient d'amortissement $coef_i$. La formule pour calculer \vec{v} est :

$$\vec{v} = \sum_{i=1..n} \left(\vec{v}_{li} \times \frac{coef_i}{\sum_{j=1..n} coef_j} \right)$$

Les MNT sont ensuite fusionnés en déplaçant le nuage de points selon les caractéristiques mises en valeur lors de l'enrichissement : respect des lignes de forces (crêtes, thalwegs, ruptures de pente), points côté, rugosité, pente. Pour cela, des méthodes de mise en cohérence du relief avec les éléments caractéristiques [Rousseau, 04] et de déplacement élastique sont employées. Elles seront sélectionnées et paramétrées en fonction du type paysager. La figure 83 illustre la manière dont sont déformés les modèles : après extraction d'un élément caractéristique (symbolisé ici par un chenal : figure 83 (a)), les points des deux éléments sont appariés (figure 83 (b)). Par la suite ces points sont déformés élastiquement avec un coefficient d'amortissement de 0.5 (dans le cas de MNT de même qualité, figure 83 (c)) afin d'obtenir une ligne finale intermédiaire entre les deux lignes d'origine (figure 83 (d)). Tous les points alentour appartenant à chacun des deux modèles mais non aux lignes reconnues comme caractéristiques, sont déformés avec ce même coefficient et selon le même vecteur appliqué sur les points homologues.

III.3 Validation de la méthodologie sur les données simulées

III.3.1 Modèle du chenal

Une première validation est effectuée sur la simulation simplifiée d'un chenal. Deux parties maritime et terrestre (voir figure 85, étape 2) sont extraites du modèle de référence puis bruitées (étape 3 de la figure 85). Les différentes variétés de bruits présentées section V.1.3.3 sont incorporées dans le modèle topographique avec les valeurs suivantes :

- bruit uniforme blanc, compris entre $\pm 0.01\text{m}$;

- un biais de 0.6m en x, 0.6m en y et de 0.8m en z ;
- des erreurs grossières avec une déformation maximale engendrée par un vecteur défini par 0.8m en x, 0.8m en y, et 0.5m en z.

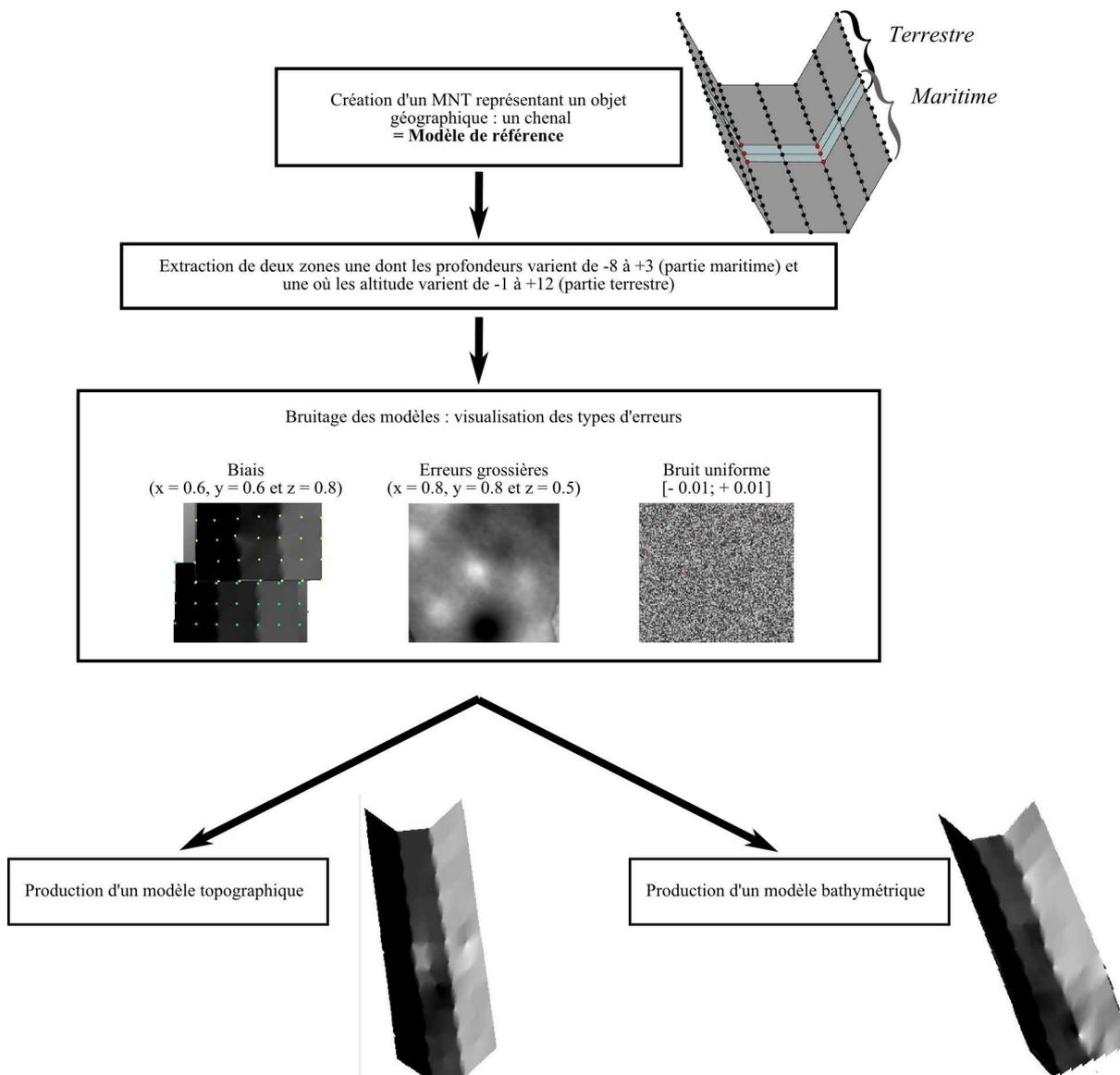


Figure 85 : Etape de production des MNT simulés simples pour la validation de la méthodologie.

Les deux modèles extraits possèdent un domaine de recouvrement complet et suffisant pour permettre une extraction d'éléments caractéristiques appartenant au chenal. Les éléments caractéristiques du chenal pour cette simulation sont les deux lignes de rupture de pente du chenal visibles sur la figure 85, étape 1.

La figure 86 affiche le modèle d'origine (représenté en estompage en **(a)**), les points issus des modèles bathymétrique en bleu et topographique en jaune (modèle **(b)**), leur estompage en **(c)** et une interpolation de ces points (figure 86 **(d)**). Sans correction planimétrique préalable, le modèle **(d)** présente un fort décalage au niveau des lignes de rupture de pente après la fusion des points et leur interpolation en grille continue. De même, l'effet de vagues irrégulières qui se détachent au niveau de la zone de recouvrement est induit par la non correction altimétrique des points bathymétriques (bleus) par rapport aux points topographiques (jaunes).

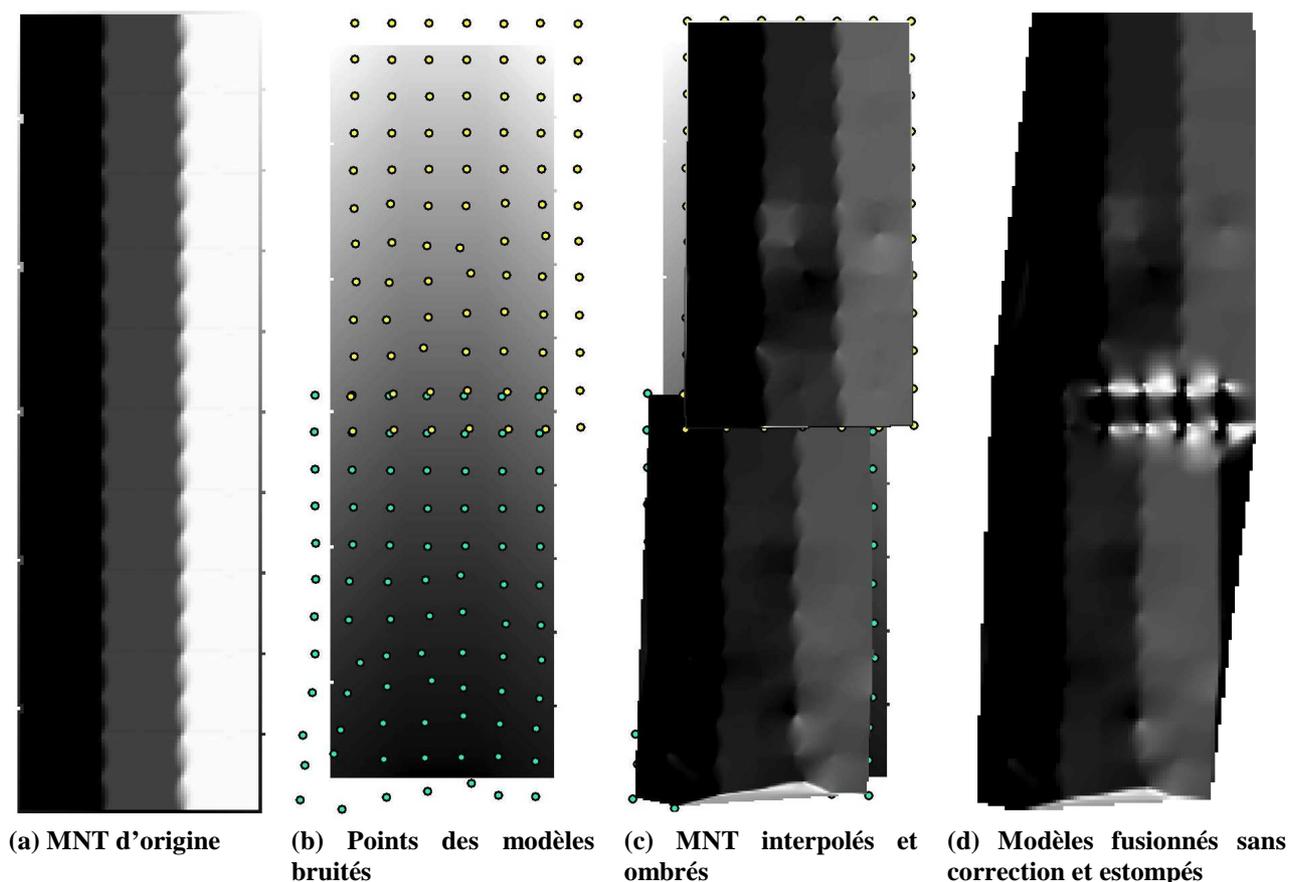


Figure 86 : Illustration du modèle d'origine en (a), des modèles topographique et bathymétrique dérivés (sous forme de points en (b) et sous forme estompée en (c)) et des points bruités interpolés puis estompés en (d).

Afin de corriger ces erreurs altimétriques et planimétriques mises en évidence sur le modèle (d), les deux lignes de rupture de pente sont donc extraites des modèles bruités topographique et bathymétrique. L'écart maximal entre ces deux lignes est calculé grâce à la distance de Fréchet (présentée chapitre II). Cette mesure est utilisée par la suite pour le calcul de la déformation élastique. Pour cette dernière étape, plusieurs méthodologies de déformation élastique ont été testées de façon à comparer les résultats et déterminer la meilleure approche vis-à-vis des MNT à fusionner.

Une première méthode réalise uniquement une déformation élastique :

- soit sur l'**ensemble des modèles**, appelée *déformation globale* : **DEG** ;
- soit sur un **rayon choisi** par l'utilisateur : *déformation locale*, **DEL32** si le rayon est égal à 32 par exemple.

Une seconde méthode calcule tout d'abord le **biais moyen** entre les données appariées et effectue avant toute déformation élastique (globale, DEG, ou locale, DEL), une *translation* préalable notée **TRANS** (ou **T**). Le biais de cette translation est calculé grâce aux couples de points homologues issus de l'appariement des lignes caractéristiques extraites des MNT. Il est exprimé sous la forme d'un vecteur moyen (direction, sens et longueur) qui est appliqué à l'ensemble des deux modèles (sous forme de nuages de points) à déformer.

Afin d'évaluer l'impact d'une translation sur le modèle fusionné final, une étude a tout d'abord été réalisée séparément sur le modèle bathymétrique (figure 87). Deux déformations sont étudiées :

- déformation élastique locale appliquée aux points compris dans un rayon de 8 (DEL8) ;

- déformation élastique locale appliquée aux points compris dans un rayon de 32 (DEL32).

La déformation élastique globale n'est pas illustrée car elle engendre une correction de toutes les données du modèle avec un poids plus ou moins important selon leur proximité avec la bande de recouvrement. C'est pourquoi elle est étudiée dans son ensemble dans le paragraphe suivant lors de la déformation et de la fusion des deux modèles.

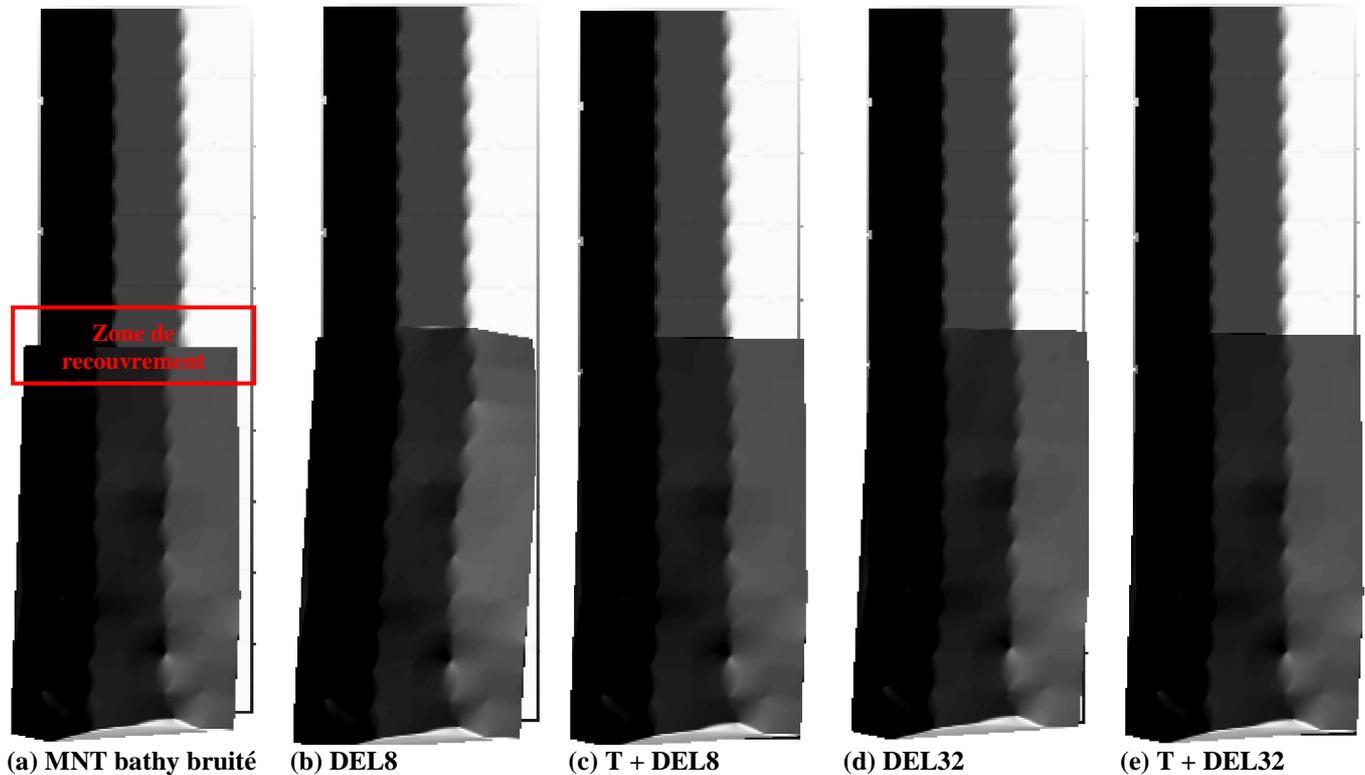


Figure 87 : Illustrations de l'impact des méthodes de déformation par translation vis-à-vis d'une méthode de déformation élastique sur le modèle bathymétrique. Les rayons choisis sont égaux à 8 et 32 dans cet exemple.

Une représentation à l'aide d'estompages des modèles étudiés a été préférée pour cette étude de manière à avoir accès visuellement aux modifications de la surface par l'ajout des erreurs ou par l'application des déformations. L'objectif est d'avoir une estimation des modifications entraînées par les différents types de traitements et d'évaluer leur pertinence vis-à-vis de la problématique d'obtention d'un modèle continu sans artéfacts de fusion.

Le bruit ajouté aux deux modèles étudiés étant similaire, la comparaison des différentes méthodes de déformation élastique ne va se faire uniquement que sur le MNT bathymétrique, les répercussions étant exactement identiques sur le modèle topographique.

La figure 87 **a**) présente le modèle d'origine auquel a été superposé le modèle bathymétrique bruité. L'aspect uniforme des trois teintes noire/grise/blanche du MNT d'origine est dû à l'absence de texture sur ce modèle non bruité. La partie bathymétrique qui lui est superposée présente par contre de fortes variations : un décalage constant de tout le MNT en raison du biais global et des erreurs locales visibles sous la forme de cuvettes ou de sommets. La figure 87 **b**) représente le modèle bathymétrique déformé élastiquement localement avec une bande de déformation de rayon égal à 8. La figure 87 **c**) est ce même modèle avec la même valeur de rayon de déformation mais ayant subi une translation préalable. De même la figure 87 **d**) est le modèle déformé avec un rayon de 32 sans translation tandis qu'en **e**) la translation est effectuée. La dépression importante située en bas à droite de tous les modèles est une erreur grossière.

Comme les points caractéristiques bathymétriques doivent être déplacés de façon à être confondus avec leurs homologues de référence, le poids choisi est de « 1 » pour le modèle de référence (aucune déformation) et « 0 » pour le modèle bathymétrique.

DEL8 comporte un impact de correction maximal vis-à-vis des autres MNT traités entre la zone de recouvrement déformée élastiquement et l'extrémité du modèle qui n'est absolument pas modifiée. En effet comme expliqué précédemment, l'algorithme utilisé pour la déformation locale a été choisi de manière à éviter les effets « marches d'escalier ». En conséquence son effet est maximal au voisinage des points caractéristiques ($R = 0$ à 1 , entourés en rouge sur la figure 88 **b**)), et décroît rapidement à partir du $\frac{1}{4}$ de la distance considérée (soit 2 pour un rayon de 8). A la moitié de la bande de déformation ($R = 4$ dans cet exemple) le coefficient de déformation n'est plus que de 50% (voir courbe de variation du coefficient d'amortissement présentée figure 84) et décroît drastiquement pour être nul à $R = 8$. Les points situés en périphérie de la zone d'influence ne sont donc pas ou peu déplacés puisqu'ils ne sont pas compris dans le rayon d'action majeur de la déformation élastique. La variation de l'impact de la modification des points selon le rayon de correction choisi est illustrée figure 88 **(d)** par la superposition des points issus des différentes déformations élastiques.

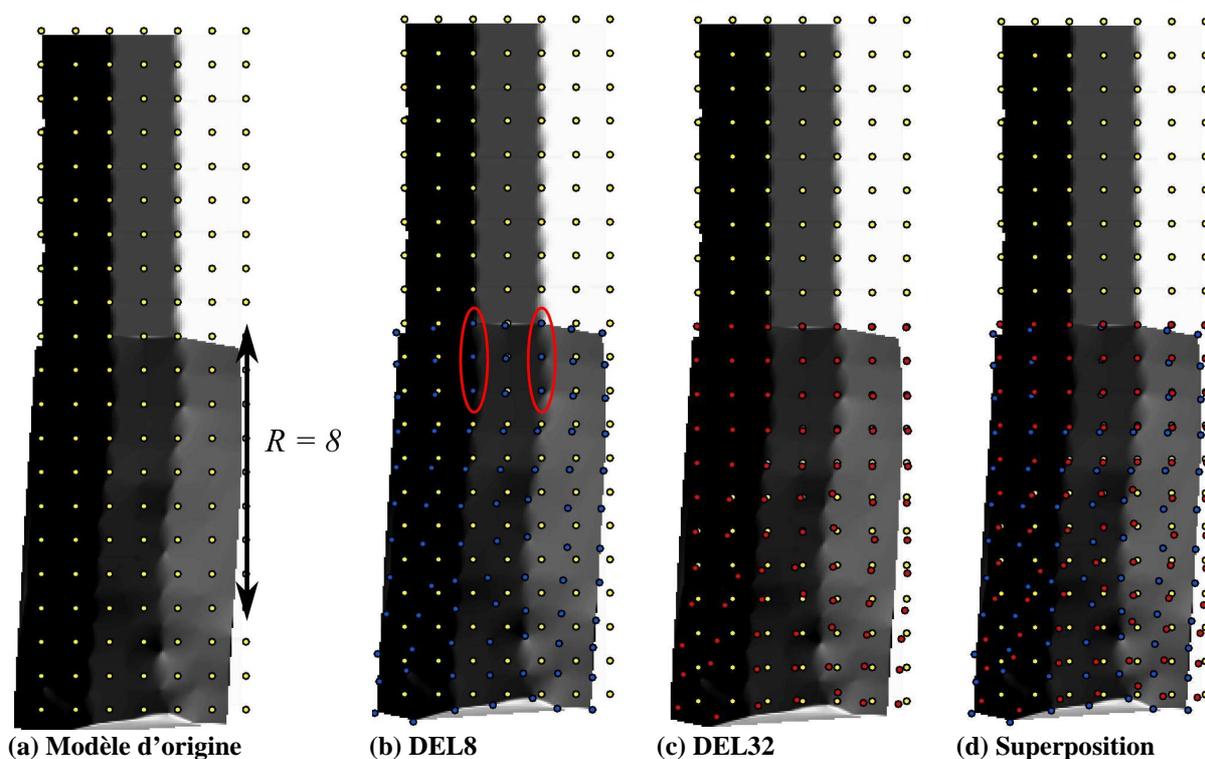


Figure 88 : Représentation du rayon d'influence de la déformation locale. En a) les points du modèle de référence ont été superposés à l'estompage du MNT bathymétrique déformé. En b) ont été rajoutés les points déformés avec une déformation locale d'un rayon égal à 8 (points bleus). Les points caractéristiques qui ont servi à la déformation sont entourés d'une ellipse rouge. En c) sont visualisés en rouge les points déformés localement avec un rayon égal à 32. d) montre la superposition de tous les points, de référence en jaune, déformés avec un rayon de 8 en bleu, et avec un rayon de 32 en rouge.

Le résultat montré figure 88 met bien en évidence que les points homologues sont toujours confondus (ellipse rouge figure 88 **b**)) quel que soit le rayon de déformation. Les points rouges et bleus en revanche ne sont déformés de manière identique que sur les trois premiers mètres (figure 88 **(d)**).

Le rayon égal à 32 permet de lisser ces déformations grâce à une bande de déformation comprenant tout le modèle. La superposition des points d'origine (en jaune) des points

déformés avec un rayon de 8 (bleu) et des points déformés avec un rayon de 32 (rouge), figure 88 (d), permet d'apprécier la variation entre les corrections effectuées. Pour une distance de 4 en partant de l'extrémité comprise dans la bande de recouvrement, les points rouges ($R = 32$) sont quasi confondus avec les points jaunes tandis que les points bleus possèdent encore un écart compris entre 0.2 et 0.3.

Les modèles translatsés figure 87 ne présentent pas de différence entre eux malgré un rayon variable : la translation permet ici de corriger l'essentiel de l'erreur due aux décalages en x, y et z. Le biais est donc parfaitement supprimé par le calcul d'un vecteur moyen et de son application à l'ensemble du modèle. Il est à noter que les erreurs grossières et de bruit blanc ne sont pas corrigées vu qu'elles influent assez peu sur les variations de positionnement calculées à partir des points homologues.

Le paragraphe suivant montre comment le déplacement des points selon un rayon variable influe sur la qualité du modèle fusionné. Les figure 89 et figure 91 suivantes illustrent l'application des méthodes de déformation présentées précédemment pour la fusion de données topographiques et bathymétriques. Le poids respectifs des données choisi est de 0,5 : aucune donnée ne prévaut.

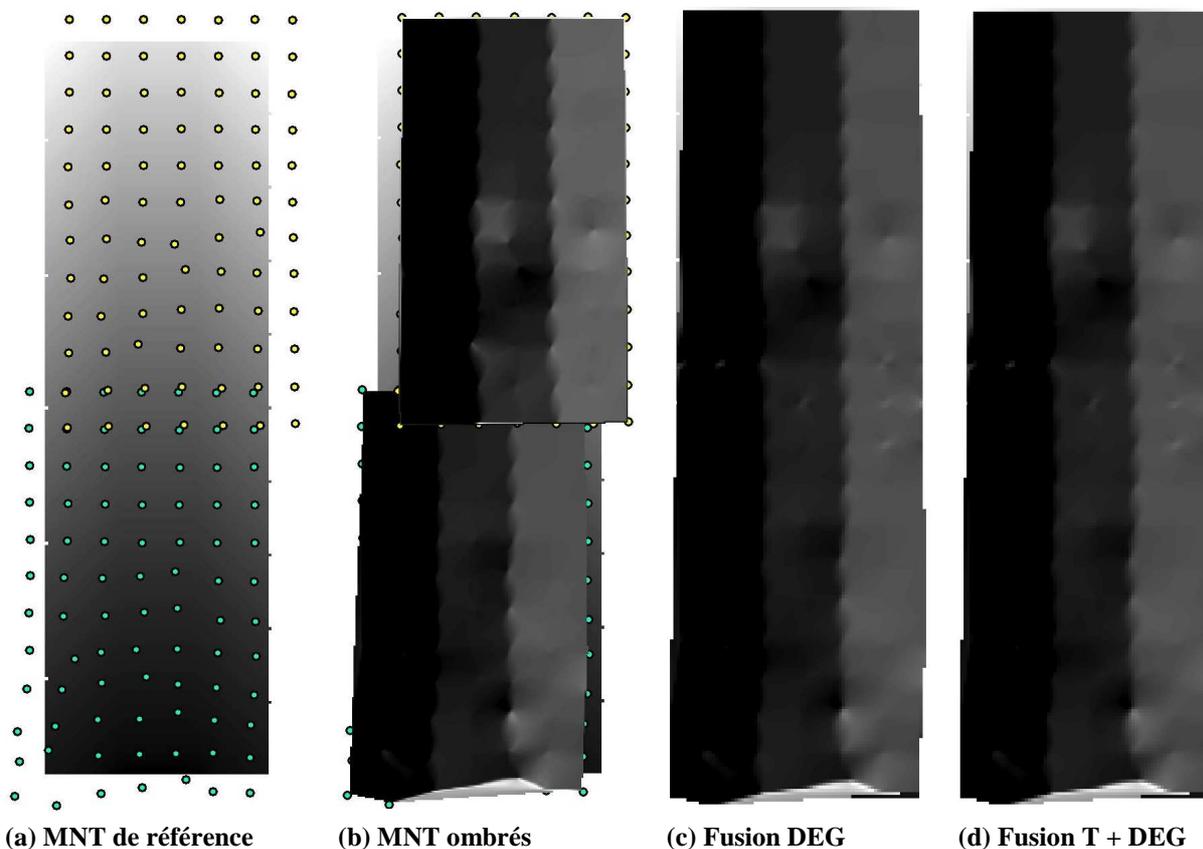


Figure 89 : Résultat de différentes méthodes de déformation élastique sur les modèles topographique et bathymétrique. En a) représentation des points de départ ; b) estompage de ces points et visualisation des erreurs introduites dans chaque modèle ; c) modèle déformé globalement sans translation ; d) modèles déformés globalement avec translation préalable.

L'intégration des modèles présentés est réalisée au moyen de trois types de déformation :

- déformation élastique appliquée à l'ensemble des modèles (DEG) ;
- déformation élastique appliquée aux points compris dans un rayon de 8 (DEL8) ;
- déformation élastique appliquée aux points compris dans un rayon de 32 (DEL32) ;

Chacune de ces déformations a été effectuée avec et sans translation préalable.

La figure 89 **(a)** représente les points topographiques (en jaune) et les points bathymétriques (en bleu) bruités. Le biais (qui représente l'erreur principale à corriger) est en x , y et z et égal à : 0.6, 0.6 et 0.8 pour la partie topographique, et -0.6, -0.6, 0.8 pour la partie bathymétrique. L'erreur en z étant la même (0.8), elle ne sera donc pas supprimée avec cette méthode. En **(b)**, l'estompage permet de visualiser les erreurs introduites dans chacun de ces modèles, notamment les erreurs grossières. Les modèles **(c)** et **(d)** suivants représentent la fusion à l'aide d'une déformation élastique effectuée sur l'ensemble des modèles. **(d)** comporte une translation effectuée auparavant. Aucune différence n'est visible entre ces deux résultats (voir figure 90).

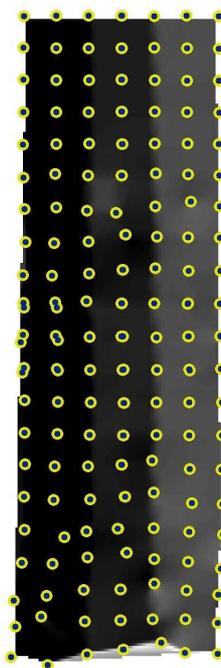
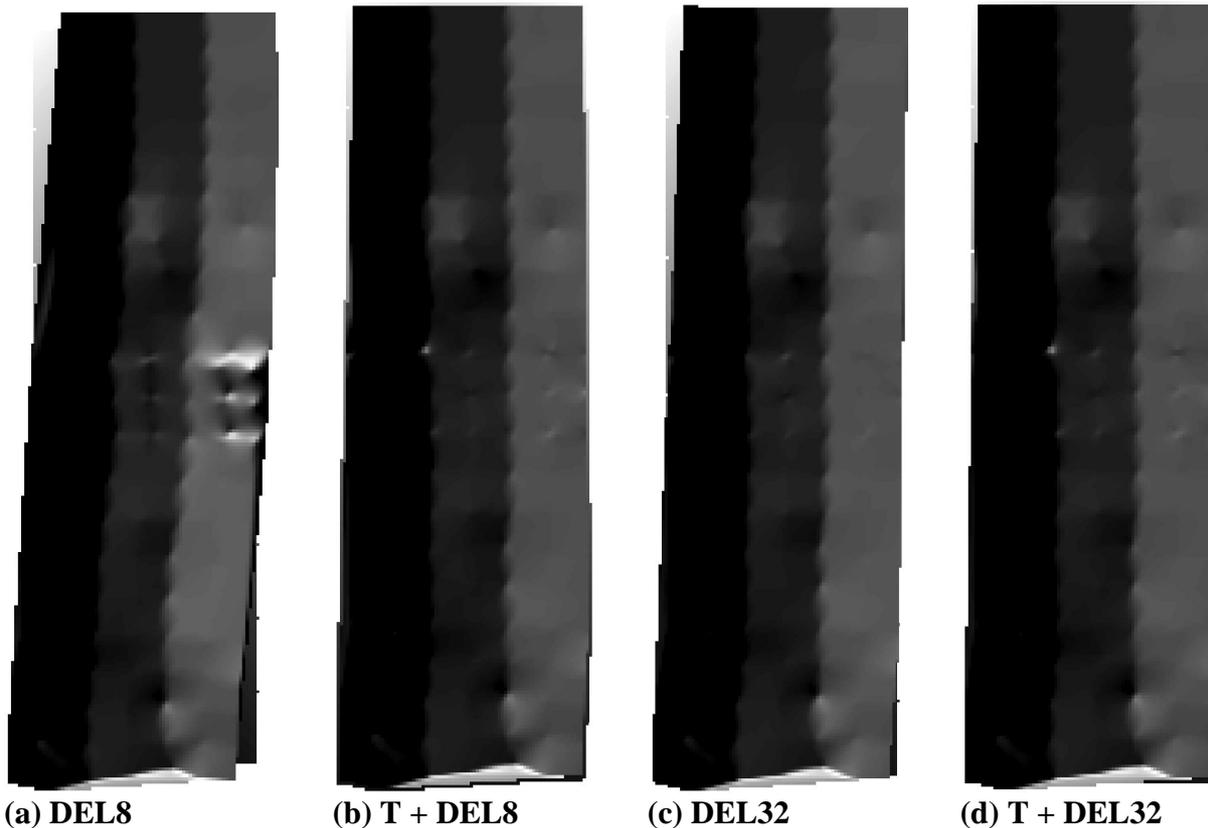


Figure 90 : Superposition des modèles intégrés sous forme de points. En jaune : modèle déformé avec une méthode de déformation élastique globale (DEG) ; en bleu : modèle translaté puis déformé avec une déformation élastique globale (T+ DEG).

Une explication de cette similitude est due au caractère « global » de DEG : la translation préalable ne change rien au MNT final obtenu.

La figure 91 présente les modèles fusionnés à l'aide d'une bande de déformation élastique à rayons variables choisis égaux respectivement à 8 (en **(a)** et **(b)**) et 32 (**(c)** et **(d)**). Le modèle intégré avec un rayon égal à 8 présente des distorsions fortes dans la zone de recouvrement, semblables à des effets de tôle ondulée. Ceci est à mettre en relation avec l'image de la figure 88 **(b)** reprise figure 92 avec agrandissement sur les points de la bande de déformation.

Malgré un rayon de 8 qui comprend les points situés à l'extrémité du MNT, la pondération fait que la déformation induite n'a pas assez de poids sur les bords pour corriger convenablement le modèle en x , y , mais également en z .



(a) DEL8

(b) T + DEL8

(c) DEL32

(d) T + DEL32

Figure 91 : Résultat de différentes méthodes de déformation élastique avec un rayon sur les modèles topographique et bathymétrique. En a) modèle déformé selon un rayon de 8 ; b) modèle translaté puis déformé selon un rayon de 8 ; c) modèle déformé selon un rayon de 32 ; d) modèle translaté puis déformé selon un rayon de 32.

Les modèles translatés, quelle que soit la valeur du rayon, présentent de grandes similitudes avec la déformation élastique globale présentée figure 89. Pour ceux non translatés en revanche, des différences notables sont remarquables.

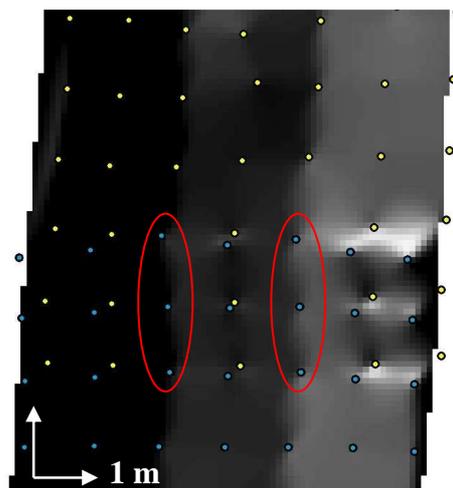


Figure 92 : Détail du déplacement des points selon un rayon $R = 8$.

Le modèle fusionné présenté figure 91 (c) ne présente pas ces mêmes artefacts car le rayon est suffisamment grand pour corriger ces valeurs.

III.3.2 Modèle Numérique de Terrain représentant un littoral

Un second modèle est utilisé pour effectuer un test de la méthodologie dans des conditions plus proches des données réelles qu'un simple chenal. Là encore deux zones (une partie maritime et une partie terrestre) sont produites de manière à étudier plus en détail les problèmes de recouvrement de données et de déformations locales et globales.

Les deux modèles produits possèdent un domaine de recouvrement total et suffisant pour permettre une extraction d'éléments caractéristiques. Pour cette simulation les éléments caractéristiques sont représentés par des lignes de rupture de pente.

Deux séries de tests sont réalisées : la première concerne la déformation, en vue d'une fusion, d'un modèle bruité sur le modèle d'origine, et la seconde la déformation conjointe du modèle bathymétrique et du modèle topographique.

III.3.2.1 Fusion d'un modèle bruité et d'un modèle d'origine

L'objectif dans cette étape est de tester le processus de fusion dans un cas idéal : un modèle « parfait » sur lequel va être déformé un MNT bruité lui-même issu de ce modèle de base. La ligne de rupture de pente du modèle de référence (qui sert de ligne caractéristique sur laquelle se base l'appariement) est la ligne d'altitude égale à zéro (ligne rouge figure 93).

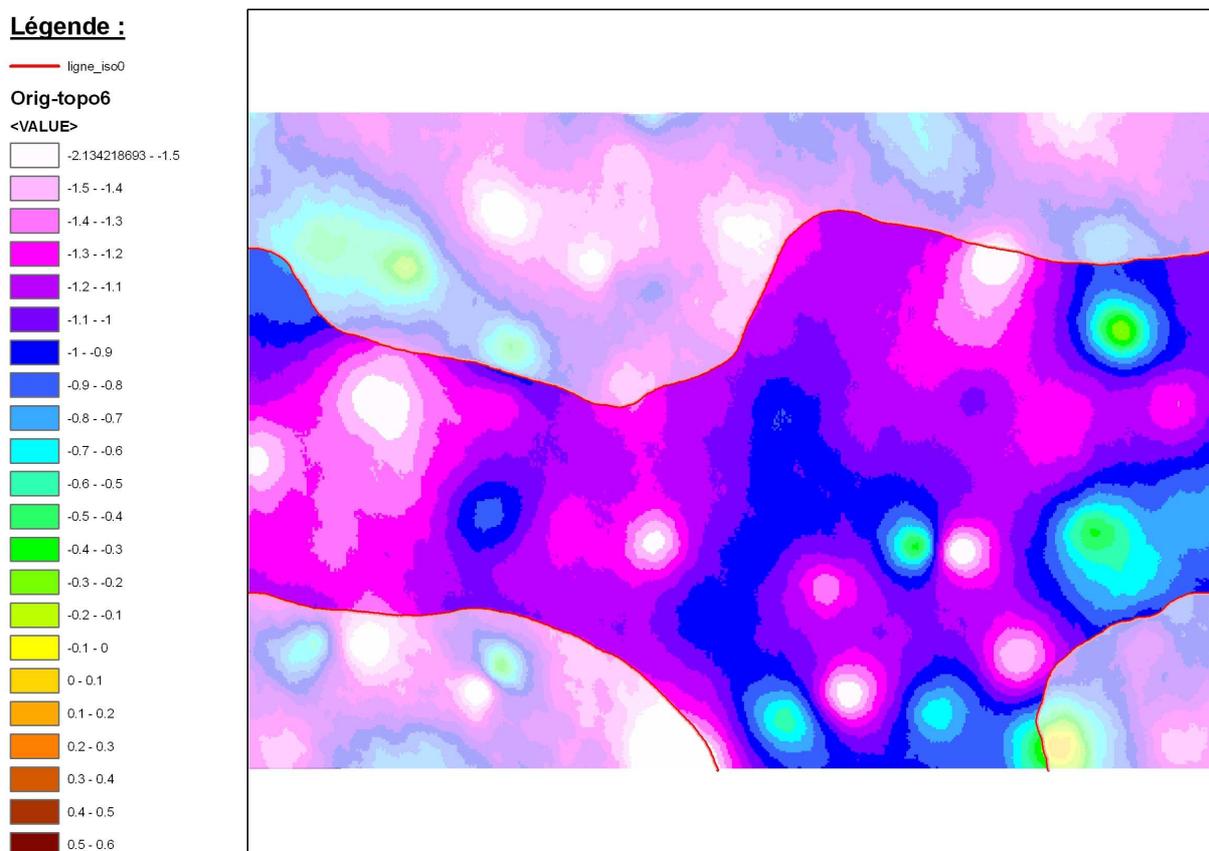


Figure 93 : Carte de différences altimétriques entre le modèle d'origine et le modèle bruité bathymétrique.

La ligne caractéristique du modèle bruité est obtenue en appliquant sur la ligne de référence extraite les mêmes paramètres de bruit utilisés pour le bruitage du MNT à déformer. Il n'y a

donc aucune incertitude quant à l'interprétation ou l'extraction des lignes qui servent au calcul du biais global et à sa correction en vue de l'intégration ultérieure.

Par ailleurs le poids choisi pour la déformation ne porte pas non plus à discussion : 1 pour le modèle d'origine et 0 pour le modèle bruité.

Une rapide analyse de la qualité respective des données est effectuée par soustraction altitudinale du modèle bruité avec celui d'origine. La figure 93 montre l'intervalle de valeurs différentielles entre les deux modèles sur l'ensemble de la zone bathymétrique avant toutes corrections. Les deux modèles comportent un intervalle d'erreur à corriger qui se situe entre [-0.2 ; -2.13].

Cartes de différences :

Une première étape évalue le biais global à l'aide de la moyenne des vecteurs de déformation en vue de la correction du modèle bathymétrique. Après calcul des points homologues sur les lignes caractéristiques extraites, le biais donne des valeurs de :

- 0,21 en X (pour 0,6 à l'origine) ;
- 0,47 en Y (0,6) ;
- 0,99 en Z (1).

Le biais moyen retrouvé est donc très proche de ce qui a été ajouté au modèle de départ afin de le bruite : 35% du bruit est récupéré en X, 78% en Y et 99% en Z. Plusieurs informations sont déductibles de ces premiers résultats :

- la méthode est capable de bien corriger les biais planimétrique et altimétrique, ce que ne font pas les autres algorithmes de fusion (voir état de l'art, chapitre II) ;
- Z et Y présentent tout deux une bonne estimation du biais ;
- la correction en X possède en revanche moins de similitude avec la valeur de départ.

Ce différentiel plus important en X s'explique de plusieurs façons.

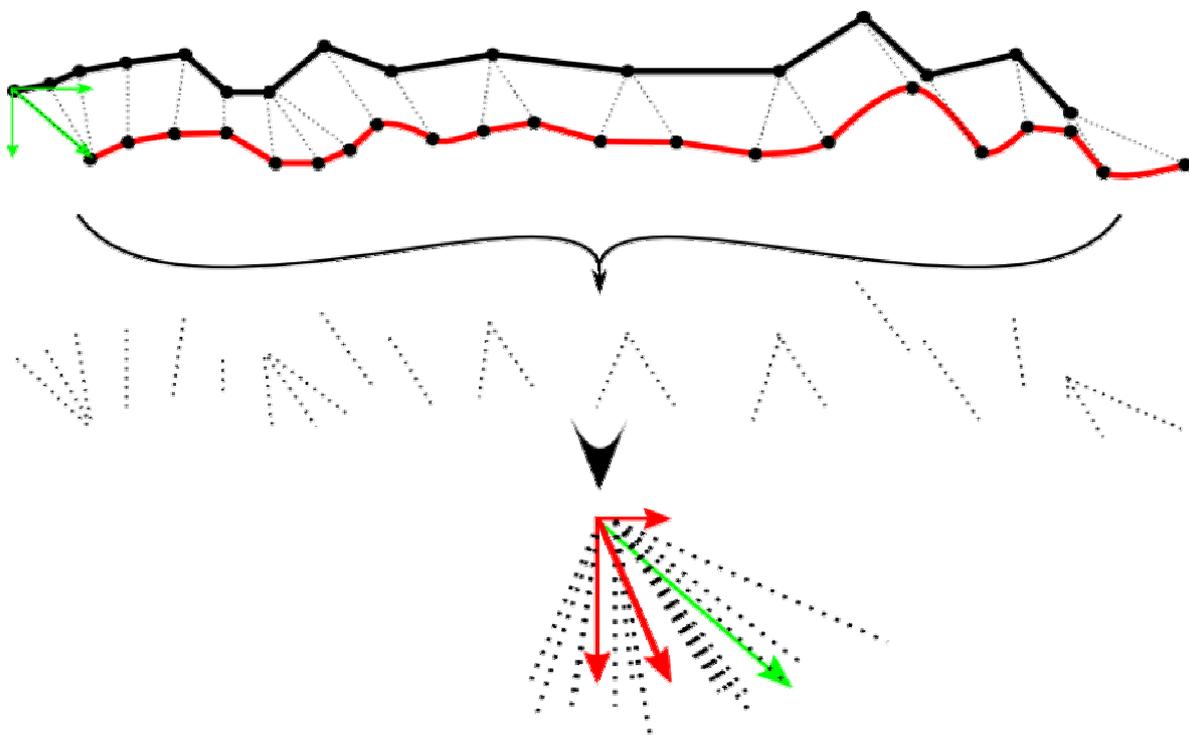


Figure 94 : Calcul du biais moyen à partir des couples de points homologues.

Tout d'abord le calcul de la distance de Fréchet et la détermination des points homologues à partir de chemin minimal est basé sur un ordonnancement des couples de points et des

distances euclidiennes entre des points. L'algorithme part donc des deux points de départ de L_1 et L_2 et construit son chemin minimal sur les couples de points ordonnés qui sont les plus proches (voir figure 94). Un biais diagonal constant est donc décelable entièrement uniquement à partir du moment où les lignes ont régulièrement des inflexions ou des changements de direction.

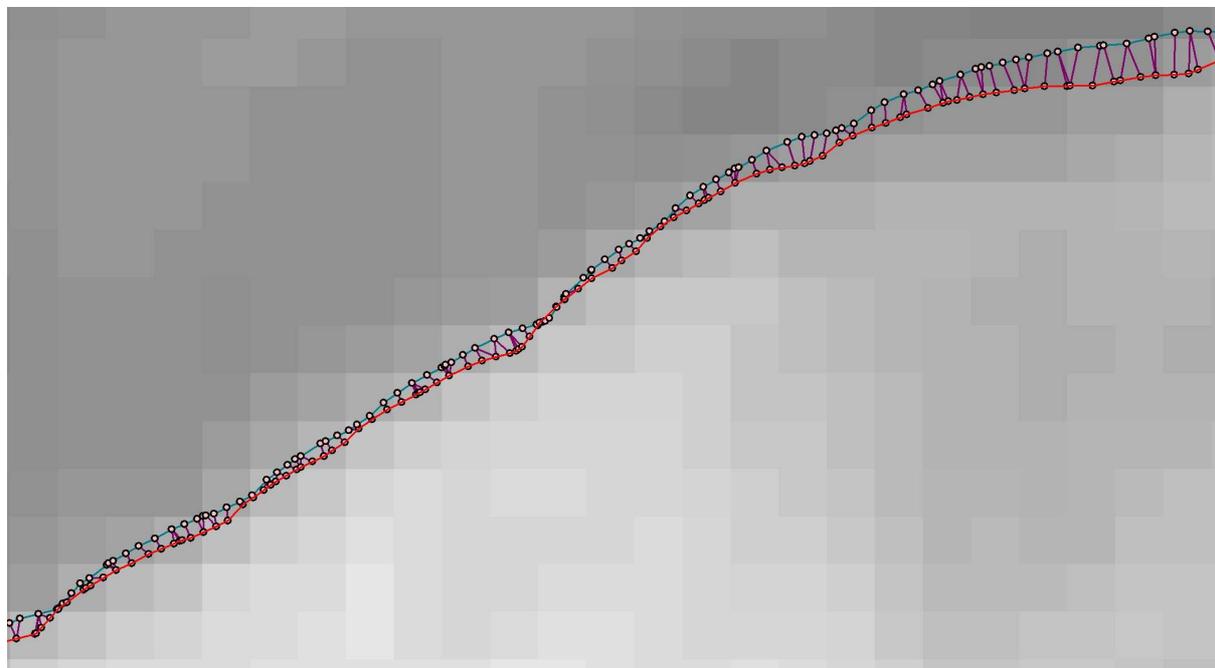


Figure 95 : Exemple de couples de points appariés entre la ligne caractéristique de référence en rouge et la ligne bruitée en bleu. Les traits violets représentent les couples.

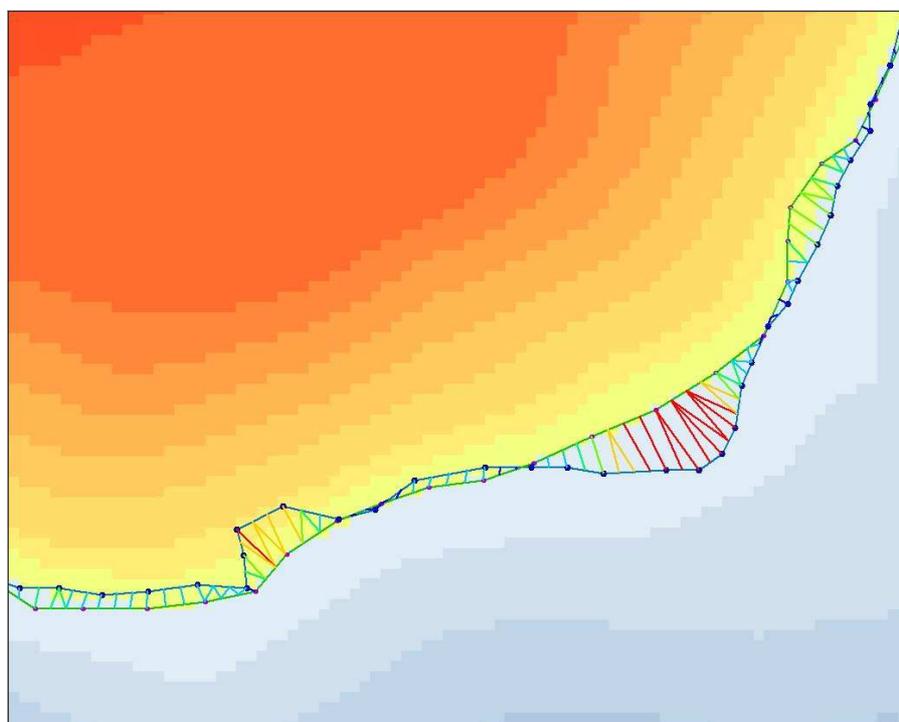


Figure 96 : Exemple de couples de points appariés entre deux lignes caractéristiques. Les traits entre les points représentent les couples. La visualisation avec le logiciel SIS® permet d'affecter à chaque trait une couleur en fonction de sa valeur : plus le trait entre les points est rouge, plus la distance (et donc le biais local) est importante.

Sur le schéma de la figure 94, les vecteurs de déformation entre les couples de points homologues sont récupérés et additionnés pour être comparés au biais d'origine (en vert sur la figure). Le vecteur final (en rouge) est proche du vecteur d'origine. L'écart s'explique du fait que les lignes n'ont pas totalement la même forme ni le même nombre de points mais surtout car elles sont majoritairement horizontales (donc parallèles à l'axe des X), ce qui biaise énormément le résultat final obtenu.

Dans l'exemple présenté figure 93, les lignes caractéristiques étant essentiellement parallèles à l'axe des X, tandis que le biais ajouté est fortement diagonal, une mauvaise estimation en X est induite par la méthode d'appariement. Le résultat s'améliore nettement en Y. En effet les points choisis comme points homologues sont les points les plus proches, qui ne sont pas toujours dans la même direction que le biais (figure 94), ce qui minimise le déplacement réel des points en X. La somme des vecteurs trouvés traduit donc bien la tendance, mais avec une valeur atténuée. Pour palier ce problème il faudrait avoir un nombre supérieur de lignes caractéristiques orientées aléatoirement de façon à éliminer ce phénomène dû à la direction préférentielle des objets appariés.

Les lignes au biais diagonal régulier figure 95 illustrent bien ce phénomène de régularité de positionnement des vecteurs entre points homologues aux extrémités hautes des lignes caractéristiques. Lorsque les lignes ne sont plus complètement horizontales ces mêmes vecteurs de déplacement varient notablement dans leur taille et leur orientation préférentielle. La figure 96 illustre un cas où les lignes ne sont plus aussi parfaitement parallèles en raison des erreurs locales fortes : dans ce cas les vecteurs de déformation trouvés entre les points homologues sont fortement dissemblables. La somme des vecteurs de correction entre points homologues va donc refléter, pour cette zone, l'impact du biais global diagonal plus les erreurs locales fortes. Néanmoins les erreurs locales étant aléatoires, globalement elles s'annulent.

Légende :

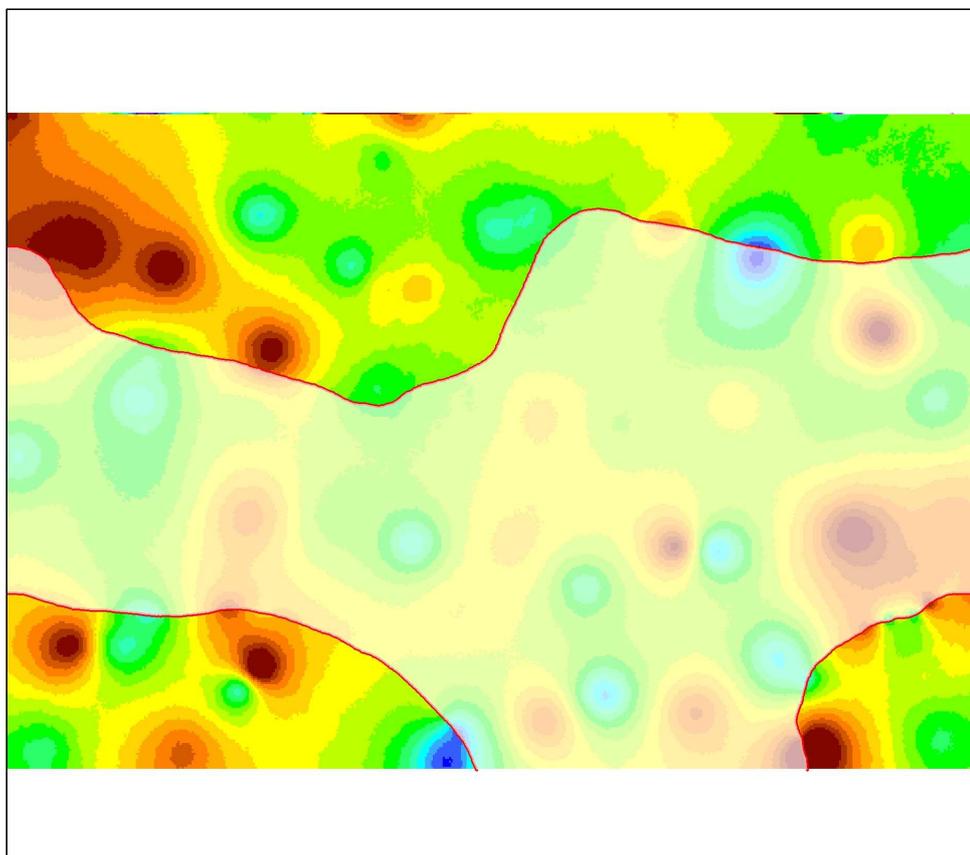
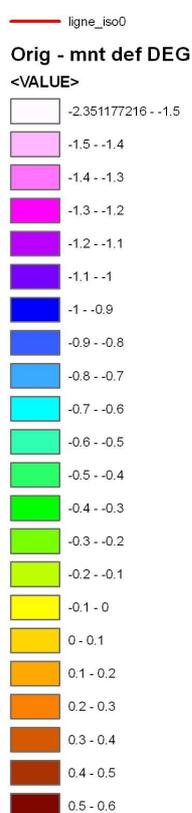


Figure 97 : Soustraction du modèle de référence avec le modèle topographique déformé rigide traduit (TRANS).

Le résultat des diverses déformations réalisées montre d'excellents résultats : les bandes de déformation élastique effectuent remarquablement bien leurs corrections sur la zone impartie, laissant inchangée la qualité (bonne ou mauvaise) du reste du modèle. La correction TRANS est également très bonne sur le modèle bruité (figure 97) : l'erreur générale du modèle située entre -0,6 et 1 a été corrigée pour n'être plus comprise qu'entre 0,1 et -0,2.

Modèles fusionnés :

Les figure 98 et figure 100 montrent des profils réalisés sur les MNT fusionnés : le modèle DEL8 présente un replat assez important suivi d'une légère augmentation d'altitude puis d'une baisse à nouveau. Ceci est dû au fait que l'algorithme déforme énormément les points jusqu'à R/2 (soit 4 mètres) mais beaucoup moins par la suite. Comme l'altitude de modèle bruité est située au dessus du modèle de référence (en raison de son fort biais positif en Z) sur la bande de déformation, les points remontent assez vite à partir de R=6. Ces points suivent ensuite le modèle bruité qui n'est plus déformé, ce qui entraîne la présence de cette « marche d'escalier » représentée par la superposition des points figure 99.

Cet effet n'est pas une erreur d'algorithme mais bien un résultat de ce qui a été demandé : la bande de déformation d'un rayon de 32m annule complètement ce phénomène de replat en raison d'une correction plus importante de la bande (DEL 32 : figure 98).

L'artéfact apparu sur DEL8 est un cas particulier qui n'est pas réaliste étant donné qu'il se produit uniquement dans le cas où les données bathymétriques sont supérieures en altitude aux données topographiques, ce qui n'arrive pas normalement.

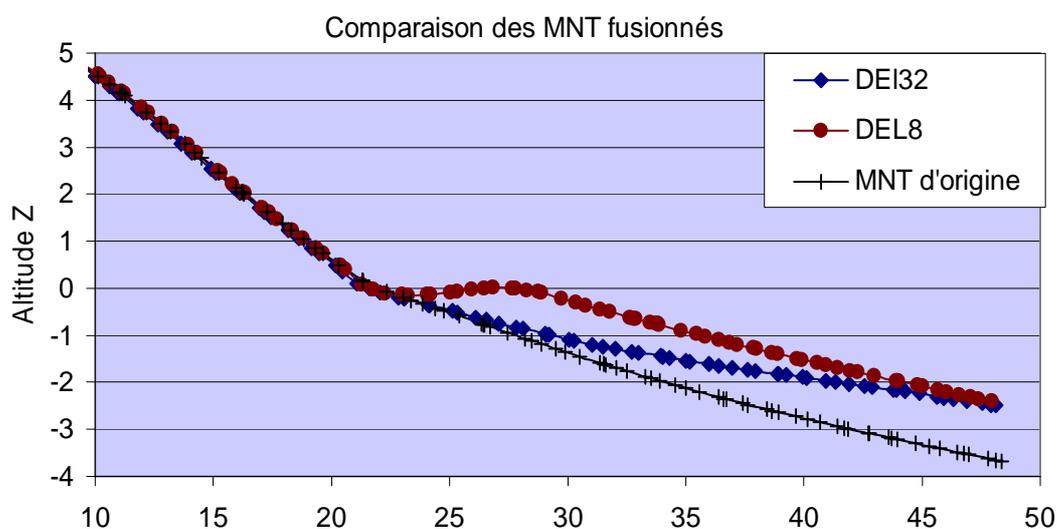


Figure 98 : Comparaison des modèles DEL8 et DEL32 avec le modèle de référence.

Tous les autres modèles comportant des corrections de biais sur leur ensemble (figure 100) ne présentent aucun artéfact de ce type : les profils allant de X=20 à X=24 (altitudes comprises entre 0.5m et -0.5m) se suivent tous (même DEL8 +T) et sont légèrement sur-correctés en profondeur par rapport au modèle d'origine.

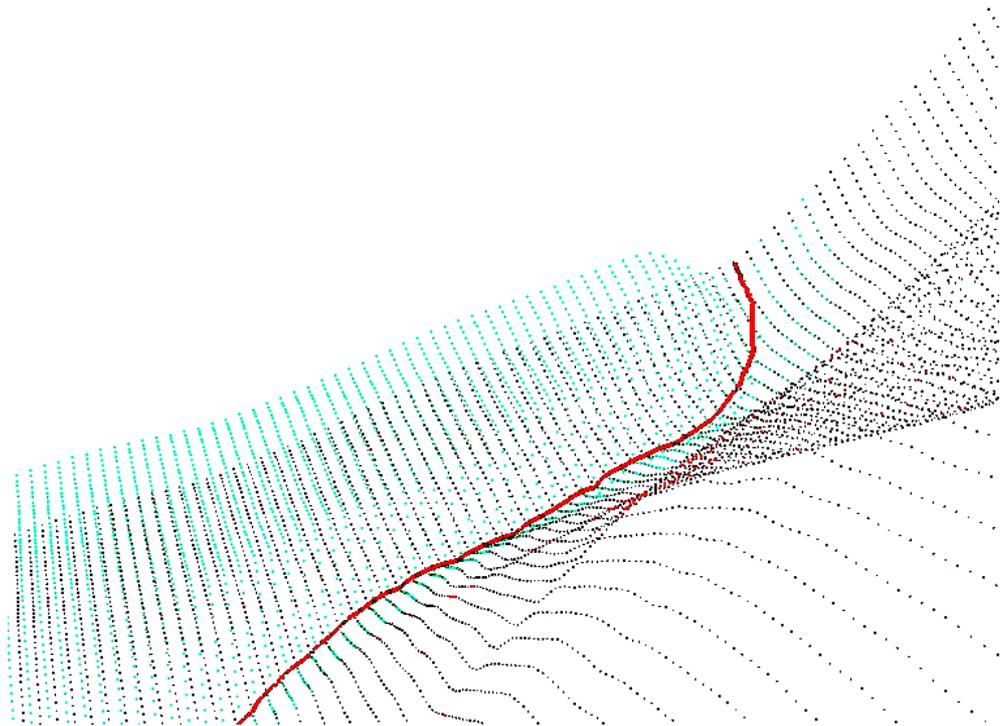


Figure 99 : Superposition du modèles de référence en rouge et du modèle bruité en bleu. Altitudes exagérées 5 fois. La ligne rouge représente la ligne caractéristique (rupture de pente du modèle d'origine) de référence.

Comparaison des MNT fusionnés

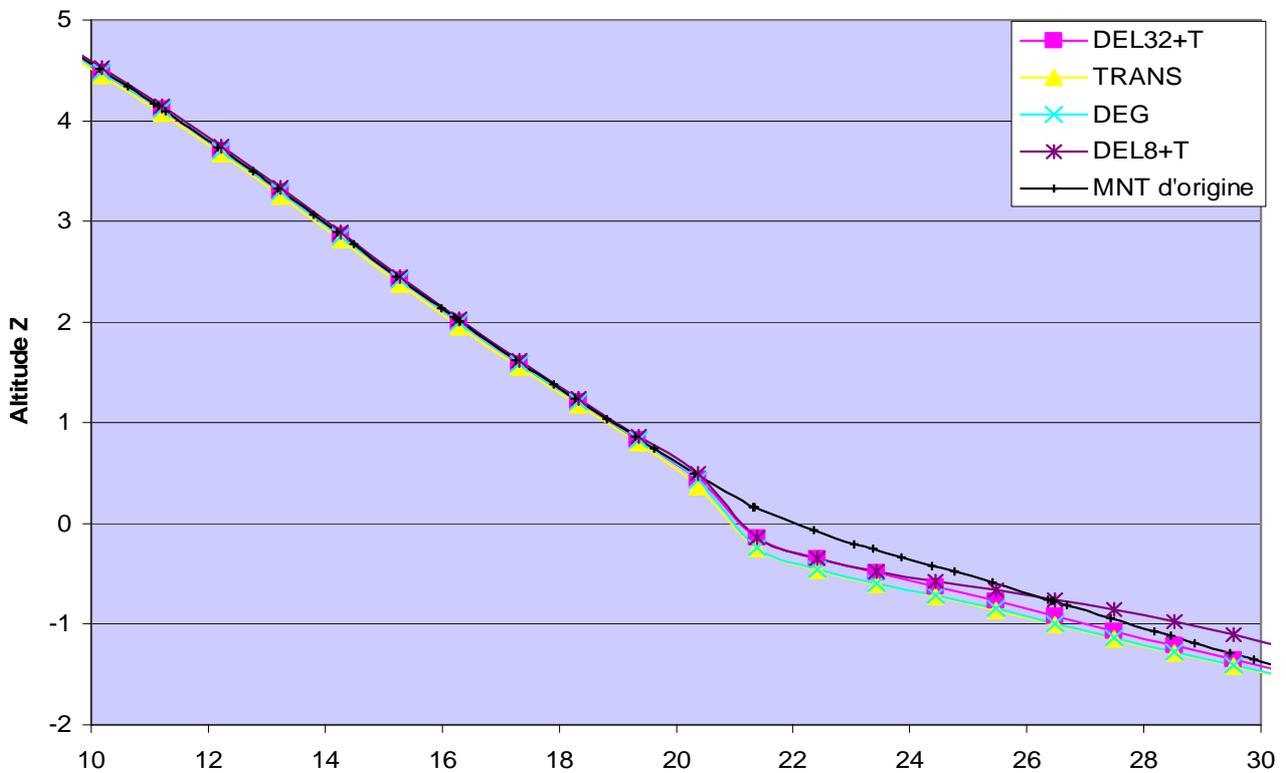


Figure 100 : Comparaison des modèles DEL32+T, TRANS, DEG et DEL8+T avec le modèle de référence.

III.3.2.2 Fusion du modèle topographique et du modèle bathymétrique

De la même manière que pour les modèles précédents, une analyse de la qualité respective des données est effectuée par soustraction altitudinale du modèle considéré (terre ou mer) avec celui de référence. La figure 101 montre la variation de cette soustraction entre modèle d'origine et modèle bathymétrique sur l'ensemble de la zone : la partie terrestre est cachée à l'aide d'un masque d'une opacité de 35%. La conservation de la partie terrestre s'explique par le fait que la continuité colorimétrique facilite la lecture, particulièrement pour comprendre des déformations locales comme DEL8 ou DEL32 avec ou non une déformation rigide (TRANS).

Toute la partie maritime est extrêmement proche du modèle initial : la couleur jaune orangée dominante sur la carte indique une variation comprise entre +0,2 et -0,1. Le MNT bathymétrique étant issu d'un modèle de référence terre/mer complet, la simulation bathymétrique se poursuit également en domaine terrestre. Comme ces valeurs n'ont pas de sens, là aussi un masque avec une transparence de 35% a été appliqué (voir figure 101). Elles ont été néanmoins conservées pour l'analyse du suivi des déformations effectuées par l'algorithme.

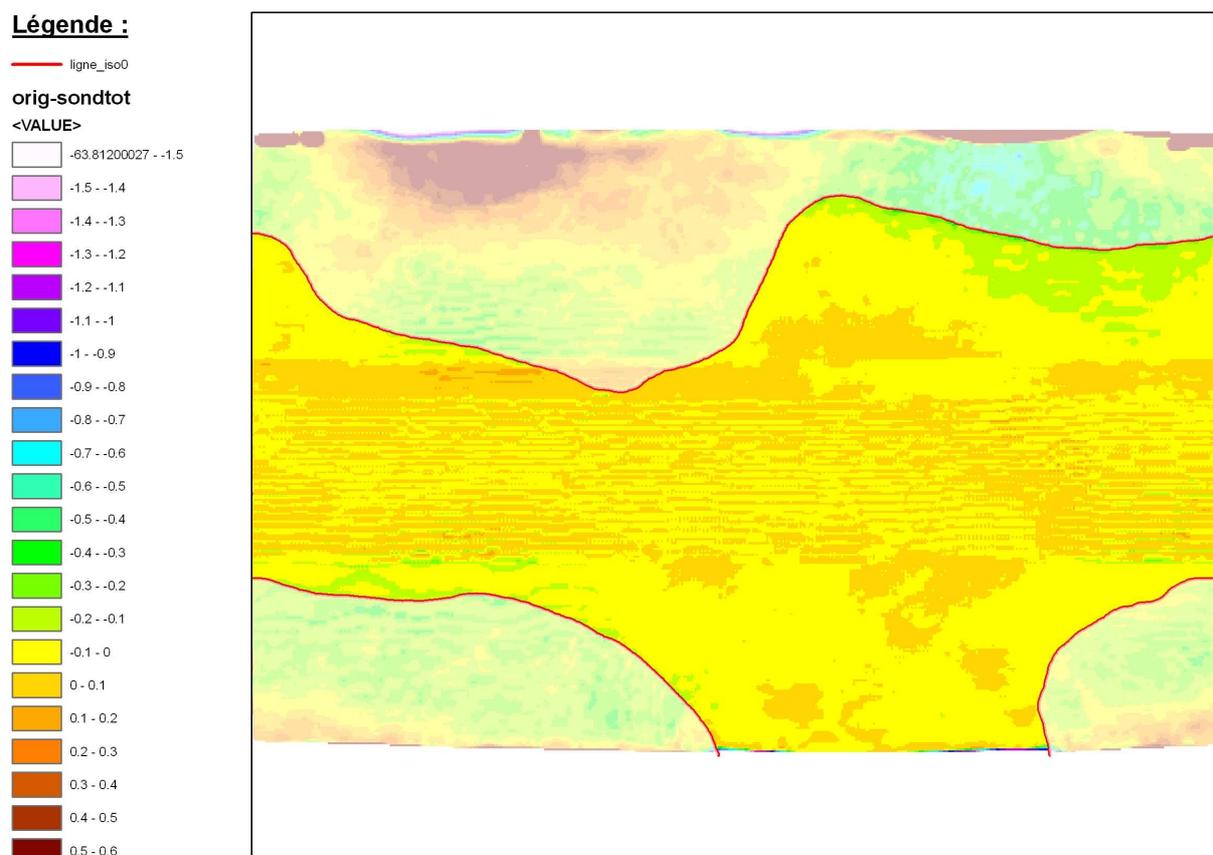


Figure 101 : Soustraction du modèle de référence et du modèle bathymétrique.

De la même manière une carte de différences est effectuée entre le modèle d'origine et le modèle topographique (figure 102). Le très fort biais introduit pour bruitez les données se retrouve sur l'ensemble du modèle : les couleurs majoritaires allant de bleu à blanc impliquent des valeurs de différences comprises entre -0,8 et >-1,5. Ces chiffres sont plus ou moins accentués par les erreurs locales insérées de manière aléatoire et comprises entre [-1 ; 1].

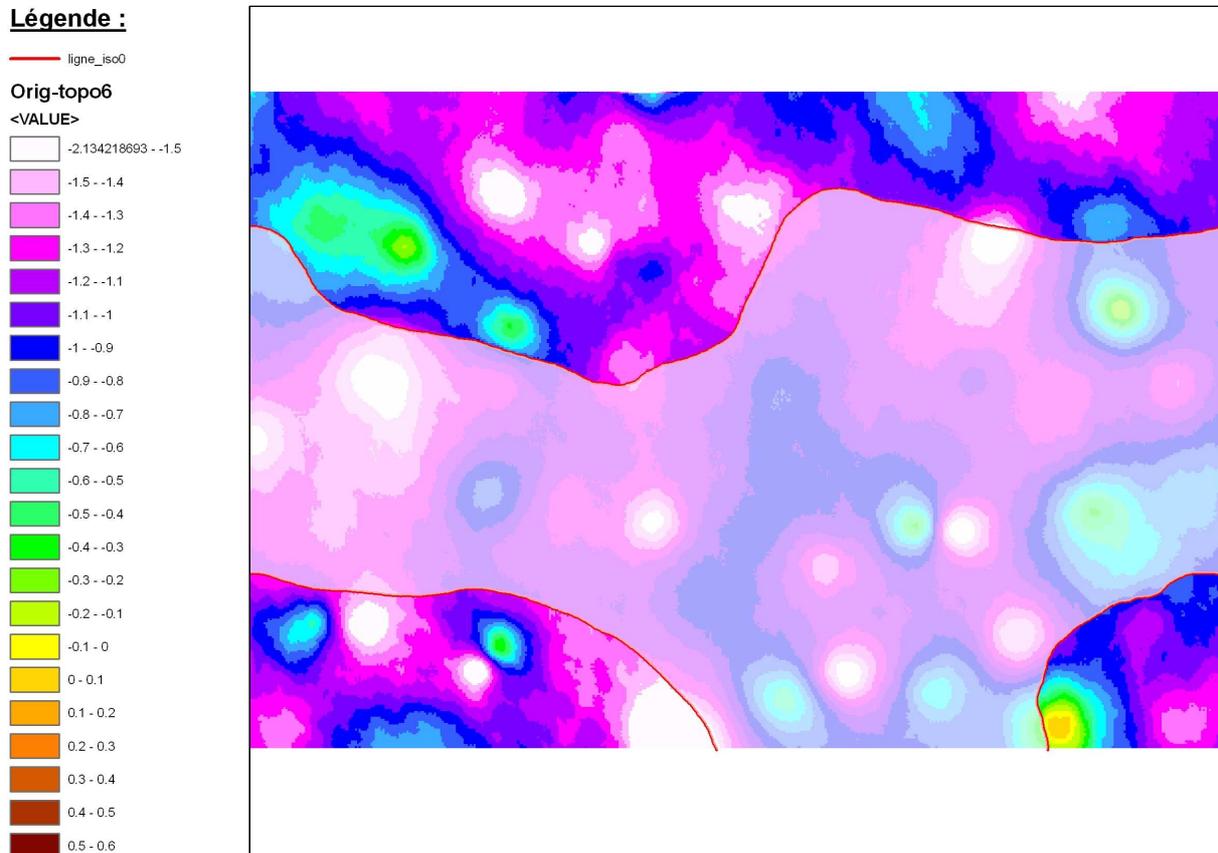


Figure 102 : Soustraction du modèle de référence avec le modèle topographique.

Les cartes de différences montrent que ces deux jeux de données, bathymétrique et topographique, possèdent une très nette différence de qualité d'un facteur de dix environ. En effet 95% des données bathymétriques sont comprises entre 0,1 et -0,1 tandis que les valeurs topographiques se situent entre -1 et >-1.5. Cette disparité induit l'utilisation d'une pondération différentielle pour la phase de déformation : 0,1 pour les données topographiques et 0,9 pour les données bathymétriques.

Chaîne de traitement des données :

La figure 103 résume le processus d'obtention des deux modèles qui vont servir aux tests de la méthodologie : le MNT de référence est divisé en une partie topographique et deux parties bathymétriques. Ces modèles sont produits avec les algorithmes de bruit décrits dans la section V.1.3.

Une évaluation qualitative de chacun des modèles a été produite par soustraction du modèle bruité au modèle de référence de manière à localiser les variations les plus importantes ainsi qu'estimer l'intervalle d'erreur. Les cartes différentielles sont représentées sous la forme de dégradés de couleur rose sur le schéma et détaillées figure 101 et figure 102 avec une échelle de couleur plus adaptée à une analyse visuelle.

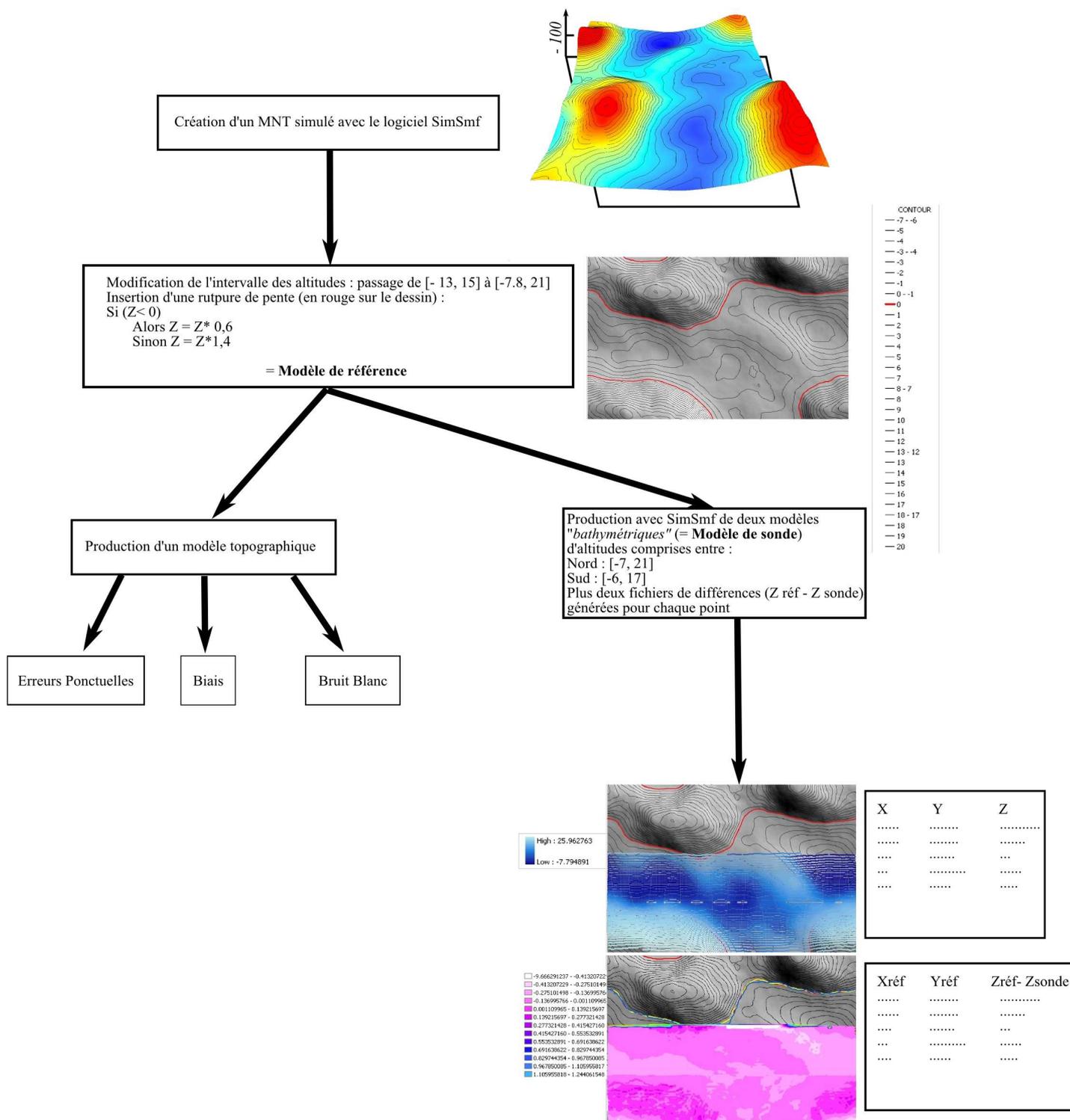


Figure 103 : Chaîne de traitement des données : production des modèles topographique et bathymétrique bruités.

En l'absence d'outils de segmentation testés et validés, les lignes caractéristiques de rupture de pente ont été extraites manuellement sur chacun des modèles terrestre et maritime. Ceci a été effectué à l'aide d'estompages de ces MNT réalisés avec ArcGis®. Des angles d'éclairage différents ont été utilisés (figure 104) de manière à souligner les lignes de rupture quelle que soit leur orientation préférentielle. La ligne caractéristique de référence du

modèle d'origine étant artificiellement la ligne du zéro, elle a été obtenue automatiquement en sélectionnant les points pour lesquels $Z = 0$.

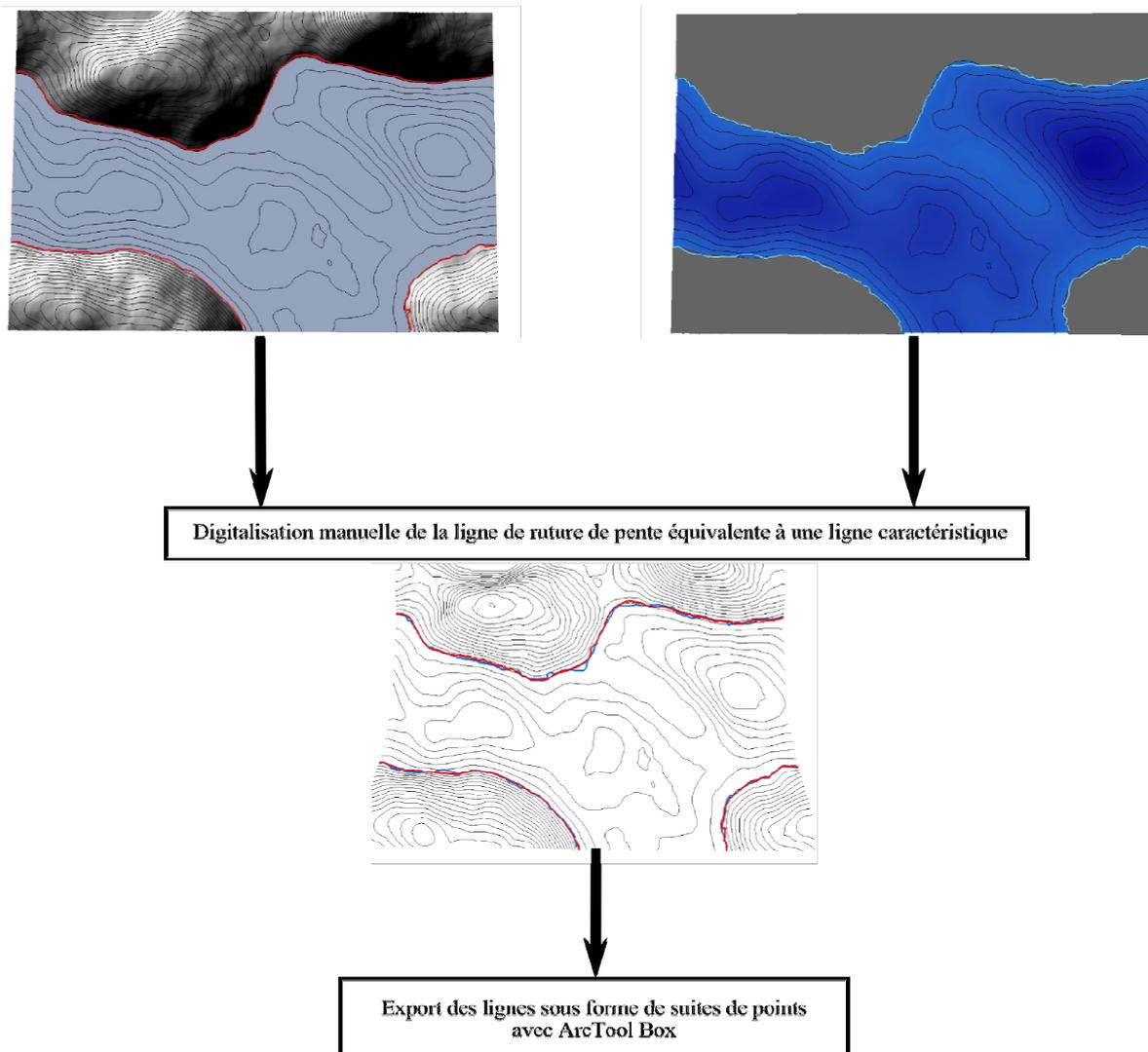


Figure 104 : Chaîne de traitement des données : digitalisation des lignes de référence et export des points pour le calcul de l'appariement.

Lorsque les lignes caractéristiques ont toutes été numérisées (référence, topographique et bathymétrique), celles-ci sont exportées sous forme d'une suite de points ordonnés regroupés par lignes dans un fichier XYZ avec un rééchantillonnage tous les mètres.

Les deux fichiers de lignes (appartenant au modèle bathymétrique et au modèle topographique dans cet exemple) sont importés dans le programme de fusion implémenté et les lignes sont appariées.

Pour chaque ligne sont calculés et exportés la distance de Fréchet discrète, la distance moyenne, et la liste de couples de points appariés (figure 105). Ces informations sont importées dans un SIG afin d'afficher les couples de points homologues avec un dégradé de couleur. Cette étape permet de visuellement évaluer si le résultat est cohérent ou pas, et de vérifier que la déformation des points homologues est conforme aux spécifications de l'utilisateur (choix du rayon de déformation, de la densité d'échantillonnage et de la valeur du seuil).

Lorsque la déformation des points homologues est validée (figure 106), ceux-ci sont utilisés pour estimer :

- une déformation moyenne calculée sur l'ensemble des points homologues. Elle doit être appliquée à tout le modèle en tant que déformation rigide (appelée TRANS dans l'exemple du chenal) ;
- un nuage de vecteurs localisés calculés à partir des différences de position d'un point par rapport à son homologue. Ces vecteurs serviront de déformation élastique localisée dans la bande de déformation. Le rayon d'investigation de cette bande est choisi par l'utilisateur. Pour comparer différents résultats, les rayons choisis pour l'exemple sont 8 et 32 (figure 106).

Figure 107, les champs de déformations (globale et élastique) sont récupérés et appliqués séparément puis conjointement sur les modèles :

- déformation élastique globale : DEG ;
- déformation avec une bande de déformation : DEL8 (R= 8) et DEL32 (R=32) ;
- déformation rigide : T pour TRANS ;
- déformation rigide + bande de déformation : T + DEL8 et T + DEL32.

Cette étape autorise l'évaluation individuelle de chacun des modes de correction effectués. Elle permet également de regarder si les déformations globale et locale sont bien en adéquation avec les attentes concernant le processus de fusion.

Chaque modèle obtenu est exporté sous forme de points XYZ dans un SIG. Ils sont ensuite interpolés en grille raster pour :

- permettre des calculs d'estompage visant à mettre en valeur les artefacts de déformation ;
- faciliter la lecture du relief ;
- autoriser des opérations entre grilles comme celle de soustraction.

En effet chaque modèle obtenu est soustrait au modèle de référence afin de produire une carte d'évaluation qualitative de la déformation et vérifier :

- la correction des biais globaux par rapport au modèle de départ ;
- le lissage ou non des biais locaux ;
- l'impact du rayon de déformation sur le résultat fusionné par rapport à la référence.

A la suite de ces traitements, les points bathymétriques et topographiques sont extraits avec un recouvrement de cinq mètres afin de produire un MNT fusionné Terre/Mer. Une visualisation complémentaire des points en 3D permet d'expliquer les apparentes incohérences des résultats obtenus.

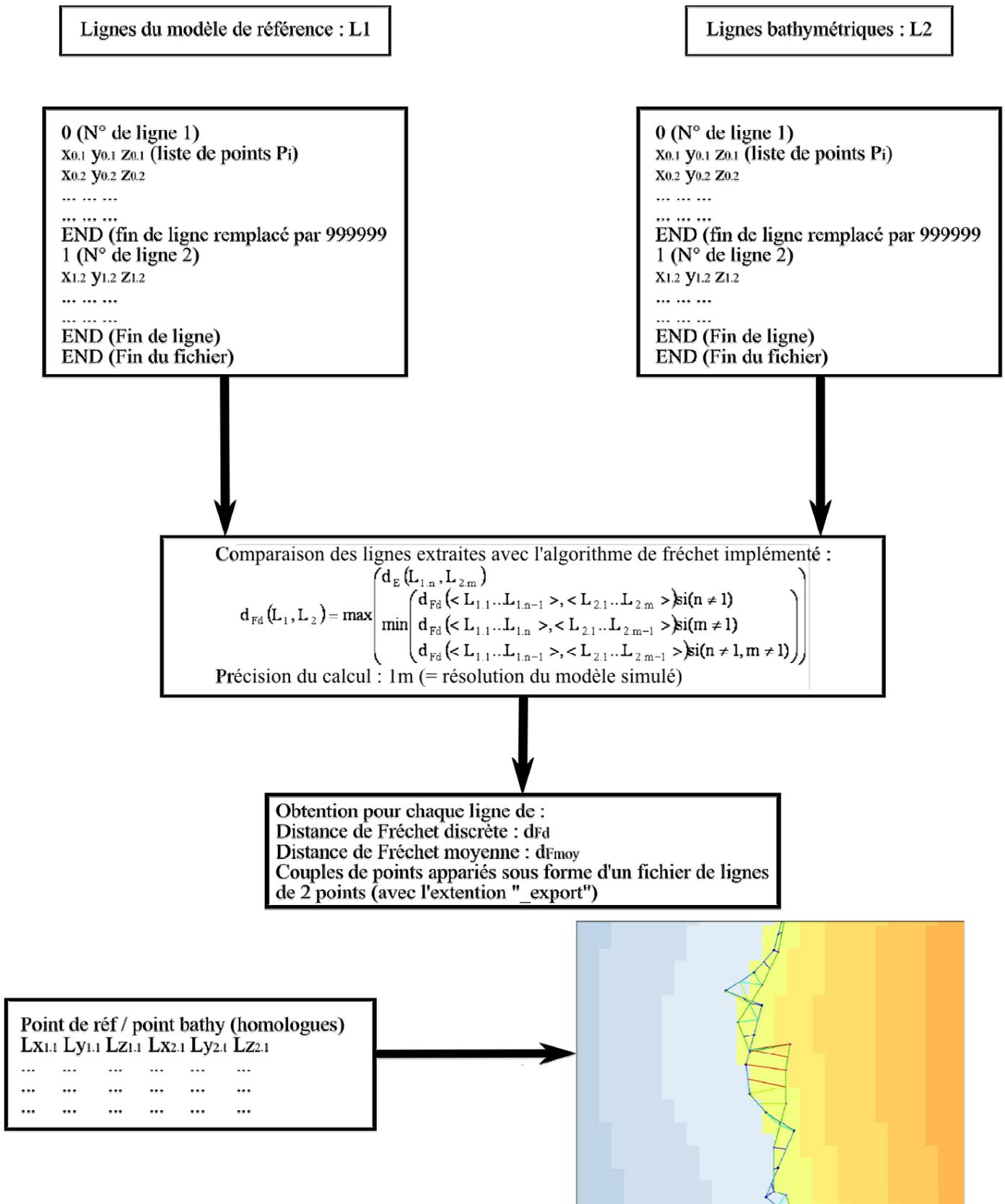


Figure 105 : Chaîne de traitement des données : calcul des points homologues.

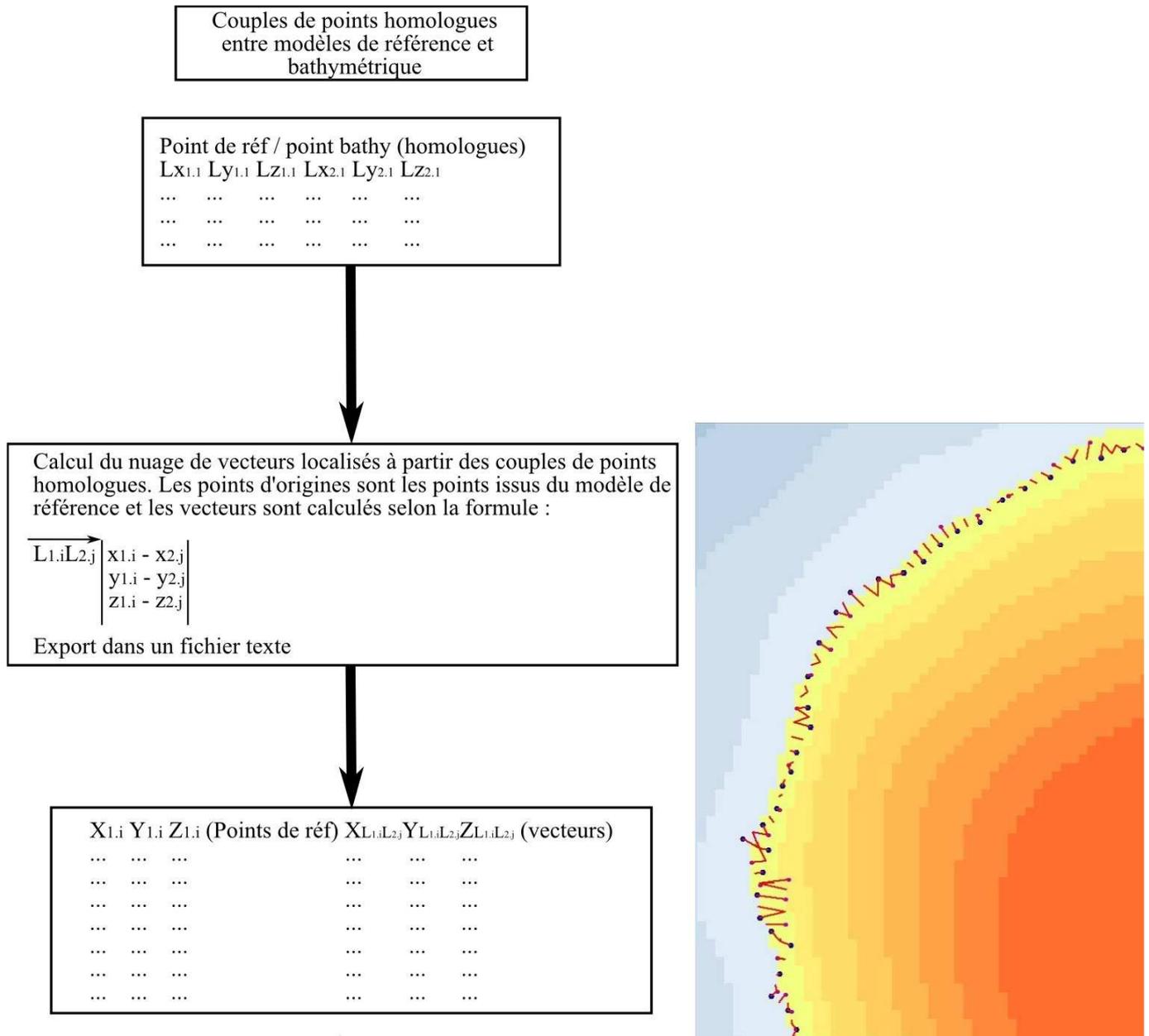
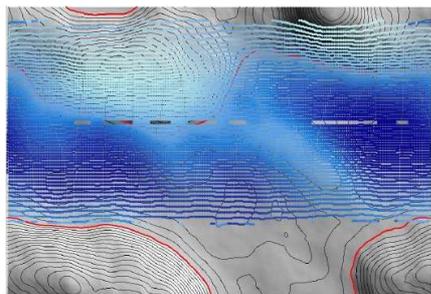


Figure 106 : Chaîne de traitement des données : calcul des déformations globale et locale.



Nuage de vecteurs localisés
 $\overrightarrow{Cph_{k,i}pf_k}$

Points de sonde

X	Y	Z
.....
.....
.....
.....
.....
.....

X _{1,i}	Y _{1,i}	Z _{1,i}	(Points P _i de réf)	X _{L_{1,i}L_{2,j}}	Y _{L_{1,i}L_{2,j}}	Z _{L_{1,i}L_{2,j}}	(vecteurs)
...	
...	
...	
...	
...	
...	
...	
...	
...	
...	

Calcul du nuage de points déformés élastiquement à partir de la formule :

$$\sum_{k=1 \dots n} \left(\overrightarrow{Cph_{k,i}pf_k} \times \frac{d_E(p_i, Cph_{k,i})^{-2}}{\sum_{l=1 \dots n} d_E(p_i, Cph_{k,i})^{-2}} \right)$$

avec $\overrightarrow{Cph_{k,i}pf_k}$: le nuage de vecteurs localisés calculés à partir des points homologues Cph_{k,i}

$d_E(p_i, Cph_{k,i})^{-2}$: distance entre le point P_i et son point homologue (Cph_{k,i})

$\sum_{l=1 \dots n} d_E(p_i, Cph_{k,i})^{-2}$: somme des distances

Obtention d'une grille de points déformés. Visualisation de la différence en chaque point entre le modèle de sonde et le modèle déformé.

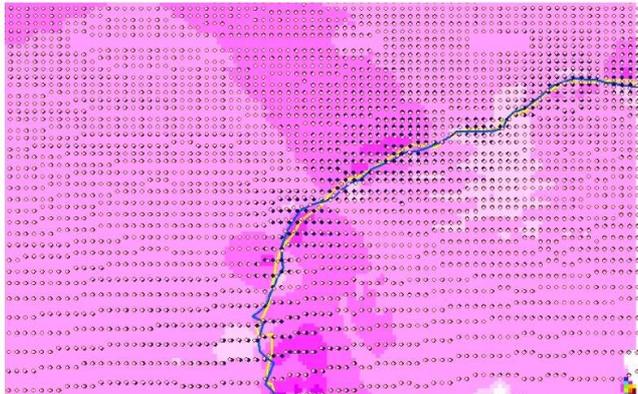
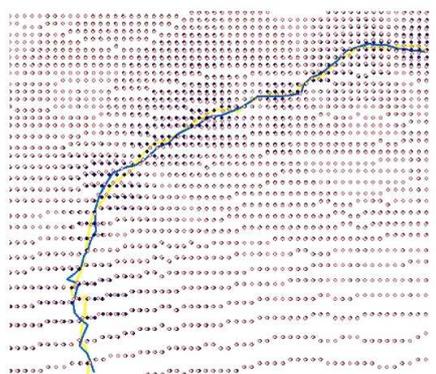


Figure 107 : Chaîne de traitement des données : déformations absolue et élastique des MNT. Visualisation du résultat dans un SIG.

Cartes de différences :

L'objectif ici n'est pas de se placer dans une optique de fusion mais bien de regarder l'impact des différents algorithmes de déformation sur des représentations plus proches de la réalité que celui du chenal ou de la déformation d'un modèle bruité sur le modèle d'origine. Cet impact est étudié au moyen d'un outil de soustraction pixel à pixel entre MNT. Il est à noter que cette méthode est rapide certes, mais elle ne met en valeur que les décalages en Z, et non ceux en X/Y. Elle constitue pour cette raison uniquement une première approximation du résultat.

A l'instar de la déformation précédente, le résultat obtenu sur les modèles DEG et TRANS ne présente pas de différences majeures sur le modèle topographique. L'élément important est que le modèle topographique est corrigé sur l'ensemble de sa surface de son biais global : d'une différence majoritairement comprise entre [-0,8 ; -1,5], l'erreur moyenne résultante a été ramenée à un intervalle de [0,1 ; - 0,3]. Le biais moyen obtenu lors de l'appariement des points homologues est de (-0,18, -1,4, -1,01). La correction rigide (TRANS) de tout le modèle avec ces valeurs corrige complètement l'erreur en Z (1), trop en Y (0,6) et pas assez en X (0,6).

Légende :

— ligne_iso0

Orig-topo TRANS

<VALUE>

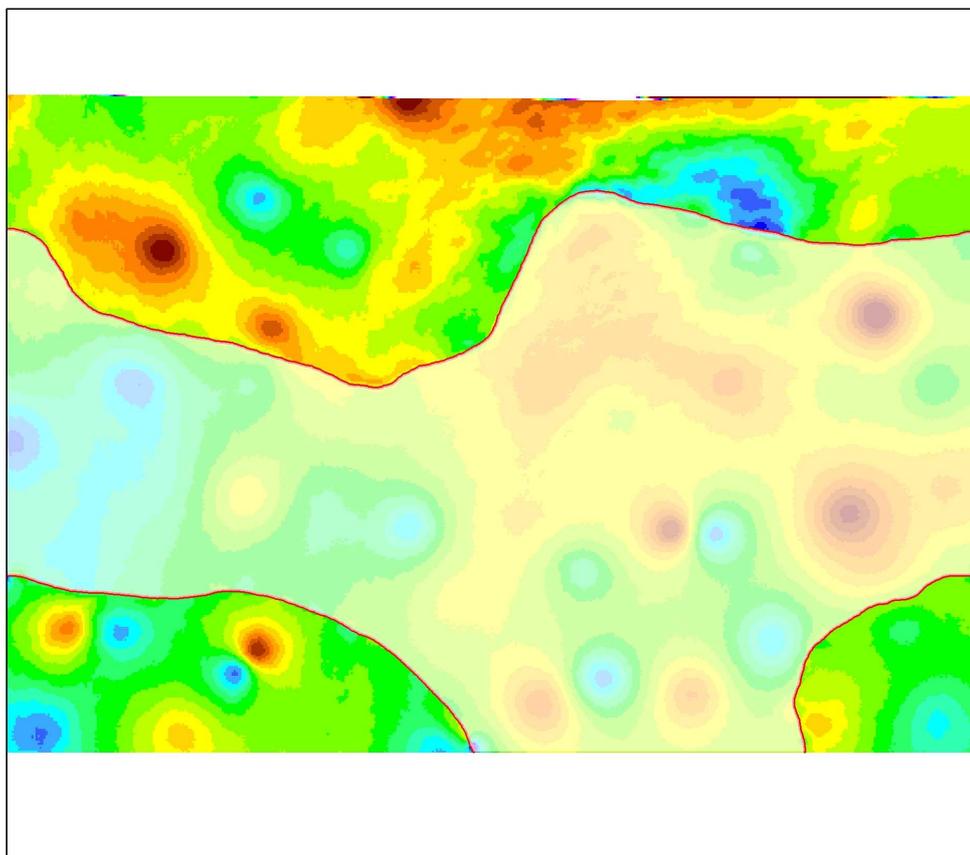
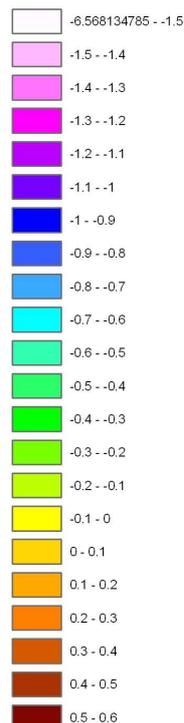


Figure 108 : Soustraction du modèle de référence avec le modèle topographique traduit.

Le modèle bathymétrique n'a pas été altéré par des corrections aussi fortes, car la pondération de 0,9 induit une très faible déformation par rapport au modèle topographique. Le résultat de la correction augmente les valeurs initialement comprises entre [+0,2 ; -0,1] à un intervalle de [0,1 ; -0,3] (figure 109) sur 80% du MNT. Néanmoins le modèle bathymétrique reste dans l'ensemble meilleur que le modèle topographique.

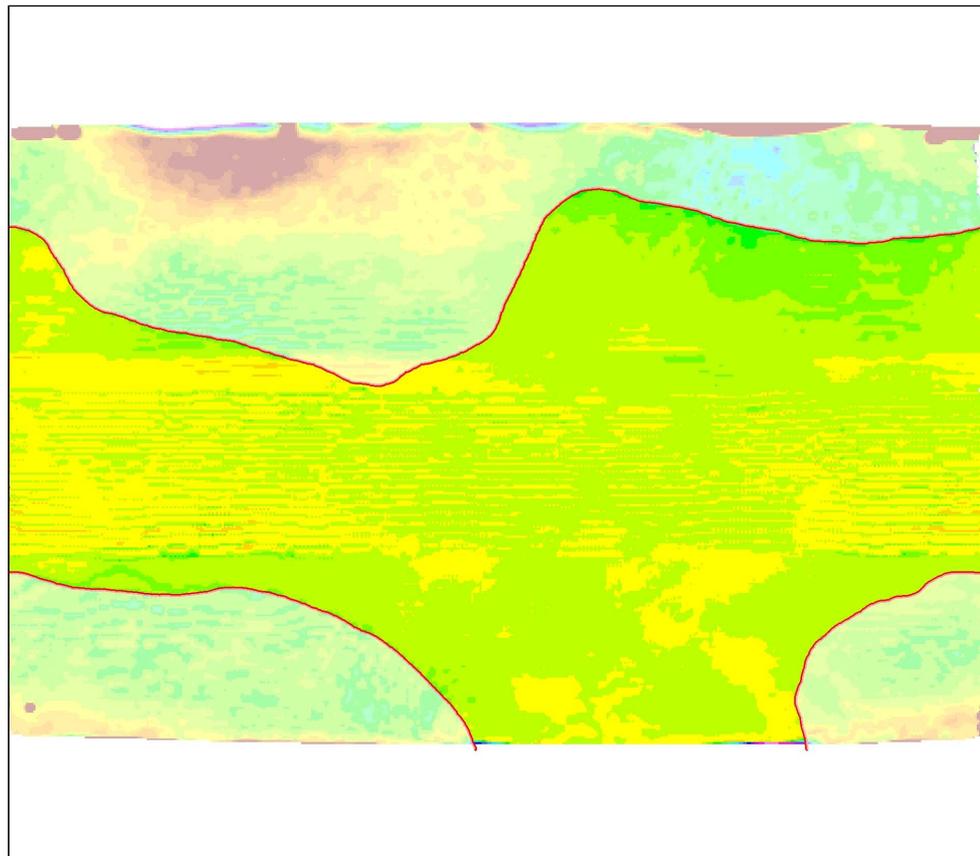
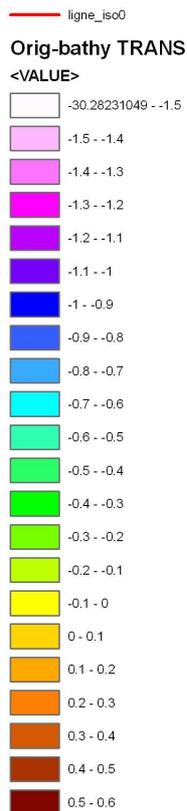
Légende :

Figure 109 : Soustraction du modèle de référence et du modèle bathymétrique translaté.

Les corrections du modèle topographique à l'aide d'une bande de déformation élastique (DEL8 et DEL32) présentent de meilleurs résultats d'un point de vue mise en pratique : le principe était de ne pas appliquer ces déformations en dehors du rayon de la bande de correction. Dans ce sens l'impératif a été scrupuleusement suivi, laissant le modèle inchangé en dehors de la bande concernée (figure 110 et figure 111) : de valeurs comprises entre -0,9 et -1,5, les points corrigés sont par la suite compris entre 0 et -0,5 sur une bande totale de 8m. Pour DEL32, la correction améliore des mêmes valeurs (figure 111) mais sur une bande de 32m. Le modèle bathymétrique subit lui aussi de faibles qui l'améliorent légèrement.

Les corrections TRANS+DEL8 et DEL32 présentent également des points importants : la translation présente l'avantage de corriger l'ensemble du modèle de manière assez correcte, tandis que la bande de déformation (exemple donné pour R=8, figure 113) corrige également les erreurs locales :

- en bas à droite un pic à 0,6 est ramené à 0 (zone 1) ;
- en haut à droite un pic à -1 est surélevé à -0,4 (zone 2).

Par contre en bas à gauche de la carte (zone 3), il apparaît que la déformation rigide par translation corrige mieux que la bande de déformation : la différence entre modèle initial et modèle déformé est de 0 en dehors de la bande et comprise entre -0,1 et -0,4 dans le rayon de correction.

Légende :

Orig-topo Loc8

<VALUE>

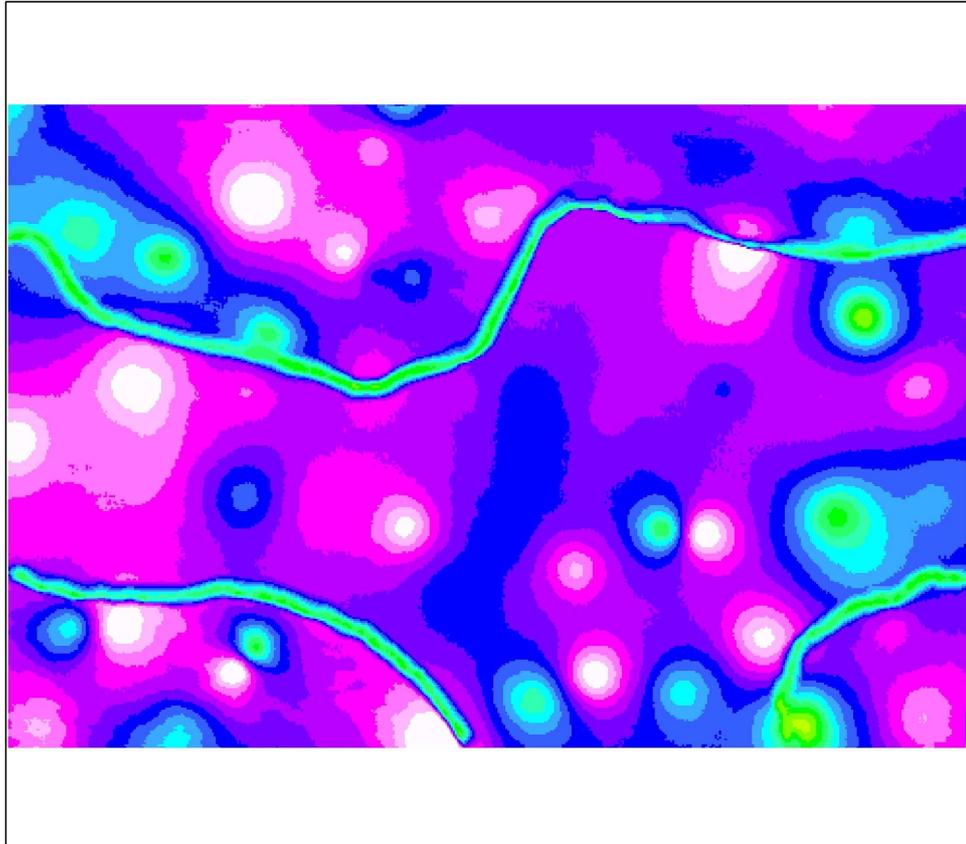
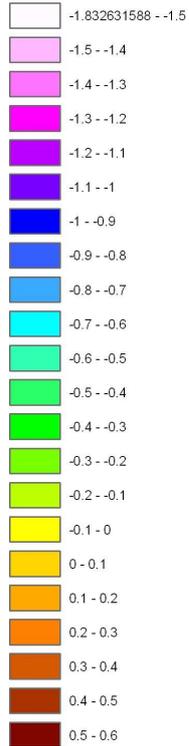


Figure 110 : Soustraction du modèle de référence avec le modèle topographique déformé élastiquement ($R=8$).

Légende :

Orig-topo Loc32

<VALUE>

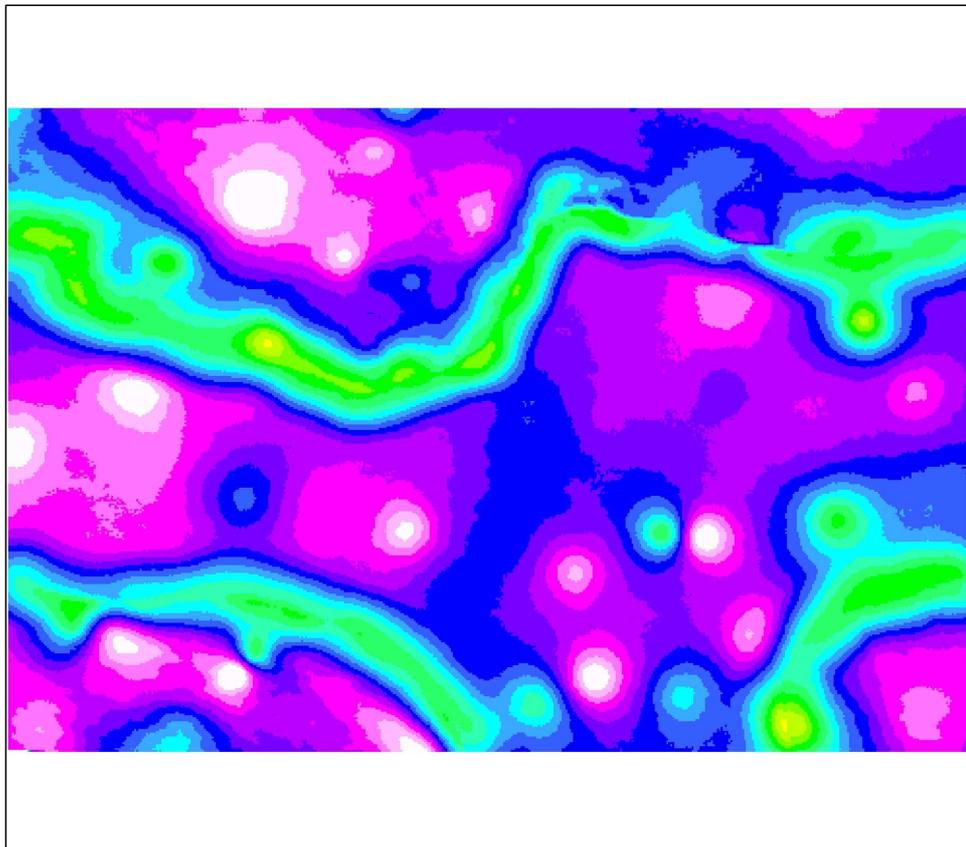
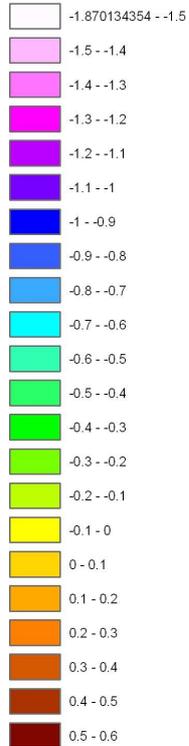


Figure 111 : Soustraction du modèle de référence avec le modèle topographique déformé élastiquement ($R=32$).

Légende :

Orig-bathy Loc32

<VALUE>

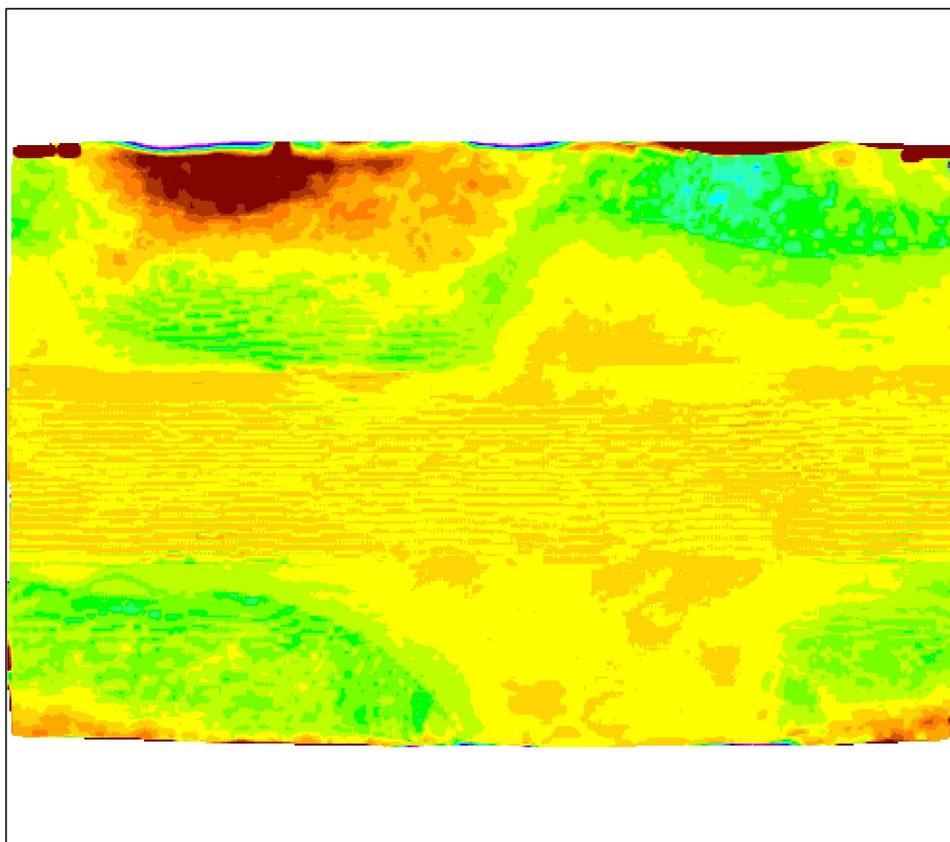
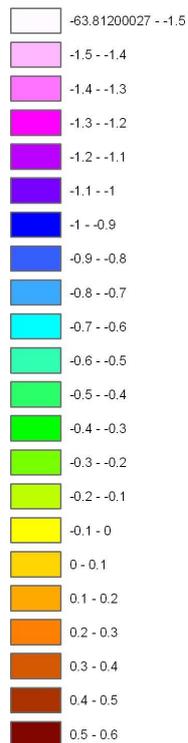


Figure 112 : Soustraction du modèle de référence avec le modèle bathymétrique déformé élastiquement avec un rayon de 32.

Légende :

Orig-topo 8 TRANS

<VALUE>

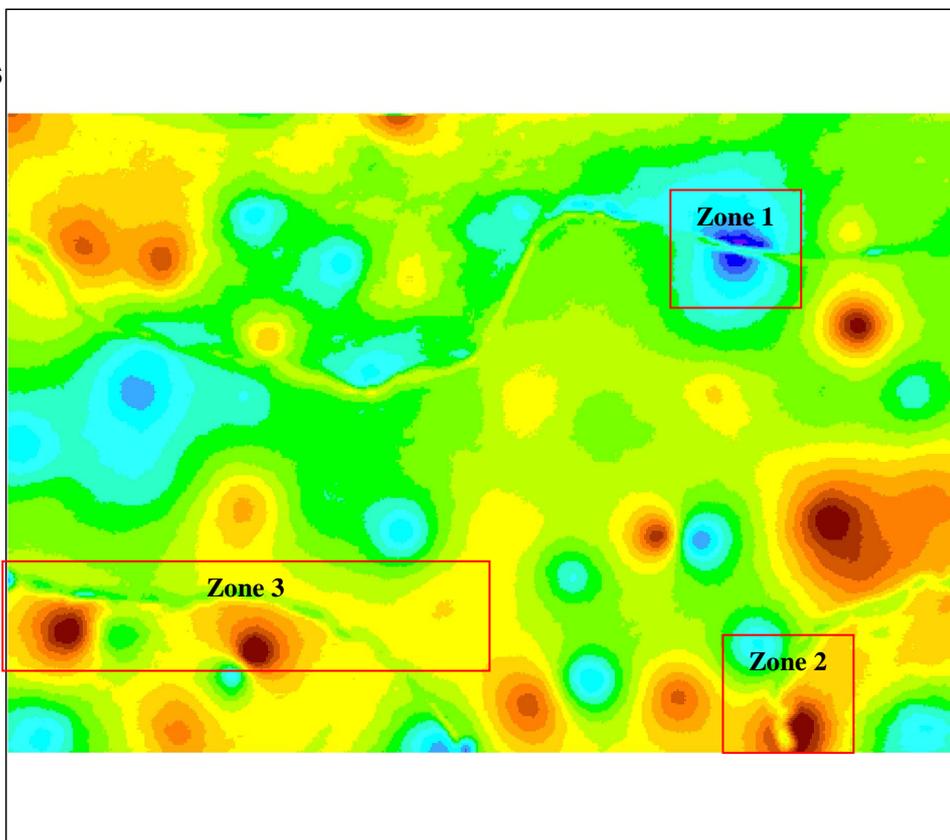
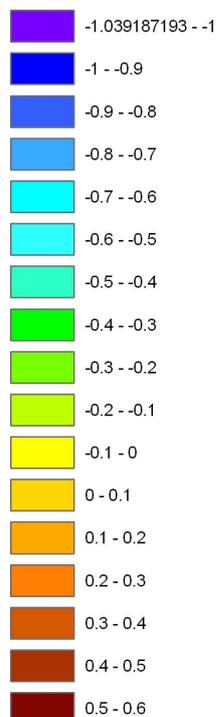


Figure 113 : Soustraction du modèle de référence avec le modèle topographique déformé élastiquement et translaté (R =8).

Modèles fusionnés :

Les modèles fusionnés terre/mer sont obtenus par extraction des points topographiques et bathymétriques sur la zone concernée avec un recouvrement de 5m (et seulement 2 pour les déformations de R=8). L'ensemble des points est ensuite interpolé en grille et en TIN de façon à visuellement apprécier le résultat final. Un profil, localisé figure 114, a été effectué pour chaque MNT obtenu de manière à comparer les déformations réalisées.

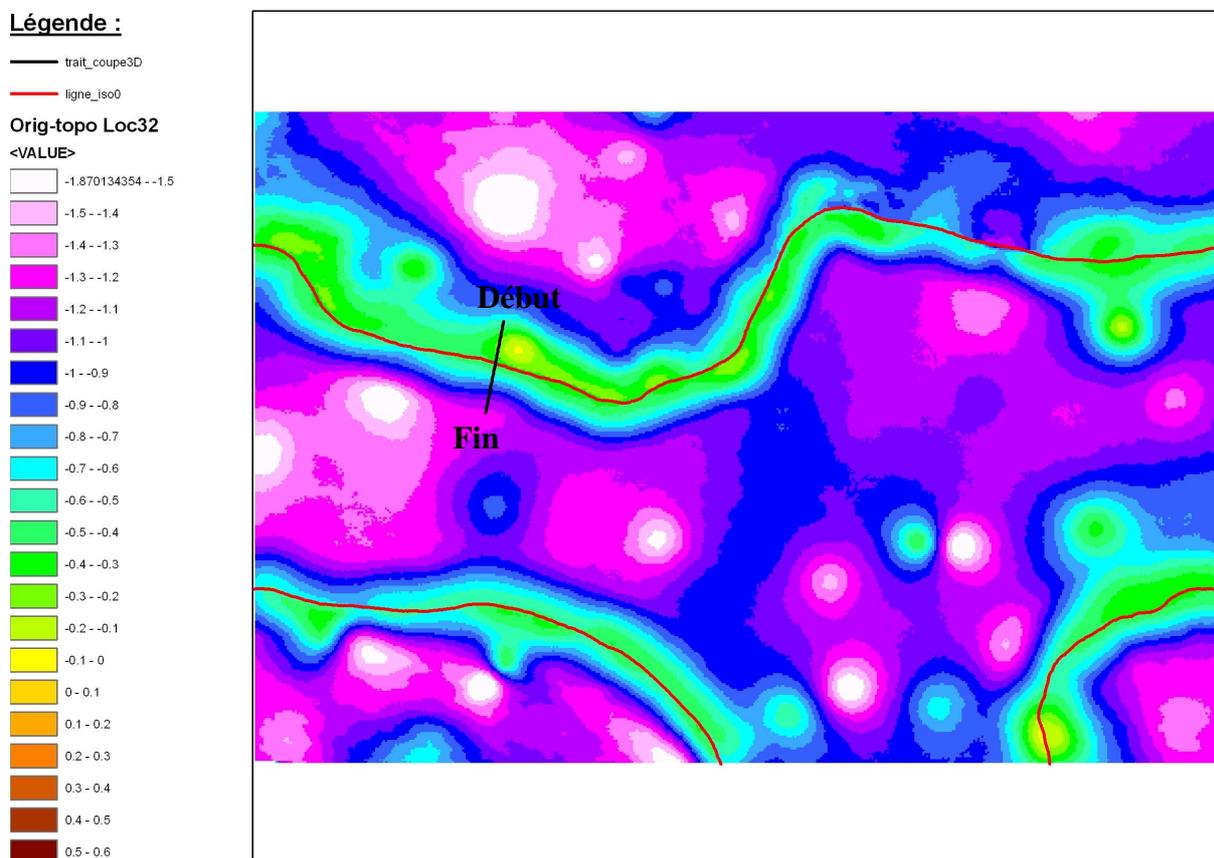


Figure 114 : Localisation du trait de coupe. La carte de soustraction entre le MNT de référence et le MNT DEL32 montre l'emplacement du trait vis-à-vis de la bande de déformation.

La figure 115 rappelle le positionnement des modèles bruités par rapport au modèle non bruité. Il apparaît très nettement que le MNT bathymétrique est confondu avec le modèle d'origine et se dégrade au niveau des altitudes comprises entre 0 à 8 mètres.

Le MNT topographique possède par contre une qualité mauvaise sur l'ensemble de la zone. La figure 116 montre les effets de la déformation avec bande de correction élastique de rayon égal à 8m : le modèle fusionné DEL8 se positionne au plus près du modèle bathymétrique et au modèle topographique et passe progressivement de l'un à l'autre dans la bande de correction élastique (pour une distance comprise entre 15 et 30 sur l'axe des x). La transition entre les deux modèles se fait néanmoins par un « ressaut » dans le modèle fusionné en raison des points de recouvrement.

En effet si la surface commune entre les points terrestres et maritimes est trop grande, il apparaît un artéfact en forme de marche d'escalier : la figure 117 montre les points topographiques (en marron sur la figure) non corrigés car à l'extérieur de la bande de correction (R=8) et possédant le biais vertical de 1m par rapport au points bathymétriques. L'exagération verticale des modèles sur les deux figures montre bien la présence d'un palier puis d'une brusque rupture de pente (figure 118) juste après la première ligne de rupture de

penne. Le choix d'un buffer de recouvrement de deux mètres entre les données minimise cet effet.

Comparaison des profils des MNT fusionnés

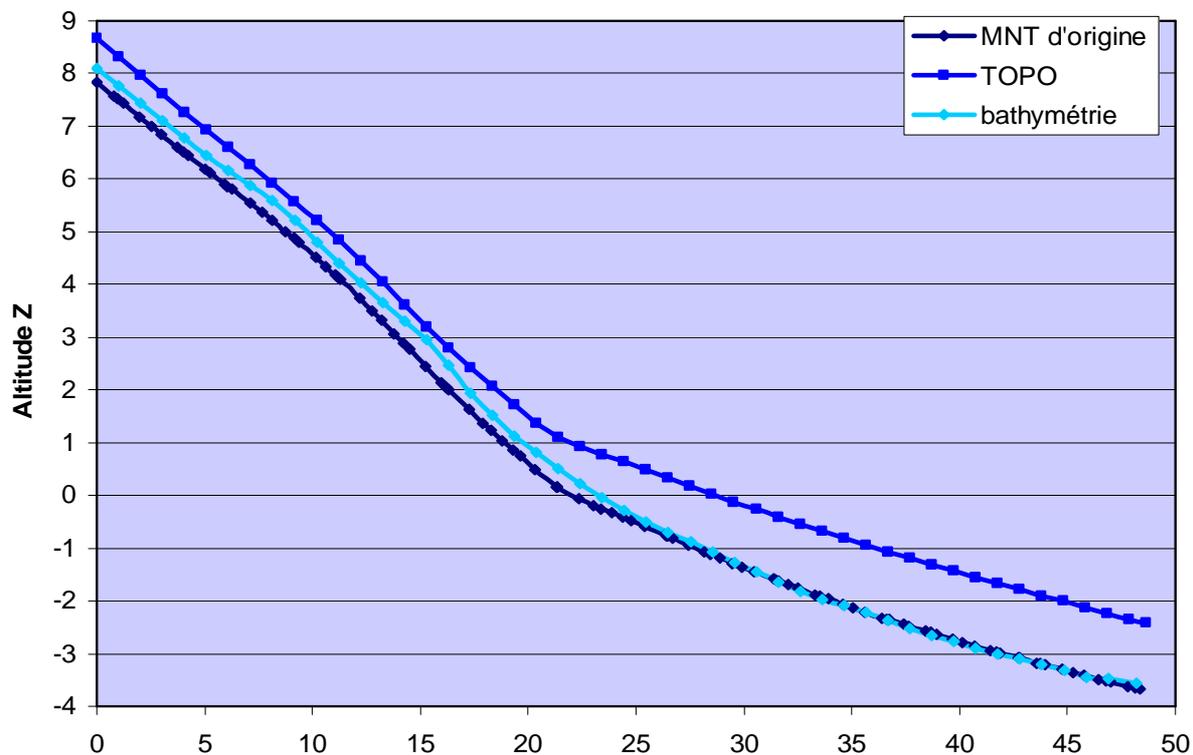


Figure 115 : Comparaison des profils du MNT de référence et des deux MNT extraits, topographique et bathymétrique.

Comparaison des profils des MNT fusionnés

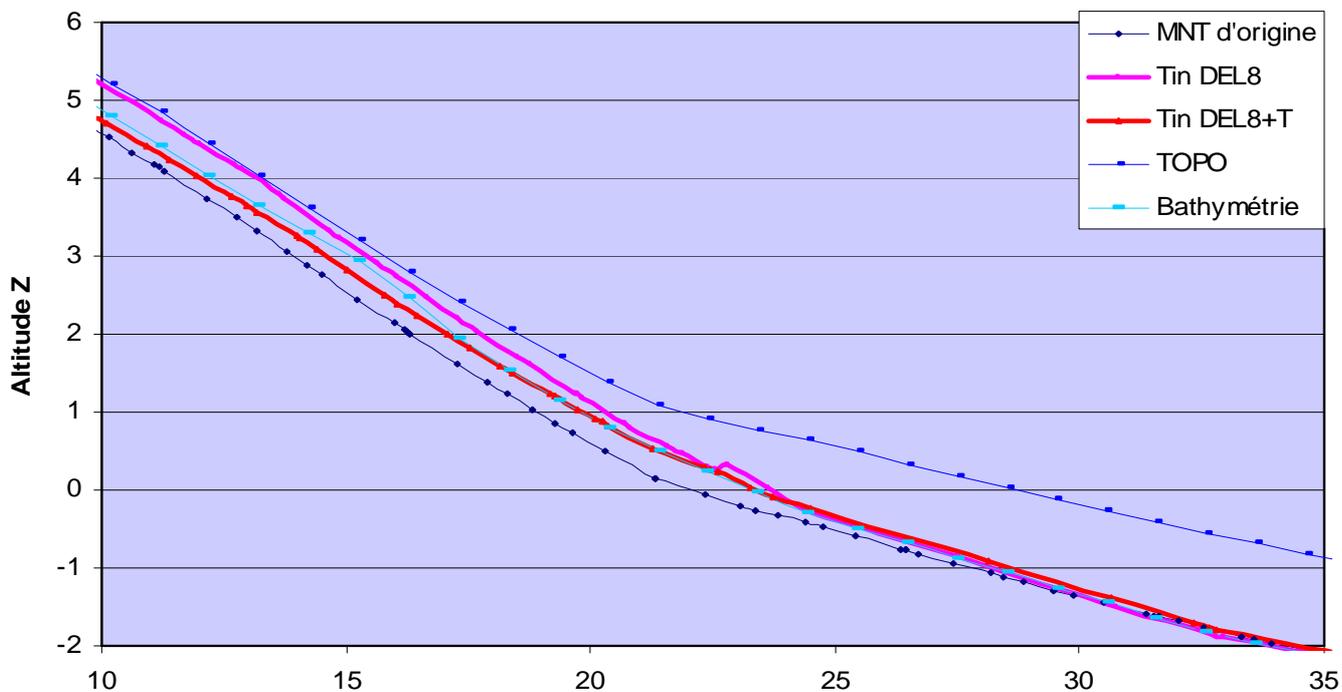


Figure 116 : Comparaison des modèles DEL8 et DEL8+T avec les modèles de référence, bathymétrique et topographique.

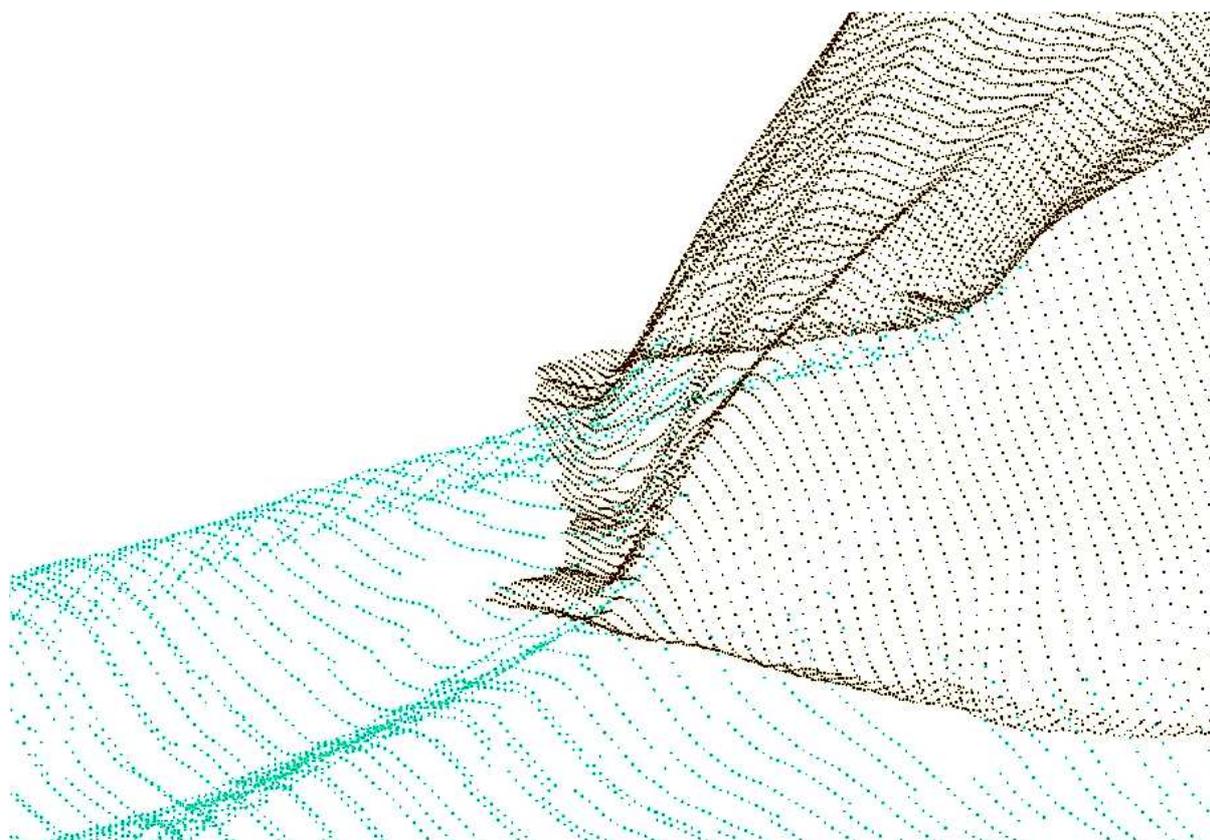


Figure 117 : Superposition du modèle topographique en marron et du modèle bathymétrique en bleu. L'insuffisance de la correction du modèle topographique induit un biais altitudinal important visible sur l'image. Altitudes exagérées 5 fois.

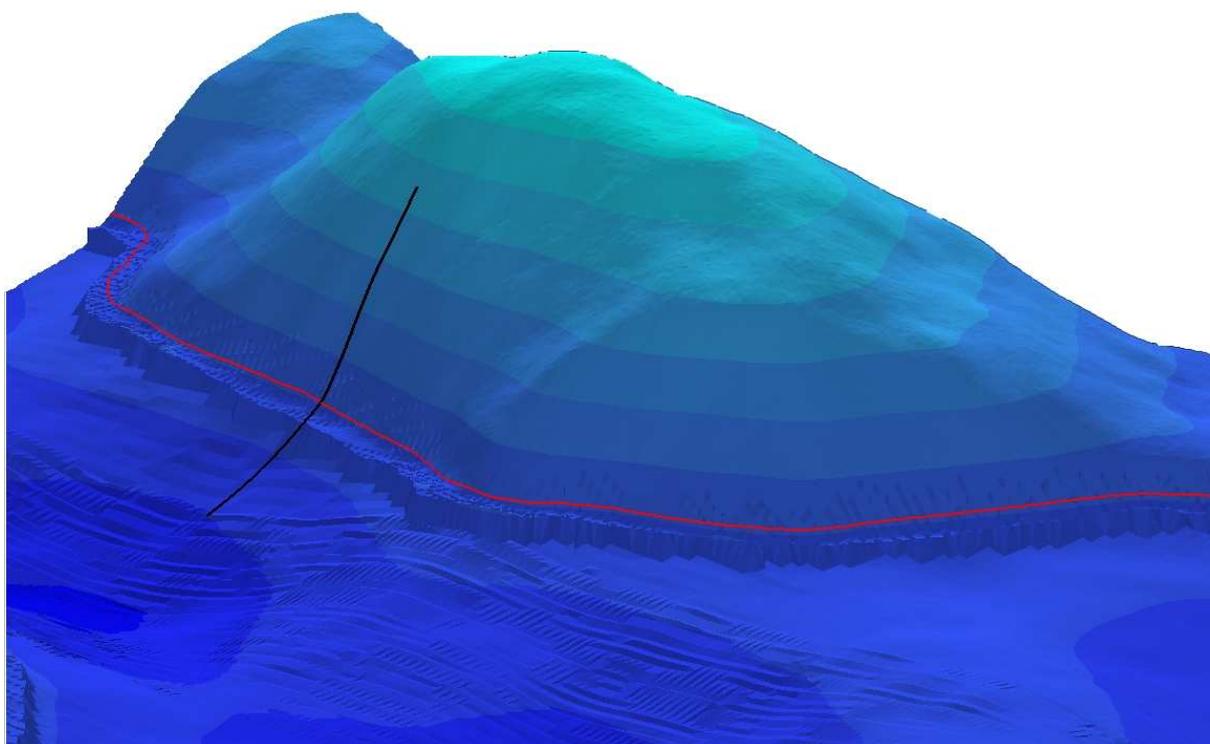


Figure 118 : Visualisation du MNT terre/mer issu de la déformation DEL8 des deux modèles terrestre et maritime bruités. Altitudes exagérées 5 fois. Le trait noir représente le trait de coupe.

Comparaison des profils des MNT fusionnés

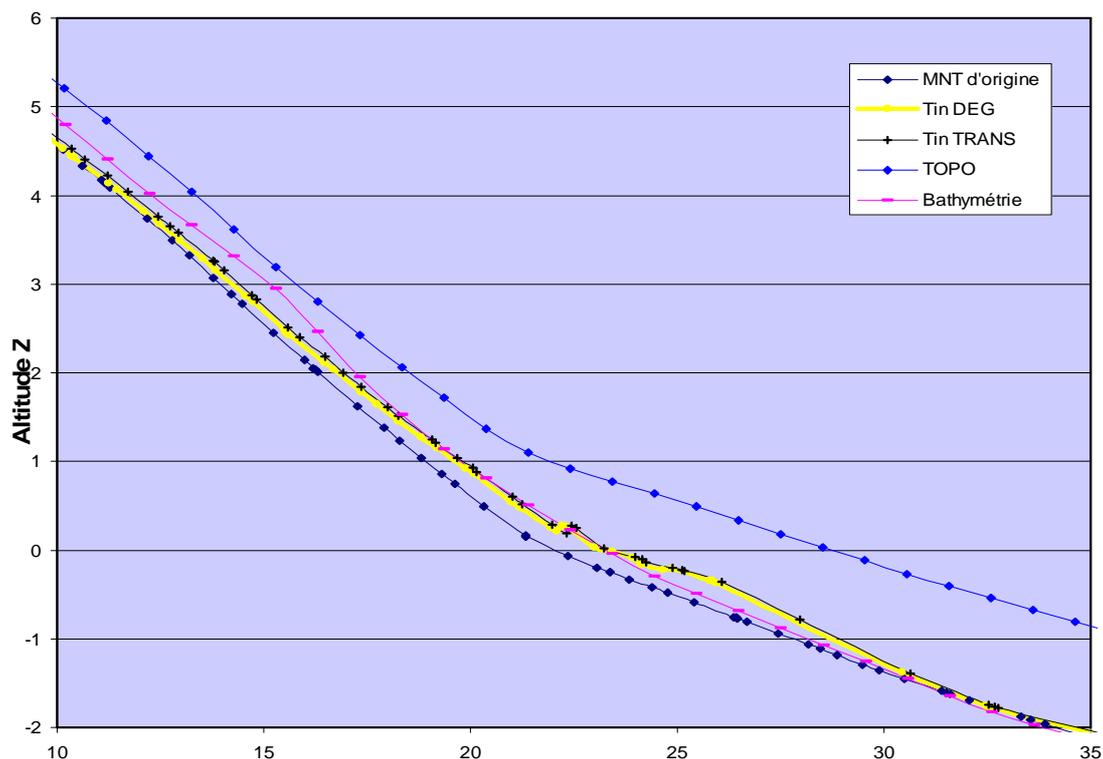


Figure 119 : Comparaison des modèles DEG et TRANS avec les modèles de référence, bathymétrique et topographique.

Comparaison des profils des MNT fusionnés

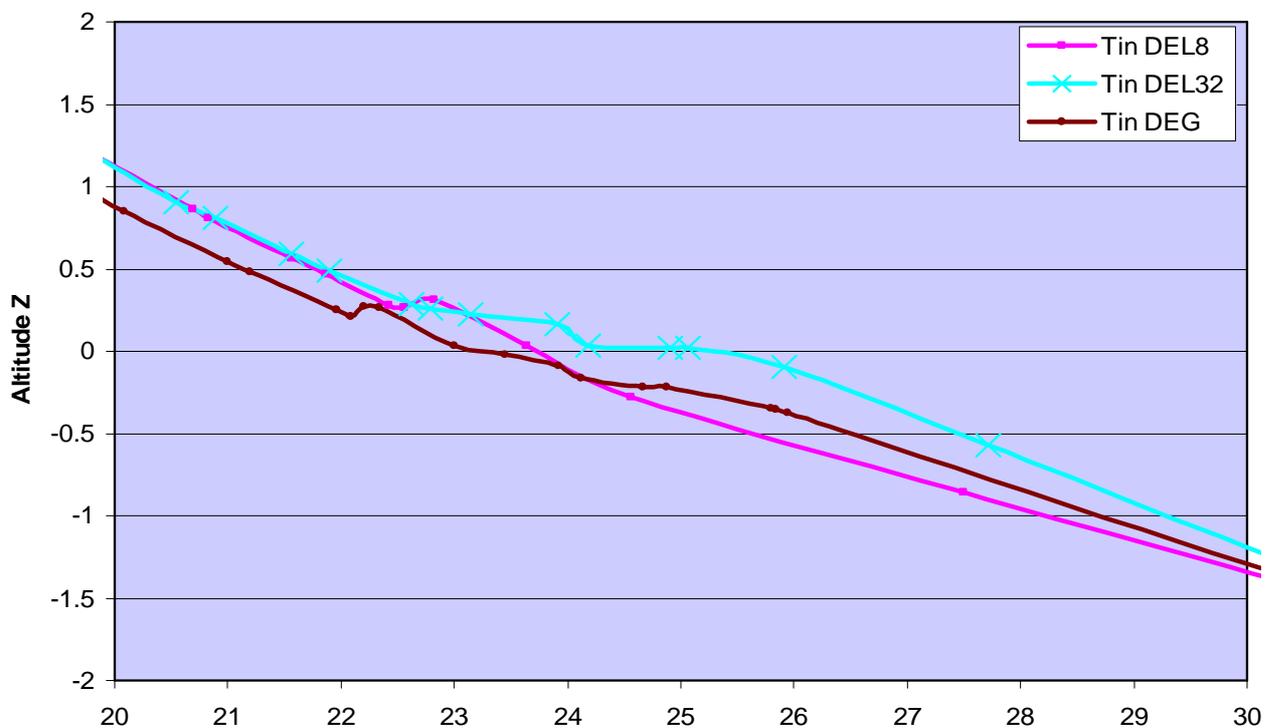


Figure 120 : Comparaison de l'artéfact des modèles fusionnés DEL8, DEL32 et DEG.

Comparaison des profils des MNT fusionnés

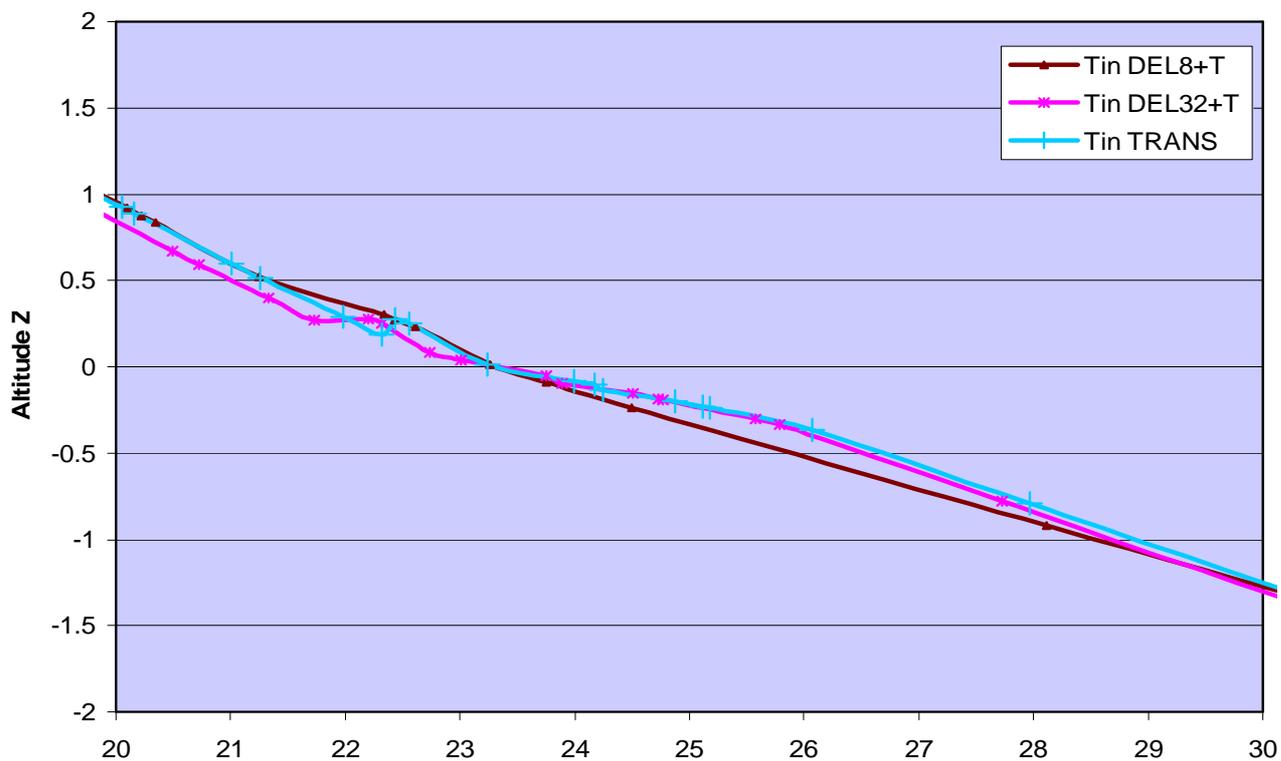


Figure 121 : Comparaison de l'artéfact des modèles fusionnés DEL8+T, DEL32+T et TRANS.

Comparaison des MNTs fusionnés

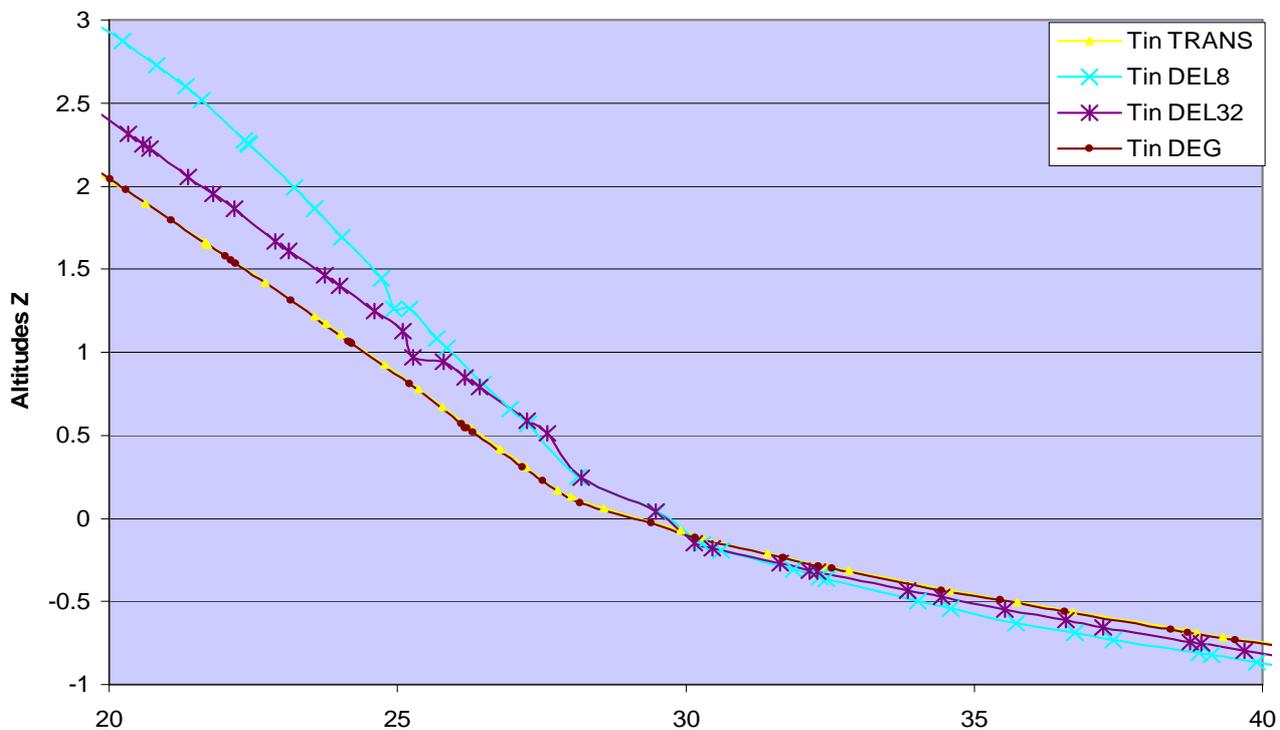


Figure 122 : Comparaison de l'artéfact sur les modèles fusionnés.

Les modèles corrigés figure 119 possèdent un passage terre/mer qui comporte un artéfact de pente bien moindre mais néanmoins présent.

Les figure 120 et figure 121 se focalisent sur les déformations engendrées autour de la ligne caractéristique : dans tous les cas la présence de points issus du buffer induit des inversions de pente ou bien une double rupture de pente qui ne correspondent à rien géomorphologiquement. C'est pourquoi le choix d'une zone de recouvrement plus petite de 0,5 m pour la bathymétrie et nul pour la topographie minimise l'artéfact (figure 122) : les seuls modèles possédant encore cet effet sont ceux qui n'ont pas été corrigés sur leur ensemble (DEL8 et DEL32) et gardent donc une erreur générale. Ceci implique que lors du processus de déformation aval, le rayon de correction doit être impérativement choisi en fonction de la zone de recouvrement existante entre les données (ou une moyenne si celle-ci n'est pas constante, ce qui arrive dans la majorité des cas). Pour que la déformation n'entraîne pas d'artéfact, le rayon choisi doit être **au moins deux fois supérieur** à la zone de recouvrement entre les données. Si cette zone n'est pas constante, les données ayant le moins de recouvrement imposent le choix du rayon.

Néanmoins, ce problème d'artéfact peut également venir de l'imprécision des lignes caractéristiques car il n'apparaît pas dans l'étude du cas précédent : déformation d'un modèle bruité sur un modèle d'origine

Au vu des résultats obtenus entre une fusion MNT de référence/MNT bruité et une fusion Terre/Mer, il est possible de valider l'algorithme et les méthodes de fusion testées :

- les corrections sont bien effectuées selon les spécifications : il n'y a pas de modification du modèle en dehors de la bande de déformation si cela n'est pas souhaité ;
- le biais calculé pour la déformation rigide du modèle (TRANS) est relativement proche (moyennant le bruit blanc et les erreurs grossières locales) du biais inséré pour bruiteur le modèle : toute correction entraîne donc une amélioration notable des modèles ;
- la correction élastique de l'ensemble du modèle (DEG) montre des résultats également satisfaisants ;
- la présence d'artéfacts sur la bande de fusion (inversion de pente et présence de replat dans le premier test, apparition d'une seconde ligne de rupture de pente dans le second) est due : soit à un cas particulier non réaliste (altitude des données bathymétriques supérieure à l'altitude des données topographiques), soit à un mauvais positionnement des lignes caractéristiques extraites.

III.4 Perspectives d'application de la méthodologie sur des données littorales

La méthodologie telle que décrite précédemment est très développée afin de résoudre le plus de problèmes possibles mais de ce fait très complexe. C'est pourquoi la totalité de ses modules n'ont pas pu être implémentés ni testés sur des données géographiques réelles issues soit de programmes comme Litto3D, soit d'acquisitions ponctuelles. En reprenant étape par étape le déroulement des processus d'intégration, voici le bilan de ce qu'il est possible d'appliquer actuellement.

III.4.1 Conversion géographique

Un changement de système de coordonnées dans un même référentiel géographique n'est pas le point qui pose le plus de problèmes bien qu'il soit possible d'être confronté à des cas particuliers assez embrassants.

Dans le cadre du LIDAR topographique acquis en Lambert 93, qui est un système projeté, et le Lidar topo-bathymétrique acquis en WGS84, il faut passer alors d'une hauteur ellipsoïdale à une altitude exprimée par rapport au marégraphe de Marseille. La conversion directe entre les deux systèmes n'est pas possible avec un logiciel comme CIRCE¹⁶ ou ArcGIS : un est projeté et l'autre pas. CIRCE impose de convertir toutes les données d'abord en Lambert II étendu pour les retransformer à nouveau en Lambert 93. Cette manipulation impose des temps de calcul très longs sur des MNT comportant plusieurs millions de points.

Un des choix possible est de regarder la différence altimétrique entre la cote du zéro hydrographique par rapport au repère fondamental dans le port de référence le plus proche (Brest ou Roscoff) [Wöppelmann et al., 1999]. Ensuite vérifier la hauteur d'eau au moment de l'acquisition et convertir les données bathymétriques de hauteur d'eau en altitude. Il est à noter que comme les ports de référence ne font pas partie du même système hydrodynamique, la correction ne peut se faire qu'avec un modèle de marée adapté à la zone. Il apparaît rapidement que ces corrections sont envisageables mais fastidieuses et sujettes à de nombreuses erreurs pour un utilisateur non familiarisé avec les différents référentiels.

III.4.2 Partition en domaines de recouvrement

Cette phase de la méthodologie n'a pas été abordée lors des tests avec les données simulées car cela aurait mis un frein aux tests des algorithmes d'intégration. En effet aucun logiciel ne permet pour le moment de comparer deux MNT afin d'en produire une partition selon le triptyque recouvrement/lacune/hors recouvrement. Pour pouvoir être appliquée, cette partie de la méthodologie doit être implémentée à la manière de l'addon codé par les élèves de l'école navale pour le SIG ArcGIS [Kervazo & Moutoulatchimy, 2006], ce qui signifie surmonter les problèmes liés aux données et à leur conversion : extensions spatiales et formats différents, variabilité temporelle, densités de données dissemblables liées aux techniques d'acquisition, etc.

III.4.3 Analyse de corrélation

Cette partie est applicable. Elle a été développée entre autre par [Guth, 2006] dans son logiciel MicroDEM via le calcul du coefficient de corrélation entre pentes moyennées de deux MNT à comparer. La correction des données s'effectue sur la planimétrie.

III.4.4 Analyse paysagère

La phase de segmentation et d'extraction de ligne caractéristique de la méthodologie peut être complétée par une analyse paysagère sur les zones de recouvrement entre modèles. Ces

¹⁶ <http://professionnels.ign.fr/ficheProduitCMS.do?idDoc=5352513>

traitements doivent se réaliser sur les zones des MNT qui ont montré une bonne corrélation suite à la première étape de la méthodologie et qui ne comportent pas ou peu de trous.

Il s'agirait dans un premier temps de lancer un (ou plusieurs) algorithme(s) de reconnaissance/extraction paysagère (extraction et pré-enrichissement, figure 123) à l'aide d'une typologie paysagère littorale à définir. Cette étape de reconnaissance des éléments caractéristiques morphologiques permet une classification du MNT selon les types paysagers (segmentation) et la mise en valeur des caractères qui leur sont propres (enrichissement).

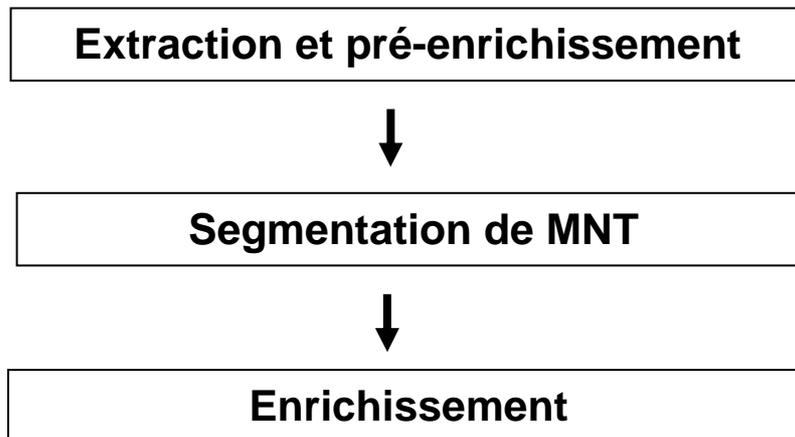


Figure 123 : Détail de la seconde phase : analyse paysagère en vue de l'extraction des lignes nécessaires au processus d'appariement.

Les études les plus exhaustives concernant la géomorphologie littorale montrent une très forte corrélation des formes paysagères présentes avec la latitude géographique [Valadas, 2005]. Ainsi la prise en compte de la localisation géographique d'un espace littoral peut être utilisée pour éviter la recherche d'éléments morphologiques non significatifs d'un environnement donné.

Cette analyse paysagère et la différenciation des traitements appliqués selon les zones identifiées à sa suite, confèrent à cette méthodologie un caractère adaptatif selon le type de relief traité, tout en restant générique.

III.4.4.1 Extraction et pré-enrichissement

Le pré-enrichissement est effectué en recherchant parmi tous les éléments de relief, ceux considérés comme les plus « pertinents » mais aussi les plus communs comme les crêtes, les thalwegs, les points hauts (sommets) et les points bas (cuvettes).

Ils sont extraits avec des paramètres très contraints pour éviter les artefacts et sont désignés en tant que support clé de l'information paysagère pour les étapes suivantes. Certains outils déjà disponibles sur des SIG comme SIS¹⁷ ou ArcGIS¹⁸ permettent déjà d'effectuer une détection de ce type d'éléments.

Tous ces éléments caractéristiques sont par la suite mis en valeur dans chacun des modèles (bathymétrique et topographique) de façon à servir de base pour le processus de segmentation.

¹⁷ <http://www.geomod.fr/logiciels/sis/sis.htm>

¹⁸ <http://www.esrifrance.fr/>

III.4.4.2 Segmentation

Cette étape vise à classer le relief étudié selon des tendances littorales : falaise, plage, zones portuaires, etc. Les classes sont à définir de manière exhaustive pour tout type de paysages littoraux présents sur terre.

III.4.4.3 Enrichissement

Après la segmentation et la répartition du relief en classes de paysage, il faut affiner la détection des éléments caractéristiques, complémentaires à ceux mis en valeur lors de la phase de pré-enrichissement. L'étape permet de ne rechercher parmi les catégories paysagères que les éléments morphologiques qui y sont effectivement présents.

L'objectif de l'enrichissement est d'utiliser ces éléments comme points d'appui essentiels pour le processus d'intégration. Cet enchaînement « segmentation/enrichissement » permet leur maintien au cours des traitements et ainsi la sauvegarde, lors de la fusion, d'informations paysagère essentielles.

L'enrichissement peut être effectué de plusieurs manières. Soit il s'agit d'intégrer au MNT existant d'autres données, soit de chercher à reconnaître des éléments caractéristiques.

Dans le premier cas il existe de nombreuses bases de données d'occupation du sol comportant des informations complémentaires au relief brut : les routes, le bâti, le réseau hydrologique ou divers autres équipements humains comme les grands ouvrages. Les bases de données vectorielles sur le territoire Français sont essentiellement produites par l'IGN :

- BD CARTO®, qui intègre les limites administratives, le réseau routier, ferré et électrique, l'hydrographie, l'occupation du sol à une précision de 50m, et la toponymie ;
- BD TOPO®, comprenant le relief, les points géodésiques, la végétation et divers bâtiments ;
- BD CARTHAGE®, focalisée sur les cours d'eau et eaux de surface à petites échelles ;

Et en partenariat avec le SHOM :

- HistoLitt®, qui représente une BD altimétrique « historique ».

Il existe également au niveau mondial différents types de bases de données, gratuites¹⁹ ou non, de précisions très variables qui permettent d'effectuer certains des enrichissements utiles. [Rousseaux, 2005] donne un exemple de tels enrichissements pour produire un MNT « réaliste et cohérent » de manière à optimiser les résultats obtenus sur une simulation de risques. [Podobnikar, 2005] utilise également cette méthode pour produire un MNT intégré de la Slovaquie.

Dans le second cas, les mesures et éléments géomorphométriques à identifier sont choisis selon la méthode de reconnaissance ou le logiciel utilisé. En effet certains outils de détection sont déjà disponibles sur des SIG comme SIS²⁰, ArcGIS²¹ ou MicroDEM²² ([Hengl & Reuter, 2008]).

¹⁹ http://www.legos.obs-mip.fr/fr/equipes/gohs/resultats/i_hydroweb

²⁰ <http://www.geomod.fr/logiciels/sis/sis.htm>

²¹ <http://www.esrifrance.fr/>

²² <http://www.usna.edu/Users/oceano/pguth/website/microdem.htm>

III.4.5 Appariement

Tous les outils d'appariement ont été implémentés. Les différents cas ont été envisagés grâce à la participation au projet d'évaluation des données SPOT pour la mise à jour du trait de côte, ce qui a permis des extensions conséquentes afin de prendre en compte le maximum de cas possibles (voir chapitre IV).

III.4.6 Corrections élastiques

D'autres modèles de corrections, qu'elles soient élastiques ou non, peuvent être envisagés pour l'intégration de modèles numériques en milieu littoral. Cela demande de plus amples études afin de déterminer quels algorithmes s'adaptent au mieux au paysage traité : plages, rocheuses, marais ou falaises par exemple. Pour des zones mal définies ou non classifiables automatiquement, il est impératif de tester différents algorithmes de déformation afin d'évaluer le degré de fiabilité du modèle obtenu. Pour ces cas, il est également envisageable d'utiliser avant toute déformation des algorithmes de rehaussement de contours (REF).

Le déformation à l'aide de coefficients 1/0 dans le cas de données de qualité très nettement supérieure vis-à-vis d'une ou des autre(s) données comparées, permet de conserver la totalité de la rugosité du modèle considéré de référence.

III.5 Bilan sur la méthodologie d'intégration

Ce chapitre présente la méthodologie d'intégration mise en œuvre afin de réaliser une fusion de données Terre-Mer. Différentes zones bretonnes où de nombreuses données disponibles permettaient cette étude ont été présentées. Devant la complexité des traitements nécessaires, le volume des données et le temps de calcul pour des opérations qui se sont révélées trop ambitieuses dans le cadre d'une seule thèse, des données simulées leur ont été préférées. Ceci a permis de valider l'enchaînement des étapes d'appariement/déformation élastique et fusion de la méthodologie mise en place.

Néanmoins, la méthodologie d'intégration a été définie de façon à prendre en compte les divers problèmes dus à la structure des données réelles et solutionner les difficultés de mise en cohérence de l'information. C'est pourquoi toute la première phase des manipulations ne se concentre que sur la préparation des données topo-bathymétriques à la fusion future : recalage géographique, mise en cohérence et étude de la répartition géostatistique de l'information au sein des données.

La seconde phase prépare au processus d'appariement de MNT : une étude numérique poussée des modèles doit mettre en valeur les éléments caractéristiques du relief analysé, support majeur de l'information paysagère. La segmentation a pour but de séparer les différents types de reliefs étudiés et d'adapter les traitements ultérieurs.

Une troisième phase applique différents types de déformation selon les valeurs de correction calculées et les catégories de terrains identifiées lors de la segmentation. Ces déformations sont soit rigides, dans le cas où un MNT nécessite une correction globale, soit élastiques pour lisser localement à l'aide d'une bande de déformation la frontière de fusion entre les modèles. Une combinaison de ces deux déformations reste également à tester sur l'ensemble des

classes de relief afin d'évaluer l'impact des corrections effectuées sur la conservation des éléments paysagers qui leurs sont propres.

Les essais n'ont été effectués que sur des données simulées car la mise en oeuvre de la méthodologie a souffert de nombreuses difficultés pour pouvoir permettre une application initialement prévue sur des données réelles :

- données trop volumineuses ;
- recouvrement insuffisant ;
- décalage non constant entre les systèmes cartographiques utilisés ;
- résolutions différentes des données délicates à gérer ;
- logiciel non adapté aux outils nécessaires à l'application de la méthodologie complète ;
- données trop complexes.

Les différents tests effectués montrent que la qualité du MNT final obtenu est considérablement dépendante des lignes caractéristiques à partir desquelles sont calculés les points homologues ainsi que des poids affectés à chaque donnée pour la déformation. La première phase de préparation et d'analyse qualitative des données est donc capitale pour obtenir un résultat satisfaisant. Dans les essais effectués, les lignes caractéristiques ont été numérisées manuellement. De ce fait les résultats sont extrêmement liés à la qualité de numérisation et des éventuelles erreurs d'interprétation de l'opérateur. Dans le cas de l'utilisation d'un logiciel d'extraction (SIS/Bathysis® pour les lignes de crêtes et thalwegs, encore en version bêta, ou le logiciel développé à l'IRENav), les lignes obtenues sont également très dépendantes de la sensibilité de l'outil pour la détection des lignes remarquables.

L'application de l'ensemble de la méthodologie nécessite des outils non encore disponibles ou en cours d'amélioration. Ceci constitue une voie d'approfondissement pour des développements ultérieurs.

CONCLUSION GENERALE

Cette thèse s'inscrit dans la problématique de l'intégration de Modèles Numériques de Terrain (MNT) appliquée à l'interface Terre/Mer. Le littoral est un milieu très sensible d'un point de vue environnemental et qui a fait l'objet de nombreux aménagements par l'homme. Dès lors qu'un écosystème concentre des enjeux environnementaux et économiques importants, il est nécessaire d'avoir un renseignement continu du relief au travers des milieux terrestre et maritime. Or ce suivi est rendu délicat de part la difficulté d'acquérir des données dans une interface dynamique soumise au phénomène de marée avec de faibles tranches d'eau, ce qui explique que ce milieu littoral ait été longtemps mal renseigné. Avec le développement de nouvelles technologies d'acquisition Lidar aéroporté et grâce à l'amélioration des levés sonar multifaisceaux, il a été possible de compléter les couvertures topo-bathymétrique manquantes et de mieux échantillonner les valeurs altimétriques. Cependant cet essor des modes d'acquisition de l'information a créé un nouveau besoin, celui d'intégrer des données multi sources en prenant en compte les spécificités des données et les attentes des acteurs du littoral.

L'objectif de cette thèse a donc été de mettre au point une méthode d'intégration globale des différents levés topo bathymétriques afin d'obtenir un Modèle Numérique de Terrain Intégré (MNTI) terre - mer continu.

L'état de l'art effectué pour le présent travail a montré qu'il existait des outils très performants de mesure de similitude et d'établissement de correspondance entre deux objets géographiques (ponctuels, linéaires ou surfaciques), mais uniquement à deux dimensions. Pour des données plus complexes telles que des données 3D (ou 2.5D pour la majorité des SIG), ces outils ne sont pas ou peu développés en raison notamment de l'absence de tables attributaires. Toute opération se fera donc uniquement sur le positionnement géographique. Une intégration se simplifie donc la plupart du temps à une mise en cohérence des données (référencement géographique, conversion de format) avant une moyenne, pondérée ou non, des points de mêmes coordonnées X,Y.

Nous avons également vu dans la littérature que de nombreux processus d'intégration de MNT ont été mis au point pour répondre à un grand nombre d'applications du domaine côtier mais ceux-ci sont rarement formalisés. En effet la fusion de deux MNT est faite le plus souvent de manière adaptative selon le contexte du programme de recherche. En complément de ces travaux une étude de qualité du modèle final obtenu est souvent réalisée soit visuellement, soit statistiquement. Cette étude permet de vérifier que ce modèle est conforme aux requis de l'application. Ces méthodes d'intégration « ad hoc » sont inadaptées à des

outils d'intégration SIG car elles sont trop spécifiques aux cas considérés. Il est en effet rarement possible de les re-utiliser dans un autre contexte d'étude ou un SIG a pour objectif de traiter un grand nombre de données diverses.

Partant de ce constat, l'objectif de cette thèse est de reprendre les algorithmes déjà existants pour la comparaison et l'appariement des données 2D et de les adapter aux traitements des différents levés topo-bathymétriques afin d'obtenir un MNT intégré terre-mer. Il a été décidé de destiner ces outils à l'appariement des éléments caractéristiques homologues, indicateurs clés de l'information de relief sur la zone de recouvrement entre les différents levés.

Pour discuter des apports et des perspectives offertes par notre méthode d'intégration des données topo bathymétriques, nous nous attarderons sur les étapes majeures de l'intégration de données : appariement des éléments caractéristiques homologues issus de chaque modèle avec les distances de Fréchet nouvellement définies et passage à un Modèle Numérique de Terrain Intégré grâce à la méthodologie globale d'intégration des données topo-bathymétriques.

1 De la reconnaissance d'éléments caractéristiques à leur appariement.

L'extraction de ces éléments caractéristiques impose la définition d'outils appropriés afin de les comparer et de les appairer entre les deux MNT. La distance de Fréchet a donc été reprise au Chapitre II et augmentée de manière à répondre au plus grand nombre de cas possibles. Ce chapitre se termine sur le test effectué sur la comparaison de traits de côte (Le Berre *et al.*, 2004) qui a permis d'identifier des cas limites pour lesquels d'autres mesures sont nécessaires : la distance moyenne de Fréchet et la distance de Fréchet discrète partielle. Par ailleurs l'algorithme a été complété de manière à rechercher le sens des lignes, la relation d'inclusion et de considérer plusieurs cas possibles d'appariement entre les lignes. C'est ainsi qu'une méthode globale d'appariement par fichier a été mise au point.

Au vu de ce chapitre, nous pouvons en déduire que l'apport de la thèse se situe au niveau algorithmique. Il concerne l'**introduction de différents types de lignes** lors du processus de comparaison : ouvertes, fermées, ou d'orientations diverses. Il a donc été impératif d'**élargir la distance de Fréchet** de départ à des lignes réelles quel que soit leur type (ouverte ou fermée) :

- distance moyenne,
- distance partielle entre lignes,

- distance entre lignes fermées,
- distance entre des extrémités de ligne uniquement (distance « deux partiel/partiel »).

Ces distances ont par la suite été incorporées aux processus d'appariement de lignes au cours de trois étapes importantes : tout d'abord rechercher la nature des lignes comparées (ouvertes ou fermées) puis vérifier leur orientation respective. Ensuite, déterminer le type d'appariement mis en jeu entre les lignes concernées via l'étude des relations entre leurs rectangles englobants. Et enfin réaliser l'appariement des lignes en accord avec le cas identifié.

2 Méthode globale d'intégration des données topo-bathymétriques

La seconde contribution concerne l'**élaboration d'une méthodologie globale d'intégration** qui vise à résoudre la plupart des problèmes liés à l'intégration de deux modèles numériques de terrain acquis de façon indépendante. Cette méthodologie d'intégration globale des MNT a été conçue pour prendre en compte tous ces éléments (typologie, algorithme d'appariement) et quels que soient les modèles de départ (qualité, résolution, système de coordonnées, date d'acquisition) afin de produire un MNTI final continu et cohérent.

La méthodologie mise en place vise tout d'abord à conditionner les données géographiques réelles de manière à solutionner les problèmes dus à la structure des jeux de données et à la mise en cohérence de l'information spatiale. Toute la première phase des manipulations se concentre donc uniquement sur la préparation des données à l'intégration future : recalage géographique, mise en cohérence et étude de qualité sur la répartition de l'information au sein des données.

L'analyse des MNT réalisée lors de la seconde phase de la méthodologie en vue de l'appariement et de la fusion reprend les lignes caractéristiques extraites lors du processus de segmentation.

Le mécanisme d'appariement exploité lors de la troisième phase de fusion est basé sur les améliorations de l'algorithme de Fréchet.

Les outils implémentés spécifiquement pour cette méthodologie (appariement et déformation élastique) ont été validés sur des données simulées simplifiées de manière à vérifier les corrections effectuées.

3 Perspectives

La richesse des données géographiques disponibles, le manque de temps et d'outils ont empêché d'effectuer une intégration sur les données disponibles collectées sur les zones d'intérêt présentées au début du Chapitre III. En effet, une validation finale de la méthodologie d'intégration passe naturellement par son application sur des cas réels. Cependant, cela nécessite auparavant l'implémentation de certains outils manquants dans la méthodologie, ainsi que la définition d'une classification paysagère sur la base des éléments caractéristiques. De ce constat il est possible de dégager des perspectives (résumées dans le tableau 19) à court terme – implémenter les outils manquants nécessaires à cette validation – et d'autres à plus long terme : valider la méthodologie globale.

Pour arriver à notre but final de validation, la première étape chronologique est tout d'abord de proposer une description géomorphométrique de chaque paysage littoral ainsi que la définition des mesures statistiques complémentaires relatives aux différents reliefs. Des mesures géomorphométriques doivent permettre la caractérisation de chaque paysage et leur identification au sein des MNT. Par là même il serait ensuite possible d'affiner les méthodes d'interpolation utilisées lors de la fusion ainsi que les algorithmes de déformation élastique. Un autre point important est de prendre en compte la taille des éléments caractéristiques morphométriques. En effet, leur détection ne va pas être la même selon le niveau de détail du MNT. En parallèle de leur inventaire pour chaque type de MNT, il faut donc s'intéresser à leur méthode de détection, et si ces éléments possèdent une taille variable ou maximale/minimale au-delà desquelles leur recherche n'aura plus de sens. Pour le cas où ils auraient une taille variable, il s'agit de tester si un algorithme de logique floue suffirait à solutionner ce point ([Scheinder, 2004], [Evans, 2003], [McMillan, 1995]).

Ensuite, la seconde étape est d'affiner certains des outils de la méthodologie avec en priorité ceux destinés à la gestion des lacunes, la segmentation et l'enrichissement du MNT (voir le chapitre III.4 pour de plus amples détails). Certains des modules déjà existants peuvent aussi être testés. [Saux et al., 2004], par exemple, ont mis au point un outil de segmentation qui, bien que non directement lié aux éléments caractéristiques définis dans ce travail, produit néanmoins une série de lignes structurantes. Il est tout à fait envisageable de le tester sur les MNT simulés (possédant un domaine de recouvrement suffisant) puis, dans le cas de résultats concluants, sur les données des Abers ou du golfe du Morbihan (données issues du programme Litto3D). La figure 124 présente la méthodologie complétée de la partie analyse paysagère manquante.

Tableau 19 : Tableaux récapitulatifs des apports de la thèse et des domaines de recherche nécessitant un approfondissement.

Domaines	Apports	Perspectives
Classification paysagère	Elaboration d'une typologie paysagère littorale complétée des éléments caractéristiques.	Définition de valeurs géomorphométriques selon les différentes classes de paysages identifiées. Prise en compte du facteur multi échelle.
Distance de Fréchet	Définition de nouvelles distances adaptées aux cas réels des lignes géographiques.	Adaptation de la distance de Fréchet aux derniers cas non encore résolus.
Déformation élastique	Test de deux méthode de déformation différentes : une locale avec un rayon de déformation défini par l'utilisateur et une globale.	Tests sur différents types de paysages et détermination de la méthode de déformation ad hoc selon le paysage considéré.
Méthodologie globale d'intégration	Elaboration d'une méthodologie globale d'intégration des MNTs.	Application de la méthodologie à des données réelles.

Ainsi, la méthode d'intégration des données topo bathymétriques détaillée dans cette thèse avec les améliorations proposées ci-dessus permettra de réaliser la fusion de données disparates tout en répondant aux besoin des acteurs du littoral : souligner les zones insuffisamment acquises ou renseignées, préciser la nature de données et leur échantillonnage. Ces étapes réalisées, la validation de la méthodologie globale - à l'aide de lignes caractéristiques (thalwegs, crêtes, ruptures de pente) extraites de MNT issus de divers programmes de suivi du littoral (recherche ou programmes nationaux comme Litto3D) - pourra être envisagée.

Devant le constat du besoin croissant pour de multiples thématiques d'une représentation continue de ce milieu sensible, extrêmement fragile et complexe qu'est le littoral, il a été entrepris un inventaire des différents processus de fusion existants. Ceux-ci ont été estimés insuffisants, trop peu précis et pas assez génériques. Nous avons donc proposé au cours de ce travail des améliorations visant à corriger ces faiblesses : prise en compte du relief intégré, adaptation aux MNT d'outils d'intégration robustes, mais destinés aux données 2D, et élaboration d'une méthodologie générique de production de MNT.

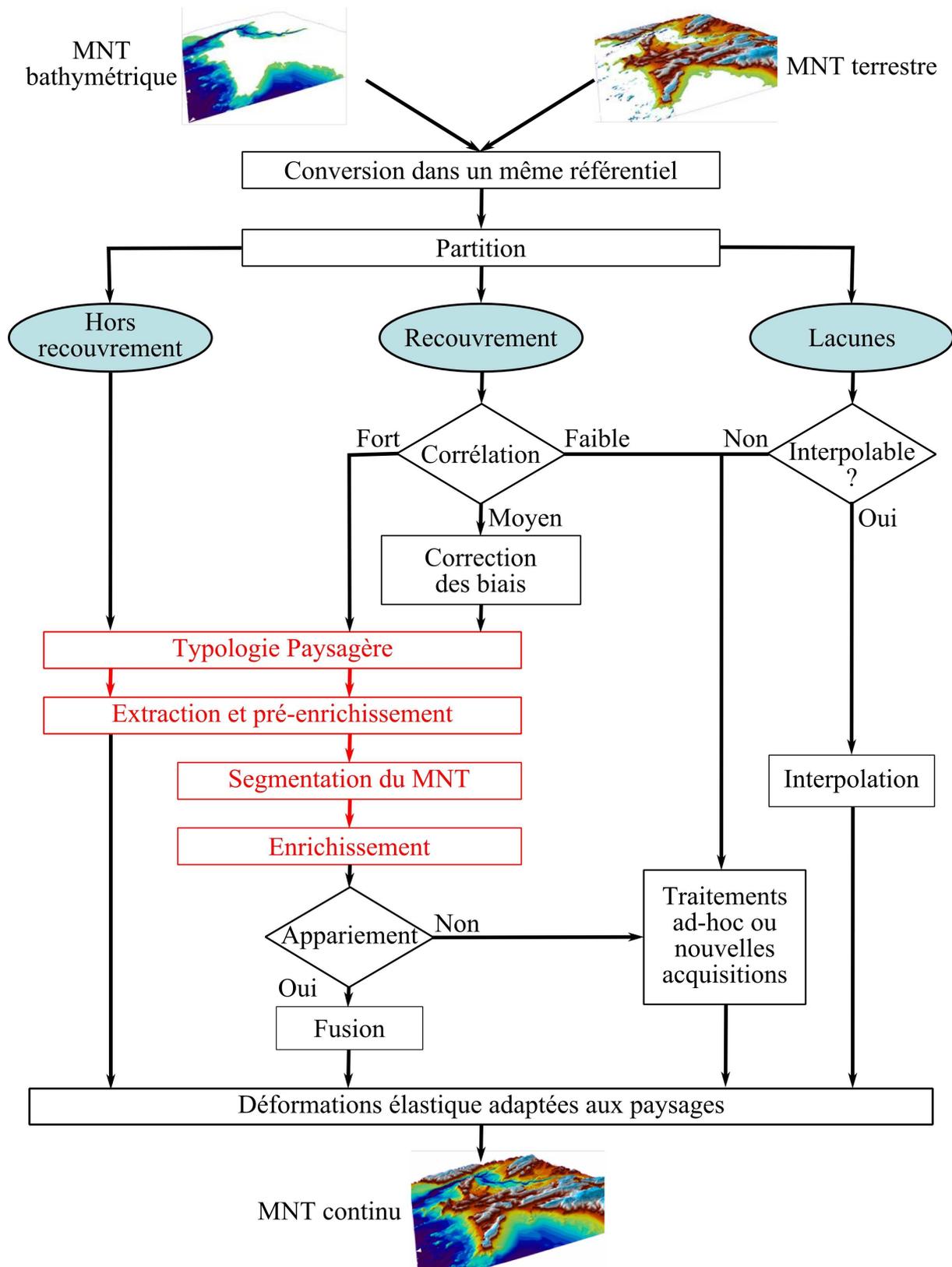


Figure 124 : Méthodologie Globale d'intégration complétée de l'analyse paysagère.

Cette méthodologie, si elle est menée jusqu'à sa validation finale, permettra d'intégrer des données multi sources, dont celles prochainement fournies par le programme d'acquisition Litto3D.

REFERENCES BIBLIOGRAPHIQUES

- [Adragna, 1997] Adragna, F. 1997. Interférométrie radar : principes, applications et limitations. *Bulletin - Société française de photogrammétrie et de télédétection*, **148**, 15 – 19.
- [Airault, 1996] Airault, S. 1996. De la base de données à la carte : une approche globale pour l'équarrissage de bâtiments. *Revue Internationale de Géomatique*, **6**(2 - 3), 203 – 217.
- [Allain *et al.*, 2000] Allain, S., Guillomont, B., Le Visage, C., Loubersac, L., & J., Populus. 2000. Données géographiques de référence en domaine littoral. *Pages 67–79 of: Shom-Ifremer (ed), CoastGIS'99 : Geomatics and coastal environment*. Populus, J. and Loubersac, L.
- [Alt & Gadau, 1995] Alt, H., & Gadau. 1995. Measuring the resemblance of polygonal curves. *Pages 102–109 of: 8th Annu. ACM Symposium on Computational Geometry*.
- [Amenta *et al.*, 1998] Amenta, N., Bernd, M., & Kolluri, D. E. 1998. The crust and the beta-skeleton: combinatorial curve reconstruction. *Graphical Models and Image Processing*, **60**, 125 – 135.
- [Arnaud & Emery, 2000] Arnaud, M., & Emery, X. 2000. *Estimation et interpolation spatiale : méthodes déterministes et méthodes géostatistiques*. Hermes, Paris.
- [Barnsley *et al.*, 1988] Barnsley, M.F., Devaney, R.L., Mandelbrot, B.B., Peitgen, H.O., Saupe, D., & Voss, R.F. 1988. *The science of Fractal images*. Springer - Verglas.
- [Beeri *et al.*, 2004] Beeri, C., Kanza, Y., Safra, E., & Sagiv, Y. 2004. Object Fusion in Geographic Information System. *In: Proceedings of the 30th VLDB Conference*.
- [Bel Hadj Ali, 1997] Bel Hadj Ali, A. 1997. *Etude de l'appariement des surfaces et implémentation des indicateurs de qualité en vue de leur sensibilité*. Rapport de DEA SIG, ENSG et Université Marne-la-Vallée, COGIT.
- [Bel Hadj Ali, 2001] Bel Hadj Ali, A. 2001. *Qualité géométrique des entités géographiques surfaciques : application à l'appariement et définition d'une typologie des écart géométriques*. Ph.D. thesis, Université Marne-la-Vallée, COGIT.
- [Bonin & Poupeau, 2005] Bonin, O., & Poupeau, B. 2005. Identification of simple landscape features on triangulated contour lines. *In: GISRUK'05*.
- [Bonin *et al.*, 2006] Bonin, O., Poupeau, B., & Fontaine, L. 2006. Identification of landscape features on triangulated contour lines: analysis refinement with terrain cross sections. *In: GISRUK'06*.
- [Bonnot-Courtois & Levasseur, 2003] Bonnot-Courtois, C., & Levasseur, J-E. 2003. Recherche d'indicateurs « naturalistes » de la limite supérieure du domaine maritime. *Cahiers Nantais*, **59**, 47 – 56.
- [Cambray, 1994] Cambray, B. de. 1994. *Etude de la modélisation de la représentation de l'information spatiale 3D dans les bases de données géographiques*. Ph.D. thesis, Université Paris 6.
- [Charleux-demargne, 2001] Charleux-demargne, J. 2001. *Qualité des Modèles Numériques de Terrain pour l'Hydrologie : Application à la caractérisation du régime de crues des bassins versants*. Ph.D. thesis, Marne-laVallée, CEMAGREF.
- [CNIG., 2003] CNIG. 2003. *Rapport final du Groupe de travail Littoral*. Tech. rept. CNIG.
- [Cocquerez *et al.*, 1995] Cocquerez, JP, Philipp, S., & Zeboudj, R. 1995. *Comparaison de méthodes de segmentation d'images*. GRETSI, Groupe d'Etudes du Traitement du Signal et des Images.

- [Comic *et al.*, 2005] Comic, L., De Floriani, L., & Papaleo, L. 2005. Morse-Smale Decompositions for Modeling Terrain Knowledge. *Lecture Notes in Computer Science*, **3693**, 426–444.
- [Coque, 1993] Coque, R. 1993. *Géomorphologie*. Armand Colin.
- [Dakowicz & Gold, 2003] Dakowicz, M., & Gold, C. 2003. Extracting meaningful slopes from terrain contours. *International Journal of Computational Geometry and Applications*, **13**(4), 339.
- [Danovaro *et al.*, 2007] Danovaro, E., De Floriani, L., & Vitali, M. 2007. Multi-resolution Morse-Smale complexes for terrain modeling. *Pages 337–342 of: Image Analysis and Processing, 2007. ICIAP 2007. 14th International Conference on.*
- [Delacourt *et al.*, 2007] Delacourt, C., Allemand, P., Berthier, E., Raucoules, D., Casson, B., Grandjean, P., Pambrun, C., & Varel, E. 2007. Remote-sensing techniques for analysing landslide kinematics. *Bulletin de la Société Géologique de France*, **178**, 89 – 100.
- [Denègre & Salgé, 1996] Denègre, J., & Salgé, F. 1996. *Les systèmes d'information géographiques*. PUF.
- [Derruau, 1988] Derruau, M. 1988. *Précis de géomorphologie*. Masson.
- [Devogele, 1997] Devogele, T. 1997. *Processus d'intégration et d'appariement de Bases de Données Géographiques : Application à une base de données routières multi-échelles*. Ph.D. thesis, IGN, COGIT.
- [Devogele, 2000] Devogele, T. 2000. Mesure d'exactitude et processus de fusion à l'aide de la distance de Fréchet discrète. *Revue internationale de Géomatique*, **10**(3-4/2000), 359–381.
- [Devogele, 2002] Devogele, T. 2002. A new Merging process for data integration based on the discrete Fréchet distance. *Pages 167–181 of: Richardson D, van Oosterom P (ed), Proceedings of the 10th International Symposium on Spatial Data Handling (SDH)*. Springer.
- [Dikau, 1989] Dikau, R. 1989. *The application of a digital relief model to landform analysis in geomorphology*. Raper, J.(Hrsg.): *Three dimensional application in Geographic Information Systems*. Vol. 5. Taylor and Francis, London, Kapitel.
- [Dimitrijévic, 2000] Dimitrijévic, I. 2000 (Septembre). *Etude des liens d'appariement topologique et géométrique dans les bases de données géographiques*. rapport de stage de DESS, Université Paris 6 (Jussieu), Laboratoire COGIT.
- [documentation française, 1997] documentation française, La. 1997. *Plans de prévention des risques littoraux (PPR) : guide méthodologique*. PPR. Ministère de l'aménagement du territoire et de l'environnement and ministère de l'équipement des transports et du logement.
- [Dupont, 1997] Dupont, S. 1997. *Génération de modèles numériques de terrain par interférométrie ROS*. Ph.D. thesis, Université de Nice - Sophia Antipolis.
- [Dupont *et al.*, 1998] Dupont, S., Demargne, L., & Durand, Ph. 1998. *Production de MNT à partir de données SAR RADARSAT*. Rapport Final, Programme AvalSAR. CNES, Toulouse.
- [Eiter & Mannila, 1994] Eiter, T., & Mannila, H. 1994. *Computing Discrete Fréchet Distance*. Rapport technique num. CD-TR 94/64. Institute of Information Systems, Technical University of Vienna.
- [EuroSDR, 2007] EuroSDR. 2007. *Land and Marine information Integration*.
- [Evans, 1980] Evans, I.S. 1980. An Integrated System of Terrain Analysis and Slope Mapping. *Zeitschrift fur Geomorphologie suppl-Bd*, **36**, 274–295.
- [Fisher *et al.*, 2004] Fisher, P., Wood, J., & Cheng, T. 2004. Where is Helvellyn? Fuzziness of multi-scale landscape morphometry. *Transactions of the Institute of British Geographers*, **29**(1), 106–128.
- [Fréchet, 1906] Fréchet, M. 1906. Sur quelques points du calcul fonctionnel. *Rendiconti del Circolo Mathematico di Palermo*, **22**, 1 – 74.

- [Gabay & Doytcher, 1994] Gabay, Y., & Doytcher, Y. 1994. Automatic adjustment of line maps. *Pages 333 – 341 of: Proceeding of the GIS/LIS'94 Annual Convention.*, vol. 1.
- [Gesch & Wilson, 2001] Gesch, D., & Wilson, R. 2001. Development of a seamless multisource topographic/bathymetric elevation model for Tampa Bay. *Marine Technology Society Journal*, **35**(4), 58 – 64.
- [Gourmelon *et al.*, 2007] Gourmelon, F., Robin, M., & Tissot, C. 2007. SIG et zones côtières. *Revue Pelagos, n° spécial, ISMAL (Alger)*, 7 – 16.
- [Guth, 2006] Guth, P.L. 2006. Geomorphometry from SRTM: Comparison to NED. *Photogrammetric Engineering and Remote Sensing*, **72**(3), 269–277.
- [Hengl & Reuter, 2008] Hengl, T., & Reuter, H.I. 2008. *Geomorphometry : concepts, software, applications*. Elsevier Science.
- [ICZMap, 2004] ICZMap. 2004. *Integrated coastal Zone Mapping*.
- [Jamet & Phalakarn, 1989] Jamet, O., & Phalakarn, B. 1989. Un algorithme rapide de mise en correspondance de cartes. *Pages 1123 –1130 of: 7eme congrès de reconnaissance des formes et intelligence artificielle*. AFCET, Barcelone.
- [Juc & Deffontaines, 2003] Juc, M., & Deffontaines, B. 2003. Constitution d'un MNT intégré terre-mer : Application à la baie du Mont Saint-Michel. *Bulletin - Société française de photogrammétrie et de télédétection*, 3 – 15.
- [Kaveh & Deffontaines, 2005] Kaveh, H.F., & Deffontaines, B. 2005. Mise à jour de MNT intégré terre - mer et application au prisme mégalidal du Mont Saint-Michel (France). *Revue Française Photogrammétrie et de télédétection*, 2 – 16.
- [Kervazo & Moutoulatchimy, 2006] Kervazo, J., & Moutoulatchimy, M. 2006. *Création d'un addon de détection de lacunes entre deux jeux de données terre-mer pour le logiciel ArcMap*. Tech. rept. Ecole Navale.
- [Kienzle, 1996] Kienzle, S.W. 1996. *Application of GIS in Hydrology and Water Resources Management*. Chap. Using DTMs and GIS to define input variables for hydrological and geomorphological analysis, pages 183 – 190.
- [Lagacherie *et al.*, 1996] Lagacherie, P., Moussa, R., cormary, D., & Molenat, J. 1996. *Application of GIS in Hydrology and Water Resources Management*. Chap. Effects of DEM data source and sampling pattern on topological parameters and on a topography-based hydrological model, pages 191 – 199.
- [Langlois, 1994] Langlois, P. 1994. *Une transformation élastique du plan basée sur un modèle d'interaction spatiale, Application à la géomatique* ». Tech. rept. MTG, université de Rouen.
- [Laurini, 1996] Laurini, R. 1996. Raccordement géométrique de bases de données géographiques fédérées. *Ingénierie des systèmes d'informations.*, **4**(3), 361 – 388.
- [Laurini, 1998] Laurini, R. 1998. Spatial multi-database topological continuity and indexing: a step towards seamless GIS data interoperability. *International Journal of Geographical Information Science: Special issue: Interoperability in GIS*, **12**(4), 373–402.
- [Le Berre *et al.*, 2004] Le Berre, I., Henaff, A., Wenzel, F., & Giraudet, J. 2004. *Cartographie synthétique de l'environnement littoral du Finistère, exploitation de SPOT pour la cartographie de l'estran, du trait de côte et de l'occupation du littoral*. Tech. rept. Rapport final, Appel à proposition CNES/IFEN “ Suivi du littoral par SPOT 5”, GEOMER laboratory/Cetmef/DDE29.
- [Le Berre *et al.*, 2005] Le Berre, I., Hénaff, A., Devogele, T., Mascaret, A., & Wenzel, F. 2005. Spot 5 : un outil pertinent pour le suivi du trait de côte ? *Noréis*, **196**(2005-3), 23–35.
- [Le Men & Jamet, 1990] Le Men, H., & Jamet, O. 1990. Interprétation automatique de l'occupation du sol sur image SPOT. *In: Symposium International de Cartographie Thématique Dérivée des Images Satellitaires*.
- [Leberl, 1990] Leberl, F. 1990. *Radargrammetric Image Processing*. Artech House.

- [Lemarié, 1996] Lemarié, C. 1996. *Etat de l'art sur l'appariement*. Tech. rept. DT/9600022. IGN.
- [Li, 1990] Li, Z. 1990. *Sampling strategy and accuracy assessment for digital terrain modeling*. Ph.D. thesis, University of Glasgow.
- [Li et al., 2005] Li, Z., Zhu, Q., & Gold, C. 2005. *Digital terrain modeling : principles and methodology*. CRC Press.
- [Longley et al., 1999] Longley, P.A., Goodchild, M.F., Maguire, D.J., & Rhind, D.W. 1999. *Geographical Information Systems: Principles, Techniques, Applications and Management*. Wiley.
- [Louvard & Grateau, 2005] Louvard, L., & Grateau, C. 2005. The Litto3D Project. In: *ICC Proceedings*. ICC.
- [Lurton, 2002] Lurton, X. 2002. *An introduction to Underwater Acoustics : Principles and Applications*. Springer.
- [MacMillan & Shary, 2008] MacMillan, R.A., & Shary, P.A. 2008. *Geomorphometry : concepts, software, applications*. Elsevier Science. Chap. Landforms and Landforms elements in geomorphometry, pages 227–254.
- [Magillo et al., 2007] Magillo, Paola, Danovaro, Emanuele, Floriani, Leila De, Papaleo, Laura, & Vitali, Maria. 2007. Extracting terrain morphology - a new algorithm and a comparative evaluation. *Pages 13 – 20 of: Braz, José, Vázquez, Pere-Pau, & Pereira, João Madeiras (eds), GRAPP (GM/R)*. INSTICC - Institute for Systems and Technologies of Information, Control and Communication.
- [Marguerit, 2004] Marguerit, C. 2004. Synthèse sur la concordance des références altimétriques en Baie du Mont-Saint-Michel. *XYZ : Rev. Assoc. Franc. Topogr.*, **26**(98), 53 – 56.
- [Mascret et al., 2006] Mascret, A., Devogele, T., Le Berre, I., & Hénaff, A. 2006. Coastline Matching Process based on the discrete Fréchet distance. In: Kainz, W., Reidl, A., & Elmes, G. (eds), *In Proceedings of the 12th International Symposium on Spatial Data Handling (SDH)*. Springer.
- [Massonet & Rabaute, 1993] Massonet, D., & Rabaute, T. 1993. Radar interferometry : limits and potential. *IEEE, Transactions on Geoscience and Remote Sensing*, **31**(2), 455 – 464.
- [McMaster, 1986] McMaster, R. 1986. A statistical Analysis of Mathematical Measures for Linear Simplification. *The American Cartographer*, **23**.
- [Monier, 1997] Monier, P. 1997. *Caractérisation du terrain en vue de son traitement numérique: application à la généralisation de l'orographie*. Ph.D. thesis, Université Louis Pasteur, Strasbourg.
- [Mustière, 1995] Mustière, S. 1995. *Mesures de la qualité de la généralisation du linéaire*. Rapport de stage du DESS de Cartographie., Université Paris I, réalisé au COGIT.
- [Mustière & Devogele, 2008] Mustière, S., & Devogele, T. 2008. Matching networks with different levels of detail. *GeoInformatica*, **12**(4), 435 – 453.
- [Nocéra, 1993] Nocéra, L. 1993. *Stéréoscopie à partir d'images de Radar à Ouverture de Synthèse : utilisation des images du capteur SIR-B*. Ph.D. thesis, INRIA, Sophia Antipolis.
- [Oliveira et al., 2006] Oliveira, De E., Populus, J., & Guillaumont, B. 2006. Using Lidar survey and satellite imagery for predictive modelling of coastal habitats - A case for seaweed in Brittany (France). *Earsel Proceedings*, **5**(2), 208–223.
- [Paquerault & Maître, 1997] Paquerault, S., & Maître, H. 1997. « La radarclinométrie ». *Bulletin SFPT*, **148**, 20 – 29.
- [Perkal, 1956] Perkal, J. 1956. On Epsilon length, *Bulletin de l'académie Polonaise des Sciences*. **4**, 399 – 403.

- [Peucker & Douglas, 1975] Peucker, TK, & Douglas, DH. 1975. Detection of surface-specific points by local parallel processing of discrete terrain elevation data. *Pages 375 –387 of: Computer Vision, Graphics and Image Processing*, vol. 4.
- [Phalakarn, 1991] Phalakarn, B. 1991. *Evaluation de la qualité des processus de segmentation d'image par mise en correspondance à une référence*. Ph.D. thesis, Université de Paris 7.
- [Pichon *et al.*, 1997] Pichon, G., Dupont, S., Nonin, P., & Renouard, L. 1997. Couvertures nuageuses : une limitation pour la génération de MNT in ROS ERS. *Bulletin - Société française de photogrammétrie et de télédétection*, **148**.
- [Pike, 1988] Pike, R.J. 1988. Geomorphometry-diversity in quantitative surface analysis. *Mathematical Geology*, **20**(5), 491–511.
- [Pike, 2000] Pike, R.J. 2000. The geometric signature: quantifying landslide-terrain types from digital elevation models. *Progress in Physical Geography*, **24**, 1 – 20.
- [Plazanet, 1996] Plazanet, C. 1996. *Enrichissement des bases de données géographiques : analyse de la géométrie des objets linéaires pour la généralisation cartographique (application aux routes)*. Ph.D. thesis, Thèse de l'université de Marne-la-Vallée, Laboratoire COGIT.
- [Podobnikar, 2005] Podobnikar, T. 2005. Production of integrated digital terrain model from multiple datasets of different quality. *International journal of geographical information science*, **19**(1), 69–89.
- [Polidori, 1995] Polidori, L. 1995. Réflexions sur la qualité des modèles numériques de terrain. *Bulletin - Société française de photogrammétrie et de télédétection*, **139**.
- [Robin, 2002] Robin, M. 2002. *Le littoral, regards, pratiques et savoirs*. Editions de l'ENS. Chap. Télédétection et modélisation du trait de côte et de sa cinématique., pages 95 – 115.
- [Rousseaux, 2005] Rousseaux, F. 2005. *Etude des modeles numeriques de terrain pour ameliorer la fiabilite des calculs d'aleas et de vulnerabilite*. Ph.D. thesis, Université de Marne-la-Vallée.
- [Ruas, 2002a] Ruas, A. 2002a. *Généralisation et représentation multiple*. Hermes Lavoisier. Chap. 2 : Pourquoi associer les représentations des données géographiques?, pages 45 – 54.
- [Ruas, 2002b] Ruas, A. 2002b. *Généralisation et représentation multiple*. Hermes Lavoisier. Chap. 4 : Les problématiques de la généralisation., pages 75 – 90.
- [RUOA, 2006] RUOA. 2006. *Dynamique et représentation du littoral*. Université de printemps du Réseau des Universités de l'Ouest Atlantique.
- [Safere seas, 2007] Safere seas. 2007.
- [Saint-Guily, 2006] Saint-Guily, B. 2006. *Encyclopédie Universalis : variabilité des courants marins*.
- [Salmeron & Milgram, 1986] Salmeron, E., & Milgram, M. 1986 (avril). Utilisation de la relaxation pour la mise en correspondance des segments d'une carte et d'une image aérienne. *Pages 32–38 of: Deuxième Colloque Image*. Semaine Internationale de l'Image Electronique, Nice.
- [Saux *et al.*, 2004] Saux, E., Thibaud, R., Li, K.J., & Kim, M.H. 2004. A new approach for a topographic feature-based characterization of digital elevation data. *Pages 73–81 of: Proceedings of the 12th annual ACM international workshop on Geographic information systems*. ACM New York, NY, USA.
- [Schneider & Wood, 2004] Schneider, B., & Wood, J. 2004. *Topological data structures for surfaces : an introduction to geographical information science*. Wiley. Chap. Construction of Metric Surface Networks from Raster-Based DEMs, pages 51–70.
- [Seatechweek, 2006] Seatechweek. 2006. *International marine science and technology week*.
- [Sederberg & Greenwood, 1992] Sederberg, T.W., & Greenwood, E. 1992. A physically based approach to 2-D shape blending. *ACM SIGGRAPH Computer Graphics*, **26**(2), 25–34.

- [Servigne, 1993] Servigne, S. 1993. *Base de données géographiques et photos aériennes : de l'appariement à la mise à jour*. Ph.D. thesis, INSA Lyon.
- [Servigne, 1994] Servigne, S. 1994. Base de données géographiques et photos aériennes : mise en correspondance. *Pages 1390–1399 of: 5th European Conference and Exhibition on Geographical Information Systems (EGIS)*.
- [Sheeren, 2005] Sheeren, D. 2005. Etude de la cohérence inter-représentations : vers une meilleure intégration des bases de données spatiales. *Bulletin Scientifique et Technique de l'IGN*, 71–80.
- [SHOM, 1998] SHOM. 1998. *Zéro hydrographique : expérience en rade de Brest*.
- [SHOM, 2009] SHOM. 2009. *Echosondeurs*.
- [SHOM & Allain, 2004] SHOM, & Allain, S. 2004 (Novembre). *Les levés hydrographiques dans les petits fonds*. Lettre du SHOM 21. SHOM.
- [SHOM & Debese, 2006] SHOM, & Debese, N. 2006 (Mai). *Etude comparative des données multifaisceaux et laser acquises sur le golfe de Morbihan*. Tech. rept. 002/06. S.H.O.M.
- [Simon, 2001] Simon, B. 2001. *Niveaux caractéristiques et coefficient de marée - Calcul direct à l'aide des constantes harmoniques*. rapport d'étude 001/01. Éditions du Service Hydrographique de la Marine.
- [Spaccapietra *et al.*, 1992] Spaccapietra, S., Parent, C., & Dupont, Y. 1992. Model Independent Assertions for Integration of Heterogeneous Schemas. *Very Large DataBases Journal*, 1(1), 81–126.
- [Sriti *et al.*, 2005] Sriti, M., Thibaud, R., & Claramunt, C. 2005. A Network-based Model for Representing the Evolution of Spatial Structures. *Proceedings of the 4th ISPRS Workshop on Dynamic and Multi-dimensional GIS*, september, 150–155.
- [Tannous *et al.*, 1997] Tannous, I., Le Goff, F., & Julien, P. 1997. Méthodologie de production de MNT par fusion de données mixtes radar/optique. *Bulletin SFPT*, 148.
- [TCI, 1999] TCI, software. 1999. *Adjust : True Rubber Sheeting inside AutoCAD*. Baker City, Oregon.
- [Thibaud, 2002] Thibaud, R. 2002. Un générateur de données bathymétriques pour l'aide au prétraitement des cartes marines. *Revue internationale de Géomatiques, SIG côtiers*, 275 – 290.
- [Tillier & Robin, 2008] Tillier, I., & Robin, M. 2008 (Janvier). Réflexions sur les conflits d'usages en zones côtières : essai de formalisations à partir d'exemples en baie de Bourgneuf (Pays de la Loire, France). *In: Le littoral, Subir, Dire, Agir*. Colloque International Pluridisciplinaire IFRÉSI/MESHS, TVES, IMN, DYRT, EREIA et CLERSE, Lille.
- [Tricart, 1995] Tricart, J. 1995. *Précis de géomorphologie : V2, géomorphologie dynamique générale*. Sedes CDU.
- [Valadas, 2005] Valadas, B. 2005. *Géomorphologie dynamique*. Armand Colin.
- [Vauglin, 1997] Vauglin, F. 1997. *Modèles statistiques des imprécisions géométriques des objets géographiques linéaires*. Ph.D. thesis, Université de Marne-la-Vallée, Laboratoire COGIT.
- [Weibel & Heller, 1991] Weibel, R., & Heller, M. 1991. *Geographical Information Systems : Principles and Applications*. Vol. 1. Longman, London. Chap. Digital Terrain Modeling, pages 269 – 297.
- [Wood, 1996] Wood, J. 1996. *The geomorphological characterization of digital elevation models, unpublished 10 Ph. D.* Ph.D. thesis, Thesis, Univ. of Leicester, Leicester, 1996.
- [Wood, 2002] Wood, J. 2002. *Landserf: visualisation and analysis of terrain models*.

- [Wöppelmann *et al.*, 1999] Wöppelmann, G., Allain, S., Bahurel, P., Lannuzel, S., Simon, B., & Brest, EPSHOM. 1999. Zéro hydrographique vers une détermination globale. *Revue XYZ*, **79**, 27 – 32.
- [Zhang & Montgomery, 1994] Zhang, W., & Montgomery, D.R. 1994. Digital elevation model grid size, landscape representation, and hydrologic simulations. *Water Resources Research*, **30**(4), 1019–1028.
- [Zhilin *et al.*, 2005] Zhilin, L., Qing, Z., & Gold, C. 2005. *Digital Terrain Modeling : Principles and Methodology*. CRC Press.

TABLE DES FIGURES

Figure 1 : Positionnement de la thèse.....	15
Figure 2 : Exemple de représentation d'un pont selon les différents types de modèles disponibles (2D, 2.5D, 2.75D et 3D, source : [Devogele, 1997])	22
Figure 3 : Représentation d'un même fleuve dans deux bases géographiques avec a) des emprises différentes, b) des niveaux de détail différents d'une base à l'autre.....	23
Figure 4 : Exemple de différences de géométries pour un fleuve selon l'échelle : a) une ligne (au Nord de Crozon) et b) un polygone (avec un zoom plus fort sur le fleuve) (Source : http://www.geoportail.fr/)	23
Figure 5 : Exemple de tables attributaires thématiques différentes d'une base de données à l'autre.	24
Figure 6 : Représentation d'un même fleuve dans deux bases géographiques avec des relations topologiques variables (absences de nœuds).....	24
Figure 7 : Illustration des différents types de relations entre bases de données.....	25
Figure 8 : Exemple de fusion entre deux bases de données terrestre et maritime.....	27
Figure 9 : Définition et illustration de la distance de Hausdorff.	29
Figure 10 : Distance de Hausdorff entre deux lignes d'emprises différentes.	29
Figure 11 : Problème de la distance de Hausdorff avec des lignes sinueuses.....	29
Figure 12 : Exemple de couple de lignes.	31
Figure 13 : Rectangle englobant simple et rectangle englobant minimum calculés à partir d'un objet. En comparaison l'enveloppe convexe a également été représentée.....	33
Figure 14 : Etapes de fusion et de déformation élastique de lignes.	36
Figure 15 : Exemple de bande de correction élastique (Source : [Devogele, 2002]).....	37
Figure 16 : Exemple de fusion + déformations élastiques (Source : [Devogele, 2000])	37
Figure 17 : Superposition des données du cadastre et de l'IGN avant tout processus de fusion et de déformation élastique.	38
Figure 18 : Etapes de production d'un MNT.	41
Figure 19 : Principe d'acquisition d'un échosondeur monofaisceau et du Lidar bathymétrique.	43
Figure 20 : Illustration des techniques d'acquisition d'échosondeur. (Source : http://www.shom.fr/fr_page/fr_act_acquisition/smf/principe.htm).....	44
Figure 21 : Systèmes de mesure du fond marin selon la profondeur. (Source : lettre du SHOM N°21)	45
Figure 22 : Représentation des différents formats de MNT possibles : points, profils, courbes de niveau, TIN (Triangular Irregular Network), Grille.....	46
Figure 23 : Illustration des deux méthodes d'interpolation utilisant les lignes structurantes. En a) interpolation à partir des lignes de niveau C_k et C_{k+1} , et b) interpolation basée sur la ligne de plus grande pente. Source [Arnaud & Emery, 2000].	49
Figure 24 : Présentation du masque appliqué sur les données topographiques et bathymétriques successivement sur une zone test (Sud de l'Angleterre : Solent, près de Portsmouth).	52
Figure 25 : Placement des profils topo-bathymétriques réalisés sur les modèles superposés et visualisation en relief avec une exagération verticale de 10.	53
Figure 26 : Dessins des profils numéro 14, 3 et 9 présentés en Figure 25.....	54
Figure 27 : Données topographiques seules et superposition avec un MNT Lidar de meilleure résolution.....	56

Figure 28 : Bilan des apports de chaque méthode dans le processus d'intégration de données.	58
Figure 29 : Couple de lignes homologues.	64
Figure 30 : Algorithme pour le calcul du chemin minimum.	65
Figure 31 : Couples de points ($L_{1, i}$, $L_{2, j}$) du chemin minimum représentés par les lignes pointillées.	65
Figure 32 : Cas pour lesquels à distances moyennes égales, les distances de Fréchet maximales correspondantes sont très différentes.	66
Figure 33 : Exemples de couples de géométries pertinentes pour le calcul des mesures dérivées de la distance de Fréchet.	67
Figure 34 : Lignes partielles homologues représentées avec leur points.	67
Figure 35 : Algorithme de recherche des points de début de L_1 et L_2 en vue d'un appariement des deux lignes fermées.	69
Figure 36 : Ligne fermées représentées avec des sens opposés.	69
Figure 37 : Etape d'inversion du sens de la ligne fermée L_2 en L_2'	70
Figure 38 : Recherche du point homologue de $L_{1,1}$ parmi les points de L_2' et permutation du point trouvé de manière à ce qu'il devienne $L_{2',1}$	71
Figure 39 : détail de l'appariement d'une ligne fermée avec une ligne ouverte.	73
Figure 40 : Mise en parallèle des relations n-m et p-p entre deux objets.	73
Figure 41 : Cas d'appariement partiel-partiel de lignes ouvertes : a) exemple impossible à résoudre en l'état actuel de nos connaissances ; b) exemple qui possède une solution unique.	74
Figure 42 : Illustration des différents positionnements des lignes lors de leur appariement. ..	75
Figure 43 : Inventaire des cas d'appariement impossibles à traiter entre une ligne ouverte et une ligne fermée. En (a), (b) et (c) dans un contexte d'appariement deux partiel/partiel ; et en (d), (e), (f) pour des cas partiel/partiel.	75
Figure 44 : Inventaire des cas impossibles à traiter entre deux lignes fermées. En (a), (b) et (c) pour un appariement deux partiel/partiel ; et en (d), (e) pour un appariement partiel/partiel.	75
Figure 45 : Exemple de lignes pour lesquelles les deux extrémités sont homologues (cas 2PP).	76
Figure 46 : Description des trois principales étapes définies pour l'appariement.	78
Figure 47 : Inventaire des appariements possibles selon la nature des lignes comparées.	79
Figure 48 : Inventaire des relations d'inclusion existant entre objets géographiques : (a) emprises équivalentes, (b) emprise incluse dans une autre, (c) emprises possédant une zone de recouvrement, (d) emprises sans intersection.	80
Figure 49 : Algorithme d'appariement total « 1-1 » entre deux lignes ouvertes. Les illustrations indiquent quel cas est alors identifié lors du calcul de la d_{pdF}	81
Figure 50 : Algorithme d'appariement de deux lignes fermées.	81
Figure 51 : Algorithme d'appariement total/partiel entre deux lignes ouvertes.	82
Figure 52 : Méthodologie de découpage d'une ligne L_1 selon une ligne L_2 d'emprise plus petite en vue du calcul de la distance de Fréchet partielle d_{pdF}	83
Figure 53 : Algorithme d'appariement partiel/partiel ou deux partiel/partiel de lignes possédant un domaine de recouvrement entre leurs emprises.	83
Figure 54 : Algorithme de test d'appariement partiel/partiel (PP) et deux partiel/partiel (2PP).	85
Figure 55 : Algorithme d'appariement global.	87
Figure 56 : Représentation des différents traits de côtes issus de la BD ORTHO [®] , des levés de terrain et photo interprétés à partir des ortho-photographies, et de l'image SOPT 5.	90

Figure 57 : Illustration des différents traits de côte numérisés au 1/1500 et 1/6000 comparés à la BD ORTHO [®] et à des levés GPS établis sur une grève.	92
Figure 58 : Comparaison des traits de côte issus de photo interprétation à partir de l'image SPOT 5 avec le trait de côte de la BD ORTHO [®] sur une côte artificielle.....	92
Figure 59 : Localisation de la première zone d'étude (encadrée en rouge sur la carte du SHOM) : zone des Abers.	97
Figure 60 : Localisation de la deuxième zone d'étude (encadrée en rouge sur la carte du SHOM) : golfe de Morbihan.	98
Figure 61 : Carte de différences exprimées en mètres entre les données issues du Lidar topo-bathymétrique et celles acquises par le Lidar topographique après correction de la hauteur ellipsoïdale.	101
Figure 62 : Carte des valeurs de différences en mètres entre les données d'acquisition sonar mono-faisceau du SHOM et celle acquises par le lidar topographique (Ifremer).....	102
Figure 63 : Carte de différences entre les données bathymétriques acquises par échosondeur mono faisceau du SHOM et par du lidar topo-bathymétrique (source : Ifremer).	103
Figure 64 : Carte de densité des données Lidar topographique. En noir homogène, la densité de la grille ré échantillonnée à 2 mètres de résolution. En extrait sur l'Aber Wrac'h, le semis de points sources.	104
Figure 65 : Carte de densité des points topo-bathymétriques (Source des données : Ifremer).	105
Figure 66 : Carte de densité des points acquis par écho sondeur mono faisceau (Source des données : Service Hydrographique et Océanographique de la Marine).....	105
Figure 67 : Carte de densité de l'ensemble des données sur la zone des Abers. Les jeux de données comprennent : la grille LIDAR topographique, le LIDAR topo bathymétrique et les points de sondes.	106
Figure 68 : Illustration d'un chenal en 3D. Les points utilisés pour construire le modèle sont représentés sur le schéma.	108
Figure 69 : MNT de référence généré à partir d'un logiciel d'émulation de reliefs fractals. La ligne rouge représente l'isohypse zéro selon laquelle les données seront par la suite subdivisées en modèles terrestre et maritime.	108
Figure 70 : Levés bathymétriques de la partie Nord obtenus à partir du MNT de référence et du logiciel SimSmf.....	109
Figure 71 : Visualisation du bruit blanc inséré dans les données par superposition des lignes de niveau de même altitude. Vue de gauche : extrait général d'une zone d'étude ; vue de droite : zoom (cadre rouge) sur une partie des courbes de niveau.	110
Figure 72 : Résultat de la soustraction effectuée entre le modèle de référence et le modèle bruité avec un bruit blanc.	110
Figure 73 : Localisation du profil topographique effectué sur les données Lidar en vue de l'estimation de l'erreur de bruit blanc.	111
Figure 74 : Profil topographique des points exportés à partir du trait de coupe présenté Figure 73.....	111
Figure 75 : Visualisation des points Lidar total et ceux situés à 2m du trait de coupe de l'étude.....	112
Figure 76 : Répartition des points extraits des traits de coupe P1-P2 et P3-P4.	112
Figure 77 : Vue de gauche : extrait du MNT de référence. Vue de droite : superposition du MNT de référence avec le MNT bruité de manière à faire apparaître la différence de contour d'une même ligne.	113
Figure 78 : Visualisation des erreurs grossières ajoutées au modèle de référence.	114
Figure 79 : Représentation de l'ensemble de la méthodologie d'intégration.....	115
Figure 80 : Détail de la première étape : comparaison des MNT et analyse de corrélation. .	116

Figure 81 : Détail de la seconde phase : segmentation en vue de l'extraction des lignes nécessaires au processus d'appariement.	117
Figure 82 : Détail de la troisième et dernière étape : appariement, fusion et déformation élastique.....	119
Figure 83 : Schéma d'intégration de deux MNT. (a) : enrichissement, (b) : appariement, (c) : fusion, (d) : déformations élastiques.	120
Figure 84 : Courbe de variation du coefficient d'amortissement entre le point p_i et l'origine d'un vecteur de déformation sur un rayon d'influence égal à 4.....	122
Figure 85 : Etape de production des MNT simulés simples pour la validation de la méthodologie.....	123
Figure 86 : Illustration du modèle d'origine en (a), des modèles topographique et bathymétrique dérivés (sous forme de points en (b) et sous forme estompée en (c)) et des points bruités interpolés puis estompés en (d).	124
Figure 87 : Illustrations de l'impact des méthodes de déformation par translation vis-à-vis d'une méthode de déformation élastique sur le modèle bathymétrique. Les rayons choisis sont égaux à 8 et 32 dans cet exemple.	125
Figure 88 : Représentation du rayon d'influence de la déformation locale. En a) les points du modèle de référence ont été superposés à l'estompage du MNT bathymétrique déformé. En b) ont été rajoutés les points déformés avec une déformation locale d'un rayon égal à 8 (points bleus). Les points caractéristiques qui ont servi à la déformation sont entourés d'une ellipse rouge. En c) sont visualisés en rouge les points déformés localement avec un rayon égal à 32. d) montre la superposition de tous les points, de référence en jaune, déformés avec un rayon de 8 en bleu, et avec un rayon de 32 en rouge.	126
Figure 89 : Résultat de différentes méthodes de déformation élastique sur les modèles topographique et bathymétrique. En a) représentation des points de départ ; b) estompage de ces points et visualisation des erreurs introduites dans chaque modèle ; c) modèle déformé globalement sans translation ; d) modèles déformés globalement avec translation préalable.	127
Figure 90 : Superposition des modèles intégrés sous forme de points. En jaune : modèle déformé avec une méthode de déformation élastique globale (DEG) ; en bleu : modèle translaté puis déformé avec une déformation élastique globale (T+ DEG).	128
Figure 91 : Résultat de différentes méthodes de déformation élastique avec un rayon sur les modèles topographique et bathymétrique. En a) modèle déformé selon un rayon de 8 ; b) modèle translaté puis déformé selon un rayon de 8 ; c) modèle déformé selon un rayon de 32 ; d) modèle translaté puis déformé selon un rayon de 32.....	129
Figure 92 : Détail du déplacement des points selon un rayon $R = 8$	129
Figure 93 : Carte de différences altimétriques entre le modèle d'origine et le modèle bruité bathymétrique.....	130
Figure 94 : Calcul du biais moyen à partir des couples de points homologues.	131
Figure 95 : Exemple de couples de points appariés entre la ligne caractéristique de référence en rouge et la ligne bruitée en bleu. Les traits violets représentent les couples.....	132
Figure 96 : Exemple de couples de points appariés entre deux lignes caractéristiques. Les traits entre les points représentent les couples. La visualisation avec le logiciel SIS® permet d'affecter à chaque trait une couleur en fonction de sa valeur : plus le trait entre les points est rouge, plus la distance (et donc le biais local) est importante.	132
Figure 97 : Soustraction du modèle de référence avec le modèle topographique déformé rigide translaté (TRANS).....	133
Figure 98 : Comparaison des modèles DEL8 et DEL32 avec le modèle de référence.	134

Figure 99 : Superposition du modèles de référence en rouge et du modèle bruité en bleu. Altitudes exagérées 5 fois. La ligne rouge représente la ligne caractéristique (rupture de pente du modèle d'origine) de référence.	135
Figure 100 : Comparaison des modèles DEL32+T, TRANS, DEG et DEL8+T avec le modèle de référence.	135
Figure 101 : Soustraction du modèle de référence et du modèle bathymétrique.	136
Figure 102 : Soustraction du modèle de référence avec le modèle topographique.	137
Figure 103 : Chaîne de traitement des données : production des modèles topographique et bathymétrique bruités.	138
Figure 104 : Chaîne de traitement des données : digitalisation des lignes de référence et export des points pour le calcul de l'appariement.	139
Figure 105 : Chaîne de traitement des données : calcul des points homologues.	141
Figure 106 : Chaîne de traitement des données : calcul des déformations globale et locale.	142
Figure 107 : Chaîne de traitement des données : déformations absolue et élastique des MNT. Visualisation du résultat dans un SIG.	143
Figure 108 : Soustraction du modèle de référence avec le modèle topographique translaté.	144
Figure 109 : Soustraction du modèle de référence et du modèle bathymétrique translaté.	145
Figure 110 : Soustraction du modèle de référence avec le modèle topographique déformé élastiquement (R =8).	146
Figure 111 : Soustraction du modèle de référence avec le modèle topographique déformé élastiquement (R =32).	146
Figure 112 : Soustraction du modèle de référence avec le modèle bathymétrique déformé élastiquement avec un rayon de 32.	147
Figure 113 : Soustraction du modèle de référence avec le modèle topographique déformé élastiquement et translaté (R =8).	147
Figure 114 : Localisation du trait de coupe. La carte de soustraction entre le MNT de référence et le MNT DEL32 montre l'emplacement du trait vis-à-vis de la bande de déformation.	148
Figure 115 : Comparaison des profils du MNT de référence et des deux MNT extraits, topographique et bathymétrique.	149
Figure 116 : Comparaison des modèles DEL8 et DEL8+T avec les modèles de référence, bathymétrique et topographique.	149
Figure 117 : Superposition du modèle topographique en marron et du modèle bathymétrique en bleu. L'insuffisance de la correction du modèle topographique induit un biais altitudinal important visible sur l'image. Altitudes exagérées 5 fois.	150
Figure 118 : Visualisation du MNT terre/mer issu de la déformation DEL8 des deux modèles terrestre et maritime bruités. Altitudes exagérées 5 fois. Le trait noir représente le trait de coupe.	150
Figure 119 : Comparaison des modèles DEG et TRANS avec les modèles de référence, bathymétrique et topographique.	151
Figure 120 : Comparaison de l'artéfact des modèles fusionnés DEL8, DEL32 et DEG.	151
Figure 121 : Comparaison de l'artéfact des modèles fusionnés DEL8+T, DEL32+T et TRANS.	152
Figure 122 : Comparaison de l'artéfact sur les modèles fusionnés.	152
Figure 123 : Détail de la seconde phase : analyse paysagère en vue de l'extraction des lignes nécessaires au processus d'appariement.	155
Figure 124 : Méthodologie Globale d'intégration complétée de l'analyse paysagère.	164

TABLE DES TABLEAUX

Tableau 1 : Matrice de distance (euclidiennes) et matrice de Fréchet des lignes de la Figure 12.....	32
Tableau 2 : Sources et techniques de calcul du positionnement d'un jeu de points de référence.	45
Tableau 3 : Caractéristiques des MNT selon les données source et les techniques utilisées, d'après [Charleux-demargne, 2001]......	51
Tableau 4 : résumé des points forts et faiblesses des différentes études présentées dans ce chapitre.....	58
Tableau 5 : Matrice de distance (euclidienne) et matrice de Fréchet des lignes de la figure 29	64
Tableau 6 : Le « chemin minimum » des lignes $(L_{1,1}, L_{2,1})$ et $(L_{1,n}, L_{2,m})$ est défini en sélectionnant les couples de réels en grisés.....	65
Tableau 7 : Evolution des valeurs de d_{dF} et d_{aF} selon le taux de ré échantillonnage des lignes de la figure 29.....	66
Tableau 8 : Matrice de Distance des lignes homologues partielles représentées figure 34.	68
Tableau 9 : Matrices de Distance et de Fréchet des lignes fermées de la figure 36.....	70
Tableau 10 : Matrice de distance résultant de l'inversion de la ligne L_2 , présentée figure 37. Le cheminement optimal est représenté en gris. Pour cet exemple $d_{Fd} = 6,08$	71
Tableau 11 : Matrice de Distance résultante de la dernière permutation, distance de Fréchet.	72
Tableau 12 : Matrice de Distance et Matrice de Fréchet des lignes L_1 et L_2 présentées figure 45.....	76
Tableau 13 : Matrice de Distance et de Fréchet des lignes de la figure 45 avec les points de la ligne L_1 inversée.....	77
Tableau 14 : Récapitulatif des distances d'appariements déjà traitées et celles introduites dans cette section.	78
Tableau 15 : Récapitulatif des cas possibles d'appariement PP et 2PP.	84
Tableau 16 : Résumé des différents appariements présentés et de leurs distances de Fréchet associées.	86
Tableau 17 : Récapitulatif des différents cas de figure rencontrés lors de la comparaison de lignes (1) pour chaque type de ligne (ouverte ou fermée) ; (2) pour chaque type de recouvrement : total, inclus ou partiel (rectangles englobants simple et étendu associés).	88
Tableau 19 : Distance en mètres entre les lignes saisies et les lignes de référence : distance de Fréchet discrète partielle moyenne suivie de la distance de Fréchet discrète partielle entre parenthèses (précision décimétrique).	91
Tableau 20 : Tableaux récapitulatifs des apports de la thèse et des domaines de recherche nécessitant un approfondissement.	163

Développement d'une approche SIG pour l'intégration de données Terre/Mer

RESUME : Le domaine maritime littoral constitue un milieu fragile, placé à l'interface Terre/Mer où s'exercent de nombreuses pressions issues des milieux terrestres, maritimes et atmosphériques. Les approches globales de Gestion Intégrée des Zones Côtières (GIZC) ont pour but de pallier les lacunes des politiques précédentes au niveau des interactions simultanées des différents acteurs du littoral. Cette approche requiert un Modèle Numérique de Terrain (MNT) littoral continu ainsi que des données maritimes et terrestres cohérentes entre elles. En raison d'acquisitions ou de productions différentes pour les MNT terrestres et maritimes, l'ensemble du relief n'est pas traité de manière ni homogène ni continue. De plus les techniques d'intégration des MNT actuelles entraînent des lissages des reliefs, voire la perte des éléments caractéristiques.

Des outils et une nouvelle méthodologie d'intégration de MNT ont donc été définis. Les outils utilisent la distance de Fréchet pour calculer une distance maximale d'écartement entre les points homologues de deux lignes. La valeur de cette distance donne une mesure de ressemblance de forme qui sert de base à leur appariement ultérieur. La méthodologie de fusion proposée dans cette thèse se veut générique tout en s'appuyant sur les lignes caractéristiques spécifiques du paysage. Elle effectue un enchaînement de traitements répartis en trois étapes majeures : 1) transformations cartographiques et analyse des zones de recouvrement ; 2) segmentation et extraction de lignes caractéristiques ; 3) appariement et fusion à l'aide de déformations élastiques. Ces outils d'intégration ont été testés sur des données réelles en 2D pour la comparaison du trait de côte, puis validés sur des MNT simulés 3D.

Mots clés : Modèle numérique de terrain (MNT), intégration terre/mer de MNT, géomorphologie littorale, déformation élastique, Systèmes d'Information Géographique (SIG), distance de Fréchet, appariement et fusion de données.

GIS approach development for seamless land/sea data integration

ABSTRACT : The coastal domain is a delicate environment standing at the land/sea border, an area where the terrestrial, the maritime and the atmospheric areas all exert their own forms of influence and pressure. Today's more comprehensive management of the coasts aims to overcome the failings demonstrated in previous policies by considering all simultaneous interactions from the various littoral actors. This approach requires a seamless Digital Elevation Model (DEM) with coherent terrestrial and maritime data. Owing to the diversity in the sources of captured data and in the methods of DEM production, the relief as a whole is rarely processed with continuity and homogeneity. Moreover, the current DEM integration techniques usually result in smoothing the relief or in the loss of the relief's main features.

That is why new tools and a whole DEM integration method were defined. These tools use the Fréchet distance to compute a maximum distance gap between homologous points from two lines. The value of this distance gives a shape similarity measure which is used as a reference in the later matching process. The merging methodology presented in this PhD purports to be generic with a specific landscape feature line basis. It performs a sequence of processing divided into three main steps: 1) cartographic transformations and overlapping area analysis; 2) segmentation and feature lines extraction; 3) DEM matching and merging with rubber sheeting deformations. Those integration tools were tested in 2D on real data for the coastline comparison, and then validated on simulated DEM.

Keywords : Digital elevation model (DEM), seamless land/sea DEM integration, coastal geomorphology, rubber sheeting, Geographical Information System (GIS), Fréchet distance, data matching, data merging.