HAL
open science

Fusion de données satellitaires pour la cartographie du potentiel éolien offshore

Mohamed Bassam Ben Ticha

To cite this version:

Mohamed Bassam Ben Ticha. Fusion de données satellitaires pour la cartographie du potentiel éolien offshore. domain_stic.inge. École Nationale Supérieure des Mines de Paris, 2007. Français. NNT : 2007ENMP1480 . tel-00198912

HAL Id: tel-00198912

https://pastel.hal.science/tel-00198912v1

Submitted on 18 Dec 2007

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.

ED n° 84 : « Sciences et technologies de l'information et de la communication »

N° attribué par la bibliothèque

|_|_|_|_|_|_|_|_|_|_|_|

T H E S E

pour obtenir le grade de :

DOCTEUR DE L'ECOLE DES MINES DE PARIS

Spécialité "Informatique temps réel, Robotique, Automatique"

présentée et soutenue publiquement par

Mohamed Bassam BEN TICHA

le 05 novembre 2007

FUSION DE DONNÉES SATELLITAIRES POUR LA CARTOGRAPHIE DU POTENTIEL ÉOLIEN OFFSHORE

Directeur de thèse: Thierry RANCHIN

Jury :

M. Jean-Paul MARMORAT ..... Président
M. Bertrand CHAPRON ..... Rapporteur
M. René GARELLO ..... Rapporteur
M. Philippe BOLON ..... Examinateur
M. Philippe REFRÉGIER ..... Examinateur
M. Nicolas FICHAUX ..... Examinateur
M. Eric PEIRANO ..... Examinateur
M. Thierry RANCHIN ..... Examinateur

Fusion de données satellitaires pour la cartographie du potentiel éolien offshore

Mohamed Bassam BEN TICHA

Remerciements

La rédaction d'une thèse est un exercice qui impose beaucoup de moments difficiles et de moments de doute mais qui aboutissent à des phases de bonheurs. Un des moments de bonheur a été de proposer ce travail à la lecture des membres du jury que je remercie d'avoir relu avec attention ce manuscrit. Un autre moment de bonheur est d'écrire ces quelques lignes pour remercier tous ceux aux côtés desquels j'ai beaucoup appris ces dernières années. Un apprentissage qui va bien au-delà de l'aspect scientifique.

D'abord je tiens à remercier mon directeur de thèse Thierry Ranchin. J'ai eu beaucoup de chance de t'avoir comme directeur de thèse et comme soutien quand il le fallait. Je tiens aussi à remercier Lucien (Wald) et François (Cauneau) pour leurs conseils avisés.

Merci à mes co-bureau grâce à qui venir travailler n'a pas été une corvée. Merci donc à Anne-Lise Beaulant, qui saura apprécier les joies de la vie toulousaine, et à Armel Oumbé, même si en matière de foot nous ayons quelques divergences qui se dissiperont quand les Aigles auront battu les Lions : il comprendra. Je tiens aussi à remercier tous ceux grâce à qui les pauses café ont été des moments agréables tous les jours. Merci à tous les membres du CEP de Sophia-Antipolis et surtout aux administratifs dont l'aide a été précieuse.

En dehors du laboratoire je remercie tous ceux dont l'amitié me rend fier et qui ont toujours été là. J'ai retrouvé une citation qui illustre bien ce que représente votre amitié.

« Ce n'est pas tant l'aide de nos amis qui nous aide que notre confiance dans cette aide. » Epicure

Un grand merci à Iyadh et Riadh, à Akrem, Mehdi, Wala et Aymen. Ces trois années, sur la Côte d'Azur, ont été très agréables, grâce à vous. J'ai une pensée à mes amis un peu partout sur cette vaste planète. Les chemins de la vie font que nous nous séparons et que nous nous retrouvons au gré des divers déménagements... notre amitié reste intacte.

Une pensée aussi à mes grands-parents toujours présents, par les forces de l'esprit, pour m'épauler. Enfin, une pensée à mes parents et mes sœurs pour leur soutien inconditionnel. Vous serez toujours tout près de moi, dans mes pensées, quelque soit la distance matérielle nous séparant.

Table des matières

1 Introduction **9**

1.1 Contexte du développement de l'éolien offshore 9

1.2 Les enjeux de la cartographie de la ressource éolienne 9

1.3 Objectif de la thèse 12

1.4 Démarche de la thèse 13

1.5 Structure du manuscrit 14

Bibliographie 15

2 Cartographie du potentiel éolien : mesures de vent, techniques de modélisation et moyens de cartographie **17**

2.1 Cartographie du potentiel éolien 17

2.1.1 L'aspect spatial 18

2.1.2 L'aspect temporel 18

2.1.3 Conclusions sur les caractéristiques des mesures de vent nécessaires pour la cartographie du potentiel éolien offshore 24

2.2 Mesures du vent en mer 25

2.2.1 Mesures sur site 26

2.2.2 Mesure du vent en mer par télédétection 27

2.2.3 Conclusions sur l'adéquation des moyens de mesure du vent en mer et la cartographie du potentiel éolien 34

2.3 Modèles pour l'estimation du potentiel éolien 35

2.3.1 Extrapolation verticale des vitesses de vent 35

2.3.2 Modèles empiriques 36

2.3.3 Modèles statistiques 37

2.3.4 Modèles physiques 39

2.4 Conclusion 42

Bibliographie 43

3 Fusion de données pour l'évaluation du potentiel éolien **51**

3.1 Fusion de données, notations et définitions 51

3.1.1 Définitions 52

3.1.2 Représentation d'une opération de fusion 52
3.2 Schéma de fusion pour la cartographie du potentiel éolien offshore 53
3.2.1 Hypothèses du schéma de fusion 54
3.2.2 Schéma de la méthode fusion 55
3.3 Fonction de transfert de la basse à la haute résolution spatiale 59
3.3.1 Analyse multi-échelle 60
3.3.2 Concept ARSIS 61
3.4 Conclusions 64
Bibliographie 64

4 Classification des champs de vent 67
Première partie : Introduction 67
Deuxième partie : A wind field classification scheme for generation of typical spatial patterns 69
4.1 Introduction 69
4.2 Classification scheme 71
4.2.1 Clustering method 73
4.2.2 Reassignment 75
4.3 Application to a case study : Irish Sea 75
4.3.1 Clustering and selection of the number of classes 75
4.3.2 Reassignment of the rare situations 76
4.3.3 Final results 76
4.4 Evaluation 77
4.5 Conclusions 79
Bibliography 83

5 Synthèse de la haute résolution spatiale 89
Première partie : Introduction 89
Deuxième partie : Fusion of SAR images and scatterometer data for wind resource assessment 91
5.1 Introduction 91
5.2 Data fusion method for offshore wind resource mapping 93
5.3 High spatial resolution wind fields synthesis 94
5.3.1 ARSIS concept 95
5.3.2 Synthesis of high spatial resolution wind fields 96
5.4 Application to some examples 97
5.4.1 Data presentation 97
5.4.2 Results 99
5.5 Quality assessment 101
5.6 Conclusion 105

Bibliography 106

6 Evaluation de la qualité de la méthode de fusion pour la cartographie du potentiel éolien 109

6.1 Introduction 109

6.2 Données utilisées et protocole d'évaluation 112

6.3 Analyse des résultats 115

6.3.1 Erreur sur les vitesses de vent 115

6.3.2 Erreur sur l'estimation des paramètres statistiques 120

6.4 Conclusions 124

Bibliographie 125

7 Conclusions et perspectives 127

A Mesures satellitaires du vent : physique de la mesure 131

A.1 Surface de la mer 131

A.2 Physique de la mesure des radars actifs 132

A.2.1 Équation radar 133

A.2.2 Réflexion spéculaire 134

A.2.3 Rétrodiffusion de Bragg 135

Bibliographie 135

B Estimateur de maximum de vraisemblance et intervalles de confiance 137

B.1 L'estimateur de maximum de vraisemblance 137

B.2 Région de confiance jointe associée une estimation 138

Chapitre 1

Introduction

1.1 Contexte du développement de l'éolien offshore

Le développement durable est défini par le rapport Brundtland [1] comme un développement qui répond aux besoins actuels sans compromettre la capacité des générations futures à satisfaire les leurs. Le volet énergétique est l'une des questions importantes dans l'optique d'un tel développement. L'humanité doit faire face à deux problématiques centrales : l'épuisement progressif des ressources fossiles et l'impact de l'utilisation d'énergies polluantes sur l'environnement. L'énergie éolienne de par son aspect renouvelable et l'absence d'émissions de gaz à effet de serre représente une des composantes d'une politique énergétique permettant de réaliser un développement durable.

L'installation de parcs éoliens offshore permet de profiter d'un vent plus régulier, plus fort qu'à terre [2] et donc d'une production plus grande et plus régulière d'énergie. Elle permet, en outre, la réduction des nuisances visuelles et sonores et des problèmes de voisinage rencontrés par l'éolien à terre. Il est, cependant, primordial d'optimiser les coûts des parcs éoliens offshore [3]. Ceci passe notamment par un choix judicieux des lieux d'implantation des parcs éoliens. Plusieurs facteurs rentrent en jeu lors du choix du site d'une ferme éolienne offshore : climatologie du vent sur la région, facilité d'accès, bathymétrie, impact environnemental, ... La climatologie du vent sur le site, c'est à dire la fréquence d'occurrence des différentes vitesses du vent selon les différentes directions, est l'un des facteurs importants. En effet, c'est cette climatologie qui conditionne directement la quantité d'énergie électrique pouvant être produite sur ce site (ou ressource énergétique). L'évaluation de la distribution spatiale de la ressource énergétique éolienne en mer est donc un enjeu important. Lors de l'étape du choix du site d'implantation, les investisseurs veulent disposer d'une cartographie précise de la ressource énergétique.

1.2 Les enjeux de la cartographie de la ressource éolienne

La ressource énergétique éolienne (ou potentiel éolien) d'un site est définie par [4] comme étant la quantité d'énergie électrique extractible durant un certain laps de temps, typiquement un an, à partir de l'énergie cinétique contenue dans le vent par une éolienne installée sur ce site.

« L'estimation de la ressource éolienne s'étend de l'estimation générale de l'énergie moyenne conte-

nue dans le vent au dessus d'une vaste région – évaluation régionale – au calcul de la production moyenne annuelle d'un aérogénérateur spécifique pour un site donné – appelé siting » [5].

Ces différents types d'estimation de la ressource éolienne sont effectués à différents moments de l'avancement d'un projet de parc éolien. La vie d'un parc éolien commence par l'étape de sélection de la région d'implantation du parc éolien. Lors de cette étape, une évaluation régionale est suffisante pour une estimation rapide de la ressource éolienne d'un site. Elle a pour but de fixer de manière approximative la région d'implantation du parc éolien. En ce qui concerne les données de vent, il est nécessaire de fournir une vitesse moyenne de vent, les directions dominantes et les valeurs extrêmes du vent sur les régions prospectées. Cette première étape est suivie par une étude de faisabilité. Un choix plus précis du site d'implantation du parc éolien est alors effectué. L'objectif est d'évaluer économiquement et techniquement le site choisi. Une estimation de la production annuelle d'énergie est effectuée. Pour cela, des données plus précises de vent sont nécessaires pour réduire les incertitudes. Des études de terrain sont, alors, effectuées pour connaître précisément les caractéristiques du site choisi. Cette étape, de choix précis du site d'implantation du parc éolien et son évaluation, est la plus coûteuse en termes de temps, d'argent et de données à collecter. Elle est suivie par la demande des permissions administratives. La dernière étape, avant l'exploitation effective du parc éolien, est la conception de ce parc et sa construction. La conception du parc nécessite des données de très grande précision pour une optimisation de la disposition des éoliennes du parc. Ceci est, généralement, entrepris en modélisant les turbulences et les sillages engendrés par les turbines.

Après la première étape de sélection de la région d'implantation du parc éolien, pour affiner le choix de l'implantation, il est nécessaire d'avoir des cartes de potentiel éolien à haute résolution spatiale. En effet, ces cartes doivent rendre compte de la variabilité du potentiel à l'échelle d'un parc éolien. L'échelle caractéristique d'un parc éolien étant de l'ordre du kilomètre, la résolution spatiale de la cartographie du potentiel éolien doit être de cette ordre (ou mieux si possible). Les estimations de potentiel éolien qui apparaissent sur ces cartes sont directement liées à la climatologie du vent sur la zone. La représentation statistique de la climatologie du vent se fait au travers de la fonction de densité de probabilité des vitesses de vent. Cette fonction donne la probabilité d'occurrence de chaque vitesse du vent. On cherche donc à retrouver la distribution statistique de la fréquence des vitesses et des directions de vent sur une zone donnée. L'estimation de la fonction de densité de probabilité associée à un site permet de résumer la variabilité temporelle du vent, sur site, au travers de quelques paramètres statistiques. A partir de cette fonction, le calcul de la production moyenne de n'importe quelle turbine installée sur un site revient à intégrer, dans le temps, le produit de cette fonction et la courbe de puissance de la turbine [5]. La courbe de puissance d'une turbine est la puissance produite par la turbine en fonction de la vitesse du vent. Elle est fournie par les constructeurs des turbines. La fonction de probabilité, généralement utilisée pour modéliser les séries temporelles de vent, est celle de Weibull à deux paramètres [5]. Pour pouvoir faire cette description statistique, il est nécessaire d'avoir une série temporelle des vitesses et directions du vent au point où la ressource doit être décrite. Cette série temporelle doit couvrir plusieurs années.

Le moyen de mesure le plus précis de vitesses et de directions du vent est d'installer un mât au point où la ressource doit être évaluée. Cependant, il est impossible, économiquement, d'installer un mât et

d'attendre plusieurs années pour acquérir une telle série de données. De plus, les mesures du mât ne sont représentatives que d'une zone limitée autour du point où elles ont été prises. Ceci est particulièrement vrai en zone côtière où la variabilité spatiale du vent est importante. Il est donc nécessaire de recourir à d'autres méthodes de mesure ou à la modélisation.

A terre, la description statistique du vent en un point est obtenue en interpolant des données disponibles à partir de stations météorologiques avoisinant le site prospecté [6]. Les données de longue durée mesurées par une station météorologique sont corrélées avec des données collectées en utilisant un mât, installé sur le site étudié, et permettant d'acquérir les statistiques de vent sur une durée limitée (au moins un an). A terre, le coût d'installation d'un mât et de l'acquisition des données reste abordable (15 000 euros en moyenne). En mer, selon [7], le coût de l'installation d'un mât pour une année de mesures s'élève à 750 000 euros. De plus, la faible densité du réseau météorologique en mer pose le problème de disponibilité de mesures de vent dans le voisinage du site prospecté. A ce problème de disponibilité de mesures sur site, s'ajoute le problème de manque de représentativité de la variabilité spatiale du vent de ces mesures ponctuelles. Ce problème se pose surtout en zone côtière, zone privilégiée d'implantation des parcs éoliens offshore.

Depuis quelques années, aux sources « classiques » de mesure de vent (stations météorologiques, instruments à bord de bateaux,...), se sont ajoutés des instruments spatioportés de télédétection : radars altimètres [8], radars diffusomètres [9], radars à ouverture synthétique [10], radiomètres [11], ... Ces sources de données présentent l'avantage d'être disponibles depuis de nombreuses années (depuis le début des années 1990 pour la plupart). Une archive de chacune de ces sources de données existe et croît de jour en jour. Il n'y a donc pas de délai d'attente comme c'est le cas pour les mesures sur site. Ces instruments de télédétection permettent d'accéder à une information spatiale sur le vent en mer. Ces cartes représentent des mesures de vitesse et de direction du vent sur un grille couvrant la zone d'acquisition des mesures. Elles permettent donc d'accéder à la variabilité spatiale du vent.

Ces différents instruments permettent d'acquérir des mesures à différentes résolutions spatiales et avec différents échantillonnages temporels. Une limitation à l'utilisation de ces données est l'absence de capteurs alliant la haute résolution spatiale et la haute résolution temporelle. Les hautes résolutions spatiales et temporelles sont relatives aux besoins de la cartographie du potentiel éolien offshore. La haute résolution spatiale est nécessaire pour atteindre la résolution de l'ordre du kilomètre nécessaire à l'industrie éolienne offshore. La haute résolution temporelle est nécessaire pour garantir que les mesures de vent sont représentatives des variations temporelles du vent. Les instruments de télédétection permettant de fournir des mesures à une résolution spatiale suffisante pour la cartographie de la ressource éolienne offshore, ont un échantillonnage temporel ne permettant pas une description statistique précise du vent. A l'opposé, il existe des capteurs dont l'échantillonnage temporel est suffisant pour cette description statistique mais dont la résolution spatiale, plus grossière, ne permet pas d'accéder aux petites échelles de la variation spatiale du vent. Aucune de ces sources ne permet donc, à elle seule d'établir une cartographie du potentiel éolien. Dans la suite, on désignera par haute résolution une résolution suffisante pour la cartographie de la ressource éolienne, avec la précision souhaitée par l'industrie éolienne offshore, et par basse résolution une résolution qui ne l'est pas. La haute résolution spatiale est, dans le cadre de cette

application, une résolution de l'ordre du kilomètre. On définit la haute résolution spatiale comme étant une résolution permettant un calcul fiable des paramètres statistiques de la climatologie du vent. Cette résolution temporelle est étudiée au chapitre 2.

Une solution pour la cartographie de la ressource éolienne, tel que proposée par [12], est de combiner différentes données disponibles pour profiter des apports de chacune d'entre elles. C'est une approche de ce type qui est retenue dans le cadre de cette thèse. Ainsi, les données à haute résolution spatiale apporteront l'information sur la variabilité du vent aux petites échelles spatiales. Les données ayant le meilleur échantillonnage temporel permettront, quant à elles, d'accéder à la composante temporelle de la variabilité du vent. Le travail exposé dans cette thèse s'appuie donc, principalement, sur le traitement de l'information et plus précisément sur la fusion de données.

1.3 Objectif de la thèse

Cette thèse s'inscrit dans l'évaluation de la ressource éolienne à haute résolution spatiale. Cette évaluation se présente sous la forme d'une description statistique de la climatologie du vent sur une zone donnée. Dans cette thèse, on propose une méthode tirant profit de différentes sources de données de vent pour arriver à cette description. On considère donc la problématique de la fusion de données, issues de différentes sources, pour la description statistique du vent.

L'objectif de cette thèse est d'étudier la pertinence de la fusion de données pour la cartographie du potentiel éolien offshore. A terme, le but est une application opérationnelle de cette méthode pour la cartographie du potentiel éolien offshore.

La méthode proposée permet une utilisation synergique de données sur le vent. Les données utilisées sont des données archivées ce qui permet d'éviter les délais d'attente imposés par l'installation d'un mât et l'attente de l'acquisition de mesures. En outre, cette méthode permet d'éviter les coûts liés à l'installation du mât et à sa maintenance. Les mesures acquises par le mât ne sont représentatives que d'une zone très limitée à proximité du point de mesure. Les données de télédétection permettent, quant à elles, de mesurer la variabilité spatiale du vent. L'objectif est de tirer profit de différents types d'instruments de télédétection pour aboutir à une cartographie de la ressource éolienne offshore.

La méthode proposée est appliquée à un cas d'étude situé en Mer d'Irlande. Un parc éolien offshore est installé au large des côtes irlandaises. Ce parc est connu sous le nom d'Arklow Bank. Cette application permet d'évaluer les performances de la méthode de fusion de données. L'évaluation des résultats permet de comparer les performances de la méthode de fusion proposée à celle de méthodes de cartographie existantes. Elle permet, en outre, de donner des pistes d'amélioration de la méthode.

Cette situation, où aucune des sources de données ne présente à elle seule toute les propriétés nécessaires pour décrire un phénomène, n'est pas propre à l'éolien. En effet, la haute résolution spatiale, quelque soit le phénomène observé, est souvent disponible aux dépens d'une répétitivité dans le temps beaucoup moins bonne que celle associée à la basse résolution spatiale. L'approche proposée dans cette thèse reste applicable dans toutes les situations où la combinaison de la résolution spatiale de certaines données et la résolution temporelle d'autres données permet la description spatio-temporelle du phéno-

mène étudié. De telles situations sont rencontrées pour la cartographie de :

- le rayonnement solaire [13],
- la pollution atmosphérique [14],
- la pluviométrie [15],
- ...

Historiquement, les premières applications de la fusion de données ont été militaires [16]. Depuis, la fusion de données s'est révélée être un cadre intéressant pour beaucoup d'autres applications telles que la robotique, la médecine, ... Chacune de ces applications a des objectifs, des capteurs (d'information, de mesure), un environnement particulier [17]. Dans le domaine de la télédétection, [18] donne la définition suivante de la fusion de données : « *La fusion de données constitue un cadre formel dans lequel s'expriment les moyens et techniques permettant l'alliance des données provenant de sources diverses* ». La définition ajoute que la fusion de données vise à l'obtention d'information de plus grande qualité ; le définition exacte de « plus grande qualité » dépendant de l'application.

Cette définition montre la nécessité de définir l'objectif visé par la fusion de données, de définir les sources de données à utiliser, les moyens et les techniques nécessaires pour effectuer la fusion. En outre, elle montre qu'il est important de pouvoir mesurer la qualité de l'information en sortie. Dans la démarche, présentée dans la section suivante et suivie par cette thèse, on essaye de répondre à ces points.

1.4 Démarche de la thèse

La première étape de la cartographie de la ressource éolienne offshore est l'analyse des différentes sources de données sur le vent en mer et de leur adéquation avec les besoins de cette cartographie. Cette analyse nous a conduit à choisir deux sources de données présentant des propriétés complémentaires pour la cartographie de la ressource éolienne. L'une de ces sources possède une résolution temporelle suffisante pour répondre aux exigences du problème de cartographie posé mais une basse résolution spatiale. L'autre source permet d'accéder à des données à haute résolution spatiale mais avec une résolution temporelle moins bonne.

Le schéma de fusion proposé dans cette thèse consiste à élaborer des fonctions de transfert de la basse à la haute résolution spatiale. Elles sont appliquées aux données à haute résolution temporelle et basse résolution spatiale pour générer une série synthétique de données à haute résolution spatiale et haute résolution temporelle. Le transfert de la basse à la haute résolution spatiale, pour des raisons physiques ne peut pas être basé sur une fonction universelle. On tire avantage du nombre limité de configurations spatiales du champ de vent au dessus d'une région donnée pour définir un nombre limité de fonctions simples appelées fonctions de transfert types. Ces fonctions sont définies pour l'ensemble des champs de vent présentant une configuration de vent similaire dans l'espace de données à basse résolution spatiale. A chaque configuration du vent à basse résolution spatiale est associée une observation à haute résolution spatiale. Cette observation à haute résolution spatiale sert à élaborer la fonction de transfert correspondant à cette configuration typique du champ de vent. Les structures à haute résolution spatiale présentes dans cette observation servent donc à inférer les structures qui doivent être présentes dans les

situations présentant la même configuration à basse résolution spatiale.

L'élaboration de ces fonctions de transfert type s'appuie sur une analyse multi-échelles des différentes données. Une telle analyse permet de relier les observations à basse et haute résolution spatiale. Pour retrouver les différentes configurations typiques du champ de vent sur une région donnée, il est nécessaire de recourir à une classification des observations.

La méthode de fusion développée dans cette thèse repose sur deux outils essentiels : la représentation multi-échelles de l'information et la classification des données. Cette méthode est appliquée à un cas pratique pour évaluer sa qualité. Une méthode d'estimation d'intervalles de confiance associés à l'estimation est proposée.

Cette démarche nous conduit à la structure du manuscrit décrite dans la section suivante.

1.5 Structure du manuscrit

Dans le chapitre 2, on présente les outils permettant de modéliser la ressource éolienne. Cette modélisation se fait au travers de cartes de paramètres statistiques décrivant la climatologie du vent. Ensuite, on présente les différentes sources de mesures du vent en mer. L'adéquation des caractéristiques de chacun de ces instruments avec la cartographie de la ressource éolienne offshore est étudiée. Les sources de données que nous utiliserons, dans la suite de ce travail, sont choisies. Un état de l'art des modèles utilisant ces données pour évaluer la ressource éolienne offshore est proposé. Il en ressort qu'aucun modèle ne répond complètement aux besoins de la cartographie de la ressource éolienne. Certains modèles se contentent d'évaluer la ressource en un point. D'autres modèles souffrent d'un manque de données nécessaires pour leur bon fonctionnement. Enfin, certains modèles sont inapplicables en pratique.

Le chapitre 3 présente la méthode proposée, basée sur la fusion de données, pour répondre aux carences constatées dans l'état de l'art. Cette méthode est basée sur le calcul de fonctions de transfert types. Ces fonctions de transfert permettent de passer de la basse résolution à la haute résolution spatiale. Elles sont appliquées aux données à haute résolution temporelle et à basse résolution spatiale pour obtenir des données synthétiques alliant la haute résolution temporelle et la haute résolution spatiale. Comme expliqué précédemment, ces fonctions de transfert type sont basées sur la séparation des champs de vent en configurations typiques. Ces configurations sont retrouvées à l'aide d'une méthode de classification. Cette méthode est présentée dans le chapitre 4. Elle permet de regrouper les champs de vent ayant le même comportement spatial en quelques situations types. Dans ce chapitre, on étudie la validité de la représentation du comportement spatial du vent par les situations types résultant de la classification.

Le chapitre 5 expose la méthode de la synthèse de la haute résolution spatiale, au travers de la génération et de l'application de fonctions de transfert types. Cette méthode est dérivée du concept ARSIS (Amélioration de la Résolution Spatiale par Injection de Structures) [19] développé dans le cadre de la fusion d'images. Il est appliqué, dans le cadre de la thèse, à la fusion des champs de vent. La qualité des résultats est évaluée à l'aide du protocole établi par [20]. Le chapitre 6 contient l'application de la méthode qu'on a élaborée à un cas pratique. Alors que l'évaluation de l'étape présentée au chapitre 5 peut s'effectuer à l'aide de données de référence, il n'est pas possible d'évaluer dans ce cas les résul-

tats obtenus. Aussi nous transposons le problème, et appliquons la méthode à des données moins bien résolues. Cette application permet une évaluation quantitative de la qualité des résultats. Finalement, les conclusions de ce travail et quelques perspectives sont données.

Bibliographie

[1] World Commission on Environment and Development. *Our Common Future (The Brundtland Report)*. Oxford University Press, Oxford, UK, 398 p., 1987.

[2] A. R. Henderson. Offshore wind in europe : The current state of the art. *Refocus*, vol. 3, n° 2, pp. 14–17, 2002.

[3] J. Beurskens et P. H. Jensen. Wind energy. En ligne, 30 p. Disponible sur : `http://www2.ewea.org/documents/20_Beurskens_Hjuler_RD.pdf` (consulté le 20-08-2007).

[4] S. Krohn. Manuel de référence sur l'énergie éolienne, 2ème partie : Définitions de l'énergie et de la puissance. En ligne. Disponible sur : `http://www.windpower.org/fr/stat/unitsene.htm` (consulté le 20-08-2007).

[5] I. Troen et E. L. Petersen. *European Wind Atlas*. Risoe National Laboratory, Roskilde, Danemark, 656 p., 1989.

[6] H. P. Frank, O. Rathmann, N. G. Mortensen et L. Landberg. The numerical wind atlas — the KAMM/WAsP method. Technical Report Risoe-R-1252(EN), Risoe National Laboratory publications, Danemark, 60 p., 2001.

[7] European Wind Energy Association. Wind energy — the facts. En ligne, 2004. Disponible sur : `http://www.ewea.org/index.php?id=91` (consulté le 20-08-2007).

[8] G. S. Brown, H. R. Stanley et N. A. Roy. The wind-speed measurement capability of spaceborne radar altimeters. *IEEE Journal of Oceanic Engineering*, vol. OE-6, n° 2, pp. 59–63, 1981.

[9] R. K. Moore et W. L. Jone. Satellite scatterometer wind vector measurements — the legacy of the Seasat satellite scatterometer. *IEEE Geoscience and Remote Sensing Society Newsletter*, pp. 18–32, septembre 2004.

[10] C. C. Wackerman, C. L. Rufenach, R. A. Shuchman, J. A. Johannessen et K. L. Davidson. Wind vector retrieval using ERS-1 synthetic aperture radar imagery. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 34, n° 6, pp. 1343–1352, 1996.

[11] P. Gloersen et F. T. Barath. A Scanning Multichannel Microwave Radiometer for Nimbus-G and SEASAT-A. *IEEE Journal of Oceanic Engineering*, vol. OE-2, pp. 172–178, 1977.

[12] N. Fichaux. *Evaluation du potentiel éolien offshore et imagerie satellitale*. Thèse de doctorat, Ecole des Mines de Paris, Paris, France, 119 p., 2003.

[13] L. Wald, T. Ranchin, M. Lefèvre, M. Albuisson et J. Remund. Increasing the spatial resolution of gridded data by fusion with other data sets. In *Proceedings of the Sixth International Conference of Information Fusion, Cairns, Queensland, Australie, 8-11 Jul 2003*, vol. 1, pp. 181–184, 2003.

[14] D. Sarigiannis, N. Soulakellis et N. Sifakis. Information fusion for computational assessment of air quality and health effects. *Photogrammetric Engineering and Remote Sensing*, vol. 70, n° 2, pp. 235–245, 2004.

[15] F. S. Marzano, M. Palmacci, D. Cimini, G. Giuliani et F. J. Turk. Multivariate statistical integration of satellite infrared and microwave radiometric measurements for rainfall retrieval at the geostationary scale. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, n° 5, pp. 1018–1032, 2004.

[16] L. Valet, G. Mauris et P. Bolon. A statistical overview of recent literature in information fusion. *IEEE Aerospace and Electronic Systems Magazine*, vol. 16, n° 3, pp. 7–14, 2001.

[17] D. L. Hall et J. Llinas. An introduction to multisensor data fusion. *Proceedings of the IEEE*, vol. 85, n° 1, pp. 6–23, 1997.

[18] L. Wald. *Data Fusion : definitions and architectures. Fusion of images of different spatial resolutions*. Ecole des Mines de Paris, Paris, France, 198 p., 2002.

[19] T. Ranchin, B. Aiazzi, L. Alparone, S. Baronti et L. Wald. Image fusion. the ARSIS concept and some successful implementation schemes. *ISPRS Journal of Photogrammetry & Remote Sensing*, vol. 58, n° 1-2, pp. 4–18, 2003.

[20] C. Thomas et L. Wald. Assessment of the quality of fused products. In *Proceedings of the 24th EARSeL Symposium "New Strategies for European Remote Sensing", 25-27 May 2004, Dubrovnik, Croatia, Oluic (Eds.), Millpress, Rotterdam, ISBN 90 5966 003 X*, pp. 317–325, 2005.

Chapitre 2

Cartographie du potentiel éolien : mesures de vent, techniques de modélisation et moyens de cartographie

Le but de la cartographie du potentiel éolien est de fournir une information, sous une forme compacte, pour l'aide à la décision lors du choix des sites d'implantation de parcs éoliens. Cette cartographie présente la distribution spatiale et temporelle du vent. Dans la première partie de ce chapitre, on décrit les moyens de présenter de manière compacte la distribution spatiale et temporelle du vent. Les données nécessaires pour arriver à cette cartographie et leurs caractéristiques sont étudiées. Dans la deuxième partie, on présente les sources de données et les instruments de mesures du vent en mer. Les techniques de modélisation pouvant être utilisées pour arriver à une cartographie du potentiel éolien sont étudiées dans la dernière partie de ce chapitre.

2.1 Cartographie du potentiel éolien

Le vent est un phénomène continu dans l'espace et dans le temps. Pour la cartographie du potentiel éolien il est nécessaire de présenter ces variations continues de manière compacte. Dans cette section, on étudie les moyens de représenter l'information sur la climatologie du vent par quelques paramètres statistiques. On décrit les caractéristiques requises pour cette représentation pour avoir une information adaptée aux besoins de l'industrie éolienne offshore. On étudie les caractéristiques que doivent avoir les données sur le vent, notamment en termes de discrétisation temporelle et spatiale, pour garantir la fiabilité de la cartographie. On s'intéresse dans le premier paragraphe à la représentation de la distribution spatiale du potentiel éolien. Dans le deuxième paragraphe, on présente les moyens permettant de décrire de manière compacte les variations temporelles du vent.

2.1.1 L'aspect spatial

Les mouvements atmosphériques se produisent sur un large domaine d'échelles spatiales allant du mouvement aléatoire de molécules jusqu'à la circulation zonale moyenne prenant en compte l'atmosphère entier [1]. Les lois « basiques » de la physique, et plus précisément celles de la mécanique des fluides, décrivent le spectre entier des mouvements atmosphériques excepté l'échelle moléculaire. Les équations exprimant ces lois n'ont pas de solution générale [1]. Il est nécessaire de spécifier les échelles spatiales isolant les phénomènes d'intérêt pour chaque cas d'étude. Les mouvements à une certaine échelle spatiale ont une durée de vie associée. Ceci a été, notamment, exprimé par l'hypothèse de Taylor de turbulence gelée.

Le tableau 2.1 donne une correspondance approximative entre les échelles spatiales des mouvements atmosphériques, leur durée de vie et les échelles d'Orlanski associées.

Les parcs éoliens étant caractérisés par des tailles kilométriques, notre domaine d'étude se situe à la limite entre les micro et les méso-échelles. Les phénomènes atmosphériques à ces échelles ont une durée de vie de quelques minutes.

Taille (km)	0.001	0.01	0.1	1	10	100	1000	10000
Période	1 s	10 s	1 mn	10 mn	2 h	15 h	7 j	70 j
Échelle	**micro-échelles**			**méso-échelles**		**grandes échelles**		

TAB. 2.1 – Échelles des mouvements atmosphériques. Selon Orlanski [2].

Le vent est une quantité caractérisée par sa vitesse et sa direction. Étant continue dans le temps et dans l'espace, on désigne sa distribution spatiale à un instant donné par champ de vent. Dans la suite, on utilisera « champ de vent » pour désigner la distribution spatiale du vecteur vent sur une région donnée.

Ces champs de vent prennent, en pratique, la forme de cartes présentant une discrétisation de la région d'étude. Des valeurs de vitesse et de direction de vent sont données pour chaque maille (terme utilisé par les numériciens) ou pixel (terme venant du traitement d'images). La taille de ces mailles (ou pixels) doit, dans le cas de l'estimation du potentiel éolien offshore, être inférieure au kilomètre. En effet, la taille d'un parc éolien offshore est de l'ordre du kilomètre. Une résolution kilométrique est donc nécessaire pour détecter les variations du potentiel éolien à l'échelle d'un parc éolien. La cartographie du potentiel éolien consiste à décrire la variabilité temporelle du vent sur chacune de ces mailles.

2.1.2 L'aspect temporel

Les variations temporelles du vent, pour un site donné, peuvent être représentées de manière compacte par une fonction de densité de probabilité. Cette fonction donne la probabilité d'occurrence de chaque vitesse du vent sur ce site. Si la forme théorique de cette fonction est connue, les variations temporelles du vent peuvent être estimées à partir d'un nombre limité de réalisations.

Des études empiriques ont démontré que la distribution des vitesses du vent à la surface, sur terre ou en mer, peut être représentée par la fonction de densité de probabilité de Weibull à deux paramètres [3,

4, 5]. Cette approximation est valide aux différentes échelles allant de l'échelle globale jusqu'à l'échelle locale [6]. Elle est devenue un standard dans le domaine de l'industrie éolienne grâce, notamment, à son utilisation pour l'élaboration d'un atlas éolien danois [7] puis de l'atlas éolien européen [8].

La fonction de densité de probabilité d'une distribution de Weibull à deux paramètres est de la forme :

$$\begin{cases} p(v) = & \frac{k}{A}\left(\frac{v}{A}\right)^{k-1} e^{-\left(\frac{v}{A}\right)^k} & v > 0 \\ & 0 & v \leq 0 \end{cases} \tag{2.1}$$

avec $p(v)$ la probabilité d'occurrence des vents de vitesse v. k est appelé paramètre de forme et A paramètre d'échelle. k est sans dimension. A est en m s^{-1}.

La fonction de distribution cumulée $P(v)$ donne la probabilité d'apparition d'un vitesse de vent inférieure à v. Cette fonction est donnée par :

$$P(v) = 1 - e^{-\left(\frac{v}{A}\right)^k} \tag{2.2}$$

La moyenne μ de v, est :

$$\mu = A\,\Gamma(1 + \frac{1}{k}) \tag{2.3}$$

où $\Gamma(x)$ est la fonction gamma :

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$$

La variance σ^2 est :

$$\sigma^2 = \mathrm{E}\left[(v-\mu)^2\right] = A^2\left[\Gamma(1+\frac{2}{k}) - \Gamma^2(1+\frac{1}{k})\right] \tag{2.4}$$

Il est intéressant pour la caractérisation de la climatologie du vent d'aller jusqu'au $4^{\text{ème}}$ moment de la distribution. Le $3^{\text{ème}}$ moment $asym(v)$ caractérise l'asymétrie de la distribution. Il est nul pour une distribution gaussienne et positif si la moyenne est plus grande que la médiane de la distribution. Le $4^{\text{ème}}$ moment $apl(v)$ mesure l'aplatissement de la distribution par rapport à une distribution gaussienne. Une valeur positive de $apl(v)$ indique que la distribution est plus pointue qu'une gaussienne.

$$asym(v) = \frac{\text{moyenne}\left((v - \text{moyenne}\,(v))^3\right)}{\text{std}^3\,(v)} \tag{2.5}$$

$$apl(v) = \frac{\text{moyenne}\left((v - \text{moyenne}\,(v))^4\right)}{\text{std}^4\,(v)} - 3 \tag{2.6}$$

Une description des caractéristiques de la distribution des vitesses de vent sur les mers et océans à une échelle globale est donnée dans [6]. La variabilité saisonnière de cette distribution, à une échelle globale, est donnée dans [9]. Cette variabilité saisonnière impose l'utilisation de données couvrant une durée d'une année complète au moins (ou un multiple entier d'une année) pour l'estimation de la distribution des vitesses du vent au site étudié. Ceci permet d'éviter un biais pouvant être introduit par une des saisons.

La densité moyenne de puissance $\overline{E}$ disponible sur une période T, sur une site, est donnée par :

$$\overline{E} = \frac{1}{2}\,\frac{1}{\mathrm{T}}\int_0^{\mathrm{T}} \rho v^3(t) dt \tag{2.7}$$

où ρ est la densité de l'air (en kg m^{-3}). $\overline{E}$ (en W m^{-2}) correspond à la densité de puissance du flux par unité de surface orthogonale au flux et v la vitesse du vent (en m s^{-1}).

$\overline{E}$ peut s'écrire en fonction des paramètres A et k :

$$\overline{E} = \frac{1}{2}\rho A^3 \Gamma\left(1 + \frac{3}{k}\right) \tag{2.8}$$

Pour l'estimation des paramètres de Weibull à partir de la série de données de vitesse de vent, il est nécessaire de recourir à un estimateur statistique. La précision des estimations des paramètres A et k dépend directement des données disponibles. Plus précisément, elle dépend de la résolution temporelle de ces données et des erreurs contenues dans ces données.

2.1.2.1 Estimateurs des paramètres de Weibull

Plusieurs estimateurs des paramètres de Weibull existent. Ceci est dû à l'interdépendance des paramètres A et k de Weibull. En effet, ces paramètres s'écrivent, en fonction des échantillons de vitesse de vent v_i :

$$k = \left[\frac{\sum_{i=1}^{n} v_i^k \ln(v_i)}{\sum_{i=1}^{n} v_i^k} - \frac{1}{n}\sum_{i=1}^{n}\ln(v_i)\right]^{-1} \tag{2.9}$$

$$A = \left[\frac{1}{n}\sum_{i=1}^{n} v_i^k\right]^{\frac{1}{k}} \tag{2.10}$$

[10] propose l'inversion suivante pour l'estimation des paramètres A et k de Weibull :

$$k \approx \left(\frac{\overline{v}}{\sigma(v)}\right)^{1,086} \tag{2.11}$$

$$A = \frac{\overline{v}}{\Gamma\left(1 + \frac{1}{k}\right)} \tag{2.12}$$

où $\overline{v}$ est la moyenne des vitesses et $\sigma(v)$ l'écart type des échantillons de vitesse du vent. Cette relation est obtenue en utilisant les équations 2.3 et 2.4. Le rapport de ces équations donne :

$$\left(\frac{\sigma}{\mu}\right)^2 = \frac{\Gamma(1+\frac{2}{k})}{\Gamma^2(1+\frac{1}{k})} - 1 \tag{2.13}$$

En utilisant l'approximation suivante :

$$\frac{\Gamma(1+\frac{2}{k})}{\Gamma^2(1+\frac{1}{k})} - 1 \approx k^{\frac{-2}{1,086}} \tag{2.14}$$

on obtient l'estimateur de l'équation 2.11 [10] .

Un autre estimateur, utilisé couramment dans l'industrie éolienne, est l'estimateur de maximum de vraisemblance. Cet estimateur résout directement les équations 2.9 et 2.10. Ceci est effectué en estimant la probabilité d'apparition de la série d'échantillons $\{v_i\}_i$ pour chaque couple de valeurs de A et k. Les valeurs de A et k correspondant à la probabilité maximale sont alors choisis. La théorie liée à cet estimateur est présentée en Annexe B de ce document.

[6] affirme que ces deux estimateurs sont équivalents. [11] a comparé le maximum de vraisemblance à d'autres méthodes. Cette comparaison n'inclut toutefois pas la première méthode (équations 2.11 et 2.12). Il en ressort que la méthode de maximum de vraisemblance est la plus performante. L'avantage de cette méthode est qu'elle peut donner des intervalles de confiances basés sur les données. L'avantage de la première méthode est qu'elle permet en cas de connaissance de la moyenne et de l'écart type de l'erreur contenue dans les données de corriger la moyenne et l'écart type des données pour obtenir une meilleure estimation des paramètres A et k. Dans la suite, on choisit l'estimateur de maximum de vraisemblance, celui le plus couramment utilisé et présentant les meilleures performances dans la plupart des cas.

2.1.2.2 Influence de l'intégration temporelle

Les mesures de vitesse et de direction du vent sur site sont des mesures intégrées sur un certain intervalle temporel. Généralement, les vitesses considérées sont des moyennes sur un intervalle de 10 minutes. En écrivant la vitesse du vent v comme somme de la vitesse moyenne $\overline{v}$ et une fluctuation v' autour de cette moyenne, [8], $\overline{E}$ s'écrit comme ;

$$\overline{E} \approx \frac{1}{2}\rho\overline{v}^3\left(1+3i^2\right) \tag{2.15}$$

où i représente l'intensité de la turbulence.

$$i = \frac{\sigma_v}{\overline{v}} \text{ avec } \sigma_v = \overline{v'^2} \tag{2.16}$$

[8] donne une relation simple pour estimer le terme i pour un terrain à rugosité homogène et des conditions neutres :

$$i = \frac{1}{\ln(z/z_0)} \tag{2.17}$$

où z est la hauteur à laquelle est mesuré le vent et z_0 la rugosité du terrain. Cette expression donne une erreur de l'ordre de 3% à une hauteur de 10 m en mer. L'influence de la turbulence sur la densité moyenne de puissance est donc limitée, surtout en mer où la rugosité est moindre que sur terre.

2.1.2.3 Pas d'échantillonnage temporel

Le pas d'échantillonnage temporel est la durée séparant deux mesures successives de la vitesse et direction du vent. L'existence de cycles journaliers impose des contraintes sur le pas d'échantillonnage. Ainsi, un pas d'échantillonnage multiple de 24 heures ne pourra pas permettre une bonne représentation de la climatologie du site. [12] montre que, hormis ces valeurs particulières, le pas d'échantillonnage peut aller jusqu'à 50 heures si la précision sur la détermination des paramètres de Weibull est relâchée jusqu'à une précision d'une décimale. L'étude présentée dans [12] se base sur l'estimation des paramètres A et k à partir de série temporelles du vent acquises sur 4 sites côtiers du sud de la France (la Grande Motte, Leucate, Sète et Saintes-Maries-de-la-mer).

2.1.2.4 Nombre d'échantillons

Dans ce paragraphe, on étudie le nombre n d'échantillons nécessaires pour avoir une erreur sur les paramètres estimés au dessous d'un certain seuil.

Le premier paramètre d'intérêt est la moyenne du vent. Soit μ_n la moyenne calculée à partir de n échantillons v_i de la vitesse du vent. Dans l'hypothèse où les échantillons dont on dispose sont indépendants et identiquement distribués, le théorème de la limite centrale s'applique. Ce théorème nous indique que μ_n suit une distribution gaussienne, centrée sur la « vraie » moyenne. En considérant que μ est la « vraie » moyenne, on cherche à avoir une estimation μ_n de cette moyenne, calculée à partir de n échantillons, avec une erreur inférieure à un certain seuil $f\mu$. Ne connaissant pas la vraie moyenne, cette propriété ne peut être établie que de manière probabiliste. On cherche donc à vérifier que la probabilité que l'erreur soit inférieure à $f\mu$, soit supérieure à un certain seuil Φ. La condition $|\mu_n - \mu|$ est en dessous d'une certaine fraction f de μ écrite en termes probabilistes si :

$$P\left(|\mu_n - \mu| < f\mu\right) \geq \Phi \tag{2.18}$$

avec Φ le seuil de confiance exigé et $P\left(|\mu_n - \mu| < f\mu\right)$ la probabilité que l'erreur sur la moyenne estimée soit inférieure à $f\mu$.

La résolution de l'équation 2.18 donne [13] :

$$n \geq \frac{2}{f^2}\left(\mathrm{erf}^{-1}\left(\Phi\right)\right)^2 \frac{\sigma^2}{\mu^2} \tag{2.19}$$

avec erf la fonction d'erreur définie par $\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}dt$.

Le deuxième membre de l'équation 2.19 ne dépend que du paramètre de forme k. Sur la figure 2.1 on a représenté le nombre d'échantillons n nécessaire pour avoir une erreur de moins de 10 %, 5 % et 1 % avec une probabilité de 90 % en fonction de k pour $1 \leq k \leq 4$. On choisit cet intervalle pour le paramètre k car c'est l'intervalle contenant les valeurs généralement calculées pour une distribution de Weibull caractérisant les vitesses du vent.

[13] établit d'autres relations similaires pour la variance, le coefficient d'asymétrie, A et k. Les relations établies pour A et k ne sont valables que dans le cas de l'utilisation de la méthode des moments II pour l'estimation de ces paramètres [13]. La méthode des moments II, pour l'estimation des paramètres A et k consiste à résoudre les équations suivantes :

$$\frac{\Gamma(1+\frac{3}{k})}{\Gamma^3(1+\frac{1}{k})} = \frac{\overline{v^2}}{\overline{v}^2} \tag{2.20}$$

$$A = \frac{\overline{v}}{\Gamma(1+\frac{1}{k})} \tag{2.21}$$

Selon [13], 250 observations indépendantes sont nécessaires pour arriver à estimer les paramètres A et k avec une incertitude de $\pm$ 90 % et un taux de confiance de 90 %. Cependant, la condition d'échantillons indépendants et identiquement distribués est difficilement vérifiable. En outre, les échantillons doivent couvrir au moins une année. Un nombre d'échantillons de 250 images couvrant une année, correspond à un pas d'échantillonnage de 35 heures. Ce pas d'échantillonnage est bien dans la gamme préconisée par [14] dont l'étude a été présentée dans le paragraphe précédent.

2.1.2.5 Influence de l'erreur sur A et k sur l'estimation de la densité moyenne de puissance

Dans ce paragraphe, on étudie l'influence de l'erreur sur les paramètres de Weibull A et k sur l'estimation de la densité moyenne de puissance d'un site. Pour cela, on utilise l'équation 2.8. Soient A la «

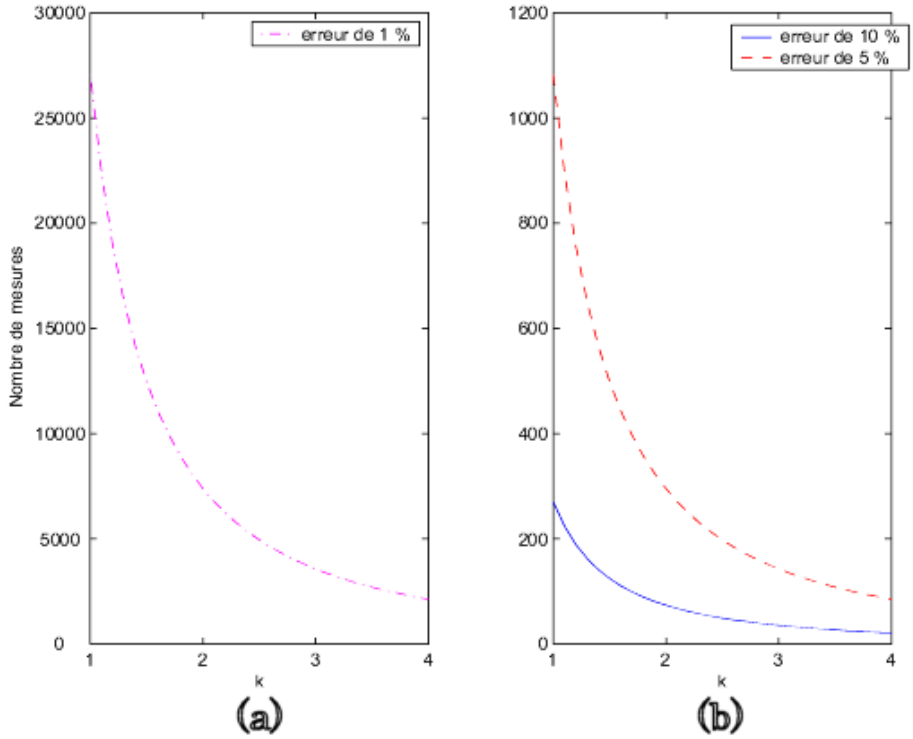


FIG. 2.1 – Nombre d'échantillons nécessaire pour avoir une erreur sur la moyenne de moins de 10 %, 5 % (b) et 1 % (a) avec une probabilité de 90 % en fonction du paramètre de forme k. Notez la différence d'échelle verticale entre la figure de gauche et celle de droite.

vraie » valeur du paramètre d'échelle et k la « vraie » valeur du paramètre de forme de la distribution du vent. Soient $(1+\alpha)A$ la valeur estimée pour le paramètre de forme et $(1+\beta)k$ la valeur estimée pour le paramètre d'échelle. α et β représentent l'erreur relative pour les paramètres estimés. L'erreur relative sur la densité moyenne de puissance est :

$$err_{\overline{E}} = (1+\alpha)^3 \frac{\Gamma\left(1+\frac{3}{(1+\beta)k}\right)}{\Gamma\left(1+\frac{3}{k}\right)} - 1 \tag{2.22}$$

La figure 2.2 représente le pourcentage d'erreur sur la densité moyenne de puissance $\overline{E}$ en fonction du pourcentage d'erreur sur la paramètre d'échelle A en considérant qu'il n'y a pas d'erreur commise sur le paramètre k. L'erreur sur $\overline{E}$ ne dépend pas de A mais uniquement de l'erreur sur le paramètre A. Il ressort de cette figure que le pourcentage d'erreur sur la densité moyenne de puissance croît rapidement en fonction de l'erreur sur A. Ceci est particulièrement sensible dans le cas où A est surestimé.

La figure 2.3 représente le pourcentage d'erreur sur la densité moyenne de puissance pour trois valeurs de pourcentage d'erreur sur le paramètre de forme k (5 %, 10 % et 20 %) en fonction de k, en considérant qu'il n'y a pas d'erreur commise sur le paramètre A. Cette figure représente cette erreur pour les valeurs usuelles de k (entre 1 et 4). Il apparaît sur cette figure que l'erreur sur le paramètre k affecte l'estimation de la densité moyenne de puissance surtout pour les faibles valeurs de k.

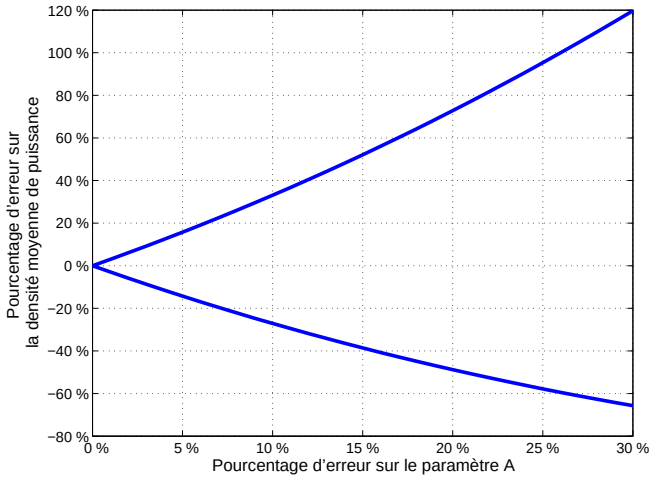


FIG. 2.2 – Influence de l'erreur sur le paramètre d'échelle A sur l'estimation de la densité moyenne de puissance.

2.1.3 Conclusions sur les caractéristiques des mesures de vent nécessaires pour la cartographie du potentiel éolien offshore

La résolution spatiale nécessaire pour rendre compte des phénomènes ayant l'échelle des parcs éoliens est de l'ordre du kilomètre. Le potentiel est donné de manière compacte sous la forme de paramètres statistiques décrivant la fonction de densité de probabilité des vitesses de vent. La précision de ces paramètres dépend de l'échantillonnage temporel des mesures. L'échantillonnage temporel peut aller jusqu'à

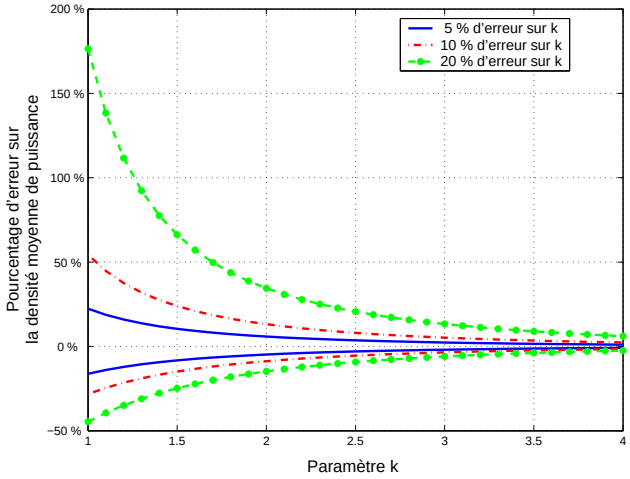


FIG. 2.3 – Influence de d'une erreur de 5 %, 10 % et 20 % sur le paramètre de forme k, en fonction de k, sur l'estimation de la densité moyenne de puissance.

50 heures [14]. Ces mesures doivent couvrir au moins une année complète pour éviter que la variabilité inter-saisonière n'influe sur l'estimation du potentiel éolien. Cependant, une durée d'un an ne permet pas de rendre compte de la variabilité inter-annuelle. La durée de vie d'un parc éolien est d'environ 20 ans. Sur une période de 20 ans, [15] relève que, sur un site en Irlande, la variabilité inter-annuelle atteint environ 14 % de la production moyenne annuelle sur cette durée. Pour garantir la qualité de l'estimation du potentiel éolien, il est donc nécessaire d'utiliser des mesures de vent couvrant plusieurs années. En considérant des vitesses de vent moyennées sur des fenêtres de 3 années, [15] indique que la variabilité descend 14 % à 4 %.

Dans cette section, on a étudié, en outre, l'influence de l'erreur sur les paramètres A et k sur l'estimation de la densité moyenne de puissance. Il a été montré que l'erreur sur le paramètre de forme A a une influence plus grande que celle du paramètre k sur la précision de l'estimation de la densité moyenne de puissance. Une erreur de 5 % sur le paramètre A induit, par exemple, 16 % d'erreur sur la densité moyenne de puissance estimée. Une erreur de 5 % sur le paramètre k, pour des valeurs de k voisinant 2, induit environ 5 % d'erreur sur l'estimation de la densité moyenne de puissance.

2.2 Mesures du vent en mer

Dans cette section on présente les différents instruments pouvant fournir des données sur le vent en mer. On étudie l'adéquation entre les caractéristiques de ces données et celles, requises pour la cartographie du potentiel éolien, étudiées dans la première section de ce chapitre.

2.2.1 Mesures sur site

2.2.1.1 Bateaux et bouées

Les mesures acquises par des instruments à bord de bateaux ont une faible couverture spatiale et une couverture irrégulière dans le temps. La mauvaise qualité de ces mesures est due à l'effet du mouvement du bateau sur les mesures enregistrées par l'anémomètre à bord du bateau. Ces mesures sont essentiellement disponibles sur les grandes routes commerciales. L'irrégularité dans le temps et dans l'espace de ces mesures limite leur apport pour la cartographie du potentiel éolien offshore.

Les mesures de bouées sont calibrées et ont une assez bonne qualité mais leur couverture spatiale est faible et non uniforme. La figure 2.4 représente la distribution des bouées recensées par le JCOMMOPS (Joint WMO-IOC Technical Commission for Oceanography and Marine Meteorology - *in situ* Observing Platform Support Centre). Cette carte donne la répartition spatiale, par pays, des bouées dérivantes et fixes envoyant des données au GTS (Global Telecommunication System) permettant de collecter les données météorologiques à l'échelle mondiale. On remarque l'irrégularité de la distribution de ces bouées et leur couverture spatiale limitée.

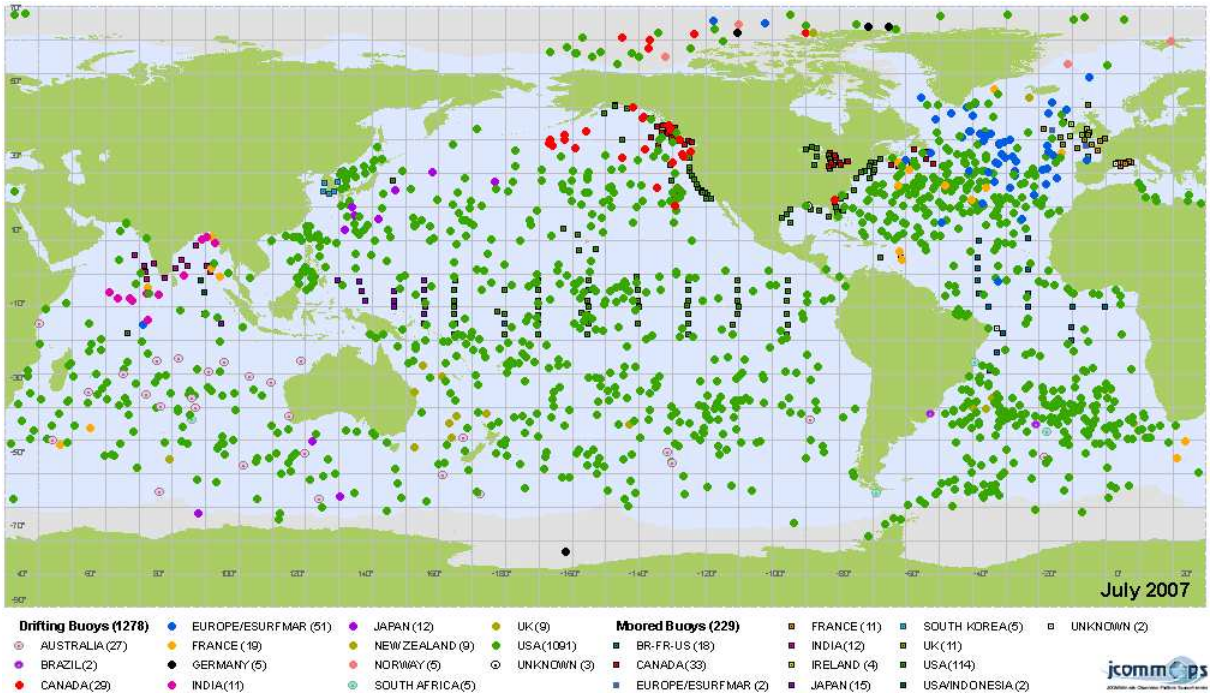


FIG. 2.4 – Distribution spatiale et répartition par pays des bouées dérivantes et fixes recensées par le JCOMMOPS Joint WMO-IOC Technical Commission for Oceanography and Marine Meteorology - *in situ* Observing Platform Support Centre) en juillet 2007. D'après [16].

2.2.1.2 Mât de mesures

C'est le moyen considéré comme le plus fiable, par l'industrie éolienne, pour la mesure du vent en un point. Il permet d'accéder à des vitesses et des directions de vent moyennées, typiquement, sur des

périodes de 10 minutes. De plus, il permet de rendre compte de la turbulence sur le site en donnant l'écart type des vitesses de vent pendant ces périodes de 10 minutes et la vitesse maximale observée [15]. Cependant, il présente un coût élevé en mer et une obligation d'attendre un an au moins. Cette durée d'un an est une durée minimale pour capturer les variations saisonnières du vent. Elle reste néanmoins insuffisante pour décrire la climatologie du vent à long terme au point où sont faites les mesures. Dans [15], la variabilité inter-annuelle est étudiée pour une station météorologique Irlandaise, celle de Malin Head. Il en résulte que l'écart type des moyennes annuelles de vent, pour une période de 20 ans, est de 5 %. La vitesse moyenne du vent sur les 20 ans est de 8,3 m s^{-1}. La moyenne annuelle la plus faible sur cette période est de 7,8 m s^{-1}. La moyenne la plus élevée est de 9,2 m s^{-1}. Ces variations se traduisent par une production d'énergie variant de -10 % jusqu'à +14 % de la production moyenne sur la période de 20 ans. Ce site est caractérisé par une moyenne de vent relativement élevée. La même variabilité, sur un site caractérisé par une moyenne de vitesse moins élevée, entraînerait une plus grande erreur sur l'estimation de l'énergie produite [15]. En considérant des vitesses de vent moyennées sur des fenêtres de 3 années, [15] indique que la variabilité descend de 10 % et 14 % à 3 % et 4 %.

Un autre inconvénient des mesures acquises par des instruments installés sur un mât est qu'elles ne sont représentatives que du point où elles ont été prises, surtout en zones côtières caractérisées par une forte variabilité spatiale du vent. Dans ce cadre, [15] préconise d'installer plusieurs mâts si l'étendue du parc est plus grande que 1 km.

2.2.2 Mesure du vent en mer par télédétection

Les instruments de mesures du vent par télédétection sont capables de mesurer le vent à distance. Cependant, certains de ces instruments (les SODAR et les LIDAR) doivent être déployés sur le site d'étude. Ces instruments sont classés dans les instruments de télédétection car ils permettent d'accéder à des mesures du vent en altitude et donc à distance de l'instrument.

2.2.2.1 SODAR

Un SODAR (SOund Detection And Ranging) est un instrument de télédétection installé sur terre (par opposition aux instruments spatioportés ou aéroportés) dont le but est la mesure des profils verticaux de vent. Il mesure les vitesses du vent à différentes hauteurs au-dessus de l'endroit où il est installé.

Un SODAR émet des ondes sonores et mesure l'onde rétrodiffusée. L'intervalle de temps entre l'émission et la réception de l'onde détermine la hauteur à laquelle la rétrodiffusion a eu lieu. Le décalage en fréquence entre l'onde émise et l'onde reçue sert à mesurer la vitesse et la direction du vent à la hauteur de rétrodiffusion grâce à l'effet Doppler.

La qualité des mesures d'un SODAR est fortement influencée par les facteurs atmosphériques autres que le vent. Ainsi, la pluie affecte fortement les mesures d'un SODAR. [17] indique que la disponibilité de mesures de bonnes qualité d'un SODAR dépend de la saison (du fait du changement des conditions climatiques) et de la hauteur. La qualité des mesures du SODAR décroît avec la hauteur. De plus, le SODAR est sensible au bruit de fond qui introduit une erreur dans l'onde sonore rétrodiffusée. Dans [17], la limite supérieure de mesure du SODAR est de 15 m s^{-1}. [18] utilise le ratio signal sur bruit

(RSB) pour filtrer les mesures SODAR. Pour un RSB de 9, la corrélation entre les mesures SODAR et les mesures du mât est de 0,78 à une hauteur de 5 m et de 0,87 à une hauteur de 10 m [18]. 25 % des mesures SODAR ont un RSB de 9 à une hauteur de 5 m et 49 % à une hauteur de 10 m. Le taux de disponibilité relevé dans [17], qui est du même ordre que celui relevé dans [18], est insuffisant pour établir une description statistique fiable de la climatologie du site. [17] préconise une utilisation du SODAR en combinaison d'un anémomètre à basse altitude. Le SODAR sert, dans ce cas, à établir des profils de vent utilisables pour extrapoler les vitesses recueillies par l'anémomètre à la hauteur de la nacelle.

L'utilisation des SODAR dans le cadre de l'estimation du potentiel éolien se limite donc à un estimation en un point, comme c'est le cas pour les mâts, et à une estimation du profil vertical du vent. La connaissance du profil vertical du vent est importante pour l'extrapolation des vitesses d'une hauteur à une autre. Toutefois, l'apport des SODAR pour la cartographie du potentiel éolien à une hauteur fixée reste limité.

2.2.2.2 LIDAR

Le principe d'un LIDAR (LIght Detection And Ranging) est semblable à celui d'un SODAR à la différence que l'onde émise est une onde lumineuse. La mesure se base sur l'effet Doppler introduit par les aérosols transportés par le vent sur l'onde rétrodiffusée.

[19] et [20] présentent une évaluation des mesures de vent recueillies par des LIDAR. Ces mesures sont perturbées par la pluie. Quand ces mesures sont exclues, [19] indique une corrélation entre les mesures LIDAR et les mesures d'un mât à proximité variant de 0,96 à 0,98 selon la hauteur. [20] étudie les mesures d'un LIDAR en offshore. Il se contente cependant d'une évaluation qualitative des résultats montrant que les profils de vent établis à partir des mesures recueillies par le LIDAR sont similaires à ceux établir en utilisant les anémomètres installés sur un mât de mesure à proximité. Il ressort de ces études que les LIDAR peuvent fournir une information de qualité sur le profil de vent [19, 20]. Les LIDAR présentent un taux de mesures « valides » plus grand que celui des SODAR. Cependant, il est nécessaire d'installer ces instruments sur site et les mesures ne sont représentatives que d'une zone limitée autour du point de mesure.

Une utilisation d'un LIDAR sur site ne permet donc pas de cartographier le potentiel éolien. Elle permet la connaissance du profil vertical. L'utilisation de LIDAR spatioportés ou aéroportés peut permettre la cartographie. Le lancement du premier LIDAR spatioporté est prévu pour 2007 [21]. Ce LIDAR sera à bord de Aeolus, une mission de l'ESA. Il permettra de fournir des profils de vent à une résolution spatiale de 200 km, les mesures étant intégrées sur 50 km. Ce LIDAR permettra de mesurer le profil de vent. Ce LIDAR a, cependant, une résolution verticale trop grossière pour le secteur éolien. La résolution spatiale horizontale est trop grossière pour la cartographie du potentiel éolien.

2.2.2.3 Radars à ouverture réelle

[22] montre la possibilité de retrouver des cartes de vecteurs de vent en mer en utilisant un radar nautique. Habituellement, ces radars sont utilisés pour détecter des objets à la surface de la mer. Au

large, ils opèrent à des fréquences de 3,0-3,1 GHz (Bande S). Ainsi, ils ont une grande couverture spatiale avec une résolution spatiale grossière. Près de la côte, ils opèrent à des fréquences de 9,3-9,5 GHz ce qui assure une plus grande résolution spatiale. La résolution spatiale est alors de l'ordre de 10 m à une distance de 750 m de l'antenne. Le radar couvre une zone d'un rayon d'environ 2 km.

Pour retrouver la direction du vent, les stries à la surface de la mer sont détectées dans les images radar. Ces stries sont alignées avec le vent. Elles ont des longueurs d'ondes plus grandes que celles des vagues de surface. En intégrant une séquence d'images radar dans le temps (32 images dans [22], ce qui équivaut à une minute environ), les signatures des vagues de surface sont éliminées. La moyenne de la séquence des images est sous échantillonnée en utilisant une pyramide gaussienne. Les stries ont des échelles spatiales caractéristiques de quelques centaines de mètres. Elles sont donc recherchées dans les plans correspondants. Pour retrouver les vitesses de vent, [22] utilise un réseau de neurones.

Les mesures d'un radar nautique installé sur une plate-forme pétrolière à 200 km au large de la Norvège ont été utilisées par [22] pour l'évaluation de leur méthode. La corrélation des directions avec celles retrouvées par un mât installé sur la plate-forme est de 0,99, le biais est de 0,6° et l'écart type est de 14,2°. Pour les vitesses de vent, si l'information sur la différence de température entre la mer et l'air est utilisée dans la phase d'apprentissage des réseaux de neurones, la corrélation est de 0,97, le biais est de 0,03 m s^{-1} et l'écart type est 0,85 m s^{-1}. Si la différence de température air-mer n'est pas utilisée la corrélation est de 0,96, le biais reste négligeable (0,01 m s^{-1}) et l'écart type est de 0,97 m s^{-1}. Cette méthode permettant de retrouver les champs de vecteurs de vent est appelée WiRAR. [23] a appliqué cette méthode à un des mesures acquises par un radar installé sur la plate-forme de recherche FINO-I située au large des côtes allemandes. La comparaison des directions retrouvées à partir des données radars et celles d'un mât situé sur la plate-forme résulte en une corrélation de 0,98 des directions à partir des deux sources, un biais de 0,3° et un écart type de 13°. Une corrélation de 0,96, un biais de 0,02 m s^{-1} et un écart type de 0,9 m s^{-1} sont trouvés pour les vitesses de vent.

Un inconvénient relevé dans [22] est l'absence de calibrage radiométrique ce qui implique une phase d'apprentissage des réseaux de neurones pour chaque installation.

Les caractéristiques des mesures acquises par ce genre d'instruments répondent aux exigences de la cartographie de la ressource éolienne offshore. Toutefois, la zone couverte par le radar (d'environ 4 km $\times$ 4 km) est limitée. Il reste aussi le problème de l'installation de ces instruments sur le site prospecté et l'attente pendant une durée minimale d'un an avant de pouvoir effectuer la cartographie.

2.2.2.4 Radiomètres radars

Ces instruments sont passifs. Les mesures sont donc grandement influencées par les nuages et la pluie. Ils ont en outre une résolution spatiale trop grossière pour la cartographie du potentiel éolien.

La possibilité de mesurer la vitesse du vent par les radars radiomètres a été démontrée par le SMMR (Scanning Multichannel Microwave Radiometer) à bord de Seasat [24]. Le SSM/I (Special Sensor Microwave/Imager) a permis d'accéder de manière opérationnelle à des mesures de vitesse de vent. L'analyse des mesures provenant de cet instrument ont montré qu'en plus de la vitesse du vent, il est possible de mesurer la direction du vent même si l'influence de la direction sur le signal est faible [25, 26]. L'ins-

trument WindSat a été lancé en 2003 pour évaluer la possibilité d'accéder de manière opérationnelle au vecteur vent.

Dans [25], 3321 vitesses de vent retrouvées grâce à SSM/I sont comparées à des mesures de bouées proches. Une erreur quadratique moyenne (RMSE) de 1,59 m s^{-1} et un biais nul sont trouvés. La corrélation entre les deux jeux de données est de 0,8. Avec une correction par la direction du vent, la RMSE est réduite à 1,3 m s^{-1} et la corrélation est de 0,85.

Les mesures recueilles par les radars radiomètres sont affectées par la pluie et les nuages. La comparaison des mesures d'un radiomètre à celles d'un instrument moins sensible à la pluie le prouve. [27] compare les vitesse de vent retrouvées grâce à WindSat et celles retrouvées grâce à un autre instrument de télédétection spatioporté. Cet instrument est QuikSCAT. C'est un radar diffusomètre. Ces instruments sont présentés dans la section 2.2.2.6 de ce chapitre. La comparaison est faite dans trois cas. Le cas 1 regroupe des paires de mesures WindSat-QuikScat situées à moins de 25 km l'une de l'autre et dont l'acquisition est faite à moins d'une heure d'intervalle. Dans le cas 2, sont éliminées les mesures où de la pluie est signalée. Dans le cas 3, ne sont gardées que les paires du cas 2 acquises à moins de 15 minutes d'intervalle. La différence moyenne est de 0,22 m s^{-1} dans le cas 1, 0,07 m s^{-1} dans le cas 2 et 0,09 m s^{-1} dans le cas 3. L'écart type de la différence est de 3,43 m s^{-1} dans le cas 1, 1,72 m s^{-1} dans le cas 2 et 1,71 m s^{-1} dans le cas 3.

Les différents radiomètres spatioportés fournissent des cartes de vitesses du vent avec une résolution spatiale de quelques dizaines de kilomètres. Cette résolution spatiale est trop grossière pour la cartographie du potentiel éolien offshore. En outre, l'influence de la pluie et des nuages sur les mesures recueillies par ces capteurs présente une grande limitation pour une utilisation systématique de ces mesures pour l'évaluation de la climatologie du vent sur une région donnée. Ces capteurs ne présentent que peu d'intérêt opérationnel pour la cartographie du potentiel éolien.

2.2.2.5 Radars Altimètres

Le premier altimètre spatioporté a été lancé en 1973 à bord de Skylab [28]. L'objectif premier des altimètres était de mesurer la hauteur de l'océan.

Les mesures altimétriques sont faites au nadir du satellite, c'est à dire à angle d'incidence nul [29]. Le mécanisme de réflexion prédominant à ces angles est présenté en annexe de ce document. La trace au sol des différents altimètres spatioportés est indiquée au tableau 2.2. Par exemple, ERS–2 couvre toute la surface planétaire en 35 jours sur une grille d'un pas de 7 km en latitude et 80 km en longitude [14].

[31] établit une relation entre le coefficient de rétrodiffusion et la vitesse du vent. Cependant, la faible couverture spatiale des altimètres, la faible répétitivité des mesures et l'absence d'information sur la direction constituent des limitations pour l'utilisation des données altimétriques pour l'établissement de statistiques de vent.

2.2.2.6 Radars diffusomètres

La mesure du vent par les diffusomètres se base sur la rétrodiffusion de Bragg présentée en annexe A de ce document. Les algorithmes, permettant de relier le coefficient de rétrodiffusion à la vitesse et

Satellite	Période de fonctionnement	Instrument	Trace (km)	Temps de revisite
GEOS–3	04/1975 - 12/1978	ALT	3,5	–
Seasat	06/1978 - 10/1978	ALT	1,7	–
Geosat	04/1985 - 09/1989	Radar Alt	1,7	17
ERS–1	07/1991 - 03/2000	RA–1	1,7	35
TOPEX/Poseidon	08/1992 - 01/2006	ALT,SSALT	2,2	10
ERS–2	04/1995 - ...	RA–1	1,7	35
Geosat follow–On	02/1998 - ...	RA	2	17
Jason–1	12/2001 - ...	Poseidon–2	2,2	10
ENVISAT	03/2002 - ...	RA–2	1,7	35

TAB. 2.2 – Altimètres disponibles. Selon [28, 30].

la direction du vent, dépendent de la bande de fréquence utilisée. [32] présente un algorithme conçu pour les instruments en bande Ku. En bande C, comme les instruments à bord de ERS–1/2, les algorithmes CMOD ont été développés. Les plus récents sont CMOD4 [33],CMOD-ifr [34] développé par l'IFREMER (Institut français de recherche pour l'exploitation de la mer) et le CMOD5 [35]. CMOD4 et CMOD-ifr se basent sur le même principe mais ont été calibrés à l'aide de données différentes. Tous ces modèles permettent d'estimer la vitesse et la direction du vent à une hauteur de 10 m au dessus de la surface de la mer. [33] indique que l'erreur atteint 1,6 m s^{-1} pour CMOD4. [36] indique que l'erreur peut atteindre 2 m s^{-1} et 20° en direction.

[14] indique que les diffusomètres sont des sources de données adaptés aux besoins de la cartographie du potentiel éolien offshore en matière de plage de vent mesuré et d'échantillonnage temporel. Actuellement, deux diffusomètres spatioportés sont en service : SeaWinds à bord de Quikscat et AMI à bord de ERS–2. Sur une période de 35 jours, on dispose au minimum de 8 mesures ERS–2 et 17 mesures Quikscat au niveau de l'équateur. Ceci correspond à un pas d'échantillonnage moyen de 33 heures. Ce pas d'échantillonnage est plus faible au latitudes moyennes. On dispose, ainsi, en moyenne de deux mesures Quikscat par jour au latitudes voisines de 50°. Au tableau 2.3 sont recensés les diffusomètres spatioportés lancés depuis Skylab. Toutefois, la basse résolution spatiale de ces capteurs ne permet pas de répondre à notre problématique. En outre, les mesures diffusométriques sont localisées loin des côtes. En effet, le signal près des côtes est contaminé par la partie réfléchie par la terre. La zone côtière étant la zone privilégiée pour l'installation de parcs éoliens, il est donc nécessaire d'extrapoler les mesures diffusométriques pour avoir une information sur le vent sur ces zones d'intérêt.

Les temps de revisite selon les différentes latitudes est donné par [40]. Les temps de revisite diminuent quand la latitude augmente. Au nord des latitudes 50 N, le temps de revisite de Quikscat est rarement supérieur à 12h. Pour la période 1996-1997, le tandem ERS et ADEOS est même caractérisé par un temps de revisite qui va rarement au dessus des 6h. Malheureusement, comme le montre le tableau 2.3 cet instrument n'est plus disponible depuis 1997. La mission qui lui a succédé (ADEOS-2) n'a été

Satellite	Période de fonctionnement	Instrument	Fréq. (GHz)	Pol.	Trace (km)	Résol. (km)	Temps de revisite
Seasat	06/1978 - 10/1978	SASS	14,6	VV,HH	2×500	50	–
ERS–1	07/1991 - 03/2000	AMI	5,3	VV	500	50	35
ERS–2	04/1995 - ...	AMI	5,3	VV	500	50	35
ADEOS	08/1996 - 06/1997	NSCAT	14,0	VV,HH	2×600	25/50	41
QuikSCAT	06/1999 - ...	SeaWinds	13,4	VV,HH	1800	50	4
ADEOS–2	12/2002 - 10/2003	SeaWinds	13,4	VV,HH	1800	50	4

TAB. 2.3 – Diffusomètres disponibles. Selon [28, 37, 38, 32, 39].

opérationnelle que quelques mois.

Les mesures de vent obtenues grâce à Quikscat ont été comparées à des mesures de bouées dans [41]. Les mesures de bouées ont été converties en mesures de vitesses à 10 m d'altitude en condition de stabilité neutre. Le biais est de 0,06 m s^{-1} et la RMSE est de 1,02 m s^{-1}. L'erreur sur la direction varie selon l'intervalle de vitesse du vent. Si tout le jeu de données est considéré, la RMSE est de 28,5° et le biais de 6,2°. La RMSE descend à 21,5° pour les vitesses supérieures à 3 m s^{-1} et à 17,3° pour les vitesses supérieures à 5 m s^{-1}.

Les mesures diffusométriques ont une résolution temporelle qui permet d'établir une cartographie fiable du potentiel éolien offshore. En effet, les caractéristiques de ces mesures répondent aux exigences en termes de résolution temporelle établis dans la première partie de ce chapitre. Cependant, la résolution spatiale de ces instruments ne permet pas de réaliser une cartographie à haute résolution spatiale (résolution de l'ordre du kilomètre).

2.2.2.7 Radars à ouverture synthétique (ROS)

Le tableau 2.4 liste les différents ROS spatioportés et quelques unes de leurs caractéristiques. La résolution spatiale des images ROS peut atteindre une dizaine de mètres. Pour le ROS à bord de TerraSAR-X, lancé en 2007, la résolution spatiale atteint 1 m. Cette haute résolution spatiale est due au mécanisme imageur des ROS. Un ROS émet une onde vers la surface, mesure l'écho rétrodiffusé et effectue une analyse de la phase du signal. C'est cette analyse en phase qui permet d'obtenir la haute résolution spatiale.

[45] donne les différentes techniques utilisées pour la mesure du vent à partir des images ROS. Certaines de ces techniques se basent sur les modèles utilisés pour les diffusomètres. En effet, les angles d'incidence des ROS et des diffusomètres étant situés dans le même intervalle (au dessus de 20°), c'est le même mécanisme de rétrodiffusion qui prédomine. Le modèle CMOD, avec ses différentes variantes, a été appliqué aux images ROS [45]. Les vitesses de vent estimées grâce à ces modèles sont des vitesses à une hauteur de 10 m du niveau de la mer. Une limitation de l'application de ce modèle est l'impossibilité de l'extraction combinée de la direction et la vitesse du vent à partir des images ROS.

L'estimation préalable de la direction à haute résolution spatiale reste difficile. [14] a développé

Satellite	Pèriode de fonctionnement	Instrument	Fréq. (GHz)	Pol.	Trace (km)	Résol. (m)	Temps de revisite
Seasat	06/1978 - 10/1978	SAR	1,28	HH	100	25	–
ERS–1	07/1991 - 03/2000	AMI	5,30	VV	100	30	35
JERS–1	02/1992 - 11/1998	SAR	1,27	HH	75	18	44
ERS–2	04/1995 - ...	AMI	5,30	VV	100	30	35
RADARSAT	11/1995 - ...	AMI	5,30	HH	100–170	10	24
ENVISAT	03/2002 - ...	ASAR	5,33	VV,HH, VV/HH, HV/HH, VH/VV	100	30	35
TerraSAR-X	06/2007 - ...	TSX-SAR	9,65	HH,VV HV,VH	10	1	11

TAB. 2.4 – Radars à Ouverture Synthétique disponibles. Selon [28, 37, 42, 43, 44].

une méthode se basant sur l'analyse multi-résolution pour l'extraction combinée de la direction et de la vitesse du vent à partir des images ROS. Cette méthode se base sur l'extraction de structures pouvant être représentatives de la direction vent à l'aide d'une transformée de Fourier. Ces directions sont, ensuite, utilisées en entrée du modèle CMOD pour le calcul des vitesses. Cette méthode ne peut, cependant, pas être utilisée dans le cas d'une forte composante azimutale du vent [14]. [46] utilise des données diffusométriques pour déterminer la direction du vent avant l'application du modèle CMOD. D'autres méthodes, liées à des phénomènes spécifiques observés dans les images ROS (stries par exemple), sont étudiées pour lier les mesures ROS et le vent. Le phénomène de coupure azimutale (*cutoff azimutal*), se traduit par une coupure du spectre de l'image ROS en azimut. [47] a montré que la valeur de la longueur d'onde de coupure peut être reliée à la vitesse du vent. Cependant, cette méthode ne permet pas de déterminer la direction du vent. En outre, cette méthode dépend de la distance en mer sur laquelle a soufflé le vent sans rencontrer d'obstacles (appelée *fetch*), et surestime les vents à faible vitesse. Face aux difficultés que pose l'estimation de la direction du vent à partir des images ROS, la direction de vent estimée par les modèles numériques ou fournie par les données de réanalyse peut être utilisée comme entrée à l'algorithme CMOD pour retrouver la vitesse du vent à partir des images ROS [48].

Les sources d'erreurs affectant les vitesses du vent retrouvées grâce aux images ROS sont :

- la précision radiométrique : pour ERS-1 et ERS-2, elle est de l'ordre de $\pm 0,4$ dB [49], ce qui se traduit par, approximativement, 0,7 m s^{-1} d'erreur sur la vitesse.
- le speckle : c'est un bruit multiplicatif présent dans les images ROS. Dans le cas de ScanSAR à bord de RADARSAT, ce bruit peut engendrer une erreur de ± 3 dB [50]. Une méthode pour réduire l'effet ce bruit est de moyenner le coefficient de rétrodiffusion sur plusieurs pixels. Ceci a pour effet de réduire la résolution spatiale. Selon [50], pour les images RADARSAT une moyenne sur 2 km $\times$ 2 km au moins doit être considérée au lieu de la résolution initiale de 100 m. Un

rapport équivalent entre les images d'origine et les images moyennées donne des moyennes sur 250 m $\times$ 250 m pour les images ERS. Pour ces images, [51] a montré que l'effet du speckle est surtout présent dans les résolutions entre 100 et 400 m. Dans la suite, on considérera des images moyennées sur 400 m. Cette résolution spatiale est meilleure que la résolution de l'ordre du kilomètre exigée par l'industrie éolienne. Elle permet aussi de diminuer significativement l'effet du speckle sur les vitesses de vent retrouvées à partir d'une image ROS.

– le bruit atmosphérique et océanique : il peut être causé par la pluie, la présence de surfactants à la surface de la mer. De plus, la bathymétrie a une influence sur les mesures ROS.
– la direction du vent utilisée en entrée : [50] étudie l'influence de l'erreur sur la direction du vent sur l'estimation de la vitesse du vent en fonction de : la vitesse du vent, la direction du vent et l'angle d'incidence. Les valeurs les plus élevées d'erreur sur la vitesse estimée surviennent pour les directions de vent à 45° de la trace du satellite. L'erreur est supérieure à 20 % pour les vitesses de vent estimées dans ce cas et peuvent atteindre plus de 30 % pour les vitesses de vent voisines de 5 m s^{-1} [50]. Si la direction du vent considérée se situe dans un intervalle de 15° autour de la trace du satellite, ou dans un intervalle de 15° autour de l'orthogonale à la trace du satellite, l'erreur sur la vitesse estimée ne dépasse pas les 10 %. Les caractéristiques de cette erreur ne varient que légèrement en fonction de l'angle d'incidence [50]. [52] a montré que, sur 56 images SAR ERS, une erreur de $\pm$ 30° dans la direction considérée pour retrouver les vitesses de vent résulte en une erreur quadratique moyenne de 1,9 m s^{-1} par rapport aux vitesses mesurées sur site.

Les vitesses de vent retrouvées grâce aux images ROS répondent aux besoins de la cartographie du potentiel éolien en termes de résolution spatiale. Cependant, la résolution temporelle des images ROS n'est pas suffisante pour établir une description fiable des variations temporelles du vent en un point donné. Dans le futur, avec le lancement de nouveaux satellites équipés de ROS, la résolution temporelle des mesures ROS pourrait devenir suffisante pour répondre aux besoins de la cartographie du potentiel éolien. Actuellement, les champs de vent retrouvés à partir des images ROS ne peuvent pas, à eux seuls, permettre d'établir une cartographie fiable du potentiel éolien.

2.2.3 Conclusions sur l'adéquation des moyens de mesure du vent en mer et la cartographie du potentiel éolien

Dans cette section, nous avons présenté les différents moyens de mesure du vent en mer. Leurs avantages et inconvénients pour la cartographie du potentiel éolien ont été décrits. Cependant, aucune des sources recensées ne permet à elle seule de cartographier le potentiel éolien offshore. Les mesures effectuées par les instruments à bord de bateaux ont une mauvaise qualité et leur couverture spatiale est limitée. Le problème de la couverture spatiale limitée et non uniforme se pose aussi pour les mesures enregistrées par des instruments fixés à des bouées. L'installation d'un mât, outre son coût élevé, ne permet pas d'accéder directement à une climatologie du vent de long-terme. Les diffusomètres mesurent la vitesse et la direction du vent avec une résolution temporelle suffisante pour établir une description statistique fiable du vent. En outre, une archive d'une dizaine d'année de mesures de diffusomètres spatioportés est disponible. Cependant, la résolution spatiale de ces instruments ne répond pas aux exigences

de l'industrie éolienne et leur utilisation en zone côtière est impossible. La résolution spatiale des ROS permet de répondre à ces exigences. Mais, ceci se fait aux dépens de la résolution temporelle. La résolution temporelle des mesures ROS est insuffisante pour cartographier d'une manière fiable le potentiel éolien. Il est donc nécessaire de recourir à des modèles pour l'estimation du potentiel éolien ou à une combinaison de ces moyens de mesure. Ces modèles peuvent être physiques ou statistiques. Les vitesses et direction du vent étant données à une certaine hauteur, différente de la hauteur de l'éolienne, il est aussi nécessaire d'utiliser un modèle d'extrapolation de ces mesures à la hauteur voulue.

2.3 Modèles pour l'estimation du potentiel éolien

2.3.1 Extrapolation verticale des vitesses de vent

Les valeurs de direction et de vitesse du vent, issues de mesures ou de modèles, sont données à une certaine hauteur (10 m pour les mesures par les instruments spatioportés par exemple). Pour extrapoler ces mesures à la hauteur d'intérêt pour l'estimation de la production d'une éolienne (hauteur de la nacelle), un modèle de profil est utilisé [8].

Les deux méthodes couramment utilisées dans la littérature sont :

– la considération d'un profil logarithmique [8] :

$$v(h) = v(h_m) \left[\frac{\ln\left(\frac{h}{z_0}\right)}{\ln\left(\frac{h_m}{z_0}\right)} \right] \tag{2.23}$$

où h est la hauteur à laquelle le vent est estimé, h_m est la hauteur des mesures et z_0 est la rugosité du terrain.

– la considération d'un profil en loi de puissance [53] :

$$v(h) = v(h_m) \left(\frac{h}{h_m} \right)^n \tag{2.24}$$

où h_m est la hauteur des mesures et n un facteur empirique exprimant les influences de la rugosité de la surface et la stabilité atmosphérique. n augmente avec l'augmentation de la rugosité et de la stabilité thermique de la couche de surface.

L'utilisation d'un profil en loi de puissance est rendue difficile à cause du choix de l'exposant n. Cette méthode est pratique quand la valeur de la rugosité z_0 (présente dans le profil logarithmique, équation 2.23) n'est pas connue. [3] a utilisé un profil en loi de puissance pour extrapoler directement le paramètre de forme A de Weibull. L'équation reliant le paramètre A à une hauteur h et le paramètre A_{h_0} calculé à une hauteur h_0 est selon [3] :

$$A(h) = A(h_0) \left(\frac{h}{h_0} \right)^n \tag{2.25}$$

Sur quelques sites, [3] a calculé une valeur moyenne de 0,23 pour n avec un écart type de 0,03.

[10] a proposé en utilisant une approche empirique la formule donnée dans l'équation 2.26 pour extrapoler le paramètre k de Weibull d'une hauteur à une autre. Cette relation est valable selon lui jusqu'à

une hauteur de 100 m au dessus du sol.

$$k(h) = k(h_m)\frac{1 - c\ln\left(\frac{h_m}{h_{ref}}\right)}{1 - c\ln\left(\frac{h}{h_{ref}}\right)} \tag{2.26}$$

où $k(h_m)$ est le paramètre de forme mesuré à la hauteur h_m, $h_{ref} = 10$ m et $c = 0,88$. [54] signale que l'utilisation de cette formule peut entraîner une sous estimation de k. [55] a proposé de modifier les paramètre c et h_{ref} pour les mettre à $c = 0,19$ et $h_{ref} = 18$

[56] a proposé l'expression suivante pour l'extrapolation directe du paramètre de forme k :

$$k(h) - k(h_m) = c_2(h - h_m)\exp\left[-\frac{h - h_m}{h_{max-}h_m}\right] \tag{2.27}$$

où h_{max} est la hauteur qui correspond au maximum du profil de k et c_2 un facteur d'échelle de l'ordre de 0,022.

En utilisant les données du mât FINO 1, installé en Mer du Nord, [57] note un comportement non-logarithmique. Sept anémomètres sont installés sur le mât FINO 1, de 30 m jusqu'à 90 m. La déviation du profil logarithmique apparaît pour les hauteurs supérieures à 45 m. Pour mieux modéliser le profil vertical du vent, [57] propose un modèle appelé ICWP (Inertially Coupled Wind Profile).

Dans la suite de ce document, on se contentera d'estimer les paramètres A et k de Weibull à une hauteur donnée. Le moyen le plus fiable pour établir des profils verticaux de vent et pour extrapoler les vitesses de vent est l'acquisition de mesures par un SODAR ou un LIDAR sur le site.

2.3.2 Modèles empiriques

Un modèle empirique permet à partir des observations de trouver une relation indépendante d'un site particulier entre les vitesses du vent à la côte, où une station météorologique est présente, et le vent en mer. [58] a comparé les mesures d'une station météorologique à Gorleston, sur la côte est anglaise, et des mesures faites en mer. Il en résulte que le ratio entre les vitesses en mer et sur terre, pour des vitesses modérées, varie entre 1,06 et 1,9 en fonction de la différence de température entre la mer et l'air et de la vitesse du vent. [59] après une campagne expérimentale a établi que la vitesse du vent pour une bouée fixée à 5 km des côtes est plus élevée de 10-15 % que la vitesse sur la côte pour des vitesses de vent inférieures à 10 m s^{-1}. Pour les vitesses supérieures à 10 m s^{-1}, très peu de différences entre le vent en mer et à la côte ont été observées. [60] a proposé une relation empirique à partir des observations entre les vitesses du vent en mer et à la côte. Cette relation suggère une augmentation de 5 % de la vitesse du vent, à une hauteur de 100 m, à 7,5 km des côtes et une augmentation de 12 % à 20 km des côtes.

[61] en utilisant des données sur différentes régions du monde a effectué une régression linéaire entre les vitesses de vent à la côte et au large. La relation établie est la suivante :

$$v_{mer} = 1,17\ v_{terre} + 1,62 \tag{2.28}$$

où v est la vitesse du vent en m s^{-1}.

Cependant toutes ces relations restent des estimations qualitatives confirmant des vitesses moyennes plus grandes en mer que sur terre. Un relation universelle entre la vitesse du vent en mer et à terre, ne

tenant pas compte des différentes conditions physiques et de la distance entre la côte et le point où les mesures doivent être extrapolées, peut induire de fortes erreurs. En effet, ce genre de relations impliquent des vitesses de vent constante sur toute la mer, sans prise en compte de la distance, en mer, sur laquelle a soufflé le vent sans rencontrer d'obstacles avant le point d'intérêt.

2.3.3 Modèles statistiques

2.3.3.1 MCP (Measure-Correlate-Predict)

Cette méthode est utilisée actuellement comme standard dans l'industrie éolienne. Elle permet d'établir la climatologie d'un site prospecté pour l'installation d'une ferme éolienne. La méthode MCP (Mesurer-Correler-Prédire) modélise la relation entre les mesures recueillies sur site, sur une durée limitée, et celles recueillies par une station météorologique proche sur la même durée. Cette modélisation est ensuite utilisée pour extrapoler les mesures de longue durée de la station météorologique. La définition d'une relation entre les mesures de vent sur deux sites peut être rendue compliquée par les effets de terrain sur le flux, les différentes situations météorologiques à grande et à petite échelle et les conditions de stabilité atmosphérique [62].

Plusieurs variantes de méthodes MCP existent. [63] propose une relation linéaire entre les vitesses sur le site cible et la station météorologique. Pour prendre en compte les directions du vent [64] a proposé une régression linéaire séparant les deux composantes du vent : composante méridienne et composante zonale.

La méthode MCP standard a été utilisée pour estimer les vitesses de vent sur deux sites offshore au large du Danemark : Rodsand et Vindeby [65]. Sur chacun de ces sites, un mât est installé. Dans les deux cas, la méthode MCP sous-estime la vitesse moyenne du vent, de 2,5 % à Vindeby [66] et de 3,4 % à Rodsand [67]. Cette méthode génère une série temporelle sur le site cible. L'erreur moyenne quadratique (RMSE) est de 1 m s^{-1} sur la série générée à Vindeby et de 1,2 m s^{-1} sur la série générée à Rodsand.

2.3.3.2 MLR (Multi-Linear Regression)

Cette méthode se rapproche de la méthode MCP. Elle s'en différencie par la possibilité de prise en compte d'autres prédicteurs que les vitesses de vent à la station météorologique. Elle peut avoir en entrée la différence de température entre l'air et la mer, par exemple.

L'application dans [65] de cette méthode pour évaluer les vitesses et direction de vent en un point en mer à partir de données sur terre montre que le paramètre le plus important est la vitesse du vent. La différence de température entre la mer et l'atmosphère, c'est à dire la stabilité, semble avoir moins d'importance. [65] explique cela par le fait que les paramètres importants pour les phénomènes météorologiques et donc pour la simulation numérique ne le sont pas toujours pour les algorithmes statistiques. Ceci peut être aussi dû au fait qu'une partie de l'information météorologique est contenue implicitement dans la méthode lors de la phase d'apprentissage.

[65] applique la méthode MLR sur deux sites danois : Rodsand et Vindeby. Sur ces deux sites l'erreur quadratique moyenne (RMSE) trouvée varie de 0,7 m s^{-1} à 1 m s^{-1}. Ces résultats sont meilleurs que

ceux issus de l'application de la méthode MCP sur les mêmes sites.

2.3.3.3 Classification And Regression Trees (CART)

La méthode CART [65] est une régression non-linéaire, basée sur la construction d'un arbre de décision. On part de l'ensemble initial de données d'apprentissage. Cet ensemble est ensuite à chaque pas de l'algorithme divisé en deux selon que chacun des paramètres est inférieur ou supérieur à un certain seuil. Cet algorithme impose donc que les données soient discrétisées, c'est à dire de diviser les vitesses et les directions de vent en intervalles. Ensuite pour transposer les mesures l'arbre généré est utilisé.

Pour illustrer cette méthode, un arbre de décision issu de CART peut conduire à dire que si la vitesse du vent est dans l'intervalle [4 m s^{-1},4,5 m s^{-1}] sur le site de mesure sur terre, elle sera dans l'intervalle [6 m s^{-1},7 m s^{-1}] en mer sur le site où la vitesse doit être estimée. De tels intervalles sont raffinés de manière itérative. La méthode CART essaye de minimiser la variance associée aux valeurs prédites par l'arbre de décision.

[65] applique la méthode CART sur deux sites danois : Rodsand et Vindeby. L'erreur quadratique moyenne (RMSE) à une hauteur de 38 m est de 0,6 m s^{-1} à Vindeby et 1,1 m s^{-1} à Rodsand. Ces résultats sont meilleurs que ceux des méthodes MCP et MLR. Cependant, les 3 méthodes ont en commun l'inconvénient de devoir acquérir des mesures sur site d'une durée minimale de un an. Cette durée est nécessaire pour ne pas introduire de biais dû à une des saisons. En outre, ces 3 méthodes ne permettent pas une cartographie du potentiel éolien. Elles permettent une estimation du potentiel éolien en un point, où des mesures de vent ont été acquises.

La méthode CART introduit la notion de classes de vent et un lien entre ces classes et la relation du vent en deux points. L'idée d'une relation entre les variations du vent et des classes de vent sera utilisée dans la méthode proposée dans le chapitre 3.

2.3.3.4 Méthode probabiliste

[14] a proposé une méthode probabiliste permettant de combiner les mesures de diffusomètres et celles de radars à ouverture synthétique. La climatologie du vent à basse résolution est calculée à partir des données de diffusomètres. Les mesures diffusométriques ne sont pas disponibles près des côtes à cause de leur contamination par la terre. Les statistiques à basse résolution spatiale, établies à partir des mesures diffusométriques, sont, donc, disponibles loin de la côte.

En un point, en utilisant les données diffusométriques disponibles au large, la direction du vent est classée en différents secteurs $\{\theta_i\}$. Une probabilité $p(\theta_i)$ d'occurrence du vent est associée à chaque secteur. Pour le transfert de ces probabilités prés la côte des probabilités de changement de direction sont utilisées. Ces probabilités $p(\theta_j/\theta_i)$ représentent les probabilités d'occurrence d'un vent de secteur θ_i au large et de secteur θ_j prés de la côte. A ces probabilités, sont liées des fonctions associant les vitesses de vent au large et à la côte. Pour obtenir la haute résolution spatiale, [14] a proposé de classer les situations de vent selon le comportement de la partie turbulente du vent.

Cette méthode présente une rigueur mathématique mais nécessite un nombre trop important de données pour la définition des probabilités jointes. Son application dans la pratique reste donc impossible.

2.3.4 Modèles physiques

Le but de cette section n'est pas d'établir un état de l'art complet de tous les modèles physiques existants mais de fournir les différents types de modèles et quelques exemples des modèles utilisés dans la communauté scientifique. Les modèles physiques peuvent être divisés en deux type principaux [68] : les modèles prognostiques et les modèles diagnostiques.

2.3.4.1 Modèles prognostiques

Ces modèles sont basés sur la résolution des équations temporelles hydrodynamiques ou thermodynamiques adaptées à l'atmosphère. On peut distinguer dans ces modèles, les modèles hydrostatiques et les modèles non-hydrostatiques. Dans les modèles hydrostatiques, l'équation du moment vertical est remplacée par l'équilibre hydrostatique. Ceci implique que les échelles du mouvement vertical du vent doivent être inférieures aux échelles horizontales considérées. Ces modèles ne peuvent donc être utilisés que pour modéliser le vent à une échelle régionale (10 km environ).

Les modèles non-hydrostatiques permettent de modéliser le vent à une échelle plus fine. Pour le fonctionnement des modèles prognostiques, il est nécessaire de définir les conditions initiales et les conditions aux limites. La performance des modèles est liée à la bonne initialisation des variables météorologiques et la bonne définition des conditions aux limites. Ces modèles ont besoin de :

- champs météorologiques 3-D sur tout le domaine de simulation pour l'initialisation des modèles.
- champs météorologiques 2-D (ou 3-D) définissant les conditions aux limites et latérales et au dessus du domaine de simulation. Les conditions aux limites doivent être définies pour toute la durée de la simulation.
- conditions aux limites en bas du domaine de simulation (orographie, rugosité, température de la surface, ...)

Ces champs, déterminant les conditions aux limites et les conditions initiales, ne sont généralement pas disponibles avec les résolutions spatiales et temporelles suffisantes pour un fonctionnement idéal des modèles. Ces modèles ne sont pas, pour l'instant, utilisables pour une estimation systématique du vent sur une région donnée. Leur utilisation pour la cartographie de la ressource éolienne offshore est donc inadaptée dans l'état actuel des connaissances.

Une comparaison de résultat d'un modèle non-hydrostatique (MC2), de vitesses retrouvées à partir d'images ROS et de mesures diffusométriques a été entreprise dans [69]. Les images ROS utilisée dans [69] ont été acquises par RADARSAT-1 et couvrent une partie du Québec. Les mesures diffusométriques sont des mesures acquises par Quikscat. MC2 est un modèle développé par le groupe de recherche en prévision numérique à "Environnement Canada" [70]. Il ressort de cette comparaison que l'erreur moyenne quadratique (RMSE) entre les mesures ROS et les mesures diffusométriques est plus faibles que la RMSE entre les sorties du modèle et les mesures diffusométriques. La RMSE entre les mesures ROS et les mesures diffusométriques est de 1,46 m s^{-1}. La RMSE entre les sorties du modèle et les mesures diffusométriques est de 2,27 m s^{-1}.

Le bon fonctionnement de ces modèles est limité par la disponibilité des conditions aux limites et par la nécessité d'une bonne initialisation de la simulation. Une utilisation systématique de ces modèles pour

évaluer le potentiel éolien est pour l'instant impossible. Ces modèles sont pour l'instant utilisés dans le cadre d'autres méthodes permettant de réduire le nombre de simulations nécessaires pour l'estimation du potentiel éolien. De telles méthodes sont présentées à la section 2.3.4.4.

Dans la catégorie des modèles prognostiques, on retrouve aussi les modèles suivants : MIUU [71], GESIMA [72], MesoMap [73], ETA [74], HIRLAM [75], Meso-NH [76], LM (DWD) [77], RAMS [78], MM5 [79], TVM [80].

2.3.4.2 Modèles diagnostiques

Ces modèles utilisent des mesures *in situ* sur la région où le vent doit être modélisé ou des données de vent géostrophique. La solution d'un modèle diagnostique essaye de se rapprocher au maximum de ces mesures, aux points où elles sont disponibles, tout en respectant une cohérence physique du champ de vent. Les modèles physiques diagnostiques sont constitués principalement par deux types d'approche :

- une simplification des solutions des équations de Navier-Stokes en état d'équilibre (ce qui exclut la composante temporelle), par exemple les modèles d'écoulement linéarisé.
- une analyse des observations météorologiques en appliquant des contraintes physiques. La contrainte choisie peut être l'équation de continuité, exprimant la conservation de masse. Ces modèles sont alors dits consistants en masse.

2.3.4.2.1 Exemples de modèles diagnostiques à équations linéarisées

2.3.4.2.1.1 WAsP

WAsP (Wind Atlas Analysis and Application Program) [81] est actuellement l'outil d'estimation du potentiel éolien le plus répandu dans l'industrie éolienne. WAsP prend en entrée une série temporelle de données de vitesses et de directions de vent recueillies à un site donné. Il élimine l'influence des conditions locales (de rugosité, de relief et d'obstacles) sur cette série. Ainsi, la série temporelle du vent géostrophique au dessus du point de mesure est construite. Les statistiques de cette série sont calculées. Ces statistiques sont considérées constantes entre le site de mesure et le site cible. Les statistiques du vent local au site cible sont obtenues en considérant l'influence des propriétés de ce site sur le vent géostrophique.

Le bon fonctionnement de WAsP nécessite que le site de mesure et le site cible soient soumis à des régimes de vent similaires et aient des conditions atmosphériques proches d'un état neutre.

En offshore, l'application de WAsP est limitée par la considération d'une rugosité constante et par l'influence du relief côtier. En effet, ce relief peut avoir une influence jusqu'à 50 km offshore.

2.3.4.2.1.2 LINCOM

LINCOM (LINearized COMputation) [82] est un modèle d'écoulement linéarisé qui résout un ensemble d'équations linéaires dans le domaine spectral (espace de Fourrier). Ces équations sont dérivées

des équations non-linéaires de la masse et du moment ainsi que l'équation de continuité, pour les fluides incompressibles. La résolution de ces équations se fait en deux étapes. La première étape consiste à résoudre les équations linéaires décrivant les perturbations en vitesse et en pression pour un terrain plat et de rugosité uniforme. La deuxième étape sert à déterminer les perturbations engendrées par les changements de relief et de rugosité. La rugosité en mer dépend de la vitesse du vent, de la distance à la côte et de la bathymétrie. En mer, LINCOM est utilisé en combinaison avec un modèle de rugosité de surface avec une procédure itérative pour parvenir à une solution stable vitesse du vent-rugosité. LINCOM ne peut pas s'ajuster à plus d'un couple vitesse-direction. Une utilisation avec plusieurs points est possible mais seulement en donnant un poids à chaque point et en ajustant la somme pondérée de ces points. Il utilise des sorties de modèles grande échelle ou des mesures locales.

2.3.4.2.2 Exemple de modèles diagnostiques à conservation de masse

Les modèles consistants en masse ne contiennent pas d'équations dépendant du temps (équations dynamiques). Ils se contentent d'imposer la non divergence du flux. Ils ont donc besoin de beaucoup d'observations pour garantir la bonne modélisation du flux. A partir des observations, les modèles diagnostiques produisent un champ de vent en état d'équilibre. L'interpolation des données d'entrée joue un rôle important dans la détermination du champ de vent final.

MINERVE [83] est un code méso-échelle à conservation de masse. Il est constitué par deux étapes principales. La première étape est une interpolation des données d'entrée sur la grille du domaine. La deuxième étape est un ajustement du champ interpolé à l'équation de continuité. Les modèles à conservation de masse sont très sensibles aux mesures météorologiques utilisées en entrée.

2.3.4.3 Données de réanalyse

Les données de réanalyse sont obtenue après une réanalyse *a posteriori* des prédictions météorologiques fournies par les modèles pour les ajuster aux observations. Des grilles de réanalyse sont produites, par exemple, par le NCEP/NCAR (National Centers for Environmental Prediction/National Center for Atmospheric Research) [84], l'ECMWF (European Centre for Medium Range Weather Forecasting) [85] et la NASA [86]. Pour le calcul de ces grilles, différentes sources de données sont utilisées. Ces sources incluent les stations météorologiques, des observations de radiosondage et des mesures satellites, ... Le but des réanalyses est de constituer à partir des différentes données collectées une archive homogène de données météorologiques grillées.

2.3.4.4 Modèles emboîtés

Ces méthodes sont constitués par l'emboîtement de modèles ayant différentes caractéristiques pour en améliorer les résultats. Un exemple de ces méthodes est la méthode KAMM/WAsP [87]. L'application de KAMM permet d'obtenir des champs de vent à basse résolution spatiale. Ces champs servent ensuite d'entrée à WAsP. KAMM (Karlsruhe Atmospheric Mesoscale Model) est un modèle atmosphérique méso-échelle non-hydrostatique [88].

Ces méthodes se basent sur la récurrence de situations similaires sur une région donnée. Elles utilisent cette récurrence pour n'effectuer qu'une seule simulation par ensemble de situations similaires. Dans la méthode KAMM/WAsP, une classification sur les paramètres météorologiques, à une échelle synoptique, est effectuée. Cette classification résulte en un ensemble limité de situations à échelle synoptique. A chacune de ces situations est associée sa fréquence d'occurrence. KAMM est ensuite utilisé pour obtenir la situation à méso-échelle correspondante à chacune de ces situations. Comme entrée, KAMM utilise les situations résultantes de la classification. En sortie, un ensemble de situations à méso-échelle est obtenu. A l'aide de la fréquence d'occurrence de ces situations les paramètres statistiques de la distribution du vent sont calculés. Finalement, WAsP utilise ces statistiques à méso-échelle pour établir les paramètres statistiques de la distribution locale du vent. Le fonctionnement de WAsP est décrit dans la section 2.3.4.2.1.1. Cette méthode a été appliquée pour établir, notamment, l'atlas éolien irlandais. Une comparaison entre la densité moyenne de puissance estimée en utilisant la méthode KAMM/WAsP et les mesures recueillies sur quelques sites en Irlande est donnée dans [87]. Dans [87], on trouve, aussi, une comparaison entre les résultats retrouvés en utilisant seulement le modèle KAMM et l'utilisation de KAMM/WAsP. La RMSE sur la densité moyenne de puissance estimée, dans le cas où KAMM est utilisé seul, à une résolution de 2,5 km, est de 251 W m^{-2} ce qui représente 53,9 % de la moyenne des densités moyenne de puissance des sites utilisé pour l'étude. La RMSE en utilisant KAMM et WAsP tombe à 91 W m^{-2} ce qui représente 23 % de la moyenne des densités de puissance.

Le modèle non-hydrostatique MC2, présenté dans la section 2.3.4.1, a été utilisé pour obtenir l'atlas éolien canadien (disponible en ligne, `http://www.windatlas.ca`) avec une résolution spatiale de 5 km. Cet atlas est disponible en ligne. Il fait clairement apparaître les limites de cette méthode en offshore avec des discontinuités apparentes et avec une absence de structures à fine échelle spatiale.

La méthode développée dans cette thèse et présentée dans le chapitre 3 s'inspire de la réduction de la climatologie de la région étudiée en quelques situations types telle qu'utilisée dans ce genre de méthodes. Cependant, au lieu d'utiliser les modèles physiques, on se basera sur les mesures fournies par les moyens de télédétection.

2.4 Conclusion

En pratique, la cartographie du potentiel éolien consiste à établir des cartes donnant la distribution spatiale de la densité moyenne de puissance à une hauteur donnée sur la région étudiée. Ces cartes sont associées à la distribution spatiale des paramètres de Weibull A et k et de la moyenne du vent. L'industrie éolienne a besoin d'une cartographie du potentiel éolien à une résolution spatiale de l'ordre du kilomètre. Pour garantir la précision de la description de la climatologie du vent, il est nécessaire de disposer de données à haute résolution temporelle. Dans ce chapitre, on a étudié les différentes caractéristiques que doivent avoir les mesures de vent pour aboutir à une cartographie à haute résolution spatiale du vent. Ces données peuvent être acquises sur le site prospecté. Cependant, l'installation d'instruments sur site, que ce soit des instruments de mesures en un point (comme des mâts équipés d'anémomètres), ou des instruments mesurant la variabilité spatiale (comme les radars à ouverture réelle) oblige l'investisseur à

attendre pendant une durée d'un an, au moins, avant de pouvoir estimer le potentiel éolien. En outre, les mesures acquises par un mât ne sont représentatives que d'une région très limitée au tour du point de mesure.

Depuis quelques années, les instruments de télédétection ont permis de constituer une archive de données sur le vent en mer. L'inventaire des différents instruments de télédétection, mesurant le vent en mer, fait apparaître qu'aucun de ces instruments ne permet, à lui seul, de fournir des données à haute résolution spatiale et haute résolution temporelle. Pour la cartographie du potentiel éolien il est donc nécessaire de recourir à des méthodes de modélisation ou à une combinaison de différentes données.

Dans ce chapitre, on a présenté les différentes classes de méthodes d'estimation du potentiel éolien existantes. Certaines méthodes (MCP, MLR, CART) se contentent d'estimer le potentiel en un point. Ces méthodes ont besoin en entrée de mesures de vent acquises sur le site prospecté sur une durée d'au moins un an. Les modèles physiques ont besoin de données qui sont, soit indisponible en mer, soit disponible à des résolutions spatiales ne permettant pas le bon fonctionnement des modèles à haute résolution spatiale. La méthode des modèles emboîtés se base sur l'application d'un premier modèle pour la cartographie du vent à basse résolution et d'un deuxième modèle pour améliorer la résolution spatiale. Dans ces méthodes, une classification des situations de vent est entreprise pour réduire le nombre d'applications du modèle à haute résolution spatiale. Ainsi, ce modèle ne sera appliqué qu'une seule fois par classe.

Dans le chapitre suivant, on propose une nouvelle méthode, permettant de répondre aux exigences de l'industrie éolienne, pour la cartographie du potentiel éolien offshore. Cette méthode se base sur la combinaison de différentes sources de données sur le vent en mer.

Bibliographie

[1] J. R. Holton. *Introduction to dynamic meteorology*. Academic Press, New York, USA, 320 p., 1972.

[2] I. Orlanski. A rational subdivision of scales for atmospheric processes. *Bulletin of American Meteorological Society*, vol. 56, pp. 527–530, 1975.

[3] C. G. Justus, W. R. Hargraves et A. Yalcin. Nationwide assessment of potential output from wind-powered generators. *Journal of Applied Meteorology*, vol. 15, pp. 673–678, 1976.

[4] J. P. Hennessey, Jr. Some Aspects of Wind Power Statistics. *Journal of Applied Meteorology*, vol. 16, pp. 119–128, 1977.

[5] K. Conradsen, L. B. Nielsen et L. P. Prahm. Review of Weibull Statistics for Estimation of Wind Speed Distributions. *Journal of Applied Meteorology*, vol. 23, pp. 1173–1183, 1984.

[6] A. H. Monahan. The Probability Distribution of Sea Surface Wind Speeds. Part I : Theory and SeaWinds Observations. *Journal of Climate*, vol. 19, pp. 497–520, 2006.

[7] E. L. Petersen, I. Troen, S. Frandsen et K. Hedegaard. *Wind atlas for Denmark. A rational method for wind energy siting.* Risoe National Laboratory, Roskilde, Danemark, 229 p., 1981.

[8] I. Troen et E. L. Petersen. *European Wind Atlas*. Risoe National Laboratory, Roskilde, Danemark, 656 p., 1989.

[9] A. H. Monahan. The Probability Distribution of Sea Surface Wind Speeds. Part II : Dataset Intercomparison and Seasonal Variability. *Journal of Climate*, vol. 19, pp. 521–534, 2006.

[10] C. G. Justus, W. R. Hargraves, A. Mikhail et D. Graber. Methods for Estimating Wind Speed Frequency Distributions. *Journal of Applied Meteorology*, vol. 17, pp. 350–353, 1978.

[11] S. C. Pryor, M. Nielsen, R. J. Barthelmie et J. Mann. Can Satellite Sampling of Offshore Wind Speeds Realistically Represent Wind Speed Distributions ? Part II : Quantifying Uncertainties Associated with Distribution Fitting Methods. *Journal of Applied Meteorology*, vol. 43, pp. 739–750, 2004.

[12] N. Fichaux, T. Poglio et T. Ranchin. Mapping offshore wind resources : synergetic potential of SAR and scatterometer data. *IEEE Journal of Oceanic Engineering*, vol. 30, n° 3, pp. 48–57, 2005.

[13] R. J. Barthelmie et S. C. Pryor. Can Satellite Sampling of Offshore Wind Speeds Realistically Represent Wind Speed Distributions ?. *Journal of Applied Meteorology*, vol. 42, pp. 83–94, 2003.

[14] N. Fichaux. *Evaluation du potentiel éolien offshore et imagerie satellitale*. Thèse de doctorat, Ecole des Mines de Paris, Paris, France, 119 p., 2003.

[15] European Wind Energy Association. Wind energy — the facts. En ligne, 2004. Disponible sur : `http://www.ewea.org/index.php?id=91` (consulté le 20-08-2007).

[16] JCOMM *in situ* Observing Platform Support Centre. DBCP status by country, July 2007 data reported on GTS via Meteo France. En ligne. Disponible sur : `http://www.jcommops.org/` (consulté le 20-08-2007).

[17] K. S. Hansen, M. S. Courtney, P. Hummelshoj, P. Friis et A. Knittel. Validation of SODAR measurements for wind power assessment. In *Proceedings of the European Wind Energy Conference, EWEC 2006, 27 February - 2 March 2006, Athens, Greece*, 2006.

[18] H. E. Jorgensen et I. Antoniou. Intercomparision of two commercially available Sodars. Technical report, Riso National Laboratory, Roskilde, Danemark, 22 p., 2002.

[19] D. A. Smith, M. Harris, A. S. Coffey, T. Mikkelsen, H. E. Jørgensen, J. Mann et R. Danielian. Wind lidar evaluation at the Danish wind test site in Høvsøre. *Wind Energy*, vol. 9, pp. 87–93, 2006.

[20] I. Antoniou, H. E. Jorgensen, T. Mikkelsen, S. Frandsen, R. Barthelmie, C. Perstrup et M. Hurtig. Offshore wind profile measurements from remote sensing instruments. In *Proceedings of the European Wind Energy Conference, EWEC 2006, 27 February - 2 March 2006, Athens, Greece*, 2006.

[21] H. Nett et M. Endemann. Atmospheric Dynamics Mission : Aeolus. In *Proceedings of the 2004 IEEE International Geoscience and Remote Sensing Symposium, 20-24 Sept. 2004*, vol. 2, pp. 1190–1195, 2004.

[22] H. Dankert, J. Horstmann et W. Rosenthal. Ocean wind fields retrieved from radar-image sequences. *Journal of Geophysical Research*, vol. 108, n° C11, pp. 33–52, 2003.

[23] H. Dankert et J. Horstmann. Wind measurements at FINO-I using marine radar-image sequence. In *Proceedings of IEEE International Geoscience and Remote Sensing Symposium, IGARSS'2006*, pp. 1296–1299, 2006.

[24] P. Gloersen et F. T. Barath. A Scanning Multichannel Microwave Radiometer for Nimbus-G and SEASAT-A. *IEEE Journal of Oceanic Engineering*, vol. OE-2, pp. 172–178, 1977.

[25] F. J. Wentz. Measurement of oceanic wind vector using satellite microwave radiometers. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 30, pp. 960–972, 1992.

[26] T. Meissner et F. Wentz. The ocean algorithm suite for the Conical-scanning Microwave Imaging/Sounder (CMIS). In *Proceedings of IEEE International Geoscience and Remote Sensing Symposium, 2002. IGARSS '02*, vol. 2, pp. 813–815, 2002.

[27] F. M. Monaldo. Evaluation of WindSat wind vector performance with respect to QuikSCAT estimates. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, pp. 638–644, 2006.

[28] D. L. Evans. Seasat—A 25–year legacy of success. *Remote Sensing of Environment*, vol. 94, n° 3, pp. 384–404, 2005.

[29] C.H. Davis. Satellite radar altimetry. *IEEE Transactions on Microwave Theory and Techniques*, vol. 40, n° 6, pp. 1070–1076, 1992.

[30] Centre National d'Etudes Spatiales. Archivage, validation et interprétation des données des satellites océanographiques. En ligne. Disponible sur : `http://www-aviso.cnes.fr` (consulté le 20-08-2007).

[31] G.S. Brown, H.R. Stanley, N.A. Roy. The wind-speed measurement capability of spaceborne radar altimeters. *IEEE Journal of Oceanic Engineering*, vol. OE-6, n° 2, pp. 59–63, 1981.

[32] M.H. Freilich. Seawinds algorithm theoretical basis document. En ligne. Disponible sur : `http://podaac.jpl.nasa.gov/quikscat/qscat_doc.html` (consulté le 20-08-2007).

[33] Ad. Stoffelen. *Scatterometry - De scatterometer*. Thèse de doctorat, Université d'Utrecht, Pays-Bas, 209 p., 1998.

[34] IFREMER CERSAT. Off-line wind scatterometer ERS products - User Manual, Version 2.1. Technical report, C2-*MUT-W-01-IF, IFREMER - CERSAT, BP 70, 29280 Plouzané, France, 86 p., 1999.

[35] H. Hersbach. *CMOD5 : An improved geophysical model function for ERS C-band scatterometry*. ECMWF Technical Memoranda, 50 p., 2003.

[36] C. Cappa S. Zecchetto. The spatial structure of the mediterranean sea winds revealed by ers1 scatterometer. *International Journal of Remote Sensing*, vol. 22, n° 1, pp. 45–70, 2001.

[37] ESA. European remote sensing satellites : A better view of the earth. En ligne. Disponible sur : `http://earth.esa.int/ers/` (consulté le 20-08-2007).

[38] NASDA. Advanced Earth observing satellite ADEOS MIDORI. En ligne. Disponible sur : `http://kuroshio.eorc.jaxa.jp/ADEOS/` (consulté le 20-08-2007).

[39] NASDA. ADEOS–2 : MIDORI–2. En ligne. Disponible sur : `http://sharaku.eorc.jaxa.jp/ADEOS2/` (consulté le 20-08-2007).

[40] M. G. Schlax, D. B. Chelton et M. H. Freilich. Sampling Errors in Wind Fields Constructed from Single and Tandem Scatterometer Datasets. *Journal of Atmospheric and Oceanic Technology*, vol. 18, pp. 1014–1036, 2001.

[41] N. Ebuchi. Evaluation of wind vectors observed by QuikSCAT/SeaWinds using ocean buoy data. In *Proceedings of IEEE 2001 International Geoscience and Remote Sensing Symposium, IGARSS'01*, vol. 3, pp. 1082 – 1085, 2001.

[42] JAXA. Japanese Earth resources satellite – 1. En ligne. Disponible sur : `http://www.eorc.jaxa.jp/JERS-1/` (consulté le 20-08-2007).

[43] Centre Canadien de Télédétection. Radarsat–1 technical specifications. En ligne. Disponible sur : `http://cct.rncan.gc.ca/radar/spaceborne/radarsat1/specs/radspec_f.php` (consulté le 20-08-2007).

[44] ESA. Envisat caring for the earth. En ligne. Disponible sur : `http://envisat.esa.int/` (consulté le 20-08-2007).

[45] F. Monaldo et V. Kerbaol. The SAR measurement of ocean surface winds : an overview. In *Proceedings of the Second Workshop on Coastal and Marine Applications of SAR, ESA SP-565*, pp. 15–32, 2004.

[46] J. Choisnard, M. Bernier, G. Lafrance. Cartographie des vents dans le golfe du Saint–Laurent à l'aide de l'imagerie Radarsat–1. *Journal Canadien de Télédétection*, vol. 30, n° 4, pp. 604–616, 2004.

[47] V. Kerbaol. *Analyse spectrale et statistique vent–vagues des images radar à ouverture synthétique (ROS) — Application aux données des satellites ERS–1/2*. Thèse de doctorat, IFREMER, Brest, France, 185 p., 1997.

[48] F. M. Monaldo, D. R. Thompson, R. C. Beal, W. G. Pichel et P. Clemente-Colon. Comparison of SAR-derived wind speed with model predictions and buoy comparisons. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, n° 12, pp. 2587–2600, 2001.

[49] E. Attema, Y.-L. Desnos et G. Duchossois. Synthetic aperture radar in Europe : ERS, Envisat, and beyond. *Johns Hopkins APL Technical Digest*, vol. 21, pp. 155–169, 2000.

[50] J. Horstmann, W. Koch, S. Lehner et R. Tonboe. Wind retrieval over the ocean using synthetic aperture radar with C-Band HH polarization. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 38, pp. 2122–2131, 2000.

[51] J. Horstmann, S. Lehner, W. Koch et R. Tonboe. Computation of wind vectors over the ocean using spaceborne syntheric aperture radar. *Johns Hopkins APL Technical Digest*, vol. 21, pp. 100–107, 2000.

[52] P. W. Vachon et E. W. Dobson. Wind retrieval from RADARSAT SAR images : Selection of a suitable C-Band HH polarization wind retrieval model. *Canadian Journal of Remote Sensing*, vol. 26, pp. 306–313, 2000.

[53] S. A. Hsu, E. A. Meindl et D. B. Gilhousen. Determining the Power-Law Wind-Profile Exponent under Near-Neutral Stability Conditions at Sea. *Journal of Applied Meteorology*, vol. 33, pp. 757–772, 1994.

[54] S. Emeis. Vertical variation of frequency distributions of wind speed in and above the surface layer observed by sodar. *Meteorologische Zeitschrift*, vol. 10, n° 2, pp. 141–149, 2001.

[55] N. Allnoch. Windkraftnutzung im nordwestdeutschen Binnenland : Ein System zur Standortbewertung für Windkraftanlagen. Technical report, Geographische Kommission für Westfalen, Münster, ARDEY-Verlag, 160 p., 1992.

[56] J. Wieringa. Shapes of annual frequency distributions of wind speed observed on high meteorologicalmasts. *Boundary Layer Meteorology*, vol. 47, pp. 85–110, 1989.

[57] J. Tambke, J. A. T. Bye, J.-O. Wolff, S. Tautz, B. Lange, M. Lange et U. Focken. Modelling offshore wind profiles using inertially coupled wave boundary layers. In *Proceedings of the European Wind Energy Conference, EWEC 2004, 22-25 November, London, U.K.*, 2004.

[58] P. E. Francis. The effect of changes of atmospheric stability and surface roughness on off-shore winds over the east coast of Britain. *Meteorology Magazine*, vol. 99, pp. 130–138, 1970.

[59] S. Sethuraman et G. S. Raynor. Comparison of mean wind speeds and turbulence at a coastal site and an offshore location. *Journal od Applied Meteorology*, vol. 19, pp. 15–21, 1980.

[60] D. Lindley, P.B. Simson, U. Hassan et D. Milborrow. Assessment of offshore siting of wind turbine generators. In *Proceedings of the 3rd International Symposium on Wind Energy Systems, 26-29 August, Cranfield, U.K.*, pp. 17–42, 1980.

[61] S. A. Hsu. *Coastal Meteorology*. Academic Press, Londres, Royaume-Uni, 260 p., 1988.

[62] A. L. Rogers, J. W. Rogers et J. F. Manwell. Vertical variation of frequency distributions of wind speed in and above the surface layer observed by sodar. *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 93, pp. 243–264, 2005.

[63] A. Derrick. Development of the measure-correlate-predict strategy for site assessment. In *Proceedings of European Community Wind Energy Conference, Lübeck-Travemünde, Germany, 8-12 March, 1993*, pp. 681–685, 1993.

[64] M. Nielsen, L. Landberg, N. G. Mortensen, R. J. Barthelmie et A. Joensen. Application of the measure-correlate-predict approach for wind resource assessment. In *Proceedings of the 2001 European wind energy conference and exhibition (EWEC '01), Copenhagen (DK), 2-6 Jul 2001*, pp. 773–776, 2001.

[65] J. L. Walmsley, R. J. Barthelmie et W. R. Burrows. The statistical prediction of offshore winds from land-based data for wind energy applications. *Boundary-Layer Meteorology*, vol. 101, pp. 409–433, 2001.

[66] R. J. Barthelmie. Monitoring offshore wind and turbulence characteristics in Denmark. In *Proceedings of the 21st British Wind Energy Association Conference, Cambridge, 1–3 September 1999, Professional Engineering Publishing, Cambridge*, pp. 311–321, 1999.

[67] R. J. Barthelmie, B. Lange et M. Nielsen. *Wind Resources at Rodsand and Omo Stalgrunde*. Report Risø-I-1456(EN), Risø National Laboratory, Roskilde, Denmark, 96 pp., 1999.

[68] C. F. Ratto, R. Festa, C. Romeo, O. A. Frumento et Galluzzi M. Mass-consistent models for wind fields over complex terrain : the state of the art. *Environmental Software*, vol. 9, pp. 247–268, 1994.

[69] Philippe Beaucage, Anna Glazer, Wei Yu, Robert Benoit, Gaëtan Lafrance et Monique Bernier. Wind assessment using synthetic aperture radar (SAR) satellite and the mesoscale compressible community (MC2) model. In *Proceedings of the European Wind Energy Conference, EWEC 2006, 27 February - 2 March 2006, Athens, Greece*, 2006.

[70] Robert Benoit, Michel Desgagné, Pierre Pellerin, Simon Pellerin, Yves Chartier et Serge Desjardins. The Canadian MC2 : A semi-lagrangian, semi-implicit wideband atmospheric model suited for finescale process studies and simulation. *Monthly Weather Review*, vol. 125, pp. 2382–2415, 1997.

[71] U. Hogstrom, R. Taesler, K . Karlsson, L. Enger et A. S. Smedman-Hogstrom. The Uppsala urban meteorological project. *Boundary Layer Meteorology*, vol. 15, pp. 69–80, 1977.

[72] H. Kapitza et D. P. Eppel. The non-hydrostatic mesoscale model GESIMA. part I : dynamical equations and tests. *Beitr. Phys. Atmosph.*, vol. 65, pp. 129–146, 1992.

[73] M. C. Brower, J. W. Zack et B. H. Baile. Validation and application of MesoMap. In *Proceedings de Windpower, AWEA ed., Boston, Massachussets*, 2000.

[74] T. Black. The new NMC mesoscale ETA model : description and forecast examples. *Weather Forecasting*, vol. 9, pp. 265–278, 1994.

[75] P. Unden, L. Rontu, H. Järvinen, P. Lynch, J. Calvo, G. Cats, J. Cuxart, K. Eerola, C. Fortelius, J. A. Garcia-Moya, C. Jones, G. Lenderlink, A. McDonald, R. McGrath, B. Navascues, N. W. Nielsen, V. Odegaard, E. Rodriguez, M. Rummukainen, R. Rõ om, K. Sattler, B. H. Sass, H. Savijärvi, B. W. Schreur, R. Sigg, H. The et A. Tijm. HIRLAM-5 scientific documentation. Technical report, SMHI, S-601 76 Norrköping, Sweeden, 144 p., 2002.

[76] J. P. Lafore, J. Stein, N. Asencio, P. Bougeault, V. Ducrocq, J. Duron, C. Fischer, P. Héreil, P. Mascart, V. Masson, J. P. Pinty, J. L. Redelsperger, E. Richard et J. Vilà-Guerau de Arellano. The Meso-NH Atmospheric Simulation System. Part I : adiabatic formulation and control simulations. *Annales Geophysicae*, vol. 16, pp. 90–109, 1997.

[77] G. Doms et U. Schättler. The nonhydrostatic limited-area model LM, part 1 : scientific documentation. Technical report, R&D Dep. of DWD, Germany, 169 p., 1999.

[78] C. J. Tremback et R. L. Walko. RAMS - Regional Atmospheric Modeling System Version 6.0 - User's Guide - Introduction. Technical report, ATMET, LLC, Boulder, Colorado, 10 p., 2005.

[79] G. A. Grell, J. Dudhia et D. R. Stauffer. A Description of the Fifth-Generation Penn State/NCAR Mesoscale Model (MM5). Technical report, NCAR Tech. Note, NCAR/TN-398 + STR, 122 p., 1995.

[80] P. Thunis et A. Clappier. Formulation and evaluation of a nonhydrostatic mesoscale vorticity model (TVM). *Monthly Weather Review*, vol. 128, pp. 3236–3251, 2000.

[81] N. G. Mortensen, L. Landberg, I. Troen, E. L. Petersen, O. Rathmann et M. Nielsen. WAsP utility programs. Technical report, Risø-I-2261(EN), Riso National Laboratory, Roskilde, Danemark, 54 p., 2004.

[82] P. Astrup, T. Mikkelsen et N. O. Jensen. A fast model for mean and turbulent wind characteristics over terrain with mixed surface roughness. *Radiat Prot Dosimetry*, vol. 73, n° 1-4, pp. 257–260, 1997.

[83] P. Geai. Méthode d'interpolation et de reconstitution tridimensionnelle d'un champ de vent : le code d'analyse objective MINERVE. Technical report, DER/HE/34-87.03, Electricité De France, Chatou, France, 177 p., 1987.

[84] E. Kalnay, M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, Y. Zhu, A. Leetmaa, B. Reynolds, M. Chelliah, W. Ebisuzaki, W. Higgins, J. Janowiak, K. C. Mo, C. Ropelewski, J. Wang, R. Jenne et D. Joseph. The NCEP/NCAR 40-Year Reanalysis Project. *Bulletin of the American Meteorological Society*, vol. 77, pp. 437–472, 1996.

[85] S. M. Uppala, P. W. Kallberg, A. J. Simmons, U. Andrae, V. da Costa Bechtold, M. Fiorino, J. K. Gibson, J. Haseler, A. Hernandez, G. A. Kelly, X. Li, K. Onogi, S. Saarinen, N. Sokka, R. P. Allan, E. Andersson, K. Arpe, M. A. Balmaseda, A. C. M. Beljaars, L. van de Berg, J. Bidlot, N. Bormann, S. Caires, F. Chevallier, A. Dethof, M. Dragosavac, M. Fisher, M. Fuentes, S. Hagemann, E. Holm, B. J. Hoskins, L. Isaksen, P. A. E. M. Janssen, R. Jenne, A. P. McNally, J.-F. Mahfouf, J.-J. Morcrette, N. A. Rayner, R. W. Saunders, P. Simon, A. Sterl, K. E. Trenberth, A. Untch, D. Vasiljevic, P. Viterbo et J. Woollen. The ERA-40 re-analysis. *Quarterly Journal of the Royal Meteorological Society*, vol. 131, pp. 2961–3012, 2005.

[86] S. Schubert, R. Rood, C.-K. Park, C.-Y. Wu, Y. Kondratyeva, A. Molod, L. Takacs, M. Seablom et W. Higgins. A multiyear assimilation with the GEOS-1 system : Overview and results. In M. J. Suarez, editor, *NASA technical report series on global modeling and data assimilation.*, vol. 6, 183 p., 1995.

[87] H. P. Frank, O. Rathmann, N. G. Mortensen et L. Landberg. The NumericalWind Atlas - the KAMM/WAsP Method. Technical report, Report Risø-R-1252((EN), Risø National Laboratory, Roskilde, Denmark, 60 p., 2001.

[88] G. Adrian et F. Fiedler. Simulation of unstationary wind and temperature fields over complex terrain and comparison with observations. *Beiträge zur Physik der Atmosphäre*, vol. 64, pp. 27–48, 1991.

Chapitre 3

Fusion de données pour l'évaluation du potentiel éolien

Après l'analyse des différentes techniques de mesure et de cartographie du potentiel éolien faite au chapitre 2, il apparaît qu'aucune des sources de données sur le vent en mer ne permet, à elle seule, d'établir une cartographie du potentiel éolien satisfaisant les besoins de l'industrie éolienne. Les techniques de modélisation, présentées au chapitre 2, sont caractérisées par des résultats insatisfaisants en mer. Ceci est dû à un manque de données sur les phénomènes influençant le vent en mer et à l'interface mer/terre qui perturbe leur bon fonctionnement. On propose, dans ce chapitre, une méthode permettant de tirer avantage de différents instruments de télédétection permettant d'accéder à des mesures de vent en mer. Cette méthode se base sur la fusion de données pour allier la haute résolution spatiale de certaines données et la haute résolution temporelle d'autres. La première section de ce chapitre est consacrée à la définition de quelques termes relatifs à la fusion de données et qui seront utilisés par la suite. Dans la section 2, on décrit les différentes étapes du schéma de fusion de données proposé. Ce schéma se base sur l'élaboration de fonctions de transfert de la basse à la haute résolution spatiale. Ces fonctions, permettant la synthèse de la haute résolution spatiale, sont décrites dans la section 3. Pour la synthèse de la haute résolution spatiale, on se place dans le cadre du concept ARSIS (Amélioration de la Résolution Spatiale par Injection de Structures) pour la fusion d'images. On présente ce concept dans la section 5. La section 4 est consacrée à la présentation de l'analyse multi-échelle qu'on utilisera dans l'implémentation de méthodes relevant du concept ARSIS.

3.1 Fusion de données, notations et définitions

Dans cette section, on définit quelques termes désignant des opérations souvent nécessaires lors d'un processus de fusion. Ensuite, pour répondre à la nécessité de trouver une représentation des processus de fusion, on décrit la représentation proposée [1]. Elle sera adoptée dans la suite du manuscrit.

3.1.1 Définitions

3.1.1.1 Fusion de données

Plusieurs définitions de la fusion de données existent. Dans le domaine de la télédétection, [2] propose la définition suivante : « la fusion de données constitue un cadre formel dans lequel s'expriment les moyens et techniques permettant l'alliance des données provenant de sources diverses ». La définition ajoute que la fusion de données vise à l'obtention d'information de plus grande qualité ; le définition exacte de « plus grande qualité » dépendant de l'application.

Tout processus de fusion prend en entrée des données venant de sources différentes ou acquises à des moments différents. Le processus de fusion doit prendre en compte ces différences. Le premier traitement des données entrant dans un processus de fusion est un traitement pour garantir la possibilité de fusion des différentes données. Ce traitement est appelé « alignement ».

Sa définition est donnée dans le paragraphe suivant.

3.1.1.2 Alignement

Les informations entrant dans un processus de fusion doivent être alignées [2]. L'alignement consiste à définir une représentation commune aux différentes informations à fusionner. Par exemple, l'alignement consiste dans le cas d'images de télédétection à trouver un système de projection spatiale commun.

L'alignement ne garantit cependant pas que les différentes données sont représentatives du même phénomène ou du même objet. Il est nécessaire donc d'« associer » les données à fusionner. La définition de l'association est donnée dans le paragraphe suivant.

3.1.1.3 Association

L'association de deux représentations, alignées, requiert que ces deux représentations réfèrent au même objet ou phénomène observé [2]. En effet, la fusion de deux représentations référentes à des objets et des phénomènes distincts a peu de signification. La nécessité d'une association des représentations peut être due aux périodes d'observation et aux changements pouvant survenir sur l'objet observé.

Le vent est un phénomène dynamique, un délai entre deux observations peut entraîner une invalidation de la fusion. Il est donc nécessaire de définir une association des représentations acquises à différentes dates.

3.1.2 Représentation d'une opération de fusion

On a choisi une représentation proposée par [1] et adoptée par [2]. Dans cette représentation de la fusion, trois types d'entrées au processus de fusion sont possibles : les sources de données à fusionner, des informations auxiliaires et des connaissances externes. Les informations auxiliaires sont des informations provenant d'un traitement spécifique d'une des sources de données ou des informations résultant d'une autre opération de fusion. Les résultats du processus de fusion peuvent aussi agir comme information auxiliaire dans le cas d'un processus de fusion itératif. Les connaissances externes représentent des connaissances *a priori* sur le processus de fusion. Elles guident le processus de fusion. Ces connaissances

externes peuvent être des lois physiques ou des lois empiriques. Le schéma représentant une opération de fusion est présenté à la figure 3.1.

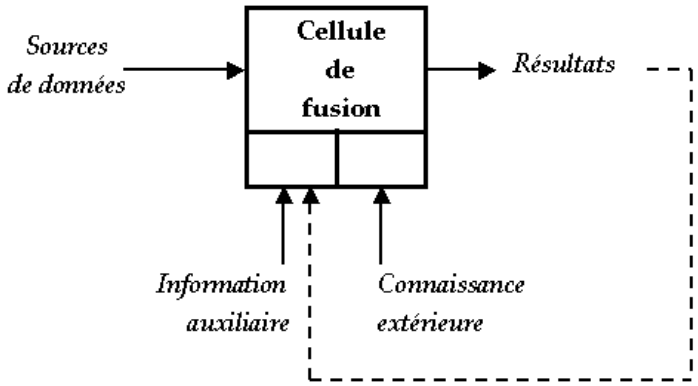


FIG. 3.1 – Représentation d'un processus de fusion. D'aprés [1]

Une méthode de fusion peut être composée de plusieurs cellules de fusion, certaines utilisant les sorties des autres.

3.2 Schéma de fusion pour la cartographie du potentiel éolien offshore

La cartographie du potentiel éolien consiste en pratique à décrire de manière compacte la climatologie du vent sur une région donnée avec une résolution spatiale permettant de faire apparaître les variations spatiales pouvant influencer la production des parcs éoliens. Ceci se fait au travers de cartes des paramètres statistiques du vent. Comme montré dans le chapitre 2, il est nécessaire pour une description précise de la climatologie du vent de disposer de mesures à hautes résolution temporelle. Les mesures diffusométriques ont une résolution temporelle satisfaisante pour cette description. Cependant, ces mesures ont une résolution spatiale trop grossière pour la cartographie du potentiel éolien. Les images ROS permettent quant à elles d'accéder à des champs de vent à haute résolution spatiale. Ces mesures, présentent l'inconvénient d'être à basse résolution temporelle et ne permettent donc pas d'établir des statistiques fiables décrivant la climatologie du vent. La fusion de données nous permet de bénéficier des avantages de chacune de ces sources. On combine, ainsi, la haute résolution temporelle des mesures diffusométriques et la haute résolution spatiale des mesures ROS.

Deux approches sont possibles pour la fusion de données en vue de la cartographie du potentiel éolien. La première approche consiste à passer par une cartographie du potentiel éolien à basse résolution spatiale, à partir des mesures diffusométriques. Cette cartographie à basse résolution spatiale est ensuite fusionnée avec les données à haute résolution spatiale fournies par les images ROS. La deuxième approche consiste à fusionner les deux séries de données (la série de données à basse résolution spatiale et la série de données à haute résolution spatiale) pour générer une série de données synthétiques ayant à la fois la haute résolution spatiale, celle des champs de vent retrouvés grâce aux images ROS, et la haute résolution temporelle, celle des mesures diffusométriques. On choisit cette deuxième approche car elle

présente l'avantage d'être indépendante de la représentation choisie pour la cartographie.

La série synthétique à haute résolution spatiale et à haute résolution temporelle est générée en utilisant une fonction de transfert permettant de passer de la basse résolution spatiale à la haute résolution spatiale. Cette fonction est appliquée à chacune des mesures à basse résolution spatiale. Pour des raisons physiques, cette fonction de transfert ne peut pas être universelle. On s'appuie sur la propriété d'invariance d'échelle pour élaborer un nombre limité de fonctions de transfert types. Les hypothèses sur lesquelles s'appuie le schéma de fusion proposé dans ce chapitre sont explicitées dans le paragraphe suivant.

3.2.1 Hypothèses du schéma de fusion

Les champs de vent mesurés par les instruments de télédétection sont des mesures moyennes et en deux dimensions de l'écoulement atmosphérique à un instant t. Ceci est compatible avec l'extension spatiale des champs considérés (de 100 km à 1000 km). Par définition, plus la résolution spatiale de l'instrument est grossière, plus la zone sur laquelle est moyennée la mesure est grande. Ces moyennes intègrent à la fois les composantes locales de la circulation atmosphérique ainsi que la turbulence et les phénomènes locaux présents dans les plus petites échelles. Considérant l'écoulement du vent de surface marine dans un secteur géographique et à un instant donné, nous considérons que les mesures diffusométriques donnent la valeur moyenne du champ de vent local (*i.e.* turbulence et circulation locale). Les champs de vent retrouvés à partir d'images ROS permettent, par différence, de retrouver la composante turbulente du champ de vent.

Un exemple d'effet local est l'effet de deux îles sur le flux. Si la dimension des îles est plus petite que la résolution spatiale des mesures diffusométriques, l'effet de ces îles ne se verra pas sur ces mesures. Une image ROS permettra d'accéder à la représentation de cet effet.

La turbulence est variable à toutes les échelles, dans l'espace et dans le temps. Selon la théorie de Kolmogorov, on peut définir trois gammes d'échelles caractérisant le transfert d'énergie turbulente entre les niveaux de production et de dissipation :

- les grandes échelles, caractérisant l'écoulement, responsables de la production d'énergie turbulente, en pleine mer, c'est l'échelle des structures météorologiques (anticyclones, dépressions, ...)
- les échelles intermédiaires, qui définissent la zone inertielle, où s'effectue le transfert de l'énergie des grandes échelles vers les petites à un taux constant,
- les petites échelles, où l'énergie est dissipée par la viscosité.

Si aucun obstacle ou cisaillement n'est présent dans la zone d'étude, à une échelle caractéristique intermédiaire entre l'échelle de génération de la turbulence et l'échelle de dissipation, les champs de vent considérés ici sont situés dans la zone inertielle. En effet, les diffusométres et les ROS permettent d'accéder aux mesures de vent dans cette gamme d'échelles.

Selon la théorie de Kolmogorov, *« la turbulence est statistiquement homogène (invariante par translation spatiale, loin des bords), isotrope (invariante par rotation), autosimilaire (invariante par dilatation des échelles), et stationnaire (invariance par échelle temporelle) ; les propriétés statistiques ne dépendent que de l'énergie dissipée ; l'énergie est transférée, sans dissipation, à travers les échelles, dans la zone*

inertielle » [3].

La théorie de Kolmogorov donne la loi de répartition des tourbillons en fonction des échelles spatiales. Il peut cependant exister des situations, dites de « cascade inverse », où la vorticité est générée à petite échelle (cyclogénèse et tornade par exemple). Cependant, notre intérêt se porte sur les statistiques de la climatologie du vent et non sur des cas particuliers. Cet aspect n'aura pas une grande influence sur la climatologie d'une zone.

En mécanique des fluides, la notion de similitude des écoulements est définie pour extrapoler les résultats ou mesures obtenus à une échelle donnée. On dit que deux écoulements sont semblables s'ils vérifient deux conditions [4] relatives à la mécanique des fluides :

- similitude géométrique : les frontières des deux écoulements (E) et (E') se déduisent l'une de l'autre par une similitude de rapport K_l. Dans notre cas, la zone est la même et seules les dates d'observation des écoulements changent. La similitude géométrique est donc vérifiée.
- similitude cinématique : en deux points homologues, les vitesses sont dans un rapport $K_v = \frac{v'}{v}$ qui est indépendant du couple (M,M') choisi. v étant la vitesse de l'écoulement au point M et v' la vitesse de l'écoulement au point M'. Cette similitude cinématique doit être vérifiée dans notre cas. Elle peut être vérifiée par l'utilisation d'une mesure, ou distance, quantifiant la similarité entre deux champs de vent.

On désigne par $SCAT_t$ un champ de vent mesuré par un diffusomètre à un instant t et par SAR_t un champ de vent mesuré par radar à ouverture synthétique au même instant. L'hypothèse qu'on considère donc pour le schéma de fusion présenté dans la suite est la suivante :

$$\boxed{\text{Si } SCAT_{t_1} \text{ est similaire à } SCAT_{t_2} \text{ alors } SAR_{t_1} \text{ est similaire à } SAR_{t_2}.}$$

Cette hypothèse implique que pour des champs de vent similaires à basse résolution, la même fonction de transfert peut être appliquée pour obtenir des champs de vent à haute résolution spatiale. La fonction de transfert peut être inférée à partir des situations où, à la fois, une mesure diffusométrique (à basse résolution spatiale) et une mesure ROS (à haute résolution spatiale) sont disponibles. Cette fonction de transfert est ensuite appliquée aux champs de vent à basse résolution similaires à la situation où la fonction de transfert a été établie.

3.2.2 Schéma de la méthode fusion

Le schéma général de la méthode de fusion est donné figure 3.2. En entrée du processus de fusion on dispose de deux sources de données. La première source de données présente des données à haute résolution spatiale et basse résolution temporelle. Il s'agit dans notre étude du ROS. La deuxième source de données présente des données à haute résolution temporelle et basse résolution temporelle. Il s'agit, ici, du diffusomètre. A la sortie du processus de fusion, on cherche à obtenir une cartographie du potentiel éolien à haute résolution spatiale avec une incertitude associée à cette cartographie.

Le schéma de fusion se base sur l'hypothèse que deux champs de vent similaires à basse résolution spatiale sont similaires à haute résolution spatiale. Pour un ensemble de champs de vent similaires, on peut donc associer un champ de vent à haute résolution spatiale. Ce dernier contient les structures à

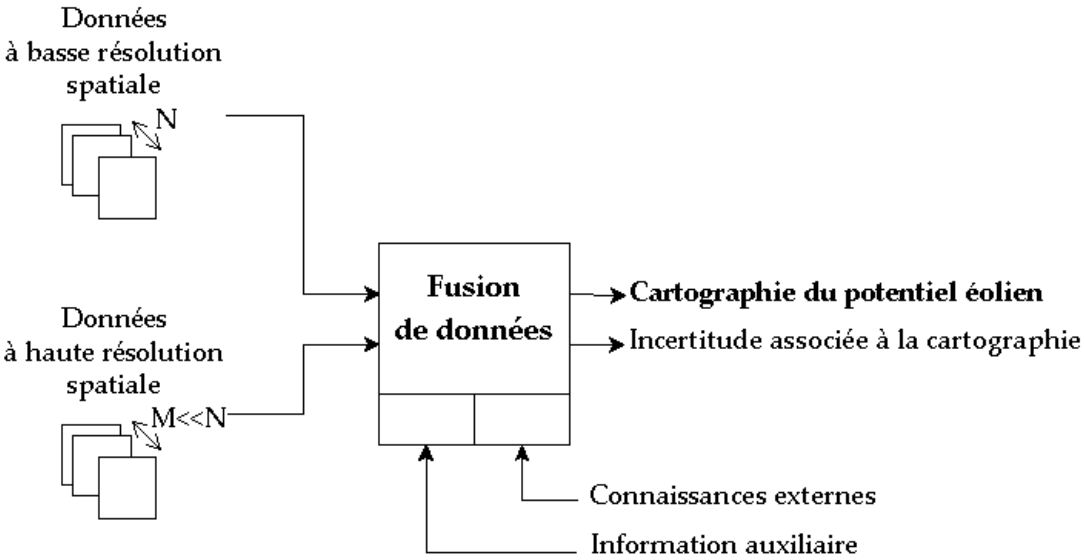


FIG. 3.2 – Schéma général de la méthode de fusion de données pour la cartographie du potentiel éolien offshore.

haute résolution spatiale qui peuvent être injectées dans les champs de vent à basse résolution spatiale. Le regroupement des champs de vent similaires se fait au travers d'une classification. Le but de la classification est de former des groupes distincts de champs de vent. Chaque groupe contient des champs de vent ayant le même comportement spatial. A chacun de ces groupes on associe une mesure à haute résolution spatiale. Cette mesure à haute résolution spatiale sert à établir une fonction de transfert pour chaque classe. Une fois qu'une fonction de transfert est associée à chaque classe, cette fonction est appliquée sur les autres mesures de la classe pour la synthèse de la haute résolution spatiale. L'ensemble des champs de vent synthétisés est alors utilisé pour établir la cartographie du potentiel éolien de la région étudiée. Les figures 3.3 et 3.4 présentent une représentation plus détaillée du schéma de fusion proposé. La figure 3.3 s'arrête au niveau de la génération d'un catalogue de fonctions de transfert types, c'est-à-dire une fonction de transfert type pour chaque classe de champs de vent. La figure 3.4 présente le reste de la méthode de fusion qui aboutit à l'établissement d'une cartographie du potentiel éolien.

Les différentes étapes de la méthode proposée ici sont la classification, l'association, la synthèse de la haute résolution spatiale et la description statistique. Ces différentes étapes sont décrites dans la suite de ce chapitre.

3.2.2.1 Classification

La première étape de la méthode de fusion est la classification de la série de données à basse résolution spatiale et haute résolution temporelle. Il résulte de cette classification une réduction de série de données en quelques situations types représentatives de la climatologie du vent à basse résolution sur la région étudiée. Chaque situation type regroupe des champs de vent similaires. Cette étape de classification utilise une mesure de similarité. Cette mesure définit de manière objective la similarité entre deux

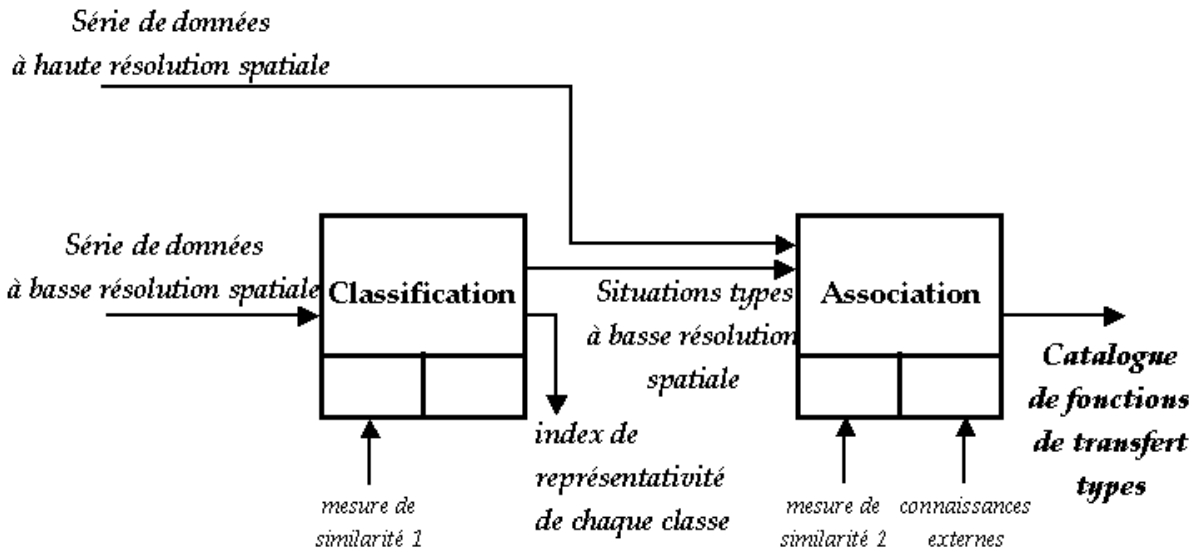


FIG. 3.3 – Première partie du schéma de fusion de données. Cette partie représente les étapes menant l'établissement d'un catalogue de fonctions de transfert types.

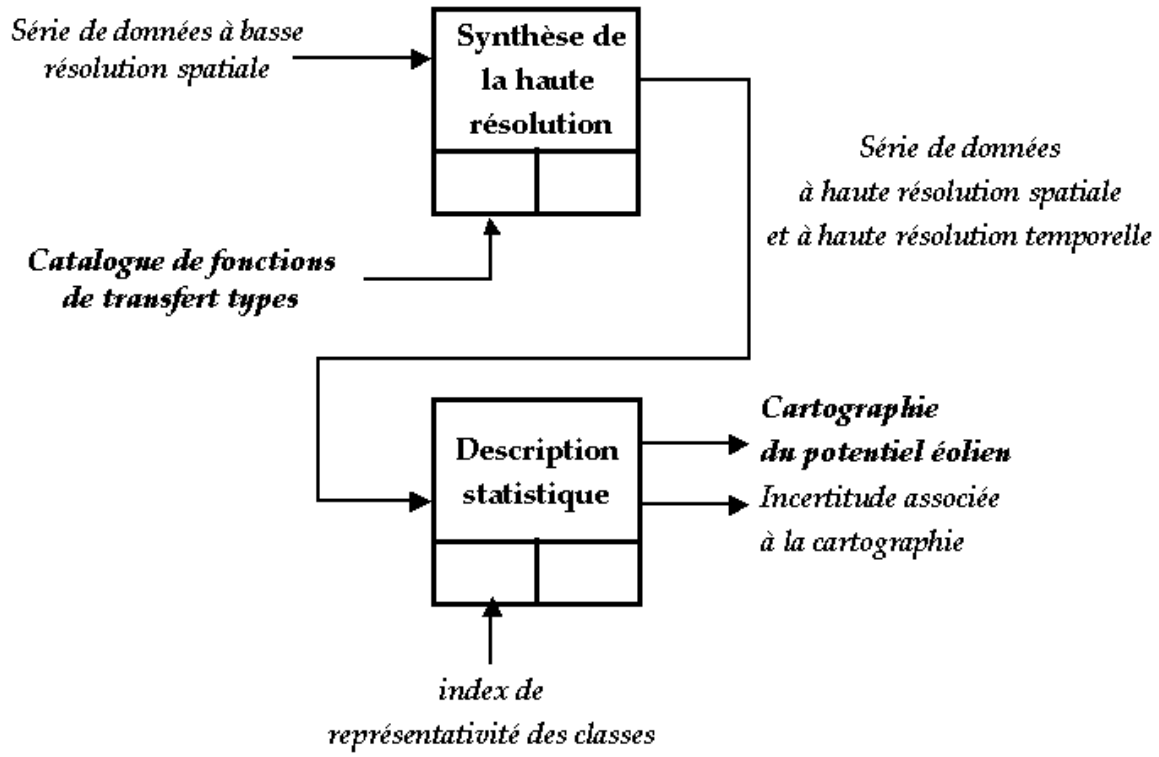


FIG. 3.4 – Deuxième partie du schéma de fusion de données. Cette partie montre l'utilisation des fonctions de transfert types pour la génération de champs de vent synthétiques à haute résolution spatiale et haute résolution temporelle. Ces champs de vent sont utilisés pour la cartographie du potentiel éolien offshore.

champs de vent. Cette étape de classification fait l'objet du chapitre 4. La classification est la fonction qui associe à chaque champ de vent une situation type :

$$\begin{aligned} Classification : \{SCAT_t\}_{t\in T} &\longrightarrow \left\{SCAT_i^{type}\right\}_{i=1..I} \\ SCAT_t &\longmapsto SCAT_i^{type} \end{aligned} \tag{3.1}$$

où $SCAT_t$ est le champ de vent mesuré par un diffusomètre à l'instant t, $SCAT_i^{type}$ est la i^{ème} situation type, T l'ensemble des dates où une mesure diffusométrique est disponible et I le nombre de situations types.

L'objectif de cette étape est de produire un ensemble de classes caractérisant le comportement spatial du vent. La similarité utilisée dans cette étape est donc une similarité du comportement spatial du vent. Pour chaque classe, on doit obtenir des champs de vent ayant les mêmes changements de direction et de vitesse. Les situations types résultant de la classification doivent représenter la climatologie de la région à l'échelle où s'effectue la classification. En utilisant ces situations types avec leur fréquence d'occurrence, on doit arriver à décrire la climatologie du vent d'une manière fidèle à basse résolution spatiale. A chacune de ces situations types est associé un index de représentativité. Cet index de représentativité traduit la fréquence d'occurrence de la situation type.

3.2.2.2 Association

La deuxième étape du schéma de fusion est l'étape d'association. Elle reçoit en entrée les situations types résultant de la classification et la série de données à haute résolution spatiale et basse résolution temporelle. Lors de cette étape, à chaque situation type à basse résolution spatiale doit être associée, au moins, une mesure à haute résolution spatiale. Pour une classe donnée, le champ de vent à haute résolution spatiale associé sert à établir la fonction de transfert type de la classe. Le choix de la fonction de transfert et son établissement sont décrits dans la section 3.3. L'étape d'association s'écrit comme suit :

$$Association : \left\{SCAT_i^{type}\right\}_{i=1..I} \longmapsto \{f_i\}_{i=1..I} \tag{3.2}$$

où $\left\{SCAT_i^{type}\right\}_{i=1..I}$ est l'ensemble des I situations types résultant de la classification des données diffusométriques et $\{f_i\}_{i=1..I}$ l'ensemble des fonctions de transfert types correspondant à ces situations types.

Soit $SCAT_{t_1}$ un champ de vent mesuré à basse résolution spatiale à l'instant t_1 et $SCAT_{t_2}$ un champ de vent à basse résolution spatiale mesuré à l'instant t_2. Soit SAR_{t_2} un champ de vent à haute résolution spatiale mesuré à l'instant t_2. Si $Classification\,(SCAT_{t_1}) = Classification\,(SCAT_{t_2}) = SCAT_i^{type}$ alors f_i, la fonction de transfert associée à la i^{ème} situation type, s'applique pour générer un champ de vent synthétique $SAR_{t_1}^*$ à haute résolution spatiale pour l'instant t_1 :

$$f_i\,(SCAT_{t_1}, SAR_{t_2}) = SAR_{t_1}^* \tag{3.3}$$

L'application de ces fonctions types se fait lors de l'étape décrite dans le paragraphe suivant.

3.2.2.3 Synthèse de la haute résolution spatiale

Une fois le catalogue de fonctions de transfert types établi, il est utilisé pour la synthèse de la haute résolution spatiale à partir de la série de données à basse résolution spatiale comme indiqué sur la figure 3.4. A la sortie de cette étape, on dispose d'une série de champs de vent synthétiques alliant la haute résolution spatiale et la haute résolution temporelle. L'application de la synthèse de la haute résolution spatiale est présentée au chapitre 5.

$$Synthese : \{SCAT_t\}_{t \in T} \times \{f_i\}_{i=1..I} \quad \longrightarrow \quad \{SAR_t^*\}_{t \in T} \tag{3.4}$$

3.2.2.4 Description statistique

L'ensemble de données résultant de la synthèse de la haute résolution spatiale est utilisée lors de l'étape de la description statistique pour établir la cartographie du potentiel éolien de la région étudiée. En pratique, cette cartographie consiste à établir des cartes de la moyenne de la vitesse moyenne du vent des paramètres de Weibull sur la région à haute résolution spatiale. A cette cartographie est associée une incertitude qui est donnée pour chaque paramètre statistique estimé.

Le choix de l'estimateur des paramètres de Weibull a été présenté dans le chapitre 2. Une méthode permettant d'associer des régions de confiances jointes pour les paramètres A et k de Weibull est présentée en annexe B.

3.3 Fonction de transfert de la basse à la haute résolution spatiale

Considérons deux mesures à basse résolution spatiale $SCAT_{t_1}$ et $SCAT_{t_2}$ acquses, respectivement, aux instants t_1 et t_2. Soit SAR_{t_1} un champ de vent à haute résolution spatiale mesuré à l'instant t_1. On cherche une estimation du champ de vent à haute résolution spatiale à l'instant t_2. Si t_1 et t_2 appartiennent à la même situation type $SCAT_i^{type}$, on génère un champ de vent synthétique $SAR_{t_2}^*$ donnant une estimation à haute résolution spatiale du champ de vent à l'instant t_2 :

$$SAR_{t_2}^* \quad = \quad f_i\left(SCAT_{t_2}, SAR_{t_1}\right) \tag{3.5}$$

où f_i est la fonction de transfert type correspondant à la situation type $SCAT_i^{type}$.

Pour établir f_i, on utilise les cas où le problème est résolu, c'est à dire les cas où à la fois une image ROS et des mesures diffusométriques, acquises dans des conditions favorables permettant leur association, sont disponibles. A l'instant t_1 on dispose à la fois d'une mesure du champ de vent à basse résolution spatiale $SCAT_{t_1}$ et d'une mesure à haute résolution spatiale SAR_{t_1}. Idéalement, la fonction de transfert f_i, correspondant à la situation type de l'instant t_1 devrait respecter la condition exprimée par l'équation 3.6. Le champ de vent synthétique $SAR_{t_1}^*$, généré en utilisant f_i, devrait ainsi être exactement égal à celui acquis à haute résolution spatiale SAR_{t_1}.

$$\begin{aligned} SAR_{t_1}^* \quad &= \quad f_i\left(SCAT_{t_1}, SAR_{t_1}\right) \\ &= \quad SAR_{t_1} \end{aligned} \tag{3.6}$$

L'inversion de la relation décrite par l'équation 3.6 ne conduit pas à une solution unique. Une solution peut être la fonction constante donnée dans l'équation 3.7. Cette fonction associe exactement le même champ de vent, à haute résolution spatiale, à tous les instants appartenant à la même situation type.

$$t \overset{f}{\longmapsto} SAR_{t_1}, \forall t \text{ tel que } classe(t) = classe(t_1) \tag{3.7}$$

Le but est de minimiser l'erreur pour l'ensemble des t appartenant à la classe et non seulement pour t_1. La fonction constante, décrite par l'équation 3.7 ne peut pas permettre d'atteindre ce but. Pour l'établissement des fonctions de transfert de la basse à la haute résolution spatiale, on utilise les hypothèses exprimées dans le paragraphe 3.2.1. Cette hypothèse stipule que l'évolution dans les échelles spatiales est similaire pour deux champs de vent appartenant à la même situation type.

Pour la fusion de champs de vent, il est intéressant d'étudier ces champs échelle par échelle. Le concept ARSIS (Amélioration de la Résolution Spatiale par Injection de Structures) [5] pour la fusion d'images définit un cadre qui répond à l'objectif d'une amélioration de la résolution spatiale des données en grilles en étudiant les relations entre l'évolution dans les échelles de deux représentations. Dans la suite, on présente le concept ARSIS et quelques méthodes s'inscrivant dans ce concept. Des exemples de fonctions de transfert sont présentées.

3.3.1 Analyse multi-échelle

L'analyse multi-échelle fournit un cadre formel pour la description hiérarchique de l'information contenue dans un signal. Une analyse multi-échelle, utilisant un algorithme pyramidal est présentée dans la figure 3.5. Dans cette figure, l'analyse multi-échelle est utilisée en combinaison de la transformée en ondelettes. La base de la pyramide représente l'image originale. Les différents plans suivants représentent les approximations successives de l'image. La différence entre deux approximations successives est contenue dans les coefficients d'ondelettes. Plusieurs implémentations de l'analyse multi-échelle existent. Elles varient selon l'utilisation d'ondelettes ou d'autres techniques (pyramide Laplacienne, filtres, ...). Dans le cas où les ondelettes sont utilisées, les propriétés de l'analyse multi-échelle dépendent du type d'ondelettes choisies.

Les transformées en ondelettes décimées sont caractérisées par une perte de la propriété de l'invariance par translation. Ceci conduit à l'apparition d'un nombre important d'artefacts lors de la reconstruction après modification des coefficients d'ondelettes. La transformée en ondelettes continue est caractérisée par une grande redondance et l'impossibilité d'une reconstruction parfaite. La transformée en ondelettes par l' « algorithme à trous » constitue un compromis entre la transformée en ondelettes continue et la transformée en ondelettes décimée [6]. On choisit l' « algorithme à trous » pour la description hiérarchique de l'information contenue dans les champs de vent. La reconstruction du champ de vent à haute résolution spatiale, après le changement de coefficients d'ondelettes, ne souffrira pas ainsi d'un déphasage qui aurait été introduit lors de l'utilisation des ondelettes orthogonales.

Le fonctionnement de l' « algorithme à trous » est le suivant. L'approximation de l'image originale est obtenue en appliquant un filtre passe bas sur cette image. L'image résultante de ce filtrage est l'ap-

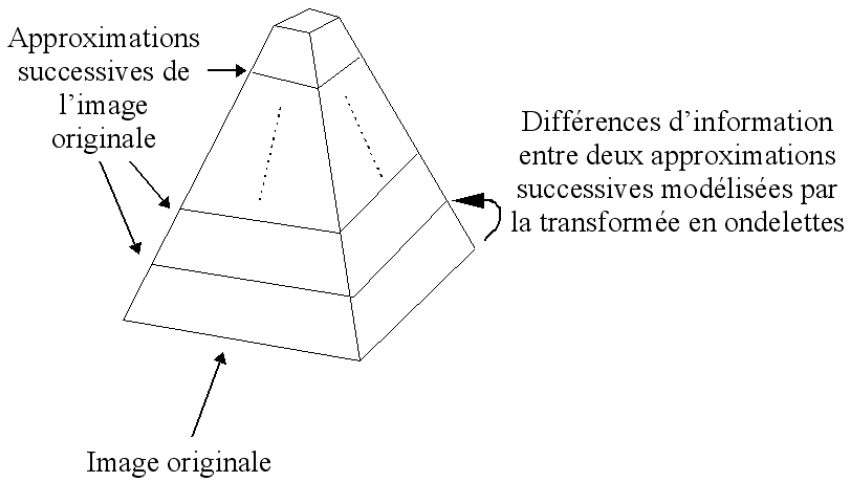


FIG. 3.5 – Représentation d'une analyse multi-échelle utilisant un algorithme pyramidal couplé à la transformée en ondelettes.

proximation de l'image originale. L'image des détails est obtenue par soustraction de l'image originale et de l'approximation. Le changement d'échelle s'effectue par la dilatation du filtre en introduisant de zéros entre les coefficients du filtre (d'où l'appellation de l'algorithme). Ce procédé est itéré pour obtenir les approximations successives de l'image et les images de détails permettant de passer d'une approximation à la suivante.

3.3.2 Concept ARSIS

Le concept ARSIS a été développé dans le cadre de fusion de données en grille [5]. Ce concept fournit un cadre formel pour le développement de méthodes de fusion d'images. Ce concept est utilisé dans plusieurs domaines tels que : l'étude de l'environnement urbain [7, 8], de la géologie [9], de la qualité de l'air [10] et de la climatologie [11].

Considérons deux représentations A et B observant le même phénomène. La représentation A étant disponible à haute résolution spatiale au travers de l'image A_h. B n'étant disponible qu'à basse résolution spatiale au travers de l'image B_l. Le concept ARSIS est basé sur l'hypothèse que les hautes fréquences manquantes dans l'image B_l sont liées à celles présentes dans l'image A_h. Une méthode relevant du concept ARSIS cherche donc à modéliser la relation entre les différentes échelles des représentations A et B. Un telle méthode effectue typiquement les opérations suivantes :

- l'extraction d'un ensemble d'informations de A,
- l'inférence de l'information manquante dans B_l en utilisant l'information extraite,
- la synthèse de l'image B_h^*.

La figure 3.6 présente le schéma général du concept ARSIS. Le modèle multi-échelle MSM effectue une description hiérarchique de l'information spatiale contenue dans l'image. Un exemple de ce modèle est l'analyse multi-résolution [12]. D'autres exemples, dont l'« algorithme à trous » sont donnés dans [13]. Le modèle inter-modalités (IMM) représente le lien entre les structures spatiales des deux

représentations A et B. Le modèle inter-modalités à haute résolution est l'inférence de l'IMM à haute résolution spatiale. Il permet la synthèse des structures à haute résolution spatiale. En utilisant le modèle multi-échelle inverse MSM^{-1}, ces détails sont injectés dans B_l. La fonction de transfert telle que explicitée dans le paragraphe précédent est dans ce cas un combinaison de la HRIMM permettant de synthétiser les structures à haute résolution spatiale et le MSM^{-1} permettant de les injecter dans B_l. Une implémentation du concept ARSIS correspond donc au choix d'un MSM, d'un IMM et d'un HRIMM.

Pour notre application le MSM est l'« algorithme à trous ». Dans la suite on donne quelques exemples d'IMM et de HRIMM.

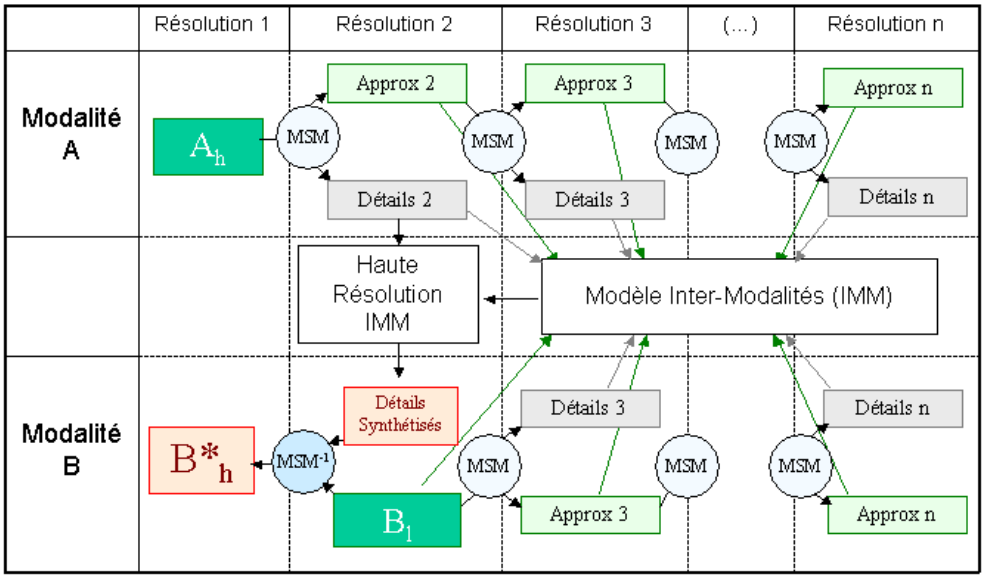


FIG. 3.6 – Schéma général de l'application du concept ARSIS utilisant un modèle multi-échelle (MSM) et un modèle inter-modalités (IMM). D'après [13].

3.3.2.1 Méthode M1

La méthode M1 s'appuie pour l'IMM et HRIMM sur le modèle identité. Les détails présents dans A_h sont considérés comme étant rigoureusement identiques à ceux qui seraient observés si B avait été acquise à haute résolution spatiale. L'image synthétisée B_h^* est donc le résultat de l'injection des structures à haute résolution spatiale présentes dans A_h dans B_l sans traitement préalable de ces structures.

$$\begin{aligned} B_h^* &= f\left(A_h, B_l\right) \\ &= B_l^{interp} + \sum_{i=1}^{n} (B_h^*)^{r_i} \\ \text{avec } (B_h^*)^{r_1} &= (A_h)^{r_1} \\ &\vdots \\ (B_h^*)^{r_n} &= (A_h)^{r_n} \end{aligned} \tag{3.8}$$

3.3.2.2 Méthode M2

Cette méthode cherche à établir une relation affine entre les détails présents dans les images B_l et A_h à partir de la résolution r_{n+1}.

$$(A_h)^{r_k} = a\,(B_l)^{r_k} + b \qquad \text{pour} \qquad k > n \tag{3.9}$$

Dans la méthode M2, le couple (a, b) est choisi de manière à ajuster les moyennes et les variances des champs $(A_h)^{r_k}$ et $(B_l)^{r_k}$, pour les résolution r_k, $k > n$.

3.3.2.3 Méthode M3

Cette méthode est similaire à la méthode M2. Elle est basé sur une relation affine entre les structures aux résolutions présentes dans B_l et A_h. Cependant, le choix de coefficient dans cette méthode est basé sur l'estimation d'une régression linéaire entre les deux représentations et non sur un ajustement des variances. La méthodes des moindres carrés est utilisée pour estimer les cofficients a et b. Il est à noter que seul le coefficient a est modifié par rapport à la méthode M2. Le coefficient b reste inchangé.

3.3.2.4 Méthodes locales

Ces méthodes n'estiment pas une fonction globale pour toute l'image comme c'est le cas pour les méthodes M1, M2 et M3. Elles estiment une relation locale entre les deux images. Cela peut se faire comme dans la méthode RWM (Ranchin-Wald-Mangolini) à l'aide d'une fenêtre glissante. Le recours à cette fenêtre conduit à une image de coefficients et non plus à un couple unique de coefficients (a, b). D'autres méthodes locales existent. Des exemples de ces méthodes sont présentés dans [13].

3.3.2.5 Conclusion sur le choix d'une méthode ARSIS

Les méthodes M2, M3 et les méthodes locales permettent d' « ajuster » la quantité d'information à haute résolution spatiale injectée dans les données à basse résolution spatiale en fonction de l'information déjà présente à basse résolution spatiale. Les méthodes M1 et M2 cherchent une relation globale, sur toute l'image ou le champ de vent dans notre cas, permettant cet ajustement. Les méthodes locales permettent un ajustement variable sur la zone couverte par les mesures. La méthode M1 considère que l'information à haute résolution spatiale est la même dans les deux observations.

Comme première approche, on choisit la méthode M1 pour le transfert de la basse à la haute résolution spatiale. Cette méthode permet de tester l'applicabilité de la méthode de fusion. En outre, elle ne nécessite pas l'application de l'analyse multi-échelles sur les données à basse résolution spatiale. L'application de cette analyse aux données peut être difficile, en pratique, si l'étendue spatiale de la zone d'étude est limitée. Sur de telles zones, les effets de bords dus à l'utilisation des ondelettes introduisent des erreurs qui rendent difficile l'établissement des relations des méthodes M1, M2 et des méthodes locales. L'étendue de la zone géographique de la zone d'étude peut être limitée par la géographie de la zone d'intérêt. En effet, si la zone est située entre deux îles proches, le nombre de mesures diffusométriques est limité. Ceci est dû au fait que les mesures diffusométriques ne sont disponibles qu'à partir de 25

km des côtes. La méthode M1, est quant à elle, applicable dans toutes les configurations géographiques. Elles ne dépend que de la disponibilité de mesures diffusométriques.

Des résultats meilleurs pour les méthodes avancées (M2, M3 ou méthodes locales) ne sont pas exclus, même s'ils doivent être démontrés. Dans cette thèse on choisit de se concentrer sur l'étude de l'applicabilité de la méthode de fusion. On utilise donc la méthode M1, applicable quelque soit la configuration géographique de la zone d'étude. En outre, le choix de la méthode M1, pour le transfert de la basse à la haute résolution spatiale, s'explique dans le contexte de l'établissement de cartes de paramètres statistiques par la considération de situation types. Les détails injectés, s'ils n'appartiennent pas à des cas extrêmes, sont représentatifs du comportement typique.

3.4 Conclusions

On a proposé une méthode pour la cartographie du potentiel éolien offshore. Cette méthode s'inscrit dans le cadre de la fusion de données. Elle permet de tirer avantage de différentes sources d'information sur le vent. Dans cette méthode, les mesures diffusométriques sont utilisées pour déterminer la climatologie du vent à basse résolution spatiale. Les mesures dérivées des images ROS sont utilisées pour inférer les variations à une résolution plus fine de la vitesse du vent. Contrairement à d'autres études qui utilisent les mesures ROS seules [14, 15], cette méthode permet d'avoir une large archive de données synthétiques pour la description statistique de la climatologie du vent. Une archive plus large permet de diminuer l'incertitude liée aux estimations. Il est toutefois nécessaire d'évaluer la qualité les champs de vent synthétiques issus de la fusion de données. Cette évaluation est donnée dans les chapitres 5 et 6.

La synthèse de la haute résolution spatiale relève du concept ARSIS pour la fusion d'images. Les méthodes relevant de ce concept modélisent le lien échelle par échelle entre deux images et utilisent cette modélisation pour augmenter la qualité de l'image à basse résolution spatiale. Différentes méthodes ont été présentées. Comme première approche, on a choisi d'appliquer le modèle identité pour la fonction de transfert de la basse à la haute résolution spatiale. L'application de cette méthode à un cas pratique est présentée au chapitre 5.

Bibliographie

[1] S. Houzelle et G. Giraudon. Contribution to multisensor fusion formalization. *Robotics and Autonomous Systems*, vol. 13, pp. 69–85, 1994.

[2] L. Wald. *Data Fusion : definitions and architectures. Fusion of images of different spatial resolutions*. Ecole des Mines de Paris, Paris, France, 198 p., 2002.

[3] P. Fischer. Turbulence et analyse en paquets d'ondelettes. *Techniques de l'Ingénieur*, vol. AF 1445, pp. 1–18, 2006.

[4] F. Cauneau. *Mécanique des Fluides*. Support de cours, Ecole des Mines de Paris, Paris, France, 165 p., 2006.

[5] T. Ranchin et L. Wald. Fusion of high spatial and spectral resolution images : the ARSIS concept and its implementation. *Photogrammetric Engineering & Remote Sensing*, vol. 66, n° 1, pp. 49–61, 2000.

[6] J.-L. Starck, J. Fadili et F. Murtagh. The undecimated wavelet decomposition and its reconstruction. *IEEE Transactions on Image Processing*, vol. 16, n° 2, pp. 297–309, 2007.

[7] V. S. Raptis, R. A. Vaughan, T. Ranchin et L. Wald. Assessment of different data fusion methods for the classification of an urban environment. In *Proceedings of the second conference F̈usion of Earth data : merging point measurements, raster maps and remotely sensed images¨; Sophia Antipolis, France, January 28-30, 1998, Thierry Ranchin and Lucien Wald Editors, published by SEE/URISCA, Nice, France*, pp. 167–182, 1998.

[8] I. Couloigner, T. Ranchin, V. P. Valtonen et L. Wald. Benefit of the future SPOT 5 and of data fusion to urban mapping. *International Journal of Remote Sensing*, vol. 19, n° 8, pp. 1519–1532, 1998.

[9] W. Yang, F. Cauneau, J.-P. Paris, et T. Ranchin. Fusion of SAR and SPOT images for the detection of geological features over the Three Gorges Dam site, China. In *Proceedings of the third conference F̈usion of Earth data : merging point measurements, raster maps and remotely sensed images¨; Sophia Antipolis, France, January 26-28, 2000, Thierry Ranchin and Lucien Wald Editors, published by SEE/URISCA, Nice, France*, pp. 137–142, 2000.

[10] L. Wald et J. M. Baleynaud. Observing air quality over the city of nantes by means of Landsat thermal infrared data. *International Journal of Remote Sensing*, vol. 20, n° 5, pp. 947–959, 1999.

[11] S. Cros, M. Schroedter-Homscheidt, T. Ranchin et L. Wald. Improvement of operational atmospheric parameters observation : increasing spatial resolution of aerosols optical depth maps by a data fusion process. In *Proceedings of the 4th annual meeting of the European Meteorological Society, 26-30 September 2004, Nice, France*, 2004.

[12] S. G. Mallat. A theory for multiresolution signal decomposition : the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, n° 7, pp. 674–693, 1989.

[13] C. Thomas. *Fusion d'images de résolutions spatiales différentes*. Thèse de doctorat, Ecole des Mines de Paris, France, 165 p., 2006.

[14] M. B. Christiansen, W. Koch, J. Horstmann, C. B. Hasager et M. Nielsen. Wind resource assessment from C-band SAR. *Remote Sensing of Environment*, vol. 105, n° 1, pp. 68–81, 2006.

[15] R. J. Barthelmie et S. C. Pryor. Can Satellite Sampling of Offshore Wind Speeds Realistically Represent Wind Speed Distributions ?. *Journal of Applied Meteorology*, vol. 42, pp. 83–94, 2003.

Chapitre 4

Classification des champs de vent

Première partie : Introduction

La première étape du schéma de fusion de données pour la cartographie de la ressource éolienne offshore décrit dans le chapitre 3 est la classification des champs de vent à basse résolution spatiale. Cette classification a pour but de regrouper les champs de vent caractérisés par un comportement spatial similaire du flux atmosphérique. Les données choisies pour la cartographie de la ressource éolienne offshore sont les données diffusométriques (à basse résolution spatiale et haute résolution temporelle) et les données ROS (à haute résolution spatiale et basse résolution temporelle). Cependant, les mesures diffusométriques en notre possession sont des mesures sous la fauchée. Ces mesures ne sont pas échantillonnées sur une grille régulière. Leur classification est possible après une interpolation. En effet, la classification suppose la comparaison de mesures disponibles aux mêmes points et donc sur une grille constante dans le temps. Pour faciliter l'évaluation de la méthode de classification, on choisit d'appliquer la classification sur des données de réanalyse fournies par le NCEP. Ces données ont une résolution de $1{,}9^\circ \times 1{,}9^\circ$. Une grille de réanalyse est disponible toutes les 6 heures.

Dans ce chapitre, on présente la méthode de classification et son application aux données NCEP. L'évaluation de la classification est entreprise en étudiant la fidélité de la climatologie du vent, à basse résolution spatiale, obtenue en utilisant les situations types, à celle obtenue en utilisant le jeu de données complet. La deuxième partie de ce chapitre est constituée par un article révisé et re-soumis au « Journal of Applied Meteorology and Climatology ». Dans cet article, un état de l'art des méthodes existantes de classification de champs de vent et plus généralement de champs météorologiques est présenté. La diversité des méthodes de classification est le résultat de la diversité des applications. En effet, la notion de similarité et le type de groupes résultant de la classification doivent être adaptés à l'utilisation finale des classes.

Le but de la classification dans notre cas est de regrouper des champs de vent similaires, au sens du comportement spatial du flux. Un nouveau schéma de classification est proposé, basé sur une mesure objective de la similarité entre deux champs de vent (équation 4.1). La classification réduit le jeu de données en quelques situations types représentant le comportement spatial du vent sur la région. Pour l'évaluation de la classification, on étudie la préservation de l'information sur le comportement spatial

par les situations types. Le nombre de situations types est choisi pour être minimal tout en maximisant la similarité des champs de vent regroupés dans la même classe.

Pour son évaluation, la méthode de classification des champs de vent est appliquée aux données de réanalyses NCEP couvrant la mer d'Irlande pendant les années 2001-2004. Les vents dominants sur cette région sont les vents de sud-ouest. Les situations types résultant de cette application sont cohérents avec ces vents dominants. En outre, cette application montre que la majorité de la variabilité sur la région peut être représentée par quelques situations types. L'incidence de l'utilisation de ces quelques situations types pour la description de la climatologie du vent sur la région est étudiée quantitativement. Cette étude montre que le comportement spatial du vent est bien restitué par les quelques situations types.

Deuxième partie : A wind field classification scheme for generation of typical spatial patterns

Mohamed Bassam BEN TICHA and Thierry RANCHIN

Soumis au **« Journal of Applied Meteorology and Climatology »** le 14/02/2007, révisé le 07/09/2007.

Abstract

For wind industry, spatial representation of wind resource is mandatory for optimizing the development of offshore wind farms. In this paper, a new automated and objective wind field classification scheme is proposed. This scheme is designed to generate typical spatial wind patterns. These patterns are intended to represent extensively the wind climatology of the area of interest. The proposed classification scheme consists of a two-stage procedure. The first stage is an agglomerative hierarchical clustering. At the end of this stage the number of classes to be retained is objectively fixed. The second stage of the classification scheme consists of reassigning the rare situations to the appropriate class. The proposed scheme is applied to a case study. The study area covers a part of U.K., Ireland and the Irish Sea. A four-year, 6-h sampled, reanalysis data set is used and the scheme applied. The wind climatology retrieved using these typical patterns is compared to the original wind climatology. Evaluation of the quality shows a good agreement between both climatologies.

4.1 Introduction

In the framework of optimized sitting of offshore wind parks, wind energy industry needs high spatial resolution mapping of wind resource. Wind resource is described by maps of mean wind speed and of probability distribution function of wind speed, by sector, at each point of the map. Generally, direction bins of 30° are used by wind energy industry. This mapping has to be achieved at a kilometer scale to capture wind variations at the scale of a wind farm. Presently, there is no data source providing wind fields at this spatial resolution and with a sufficient temporal sampling for establishing reliable wind statistics. A solution is to associate typical high spatial resolution wind fields to typical large scale wind configurations. Previous studies outlined that local wind is drived by larger scale wind field [1], *i.e.* for similar large-scale wind fields, the local wind field spatial behavior is similar. A classification of large scale wind fields into typical situations is, therefore, needed to group similar wind fields and associate to each typical situation, a typical wind behavior at high spatial resolution. In this paper, we focus on the large-scale classification of wind fields. The classification is evaluated by quantifying the fidelity of the wind statistics computed from the typical situations to the wind statistics computed from the whole data set.

Meteorological fields over a given region vary continuously in time and are rarely exactly the same at two different moments. However, for many meteorological parameters, recurrent spatial patterns are often observed. Classification of meteorological situations aims at retrieving these patterns and at associating each individual situation with the corresponding typical pattern. Initial classifications of me-

teorological situations focused on large-scale weather regimes. An example of these classifications is the traditional “Grosswetterlagen classification” of the German Weather Service [2]. These weather regimes were generated using a subjective classification method. It consists of manual procedures based on visual analysis of synoptic maps. Another well known subjective classification is the “Lamb classification” of British Isles weather types [3]. The increasing of computational resources led to automatic and more specific classifications focusing on particular meteorological fields at finer scales. Of interest in this paper are wind vector fields.

The recurrence of similar situations is used to represent the wind climatology of a region with few typical patterns carrying most of the information about spatial variations of wind over the interest region. [1] grouped subjectively large-scale daily weather patterns and established relationships between the retained typical patterns and local wind flow. This study showed qualitatively that local flow is in a large part related to large-scale meteorological situations. However, for a quantitative evaluation of wind climatology of a site there is a need to a more objective classification method. The need for a quantitative evaluation of wind climatology arises in many applications such wind energy resource assessment, pollutants transport and air quality assessment. But, measurements are, most often, local point measurements that cannot report on the spatial variation of wind over large areas.

Downscaling methods have been used to achieve this mapping. [4] proposed a downscaling procedure based on three stages. Having a time series representative for a larger area, the first task is to group the large amount of data into defined classes. Then, a numerical model performs simulations for each of these classes. Finally, a post-processing program weights and sums the particular simulation results. [5] proposed to merge the high spatial resolution information given by some remote sensing instruments (synthetic aperture radars) and the high temporal resolution given by other remote sensing instruments (scatterometers) to achieve offshore wind resource mapping. These methods have in common a need for a classification of wind fields.

The typical wind situations of interest in wind energy sector are mesoscale wind fields. They have the finer scale with large temporal data sets. For mesoscale wind field classification, [6] developed a method based on principal components analysis (PCA). Classification schemes based on PCA exhibit wind flow patterns having the largest contribution to the temporal variance of wind fields over the studied region. But, the orthogonality of the principal components (both in time and space) have no physical interpretation and can lead to errors when looking for phenomena related to each wind typical pattern. [7] proposed a method based on PCA to reduce the dimensionality of data and, then, perform a cluster analysis based on the first principal components. [8] used PCA combined to a clustering algorithm to establish wind patterns. They study the impact of these patterns on air quality.

In wind energy sector, for the construction of numerical wind atlases, the first classification method was based on the wind rose at a point in the region (or the mean wind speed over the region) [9, 10]. A subdivision of the wind rose at this point is then performed. The result is situations with a constant direction and speed over the whole region. This method does not exhibit the spatial behavior of wind over the region.

An automated classification scheme for wind fields was proposed by [11]. This scheme is based on

the spatial similarity between wind fields. [12] extended this scheme by introducing a two-stage classification. The first stage determines the number of classes to be retained. The second stage determines the members of the classes. They proposed, in addition, to exclude rare situations from the classification to not disturb the final classes.

The scheme proposed in this paper is inspired from this latter classification scheme. This scheme aims at providing typical spatial patterns of wind behavior. Some modifications are done to fit the requirements of wind resource assessment. This scheme is based on a two-stage classification scheme. An objective criterion is taken for the determination of the number of classes. The first stage distinguishes rare situations and frequent typical situations. The processing of rare situations is different than the one proposed by [12]. Instead of excluding these situations, we propose a different strategy allowing associating them to a class. This association is done for practical reasons. The elements of the same class resulting from the classification scheme are intended to be processed similarly for the production of wind resource maps, *i.e.* the same wind field fine-scale structures are associated with all the elements of the class. Keeping the rare situations alone would impose a separate processing for each of these situations. This processing could be tiresome compared to the willed profit. Associating each rare situation to a class will make easier the processing of the results of the classification. A different computation of the representative of each class is proposed to avoid errors that could be introduced by rare situations.

In section 4.2, we present the classification scheme. In section 4.3, we present the application of the classification scheme to a case study covering the Irish Sea. The different wind patterns are interpreted. In section 4.4, evaluation of the results of the classification scheme is undertaken. The evaluation focuses on the representativeness of the retained typical patterns regarding the original temporal time series at each point of the grid and the statistical representation of wind climatology as computed from the original data set. Conclusions of this work are then presented.

4.2 Classification scheme

The classification method aims at defining a limited number of classes or clusters of wind fields such as the elements of the same class are as similar as possible. The notion of similarity between two wind fields is formally expressed using an objective distance measure. Many distances have been used for the classification of wind fields. [13] investigated the quality of some circulation similarity measures. Some distances measure the similarity between fields, *e.g.* correlation coefficient. Other distances measure the dissimilarity of wind fields. We use the distance mentioned by [13] and used by [11]. This distance was used at the U.S. National Centers for Environmental Prediction (NCEP) to characterize forecast errors [13]. [13] evaluated this distance as the best performing when measuring the similarity of two circulation patterns.

In the following, u is the zonal component of wind vector, v is its meridional component and s the wind speed. F_t designates a wind field at time t, *i.e.* wind vectors at different points of a grid. Each point is characterized by its coordinates (i, j). $\{.\}$ indicates a set. The aim of the classification is to obtain a reduced set of wind fields $\{W_k\}_k$ the most representative of the real wind field set $\{F_t\}_t$.

The distance between wind fields F_{t_1} and F_{t_2} is defined as

$$d(F_{t_1}, F_{t_2}) = \frac{1}{nm} \sum_{i=1}^{n} \sum_{j=1}^{m} \left[(\widetilde{u_{t_1,ij}} - \widetilde{u_{t_2,ij}})^2 + (\widetilde{v_{t_1,ij}} - \widetilde{v_{t_2,ij}})^2 \right]^{\frac{1}{2}} \tag{4.1}$$

where n is the latitudinal grid size and m its longitudinal size. $\widetilde{u_{t,ij}}$ and $\widetilde{v_{t,ij}}$ are respectively the scaled zonal and the normalized meridional dimensionless components of wind vector. Equation (4.2) is used to obtain the scaled components $\widetilde{u_{t,ij}}$ and $\widetilde{v_{t,ij}}$ from the initial wind vector components $u_{t,ij}$ and $v_{t,ij}$. The scaling factor s_t is given by Eq. (4.3). s_t is the mean wind speed over the grid points at time t. The use of this scaling factor implies that the distance given in Eq. (4.1) does not separate wind fields by the mean wind speed intensity over the whole area. It separates wind fields regarding the relative wind speeds and directions at the different points of the grid.

$$(\widetilde{u_{t,ij}}, \widetilde{v_{t,ij}}) = \frac{1}{s_t} (u_{t,ij}, v_{t,ij}) \tag{4.2}$$

$$s_t = \frac{1}{nm} \sum_{i=1}^{n} \sum_{j=1}^{m} \left(u_{t,ij}^2 + v_{t,ij}^2 \right)^{\frac{1}{2}} \tag{4.3}$$

Scaling of wind fields is necessary to focus on the relative wind change of speed and direction and not on the intensity of the wind speed. Thus, the similarity measure will focus on the spatial behavior of wind flow over the considered region.

[12] used this similarity measure in the implementation of a two-stage classification procedure of wind fields. The first stage of this procedure consists of an agglomerative hierarchical clustering algorithm. A hierarchical clustering is a sequence of partitions in which each partition is nested into the next partition in the sequence [14]. As a first step, a hierarchical clustering analysis considers each wind field as a one-element cluster. At each step of the analysis, two clusters are merged until only one cluster remains. This algorithm needs neither an *a priori* knowledge of the number of classes nor an *a priori* knowledge of the typical flow patterns. It suits, therefore, the goal of an objective generation of wind patterns. Many different agglomerative hierarchical clustering algorithms exist. These algorithms are based on different manners of merging clusters from one step to another. The most widely used are the single-link clustering, the complete-link clustering, the pair group methods and the Ward's method [14]. Ward's method aims at generating clusters with similar sizes. This constraint is unrealistic in the case of wind fields classification. Typical patterns occur more frequently than others and are representative of the dominant wind over the region. Thus, a method aiming at equalizing the sizes of clusters is not suitable for the generation of typical wind flow patterns. Single-link clustering has the tendency to form chains [14] that do not correspond to the intuitive classification of wind fields. [12] used a complete linkage method. However, complete-link clustering is sensitive to the presence of outliers in data [14]. In addition, the stop criterion used by [12] remains subjective and does not insure a complete description of wind climatology over the region. It gives the most important patterns but does not describe extensively the wind climatology of the studied area. The second stage of the procedure proposed by [12] consists of a k-means classification algorithm. This algorithm aims at reclassifying wind fields using the patterns obtained by the first stage as seeds. A threshold of similarity is fixed. The wind fields that are

more distant than this threshold to all seeds are assigned to a special set of indeterminate wind fields or outliers. This exclusion makes the centroids of the retained classes more realistic than they could be if disturbed by the outliers. This is useful to an interpretation of the resulting patterns. The classification procedure proposed by [12] gives wind field patterns that can be easily associated to meteorological situations. However, it remains inadequate for an automatic processing of the wind fields according to their class membership.

In the following, we present a similar two-stage classification procedure. This procedure is inspired from the procedure proposed by [12]. It aims at generating wind field patterns having a more extensive wind climatology description. A different clustering method is used to avoid the sensitivity of the complete-linkage method to the outliers. The first stage consists of an agglomerative hierarchical clustering algorithm. While the complete link clustering is used by [12], we use the centroid method. This method is less sensitive to outliers. The reported disadvantage of this method is that it generates a large number of small classes containing outliers, *i.e.* wind fields occurring rarely [7]. In our case, this could be seen as an advantage. In fact, these small clusters will contain wind fields that occur rarely and are similar to a low number of wind fields. This could be seen as an automatic exclusion of outliers. The algorithm of the centroid method is detailed in the section 4.2.1. The classification procedure is intended to generate classes whose members could be processed similarly. A great number of small classes, implying separate processing for each class, could be tiresome compared to the expected improvement of the results quality. The second stage of our classification procedure aims at reassigning these rare situations to more significant classes. The representative of a class, *i.e.* the typical pattern, is chosen in a manner such it is not affected by these rare situations. A detailed presentation of this second stage is given, in section 4.2.2.

4.2.1 Clustering method

Using the measure defined in Eq. (4.1), the distance between each pair of wind fields is computed. If N is the number of wind fields, we obtain $\frac{N(N-1)}{2}$ distances between wind fields. These distances form a table of distances between each pair of wind fields. An agglomerative hierarchical clustering method is used as a first stage of the classification process. This method is the unweighted pair group method centroid known also as the centroid method [14]. It consists of an iterative grouping of wind fields. At the first step of the algorithm, each of the N wind fields is considered as a one-element cluster. At each step, the two most similar clusters are grouped. For the centroid method, the similarity of two clusters is defined as the similarity measure of the centroids of the clusters. We use the similarity measure defined in Eq. (4.1). The two most similar clusters are those having the minimal value in the distance table. They define a new cluster. At each step, the number of clusters is reduced by one. This is done until only one-cluster, grouping all data, remains. In the literature, the use of the centroid method is restricted to the use of the Euclidean distance [14]. This is due to the manner of updating the distances table. In our implementation of this method, we recompute the centroids at each step of the algorithm and recompute a new distance table. This table is obtained by computing the distances between each pair of centroids of the remaining clusters. We can, therefore, use the distance defined in Eq. (4.1).

The centroid method results in a tree representing the clustering analysis of the data. This tree is called dendrogram. An example of such a dendrogram is given in Figure 4.1. This illustrative dendrogram is made using ten wind fields. At the top of this tree, there is only one cluster, grouping all the data set. The centroid of this cluster is the mean of all wind fields. Let denote the level at the top of the dendrogram as Level 1.

The dendrogram could be cut at different levels. These levels correspond to the different steps of the clustering algorithm. They are shown in Figure 4.1. At each step of the centroid method the number of clusters is reduced by one. At level 1 we have only one cluster, at level 2 we have 2 clusters, ... Thus the level of cutting of the dendrogram defines the number of clusters resulting from the clustering method. Each cluster is characterized by the number of elements it contains. This number is the representativeness index of the cluster.

The number of clusters is fixed by the level of cutting the dendrogram. To fit the requirement of an objective choice of the number of clusters, we have to fix objectively the level of cutting the tree.

We aim at obtaining classes having significant representativeness indexes, *i.e.* classes having a significant size. At the same time, we aim at separating clearly the different wind fields. Using the centroid method, at each level of cutting of the dendrogram, we obtain a large number of small clusters containing one or two elements and few clusters having a significant number of elements [14]. An objective criteria for the choice of the dendrogram cutting level is the level maximizing the number of the clusters containing a significant number of elements.

In this study, we consider that a significant representativeness is of 1%. A cluster is considered having a significant representativeness if it groups at least 1% of the whole wind fields. For each level of the dendrogram, the number of clusters having a representativeness index greater than 1% have to be computed. The level maximizing this number is retained as the dendrogram cutting level.

The choice of the value of 1% as a threshold of representativeness does not influence the method itself. An additional criterion for the choice of the number of clusters focuses on the different direction sectors represented by the selected clusters. This criterion relaxes the threshold of 1% according to the characteristics of the wind climatology of the studied region. It remains an objective criterion. It is determined using the results of the clustering analysis. This additional criterion is needed to describe the spatial behavior of wind for all direction sectors. This step could be seen as a relaxation step of the 1% threshold where following ranked clusters are retained as final classes until all wind direction sectors are represented. Therefore, at each point of the grid, each wind direction sector should be represented in at least one cluster. As wind energy sector uses direction sectors of 30 degrees, we use this convention in our study. The clusters are ranked according to their representativeness. The remaining clusters are considered as rare situations. The members of these classes are outliers. This process generates many outliers. In fact, rare situations have great distance values to the other wind fields and are therefore put in separate clusters containing few wind fields. In the next step, we propose to attach these rare situations to more recurrent situations.

4.2.2 Reassignment

The centroid method results in a large number of small size clusters. These clusters contain outliers. Outliers, *i.e.* rare situations, are reassigned to the clusters, retained from the first stage of classification, having the most similar representatives. The distance of each wind field, reassigned to a rare situation, to all the centroids of the retained clusters is computed. The wind field is then assigned to the class of the most similar wind field.

The rare situations assigned to a cluster are not included in the computation of the representative of the class. Thus, they will not introduce error on the processing of the class. Most of the errors will concern these situations. Being rare, less error is introduced than the error that would be introduced if these wind fields are introduced in the computing of the centroids.

This reassignment is of interest in application where there is a need to associate a local scale configuration to each wind field.

In next section, we present the application of this scheme to a practical case study. We present the results of each stage of the classification scheme.

4.3 Application to a case study : Irish Sea

The classification scheme described in section 4.2 was applied to reanalysis data over a region covering the Irish Sea. This reanalysis data are from the National Centers for Environmental Prediction (NCEP) and National Center for Atmospheric Research (NCAR). The aim of reanalysis projects is to create consistent data sets covering long periods. The NCEP/NCAR data covers 1948 to present [15]. The data used in this study covers the period from the 01 January 2001 to the 31 December 2004. 10 m height wind speeds and directions are considered in this study. Each day, reanalysis grids are available at 00, 06, 12, and 18 Greenwich Mean Time (GMT). The studied area and the grid points used in the classification are represented in Figure 4.2. This area is characterized by dominant south-westerly winds [9].

4.3.1 Clustering and selection of the number of classes

The distances between each pair of wind fields are computed as indicated in section 4.2. The dendrogram is computed. In order to choose the appropriate level of cutting the dendrogram, we apply the method described in section 4.2.1. For each level of cutting, we compute the number of clusters having a representativeness rate greater than 1%. They are presented in Figure 4.3. We obtain a maximum number of clusters having a representativeness of 1% when choosing to cut the dendrogram at the level 826. From these 826 clusters, only 22 clusters have a representativeness index greater than 1%.

These 22 clusters do not represent all wind direction sectors. Therefore, we have to add less representative clusters until representing all wind direction sectors. This task is equivalent to a relaxation of the constraint on the representativeness index as proposed in section 4.2.2.

When all wind direction sectors are represented, for each point, we reach a total number of 36 clusters. The minimal representativeness rate of these clusters is of 0.3%. The representativeness index

of 1% have been relaxed to 0.3% using the constraint of representing all wind directions. These 36 clusters have a total representativeness rate of 72%, *i.e.* 72% of the wind fields belong to a cluster.

4.3.2 Reassignment of the rare situations

In Table 4.1 are shown the representativeness rates of the 36 clusters before the reassignment and their representativeness rate after the reassignment of the rare situations. This denotes a great disparity of the number of rare situations assigned to each cluster.

The effect of these rare situations on each class will be assessed in section 4.4. The great number of rare situations, representing 27.8% of the total wind fields, are characterized by low wind speeds. This could be explained by less correlation between different grid points at low speed wind fields. These situations correspond usually to transition situations between two stable situations.

4.3.3 Final results

The classification process generated 36 classes. The 36 representatives of these classes, associated to the occurrence rates of the corresponding classes, are given in Figure 4.4. The representative of the class is the mean wind field of class members without the reassigned rare cases.

As expected and in accordance with [9], the most frequent classes are those characterized by a south-westerly wind, the dominant wind over this region. For each point of the grid, all wind directions are represented in these classes.

To help in the analysis of the results, we retained two points in our study : an inland point and an offshore point. The inland point is the point located at the longitude 7.5 W and the latitude 54.3 N. In the following, we will refer to this point as the point (1,1). The second point is the central point of the grid of Figure 4.2. It is located at the longitude 5.6 W and the latitude 52.4 N. In the following, we will refer to this point as the point (2,2). The mean wind speed and direction of each class, for these two points, and the corresponding standard deviation are given in Table 4.2.

To characterize the different classes, we computed the mean wind direction and speed of each class and for each point of the grid. These mean values are slightly different from those of the representative situations as the outliers were not included in the computation of the representative of the classes. They are included in the computation of the mean values. We computed, in addition, the standard deviation of the speed and the direction within the class for each point of the grid. In Table 4.2, results are only presented for Point (1,1) and Point (2,2).

In Table 4.2, for each point, the high mean wind speeds tends to be associated with low wind direction standard deviations. This means that the classification scheme adapts the direction bins to speed magnitude, *i.e.* the higher is the wind speed, the smaller is the wind direction standard deviation. The second property of interest is the relationship between the different points of the grid when classified. It is useful to know the importance of each point in the classification. Although the results on only two points are presented here, the conclusions are the same for the other points. Classes with higher wind speed on point (1,1) than on point (2,2) are classes 7, 8, 10, 12, 14, 17 and 20. Classes 7, 8 and 10 are characterized by higher wind direction standard deviation on point (2,2) than on point (1,1). This is in

accordance with the tendency described previously. For classes 12, 14, 17 and 20 the direction standard deviation on point (1,1) within each class is slightly higher than the direction standard deviation on point (2,2). No class with higher wind speed on point (2,2) than on point (1,1) is characterized by higher wind direction standard deviation on point (2,2). The higher direction standard deviation on point (1,1) is due to an implicit greater weight of the point (2,2) in the classification. Thus, the precision of the wind direction bins on point (2,2) is higher. In fact, point (2,2) being the central point, its correlation to the other points of the grid is greater than the one of point (1,1). Implicitly, the weight of the central point is greater than the weight of any of the other points of the grid.

The classification scheme gives more importance to the central point of the grid. This is of interest when focusing on the central region of the grid. A weighting of the different points could be done to give different importance to the different points of the grid. The most representative wind patterns generated by the classification scheme correspond to the dominant wind directions over the area.

In the following section, we quantitatively evaluate the results of the classification. This evaluation is based on the comparison of parameters, describing wind climatology over the studied area, computed using the wind patterns generated by the classification scheme and the original data set.

4.4 Evaluation

The classification scheme aims at generating typical spatial patterns. Therefore, the evaluation focuses on the fidelity of the typical spatial behaviors retained by the classification scheme to the actual spatial behavior of wind over the region described by the complete data set, *i.e.* on the description of wind climatology at each grid point.

As described in section 4.2 the typical wind fields F_t are classified after a scaling by the factor s_t given in Eq. (4.3). This scaling is performed to focus on the spatial behavior and the relative change of wind speed and direction. It permits to give less importance to the mean wind speed intensity over the area. s_t is the mean wind speed over the region of interest.

The typical patterns, obtained by the classification scheme, represent the ratio of the wind speed at each point of the grid and the mean wind speed over the area. They report also the deviation of the wind direction at each point from the mean wind direction over the area. Knowing the class, represented by the typical pattern W_k, to which a wind field belongs and the mean wind speed over the region s_t, a simple rescaling of the typical pattern W_k by s_t is expected to give the information about the contribution of each grid point to this mean and the different wind direction changes.

For tevaluating the classification, we generate a synthetic data set based on the typical patterns set $\{W_k\}_k$ and the time series of wind speed spatial mean s_t over the whole area. Lets denote F_t^* the synthetic wind field obtained by the scaling of W_k by s_t :

$$F_t^* = s_t \times W_k \tag{4.4}$$

where W_k is the typical pattern of the wind field class at time t. The typical wind pattern W_k is considered representative of the spatial behavior of wind field at time t if $\| F_t^* - F_t \|$ is minimal.

Using the synthetic wind fields set $\{F_t^*\}_t$, we evaluate the classification regarding two criteria. The first criterion focuses on the error on each time step t. This is performed by comparing at each time step F_t^* and F_t. We compute the bias and the standard deviation of the instantaneous error series.

This error is the difference between the wind speeds at each time t. The error series on speed and direction follows a normal distribution for all grid points. As shown in Table 4.3, the biases on the wind speed and the wind direction are both negligible for all grid points. The synthetic time series could, therefore, be considered as unbiased. Table 4.3 shows a good agreement between the initial time series and the synthetic ones. The standard deviation of the error on the synthetic wind speed varies between 0.8 m s^{-1} and 1.5 m s^{-1}. Compared to the performed reduction of data size, this error distribution can be considered as satisfactory. The standard deviation of the error on wind direction varies between 14 degrees and 32 degrees. We used direction sectors of 30 degrees to choose the number of the classes to be retained. This is the conventional direction sector extension used in wind energy sector [16]. The error standard deviation of 14 degrees present in Point (2,2) implies that, for most cases, the direction sector given by the synthetic wind field F_t^* is the same as the direction sector given by the original field F_t. The error is minimal at the central point of the grid, Point (2,2), for both of speed and direction. This confirms the conclusions of the previous section. The central point of the grid had indeed a greater importance than the other points.

The second error criterion is the error on the statistical parameters used to describe the wind climatology of a region. Usually, the two-parameter Weibull PDF is used to describe the wind climatology of a site [17]. The Weibull PDF gives the probability of occurrence of the different wind speeds at a given location. This function is given in Eq. (4.5). It is characterized by its scale parameter A and its shape parameter k. A is analogue to the first moment and k is analogue to a second moment of the PDF.

$$p(v) = \frac{k}{A}\left(\frac{s}{A}\right)^{k-1} e^{-\left(\frac{s}{A}\right)^k} \tag{4.5}$$

The mean wind speed and the parameters A and k estimated from the original data set and from the synthetic one are given in the Table 4.4. This table shows, in general, an underestimation of the mean wind speed and the scale parameter A by using the synthetic data set $\{F_t^*\}$. This could be explained by an underestimation of the very high speeds in the synthetic wind fields. If a point of the grid, at time t, has a very high speed compared to the other points, this will not be reported in the synthetic wind field. This is due to the fact that any of the typical patterns of Figure 4.4 shows a very large deviation of the wind speed at a point of the grid from the mean wind speed over the region. The only point showing a slight overestimation of the parameter A is the central point. This overestimation of 0.8 % could be considered as negligible. Table 4.4 shows that the shape parameter k is overestimated for all grid points. This parameter indicates the dispersion of the wind speeds around a certain value. Using few typical wind field patterns tightens the wind speeds around a certain value. Thus, the shape parameter k increases. The overestimation of k has a maximum value of 15.2 % and a value of 10.6 % at the central point. Such errors on the parameter k do not imply great errors on the wind climatology description. They show that the extreme events are not well reported. This is the case when we reduce the data set dimension and report the information by few typical cases.

4.5 Conclusions

This study proposes a new classification scheme designed to generate typical situations representing wind climatology over a given area and evaluates its quality compared to the original data set. We applied this classification scheme to a four-year reanalysis data series covering Ireland, the Irish sea and a part of the UK. The typical patterns generated by this classification scheme are in agreement with the wind climatology in the different directions originally evaluated by [9].

To assess the validity of these typical situations as representatives of wind climatology of the area, we compared wind climatology description based on the original data set to wind climatology description based on the typical patterns. The typical patterns were used in this validation as indicators of the spatial behavior of wind over the area.

It was shown that the 36 typical situations, obtained from classification, report on most of the information about the spatial variations of wind field over the area. The study of the 36 classes enhanced the greater weight given to the central point. Additionally, this latter has a small standard deviation of the direction within the same class.

The 36 typical situations, describing the time series of the mean wind speed over the area, allow a good representation of wind climatology at the different grid points. Synthetic time series of wind speed and direction were generated using the 36 typical patterns and the time series of the mean wind speed over the area. These time series were compared to the original time series. The quality of the results was quantitatively assessed using:

- the quality at each time step,
- the statistical quality of the whole time series.

This study showed a good agreement between the initial times series and the synthetic ones. The error, *i.e.* the difference between the synthetic and the original time series, is characterized by a standard deviation of 0.8 m s^{-1} on the speed and 14 deg. at the grid central point. It was shown that the speed and the direction synthetic time series are both unbiased.

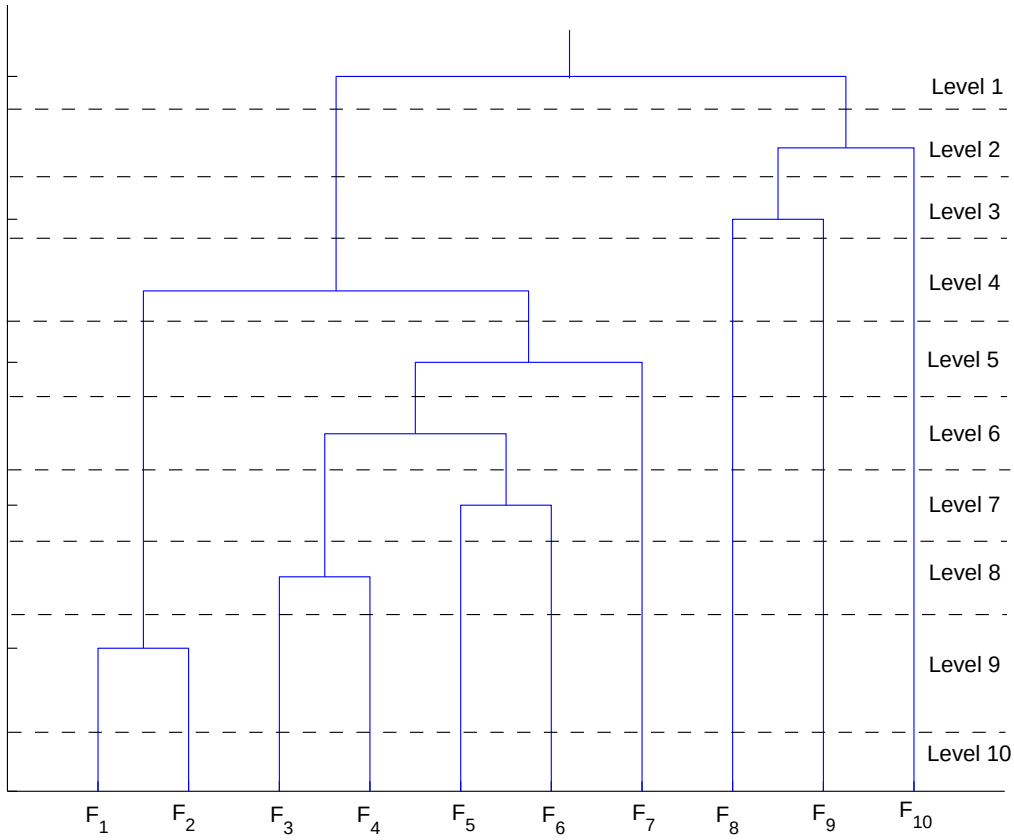


Figure 4.1: An example of a dendrogram resulting from a hierarchical clustering analysis of 10 wind fields.

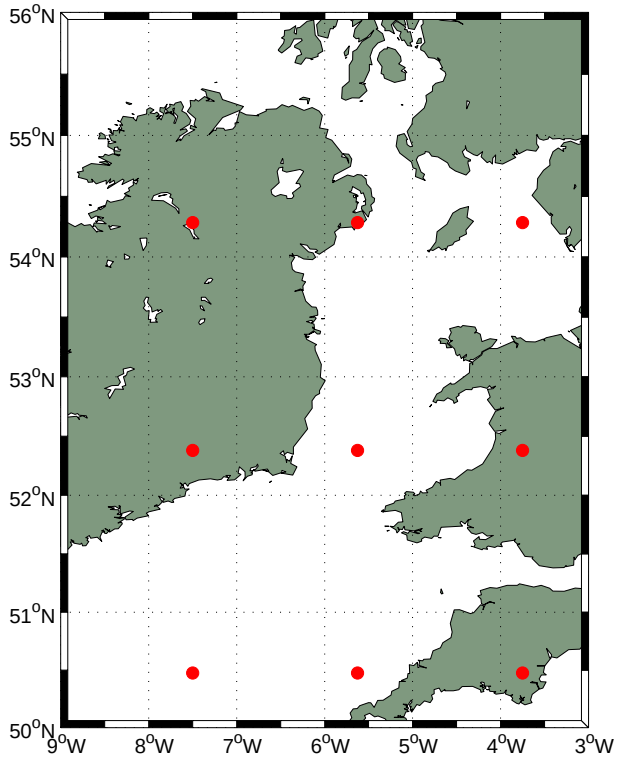


Figure 4.2: The studied area. The 9 points are the grid points used in this study.

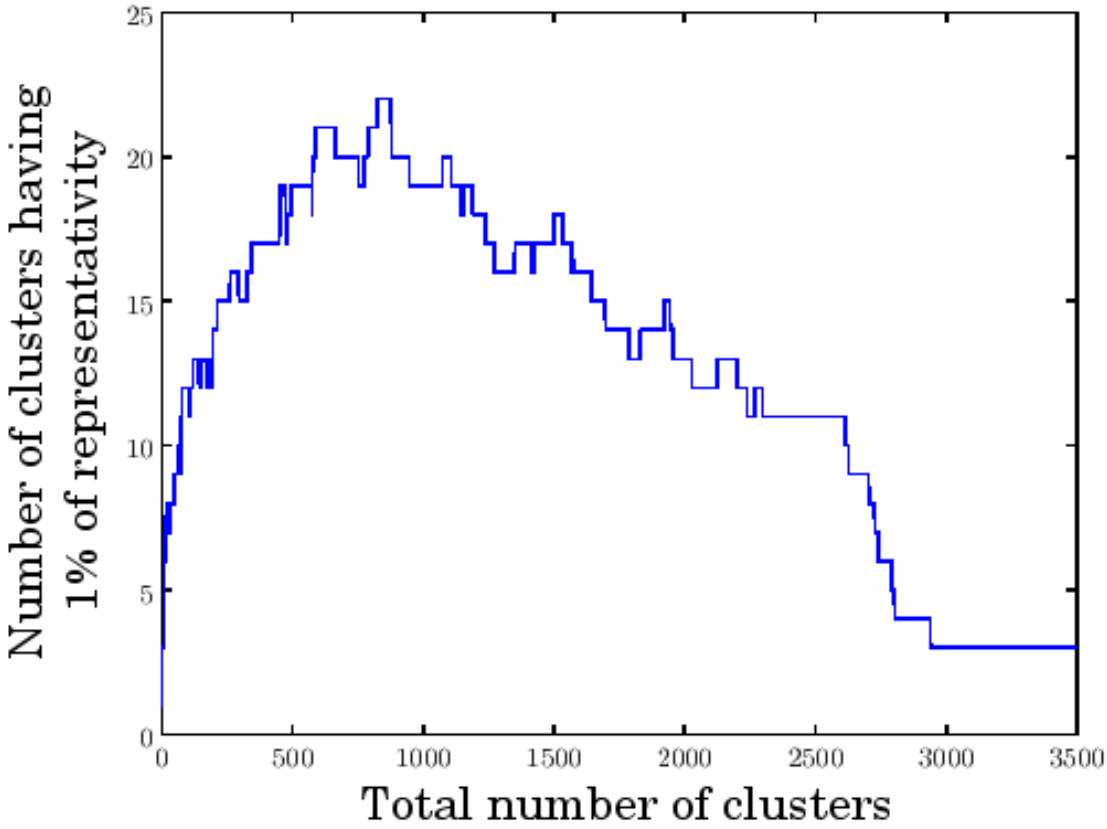


Figure 4.3: The number of clusters having a representativity greater than 1 % in function of the total number of clusters.

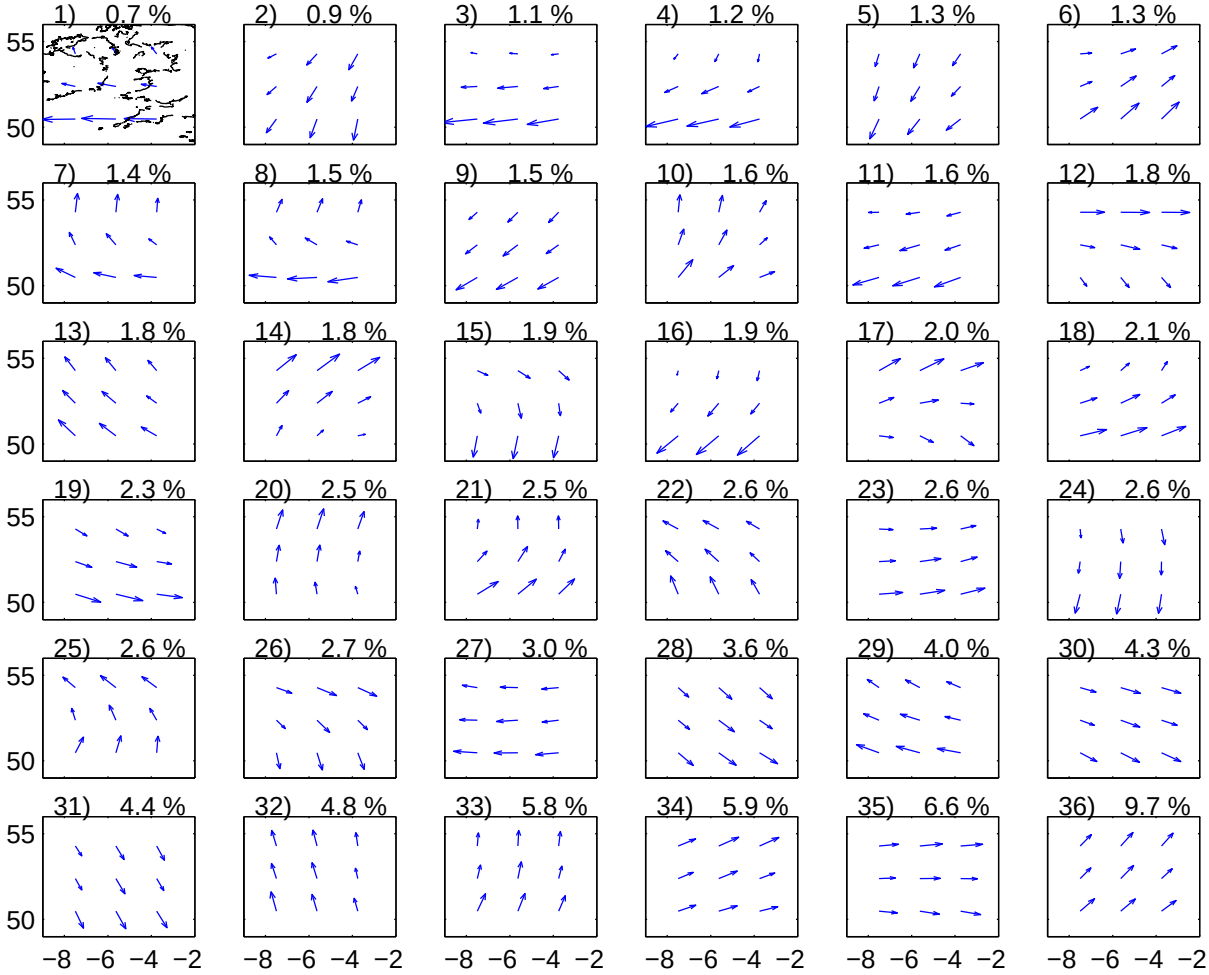


Figure 4.4: The representatives of the 36 typical situations with their occurrence rates.

Bibliography

[1] R. Guardans and I. Palomino. Description of wind field dynamic patterns in a valley and their relation to mesoscale and synoptic-scale meteorological situations. *Journal of Applied Meteorology*, vol. 34, n° 1, pp. 49–67, January 1995.

[2] F. Baur, P. Hess and H. Nagel. *Kalender der Grosswetterlagen Europas 1881-1939*. Bad Homburg, 1944.

[3] H. H. Lamb. *British Isles Weather types and a register of daily sequence of circulation patterns, 1861-1971*. Geophysical Memoir, Vol. 116, HMSO, London, United Kingdom, 1972.

[4] H.-T. Mengelkamp. Wind climate simulation over complex terrain and wind turbine energy output estimation. *Theoretical and Applied Climatology*, vol. 63, n° 3-4, pp. 129–139, 1999.

[5] M. B. Ben Ticha and T. Ranchin. A case based reasoning data fusion scheme: application to offshore wind energy resource mapping. In *The 9th International conference on information fusion*, Florence, Italy, 2006. paper # 206, IEEE Catalog 06EX1311C.

[6] D. M. Hardy and J. J. Walton. Principal components analysis of vector wind measurements. *Journal of Applied Meteorology*, vol. 17, pp. 1153–1162, aug 1978.

[7] L. S. Kalkstein, G. Tanb and J. A. Skindlov. An evaluation of three clustering procedures for use in synoptic climatological classification. *Journal of Applied Meteorology*, vol. 26, pp. 717–730, juin 1987.

[8] S. Beaver and A. Palazoglu. Cluster analysis of hourly wind measurements to reveal synoptic regimes affecting air quality. *Journal of Applied Meteorology and Climatology*, vol. 45, pp. 1710–1726, 2006.

[9] H. P. Frank and L. Landberg. Modelling the wind climate of ireland. *Boundary-Layer Meteorology*, vol. 85, pp. 359–378, 1997.

[10] N. J. Cutler. Class generation for numerical wind atlases. Technical report, Riso National Laboratory, Wind Energy Department, Roskilde, Denmark, 212 p., 2005.

[11] R. O. Weber and P. Kaufmann. Automated classification scheme for wind fields. *Journal of Applied Meteorology*, vol. 34, pp. 1133–1142, May 1995.

[12] P. Kaufmann and R. O. Weber. Classification of mesoscale wind fields in the MISTRAL field experiment. *Journal of Applied Meteorology*, vol. 35, pp. 1963–1979, novembre 1996.

[13] Z. Toth. Intercomparison of circulation similarity measures. *Monthly Weather Review*, vol. 119, n° 1, pp. 55–64, January 1991.

[14] A. K. Jain and R. C. Dubes. *Algorithms for clustering data*. Prentice Hall, New Jersey, 1988.

[15] R. Kistler, E. Kalnay, W. Collins, S. Saha, G. White, J. Woollen, M. Chelliah, W. Ebisuzaki, M. Kanamitsu, V. Kousky, H. van den Dool, R. Jenne and M. Fiorino. The NCEP-NCAR 50-Year Reanalysis: Monthly Means CD-ROM and Documentation. *Bulletin of the American Meteorological Society*, vol. 82, pp. 247–268, février 2001.

[16] I. Troen and E. L. Petersen. *European Wind Atlas*. Risoe National Laboratory, Roskilde, Danemark, 1989.

[17] E. G. Pavia and J. J. O'Brien. Weibull statistics of wind speed over the ocean. *Journal of Applied Meteorology*, vol. 25, pp. 1324–1332, octobre 1986.

Table 4.1: representativeness index of each class, before and after the reassignment of rare situations.

	representativeness rate	
Class	after reassignment	before reassignment
1	0.7 %	0.4 %
2	0.9 %	0.3 %
3	1.1 %	0.8 %
4	1.2 %	0.7 %
5	1.3 %	0.8 %
6	1.3 %	0.5 %
7	1.4 %	0.5 %
8	1.5 %	0.3 %
9	1.5 %	1.0 %
10	1.6 %	0.9 %
11	1.6 %	1.0 %
12	1.8 %	1.1 %
13	1.8 %	1.2 %
14	1.8 %	1.0 %
15	1.9 %	0.9 %
16	1.9 %	0.8 %
17	2.0 %	0.8 %
18	2.1 %	1.2 %
19	2.3 %	1.0 %
20	2.5 %	1.5 %
21	2.5 %	1.1 %
22	2.6 %	1.7 %
23	2.6 %	2.0 %
24	2.6 %	2.3 %
25	2.6 %	1.2 %
26	2.7 %	1.7 %
27	3.0 %	2.1 %
28	3.6 %	2.6 %
29	4.0 %	2.9 %
30	4.3 %	3.9 %
31	4.4 %	3.4 %
32	4.8 %	4.3 %
33	5.8 %	5.3 %
34	5.9 %	5.5 %
35	6.6 %	6.0 %
36	9.7 %	9.4 %
Total	100.0 %	72.2 %

Table 4.2: For two points of the grid; mean wind speed (in m s^{-1} and direction (in °) in each wind class/ Std deviation within the class.

	Point (1,1)		Point (2,2)	
Class	Mean speed ± Std dev	Mean direction ± Std dev	Mean speed ± Std dev	Mean direction ± Std dev
1	1.9 ± 0.9	182 ± 59	3.7 ± 2.1	101 ± 19
2	3.3 ± 1.9	72 ± 43	5.0 ± 2.5	26 ± 17
3	2.0 ± 1.0	121 ± 59	4.8 ± 1.9	81 ± 16
4	1.9 ± 0.9	17 ± 63	3.8 ± 1.9	57 ± 15
5	5.5 ± 2.4	16 ± 19	6.3 ± 3.1	26 ± 12
6	3.8 ± 2.4	300 ± 66	5.8 ± 2.7	227 ± 21
7	4.9 ± 1.8	191 ± 18	3.2 ± 1.8	151 ± 27
8	3.2 ± 1.5	211 ± 31	2.0 ± 1.3	132 ± 48
9	4.5 ± 1.9	48 ± 25	6.8 ± 3.1	49 ± 12
10	6.1 ± 2.6	179 ± 17	5.1 ± 2.8	204 ± 21
11	3.9 ± 1.7	100 ± 30	7.2 ± 3.2	68 ± 10
12	5.4 ± 2.0	270 ± 17	4.3 ± 1.9	285 ± 15
13	6.2 ± 3.0	153 ± 17	6.3 ± 2.9	136 ± 9
14	6.1 ± 2.2	225 ± 15	4.2 ± 1.8	224 ± 12
15	3.2 ± 1.5	281 ± 41	4.3 ± 2.1	349 ± 17
16	2.0 ± 1.3	340 ± 81	4.1 ± 2.1	34 ± 18
17	5.5 ± 2.1	227 ± 17	3.5 ± 1.8	259 ± 16
18	4.3 ± 2.6	215 ± 81	7.2 ± 4.1	238 ± 30
19	4.4 ± 2.4	327 ± 49	6.0 ± 3.2	294 ± 19
20	6.2 ± 2.4	198 ± 15	5.0 ± 2.3	188 ± 11
21	4.3 ± 2.2	162 ± 53	7.2 ± 3.6	207 ± 13
22	7.0 ± 3.2	113 ± 24	7.7 ± 3.6	136 ± 15
23	5.2 ± 2.7	284 ± 37	7.9 ± 3.6	261 ± 10
24	4.2 ± 2.5	352 ± 33	7.1 ± 2.8	3 ± 11
25	6.5 ± 2.9	122 ± 39	6.6 ± 3.6	159 ± 19
26	4.4 ± 1.9	283 ± 31	4.7 ± 1.9	320 ± 12
27	5.7 ± 2.4	95 ± 23	7.2 ± 3.3	85 ± 11
28	5.4 ± 2.5	322 ± 28	7.0 ± 2.9	315 ± 11
29	5.7 ± 2.5	130 ± 19	7.4 ± 3.0	111 ± 9
30	6.6 ± 3.0	288 ± 21	7.9 ± 3.3	295 ± 8
31	4.9 ± 2.4	326 ± 31	6.9 ± 2.9	337 ± 11
32	7.4 ± 2.9	166 ± 14	7.8 ± 2.9	166 ± 11
33	7.3 ± 3.2	183 ± 22	9.3 ± 3.2	189 ± 9
34	7.5 ± 2.9	243 ± 14	8.0 ± 3.1	243 ± 7
35	7.6 ± 3.1	264 ± 17	8.0 ± 3.2	270 ± 9
36	7.6 ± 3.0	219 ± 14	8.8 ± 3.1	218 ± 9

Table 4.3: Errors on the synthetic temporal series.

	Wind speed		Wind direction	
	Bias (m s^{-1})	Std. dev. (m s^{-1})	Bias (°)	Std. dev. (°)
Point (1,1)	-0.2	1.4	1	32
Point (1,2)	-0.2	1.2	0	27
Point (1,3)	-0.3	1.4	0	30
Point (2,1)	0.0	1.0	0	23
Point (2,2)	0.0	0.8	0	14
Point (2,3)	-0.1	1.0	-1	22
Point (3,1)	-0.2	1.5	0	27
Point (3,2)	-0.2	1.2	0	23
Point (3,3)	-0.3	1.5	-1	27

Table 4.4: Statistical parameters and errors on these parameters induced by the use of the synthetic data set. A and the mean wind speed are m s^{-1}. k is a dimensionless parameter.

	Original data set			Synthetic data set			Error		
	Mean	A	k	Mean	A	k	Mean	A	k
Point (1,1)	5.8	6.54	1.99	5.6	6.29	2.11	-4.2 %	-3.9 %	6.1 %
Point (1,2)	6.7	7.56	2.08	6.5	7.37	2.19	-2.7 %	-2.4 %	5.2 %
Point (1,3)	6.5	7.30	2.04	6.2	6.99	2.18	-4.5 %	-4.2 %	6.6 %
Point (2,1)	5.6	6.33	1.98	5.6	6.32	2.28	-0.4 %	-0.1 %	15.2 %
Point (2,2)	6.8	7.73	2.11	6.9	7.79	2.33	0.6 %	0.8 %	10.6 %
Point (2,3)	5.4	6.08	1.97	5.3	6.01	2.26	-1.6 %	-1.2 %	15.1 %
Point (3,1)	7.8	8.81	2.19	7.6	8.56	2.31	-3.1 %	-2.9 %	5.7 %
Point (3,2)	7.8	8.80	2.17	7.6	8.61	2.30	-2.3 %	-2.1 %	6.0 %
Point (3,3)	7.5	8.49	2.14	7.2	8.12	2.27	-4.4 %	-4.3 %	5.8 %

Chapitre 5

Synthèse de la haute résolution spatiale

Première partie : Introduction

Dans ce chapitre, on présente la synthèse de la haute résolution spatiale (voir figure 3.4, étape synthèse de la haute résolution). Cette synthèse s'effectue à l'aide de fonctions de transfert types. Pour chaque classe établie dans le chapitre 4, on associe une fonction de transfert type (voir figure 3.3, étape d'association). Cette fonction permet de générer des champs de vent à haute résolution spatiale à partir des champs à basse résolution spatiale appartenant à la situation type à laquelle est associée cette fonction. Elle est établie à l'aide d'un champ de vent à haute résolution spatiale correspondant à la classe. Chaque champ de vent acquis à haute résolution spatiale est associé au champ de vent à basse résolution spatiale mesuré à la date la plus proche. Pour l'exemple considéré, une mesure diffusométrique et disponible toutes les 12 heures. La différence temporelle entre les mesures à haute et à basse résolution spatiale n'excède pas les 6 heures. Le champ de vent à haute résolution spatiale est attribué à la même classe que le champ à basse résolution spatiale associé.

Après avoir associé un champ de vent à haute résolution spatiale à une situation type, la synthèse de la haute résolution spatiale peut être appliquée aux champs de vent à basse résolution spatiale appartenant à cette situation type. Pour générer les champs de vent à haute résolution spatiale, une fonction de transfert de la basse à la haute résolution spatiale est utilisée. En se basant sur des méthodes relevant du concept ARSIS, présenté au chapitre 3, on utilise une représentation hiérarchique de l'information permettant d'extraire les détails présents dans le champ de vent à haute résolution spatiale et de les injecter dans les champs de vent à basse résolution spatiale de la classe considérée.

Dans ce chapitre, on présente l'établissement de ces fonctions de transfert type et des exemples d'application de ces fonctions. La deuxième partie de ce chapitre est constituée par un article révisé et resoumis au journal « IEEE Transactions on Geoscience and Remote Sensing ». Dans cet article, on applique la synthèse de la haute résolution spatiale à quelques cas où un champ de vent référence existe. Ce champ de vent référence sert à évaluer quantitativement la qualité des champs de vent synthétisés. L'évaluation de la qualité est entreprise dans des cas favorables à la fusion des données diffusométriques et des données ROS ainsi que dans des cas défavorables. Les cas favorables sont les cas où les situations types de vent sont stables. Comme attendu, la qualité des champs de vent fusionnés est moindre dans les

cas défavorables. Toutefois, dans tous les cas, la qualité des champs de vent fusionnés est satisfaisante pour l'estimation de la ressource éolienne offshore.

Deuxième partie : Fusion of SAR images and scatterometer data for wind resource assessment

Mohamed Bassam BEN TICHA and Thierry RANCHIN

Soumis au **« IEEE Transactions on Geoscience and Remote Sensing »** le 29/05/2007, révisé le 01/09/2007.

Abstract

High spatial resolution images provided by spaceborne synthetic aperture radars (SAR) are of great interest for monitoring wind fields offshore. However, the temporal sampling of these measurements is not sufficient to describe the wind climatology of a site. Scatterometer data have a lower spatial resolution but have a sufficient temporal sampling to describe accurately the wind climatology of a site. In this paper, we use an image fusion method based on the ARSIS concept (from its French name "Amélioration de la Résolution Spatiale par Injection de Structures") [1] to benefit from high spatial resolution of SAR images and temporal sampling of scatterometer data. The proposed method is applied to some case studies. The results are quantitively assessed.

5.1 Introduction

In the framework of offshore wind energy development, wind resource mapping is essential. An accurate wind resource knowledge allows a better sitting of wind turbines and thus improve their energy yield. Wind resource is the quantity of available wind energy in this area. This resource is related to the wind climatology of the area. Wind climatology is described by maps of mean wind speed and probability distribution function of wind speed at each point of the map. The probability distribution function suiting the temporal variations of wind and used by the wind industry is the two-parameter Weibull distribution [2].

A kilometre-scale mapping of offshore wind resource is needed by offshore wind industry. This spatial resolution allows description of the behaviour of wind at the scale of a wind park. For an accurate estimation of wind climatology, high temporal resolution data is needed to capture temporal wind variations.[3] showed that a temporal resolution of 50 hours allows assessing wind climatology of a site with a sufficient precision for offshore wind industry. Offshore arises the problem of lack of *in situ* measurements and the cost of installing measurement devices. Fortunately, some spaceborne remote sensing instruments provide relevant information on offshore wind. Of particular interest are scatterometers [4] and synthetic aperture radars (SAR) [5]. The measurements acquired by these instruments are used to retrieve 10-m-height wind data through the use of CMOD algorithms [4].

Since 1991, there have been continuously spaceborne scatterometers observing the Earth:

- AMI on board ERS-1 from August 1991 until June 1996;
- AMI on board ERS-2 from March 1996 until January 2001;

- NSCAT on board ADEOS from September 1996 until June 1997;
- SeaWinds on board ADEOS-2 from December 2002 until October 2003;
- SeaWinds on board QuikScat from July 1999 until now.

CMOD models (CMOD-4 [6], CMOD-Ifremer [7], CMOD-5 [8]) are used to retrieve jointly wind speed and direction from backscatters. These models were empirically established by linking collocated scatterometer data and *in situ* data or model wind fields. Thanks to a large swath, scatterometer measurements of wind speed and direction are quite frequent (2 measurements per day taken by QuikScat at each point on the mid-latitudes). The spatial resolution of NSCAT and AMI is of 50 km. SeaWinds has a spatial resolution of 25 km. These resolutions are too coarse to achieve the kilometre-scale mapping of offshore wind resource. Presently, most of the offshore wind farms are built at less than 20 km from the coastline. Scatterometers are characterized by a lack of data over coastal areas due the contamination of measurements by land backscatter in these areas. As an example, the nearest measurements to the coast given by SeaWinds are located 25 km offshore. Scatterometers have sufficient temporal resolution but insufficient spatial resolution for wind resource assessment.

High spatial resolution wind fields are retrieved using SAR images. They are present in several missions (such as ERS-1, ERS-2, Radarsat, Envisat and TerraSAR-X). The typical spatial resolution of SAR images is of a few tens of meters. The wind fields retrieved using these images have, however, a spatial resolution of a few hundreds of meters. Indeed, the measurements taken by SAR have to be averaged over several pixels [9]. This is due to noise (speckle) in SAR images that affects the backscattering coefficient. Wind speed is also retrieved from SAR images using CMOD models. In order to retrieve wind speeds from SAR images wind direction is needed as input. Mainly two approaches are used to get information about wind direction. The first is based on the retrieval of the wind direction from meteorological features that could be present in the SAR image [10, 11, 9, 12, 13]. The second approach consists on using wind directions provided by external sources (numerical weather prediction models or buoys for example) [14]. The spatial resolution of the resulting wind fields suits the requirements of wind resource mapping offshore. Wind fields retrieved using SAR images can show phenomena not revealed by scatterometer measurements. As an example, wind fields retrieved from SAR images can show gap flows, characterized, typically, by an acceleration of wind over the sea between two close islands. If the gap between the islands is less than 25 km (scatterometer's spatial resolution), scatterometer measurements will not reveal this acceleration whereas this distance is sufficient to install a wind park. Nevertheless, high spatial resolution of SAR wind fields is associated to a low temporal sampling. In fact, at a given location, the temporal gap between two successive SAR images is too long to establish reliable wind statistics describing the wind climatology of the site. As an example, in [3], an average of 1.27 to 4.75 ERS-1 images per month over the French gulf of Lion is reported. This temporal sampling is not sufficient to establish a reliable statistical description of wind over an area [15]. Even with more frequent Envisat images acquisition, thanks to a wide swath mode, the temporal sampling of SAR images remains insufficient.

To achieve wind resource mapping offshore, a solution is to combine the two data sets to benefit

of the high spatial resolution of SAR and the high temporal resolution of scatterometers. The aim of the fusion method, presented in this paper, is to generate a synthetic data set allying the high temporal resolution of scatterometers and the high spatial resolution of SAR. The SAR wind fields are used to infer the high spatial structures that should be present in each scatterometer measurement.

In the section 2, we shortly present the data fusion method [16]. One of the components of the method is the synthesis of high spatial wind fields. In this paper, the focus will only be on this part of the method. The synthesis is presented in section 3. In section 4, we expose the application of the synthesis of high spatial resolution wind fields using scatterometer data and SAR wind fields. In section 5, we evaluate the results of this application and quantify the errors of the chosen implementation. Finally, some conclusions and prospects are given.

5.2 Data fusion method for offshore wind resource mapping

According to Kolmogorov's theory [17], the atmospheric large scale eddies are the source of the energy. This energy is transferred from large to small scales where it is dissipated through an intermediate range of scales, called inertial range, in which the energy is conserved and transferred at a constant rate. SAR and scatterometers measurements are located in this inertial range. The high spatial resolution structures, representing local effects, in SAR wind fields are, therefore, drived by low spatial resolution (or large scale) structures present in scatterometer measurements.

The fusion method presented in this paper is based on the assumption that if two wind fields are similar at low spatial resolution, they belong to the same wind situation, also called typical situation in Fig. 1, and the finer scale structures of these wind fields are similar. Based on this assumption, it is only necessary to have one high spatial resolution measurement to infer the high spatial structures (representing local effects and turbulence) of all similar wind fields. The fusion method aims at generating a synthetic data set of wind fields at high spatial resolution and high temporal resolution. Hence, the high spatial structures contained in a SAR wind field are injected on scatterometer measurements belonging to the same typical situation.

The data fusion method scheme is presented in Fig. 5.1. The first step of the method is the classification of low spatial resolution wind fields. The classification method is detailed in [18, 16]. The classification results on several classes grouping similar wind fields. Each class represents a typical behaviour of wind flow over the studied area. Therefore, under the assumption of the conservation of similarity when the spatial resolution increases, each class could be associated to typical high spatial resolution structures.

The classification is followed by the association of a SAR wind field to each class. This high spatial resolution wind field, obtained from SAR, is used to extract the high resolution structures associated to the class. The step of association results, as represented in Fig. 5.1, in a set of high spatial resolution typical fields. For each class, a typical field is associated. This typical field is representative of the high spatial resolution structures that should be "added" to each low spatial resolution wind field to generate a synthetic high spatial resolution wind field. This synthetic data set of wind fields allies the high spatial

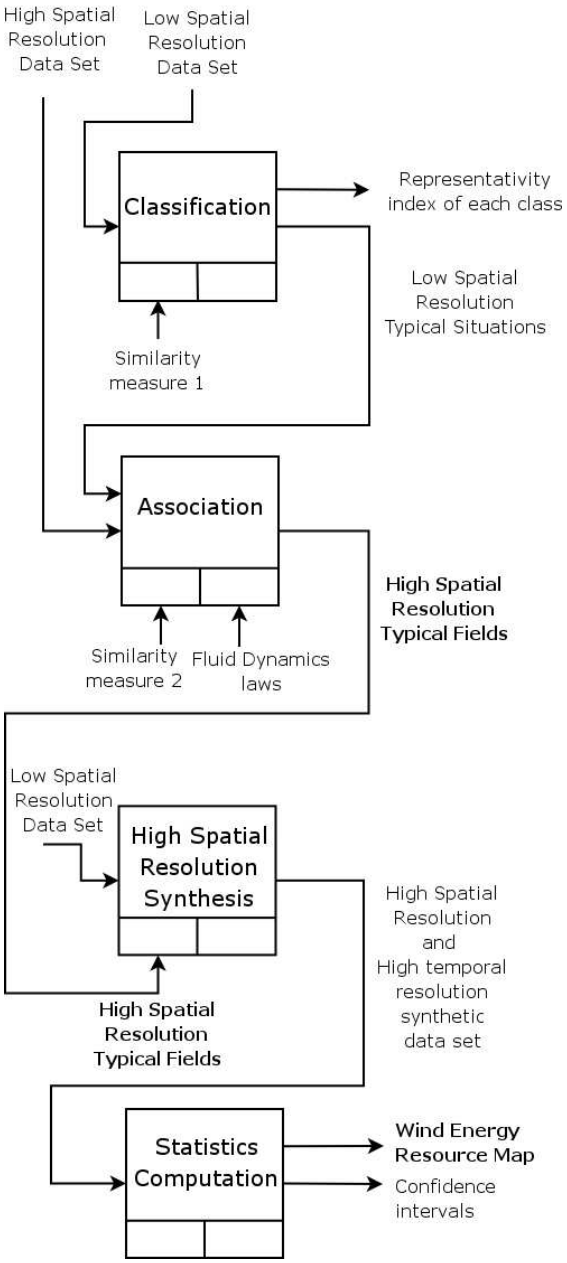


Figure 5.1: Scheme of the data fusion method for wind resource mapping offshore.

resolution and the high temporal resolution. It is used to compute the statistical parameters representing the wind climatology. In the next section, we focus on the high spatial resolution synthesis.

5.3 High spatial resolution wind fields synthesis

This principle of similarity of the structures contained in the wind fields at each scale induces a necessity of a scale-by-scale description of the information content [19]. This description is then used to increase the spatial resolution of the wind field. This principle suits the ARSIS (from its French name “Amélioration de la Résolution Spatiale par Injection de Structures”) concept of image fusion used in several areas such as urban environment study [20, 21], geology [22], air quality [23] and climatology [24]. This con-

cept will be used to define the method of high spatial resolution wind fields synthesis. In the following, the ARSIS concept is first presented and the choices conducting to the method for high spatial resolution wind fields synthesis expressed.

5.3.1 ARSIS concept

Lets denote the acquired image of lowest spatial resolution by B_l, and the image of highest spatial resolution by A_h. The subscripts l and h denote the spatial resolution of observations B or A, *i.e.*, low and high resolution, respectively. The fusion methods aim at constructing a synthetic image B_h^*, which is close to what would be observed by a similar sensor with a better spatial resolution. The methods should perform a high-quality transformation of the content of B_l, when increasing the spatial resolution from l to h. The general problem is relevant to the fusion of representations and is the creation of a new representation B^* from the original representations A and B [25]:

$$B^* = f(A, B) \tag{5.1}$$

It may be seen as the inference of the information that is missing in the image B_l for the construction of the synthesized images B_h^*.

The ARSIS concept is based on the assumption that the missing information is linked to the high frequencies of the representations A and B. It searches a relationship between the high frequencies in B and A and models such a relationship. A method belonging to the ARSIS concept performs typically the following operations:

(i) the extraction of a set of information from A;

(ii) the inference of the information that is missing in the image B_l using this extracted information;

(iii) the construction of the synthesized image B_h^*. The most recent methods perform a scale-by-scale description of the information content of both images and synthesis of the high-frequency information missing to transform the low spatial resolution image into high spatial resolution image.

Fig. 5.2 presents the general scheme that applies in the case of use of a multiscale model [26]. This case is used in the following for a better description of the ARSIS concept. Similar schemes can be used in other cases, where other tools or strategies are used. Input to the fusion process are the images A at high spatial resolution (A_h, resolution 1) and the image B at low spatial resolution (B_l, resolution 2).

Three models appear in this scheme of Fig. 5.2. The Multiscale Model (MSM) performs a hierarchical description of the information content relative to spatial structures in an image. An example of such a model for remotely sensed images is the combination of the wavelet transform and multiresolution analysis [27]. [26] provide practical details for the implementation of the algorithm of Mallat combined with a Daubechies wavelet. When applied to an image, the MSM provides one or more images of details, that are the high frequencies, and one image of approximation, i.e. the lower frequencies. As an example, assume an image at 1 m resolution. The first iteration of the MSM gives one image of the structures comprised between, say, 1 and 2 m (details image) and one image of the structures larger than 2 m

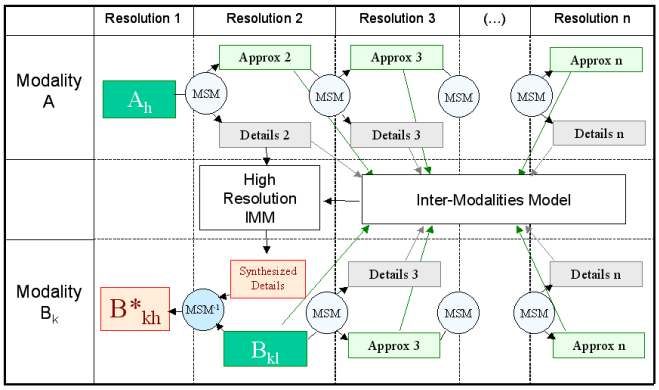


Figure 5.2: General scheme for the application of the ARSIS concept using a multiscale model (MSM) and its inverse (MSM^{-1}).

(approximation image). The spatial variability within an image can thus be modelled and the model can be inverted (MSM^{-1}) to perform a synthesis of the high-frequency information to retrieve the original image at 1 m.

The Inter-Modality Model (IMM) deals with the transformation of spatial structures with changes in modality. It models the relationships between the details or approximation observed in the representation A and those observed in the representation B. The IMM may relate approximations and/or details for one or more resolutions.

The High Resolution Inter-Modality Model (HRIMM) performs the transformation of the IMM with the change in resolution. This operation is not obvious. Implementing a method derived from this concept, consists on the choice of a MSM, an IMM and a HRIMM.

5.3.2 Synthesis of high spatial resolution wind fields

Let SAR_t denote the wind field at high spatial resolution at the date t and $SCAT_t$ the wind field at low spatial resolution at the same date. The aim of the fusion method is to generate a synthetic high spatial resolution wind field SAR_t^* for dates where no high spatial resolution SAR images are available. This synthetic wind field is based on $SCAT_t$ and on a SAR image acquired at a different date but likely containing similar high spatial structures. Let t_2 be the date without corresponding high spatial resolution wind field and t_1 a date corresponding to the same typical situation and corresponding to the moment of the acquisition of a SAR image. Under the assumption that the high spatial resolution structures present in two images belonging to the same typical situation are similar, the high spatial structures present in SAR_{t_1} are similar to those that would be observed if SAR_{t_2} were acquired. The synthetic high spatial resolution wind field $SAR_{t_2}^*$ is expressed as function of $SCAT_{t_2}$ and SAR_{t_1}:

$$SAR_{t_2}^* = f(SCAT_{t_2}, SAR_{t_1}) \tag{5.2}$$

Ideally, we should have:

$$SAR_{t_2}^* = f(SCAT_{t_2}, SAR_{t_2}) \tag{5.3}$$

if SAR_{t_2} were available.

[26, 1] present some examples of implementation of the ARSIS concept. Here, the selected MSM is the Undecimated Wavelet Transform (UWT) [28]. The decimated bi-orthogonal wavelet (DWT) transform is characterized by a loss of the translation-invariance property leading to a large number of artefacts when an image is reconstructed after modification of its wavelet coefficients. The continuous wavelet transform (CWT) is characterized by a great amount of redundancy and by the impossibility of a perfect reconstruction. The UWT is intermediate between the DWT and the CWT. For the IMM and the HRIMM models, as a first approach, we selected the identity model. This model considers that the high spatial resolution structures of the two dates t_1 and t_2 are exactly identical if t_1 and t_2 belong to the same class. More sophisticated models exist. Such models try to identify a relationship between the details of the images [26, 29, 1]. But, in our application, and based on the similarity assumption these models should provide relevant results.

The fusion can only be undertaken over the intersection of the regions where SCAT and SAR measurements are available. For scatterometer measurements we consider the convex hull of the covered zone. This is due to the necessity of interpolating SCAT measurements for the fusion. We use a cubic interpolation.

5.4 Application to some examples

5.4.1 Data presentation

The classification scheme was applied on data covering the Irish Sea [18]. To illustrate the fusion scheme, we choose two classes resulting from this classification, hereafter denoted Class 1 and Class 2. Class 1 is characterized by southwesterly winds, *i.e.* the dominant winds over the Irish Sea. Class 2 is characterized by easterly winds. To retrieve wind speeds from the SAR images we used the empirical algorithm CMOD-4 algorithm. As input for wind directions, we use the directions retrieved using the features present in SAR images [13].

Two SAR images acquired by ERS-2 are available for Class 1. Two SAR images, acquired by the same satellite, correspond to Class 2 situations. The images used are listed in Table 5.1. In addition to the date and time of acquisition of SAR images and scatterometer measurements, the closest wind speed and direction measurements taken by a mast installed offshore, near the Irish coast, are given. The mast coordinates are (Longitude: 05.92 W, Latitude: 52.89 N). The measurements of the mast are taken with a maximal gap of ± 5 minutes relatively to the SAR images acquisition time. These measurements are 10-minute averaged wind speeds and directions. As shown in Table 5.1, the temporal gap between the acquisition of SAR and scatterometer measurements varies from less than 3 hours on August 18th, 2003 to more than 5 hours on March 19th, 2001. The two situations of Class 1 (images 5.3(a) and 5.3(b)) correspond to stable meteorological situations. The same situation remains for the following six hours. In the case of Class2 (images 5.3(c) and 5.3(d)) the images are corresponding to unstable meteorological situations and are, therefore, less reliable for representing the class. For March 19th, 2001, the situation was initially classified as outlier. It was in a second step affected to this typical situation as explained in the classification scheme presented in [18]. The situation of June 28th, 2003 was also classified as outlier

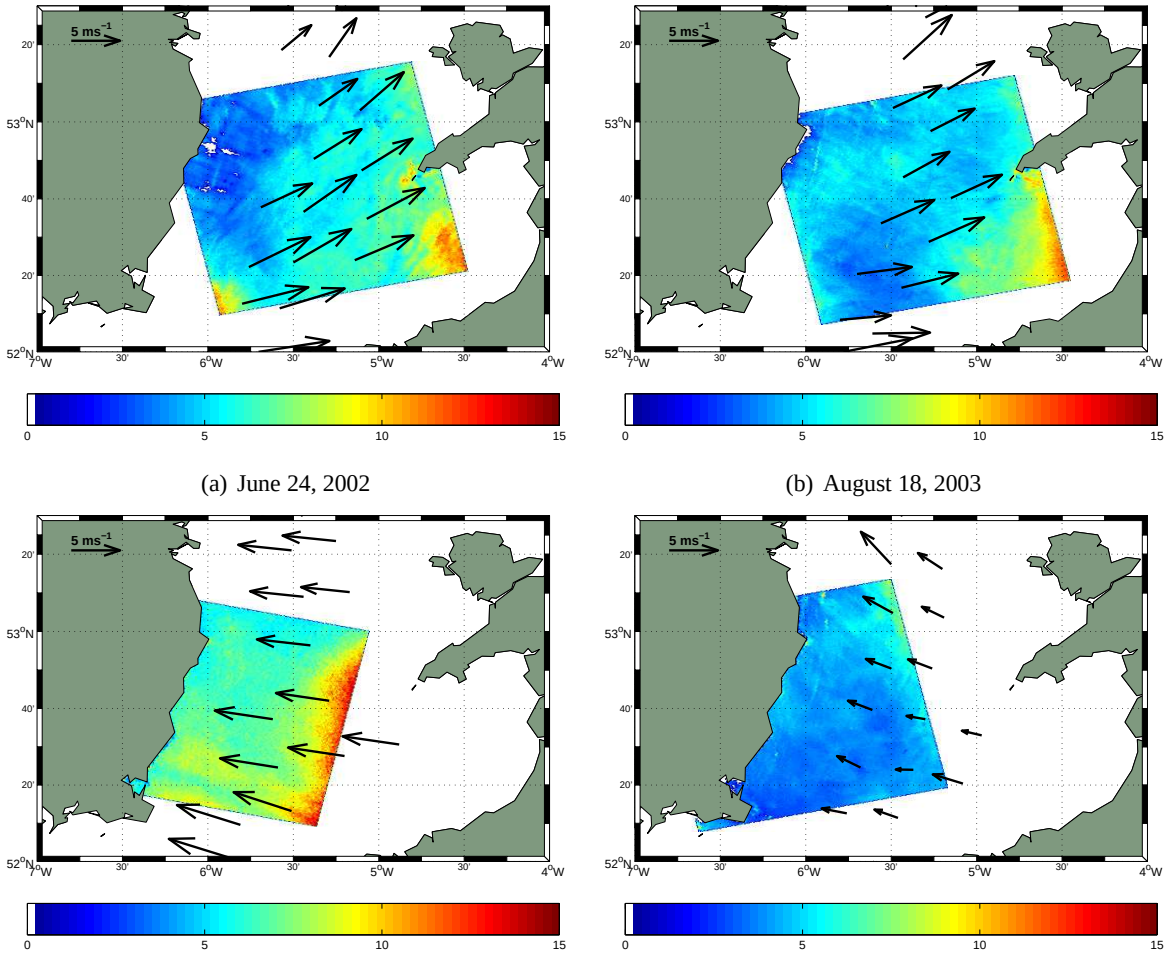
Table 5.1: Speed and direction values are averaged over 10 mn bins. Mast coordinates: Latitude 52.89 N, Longitude 05.92 W

SAR acquisition time	SCAT acquisition time	Mast measured direction	Mast measured speed
Class 1			
24-06-2002 22h22	24-06-2002 18h21	305 deg.	4.6 m s^{-1}
18-08-2003 22h22	18-08-2003 19h44	271 deg.	7 m s^{-1}
Class 2			
19-03-2001 11h20	19-03-2001 05h43	112 deg.	7.7 m s^{-1}
28-06-2003 22h25	28-06-2003 18h21	120 deg.	4.1 m s^{-1}

in a first step. In addition, after 6 hours the situation classification changed to another typical situation. Hence, Class 1 cases are favourable cases for high spatial resolution wind fields synthesis and Class 2 cases are unfavourable cases.

The wind fields retrieved from SAR images listed in Table 5.1 are presented in Fig. 5.3. Arrows represent the wind fields measured by SeaWinds scatterometer on board QuikScat. The pseudocolor images represent the SAR based wind fields. The four wind fields represent a variety of wind conditions going from moderate wind speeds to relatively high wind speeds.

The chosen spatial resolution of SAR based wind fields is of 400 m. The spatial resolution of SCAT measurements acquired by Quikscat is of 25 km. The goal of the fusion algorithm is to generate synthetic 400 m wind fields for dates where only SCAT measurements where acquired and no SAR image is available. In the following, for each date, we use the QuikScat measurement and the SAR wind field of the other date belonging to the same Class to generate a synthetic wind field at high spatial resolution. The original wind field is used as reference to assess the quality of the fused wind field. For example, the high spatial resolution structures extracted from the wind field obtained from the SAR image of June 24th, 2002 are injected in the wind field measured by the scatterometer on August 18th, 2003. The resulting synthetic wind field is compared to the wind field obtained from the SAR image of August 18th, 2003.



(a) June 24, 2002

(b) August 18, 2003

(c) March 19, 2001

(d) June 28, 2003

Figure 5.3: Wind fields retrieved using CMOD-4 from SAR image and the corresponding scatterometer measurements for the different dates of Table 5.1. 5.3(a) and 5.3(b) are belonging to Class 1. 5.3(c) and 5.3(d) are belonging to Class 2. Speeds are in m s^{-1}.

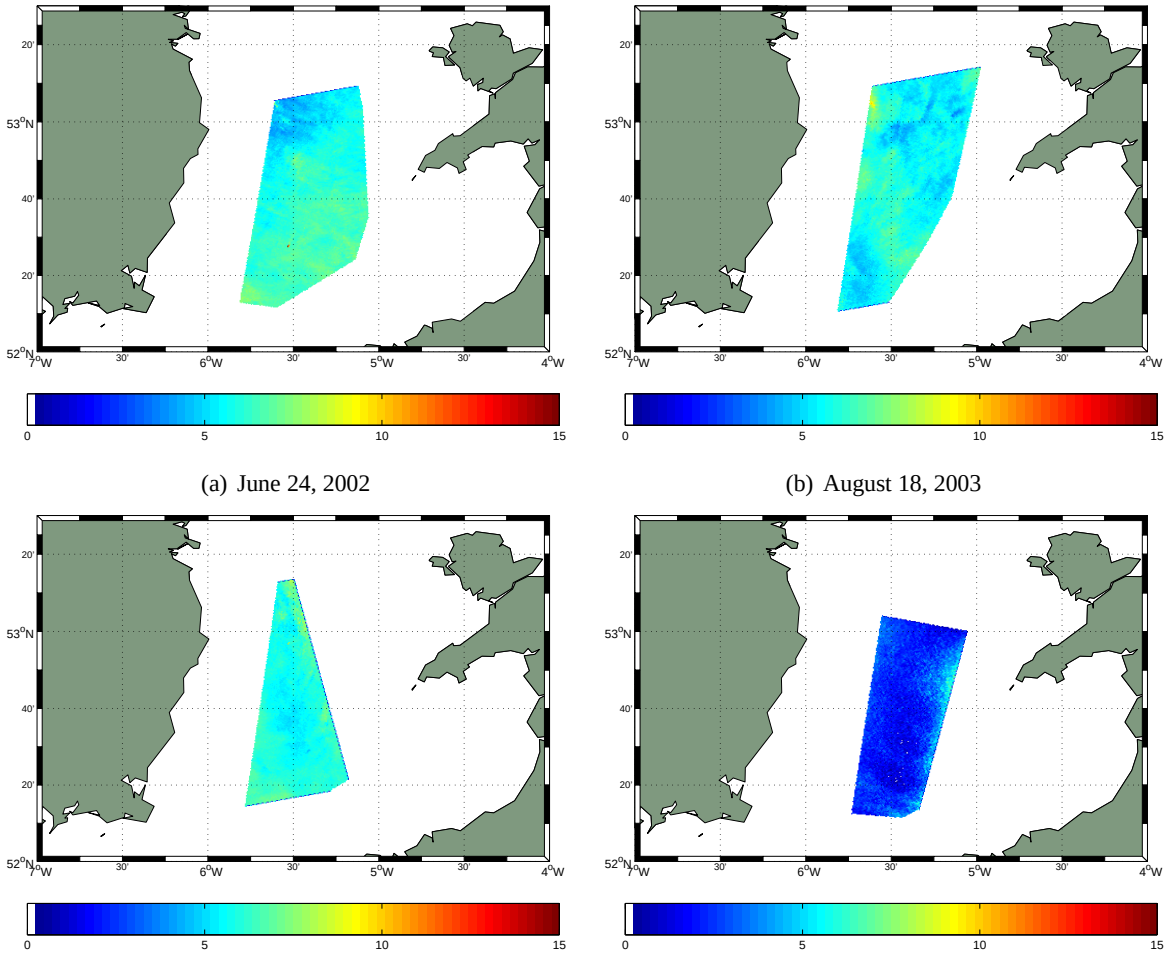
5.4.2 Results

The synthesized wind fields for each of the dates given in Table 5.1 are given in Fig. 5.4. These synthetic wind fields cover the convex hull of scatterometer measurements of each date. In Fig. 5.5, are represented the relative errors contained in the synthetic wind fields (relative to the wind speeds in the reference wind fields).

The synthetic wind field generated for June 24th, 2002, represented in Fig. 5.4(a) shows the same spatial distributions of wind speeds as the reference wind field represented in Fig. 5.3(a). It is characterized by low wind speeds ($\sim$ 5 m s^{-1}) on the north-western part of the wind field and moderate wind speeds ($\sim$ 7-8 m s^{-1}) on the remaining part of the wind field. It is in accordance with wind speeds present in the reference wind field of the Fig. 5.3(a).

The synthetic wind field generated for August 18th, 2003 (Fig. 5.4(b)) is characterized by slightly higher speeds than those of the reference wind field (Fig. 5.3(b)).

The synthetic wind fields generated for March 19th, 2001 (Fig. 5.4(c)) and for June 28th, 2003 (Fig.



(a) June 24, 2002

(b) August 18, 2003

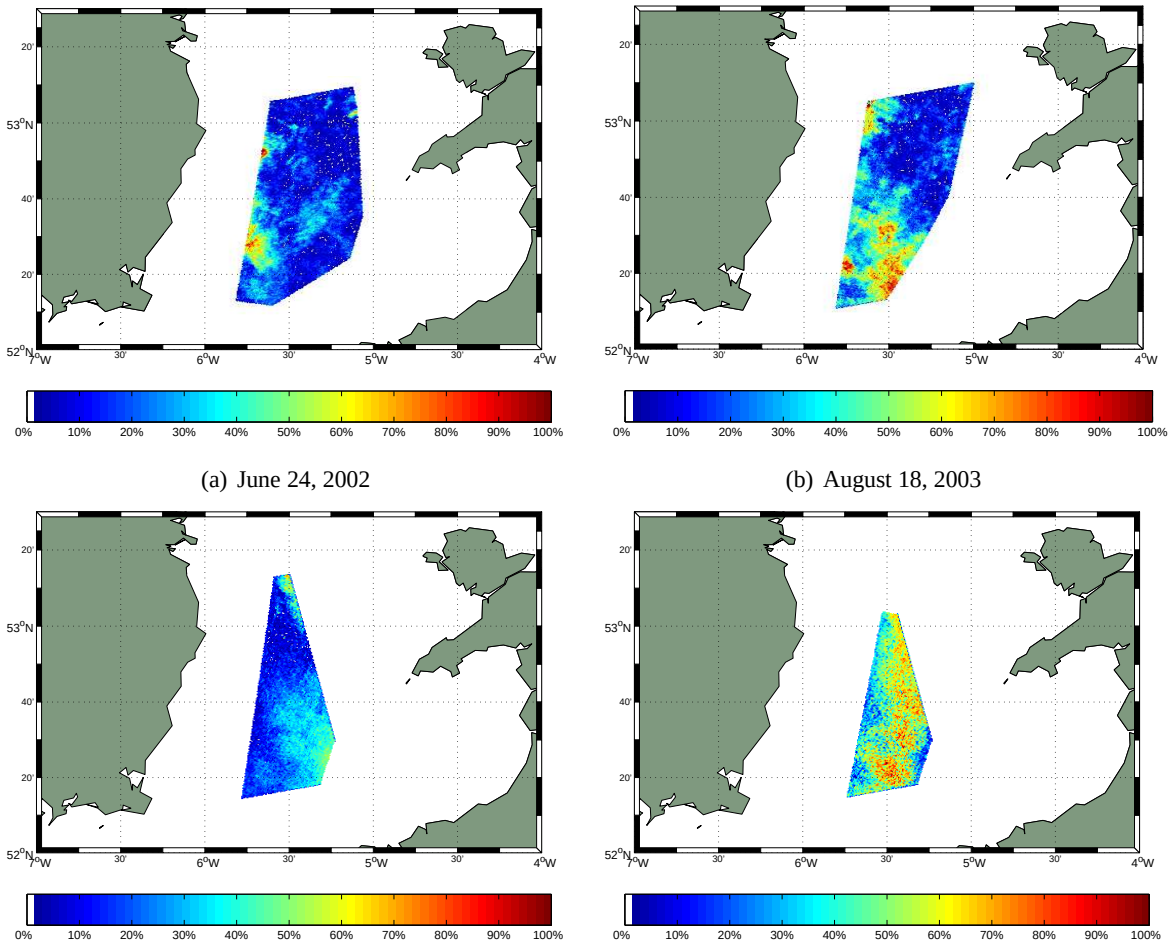
(c) March 19, 2001

(d) June 28, 2003

Figure 5.4: Synthetic wind fields for the different dates of Table 5.1. 5.4(a) is the result of the fusion of the SAR acquired on August 18, 2003 and QuikScat measurements of June 24, 2002. 5.4(b) is the result of the fusion of the SAR acquired on June 24, 2002 and QuikScat measurements of August 18,2003. 5.4(c) is the result of the fusion of the SAR acquired on June 28, 2003 and QuikScat measurements of March 19, 2001. 5.4(d) is the result of the fusion of the SAR acquired on March 19, 2001 and the QuikScat measurements of June 28, 2003. Speeds are in m s^{-1}.

5.4(d)) show wind speeds with small variations. The low wind speeds of the synthetic wind field of June 28, 2003 (Fig. 5.4(d)) are in accordance with the reference wind field (Fig 5.3(d)). The wind speeds in the synthetic wind field generated for March 19, 2001 (Fig. 5.4(c)) are higher. They are in accordance with the wind speeds in the reference wind field (Fig. 5.3(c)).

The visual analysis of the synthetic wind fields shows that the range of wind speeds in each synthetic wind field is in accordance with the range of wind speeds in its correspondent reference wind fields. However, some differences, in the wind field structures, between the synthetic and the reference wind fields appear. These differences can be attributed to the temporal differences between the synthesized and the reference wind fields. In the next section, we use quantitative criteria to assess the differences between these wind fields.



(a) June 24, 2002

(b) August 18, 2003

(c) March 19, 2001

(d) June 28, 2003

Figure 5.5: Relative error on the synthetic wind fields compared to the reference wind fields of Fig. 5.3.

5.5 Quality assessment

In order to quantitively assess the quality of the fused wind fields, we use the protocol defined in [30] and derived from [29]. The distances retained for the comparison of two wind fields are those of [31]. This protocol states that the assessment of the quality of the resulting synthetic wind fields is achieved by the assessment of two properties: the coherence property and the synthesis property.

1. **Coherence property** The coherence property states that once the synthetic wind field SAR_t^* degraded to its original resolution, it should be as similar as possible to the original wind field $SCAT_t$.

2. **Synthesis property** The synthesis property states that the resulting wind field SAR_t^* should be as close as possible to the wind field that would be retrieved from SAR if an image were acquired at high spatial resolution at the date t.

In order to assess the results of the fusion, we compare the resulting fused wind fields to the reference wind fields. The difference of the resulting fused wind fields and the reference wind fields is not only

explained by the error due to the fusion scheme. The temporal difference between the acquisition of the SAR and the SCAT can contribute to this difference.

In the case of lack of "truth wind fields", the extrapolation of quality assessment to high resolution is proposed by [29]. In the case studies presented here, although the reference wind fields differ from the truth wind fields due to the temporal gap between the acquisition of scatterometer measurements and SAR images, the SAR wind field represent a good reference especially in the cases where a relative high speed wind exists. In such situations the wind field is persistent and the temporal gap between the SAR and scatterometer measurements could be considered less influent than it is for low wind speed situations. As stated in section 5.4.1, the wind fields of Class 1 are persistent and the temporal gap between the SAR images and QuikScat measurements is smaller than for wind fields of Class 2. The quality assessment of the fused wind fields of Class 1 is, therefore, more significant than that of Class2.

The quality assessment proposed in [31] is composed of different parameters: the bias, the correlation coefficient, the standard deviation of the error and the difference between the variances of the reference image and the fused image. In addition, we give the root mean squared error (RMSE). The RMSE is used in many studies to quantify the error on wind speeds retrieved using remotely sensed data.

The relative bias is defined as:

$$\text{bias}_{\text{rel}} = \frac{\mu\left(SAR^{*} - SAR^{ref}\right)}{\mu\left(SAR^{ref}\right)} \tag{5.4}$$

where $\mu(.)$ denotes the mean of the wind field. The exponent * indicates the synthesized wind field and the exponent ^{ref} indicates the reference wind field.

The relative standard deviation of the difference between the synthetic wind field and the reference is given by:

$$\sigma_{\text{rel}} = \frac{\sigma\left(SAR^{*} - SAR^{ref}\right)}{\mu\left(SAR^{ref}\right)} \tag{5.5}$$

where $\sigma(.)$ denotes the standard deviation of the wind field.

The relative difference between the variances of the reference wind field and the fused wind field is:

$$\text{varDiff}_{\text{rel}} = \frac{\text{var}\left(SAR^{*}\right) - \text{var}\left(SAR^{ref}\right)}{\text{var}\left(SAR^{ref}\right)} \tag{5.6}$$

The correlation coefficient between the reference wind field and the fused wind field is defined as:

$$\text{cc} = \frac{\text{cov}\left(SAR^{*}, SAR^{ref}\right)}{\sigma\left(SAR^{*}\right)\sigma\left(SAR^{ref}\right)} \tag{5.7}$$

where cov$(.,.)$ is the covariance between two variables.

The RMSE is given by:

$$\text{RMSE} = \left(\text{var}\left(SAR^{*}\right) + \left(\mu\left(SAR^{*}\right) - \mu\left(SAR^{ref}\right)\right)^{2}\right)^{0.5} \tag{5.8}$$

For the assessment of the coherence property, the synthetic wind field SAR_t^* degraded to low spatial resolution l and $SCAT_t$ have to be compared. To degrade the synthetic wind field SAR_t^* we use the UWT. This is the the same MSM chosen for the fusion method. The spatial validity of the comparison

Table 5.2: Number of pixels where the assessment of the coherence property is valid. A pixel is a 400 m $\times$ 400 m area.

Date	number of pixels
Class 1	
24-06-2002	5331
18-08-2003	3555
Class 2	
19-03-2001	503
28-06-2003	110

between the synthesized wind field SAR_t^* degraded from resolution h to resolution l and the reference wind field, at resolution l, $SCAT_t$ is limited by the edge effects due to the limited area covered by the synthetic wind field. The wide swath mode images provided by Envisat should help to overcome this problem. In fact, the spatial coverage of such images being larger the valid area of comparison will be larger. In Table 5.2 are given the number of pixels remaining after the elimination of an area equivalent to the half of the wavelet filter length. This area is considered as the valid area for assessing the coherence property. The areas remaining for June 28th, 2003 and for March 19th, 2001 (Class 2 dates) are very limited. The results for these two dates can not be considered as statistically significant and will not be commented. In Table 5.3 are given the measurements of the difference between the interpolated measurements $SCAT_t$ and the wind field SAR_t^* degraded to the resolution of scatterometer measurements. For the two wind fields of June 24th, 2002 and August 18th, 2003 the quality assessment is globally good expected the correlation coefficient and the relative variance difference. The difference between the synthesized wind field SAR_t^* degraded from resolution h to resolution l and $SCAT_t$ can be explained by the use of the UWT as a MSM. In fact, UWT is characterized by a redundancy in the scales. Thus, when applying the UWT the decorrelation of scales is not respected and some interactions between different scales occurs. The low spatial resolution is affected by the introduced details. The protocol, used here, was established in cases where the ratio in spatial resolution is lower than the ratio between SAR and SCAT measurements. The ratio of SAR based wind fields and SCAT spatial resolutions is of 64 whereas the maximal ratio between images used with protocol was, until this application, of 10 [26].

The assessment of the synthesis property is given in Table 5.4. This assessment is done over the whole area where wind speeds are synthesized and for all cases. The reference wind field is the original SAR based wind field. As expected the error was larger on the wind fields of Class 2. This large error can be explained by the large temporal gap between the scatterometer and the SAR image for March 19th, 2001 and by the unstable meteorological conditions. This large error can also be explained by the results of the classification. In fact, both of the two situations of Class 2 were classified as outliers in the first step of the classification scheme. The occurrence rate of Class 1 is 6.6 % . The occurrence rate of Class 2 is 4 % [18]. This means that these situations are less representative of the typical situations grouping

Table 5.3: Quality assessment of the synthesized wind fields. Verification of the coherence property comparing SAR_t^* wind field degraded to resolution l and the $SCAT_t$ images.

Date	bias_{rel}	σ_{rel}	$\text{varDiff}_{\text{rel}}$	cc	RMSE (m s^{-1})
Class 1					
24-06-2002	-10.71 %	3.93 %	-27.30 %	0.86	0.72
18-08-2003	-7.34 %	7.29 %	-76.82 %	-0.23	0.57
Class 2					
19-03-2001	-34.40 %	14.95 %	1135 %	-0.20	0.89
28-06-2003	-34.37 %	27.28 %	-60 %	-0.70	0.28

them. This also means that these two situations are less similar each other than the situations of Class 1.

The low correlation coefficients between $SCAT_t$ and the degraded SAR_t^* can be explained by the sensitivity of this parameter to any spatial shift. Due the dynamic nature of wind field and the temporal gap between the acquisition of SAR images and scatterometer measurements such shifts are expected to be present in the wind fields. These low correlation coefficients do not mean, however, that the injected structures are not valid. This is shown by the low RMSE value.

Although any adjustment of the quantity of injected structures was undertaken by the fusion method, the relative variance difference between SAR_t^* and the reference SAR based wind field is very low for the images of Class 1. The only date where this distance measures a large difference is March 19th, 2001. This is due to the low variance of the injected structures. In fact, the wind field from which these structures were extracted (June 28th, 2003) is characterized by a very low wind speeds. In addition, a few variation is revealed by the SAR wind field of June 28th, 2003 given in Image 5.3(d).

The reported error on wind speeds retrieved by applying CMOD-4 on SAR images is of 20% and the RMSE is of 2 m s^{-1}. The RMSE of the synthesized wind fields exceeds this reference value of 2 m s^{-1} only for the images of Class 2. As explained before, these images are particular cases. A larger error on these dates is not surprising. The RMSE remains, however, near the value of 2 m s^{-1}.

The results of the quality assessment differ according to the class representativity of the "instantaneous" wind fields. The synthesized high spatial resolution wind fields corresponding to the dates of Class 1, which present persistent wind situations, are characterized by less difference with the reference wind fields than the synthesized wind fields corresponding to the dates of Class 2. The two dates of Class 2 are less representative because they were initially classed as outliers and reassigned to this Class by the classification scheme [18].

Table 5.4: Quality assessment of the synthesized wind fields. Verification of the synthesis property comparing SAR_t^* wind field and the reference wind field.

Date	bias_{rel}	σ_{rel}	$\text{varDiff}_{\text{rel}}$	cc	RMSE (m s^{-1})
Class 1					
24-06-2002	13.88 %	14.08 %	-1.46 %	0.46	1.05
18-08-2003	24.41 %	19.44 %	-3.78 %	0.08	1.42
Class 2					
19-03-2001	-22.65 %	19.03 %	-88.73 %	0.11	2.26
28-06-2003	-49.45 %	23.19 %	0.97 %	0.30	2.37

5.6 Conclusion

High spatial resolution wind measurements, associated to a high temporal repetitiveness, are needed by offshore wind industry. No data source can, presently, achieve both the high temporal repetitiveness and the high spatial resolution. In order to take advantage from the high temporal repetitiveness of scatterometer measurements and the high spatial resolution of SAR based wind fields, we proposed an information fusion method. This method aims at generating a synthetic data set allying high spatial resolution and high temporal resolution.

In this paper, we focused on a component of this fusion method: the synthesis of high spatial resolution wind fields. The high spatial resolution wind field synthesis scheme considers that similar wind fields at low spatial resolution are characterized by the same high spatial resolution structures.

The results of the application of the synthesis of high spatial resolution to some case studies are satisfying. The method was applied in favourable and unfavourable cases. As expected, in the unfavourable cases, the results are worst than in favourable case. However, even in unfavourable cases the error is not excessive. Different error distances have been used to evaluate the fused wind fields. The only distance showing unsatisfying results is the correlation coefficient. This is due to a temporal difference between the reference wind fields and the fused wind fields. In fact, this temporal difference induces a spatial shift of the structures at high spatial resolution present in the wind field. The correlation coefficient parameter is very sensitive to such shifts.

The applicability of data fusion for the generation of synthetic high spatial resolution wind fields has been shown. The method presented in this paper is independent of the geographic location. It could be applied everywhere scatterometer measurements and SAR images are acquired.

Future works will focus on more sophisticated IMM and HRIMM models. More advanced models, performing an adjustment of the injected structures, are expected to improve the results of the fusion.

The sources of difference between the reference image and the fused image are:

- the temporal gap between the SAR acquisition time and the SCAT acquisition time. During this time, meteorological phenomena can appear and change dramatically the wind field;
- the difference between the SCAT measurements and the SAR measurements when degraded to low spatial resolution;
- the fusion model considering the same high spatial structures for two individual wind fields. The model applied in this paper is the simplest model. [26] and [1] report the worst quality for this model;
- edge effects due to the use of wavelets. The practical application of UWT is done with filters. The size of the filter introduces a border effect in the image.

Bibliography

[1] T. Ranchin, B. Aiazzi, L. Alparone, S. Baronti and L. Wald. Image fusion. the ARSIS concept and some successful implementation schemes. *ISPRS Journal of Photogrammetry & Remote Sensing*, vol. 58, n° 1-2, pp. 4–18, 2003.

[2] I. Troen and E. L. Petersen. *European Wind Atlas*. Risoe National Laboratory, Roskilde, Danemark, 656 p., 1989.

[3] N. Fichaux, T. Poglio and T. Ranchin. Mapping offshore wind resources: synergetic potential of SAR and scatterometer data. *IEEE Journal of Oceanic Engineering*, vol. 30, n° 3, pp. 48–57, 2005.

[4] A. Stoffelen. *Scatterometry - De scatterometer*. Thèse de doctorat, University of Utrecht, Netherlands, 209 p., 1998.

[5] C. C. Wackerman, C. L. Rufenach, R. A. Shuchman, J. A. Johannessen and K. L. Davidsoon. Wind vector retrieval using ERS–1 synthetic aperture radar imagery. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 34, n° 6, pp. 1343–1352, 1996.

[6] A. Stoffelen and D. Anderson. Scatterometer data interpretation: Derivation of the transfer function CMOD4. *Journal of Geophysical Research*, vol. 102, n° C3, pp. 5767–5780, 1997.

[7] Y. Quilfen, B. Chapron, T. Elfouhaily, K. Katsaros and J. Tournadre. Observation of tropical cyclones by high-resolution scatterometry. *Journal of Geophysical Research*, vol. 103, n° C4, pp. 7767–7786, 1998.

[8] H. Hersbach, A. Stoffelen and S. de Haan. An improved C-band scatterometer ocean geophysical model function: CMOD5. *Journal of Geophysical Research*, vol. 112, n° C03006, 2007.

[9] J. Horstmann, W. Kock, S. Lehner and R. Tonboe. Wind retrieval over the ocean using synthetic aperture radar with C-band HH polarization. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 38, n° 9, pp. 2122–2131, 2000.

[10] T.W. Gerling. Structure of the surface wind field from the seasat SAR. *Journal of Geophysical Research*, vol. 91, n° C2, pp. 2308–2320, 1986.

[11] F. Fetterer, D. Gineris and C. C. Wackerman. Validating a scatterometer wind algorithm for ERS-1 SAR. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 36, n° 1, pp. 479–492, 1998.

[12] N. Fichaux and T. Ranchin. Combined extraction of high spatial resolution wind speed and wind direction from sar images: A new approach using wavelet transform. *Canadian Journal of Remote Sensing*, vol. 28, n° 3, pp. 510–516, 2002.

[13] B. Furevik, O. Johannessen and A. D. Sandvik. SAR-retrieved wind in polar regions - comparison with in situ data and atmospheric model output. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 40, n° 8, pp. 1720–1732, 2002.

[14] F. M. Monaldo, D. R. Thompson, R. C. Beal, W. G. Pichel and P. Clemente-Colon. Comparison of SAR-derived wind speed with model predictions and buoy comparisons. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, n° 12, pp. 2587–2600, 2001.

[15] R. J. Barthelmie and S. C. Pryor. Can satellite sampling of offshore wind speeds realistically represent wind speed distributions. *Journal of Applied Meteorology*, vol. 42, n° 1, pp. 83–94, 2003.

[16] M. B. Ben Ticha and T. Ranchin. A case based reasoning data fusion scheme: application to offshore wind energy resource mapping. In *Proceedings of the 9th International conference on information fusion*, Florence, Italy, 2006.

[17] H. Tennekes and J. L. Lumley. *First Course in Turbulence*. The MIT Press, Cambridge, U.S.A., 300 p., 1972.

[18] M. B. Ben Ticha and T. Ranchin. Wind fields typical spatial patterns generation. *submitted to the Journal of of Applied Meteorology*, 2007.

[19] M. Farge, N. Kevlahan, V. Perrier and E. Goiranf. Wavelets and turbulence. *Proceedings of the IEEE*, vol. 84, n° 4, pp. 639–669, 1996.

[20] V. S. Raptis, R. A. Vaughan, T. Ranchin and L. Wald. Assessment of different data fusion methods for the classification of an urban environment. In *Proceedings of the second conference F̈usion of Earth data: merging point measurements, raster maps and remotely sensed images, Sophia Antipolis, France, January 28-30, 1998, Thierry Ranchin and Lucien Wald Editors, published by SEE/URISCA, Nice, France*, pp. 167–182, 1998.

[21] I. Couloigner, T. Ranchin, V. P. Valtonen and L. Wald. Benefit of the future SPOT 5 and of data fusion to urban mapping. *International Journal of Remote Sensing*, vol. 19, n° 8, pp. 1519–1532, 1998.

[22] W. Yang, F. Cauneau, J.-P. Paris, and T. Ranchin. Fusion of SAR and SPOT images for the detection of geological features over the Three Gorges Dam site, China. In *Proceedings of the third*

conference F̈usion of Earth data: merging point measurements, raster maps and remotely sensed images; Sophia Antipolis, France, January 26-28, 2000, Thierry Ranchin and Lucien Wald Editors, published by SEE/URISCA, Nice, France, pp. 137–142, 2000.

[23] L. Wald and J. M. Baleynaud. Observing air quality over the city of nantes by means of Landsat thermal infrared data. *International Journal of Remote Sensing*, vol. 20, n° 5, pp. 947–959, 1999.

[24] S. Cros, M. Schroedter-Homscheidt, T. Ranchin and L. Wald. Improvement of operational atmospheric parameters observation: increasing spatial resolution of aerosols optical depth maps by a data fusion process. In *Proceedings of the 4th annual meeting of the European Meteorological Society, 26-30 September 2004, Nice, France*, 2004.

[25] L. Wald. *Data Fusion : definitions and architectures. Fusion of images of different spatial resolutions*. Presses de l'Ecole, Ecole des Mines de Paris, Paris, France, 200 p., 2002.

[26] T. Ranchin and L. Wald. Fusion of high spatial and spectral resolution images: the ARSIS concept and its implementation. *Photogrammetric Engineering & Remote Sensing*, vol. 66, n° 1, pp. 49–61, 2000.

[27] T. Ranchin. Wavelets, remote sensing and environmental modelling. In *Proceedings of the 15th IMACS World Congress on Scientific Computation, Modelling and Applied Mathematics, Volume 6: Applications on Modelling and Simulation*, pp. 27–34, 1997.

[28] P. Dutilleux. An implementation of the "algorithme a trous" to compute the wavelet transform. In J. M. Combes, A. Grossman and Ph. Tchamitchian, editors, *Wavelets: Time-Frequency Methods and Phase Space*, pp. 298–304. Springer, Berlin, 1989.

[29] L. Wald, T. Ranchin and M. Mangolini. Fusion of satellite images of different spatial resolutions: assessing the quality of resulting images. *Photogrammetric Engineering & Remote Sensing*, vol. 63, n° 6, pp. 691–699, 1997.

[30] C. Thomas and L. Wald. Assessment of the quality of fused products. In *Proceedings of the 24th EARSeL Symposium "New Strategies for European Remote Sensing", 25-27 May 2004, Dubrovnik, Croatia, Oluic (Eds.), Millpress, Rotterdam, ISBN 90 5966 003 X*, pp. 317–325, 2005.

[31] C. Thomas and L. Wald. Comparing distances for quality assessment of fused images. In *Proceedings of the 26th EARSeL Symposium "New Developements and Challenges in Remote Sensing", 29-31 May 2006, Warsaw, Poland, ISBN 978-90-5966-053-3*, pp. 101–111, 2007.

Chapitre 6

Evaluation de la qualité de la méthode de fusion pour la cartographie du potentiel éolien

6.1 Introduction

La méthode de fusion de données pour la cartographie du potentiel éolien offshore présentée au chapitre 3 génère une série synthétique de champs de vent qui sont destinés à décrire statistiquement la climatologie du vent sur la région étudiée. Pour l'évaluation de la méthode, deux aspects sont donc à étudier :

- la qualité des champs de vents synthétiques considérés individuellement, c'est-à-dire l'erreur sur les vitesses de vent estimées à chaque instant t où un champ de vent est synthétisé,
- la qualité des statistiques dérivées de ces champs, c'est à dire l'erreur sur l'estimation finale de la densité moyenne de puissance.

Ces deux aspects sont liés. En effet, l'erreur sur les champs de vent synthétisés conditionne l'erreur sur l'estimation finale de la densité moyenne de puissance et des paramètres de Weibull. L'estimation de la qualité individuelle des champs de vent synthétisés est nécessaire pour comprendre les sources d'erreur dans les champs synthétisés. Toutefois, elle ne peut pas conduire à une conclusion finale sur la qualité de la description statistique de la climatologie du vent.

Lors de l'évaluation de la qualité de la cartographie du potentiel éolien, il est nécessaire de choisir une référence à laquelle sont comparés les résultats. Dans le chapitre 5, on a évalué la qualité des champs de vents synthétisés à partir d'images ROS et de mesures diffusométriques dans quelques cas, où une image ROS de référence est disponible. Cependant, le nombre de tels cas est limité. Dans ce chapitre, on se propose d'étudier la qualité de la méthode dans un cadre plus général où des données pouvant être prises comme référence ne sont pas disponibles.

En l'absence de cartographie du potentiel éolien à haute résolution spatiale de référence, on choisit pour valider la méthode proposée de transposer l'application du problème à plus basse résolution spatiale, s'inspirant en cela de ce qui est proposé dans la littérature [1].

Dans le cadre de la fusion d'images [1], l'évaluation de la qualité se fait en déplaçant la fusion dans les échelles de façon à avoir une référence à laquelle peut être comparé le produit de la fusion. Pour illustrer cette évaluation de la qualité, on considère le schéma de fusion de la figure 6.1. On dispose de deux images A_{r_1} et B_{r_2}. Dans la figure 6.1, une représentation pyramidale des observations A et B est donnée. L'indice r_k représente la résolution spatiale de l'image. r_1 correspond à la haute résolution spatiale. L'image synthétisée $B^*_{r_1}$ a une résolution r_1. Aucune référence n'est disponible à cette résolution. L'évaluation du résultat de la méthode à la résolution r_1 est donc impossible.

Pour l'évaluation de la fusion, [1] applique la fusion à A_{r_2} et B_{r_3}. L'application de la fusion et l'évaluation de l'image résultant de ce processus sont présentés dans la figure 6.2. L'image $B^*_{r_2}$, synthétisée en appliquant la fusion à A_{r_2} et B_{r_3}, a une résolution r_2. Cette image peut être comparée à B_{r_2}. Le bilan de qualité établi à cette résolution est extrapolé à la résolution r_1. Le ratio des résolutions utilisées pour l'étude de la qualité $\frac{r_3}{r_2}$ est égal au ratio des résolutions des données d'origine $\frac{r_2}{r_1}$. [1] étudie la validité de cette extrapolation. Il en ressort que souvent cette hypothèse est vérifiée et que le bilan de qualité établi à la résolution r_2 est valide à la résolution r_1. Des cas d'images favorables et d'images défavorables, pour la validité de l'extrapolation du bilan de qualité au travers des échelles, ont été étudiés dans [1]. Les images signalées comme défavorables sont celles couvrant des zones urbaines et celles fortement bruitées. Dans le cas des images couvrant des zones urbaines, l'apparition de structures à la résolution caractéristique des bâtiments induit une grande différence entre une échelle et la suivante. Ceci peut expliquer la limitation de la validité de l'hypothèse d'extrapolation du bilan de qualité au travers des échelles. Dans notre cas, on s'intéresse à des champs de vent. Ces champs sont constitués par une variété de structures situées aux différentes échelles. Selon la théorie de Kolmogorv, la densité spectrale de ces champs est en $k^{-\frac{5}{3}}$. Il n'y a donc pas de discontinuité dans les échelles comme celle pouvant apparaître dans les zones urbaines. Notre cas d'étude peut donc être considéré dans les cas favorables pour l'hypothèse d'extrapolation dans les échelles.

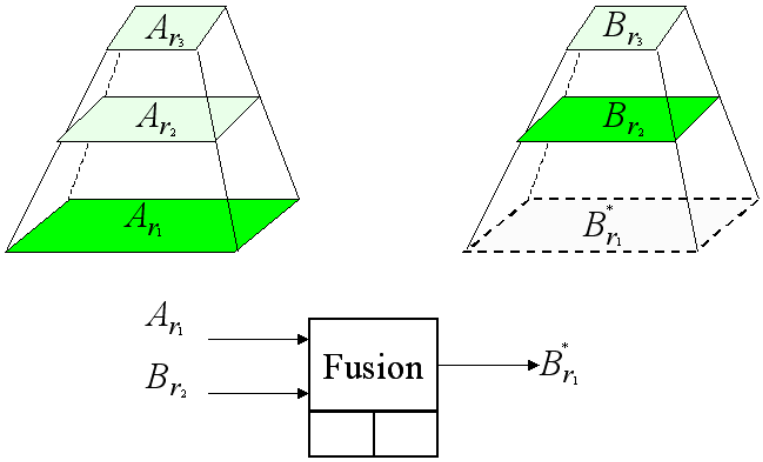


FIG. 6.1 – Schéma de fusion entre une image A_{r_1} à haute résolution spatiale et une image B_{r_2} à basse résolution spatiale. La représentation pyramidale des observations A et B est donnée en haut. Les images en vert foncé sont celles entrant dans le processus de fusion. L'image $B^*_{r_1}$ est synthétisée à la résolution spatiale r_1. Aucune référence n'est disponible pour l'évaluation du résultat de la fusion.

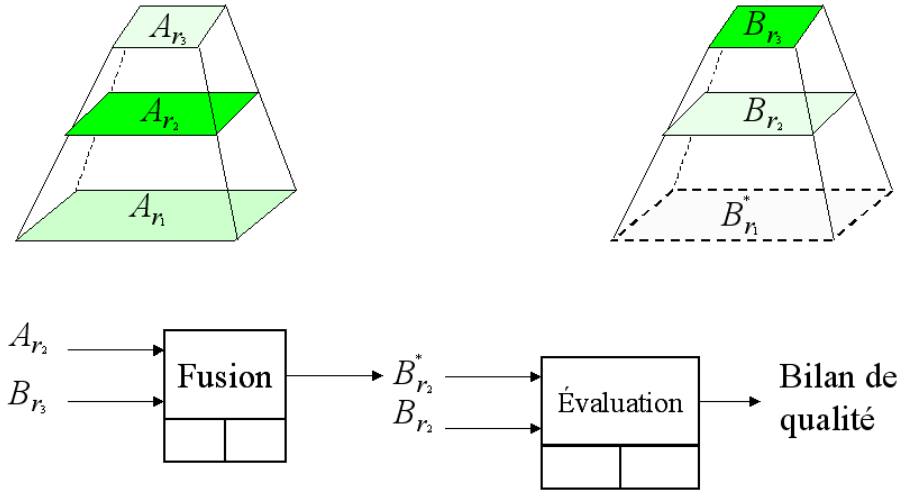


FIG. 6.2 – Déplacement de l'application de la fusion à fin de pouvoir évaluer la méthode. Contrairement à la fusion illustrée dans la figure 6.1, ce ne sont pas les images A_{r_1} et B_{r_2} qui sont fusionnées mais A_{r_2} et B_{r_3}. L'image synthétisée $B^*_{r_2}$ a une résolution r_2 et peut être ainsi comparée à l'image B_{r_2}. Le bilan de qualité, ainsi établi, est extrapolé à la fusion initiale illustrée par la figure 6.1.

La qualité de la fusion est quantifiée au travers de distances mesurant la similitude entre l'image résultant de la fusion et l'image référence. [1] a recensé les différentes distances utilisées dans la littérature pour la mesure de la qualité des images fusionnées et en a retenu quatre :

– le biais relatif entre l'image résultant de la fusion et l'image de référence donné par l'équation 6.1.

$$\text{biais}_{\text{rel}} = \frac{\mu\left(B^* - B^{ref}\right)}{\mu\left(B^{ref}\right)} \tag{6.1}$$

où $\mu(.)$ indique l'opérateur moyenne, B^* l'image synthétisée et B^{ref} l'image référence. Cette distance permet de mesurer une erreur systématique dans l'image fusionnée. Dans notre cas, où des champs de vent sont considérés, le biais peut être dû à une tendance d'une des sources de données à surévaluer les vitesses de vent.

– l'écart type relatif de la différence entre l'image fusionnée et l'image référence donné par l'équation 6.2,

$$\sigma_{\text{rel}} = \frac{\sigma\left(B^* - B^{ref}\right)}{\mu\left(B^{ref}\right)} \tag{6.2}$$

où $\sigma(.)$ indique l'opérateur écart type. Cette distance mesure la dynamique de l'erreur relative.

– la différence relative entre les variances de l'image fusionnée et l'image référence donnée par l'équation 6.3.

$$\text{varDiff}_{\text{rel}} = \frac{\text{var}\left(B^*\right) - \text{var}\left(B^{ref}\right)}{\text{var}\left(B^{ref}\right)} \tag{6.3}$$

où var$(.)$ indique l'opérateur variance. Cette distance permet de mesurer la différence entre la variabilité contenue dans l'image de référence et celle synthétisée. Elle permet ainsi de savoir si trop

(ou pas assez) d'information à haute résolution spatiale a été introduite dans l'image fusionnée.

– le coefficient de corrélation entre l'image fusionnée et l'image référence donné par l'équation 6.4.

$$\text{cc} = \frac{\text{cov}\left(B^*, B^{ref}\right)}{\sigma\left(B^*\right)\sigma\left(B^{ref}\right)} \tag{6.4}$$

où cov$(.)$ indique l'opérateur covariance. Cette distance permet de mesurer la concordance entre les structures spatiales présentes dans l'image fusionnée et celles présentes dans l'image de référence. Cette distance est très sensible à un décalage spatial des structures. Comme montré dans l'exemple montré dans le chapitre 5, s'il existe un décalage temporel entre le champ fusionné et le champ référence, cette distance peut donner de mauvais résultats.

Pour comparer les champs de vent on retient ces quatre distances et on rajoute la différence quadratique moyenne (RMSE) souvent utilisée pour quantifier l'erreur contenue dans des champs de vent. L'expression de cette distance est donnée dans l'équation 6.5

$$\text{RMSE} = \left(\text{var}\left(B^*\right) + \left(\mu\left(B^*\right) - \mu\left(B^{ref}\right)\right)^2\right)^{0.5} \tag{6.5}$$

L'application du protocole d'évaluation, au cas d'étude choisi est limité par la géographie de la région choisi (la Mer d'Irlande) qui se situe entre deux îles proches. Cette région a été choisie en raison de la présence d'un parc éolien offshore déjà installé près des côtes irlandaises. Cependant, la résolution spatiale des mesures diffusométriques (25 km) ne permet d'avoir que 3 points de mesures au maximum sur la même latitude en mer d'Irlande. L'application de la transformée en ondelettes sur ces données pour dégrader leur résolution introduirait beaucoup d'erreurs dues aux effets de bord. L'application du protocole d'évaluation décrit dans [1] tel quel est donc impossible. Pour obtenir une évaluation de la qualité de notre approche, nous sommes donc amenés à utiliser l'approche proposée à la figure 6.2.

6.2 Données utilisées et protocole d'évaluation

Pour l'évaluation de la méthode de fusion, on transpose le problème à des échelles où des données de référence sont disponibles. La transposition consiste à utiliser les grilles de réanalyse NCEP/NCAR comme des données à basse résolution spatiale. La haute résolution spatiale est incarnée par les grilles de réanalyse ECMWF. La résolution spatiale des grilles NCEP/NCAR est d'environ 1,9°×1,9°. La résolution spatiale des grilles de réanalyse ECMWF est de 0,25°×0,25°. Le ratio entre la résolution spatiale les grilles de réanalyse NCEP/NCAR et celle des grilles de réanalyse ECMWF est de 8. Ce ratio est inférieur au ratio existant entre la résolution spatiale des données diffusométriques et celle des champs de vent retrouvés à partir d'images ROS (ratio 64). Cependant, cette transposition permet d'évaluer l'influence de certaines caractéristiques des données sur la méthode de fusion. Les données utilisées dans cette étude couvrent quatre années de 2001 jusqu'à 2004. Toutes les 6 heures, une grille NCEP/NCAR et une grille ECMWF sont disponibles. Dans la transposition considérée, les grilles de réanalyse NCEP/NCAR jouent le rôle des données diffusométriques (basse résolution spatiale et haute résolution temporelle). Un sous-ensemble des grilles de réanalyse ECMWF simule les données ROS (haute résolution spatiale et basse

résolution temporelle). L'ensemble complet des grilles de réanalyse ECMWF (ayant la même haute résolution temporelle que les grilles de réanalyse NCEP/NCAR) sert de référence à laquelle est comparé le résultat de la fusion. Le sous-ensemble de grilles ECMWF choisi, pour simuler les données, ROS doit avoir une basse résolution temporelle. On choisit de retenir un nombre de grilles de réanalyse correspondant au nombre de situations types de vent ($M = 36$) établi au chapitre 4. Pour chaque situation type, une seule grille ECMWF est retenue. Cette grille est choisie aléatoirement parmi les dates appartenant à cette situation type. On dispose donc de 4 grilles de réanalyse NCEP/NCAR par jour pendant quatre années (le nombre total est $N = 5844$) avec une résolution spatiale de 1,9°×1,9° qui sont fusionnées avec 36 grilles de réanalyse ECMWF à une résolution spatiale de 0,25°×0,25° pour obtenir un ensemble de N champs de vent synthétiques. La cartographie obtenue grâce à cet ensemble de données synthétiques est comparée à celle établie à partir de l'ensemble complet des N grilles de réanalyse ECMWF. Le tableau 6.1 donne les règles de transposition utilisées et les caractéristiques de chaque source de données.

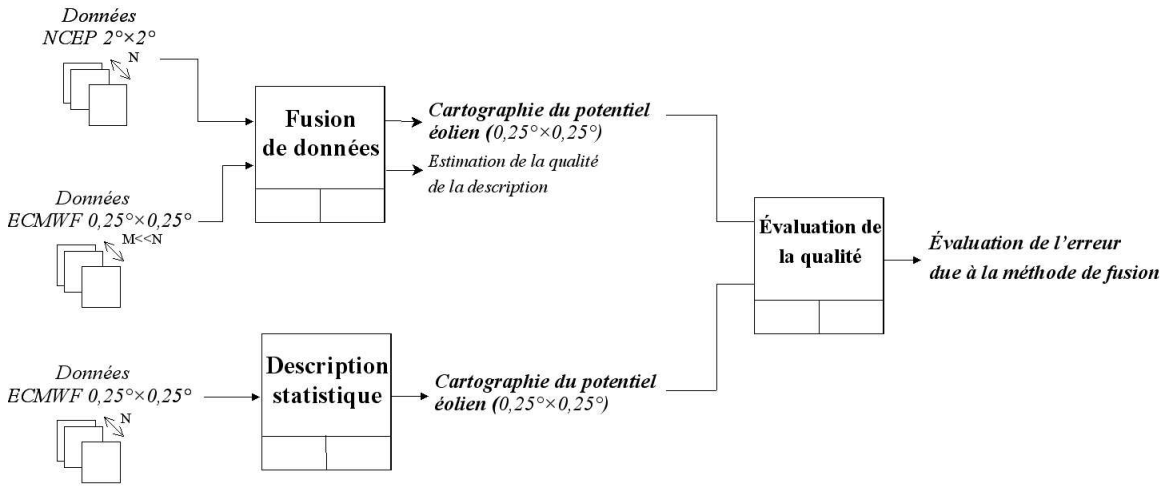


FIG. 6.3 – Schéma d'évaluation de la qualité de la méthode de fusion.

Le schéma représentant l'évaluation de la méthode de fusion est donné en figure 6.3. Le choix du sous-ensemble de grilles de réanalyse ECMWF, simulant la basse résolution temporelle, impliquant un choix aléatoire d'une grille pour chaque classe, on réitère ces évaluations plusieurs fois. Ceci permet de tester l'effet du choix des données à haute résolution spatiale sur le résultat de la fusion. La classification décrite au chapitre 4 produit des classes dont l'index de représentativité varie entre 0,7 % et 9,7 %. Le nombre de l'ensemble complet des grilles ECMWF étant $N = 5844$, l'effectif maximal d'une classe est de 565 grilles. On choisit un nombre d'itérations de l'application de la méthode de fusion de 100. Le choix aléatoire de la grille ECMWF pour chaque classe permettra pour ce nombre là de parcourir un nombre statistiquement significatif de grilles appartenant à la classe.

Outre l'évaluation du résultat final de la fusion (la cartographie du potentiel éolien au travers de paramètres statistiques), on évalue individuellement les champs de vent synthétiques générés. Le schéma de cette évaluation est donné dans la figure 6.4. Les données utilisées sont identiques à celles utilisées dans la figure 6.3. Dans ce schéma, l'évaluation de la qualité porte sur la série de données synthétiques générée par la fusion de données et utilisée pour établir la cartographie du potentiel éolien. Dans le cha-

TAB. 6.1 – Transposition du problème et les caractéristiques des différentes données utilisées.

	Données à haute résolution spatiale et basse résolution temporelle	Données à basse résolution spatiale et haute résolution temporelle	Données référence
Problème originel			
Source	mesures ROS	mesures diffusométriques	INDISPONIBLES
Résolution spatiale	400 m	25 km	-
Résolution temporelle	quelques jours	2 mesures/jour (aux latitudes moyennes)	-
Problème transposé			
Source	sous-ensemble de réanalyses ECMWF	réanalyses NCEP/NCAR	réanalyses ECMWF
Résolution spatiale	$0{,}25^\circ \times 0{,}25^\circ$	$1{,}9^\circ \times 1{,}9^\circ$	$0{,}25^\circ \times 0{,}25^\circ$
Résolution temporelle	irrégulière 1 grille/situation type	1 grille/6 heures	1 grille/6 heures

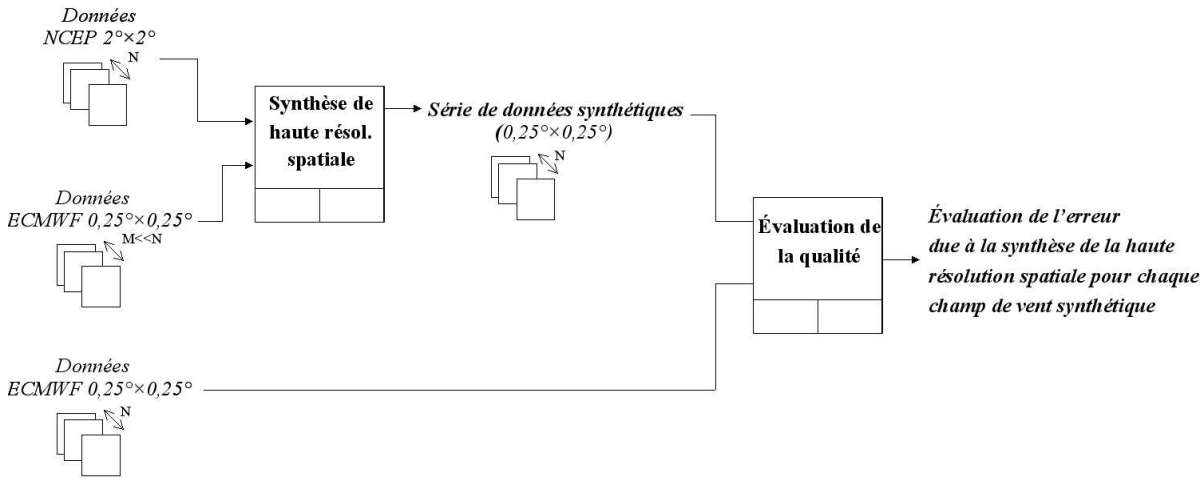


FIG. 6.4 – Schéma d'évaluation de la qualité de la méthode de fusion.

pitre 5, cette évaluation est fait au travers de quelques cas, où des champs de vent ROS sont disponibles. Ici, en utilisant une transposition dans les échelles, une évaluation systématique est possible.

Notre cas d'étude se situe en Mer d'Irlande. La figure 6.5 donne les points de la grille de réanalyse NCEP/NCAR considérés. Ces points sont superposés à un exemple de grille de réanalyse ECMWF. Comme expliqué dans le chapitre 5, la fusion ne peut être effectuée qu'à l'intérieur de la zone convexe des points de la grille NCEP/NCAR. Ceci est du à l'interpolation des valeurs NCEP/NCAR avant d'injecter

les structures à haute résolution spatiale. La méthode de synthèse de la haute résolution spatiale utilisée dans ce chapitre est la même que celle dans le chapitre 5. C'est la méthode UWT-M1.

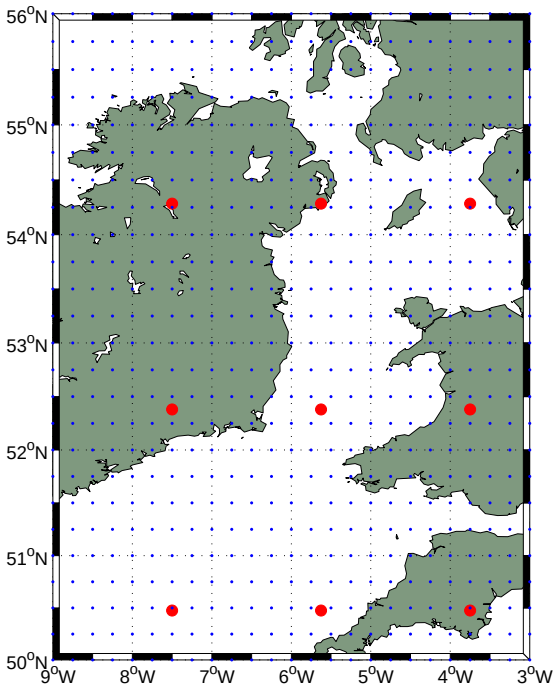


FIG. 6.5 – Région considérée pour l'application de la méthode de fusion. Les points rouges représentent les points de la grille de réanalyse NCEP/NCAR retenus. Les points bleus représentent les points de la grille de réanalyse ECMWF.

6.3 Analyse des résultats

6.3.1 Erreur sur les vitesses de vent

Dans cette section, on compare les vitesses de vent synthétisées issues de l'application de la fusion et les vitesses de vent référence. Le schéma de l'application de la fusion et de l'évaluation est présentée dans la figure 6.4. La figure 6.6 représente la distribution de l'erreur sur l'ensemble des champs de vent pour une application donnée de la méthode de fusion. Ces erreurs couvrent donc 4 années avec une grille synthétisée toutes les 6 heures. En réitérant l'application de la méthode en changeant le jeu de données simulant la haute résolution spatiale, la distribution de l'erreur garde les mêmes caractéristiques avec une faible variation des 4 premiers moments caractérisant la distribution. La moyenne de l'erreur est de 0,2 m s $^{-1}$. L'écart type est de 1,6 m s $^{-1}$. Le coefficient d'asymétrie est de 0,02. La faible valeur de ce coefficient confirme l'impression visuelle de symétrie de la distribution de l'erreur. Le coefficient d'aplatissement est de 0,9. Cette valeur indique que la distribution de l'erreur est plus pointue qu'une gaussienne. La faible variation de ces moments statistiques, et donc de la distribution, en changeant le choix du jeu de données simulant la haute résolution spatiale, prouve la stabilité de la méthode.

Outre cette série temporelle d'erreurs, il est intéressant d'étudier l'erreur pour chaque date. On cal-

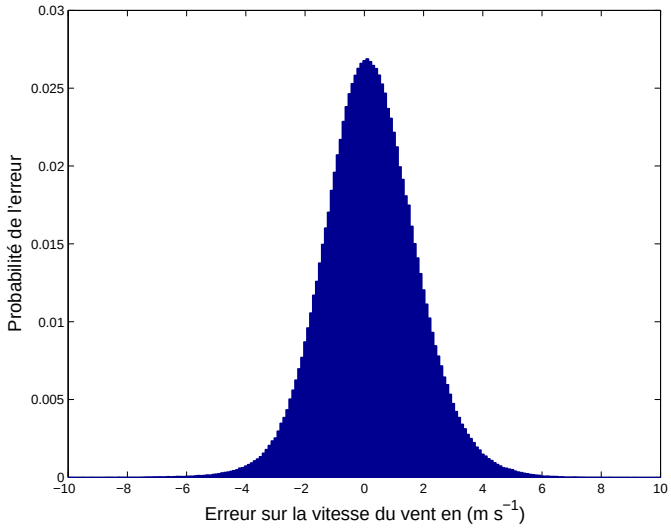


FIG. 6.6 – Distribution de l'erreur sur les vitesses de vent. L'histogramme est représenté avec des intervalles de vitesse de 0,1 m s^{-1}. La moyenne des erreurs est de 0,2 m s^{-1}. L'écart type est de 1,6 m s^{-1}. Le coefficient d'asymétrie est de 0,02. Le coefficient d'aplatissement est de 0,9. 81 % des erreurs sont dans l'intervalle [-2 m s^{-1},2 m s^{-1}].

cule ici la différence entre le champ de vent à une date t et le champ de vent référence à la même date de façon globale. On utilise ici les 5 distances présentées dans la section précédente pour la comparaison des champs de vent. La distribution des RMSE est donnée dans la figure 6.7. On remarque que 85 % des RMSE sont au dessous de 2 m s^{-1}. La valeur globale de la RMSE est de 1,6 m s^{-1}. Dans le chapitre 5, le calcul de la RMSE a donné des valeurs de 1,05 m s^{-1} et 1,42 m s^{-1} dans le cas favorable à la fusion et des valeurs de 2,26 m s^{-1} et 2,37 m s^{-1} dans le cas défavorable à la fusion. Les résultats de la fusion montrent que des cas défavorables à la fusion, avec une RMSE assez large, existent. Cependant, de tels cas sont assez rares.

La figure 6.8 montre la distribution des coefficients de corrélation calculés entre les données synthétisés et les champs de vent ECMWF références. Cette figure confirme l'explication avancée dans le chapitre 5. Nous avions imputé les mauvais coefficients de corrélation calculés à la différence temporelle entre les champs de vent synthétisés et les champs de vent références. En effet, ici avec les jeux de données NCEP et ECMWF, les données étant disponibles aux mêmes heures, on a éliminé le facteur de différence horaire entre les champs synthétisés et les champs références. Les coefficients de corrélation sont nettement améliorés avec 87 % des coefficients au dessus de 0,5 et 47 % au dessus de 0,8.

La figure 6.9 montre l'histogramme des biais relatifs sur les différents champs de vent synthétisés. Pour 85 % des champs de vent synthétisés, la valeur absolue du biais relatif aux champs de vent références est inférieur à 20 % et pour 55 % des champs de vent la valeur absolue du biais relatif est inférieure à 10 %. La figure 6.10 montre la distribution de la différence relative des variances des champs de vent synthétisés et des champs de vent références. Cette distance permet de mesurer si trop d'information à haute résolution spatiale a été injectée (et donc la variance du champ de vent synthétisé est augmentée par rapport au champ référence) ou pas assez (la variance du champ synthétisé est dans ce cas plus faible

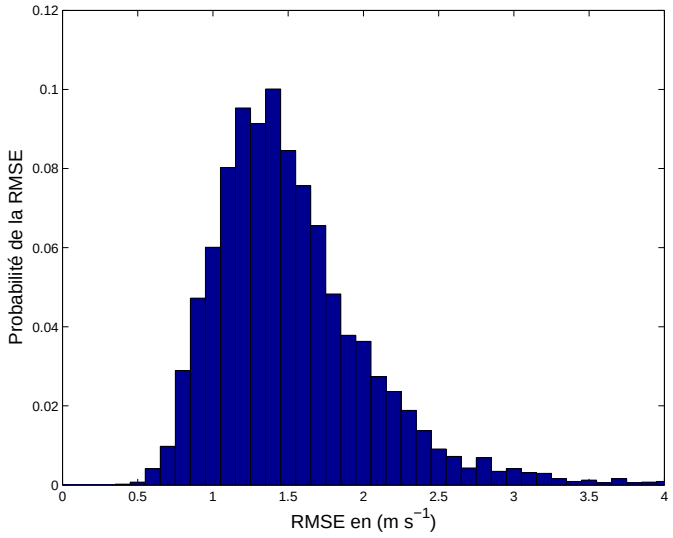


FIG. 6.7 – Distribution des RMSE. RMSE globale de 1,6 m s^{-1}. 85 % des RMSE sont au dessous de 2 m s^{-1}. L'histogramme est tracé avec des intervalles de 0,1 m s^{-1} pour les RMSE.

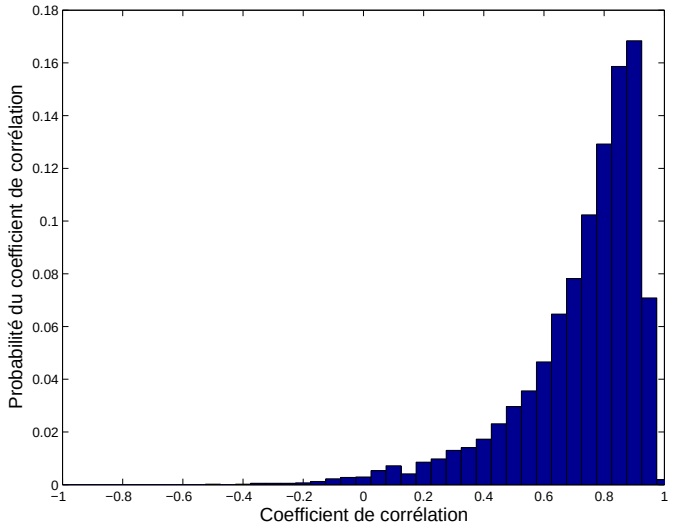


FIG. 6.8 – Distribution des coefficients de corrélation. 47 % des coefficients de corrélation sont au-dessus de 0,8 et 87 % au-dessus de 0,5. L'histogramme est tracé avec des intervalles de 0,05 pour le coefficient de corrélation.

que celle du champ référence). Il ressort de la figure 6.10 qu'il y a plus de cas où la variance des champs synthétisés est haute par rapport aux champs référence. Ceci peut s'expliquer par la présence de situations calmes sans structures présentes. Les structures à haute résolution spatiale injectées représentent un surplus d'information injectée. Ce phénomène peut être corrigé par des méthode de synthèse de la haute résolution spatiale plus évoluées que celle utilisée ici. Ces modèles auront pour rôle d'ajuster la quantité d'information ajoutée. Des exemples de ces modèles ont été présentés au chapitre 3.

L'histogramme de l'écart type de l'erreur sur les différents champs de vent synthétisés est présenté dans la figure 6.11. Cet histogramme représente la distribution des écarts types de l'erreur contenue dans

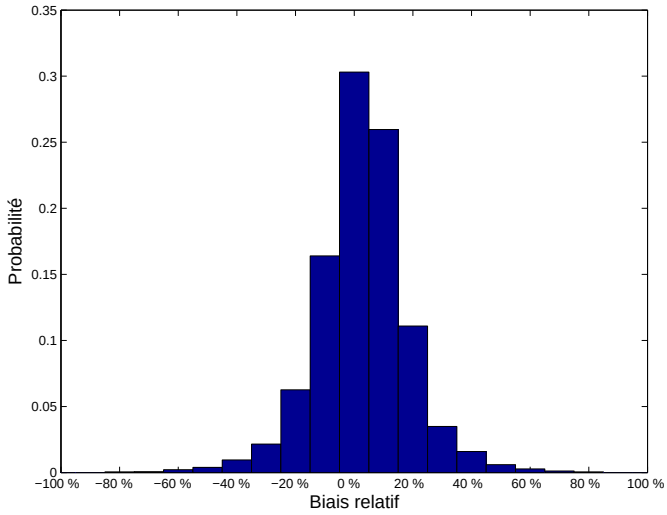


FIG. 6.9 – Distribution du biais relatif. 84 % des biais relatifs sont entre -20 % et 20 %. L'histogramme est tracé avec des intervalles de 10 % pour le biais relatif.

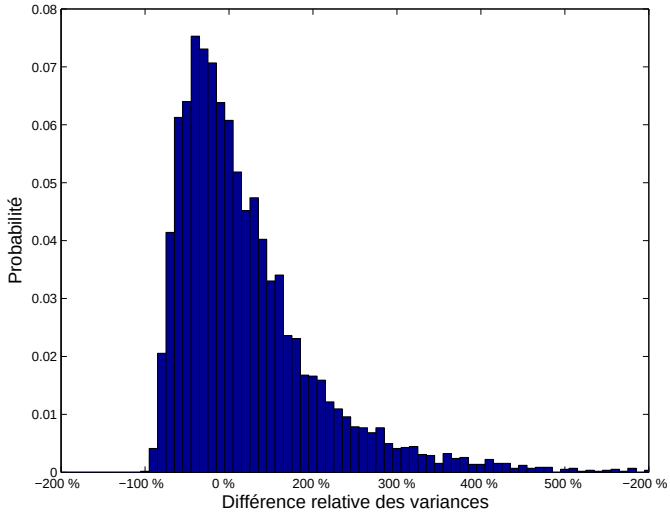


FIG. 6.10 – Distribution de la différence de variance. 58 % des différences sont entre -50 % et 50 %. 86 % entre -100 % et 100 %.

chaque champ de vent synthétique relativement à la moyenne du champ de vent référence. Il ressort de cet histogramme que 49 % des écarts types sont inférieurs à 20 %. On remarque sur cet histogramme la décroissance rapide de la probabilité d'apparition de grands écarts types. Les écarts types sont concentrés en dessous de 30 %. Des valeurs de l'écart type, allant jusqu'à 20 %, représentent une erreur raisonnable dans le champ de vent synthétisé. Les valeurs plus grandes que 50 % représentent de grandes erreurs. Ces grandes valeurs appairassent, essentiellement, pour de faibles valeurs de la vitesse du vent où une petite erreur absolue sur la vitesse représente une grande erreur en relatif (par rapport à la moyenne).

On s'intéresse maintenant à la répartition géographique de ces erreurs. La figure 6.12 montre la moyenne de l'erreur. Pour chaque point de la grille, on a représenté la valeur absolue de l'erreur (figure 6.12(a)) et le pourcentage de cette erreur par rapport à la moyenne réelle, en utilisant les données réfé-

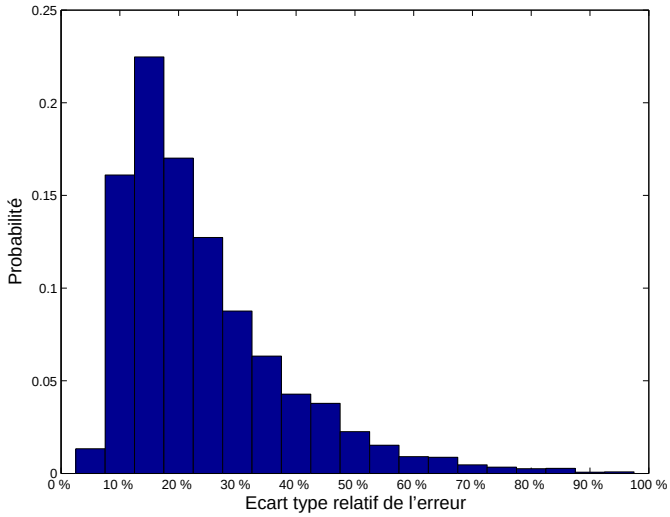
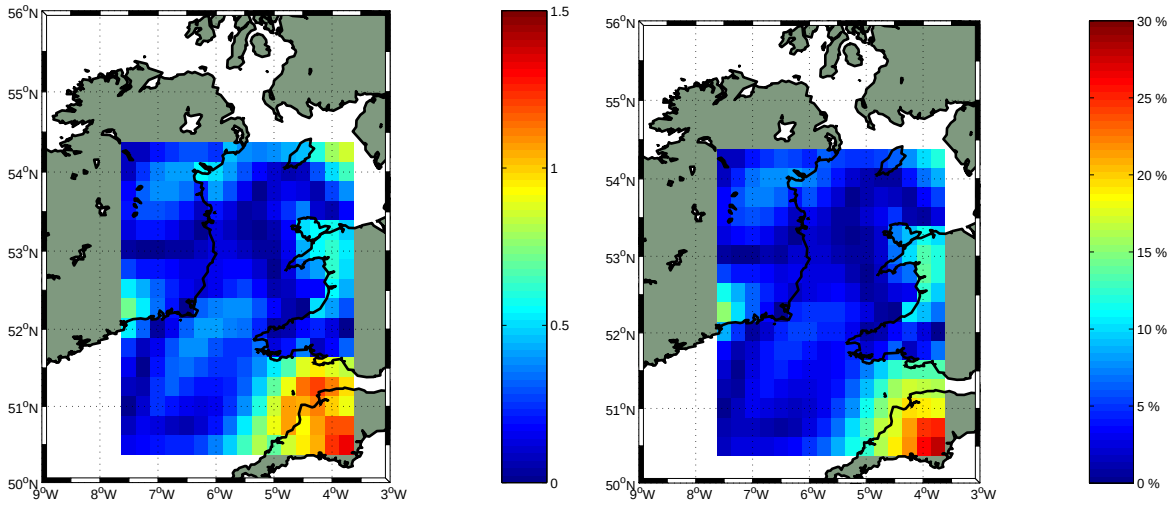


FIG. 6.11 – Distribution de l'écart type relatif de l'erreur. 49 % des écarts types relatifs sont entre inférieurs à 20 %. 74 % sont inférieurs à 30 %. L'histogramme est tracé avec des intervalles de 5 % pour l'écart type relatif de l'erreur.

rence (figure 6.12(b)). On remarque que les valeurs les plus élevées (supérieures à 0,5 m s^{-1}) de l'erreur moyenne se situent sur les pixels couvrant la terre. Les valeurs maximales de l'erreur sont atteintes pour la région la plus au sud et à l'est couverte par les données synthétisées. Cette erreur peut s'expliquer par l'utilisation des données NCEP/NCAR dont la résolution (1,9°×1,9°) ne permet pas de restituer le passage terre/mer. En effet, la maille NCEP/NCAR, couvrant cette zone, contient une grande partie de mer. Les vitesses sur mer étant plus élevées, ceci conduit à une surestimation de la moyenne du vent sur terre. La figure 6.13 montre la différence entre la moyenne des grilles NCEP/NCAR et celle des grilles ECMWF. Cette figure confirme que l'erreur sur les pixels couvrant la terre vient de la nature des données NCEP/NCAR. On remarque que la méthode de fusion a permis de réduire l'erreur moyenne maximale de 3 m s^{-1} à 1,5 m s^{-1}. L'application de la méthode de fusion avec les mesures diffusométriques et le mesures ROS ne posera pas ce genre de problème car l'interface terre/mer ne sera pas présente dans ce genre de données. En effet, ces deux sources de données couvrent uniquement la mer. A l'intérieur d'une « empreinte » de mesure diffusométrique (résolution spatiale de 25km) il n'existe pas de phénomène comparable à l'interface terre/mer existant à l'intérieur d'une maille de la grille de réanalyse NCEP/NCAR. En pleine mer, on remarque une erreur moyenne assez faible (inférieure à 0,5 m s^{-1}).

Dans cette section, on s'est intéressé à une application de la méthode de fusion pour l'étude de la qualité de la série temporelle synthétisée. Dans la section suivante on s'intéresse à la qualité des paramètres de Weibull calculés à partir de ces vitesses. On étudie, en outre l'influence du choix de la grille à haute résolution spatiale pour chaque classe. Ceci est fait en réitérant l'application de la méthode de fusion avec différents sous-ensembles de grilles ECMWF simulant le jeu de données à haute résolution spatiale.



(a) Erreur moyenne absolue des champs de vent synthétisés (en m s^{-1}).

(b) Erreur moyenne relative des champs de vent synthétisés.

FIG. 6.12 – Distribution spatiale de l'erreur moyenne des champs de vent synthétisés.

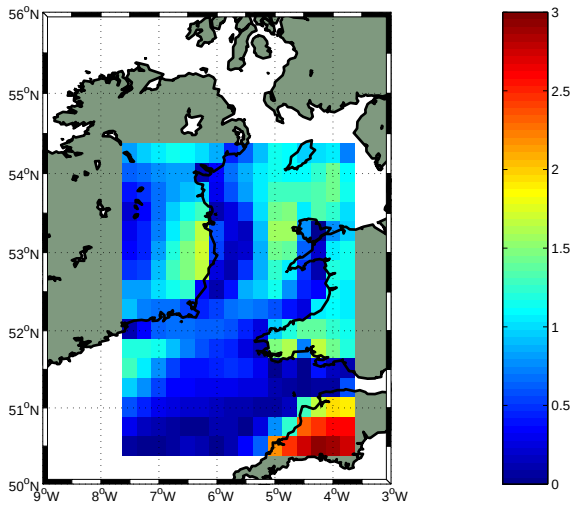
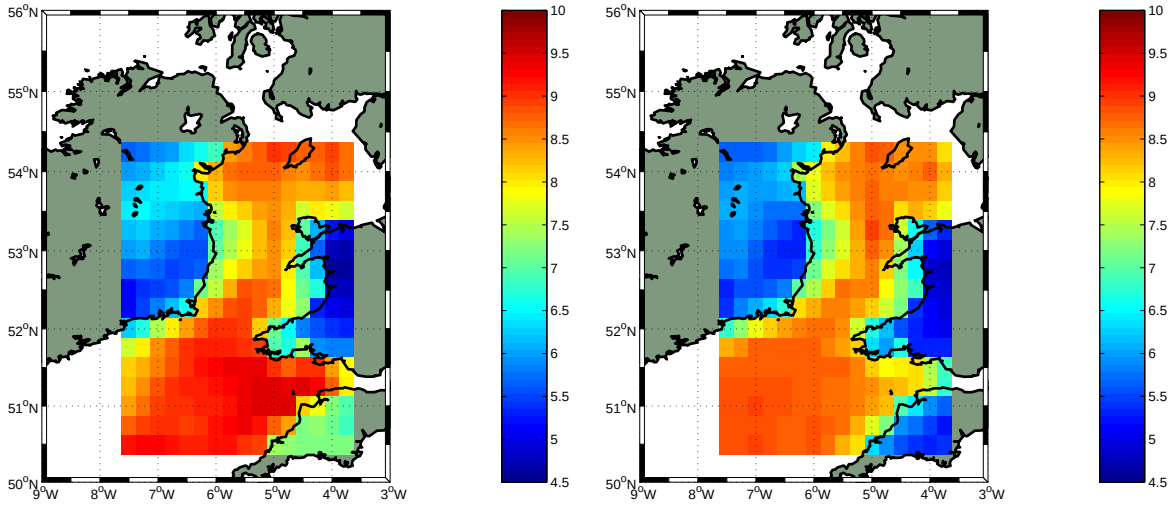


FIG. 6.13 – Différence moyenne entre les grilles de réanalyse NCEP/NCAR interpolées et les grilles de réanalyse ECMWF (en m s^{-1}). Notez le changement d'échelle pour les vitesses de vent par rapport à la figure 6.12.

6.3.2 Erreur sur l'estimation des paramètres statistiques

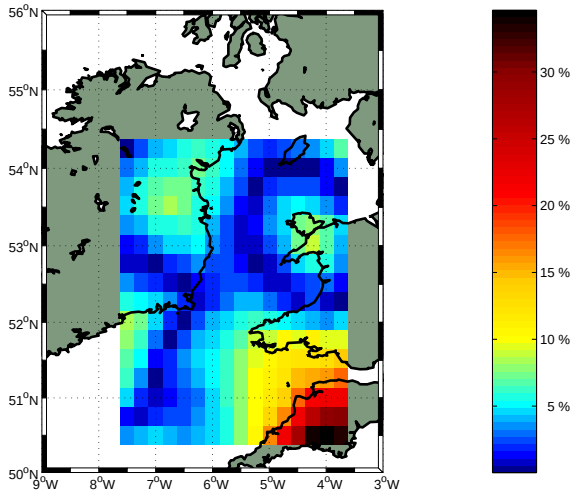
Pour l'estimation des paramètres A et k, on utilise l'estimateur de maximum de vraisemblance présenté dans le chapitre 2 et dont la théorie est décrite dans l'Annexe B. La figure 6.14 représente la cartographie du paramètre A de Weibull estimé à partir des données synthétiques et des données référence ainsi que la carte de l'erreur relative contenue dans la cartographie à partir des données synthétiques. On remarque sur cette figure que les structures spatiales présentes dans la carte du paramètre A estimé à partir des données références (figure 6.14(b)) sont bien restituées dans la carte estimée à partir des données synthétiques (figure 6.14(a)).

La carte de l'erreur relative $\frac{(A^{synth}-A^{ref})}{A^{ref}}$ indique que l'erreur est essentiellement concentrée près des côtes. Ceci est particulièrement le cas pour la région la plus au sud et à l'est. Ceci peut s'expliquer par la nature des données NCEP/NCAR dont la résolution ne permet pas d'intégrer le passage terre/mer. Ainsi, les valeurs de vitesse de vent données pour ce point dans les grilles NCEP/NCAR interpolées sont caractérisées par une surestimation de la vitesse sur terre. Ceci conduit donc à une surestimation du paramètre A sur les points situés près de la terre dans les données synthétisées. Excepté ces points, l'erreur sur la région étudiée se situe à un seuil inférieur à 10 %.



(a) Valeurs du paramètre A à une altitude de 10 m calculées à partir des données synthétisées.

(b) Valeurs du paramètre A à une altitude de 10 m calculées à partir des données référence.

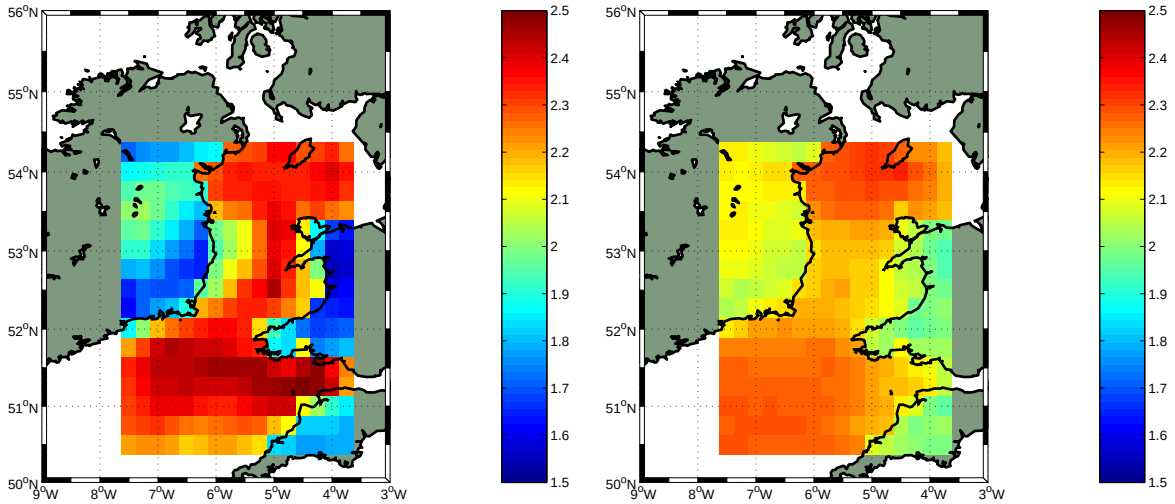


(c) Erreur commise sur le paramètre A en utilisant les données synthétisées.

FIG. 6.14 – Estimation du paramètre A à partir des données synthétisées et des données références et l'erreur relative commise en considérant les données synthétisées

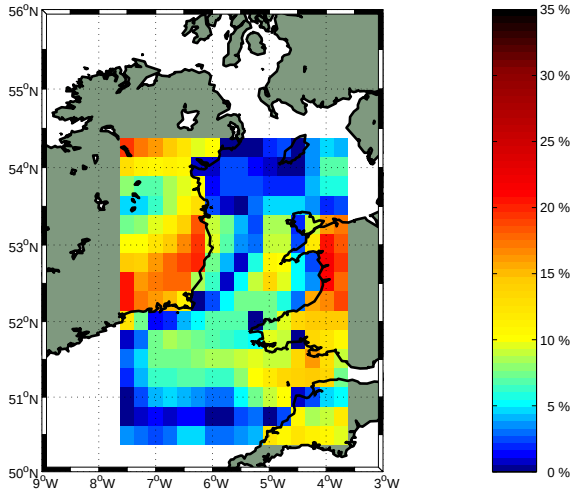
La figure 6.15 représente la cartographie du paramètre k de Weibull estimé à partir des données synthétiques, la cartographie de ce paramètre estimé à partir des données référence et l'erreur relative sur ce paramètre commise lors de la cartographie à partir des données synthétisées. On remarque que la

carte établie à partir des données synthétiques (figure 6.15(a)) introduit plus de structures spatiales que celle établie en utilisant les données référence (figure 6.15(b)). L'erreur maximale sur le paramètre k est de 20 %. Comme indiqué dans le chapitre 2, l'erreur sur le paramètre k a une incidence moins grande qu'une erreur sur le paramètre A sur l'estimation finale de la densité moyenne de puissance.



(a) Valeurs du paramètre k à une altitude de 10 m calculées à partir des données synthétisées.

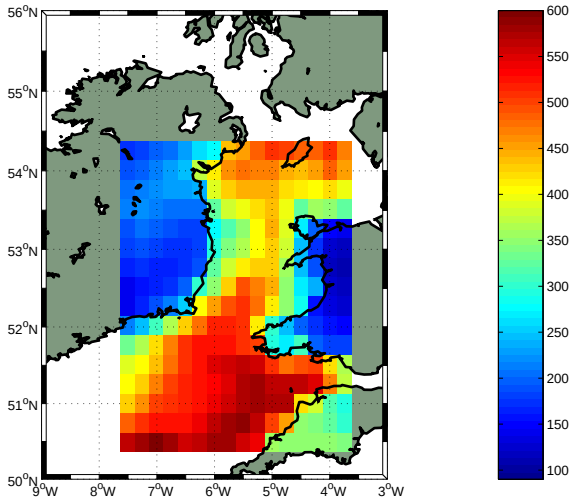
(b) Valeurs du paramètre k à une altitude de 10 m calculées à partir des données référence.



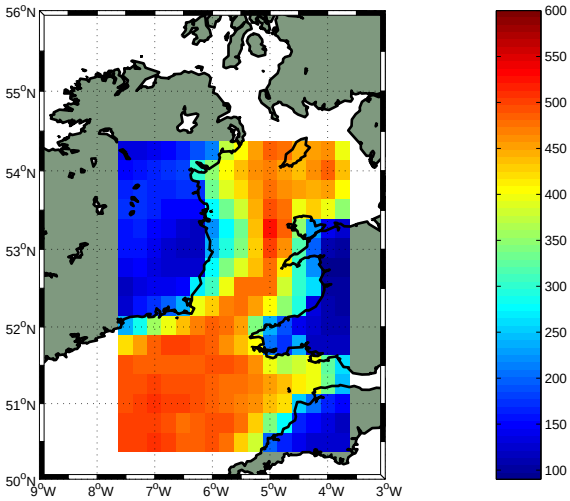
(c) Erreur commise sur le paramètre k en utilisant les données synthétisées.

FIG. 6.15 – Estimation du paramètre k à partir des données synthétisées et des données références et l'erreur relative commise en considérant les données synthétisées

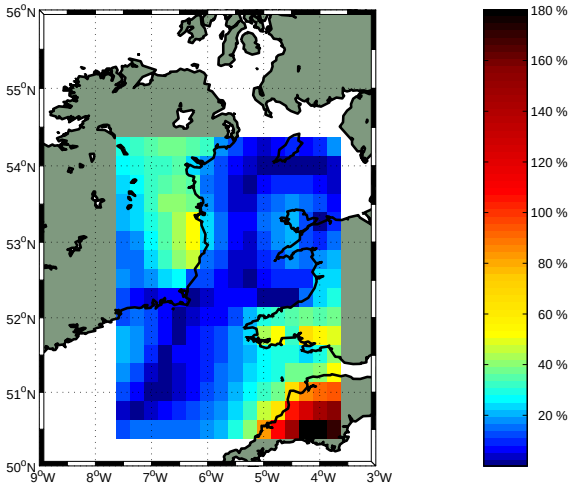
La figure 6.16 représente la cartographie de la densité moyenne de puissance en utilisant les données synthétique (figure 6.16(a)), à partir des données référence (figure 6.16(b)) ainsi que l'erreur relative introduite par l'utilisation des données synthétiques (figure 6.16(c)). L'erreur sur la densité moyenne de puissance présente les même caractéristiques que l'erreur sur le paramètre A de Weibull. Les erreurs sont plus importantes sur la région la plus au sud et à l'est. L'erreur atteint 160 % sur certains points sur terre. Ceci est dû, comme expliqué précédemment, à la nature des données de réanalyse NCEP/NCAR



(a) Estimation de la densité moyenne de puissance en (W m^{-2}) à une altitude de 10 m en utilisant les données synthétisées.



(b) Estimation de la densité moyenne de puissance en (W m^{-2}) à une altitude de 10 m en utilisant les données référence.



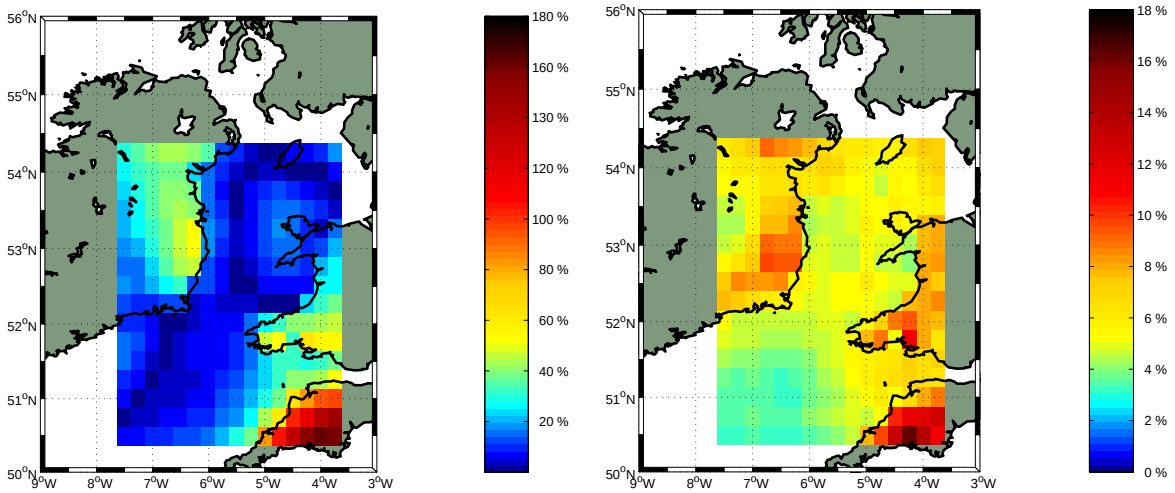
(c) Erreur relative commise sur l'estimation de la densité moyenne de puissance en utilisant les données synthétisées.

FIG. 6.16 – Estimation de la densité moyenne de puissance à partir des données synthétisées et des données référence. L'erreur relative sur cette estimation est donnée.

qui couvrent à la fois la terre et la mer dans une seule maille. En pleine mer, l'erreur n'excède pas les 30 %. [2] relève une RMSE de 251 W m^{-2} pour l'application de WAsP en Irlande. Cette RMSE représente 53,9 % de la moyenne des sites étudiés. En combinant WAsP et KAMM (méthode KAMM/WAsP présentée dans le chapitre 2), [2] relève une RMSE de 91 W m^{-2}. Cette RMSE représente 23 % de la moyenne des densité de puissance des sites considérés dans [2]. Pour comparer nos résultats à ceux de [2], nous calculons la RMSE des densités de puissance estimées en utilisant les données synthétiques. La RMSE est de 74 W m^{-2}. Cette RMSE représente 23 % de la moyenne des densités de puissance de la région couverte par la carte. En ne considérant que les points en mer, la RMSE est de 69 W m^{-2}, ce qui représente 16 % des densités moyenne de puissance en mer. La valeur absolue de la RMSE ne diminue pas d'une manière sensible en considérant que les points en mer. Ceci s'explique par des densités

moyennes de puissance plus faible sur terre que sur mer. La valeur relative de la RMSE diminue quant à elle de manière sensible.

Pour l'étude de l'influence du sous-ensemble de données ECMWF utilisé lors de la fusion sur les résultats, on réitère l'application de la fusion plusieurs fois avec, pour chaque itération, un sous-ensemble de grilles de réanalyse ECMWF différent. On choisit un nombre d'itérations de 100. Chaque classe regroupant au maximum 565 champs de vent, 100 itérations permettront de parcourir un nombre significatif de ces champs de vent. L'erreur en zone côtière est systématique et n'est pas due à un choix particulier du sous-ensemble de grilles ECMWF simulant les données ROS. La figure 6.17 montre que la moyenne des erreurs, sur les 100 itérations, est plus élevée sur terre et sur les zones côtières. L'écart type de l'erreur (figure 6.17(b)) est faible en mer et ne dépasse pas les 8 %. Ceci montre que le choix de la grille ECMWF pour chaque situation type n'influe que légèrement sur le résultat final. L'essentiel de l'erreur vient des caractéristiques des mesures à basse résolution spatiale. L'écart type est plus élevé sur terre. Ceci montre une sensibilité plus grande sur ces zones que sur mer au sous-ensemble de grilles à haute résolution spatiale utilisées lors de la fusion.



(a) Moyenne de l'erreur relative sur la densité moyenne de puissance.

(b) Ecart type de l'erreur relative sur la densité moyenne de puissance.

FIG. 6.17 – Statistiques de l'erreur sur la densité moyenne de puissance en considérant 100 itérations du chois alétaoire des représentats ECMWF.

6.4 Conclusions

Dans ce chapitre, nous avons étudié la qualité de la méthode de fusion. Pour pouvoir disposer de données de référence, nous avons transposé le problème dans les échelles et utiliser des données de réanalyse pour simuler les données d'origine (mesures diffusométriques et mesures ROS). Les données de réanalyse NCEP/NCAR ont été utilisées pour simuler les mesures diffusométriques (haute résolution temporelle et basse résolution spatiale). Un sous-ensemble des données de réanalyse ECMWF a été utilisé pour simuler les mesures ROS (haute résolution spatiale et basse résolution temporelle). L'ensemble

complet des données de réanalyse ECMWF (haute résolution spatiale et haute résolution temporelle) a été utilisé comme référence à laquelle ont été confrontés les résultats de la fusion.

Deux aspects de la qualité des produits de la fusion ont été étudiés : la qualité des vitesses de vent produites et la qualité des paramètres statistiques. Pour la synthèse de la haute résolution spatiale, on a utilisé la méthode UWT-M1 relevant du concept ARSIS. Malgré la simplicité de cette méthode, les résultats sont bons (RMSE de 1,6 m s^{-1}) comparés aux erreurs généralement admises pour les mesures de vent par télédétection (RMSE de 2 m s^{-1}). Il a été montré que les erreurs les plus importantes dans les produits fusionnés sont dues à l'interface terre/mer. Ce problème ne se posera pas lors de l'application de la méthode aux mesures diffusométriques et aux mesures ROS. Les paramètres de Weibull évalués à partir des données fusionnées permettent de restituer les structures spatiales à haute résolution caractérisant la densité moyenne de puissance sur la région. Les erreurs se concentrent sur la zone côtière. Ceci est du, comme évoqué précédemment, aux erreurs à l'interface terre/mer introduites par les données de réanalyse NCEP/NCAR.

Le ratio des résolutions spatiales entre les produits de réanalyse fusionnés est de 8. Le ratio entre les résolutions spatiales des mesures diffusométriques et des mesures ROS est 64. L'effet de ces ratio différents n'a pas pu être étudié. Cependant, l'application de la synthèse de la haute résolution spatiale à quelques cas (présentés dans le chapitre 5) montre des résultats consistants avec les résultats de l'évaluation de la qualité de la méthode de fusion appliquée aux données de réanalyse. Ceci laisse à penser que la différence des ratios des résolutions spatiales a une influence faible sur la validité des résultats.

Pour la validation de l'application de la méthode de fusion, que nous avons proposée, avec les mesures diffusométriques et les mesures ROS, il est possible d'utiliser les données de production du parc éolien offshore d'Arklow Bank, situé au large des côtes irlandaises. Malheureusement, nous n'avons pas eu accès à ces données. Ces données permettraient de valider les résultats de la fusion en un point.

Bibliographie

[1] C. Thomas et L. Wald. Assessment of the quality of fused products. In *Proceedings of the 24th EARSeL Symposium "New Strategies for European Remote Sensing", 25-27 May 2004, Dubrovnik, Croatia, Oluic (Eds.), Millpress, Rotterdam, ISBN 90 5966 003 X*, pp. 317–325, 2005.

[2] H. P. Frank, O. Rathmann, N. G. Mortensen et L. Landberg. The NumericalWind Atlas - the KAMM/WAsP Method. Technical report, Report Risø-R-1252((EN), Risø National Laboratory, Roskilde, Denmark, 60 p., 2001.

Chapitre 7

Conclusions et perspectives

Pour un choix judicieux des sites d'implantation de parcs éoliens, l'industrie éolienne a besoin d'une cartographie à haute résolution spatiale du potentiel éolien offshore. Pour établir cette cartographie, on a besoin de données sur le vent alliant une haute résolution spatiale à une haute résolution temporelle. La haute résolution temporelle est nécessaire pour garantir la précision des estimations du potentiel éolien. La haute résolution spatiale est nécessaire pour détecter les variations du potentiel éolien à l'échelle du parc éolien. La revue des sources de données sur le vent mer, présentée au chapitre 2, a montré qu'aucune source de donnée ne permet, à elle seule, d'accéder à des cartes de vent (mesures du vent à différents points d'une grille à un instant donné) à haute résolution spatiale et à haute résolution temporelle. Certaines mesures, faites sur site, ne permettent d'accéder qu'à des informations sur le vent sur une zone limitée autour du point de mesure. Certains instruments de télédétection, spatioportés, permettent d'accéder à des cartes de vent. Cependant, aucun de ces instruments ne permet d'accéder à des cartes de vent à haute résolution spatiale et à haute résolution temporelle. Il est donc nécessaire de recourir à des modèles ou à l'utilisation combinée de différentes données. Deux types principaux de modèles existent : les modèles statistiques et les modèles physiques. Les modèles physiques dont la résolution répond aux besoins de l'industrie éolienne offshore (résolution de l'ordre du kilomètre) sont encore dans le domaine de la recherche et certaines données nécessaires à leur bon fonctionnement ne sont pas disponibles. Les modèles statistiques, actuellement utilisés, se basent sur des mesures sur site, acquises pendant un an, et une régression entre ces mesures et des mesures sur une station météorologique proche couvrant une plus longue durée. Ces modèles n'estiment le potentiel éolien que sur une zone limitée autour du point de mesure. En outre, ils obligent l'investisseur à attendre pendant la durée de la campagne de mesure.

La méthode proposée dans cette thèse tire profit de la haute résolution spatiale de certaines données de télédétection et de la haute résolution temporelle d'autres données de télédétection. Cette méthode s'inscrit dans le cadre de la fusion de donnée. En combinant les deux sources de données, la méthode génère une série synthétique de données alliant la haute résolution spatiale et la haute résolution temporelle. La méthode se base sur l'occurrence de configurations similaires de vent sur une région donnée. Ces configurations similaires sont appelées situations types. L'hypothèse utilisée dans cette thèse est la conservation de la similarité au travers des échelles, c'est à dire que si deux champs de vent sont similaires à basse résolution spatiale, ils sont similaires à haute résolution spatiale. Cette hypothèse est valide

pour les plages de résolution spatiales considérées. Donc pour les différentes configurations de champ de vent appartenant à la même situation type les structures à haute résolution spatiale sont similaires. Il suffit donc pour chaque situation type d'avoir une mesure des structures de vent à haute résolution spatiale pour inférer les structures à haute résolution spatiale pour tous les champs de vent appartenant à cette situation type. Les structures à haute résolution spatiale sont extraites des données de télédétection à haute résolution spatiale et sont injectées dans les données à basse résolution spatiale et haute résolution temporelle.

Pour l'établissement des situations types, on a élaboré une méthode de classification de champs de vent. Cette méthode permet de regrouper les champs de vent ayant un comportement spatial similaire. La classification permet de résumer l'information sur la variation spatiale contenue dans les champs de vent en quelques situations types. Les méthodes de classification de champs de vent, décrites dans la littérature, ont des objectifs différents de l'objectif recherché dans le cadre de la fusion ou sont faites à des échelles différentes. Nous avons, à partir des méthodes existantes, proposé une méthode de classification adaptée à notre méthode de fusion de données. L'application de la méthode de classification sur un cas d'étude, présentée dans le chapitre 4, montre que la majeure partie de la variation spatiale des champs de vent peut être représentée par quelques situations types. Pour chacune de ces situations types, on associe un champ de vent mesuré à haute résolution spatiale. Les structures présentes dans ce champ de vent sont ensuite injectés dans tous les champs de vent appartenant à la même situation type. La synthèse de la haute résolution spatiale, qui est constituée par cette injection de structures à haute résolution spatiale dans les champs de vent à basse résolution spatiale, est présentée dans le chapitre 5.

L'application complète de la méthode de fusion a été effectuée, avec d'autres données, et à d'autres résolutions spatiales que celles des données de télédétection. Cette application est présentée au chapitre 6. On a choisi ces données pour pouvoir disposer d'une référence lors de l'évaluation de la méthode de fusion. Il ressort de cette évaluation que les résultats de la fusion sont satisfaisants, malgré la simplicité de la fonction de transfert de la basse à la haute résolution spatiale utilisée. Ce modèle considère que les structures à haute résolution spatiale sont exactement identiques pour tous les champs de vent appartenant à la même situation type et ne fait aucun ajustement de ces structures.

Dans cette thèse, nous avons montré la pertinence de l'utilisation de la fusion de données de télédétection pour la cartographie du potentiel éolien offshore. La méthode de fusion de données a été conçue pour être indépendante de la zone géographique prospectée. L'application de cette méthode est, toutefois, limitée par la disponibilité de mesures diffusométriques. Ces mesures ne sont présentes que loin des côtes (à 25 km des côtes dans le cas de Quikscat). Leur extrapolation est donc nécessaire pour pouvoir appliquer la méthode de fusion en zone côtière. Cependant plusieurs questions restent encore en suspens.

Le choix du nombre de situations types est fait dans notre méthode de classification de manière automatique. Ce nombre est choisi de telle sorte qu'il sépare les données en un maximum de classes tout en préservant une représentativité minimale pour chaque classe. La représentativité d'une classe est son taux d'occurrence. Ce choix est justifié par la volonté de séparer au maximum les classes pour garantir que les éléments d'une même classe sont similaires et par la volonté de réduire le nombre de classes. Nous n'avons pas étudié, ici, l'influence du nombre de classes sur le résultat de la fusion. Il

serait intéressant de modifier les critères de sélection du nombre de classes pour étudier l'impact que peut avoir un tel changement. En pratique, le nombre de classes est limité par le nombre de données à haute résolution spatiale disponibles.

La méthode de classification retenue dans cette thèse est une méthode qui s'intéresse à la similarité entre les comportements spatiaux des champs de vent. Une autre approche peut être considérée. Elle consiste à comparer l'évolution dans les échelles des champs de vent. On pourrait comparer les coefficients d'ondelettes issus d'une analyse multi-échelles des champs de vent au lieu de comparer les champs de vent eux-mêmes. A priori, les deux approches sont équivalentes. En effet, l'hypothèse de la conservation de la similarité au travers des échelles garantit que si deux champs de vent sont similaires les structures à plus haute résolution sont similaires.

Pour la synthèse de la haute résolution spatiale, on a appliqué le modèle identité. D'autres modèles existent et ont été présentés au chapitre 3. L'application de modèles plus élaborés peut améliorer les résultats de la méthode de fusion.

Pour la validation de la méthode, on a utilisé des données de réanalyse. Nous n'avons pas pu avoir accès à la série temporelle des données de vent du parc d'Arklow Bank situé en Mer d'Irlande. Une comparaison des résultats de l'application de la méthode de fusion à haute résolution spatiale (avec les données de télédétection) avec les données de production du parc serait un moyen de validation, même si cette comparaison ne porte que sur un point. Pour une validation plus extensive de la méthode de fusion, la méthode de fusion peut être appliquée sur une zone où une grille de bouées est installée. La validation de la méthode de fusion pour la cartographie du potentiel éolien, basée sur la transposition de l'application de la méthode à des résolutions spatiales plus basse, a montré que la qualité de la cartographie est comparable à la méthodes KAMM/WAsP. Les résultats sont meilleurs que la méthode KAMM/WAsP si seule la zone offshore est considérée.

Pour l'application complète de la méthode aux données diffusométriques et données ROS, il est nécessaire d'avoir des mesures diffusométriques sur une grille régulière, constante dans le temps. De tels champs de vent peuvent être produits en utilisant une interpolation appropriée. Pour arriver à une cartographie du potentiel éolien en zone côtière, il est, en outre, nécessaire d'extrapoler les mesures diffusométriques près des côtes. Ces opérations d'extrapolation et d'interpolation doivent tenir compte des caractéristiques physiques du phénomène considéré. Nous nous sommes concentrés dans cette thèse sur la méthode de fusion. Une étude des opérations d'interpolation et d'extrapolation des champs de vent doit être entreprise pour arriver à la cartographie finale du potentiel éolien à haute résolution spatiale.

Annexe A

Mesures satellitaires du vent : physique de la mesure

La télédétection ne permet pas d'accéder directement aux champs de vent. Elle permet cependant d'y accéder à travers leurs conséquences sur la surface de la mer (diffusomètres et ROS) ou sur les aérosols (LIDAR et SODAR). On commence donc par l'étude de la surface de la mer et la manifestation du vent sur celle-ci. Ensuite, on décrit la physique de la mesure des intruments de télédétection.

A.1 Surface de la mer

La surface marine subit des déformations causées par différents phénomènes : l'action du vent, les interactions terre–lune et terre–soleil, les séismes, ... Ces déformations se manifestent par la présence d'ondes à cette surface : les vagues. Ces ondes ont différentes échelles spatiales et temporelles selon les processus mis en jeu pour leur formation. Les ondes se propageant à la surface de la mer appartiennent à deux grandes catégories [1] :

- ondes longues comme les ondes de marée par exemple ;
- ondes de surface comme la houle.

Sur la figure A.1, on peut voir la classification des vagues faite par [2]. Notre intérêt se porte sur les vagues pouvant représenter le vent.

Le vent soufflant à la surface de l'eau exerce sur la pellicule d'eau superficielle une « force de frottement » qui dépend de la densité de l'air et de la vitesse du vent [1]. Elle dépend également du caractère lisse ou rugueux de la mer. Un autre paramètre décrivant l'impact du vent sur la mer est la distance (et par conséquence la durée) sur laquelle le vent a soufflé sans rencontrer d'obstacle [3]. Cette distance est appelée *fetch*.

Les vagues ainsi générées se propagent à la surface de la mer. Ces vagues peuvent être modélisées par la superposition d'ondes monochromatiques. Ces ondes sont de la forme :

$$h(t,x) = A\cos\left(\Omega t - Kx + \phi\right) \tag{A.1}$$

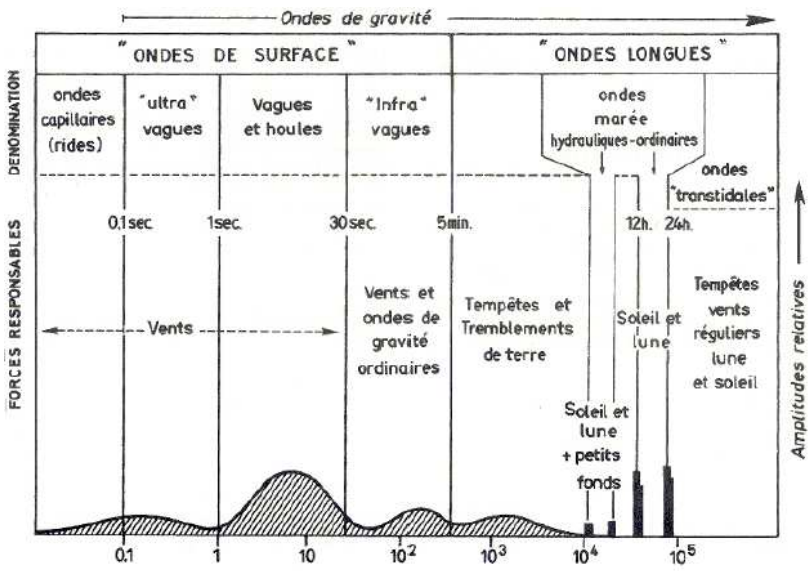


FIG. A.1 – Classification des vagues selon leur période et les phénomènes les générant. Selon [2]

h étant la hauteur de la mer par rapport au niveau moyen. La résolution de l'équation de dispersion pour cette onde, en supposant la densité de l'air négligeable par rapport à celle de l'eau, donne la vitesse de phase suivante [1] :

$$c^2 = (\frac{g\lambda}{2\pi} + \frac{2\pi}{\lambda\rho}T) \tag{A.2}$$

avec : ρ : masse volumique de l'eau
g : la constante de gravité
λ : la longueur d'onde de la vague
T : la tension superficielle de l'eau

Cette vitesse passe par un minimum pour $\lambda_m = 2\pi\sqrt{\frac{T}{g\rho}}$. En effectuant les calculs pour les valeurs usuelles de $T = 74\ 10^{-3}$N.m^{-1}, $g = 9,81$m.s^{-1} et $\rho = 10^3$kg.m^{-3} on obtient $\lambda_m = 1,73$ cm ce qui correspond à une vitesse $c_m = 23,2$ cm.s^{-1}.

On peut ainsi dégager trois catégories de vagues selon le processus qui régit leur évolution [1] :

– $\lambda < \lambda_m/3$ = 0,57 cm : vagues capillaires ;
– $\lambda_m/3 < \lambda < 3\lambda_m$: vagues de transition capillarité–gravité ;
– 5,19 cm = $3\lambda_m < \lambda$: vagues de gravité.

C'est donc les vagues de capillarité et celle de transition qui peuvent nous indiquer le comportement du vent près de la surface de la mer.

A.2 Physique de la mesure des radars actifs

On s'intéresse, dans ce paragraphe, aux différents mécanismes d'interactions entre les vagues et le rayonnement émis par les radars.

A.2.1 Équation radar

Un radar émet une onde électromagnétique de puissance P_e en direction de la cible. Une partie P_r de ce signal est renvoyée vers le radar. P_r dépend des paramètres logistiques caractérisant l'instrument, de l'onde émise et de la surface visée. L'équation radar permet de dissocier le rôle de la surface dans la rétrodiffusion des autres paramètres [4].

La grandeur mesurée par le radar est la section efficace de rétrodiffusion. Elle est définie comme [5] :

$$\sigma = \lim_{R \to +\infty} 4\pi R^2 \frac{E_r^2}{E_i^2} \tag{A.3}$$

avec : R : la distance radar—cible
E_i : le champ électrique incident, émis
E_r : le champ électrique réfléchi, mesuré

σ est l'aire qui, interceptant la même puissance et la diffusant de façon isotropique produit le même écho que celui observé [4]. Cette grandeur dépend donc de la nature de diffusion de la surface observée.

De façon générale, on utilise la section de rétrodiffusion normalisée définie par :

$$\sigma_0 = \int_S \sigma dS \tag{A.4}$$

L'équation radar relie la puissance reçue aux paramètres de l'instrument et à σ_0 [6]. Selon [6], cette équation est de la forme :

$$P_r = \frac{\lambda^2}{(4\pi)^3} \iint \frac{P_e G_e G_r \sigma_0}{R^4} dS \tag{A.5}$$

avec : P_r : la puissance reçue
λ : la longueur d'onde
P_e : la puissance émise
G_e : le gain de l'antenne à l'émission
G_r : le gain de l'antenne à la réception
R : la distance cible-radar

Dans le cas des satellites d'observation de la Terre, l'émission et la réception se font sur la même antenne. Les gains en émission et en réception sont donc identiques. Pour résoudre l'équation radar, on peut supposer que σ_0 ne varie pas sur la zone d'intérêt [7] :

$$\sigma_0 = \frac{(4\pi)^3 R^4 P_r}{\lambda^2 G^2 S P_e} \tag{A.6}$$

Selon [7], σ_0 est exprimée en dB, c'est à dire en $10\log(\sigma_0)$.

[5] signale que la résolution exacte des équations de rétrodiffusion est impossible pour des surfaces telles que l'océan. Le recours à des approximations est donc nécessaire. Le domaine de validité de ces approximations varie selon l'angle d'incidence. La rétrodiffusion n'est pas de même nature selon que les mesures soient faites au nadir (incidence normale) ou loin de la normale. Sur la figure A.2 on peut voir les différents phénomènes prépondérants selon l'angle d'incidence. Ainsi, dans le domaine utilisé par les instruments de télédétection, on distingue deux cas particuliers qui sont la réflexion spéculaire et la rétrodiffusion de Bragg.

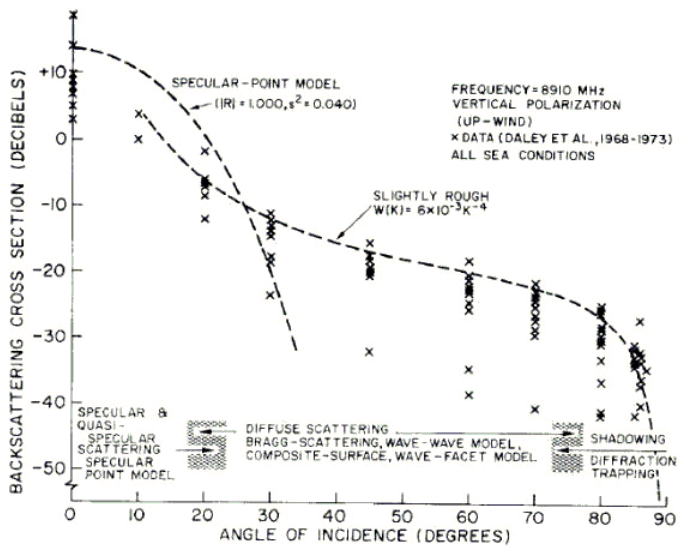


FIG. A.2 – Phénomène de rétrodiffusion prépondérant selon l'angle d'incidence. Selon [5]

A.2.2 Réflexion spéculaire

La réflexion spéculaire se fait par les facettes de la surface normale à l'onde incidente. La surface observée est ainsi modélisée par les plans qui lui sont tangents. Cette approximation est valide si le rayon de courbure de la surface est plus grand que la longueur d'onde émise. Ce modèle a été développé à partir des lois de l'optique physique. La condition d'applicabilité de ce modèle est la suivante :

$$\rho k \cos^3 \theta \gg 1 \tag{A.7}$$

avec : ρ : le rayon de courbure de la surface

k : le nombre d'ondes

θ : l'angle d'incidence

Cette condition montre que la validité de ce modèle se situe près d'une incidence normale ($\theta \approx 0$). En effet, plus l'angle d'incidence augmente plus la probabilité de trouver une « facette » normale à l'onde diminue [7]. Ceci est dû à la limitation de la pente des vagues de la mer.

Soit ζ la hauteur de la surface de la mer et $p_s(\zeta_x, \zeta_y)$ la fonction de densité de probabilité des pentes à la surface ($\zeta_x = \frac{\partial \zeta}{\partial x}$) et ($\zeta_y = \frac{\partial \zeta}{\partial y}$).

[8] montre que :

$$\sigma_0 = \pi \sec^4 \theta \; p_s(-\tan\theta, 0) \; |R(0)|^2 \tag{A.8}$$

avec : $R(0)$: Le coefficient de Fresnel de réflexion de la surface à incidence normale

Pour une surface isotropique et de distribution gaussienne, s^2 étant la pente quadratique moyenne, cette équation devient :

$$\sigma_0 = \frac{|R(0)|^2}{s^2} \sec^4 \theta \exp\left(\frac{-\tan^2 \theta}{s^2}\right) \tag{A.9}$$

La réflexion, dans ce modèle, dépend donc de la pente qui est elle même dépendante du vent local. On peut ainsi accéder à l'information sur le vent à travers le coefficient de rétrodiffusion.

[9] indique que ce modèle n'est plus valide pour les angles d'incidence plus grands que 15°. Cependant, [7] indique que ce mécanisme a une contribution non négligeable à la rétrodiffusion jusqu'à un angle d'incidence de 30°. Au delà de cet angle, la rétrodiffusion est décrite par la rérodiffusion de Bragg présentée ci-dessous.

A.2.3 Rétrodiffusion de Bragg

Ce modèle s'appuie sur la méthode des « petites perturbations » [5]. Cette méthode suppose que la surface est faiblement rugueuse [10]. Ce critère rend ce modèle invalide aux incidences normales [10]. [7] affirme que la rétrodiffusion de Bragg est prédominante entre 20° et 70°.

La rétrodiffusion de Bragg consiste en une résonance entre l'onde émise et les vagues à la surface de la mer. Cette résonance a lieu pour les vagues de longueur d'onde :

$$\lambda_v = \frac{n\lambda}{2\sin\theta} \tag{A.10}$$

Cependant, c'est pour $n = 1$ qu'on obtient la résonance dominante [5].

Au premier ordre on peut écrire [5] :

$$\sigma_0(\theta) = 4\pi k^4 \cos^4\theta |g_{ij}(\theta)|^2 W(2k\sin\theta, 0) \tag{A.11}$$

où W est la densité spectrale bidimensionnelle (cartésienne) des nombres d'ondes de la rugosité de surface. Les indices ij indiquent la polarisation de l'onde émise et celle reçue. g_{ij} sont les coefficients de rétrodiffusion au premier ordre.

$$g_{HH}(\theta) = \frac{\epsilon_r - 1}{\left[\cos\theta + (\epsilon_r - \sin^2\theta)^{\frac{1}{2}}\right]^2} \tag{A.12}$$

$$g_{VV}(\theta) = \frac{(\epsilon_r - 1)\left[\epsilon_r(1+\sin^2\theta) - \sin^2\theta\right]}{\left[\epsilon_r\cos\theta + (\epsilon_r - \sin^2\theta)^{\frac{1}{2}}\right]^2} \tag{A.13}$$

où ϵ_r est la permittivité de la mer.

Dans ce modèle de réflexion, à une fréquence donnée du radar correspond une longeur d'onde de la surface marine responsable de l'imagerie. Dans le cas où l'onde émise par le radar est dans le domaine des micro-ondes, la réponse de la mer se fera à travers les vagues dont la longueur d'onde est de quelques centimètres. L'étude faite au paragraphe A.1 a montré que l'évolution de ces ondes est régie par le vent. Le coefficient de rétrodiffusion dépend donc du vent local.

Bibliographie

[1] H. Lacombe. *Cours d'océanographie physique*. Gauthier-Villars, Paris, France, 392 p., 1965.

[2] W. Munk. Origin and generation of waves. In *Proceedings of the first conference on coastal engineering, The engineering Foundation ed., Berkeley*, pp. 1–4, 1951.

[3] D. Hauser, H. Dupuis, X. Durrieu de Madron, C. Estournel, C. Flamant, J. Pelon et P. Queffeulou. La campagne FETCH. *La Météorologie*, vol. 8, n° 29, pp. 14–31, 2000.

[4] M. C. Girard et C. M. Girard. *Traitement des données de télédétection*. Dunod, Paris, France, 350 p., 1999.

[5] G. R. Valenzuela. Theories for the interaction of electromagnetic and ocean waves - a review. *Boundary Layer Meteorology*, vol. 13, pp. 61–85, 1978.

[6] A. K. Fung et F. T. Ulaby. *Manual of Remote Sensing - Second edition - Volume I : Theory, Instruments and Techniques, Ed. : D.S. Simonett*, pp. 115–164. American Society of Photogrammetry, Falls Church, Va, USA, 1983.

[7] Ad. Stoffelen. *Scatterometry - De scatterometer*. Thèse de doctorat, Université d'Utrecht, Pays-Bas, 209 p., 1998.

[8] D. E. Barrick. Rough surface scattering based on the specular point theory. *IEEE Transactions on Antennas and Propagation*, vol. AP-16, n° 4, pp. 449–454, 1968.

[9] D. E. Barrick. Wind dependence of quasi-specular microwave sea scatter. *IEEE Transactions on Antennas and Propagation*, vol. AP-22, n° 1, pp. 135–136, 1974.

[10] V. Kerbaol. *Analyse spectrale et statistique vent–vagues des images radar à ouverture synthétique (ROS) — Application aux données des satellites ERS–1/2*. Thèse de doctorat, IFREMER, Brest, France, 185 p., 1997.

Annexe B

Estimateur de maximum de vraisemblance et intervalles de confiance

Dans cette annexe, on présente le principe de l'estimateur de maximum de vraisemblance utilisé dans cette thèse pour estimer les paramètres de Weibull A et k à partir des mesures de vitesse du vent. Cet estimateur permet, en outre, d'associer des intervalles de confiances aux estimations des paramètres de Weibull.

B.1 L'estimateur de maximum de vraisemblance

L'estimateur de maximum de vraisemblance est basé sur le calcul de la probabilité jointe des paramètres A, k et des observations. L'expression de cette probabilité est donnée dans l'équation B.1.

$$L(k, A, v_1, ..., v_n) = \prod_{i=1}^{n} p(k, A, v_i) \tag{B.1}$$

$L(k, A, v_1, ..., v_n)$ donne la probabilité d'observer la série temporelle de vitesses du vent $(v_1, ..., v_n)$ dans un site caractérisé par une distribution du vent suivant une distribution de Weibull de paramètres A et k. Cette expression suppose que les observations sont indépendantes et identiquement distribuées.

Les valeurs du couple (A, k) maximisant L sont les valeurs les plus plausibles pour les paramètres A et k caractérisant la distribution du vent sur le site. La maximisation de L conduit aux équations suivantes :

$$\widehat{k} = \left[\frac{\sum_{i=1}^{n} v_i^k \ln(v_i)}{\sum_{i=1}^{n} v_i^k} - \frac{1}{n} \sum_{i=1}^{n} \ln(v_i) \right]^{-1} \tag{B.2}$$

$$\widehat{A} = \left[\frac{1}{n} \sum_{i=1}^{n} v_i^k \right]^{\frac{1}{k}} \tag{B.3}$$

Cependant, le couple (A, k) estimé par le maximum de vraisemblance peut différer du couple réel caractérisant la distribution spécialement dans le cas où le nombre d'échantillons est petit.

B.2 Région de confiance jointe associée une estimation

Soit α la probabilité que l'estimation du couple (A, k) soit erronée. La région de confiance avec une probabilité $p = (1 - \alpha)$, est la région de l'espace des A et k contenant la vraie valeur du couple (A, k) avec une probabilité p.

La dimension de la région de confiance renseigne sur les incertitudes liées à l'estimation. Si par exemple, une région de confiance à 99 % couvre seulement une petite partie de l'espace des A et k, ceci signifie que très peu d'incertitude est associée à l'estimation des paramètres A et k avec une probabilité de 0,99. Au contraire, si la région de confiance couvre une grande partie de l'espace des A et k, cela signifie qu'une grande incertitude est associée aux valeurs estimées. Cette incertitude est due aux données.

Une méthode pour déterminer des régions de confiance jointes des paramètres A et k est de calculer le ratio de vraisemblance $R(A, k)$.

Le ratio de vraissemblance est donné par l'équation B.4. Ce ratio est interprété comme suit : avoir un ratio $R(A, k) = 0.1$ signifie qu'il est 10 fois plus probable que les paramètres estimés $(\widehat{A}, \widehat{k})$ soient les « vrais » paramètres qu'il n'est probable que (A, k) le soient.

$$R(A, k) = \frac{L(k, A, v_1, ..., v_n)}{L(\widehat{k}, \widehat{A}, v_1, ..., v_n)} \tag{B.4}$$

Asymptotiquement, $-2 \ln [R(A, k)]$ suit une distribution du Chi-2.

En utilisant cette propriété donnant la distribution de l'expression $-2 \ln [R(A, k)]$, on peut établir une région de confiance jointe sur les paramètres A et k. Une région de confiance à $100(1 - \alpha)\%$ des paramètres (A, k) est l'ensemble des couples (A, k) tels que :

$$-2 \ln [R(A, k)] \leq \chi^2_{(1-\alpha),1} \tag{B.5}$$

c'est-à-dire la région de l'espace des paramètres (A, k) contenant la « vraie » valeur du couple (A, k) avec une probabilité de $(1 - \alpha)$.

FUSION DE DONNÉES SATELLITAIRES POUR LA CARTOGRAPHIE DU POTENTIEL ÉOLIEN OFFSHORE

Résumé

L'énergie éolienne est une des composantes d'une politique énergétique permettant de réaliser un développement durable. Ces dernières années, des parcs éoliens offshore ont été installés. Ces parcs bénéficient d'un vent plus fort et plus régulier en mer que sur terre. Pour un choix judicieux des lieux d'implantation des parcs éoliens, il est nécessaire de disposer d'une cartographie du potentiel éolien. Ces cartes doivent être à haute résolution spatiale pour détecter les variations du potentiel à l'échelle d'un parc éolien. La cartographie du potentiel éolien se fait au travers de la description de la variation spatiale des paramètres statistiques caractérisant la climatologie du vent. Pour une estimation précise de ces paramètres statistiques, il est nécessaire d'avoir des mesures de vitesse et de direction du vent à haute résolution temporelle. Cependant, aucune source de données, actuelle, n'allie la haute résolution spatiale et la haute résolution temporelle. On propose une méthode de fusion de données permettant de tirer profit de la haute résolution spatiale de certains instruments de télédétection (les radars à ouverture synthétiques) et de la haute résolution temporelle d'autres instruments de télédétection (les radars diffusomètres). La méthode de fusion est appliquée à un cas d'étude et les résultats sont évalués. Les résultats montrent la pertinence de la fusion de données pour la cartographie du potentiel éolien offshore.

Mots clés : potentiel éolien offshore, radars à ouverture synthétique, radars diffusomètres, fusion de données, concept ARSIS

REMOTELY SENSED DATA FUSION FOR OFFSHORE WIND ENERGY RESOURCE MAPPING

Abstract

Wind energy is a component of an energy policy contributing to a sustainable development. Last years, offshore wind parks have been installed offshore. These parks benefit from higher wind speeds and lower turbulence than onshore. To sit a wind park, it is necessary to have a mapping of wind resource. These maps are needed at high spatial resolution to show wind energy resource variations at the scale of a wind park. Wind resource mapping is achieved through the description of the spatial variations of statistical parameters characterizing wind climatology. For a precise estimation of these statistical parameters, high temporal resolution wind speed and direction measurements are needed. However, presently, there is no data source allying high spatial resolution and high temporal resolution. We propose a data fusion method taking advantage of the high spatial resolution of some remote sensing instruments (synthetic aperture radars) and the high temporal resolution of other remote sensing instruments (scatterometers). The data fusion method is applied to a case study and the results quality is assessed. The results show the pertinence of data fusion for the mapping of wind energy resource offshore.

Key Words : offshore wind resource, synthetic aperture radars, scatterometers, data fusion, ARSIS concept

Laboratoire d'accueil	:	Centre Energétique et Procédés - Ecole des Mines de Paris Rue Claude Daunesse - B.P. 207 - F-06904 Sophia Antipolis Cedex
Thèse présentée par	:	Mohamed Bassam BEN TICHA **le** : 05 novembre 2007
Discipline	:	« Informatique temps réel, Robotique, Automatique » - Ecole des Mines de Paris