

RAPPORT DE STAGE ÉLÈVE-INGÉNIEUR

MISE EN PLACE D'UNE BASE DE DONNÉES GÉOSPATIALES SUR LES MÂTS ÉOLIENS ET LEURS PARCS DE RATTACHEMENT

Lieu du stage : ABO Wind – Toulouse – France

Dates du stage: 04/09/2023 - 05/01/2024

Nom du service : Informatique

Responsable du service : Aurélie ROQUIER

Stage supervisé par : Marion DUJARDIN



Remerciements

Je tiens à exprimer ma gratitude envers toutes les personnes qui ont participé de manière directe ou indirecte à la réussite de ce stage. Mes remerciements s'adressent à :

- Ma tutrice de stage Mme Marion DUJARDIN, géomaticienne, pour son accueil chaleureux et son soutien tout au long du stage. Elle m'a accordé sa confiance et m'a permis d'apprendre énormément à ses côtés.
- ❖ Tous les membres du service IT notamment Mme Aurélie ROQUIER, M. Laurent CARDONA, Mme Pascaline MOKRANI, M. Teddy YVART et M. Thomas FOURGS pour leur hospitalité qui a facilité mon intégration, leur bonne humeur favorisant une ambiance agréable au sein du bureau, et leur disponibilité qui a grandement contribué à rendre mon expérience professionnelle des plus enrichissantes.
- Mon établissement l'EPF pour l'opportunité offerte par la réalisation de ce stage, ainsi que pour son accompagnement tout au long de la recherche de celui-ci.
- Mes parents ASSELE Séraphin et ASSELE Élise, ainsi que mes nombreux frères et sœurs pour leur assistance et leur soutien indéfectible.
- Mes amis et tous mes proches pour l'accompagnement et les encouragements durant ces quatre mois de stage.
- Les relecteurs et toutes les personnes qui ont apporté leur contribution à l'amélioration de la rédaction de ce rapport de stage.

MISE EN PLACE D'UNE BASE DE DONNÉES GÉOSPATIALES SUR LES MÂTS ÉOLIENS ET LEURS PARCS DE RATTACHEMENT

Résumé

Connaître le contexte éolien d'un site avant de s'y engager revêt une importance cruciale pour la réalisation des parcs éoliens chez ABO Wind. Grâce aux Systèmes d'Information Géographique connectés aux bases de données, l'obtention rapide d'un grand nombre d'informations permet d'accroître l'efficacité des investigations. Cependant, les nombreuses bases de données existantes sur les mâts éoliens et leurs parcs présentent des différences notables. L'objectif principal de ce travail, réalisé en plusieurs étapes, était d'unifier ces bases afin de construire une base de données géospatiales fiable, tout en la mettant à disposition des utilisateurs via une interface. De l'inspection des données recueillies à la visualisation de ces informations, en passant par diverses étapes de traitement et de comparaison, plusieurs outils de gestion et d'analyse de données ont été utilisés.

Mots clés : mât éolien, parc éolien, base de données, données géospatiales, analyse, traitement, visualisation.

Abstract

Knowing the wind context of a site before committing to it is of crucial importance to ABO Wind's wind farm development. Thanks to Geographic Information Systems connected to databases, a large amount of information can be obtained quickly, thus increasing the efficiency of investigations. However, the many existing databases on wind towers and their farms present some notable differences. The main aim of this multi-stage project was to unify these databases in order to build a reliable geospatial database, while making it available to users via an interface. From the inspection of collected data to the visualization of this information, via various processing and comparison stages, several data management and analysis tools were used.

Keywords: wind tower, wind farm, database, geospatial data, analysis, processing, visualization.

Sommaire

Remerciements	2
Résumé	3
Liste des figures :	6
Liste des tableaux :	6
Glossaire	7
INTRODUCTION	10
I- PRESENTATION DE L'ENTREPRISE	11
1. Le groupe ABO Wind	11
1.1. Histoire et chiffres clés du groupe	11
1.2. La filiale française	12
1.2.1. Ses chiffres clés	12
1.2.2. Les concurrents de ABO Wind France	13
2. La structure d'accueil	13
2.1. L'agence de Toulouse	13
2.2. Le service informatique	14
II- PROJET D'ÉTUDE	15
1. Présentation du projet	15
1.1. Le projet éolien	
1.2. La gestion actuelle des données sur les éoliennes chez ABO Wind	16
2. Présentation des outils utilisés	17
3. Présentation du travail effectué	18
3.1. Traitement des données	18
3.2. Remplissage des tables finales	
3.3. Intégration des données dans la base de données PostgreSQL	26
3.3.1. Configuration des champs : nomenclatures et types de données	26
3.3.2. Création de la base de données	27
3.3.3. Intégration de la donnée dans les tables	
3.4. Interfaçage avec QGIS	31
3.4.1. Création des enveloppes	31
3.4.2. Connexion à la base de données sur QGIS	
3.4.3. Mise en forme de la visualisation des données	33
3.4.4. Phase de test	35
3.5. Rédaction des livrables	35

CONCLUSION		37
1. Conclusion technique		37
2. Conclusion personnelle		37
ANNEXE A : Évolution du groupe ABO W	/ind de 1996 à 2021	38
ANNEXE B : Quelques Transformers FMI	E utilisés	39
ANNEXE C : Visualisation globale de que	elques Workflows des traitements FME	40
ANNEXE D : Tableau comparatif des cha	mps des tables entre les BDD	41

<u>Liste des figures :</u>

Figure 1: Organigramme de la filiale française	12
Figure 2: Les domaines d'application des actions du service IT	14
Figure 3: Organigramme du service IT	14
Figure 4: Les étapes du projet éolien chez ABO Wind	15
Figure 5: Schéma de l'inspection des champs sélectionnés dans les BDD	19
Figure 6: Fusion des données des tables "parc"	21
Figure 7: Jointures entre les tables	22
Figure 8: Visualisation des changements	23
Figure 9 : Cas de traitement sans code ICPE	
Figure 10: Autre cas de traitement sans le code ICPE	24
Figure 11: Aperçu du remplissage du classeur des parcs éoliens	25
Figure 12: Aperçu du remplissage du classeur des mâts éoliens	25
Figure 13: Nomenclature des ID des parcs et mâts	26
Figure 14: MCD de la BDD	28
Figure 15: Aperçu du code SQL généré	28
Figure 16: Arborescence de la BDD dans PgAdmin4	29
Figure 17: Intégration des données sur les parcs éoliens	30
Figure 18: Intégration des données sur les mâts éoliens	31
Figure 19: Workflow de la création des enveloppes parcs	32
Figure 20: Configuration de la connexion à la BDD PostGIS	32
Figure 21: Affichage standard des données sur QGIS	
Figure 22: Configuration de la symbologie des mâts dans QGIS	33
Figure 23: Personnalisation de la symbologie des parcs dans QGISQGIS	34
Figure 24: Visualisation après mise en forme de l'affichage des données dans QGIS	34
Figure 25: QR Code du fichier guide du test	35
Figure 26: Aperçu du catalogue de données	36

<u>Liste des tableaux :</u>

Tableau 1: Définitions des termes clés	7
Tableau 2: Significations des acronymes	
Tableau 3: Chiffres clés du groupe ABO Wind	
Tableau 4: Quelques chiffres clés de ABO Wind France	
Tableau 5: Les régions sélectionnées	16
Tableau 6: Outils utilisés	17
Tableau 7: Récapitulatif des tables contenues dans les BDD	20
Tableau 8: Configuration des champs de la table « parc »	27
Tableau 9: Configuration des champs de la table « mat »	27

Glossaire

Les termes clés :

Tableau 1: Définitions des termes clés

Terme	Définition	
Energies renouvelables	Source d'énergie obtenue à partir de ressources naturelles régénératives, telles que l'énergie solaire, éolienne, hydraulique, géothermique, et biomasse, contribuant à la réduction des émissions de gaz à effet de serre.	
Mât éolien	Ou simplement mât, ici sont des appellations de l'éolienne qui est une machine qui utilise l'énergie cinétique du vent pour produire de l'électricité.	
Parc éolien	Ensemble d'éoliennes installées dans une zone spécifique pour produire de l'énergie éolienne à grande échelle.	
Repowering	En français renouvellement, consiste à remplacer tout ou partie d'anciennes infrastructures énergétiques par de nouvelles, plus puissantes et/ou plus efficaces, se traduisant par une augmentation de l'énergie produite.	
Base de données	Ensemble structuré de données organisées et stockées de manière à être facilement consultées, gérées et mises à jour.	
Données géospatiales	Informations liées à la localisation géographique, généralement représentées par des coordonnées géographiques permettant une analyse spatiale.	
Traitement	Ensemble d'opérations appliquées aux données pour les préparer, les analyser ou les transformer de manière à répondre à des objectifs spécifiques.	
Transformers	Dans FME, ce sont des modules qui effectuent des opérations spécifiques sur les données spatiales ou attributaires. Ils peuvent modifier, enrichir, filtrer ou transformer les données de diverses manières. Exemple: « FeatureMerger » (pour la fusion des données) et « FeatureJoiner » (pour la jointure des données).	
Aktiengesellschaft	Forme de société par action en droit des pays germanophones (Allemagne, Suisse et Autriche)	
Reader (lecteur)	Dans FME, c'est un composant qui lit des données spatiales à partir de différentes sources. Il peut s'agir de fichiers, de bases de données, de services Web, etc. Les lecteurs sont utilisés pour extraire des données en début de flux de travail.	
Writer	Dans FME, c'est un composant qui écrit des données spatiales dans un format spécifique ou dans une destination particulière. Il est utilisé pour exporter les résultats du flux de travail vers des fichiers, des bases de données, ou d'autres systèmes de stockage.	
Workflow	Représente l'ensemble des étapes, opérations et transformations effectuées sur les données au sein de FME. Il décrit la séquence d'actions pour atteindre un objectif spécifique dans la manipulation des données.	

Table	Un ensemble de données organisées sous forme d'un tableau où les colonnes (champ) correspondent à des catégories d'information et les lignes à des enregistrements, également appelés entrées.
Géorisques	Site édité par le ministère du Développement durable, il rassemble les informations géographiques sur les risques naturels et technologiques dans un portail national.
PostGIS	Extension géospatiale pour le système de gestion de base de données relationnelle PostgreSQL. Il étend les fonctionnalités de PostgreSQL pour permettre le stockage, la manipulation et l'analyse de données géospatiales.
Encodage	Représentation des caractères textuels sous forme binaire dans un système informatique. Il existe plusieurs standards d'encodage, chacun attribuant une valeur numérique à chaque caractère.

Les acronymes :

Tableau 2: Significations des acronymes

Acronyme	Signification
ABO	Issue des noms Ahn et Bockholt
BDD	Base de données
FME	Feature Manipulation Engine
ETL	Extraction, Transformation, Chargement
SGBDR	Système de gestion de Base de Données Relationnelles
IT	Informatique
SARL	Société à Responsabilité Limité
AG	Aktiengesellschaft
EDF	Electricité de France
QGIS	Quantum Geographic Information System
MCD	Modèle Conceptuel des Données
MLD	Modèle Logique des Données
SQL	Structured Query Language
SIG	Système d'Information Géographique
EPSG	European Petroleum Survey Group
ICPE	Installations Classées pour la Protection de l'Environnement
DREAL	Directions Régionales de l'Environnement, de l'Aménagement et du Logement
ID	Identifiant
Cf	Confer
OREOL	Outil de Référencement des EOliennes
CSV	Comma Separated Values
GPS	Global Positioning System
UTF-8	Universal Character Set Transformation Format - 8 bit

INTRODUCTION

À l'ère de la transition énergétique, les énergies renouvelables s'imposent comme la clé d'un avenir plus durable. Parmi ces sources d'énergie propre, l'énergie éolienne se distingue par son potentiel à la fois puissant et respectueux de l'environnement. Pourtant, la réalisation d'un parc éolien, du concept à la concrétisation, implique un ensemble complexe d'étapes et de paramètres. Par ailleurs, de nos jours, les systèmes d'informations géographiques connectés aux bases de données permettent d'accéder à une multitude d'informations en un simple clic. La possession et la gestion de l'information joue donc un rôle crucial et les bases de données dédiées aux parcs et mâts éoliens se multiplient, chacune apportant sa contribution spécifique. Cependant, cette diversité peut engendrer des défis en termes de cohérence, de fiabilité, et d'efficacité opérationnelle. La nécessité d'une base de données unique et fiable s'est donc fait ressentir au sein de l'entreprise ABO Wind France.

Dès lors, comment concevoir une base de données géospatiales efficace pour centraliser et gérer l'information liée aux parcs et mâts éoliens ?

Pour répondre à ce problème, la première section de ce rapport exposera les contours de l'entreprise ABO Wind. Ensuite, la mission assignée pendant le stage sera présentée, suivie d'une description des tâches effectuées.

I- PRESENTATION DE L'ENTREPRISE

1. Le groupe ABO Wind

ABO Wind couvre l'intégralité de la chaîne de valeur des projets de production d'énergie renouvelable, que ce soit à travers des prestations clés en main ou des prestations spécifiques dédiées. Le groupe s'engage dans la planification, la construction, l'exploitation, et l'entretien de parcs éoliens et solaires, ainsi que dans le développement de projets d'énergie hybride, de stockage de batteries et d'hydrogène.

1.1. Histoire et chiffres clés du groupe

En 1996, Jochen Ahn et Matthias Bockholt fondent la "Planungsgesellschaft zur Nutzung der Windkraft & anderer regenerativer Energien" (Société de développement de l'utilisation de l'énergie du vent et d'autres énergies renouvelables). Par la suite, le nom de l'entreprise, initialement complexe et peu favorable au référencement, a été modifié. L'année 2000 marque une évolution significative alors que l'entreprise se transforme en une société par actions sous le nom de ABO Wind, le terme "ABO" étant un acronyme dérivé des noms de ses fondateurs, Ahn et Bockholt.

L'expansion à l'échelle internationale débute en 2001 avec la création de la filiale ABO Wind España S.A.U. à Valence, en Espagne. En 2002, l'entreprise élargit davantage sa présence mondiale avec l'établissement de sa filiale française. L'entreprise a ensuite ouvert des filiales dans d'autres pays européens, en Amérique, en Afrique et en Asie.

En 2010, ABO Wind franchit une nouvelle étape majeure en créant ABO Invest AG, une société en actions. Cette entité a pour objectif de développer des centrales éoliennes à l'échelle internationale, en collaboration avec des citoyens et des fondations. Cette initiative souligne l'engagement de l'entreprise en faveur de projets éoliens durables et participatifs, renforçant ainsi sa position en tant qu'acteur clé du secteur des énergies renouvelables à l'échelle mondiale.

Tableau 3: Chiffres clés du groupe ABO Wind

Catégorie	Valeur	
Pays d'implantation	16	
Agence	40	
Employés	> 1000	
Projets développés et vendus	> 5 GW	
Ventes en 2022	232 millions d'euros	
Chiffre d'affaires total en 2022	308 millions d'euros	
Fonds propres	170 millions d'euros	
Projets en développement actuellement	> 22 GW	
Production annuelle d'électricité des centrales installées	> 5 milliards de kilowattheures	
Emissions de CO2 évitées annuellement	> 2 millions de tonnes	

1.2. La filiale française

Établie en 2002, ABO Wind France, en tant que SARL, opère à travers quatre agences implantées respectivement à Toulouse, Nantes, Orléans et Lyon. Elle représente actuellement la filiale la plus importante après la société mère allemande, ABO Wind AG.

Fonctionnant selon le modèle de l'entreprise mère, la filiale française applique les pratiques éprouvées d'ABO Wind AG sur une vaste portion du territoire français. En effet, elle intervient sur toute la chaîne de valeur des projets de productions d'énergie renouvelable. Soit sous forme de prestation clés en main, soit sous forme de prestation spécifique dédiée. Ainsi, une équipe dédiée à la conduite de travaux se répartit dans les 4 agences en France. L'équipe veille à la réalisation du chantier dans le respect des normes de qualité, des calendriers et des coûts prévisionnels, dans une démarche de développement durable.

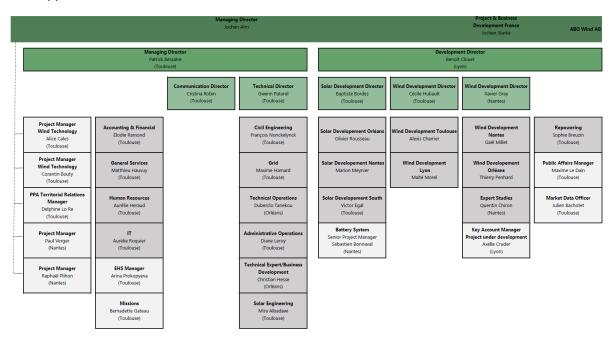


Figure 1: Organigramme de la filiale française

1.2.1. Ses chiffres clés

Tableau 4: Quelques chiffres clés de ABO Wind France

Catégorie	Valeur
Employés	> 180
Pourcentage de femme	41 %
Pourcentage d'homme	59 %
Parcs éolien mis en service	38
Projets éoliens en développement	6
Puissance installée	> 400 MW
Projets en développement	1,6 GW

1.2.2. Les concurrents de ABO Wind France

ABO Wind France s'affirme comme un acteur clé du développement des énergies renouvelables en France, évoluant dans un environnement concurrentiel marqué par la présence d'autres entreprises, tant nationales qu'internationales.

Parmi les principaux concurrents d'ABO Wind France figurent :

- **EDF Renouvelables :** En tant que filiale du groupe EDF, EDF Renouvelables occupe la première place sur le marché français du développement des énergies renouvelables, intervenant dans des segments variés tels que l'éolien, le solaire, la biomasse et l'hydroélectricité.
- **ENGIE Green :** Appartenant au groupe ENGIE, ENGIE Green représente un autre acteur majeur du secteur des énergies renouvelables en France, opérant dans des domaines variés tels que l'éolien, le solaire, la biomasse et le stockage d'énergie.
- Vinci Energies: En tant que filiale du groupe Vinci, Vinci Energies se distingue en tant qu'entreprise multiservices impliquée dans le développement, la construction et l'exploitation de projets d'énergies renouvelables.
- Neoen: Entreprise française spécialisée dans le développement et l'exploitation de projets d'énergies renouvelables, Neoen se positionne dans des secteurs tels que l'éolien, le solaire et le stockage d'énergie.
- **TotalEnergies :** En tant que groupe pétrolier et gazier français, TotalEnergies entreprend une diversification vers les énergies renouvelables, participant activement dans des domaines tels que l'éolien, le solaire, la biomasse et l'hydroélectricité.

2. La structure d'accueil

C'est donc au sein du service Informatique à Toulouse que le stage en tant qu'élève ingénieur a été effectué.

2.1. L'agence de Toulouse

L'agence de Toulouse, en tant que première et principale agence, est constituée d'une équipe de 88 collaborateurs, incluant stagiaires et alternants. Elle présente une structure à taille humaine, ce qui favorise un environnement social positif. Les collaborateurs sont encouragés à interagir et à partager leurs connaissances. L'entreprise organise également des événements réguliers pour favoriser le team building.

L'agence ABO Wind de Toulouse offre des conditions de travail attractives à ses collaborateurs et comprend divers services tels que la communication, la comptabilité, la construction, le développement éolien, le développement photovoltaïque, l'environnement de travail, l'exploitation administrative, l'exploitation technique, la santé, la sécurité et l'environnement (HSE), l'hydrogène, l'informatique, l'ingénierie photovoltaïque, le raccordement, le Repowering et les ressources humaines.

Le groupe ABO Wind dispose de deux services informatiques, l'un basé en Allemagne et l'autre en France. Le service allemand gère les filiales étrangères, tandis que le service français implanté à l'agence de Toulouse, gère les agences du territoire français.

2.2. Le service informatique

Le service IT a pour missions fondamentales d'assurer la fiabilité, l'administration, la sécurité, et la supervision des systèmes informatiques de l'entreprise. Il est également chargé de fournir les outils nécessaires au développement de l'activité, comprenant du matériel, des logiciels, des données, formations et accompagnement, afin de répondre aux besoins opérationnels des collaborateurs.



Figure 2: Les domaines d'application des actions du service IT

Le département informatique de l'agence regroupe un éventail de professionnels.

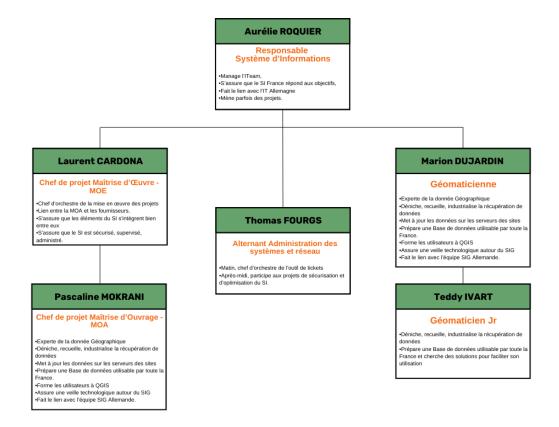


Figure 3: Organigramme du service IT

II- PROJET D'ÉTUDE

1. Présentation du projet

1.1. Le projet éolien

La création d'un parc éolien peut être entreprise de manière complète, de la conception initiale à la mise en service, et il est également envisageable de renouveler un parc existant. Dans les deux cas, plusieurs étapes sont à prendre en compte.

Le processus de réalisation d'un projet éolien est une démarche complexe qui passe par différentes phases :

- **Prospection :** identification des sites potentiels pour l'installation de parcs éoliens.
- **Développement :** réalisation des études de faisabilité techniques et économiques.
- **Construction :** installation des éoliennes et des infrastructures associées.
- **Exploitation**: production d'électricité et maintenance des éoliennes.

Les éoliennes ont une durée de vie moyenne de 20 à 25 ans. Lorsqu'elles arrivent en fin de vie, elles peuvent être remplacées par de nouvelles. Ce processus appelé le renouvellement ou Repowering des parcs éoliens comprend les étapes suivantes :

- Étude de faisabilité : détermination la faisabilité du projet et de définir les besoins du projet.
- **Conception**: conception du projet en détail, en tenant compte des aspects techniques, économiques et réglementaires.
- **Démantèlement des éoliennes existantes :** retrait des éoliennes existantes en respectant les normes environnementales.
- Installation des éoliennes : installation des nouvelles éoliennes.

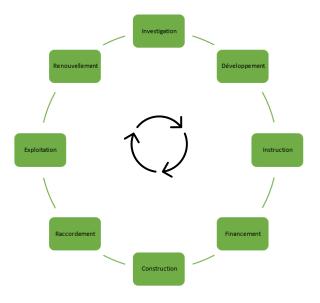


Figure 4: Les étapes du projet éolien chez ABO Wind

Il est donc crucial de comprendre le contexte éolien d'un site avant de s'y engager, et dans le cadre d'une démarche de renouvellement, la connaissance des années de service des parcs existants par exemple, revêt une importance primordiale. Ainsi, à chaque étape des processus de réalisation et de renouvellement, disposer de données sur les parcs et les mâts éoliens existants s'avère essentiel pour assurer une plus grande efficacité. Cela s'explique par le fait que ces informations fournissent une base solide pour prendre des décisions éclairées, anticiper les défis potentiels, et exploiter pleinement les opportunités liées à l'évolution des projets éoliens. Elles fournissent une panoplie d'informations essentielles allant au-delà de la simple localisation des éoliennes. Elles englobent des détails cruciaux tels que les caractéristiques techniques des mâts, les dates d'installation, les données de production énergétique, les propriétaires des parcs, les développeurs impliqués, et les intervenants clés dans chaque projet. Ces données constituent ainsi une source inestimable.

1.2. La gestion actuelle des données sur les éoliennes chez ABO Wind

Au sein d'ABO Wind, les différents services impliqués dans les processus de réalisation et de renouvellement des parcs éoliens se réfèrent principalement à trois bases de données distinctes :

- La BDD de la DREAL: Cette base provient de la DREAL et est une source administrative. Elle est structurée en fonction des régions et chaque donnée dépend de la personne responsable de cette donnée. Cette base est considérée comme la plus précise et à jour, car elle présente les projets autorisés, refusés et en instruction.
- La BDD de Géorisques: Cette base est construite autour du volet ICPE. Chaque parc éolien y
 est classifié en tant qu'installation avec une classification environnementale spécifique. Cette
 base offre des informations plus détaillées et est géographiquement organisée selon les ICPE.
 Elle est mise à jour par les propriétaires des parcs éoliens via la plateforme OREOL (outil de
 référencement des éoliennes).
- La BDD de Wind Power: Dérivée de Wind Power, cette base répertorie les parcs éoliens en France. Elle est initialement gratuite et élaborée à partir des recherches d'un tiers pendant son temps libre. Celle-ci est désormais payante.

Étant donné la distinction entre ces bases de données, des interrogations sur les divergences entre elles sont inévitables. Afin de cibler les territoires couverts par les services de développement, onze régions spécifiques de France ont été sélectionnées (cf. tableau 5).

Nom	Code
Nouvelle-Aquitaine	75
Occitanie	76
Grand Est	44
Hauts-de-France	32
Auvergne-Rhône-Alpes	84
Bourgogne-Franche-Comité	27
Normandie	28
Pays de la Loire	52
Centre-Val de Loire	24
Provences-Alpes-Côte d'Azur	93
Bretagne	53

Tableau 5: Les régions sélectionnées

Une première analyse des BDD de la DREAL et de Géorisques a été effectuée. Pour ce faire, une comparaison des informations a été réalisé principalement sur le nombre d'éoliennes recensées, leurs caractéristiques, leur état administratif, les coordonnées géographiques, les dates de mise en service et leurs hauteurs. Cela a permis de se rendre compte des différences notables entre celles-ci.

Afin de créer une source unique, fiable, cohérente et de faciliter les opérations entre les services de l'entreprise, il était impératif de trouver des solutions pour unifier ces bases de données tout en maximisant la pertinence des données.

2. Présentation des outils utilisés

Tableau 6: Outils utilisés

Logiciel	Description
Microsoft Excel	Excel est un logiciel développé par Microsoft. Il offre une interface composée de cellules organisées en lignes et colonnes, permettant aux utilisateurs d'effectuer des opérations mathématiques, statistiques et logiques. Excel est également utilisé dans la manipulation et la visualisation de données.
Xmind	Xmind est un logiciel de carte mentale et de brainstorming, développé par Xmind Ltd. En plus des éléments de gestion, le logiciel peut être utilisé pour capturer des idées, clarifier la pensée, gérer des informations complexes et promouvoir la collaboration d'équipe.
FME by Safe Saftware	FME est un logiciel de type ETL développé par la société canadienne SAFE Software, spécialisé en donnée géographique vectorielle et image. Il est utilisé dans tous les secteurs d'activité de l'information géographique.
QGIS	QGIS est un logiciel open source de SIG. Il a été développé par une communauté de développeurs bénévoles et professionnels et de visualiser, de créer, de gérer et d'analyser des données géospatiales
JMerise JMERISE	Développé par un professeur, JMerise est un outil de modélisation de bases de données qui vise à simplifier le processus de conception en utilisant une approche visuelle et en offrant des fonctionnalités de génération de scripts SQL.
PgAdmin 4	PgAdmin 4 été développé par la communauté open source PostgreSQL. C'est un outil graphique multi-plateforme et le principal outil de gestion open source des bases de données PostgreSQL. Il fournit une interface graphique pour la création, la maintenance et l'utilisation de base de données.

2023/2024

3. Présentation du travail effectué

3.1. Traitement des données

Le traitement des données passe par une succession d'opérations notamment de fusions et de jointures. L'ensemble de ces traitements sur les données a été réalisé à l'aide du logiciel FME. Ainsi, du traitement initial jusqu'à l'intégration des données dans la base de données PostgreSQL, et même pour l'interfaçage QGIS, les « Transformers » du logiciel FME ont été utilisés.

> Analyse des données

À la suite de la première analyse réalisée par la tutrice, et en collaboration avec certains services concernés, une première sélection des champs a été effectuée. Cependant, toutes les bases de données et leurs contenus n'avaient pas été pris en compte, d'où la nécessité d'une seconde analyse plus approfondie.

Cette seconde analyse a consisté à examiner de manière exhaustive les champs contenus dans les BDD de la DREAL, Géorisques et Wind Power. Les structures des bases de données étant différentes, les approches devaient être adaptées. En effet, les bases issues de Géorisques et de la DREAL sont mises à disposition par région. Une analyse a été réalisée donc par région pour ces sources de données. En revanche, la BDD Wind Power présente des informations globales sans distinction régionale. Pour évaluer ces différences, une comparaison détaillée des contenus des BDD de la DREAL et Géorisques a été initiée dans un premier temps à travers un tableau comparatif dans un tableur Excel. L'objectif premier était d'identifier les similitudes entre les bases de données et sélectionner les champs d'intérêt.

Puis, une comparaison globale des trois bases de données a été réalisée grâce à une schématisation afin d'obtenir une vue d'ensemble plus claire. Ainsi, la réalisation d'un schéma récapitulatif (*cf. figure 5*) a été entreprise. Celui-ci illustre les liens entre les bases de données sources et les champs préalablement sélectionnés, en incorporant les différentes nomenclatures utilisées dans ces bases.

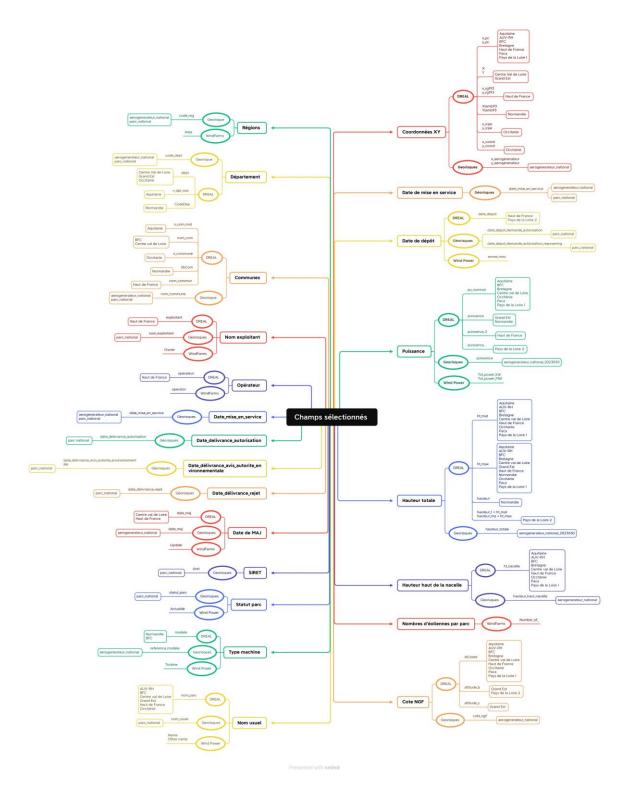


Figure 5: Schéma de l'inspection des champs sélectionnés dans les BDD

À la suite de cette inspection, plusieurs décisions ont été prises :

- Donner la priorité aux bases de données de la DREAL et de Géorisques en tant que principale, tout en intégrant celle de Wind Power en tant que base complémentaire.
- Prendre en compte les deux tables « parc » et « mat » des BDD,

- Opter pour une base de données relationnelle comprenant deux tables : une table « mat » et une table « parc » reliées par l'identifiant du parc.

Enfin, Les champs sélectionnés ont été révisés afin de garantir une pertinence adéquate.

Collecte des données à jour

Après avoir établi ces directives, il était nécessaire d'effectuer le téléchargement des données sur les sites concernés afin de construire une base de données avec des informations aussi actualisées que possible.

Le téléchargement des données depuis le site de Géorisques a été réalisé de manière simple, par région.

Pour ce qui est des données de la DREAL, étant donné que les sites DREAL sont organisés par région, il a été nécessaire de visiter chaque site individuellement, de les parcourir et d'extraire les données. Puisque la publication des données dépend du responsable, les formats pouvaient varier. Certaines données pouvaient être récupérées directement depuis le site dans le format souhaité. Tandis que pour les données en format shapefile, il était nécessaire de les charger sur QGIS et de passer ensuite par une opération permettant d'extraire ces données en format CSV.

Il est important de souligner que les données provenant du site Géorisques étaient structurées avec deux tables distinctes, à savoir « parc » et « mat », et elles étaient constituées par région. En revanche, cette organisation n'était pas présente dans les données provenant du site de la DREAL (cf. tableau 7).

	BDD DREAL		BDD Géorisques	
Région	TABLE			
	parc	mat	parc	mat
Nouvelle-Aquitaine	/	~	/	/
Occitanie	/	~	/	~
Grand Est	/	~	~	/
Hauts-de-France	~	/	~	~
Auvergne-Rhône-Alpes	/	~	~	/
Bourgogne-Franche-Comité	×	~	~	~
Normandie	×	~	/	/
Pays de la Loire	/	~	~	~
Centre-Val de Loire	×	~	~	/
Provences-Alpes-Côte d'Azur	/	~	~	-
Bretagne	/	/	~	/

Tableau 7: Récapitulatif des tables contenues dans les BDD

Liaison des données

L'objectif ici était de prendre en compte les données sur les parcs et les mâts éoliens issues des deux bases de données, de les recroiser et de les comparer en vue de les unifier dans une seule base. L'analyse des données a révélé que le champ du code ICPE serait celui permettant de relier ces tables.

Cependant, il est important de noter que ce champ n'est pas présent dans toutes les tables de la base de données de la DREAL. Par conséquent, il a été nécessaire de procéder de deux manières distinctes :

• Croisement des tables avec le code ICPE :

Dans la BDD Géorisques, toutes les tables « parc » et « mat » des régions sélectionnées comportent un champ incluant le code ICPE des parcs. Cependant, pour la BDD de la DREAL, seules les tables « parc » des régions de la Nouvelle-Aquitaine, l'Occitanie, Hauts-de-France, Pays de la Loire et du Grand Est possèdent le champ « code ICPE ».

- Fusions des tables parcs

Le traitement a débuté par la fusion des tables « parc » des bases de données de la DREAL et de Géorisques (cf. figure 6) pour les régions possédant le code ICPE dans leurs tables. Cette fusion a été réalisée à l'aide du Transformers « Feature Merger » de FME. En ce qui concerne sa configuration, la table issue de la base de données de Géorisques a été désignée comme le "Requestor", tandis que la table issue de la base de données de la DREAL a été désignée comme le "supplier". Ce choix de configuration s'explique par le fait que les tables de la base de données de Géorisques sont plus complètes que celles de la DREAL. À la suite de cette fusion, plusieurs résultats ont été obtenus :

- Des parcs fusionnées ("merged") issus des tables des deux BDD.
- Des parcs non fusionnés issus de la BDD de la DREAL ("unmerged supplier").
- Des parcs non fusionnés issus de la BDD de Géorisques ("unmerged requestor").

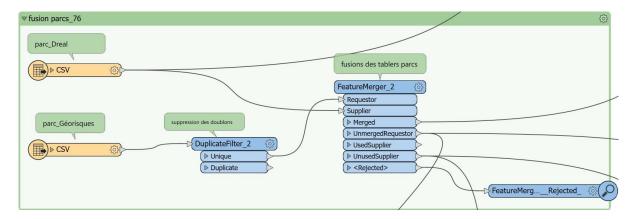


Figure 6: Fusion des données des tables « parc »

Jointures entre les tables « mat » et les tables « parc »

Une fois les tables des parcs ont été fusionnées, des jointures ont été réalisées avec le code ICPE entre les données obtenues et les tables des mâts éoliens (cf. figure 7). À noter que comme le code ICPE est associé au parc éolien, les tables « mat » de la base de données de la DREAL ne comprennent pas de

champ contenant ce code. Ainsi, une jointure préalable entre la table des parcs et la table des mâts éoliens a été nécessaire pour lier le code ICPE des parcs aux mats éoliens qu'ils possèdent.

Plus spécifiquement, en considérant en premier lieu les parcs fusionnés, l'utilisation du Transformers « FeatureJoiner » de FME a permis d'effectuer des jointures entre ces parcs et la table des mâts provenant de la BDD de la DREAL d'une part, et entre ces parcs et la table de la BDD de Géorisques d'autre part. Les parcs fusionnés ont été pris comme la source contenant les entités à joindre ("left"). Les mats, quant à eux, ont été associés à la source contenant les entités à utiliser pour la jointure ("right"). Ce choix a été fait pour une meilleure cohérence dans la lecture, étant donné que les mâts éoliens font partie intégrante des parcs. Les résultats de ces jointures incluent :

- Des données jointes ("merged").
- Des données non jointes issues de la fusion des parcs ("unjoined left").
- Des données non jointes issues des tables des mats ("unjoined right").

Par la suite, les données non jointes ont été prises en considération et ont été jointes respectivement aux parcs non fusionnés. Autrement dit, des jointures ont été effectuées entre les parcs non fusionnés provenant de la BDD de la DREAL et la table « mat » de la même base, de même que pour les parcs non fusionnés issus de la BDD de Géorisques et la table « mat » de la même base.

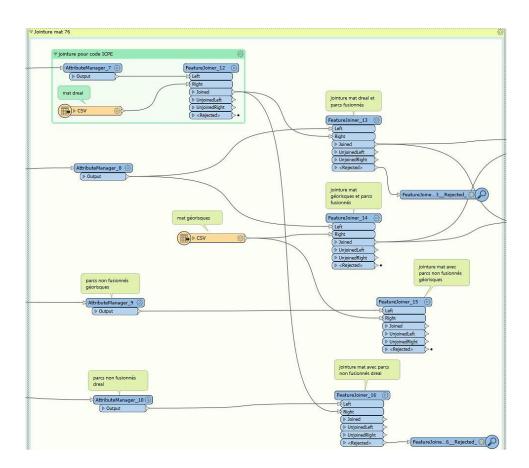


Figure 7: Jointures entre les tables

Détection des différences

Dans l'objectif de faciliter la comparaison et l'unification des données croisées, le Transformers « ChangedDetector » de FME a été utilisé pour détecter les modifications entre les ensembles de données. Les données prises en compte étaient celles issues des deux jointures avec les parcs fusionnés. Ce processus a permis de distinguer :

- Les entités ajoutées ("inserted") : entités présentes dans le deuxième jeu de données mais absentes dans le premier.
- Les entités supprimées ("deleted") : entités présentes dans le premier jeu de données mais absentes dans le deuxième.
- Les entités inchangées ("unchanged") : entités présentes dans les deux jeux de données et n'ayant subi aucune modification.

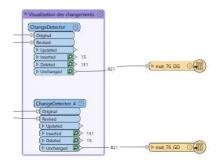


Figure 8: Visualisation des changements

• Croisement des tables sans le code ICPE :

Dans les cas où les tables ne comportaient pas le champ du code ICPE, il s'agissait des régions dépourvues de tables « parc » dans la BDD de la DREAL, à savoir : Bourgogne-Franche-Comité, Normandie et Centre-Val de Loire. Dans ces situations, l'approche adoptée a consisté à utiliser les tables de la BDD de Géorisques. La procédure a ainsi impliqué une jointure directe entre les tables « parc » et « mat » de cette base de données.

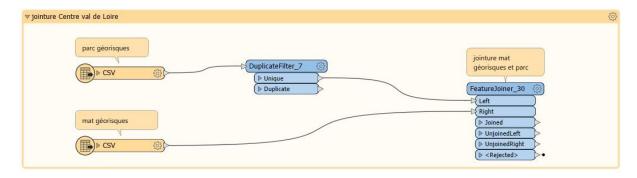


Figure 9 : Cas de traitement sans code ICPE

En ce qui concerne le croisement entre les tables « mat » des deux BDD pour la comparaison des données, cela ne pouvait pas se faire directement en raison des différences entre ces tables. Afin de surmonter cette difficulté, une alternative a été envisagée. En procédant par les noms des parcs contenus dans la table « mat », une standardisation en majuscules a été appliquée au champ correspondant. Une jointure avec la table « parc » a ensuite été réalisée pour faciliter le tri, suivi de l'utilisation d'un « ChangeDetector » pour mettre en évidence les divergences.

Il est important de noter que cette méthode n'était pas applicable à l'ensemble des régions, et elle n'a pas permis une comparaison exhaustive. Ainsi, la comparaison a exigé une approche manuelle, impliquant l'examen individuel des éoliennes.

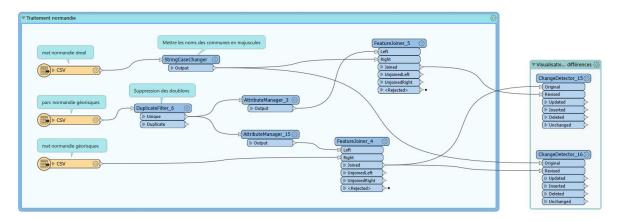


Figure 10: Autre cas de traitement sans le code ICPE

En amont, les tables de la région de PACA étaient présentes dans les deux bases, cependant, le code ICPE n'était disponible que dans les tables provenant de Géorisques. Pour cette région, aucun croisement spécifique n'a été réalisé.

3.2. Remplissage des tables finales

Des classeurs Excel ont été préparés pour les parcs et les mâts éoliens, chacun composé de onze feuilles correspondant aux différentes régions. Sur ces feuilles, des tableaux ont été créés avec des nomenclatures de colonnes correspondant à celles de la BDD finale. Cela a permis d'y intégrer les données traitées, comparées et vérifiées dans certains cas.

Ainsi, une fois les données fusionnées et jointes, il a été nécessaire de les extraire, de les inspecter, de les comparer, puis de les organiser dans les feuilles des classeurs « mat » et « parc » préalablement configurées. En ce qui concerne les données relatives aux parcs éoliens, l'intégration a débuté par les parcs fusionnés, suivis des non fusionnés après une comparaison approfondie. Pour les données relatives aux mâts éoliens, leur intégration s'est effectuée après une analyse comparative des résultats des différentes jointures. Les données sur les mâts issues d'autres jointures ont également été prises en compte et ajoutées par la suite.

code_icpe -i nom_parc	code_com nom_com	▼ code_dpt ▼ code	reg 💌 nb_eol	▼ pu_totale	▼ dt_depot ▼
6306501 IEL EXPLOITATION 28	85162 RIVES-D'AUTISE	85	52	9 18	28/12/2011
6306502 IEL EXPLOITATION 29	85306 XANTON CHASSENON	85	52	3 6	28/12/2011
6306511 FERME EOLIENNE DU CANAL DE GARGOUILLEAU	85105 LE GUE-DE-VELLUIRE	85	52	5 10	30/06/2005
6306514 Ferme Eolienne des Hautes Landes	44048 COUFFE	44	52	5 10	29/12/2011
6306530 EDF EN FRANCE EDF énergies renouvelables	85029 BOUIN	85	52	5 12	12/12/2001
6306559 Parc éolien de Maché	85130 MACHE	85	52	4 8	07/04/2005
6306585 Ferme Eolienne d'Erbray SAS	44054 ERBRAY	44	52	3 6.89	25/04/2012
6306586 FERME EOLIENNE DE LA LANDE	53072 COMMER	53	52	6 12	16/05/2012
6306587 I	53113 HAMBERS	53	52	4 8.2	11/12/2009
6306603 Parc éolien de Beauséjour	44118 PANNECE	44	52	5 10	19/09/2006
6306605 Parc éolien de La Vallière	44118 PANNECE	44	52	4 8	19/09/2006
6306606 Eoliennes de Soulvache SAS	44200 SOULVACHE	44	52	4 8	24/04/2007
6306627 PARC EOLIEN DU HAM	53112 LE HAM	53	52	3 6	25/05/2011
6306628 EOLIENNES de La Haie Traversaine	53111 LA HAIE TRAVERSAINE	53	52	8 16	02/05/2006
6306630 CENTRALE EOLIENNE DE BAIS ET TRANS	53266 TRANS	53	52	3 6	14/12/2009
6306631 Ferme eolienne Lassay	53127 LASSAY LES CHATEAUX	53	52	3 6.89	15/02/2007
6306632 FERME EOLIENNE DE TRANS ET COURCITE	53083 COURCITE	53	52	3 10.35	25/06/2010
6306637 Société d'Energie Eolienne de Campbon	44025 CAMPBON	44	52	5 12	23/10/2006
6306641 PARC EOLIEN NORDEX XIX SAS	44205 LES TOUCHES	44	52	6 15	24/12/2009
6306642 Eoliennes Pouille les Coteaux SNC	44096 MESANGER	44	52	4 9.2	08/06/2007
6306643 Parc éolien du Bois Gautier SAS	44031 LA CHAPELLE-GLAIN	44	52	5 11.75	23/05/2007
6306644 Parc de la Renardière	44075 ISSE	44	52	6 12	15/05/2007
6306645 PARC DE BEAUMONT	44075 ISSE	44	52	6 12	20/06/2007
6306646 FERME EOLIENNE DE BENET SAS	85020 BENET	85	52	5 10	16/09/2004
6306647 LONGEVILLE	85127 LONGEVILLE SUR MER	85	52	5 12.5	04/12/2007
6306649 FP LUX WIND MONTAFILANT	44197 SION LES MINES	44	52	5 10	14/06/2007
6306650 LHI GREEN INFRASTRUCTURE WIND FRANKREICH	44091 MARSAC SUR DON	44	52	5 10	27/05/2009
6306663 FERME EOLIENNE DE TASSILLE	72348 TASSILLE	72	52	4 8	28/10/2008
parc_paysDeLaLoire parc_n.aquitaine parc_occitanie	parc_hautDeFrance parc_GrandEst parc_Breta	gne parc (+) :	•		

Figure 11: Aperçu du remplissage du classeur des parcs éoliens

code_icpe 🔻 id_mat	▼ id_parc	num_mat 🔻	code_com 💌 nom_com	code_dept pu_unit	▼ ht_max	✓ ht_mat	▼ ht_r	nacelle 🕶 dian	_roto 🔻
6306501 M520011	PC52001	E1	85162 RIVES-D'AUTISE	85	2	150	98	100	104
6306501 M520012	PC52001	E2	85162 RIVES-D'AUTISE	85	2	150	98	100	104
6306501 M520013	PC52001	E3	85162 RIVES-D'AUTISE	85	2	150	98	100	104
6306501 M520014	PC52001	E4	85162 RIVES-D'AUTISE	85	2	150	98	100	104
6306501 M520015	PC52001	E5	85162 RIVES-D'AUTISE	85	2	150	98	100	104
6306501 M520016	PC52001	E6	85162 RIVES-D'AUTISE	85	2	150	98	100	104
6306501 M520017	PC52001	E7	85162 RIVES-D'AUTISE	85	2	150	98	100	104
6306501 M520018	PC52001	E8	85162 RIVES-D'AUTISE	85	2	150	98	100	104
6306501 M520019	PC52001	E9	85163 RIVES-D'AUTISE	85	2	150	98	100	104
6306502 M520021	PC52002	E1	85306 XANTON CHASSENON	85	2	150	98	100	104
6306502 M520022	PC52002	E2	85306 XANTON CHASSENON	85	2	150	98	100	104
6306502 M520023	PC52002	E3	85306 XANTON CHASSENON	85	2	150	98	100	104
6306511 M520033	PC52003	E3	85105 LE GUE-DE-VELLUIRE	85	2	125	80	80	90
6306511 M520034	PC52003	E4	85105 LE GUE-DE-VELLUIRE	85	2	125	80	80	90
6306511 M520032	PC52003	E2	85105 LE GUE-DE-VELLUIRE	85	2	125	80	80	90
6306511 M520031	PC52003	E1	85105 LE GUE-DE-VELLUIRE	85	2	125	80	80	90
6306511 M520035	PC52003	E5	85303 VIX	85	2	125	80	80	90
6306514 M520041	PC52004	E1	44048 COUFFE	44	2	125	80	82	90
6306514 M520042	PC52004	E2	44096 MESANGER	44	2	125	80	82	90
6306514 M520043	PC52004	E3	44096 MESANGER	44	2	125	80	82	90
6306514 M520044	PC52004	E4	44048 COUFFE	44	2	125	80	82	90
6306514 M520045	PC52004	E5	44048 COUFFE	44	2	125	80	82	90
6306530 M520051	PC52005	E1	85029 BOUIN	85 2.4		102	60	60	84
6306530 M520052	PC52005	E2	85029 BOUIN	85 2.4		102	60	60	84
6306530 M520054	PC52005	E4	85029 BOUIN	85 2.4		102	60	60	84
6306530 M520053	PC52005	E3	85029 BOUIN	85 2.4		102	60	60	84
6306530 M520055	PC52005	E5	85029 BOUIN	85 2.4		102	60	60	84
6306559 M520062	PC52006	E2	85130 MACHE	85	2	125	80	78	90
mat_75_NA	mat_PdL	mat_76 Feuil1	mat_Bretagne_53 mat_BFC_27	mat_Normandie_28	+				

Figure 12: Aperçu du remplissage du classeur des mâts éoliens

Il convient de noter que malgré la fusion et la jointure des tables de toutes les régions lorsque cela était possible avec le code ICPE, le remplissage des données pour l'ensemble de ces régions n'a pas été réalisé en raison de contraintes temporelles, les priorités étant données aux comparaisons et aux analyses. Ainsi, les régions du Pays de la Loire, de Provence-Alpes-Côte d'Azur et de la Nouvelle Aquitaine ont été complètement remplies, tandis que pour les autres régions, les tables des parcs ont été préalablement remplies, car les fusions nécessitent moins de comparaisons. Les données sur les éoliennes restent à être comparées entre les BDD.

Création des ID

Comme évoqué précédemment, étant donné la distinction des bases de données, l'absence d'un identifiant unique pour chaque parc ou mât éolien dans n'importe quelle base était évidente. Bien que le code ICPE permettait d'établir un lien entre les BDD, il n'était pas référencé pour tous les parcs et dans toutes les tables. Cela a conduit à la nécessité de créer des identifiants uniques (*cf. figure 13*) pour les parcs (id_parc) et les mâts éoliens (id_mat) dans la BDD de l'entreprise. Ainsi, des nomenclatures ont été développées pour les parcs en fonction de leur région respective et pour les mâts éoliens en fonction du parc auquel ils sont associés.

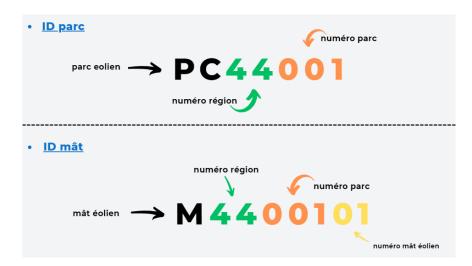


Figure 13: Nomenclature des ID des parcs et mâts

3.3. Intégration des données dans la base de données PostgreSQL

3.3.1. Configuration des champs : nomenclatures et types de données

Afin de garantir une compréhension optimale des champs des tables, un nom unique a été attribué à chaque champ, tout en respectant une limite de 10 caractères. Cette restriction vise à favoriser la lisibilité, à optimiser l'utilisation de l'espace mémoire, et à respecter la limite des caractères qui peuvent s'afficher dans QGIS.

Parallèlement, la définition anticipée des types de champs avant leur intégration dans la BDD joue un rôle crucial dans l'établissement d'une base solide et efficiente. Cette approche a pour objectif de garantir la qualité et l'intégrité des données stockées. Dans le cadre de ce projet, la définition des types de données a été effectuée en prenant en considération les données présentes dans les tables sources utilisées.

<u>Tableau 8:</u> Configuration des champs de la table « parc »

TABLE PARC										
N°	Champ	Nomenclature	Туре	Précision						
1	identifiant du parc	id_parc	VARCHAR	10						
2	code icpe du parc	code_icpe	VARCHAR	10						
3	nom du parc	nom_parc	VARCHAR	250						
4	code du département	code_dpt	VARCHAR	2						
5	code de la région	code_reg	VARCHAR	2						
6	Nombre d'éolienne par parc	nb_eol	INTEGER	-						
7	puissance totale du parc	pu_totale	NUMERIC	(5,2)						
8	date de dépôt de la demande d'autorisation	dt_depot	DATE	-						
9	date de début de l'enquête publique	dt_db_enq	DATE	-						
10	date de fin de l'enquête publique	dt_fin_enq	DATE	-						
11	nom du commissaire enquêteur	n_com_enq	VARCHAR	50						
12	avis du commissaire enquêteur	avis_com	VARCHAR	250						
13	date de délivrance de l'avis de l'autorité environnementale	dt_ae	DATE	-						
14	date de décision	dt_dec	DATE	-						
15	etat administratif	etat_adm	VARCHAR	10						
16	contentieux	contentieu	BOOLEAN							
17	en service	en_service	VARCHAR	4						
18	date de mise en service	dt_mes	DATE							
19	developpeur	dvloppeur	VARCHAR	250						
20	société projet (spv)	n_spv	VARCHAR	250						
21	siret de la société projet	siret_spv	VARCHAR	50						
22	Propriétaire du parc	n_proprio	VARCHAR	250						
23	Date de cessation d'activité	dt_ces_act	DATE	-						
24	date de démolition du parc	dt_demol	DATE							
25	Renouvellement_dépot de la demande	dt_depo_rp	DATE	-						
26	Renouvellement_date de délivrance de l'avis de l'autorité environnementale	dt_ae_rp	DATE							
27	Renouvellement_date de décision	dt_dec_rp	DATE	-						
28	Renouvellement notable	rp_notable	BOOLEAN							
29	date de mise à jour	dt_maj	DATE	-						
30	nom de l'administrateur	n_admin	VARCHAR	3						
31	lien vers la source de d'information	source	VARCHAR	250						

Tableau 9: Configuration des champs de la table « mat »

TABLE MAT EOLIEN												
N°	Champ	Nomenclature	Туре	Précision								
1	identifiant du mat	id_mat	VARCHAR	10								
2	identifiant du parc	id_parc	VARCHAR	10								
3	numéro du mat	num_mat	VARCHAR	3								
4	code insee	code_com	VARCHAR	7								
5	nom de la commune	nom_com	VARCHAR	250								
6	puissance unitaire	pu_unit	NUMERIC	(5,2)								
7	hauteur maximale	ht_max	NUMERIC	(5,2)								
8	hauteur du mat	ht_mat	NUMERIC	(5,2)								
9	hauteur de la nacelle	ht_nacelle	NUMERIC	(5,2)								
10	diamètre du rotor	diam_rotor	NUMERIC	(5,2)								
11	cote NGF	cote_ngf	NUMERIC	(5,2)								
12	constructeur de l'éolienne	construct	VARCHAR	100								
13	référence modèle de l'éolienne	ref_model	VARCHAR	100								
14	coordonnée x	x_coord	FLOAT	-								
15	coordonnée y	y_coord	FLOAT	-								
16	epsg	epsg	INTEGER	-								
17	etat du mat	etat_mat	VARCHAR	4								
18	date de mise à jour	dt_maj	DATE	•								
19	nom de l'administrateur	n_admin	INTEGER	3								

3.3.2. Création de la base de données

Génération du script SQL avec JMerise

Dans un premier temps, l'utilisation de JMerise a été entreprise pour la conception du modèle conceptuel de données (MCD). Deux entités distinctes, à savoir « mat » et « parc », ont été créées au sein de ce MCD. Ensuite, des champs ont été ajoutés à chaque entité avec une configuration

minutieuse des paramètres, comprenant la spécification des noms, types de données et contraintes associées (clés primaires ou étrangères). Après cette étape, l'instauration de l'association entre les deux entités a été réalisée, et les règles de liaison ont été définies en utilisant les cardinalités. Plus précisément, cela a été exprimé par la règle selon laquelle un parc peut posséder un à n éoliennes, tandis qu'une éolienne ne peut appartenir qu'à un et un seul parc. S'appuyant sur ce MCD élaboré (cf. figure 14), JMerise a ensuite automatisé la génération du MLD (cf. figure 15) ainsi qu'un script SQL complet prêt à être exécuté sur le SGBDR.

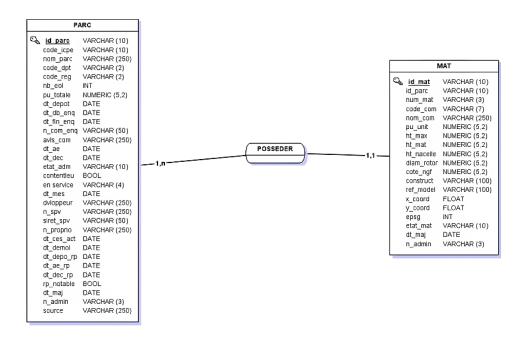


Figure 14: MCD de la BDD

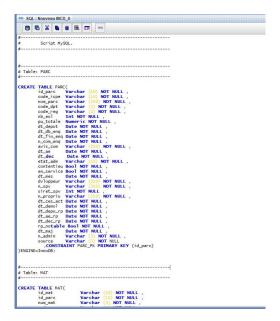


Figure 15: Apercu du code SQL généré

Configuration sur PgAdmin4

Une fois le code SQL généré, la création des tables dans PGAdmin4 a été entreprise. L'accès au serveur de l'entreprise a permis la création d'une BDD PostgreSQL avec l'intégration de l'extension PostGIS, nommée « bdd_abo_eol ». Des modifications ont été apportées au code, incluant l'ajout de lignes pour créer le schéma « eol » et accorder diverses permissions à l'utilisateur « postgres » sur le schéma et son contenu. Étant donné que les données destinées à être intégrées à la base possèdent une référence spatiale telles que des coordonnées GPS, celles-ci vont être stockées dans un champ appelé « geom » stockant la codification des géométries de la table. Un champ « geom » a donc été ajouté à chaque table dans le code SQL. À la suite de ces ajustements, le code SQL a été collé et exécuté, entraînant la création des tables « parc » et « mat » dans le schéma « eol » (cf. figure 16).



Figure 16: Arborescence de la BDD dans PgAdmin4

3.3.3. Intégration de la donnée dans les tables

À la suite de la mise en place de la base de données et des tables, l'intégration des données devient réalisable. À cet effet, l'utilisation de Transformers FME a été privilégiée. Il est à noter que le processus d'intégration des données pour la table « parc » a différé de celui de la table « mat », en raison de la présence de coordonnées dans cette dernière.

Intégration des données des parcs éoliens

Pour la table « parc », le chargement du tableau Excel des données a été effectué via le « Reader » FME. Ensuite, l'utilisation du Transformers « AttributeManager » a permis de préciser adéquatement les types de champs afin de les aligner avec ceux spécifiés lors de la création des tables dans la base de données. La définition d'une clé primaire unique est essentielle pour identifier de manière univoque chaque ligne de la table. Ainsi, un champ a été ajouté avec spécification de sa valeur par incrémentation. Après cette étape, l'utilisation du Transformers « Attribute Encoder » a été entreprise pour gérer l'encodage de toutes les entités, les convertissant en UTF-8. L'usage de l'UTF-8 a été privilégié en raison de sa capacité à représenter un large jeu de caractères, assurant ainsi une prise en charge adéquate de diverses données textuelles.

Une fois les données chargées et configurées, l'intégration dans la BDD a été entreprise. Une connexion à la base depuis FME était nécessaire, et cette opération a été facilitée par un Writer minutieusement configuré. Une fois que le format PostGIS a été spécifié, il a fallu configurer les paramètres de connexion à la base PostgreSQL, incluant l'hôte, le port, le nom de la BDD, et le mot de passe. Une fois la connexion établie, plusieurs options ont été configurées pour préciser les détails spécifiques à la BDD concernant la manière dont les données devaient être écrites. Parmi ces options, la table « parc » de destination a été spécifiée, l'opération (dans ce cas, une insertion) a été définie, ainsi que la méthode d'exécution. Le Writer a été ajouté en précisant la table « parc » comme destination. Ensuite, le « Writer » a été connecté aux configurations précédemment établies sur les données chargées, et enfin, l'exécution a été lancée pour effectuer l'intégration.

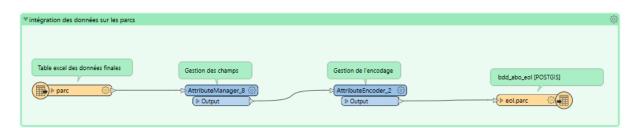


Figure 17: Intégration des données sur les parcs éoliens

> Intégration des données sur les mats éoliens

Concernant l'intégration des données de la table « mat », la procédure a suivi la même approche que celle utilisée pour la table « parc ». Cependant, la table « mat » présente une particularité car elle possède des coordonnées géographiques. Afin de traiter cette spécificité, le Transformers « Reprojector » a été utilisé pour appliquer le système de coordonnées de référence RGF93/Lambert 93 (code EPSG :2154).

Quant à l'ajout du Writer pour la connexion à la base de données, les configurations ont été ajustées : la table de destination spécifiée a été la table « mat ». Les options, telles que l'opération (insertion) et la méthode d'exécution, ont été configurées de la même manière que pour la table « parc ». Enfin, le Writer a été connecté aux configurations préalablement établies sur les données chargées, et l'exécution a été lancée pour effectuer l'intégration des données.

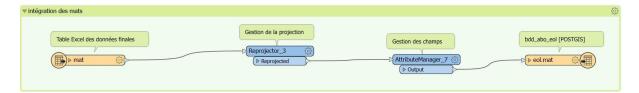


Figure 18: Intégration des données sur les mâts éoliens

3.4. Interfaçage avec QGIS

Une fois la BDD créée et les données intégrées, la prochaine étape consiste à les rendre accessibles aux utilisateurs. Pour ce faire, il a été décidé d'utiliser l'interface du logiciel QGIS. Ce choix est judicieux car il permet de protéger les données (car les données seront accessibles en consultation et non en modification), tout en offrant aux utilisateurs une interface visuellement cartographique, attrayante et plus simple à utiliser.

3.4.1. Création des enveloppes

Il est connu que les parcs éoliens englobent les mâts éoliens, c'est-à-dire qu'un parc éolien comprend les éoliennes qui le composent. La symbolisation des éoliennes peut être représentée par des points simples, tandis que pour les parcs, une enveloppe autour des éoliennes est nécessaire. Sur QGIS, cela implique la création de zones autour des éoliennes, suivie d'un regroupement par champ. Cependant, il peut arriver que si les éoliennes ont des espaces trop importants, elles ne soient enveloppées que partiellement, ce qui peut conduire à la création incomplète des parcs. C'est pour remédier à cette situation que la création d'enveloppes convexes a été privilégié. Grâce à ces figures géométriques, tous les points de l'ensemble sont à l'intérieur ou sur la frontière de l'enveloppe.

Pour plus d'efficacité et de pertinence, FME a été utilisé. Initialement, un « AttributeKeeper » a été employé pour prendre en compte les coordonnées géographiques des éoliennes. Ensuite, un « VertexCreator » suivi d'un « Reprojector » ont été utilisé pour mettre les coordonnées dans le bon système. Les zones tampons autour des éoliennes ont été générées à l'aide du Transformers « Bufferer ». Ces zones tampons ont été liées par l'identifiant du parc grâce à un « Aggregator ». Le Transformers « HullReplacer », a permis la création des zones convexes. Après ces étapes, une fusion entre les géométries créées et les parcs a été réalisée à l'aide du « FeatureMerger ». Enfin, cette fusion a été intégrée à la base de données PostGIS.

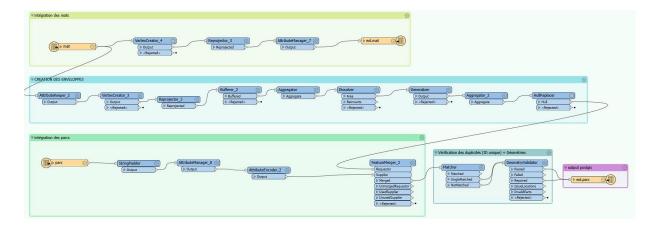


Figure 19: Workflow de la création des enveloppes parcs

3.4.2. Connexion à la base de données sur QGIS

L'affichage des données sur QGIS nécessite une connexion à la BDD. Il est possible de le faire sur QGIS car il prend en charge une variété de formats de bases de données, y compris PostgreSQL/PostGIS. Pour se connecter à une base de données PostGIS, il faut fournir les informations suivantes : l'hôte, le port, le nom et le mot de passe de la BDD.

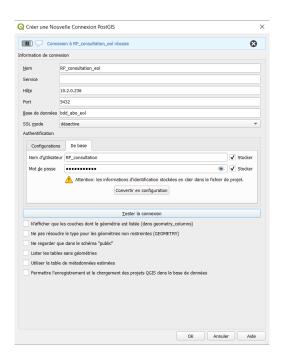


Figure 20: Configuration de la connexion à la BDD PostGIS

Une fois la connexion à la base de données établie, les tables s'affichent, et il suffit de les intégrer en tant que couches en effectuant un double-clic sur chacune d'elles.

L'affichage des couches est standard, sans fond de carte ni symbologies illustratives. Cela a nécessité une mise en forme ultérieure pour améliorer la visualisation.

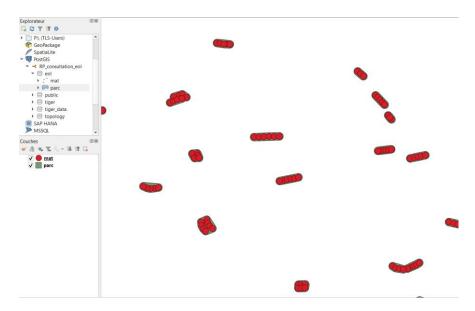


Figure 21: Affichage standard des données sur QGIS

3.4.3. Mise en forme de la visualisation des données

Dans un premier temps, des fonds de carte ont été ajoutés pour contextualiser géographiquement les données à visualiser. Ensuite, la configuration des symbologies des mâts et parcs a été nécessaire pour assurer une compréhension visuelle aux utilisateurs.

Concernant la symbologie des mâts éoliens, le choix s'est porté sur un symbole SVG représentant les pales et un symbole simple en arrière-plan pour renforcer la visibilité. Des couleurs non agressives ont été sélectionnées dans le but d'améliorer l'esthétique visuelle.

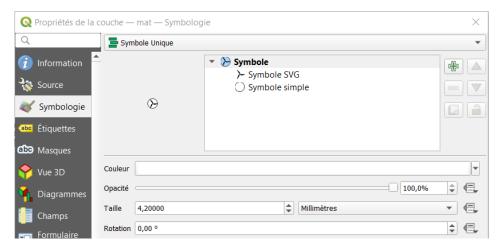


Figure 22: Configuration de la symbologie des mâts dans QGIS

En ce qui concerne la symbologie des parcs, une approche de catégorisation a été choisie pour distinguer les parcs selon leurs statuts administratifs. Diverses couleurs, accompagnées d'une légende, ont été intégrées dans le but de clarifier et de maintenir une compréhension intuitive pour les utilisateurs.

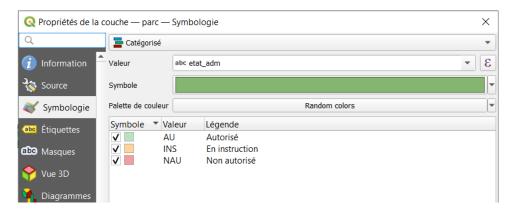


Figure 23: Personnalisation de la symbologie des parcs dans QGIS

Après l'achèvement de toutes ces configurations, le résultat est désormais plus visuel et agréable (cf. figure 24).



Figure 24: Visualisation après mise en forme de l'affichage des données dans QGIS

3.4.4. Phase de test

Le 13/12/2023, une évaluation de la base de données a été réalisée via l'interface QGIS par quelques utilisateurs. L'objectif était de leur présenter l'interface, de les familiariser avec l'interaction des données via les requêtes SQL depuis QGIS, d'explorer divers aspects des données, et de recueillir leurs retours afin d'apporter des améliorations pour mieux répondre à leurs attentes.

Au préalable, un formulaire avait été élaboré pour recueillir les besoins et les requêtes que les utilisateurs envisageaient pour la base de données. En se basant sur les réponses obtenues, un guide de test a été rédigé (*cf. figure 24*). Étant donné que tous les utilisateurs n'étaient pas familiers avec le logiciel QGIS, le guide devait être explicite, détaillé et clair. De plus, un fichier QGIS incluant les fonds de carte (image satellite et fond openstreetmap) sans les couches a été préparé.



Figure 25: QR Code du fichier guide du test

Pendant le test, les utilisateurs ont suivi progressivement le guide à leur propre rythme. Ma présence consistait à répondre aux questions et à débloquer les utilisateurs si nécessaire. Les remarques instantanées ont été faites et précisées sur le chat de la réunion.

Afin de prendre en compte les retours des utilisateurs sur des aspects spécifiques, un formulaire a été élaboré. Les réponses ont ensuite été prises en compte pour améliorer le travail fourni.

3.5. Rédaction des livrables

Au terme de ce stage au sein d'ABO Wind France, il était essentiel de consolider les résultats des travaux par la rédaction de livrables. Ces documents ont pour objectif de détailler le travail effectué, d'apporter des explications claires et de proposer un plan de maintenance pour assurer la pérennité de la base de données géospatiale nouvellement créée. Cela permettra également à la prochaine personne qui reprendra le travail de mieux saisir ce qui a été déjà réalisé et ce qui reste à construire.

Les livrables étant les suivants :

• Livrable n°1 : le Descriptif de la mise en place de la BDD se concentre sur la description exhaustive du travail accompli tout au long du stage. Il couvre la méthodologie adoptée, les étapes de traitement des données, les choix techniques effectués (comme l'utilisation de FME et QGIS), et les résultats obtenus. Ce document vise à offrir une vision complète du processus,

de l'intégration des données à la visualisation dans QGIS. La proposition d'un plan de maintenance visant à assurer la durabilité et la mise à jour continue de la BDD au fil du temps y est inclus.

- Livrable n°2: le Support d'explications des requêtes SQL, a été élaboré dans le but d'offrir une assistance claire aux utilisateurs pour comprendre la nature des requêtes SQL et leurs relations avec le logiciel QGIS. Cette ressource vise à simplifier la compréhension des aspects techniques de l'utilisation de requêtes SQL dans le contexte spécifique de QGIS, facilitant ainsi une utilisation plus efficace et autonome de la BDD.
- Livrable n°3: le Catalogue de données, ce document offre une vue d'ensemble exhaustive de la structure de la base, décrivant chaque table, ses champs et les types de données utilisés. Il s'accompagne de commentaires expliquant la signification de chaque élément.

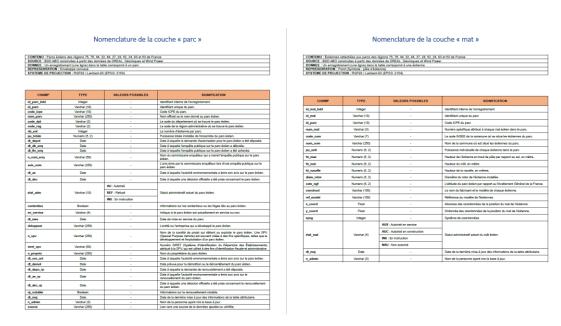


Figure 26: Aperçu du catalogue de données

CONCLUSION

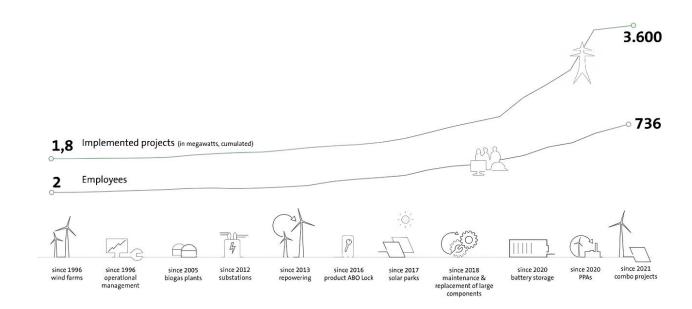
1. Conclusion technique

En somme, l'objectif était de consolider des données provenant de BDD distinctes pour établir une BDD géospatiales cohérente et fiable, englobant les informations sur les éoliennes et leurs parcs associés dans onze régions préalablement sélectionnées. Les principales tâches comprenaient l'analyse des données existantes, le traitement des données via des fusions et jointures avec le logiciel FME, la comparaison et la mise en forme des données, la création de la base de données avec la génération du code SQL via JMerise, l'intégration des données avec FME, et enfin l'interfaçage QGIS, incluant la mise en forme des fonds de carte et des symbologies. Ces tâches ayant été réalisé sur les données de quelques régions, il reste donc un travail à effectuer pour prendre en compte toutes les régions et ainsi avoir la base de données complète. Ces opérations ont été effectuées sur les données de quelques régions, laissant encore un travail à accomplir pour les régions restantes et ainsi obtenir une BDD géospatiales complète.

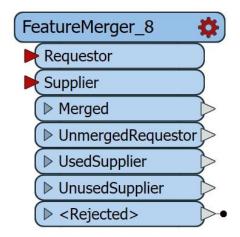
2. Conclusion personnelle

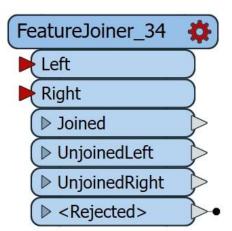
Au terme de ce stage, j'ai pu acquérir une diversité de compétences et de connaissances, allant de la manipulation de logiciels spécialisés tels que QGIS, PgAdmin4 et FME, à la conception de bases de données géospatiales en passant par le langage SQL. Cette expérience m'a également permis d'approfondir ma compréhension pratique des systèmes de gestion de bases de données relationnelles et d'explorer le domaine de la géomatique, un domaine que je ne connaissais pas auparavant. Outre les compétences techniques, j'ai renforcé ma capacité à organiser et planifier des tâches. De plus, j'ai également pu renforcer mes aptitudes en communication et ma capacité à travailler de manière synergique. Ainsi, ce stage a été une opportunité enrichissante qui m'a permis de développer des compétences techniques, de découvrir un nouveau domaine professionnel et d'appréhender les défis liés à la gestion des données géospatiales dans le secteur éolien. Je suis reconnaissante pour cette opportunité d'apprentissage et de croissance personnelles au sein d'ABO Wind France.

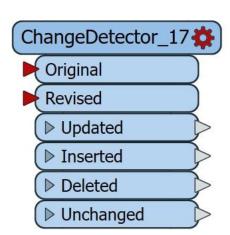
ANNEXE A : Évolution du groupe ABO Wind de 1996 à 2021



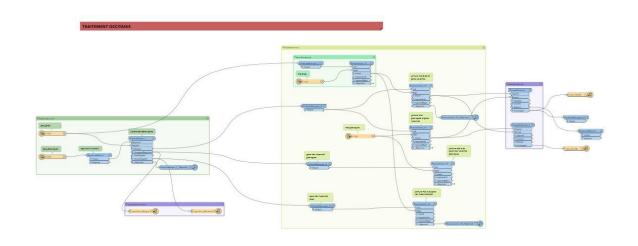
ANNEXE B: Quelques Transformers FME utilisés

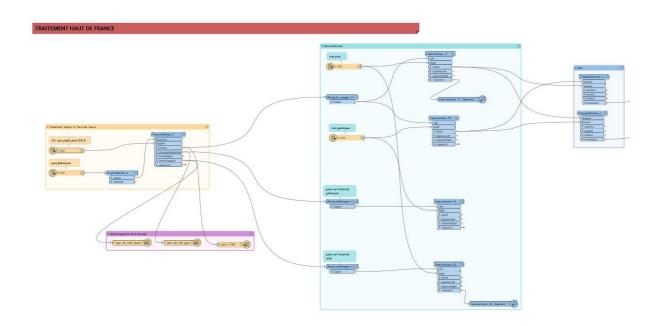






ANNEXE C : Visualisation globale de quelques Workflows des traitements FME





ANNEXE D : Tableau comparatif des champs des tables entre les BDD

ATTRIBUTS DES BDD														
TABLES PARCS_DREAL PAR REGION								BDD GEORISQUES						
AQUITAINE	AUV-RH	BFC	BRETAGNE	CENTRE VAL DE LOIRE	GRAND EST	HAUT DE France	NORMANDIE	OCCITANIE	PACA	PAYS DE	LA LOIRE 2	Aerogenerateur national	Parc national	Poste livraison national
alt_base	alt_base	alt_base	alt_base	AC_NON_RAC	ALTITUDE_B	alt_base	codeCom	alt_base	ALT_BASE	alt_base	altitude_b	code_dept	code_dept	code_dept
date_demol	date_demol	date_demol	datdecisi0	ALT_BASE	ALTITUDE_S	code_arron	codeDep	c_mat_diff	DATE_DEMOL	date_demol	date_depot	code_insee	code_insee	code_insee
date_prod	date_prod	date_prod	datdecisio	DATE_DEC	DEPT	code_com	hauteur	d_service	DATE_PROD	date_prod	date_derni	code_reg		code_reg
date_real	diam_rotor	date_real	date_demo0	DATE_DEMOL	EPSG	code_icpe	idMat	date_demol	DATE_REAL	date_real	date_reali	constructeur	contentieux	date_maj
diam_rotor	en_service	diam_rotor	date_demol	DATE_MAJ	ID_EOL	contentieu	idParc	date_prod	DIAM_ROTOR	diam_rotor	demandeur	cote_ngf	date_arrete_prefectoral_complementaire	epsg
en_service	etat_icpe	etat_mat	date_prod	DATE_PROD	ID_MAP	date_crea	libCom	date_real	EN_SERVICE	en_service	etat_aut_1	date_maj	date_cessation_activite	id_parc
etat_mat	etat_mat	ht_mat	date_prod_	DEPT	ID_S3IC	date_decis	modele	dept	ETAT_MAT	etat_au_ma	etat_autor	date_mise_en_service	date_debut_construction	id_poste_livraison
gid	gid	ht_max	date_real	DETRUIT	MANDATAIRE	date_depot	puissance	diam_rotor	HT_MAT	etat_mat	etat_eolie	diametre_rotor	date_debut_demantelement	libelle
ht_mat	ht_mat	ht_nacelle	diam_rotor	DIAM_ROTOR	NOM_PARC	date_maj	sourceGeom	e_adm_mat	HT_MAX	gid	etat_servi	epsg	date_delivrance_autorisation	nom_commune
ht_max	ht_max	id_au	en_service	EN_COURS	PUISSANCE	date_prod	Xlamb93	en_service	HT_NACELLE	ht_mat	gid	hauteur_mat_nacelle	date_delivrance_autorisation_repowering	x_poste_livraison
ht_nacelle	ht_nacelle	id_mat	etat_mat	EN_SERVICE	TYPE_ETAT	date_real	Ylamb93	etat_mat	ID_AU	ht_max	hauteur1	hauteur_totale		y_poste_livraison
id_au	id_au	id_parc	etat_mat_l	ETAT_AU	X	departemen		gid	ID_MAT	ht_nacelle	hauteur_ma	id_aerogenerateur	date_delivrance_avis_autorite_environnementale_repowering	
id_mat	id_mat	insee_com	ht_mat	ETAT_ICPE	Υ	diam_rotor		ht_mat	ID_PARC	id_au	id_mat	id_parc	date_delivrance_recipisse_prefecture	
id_parc	id_parc	leg_mat	ht_max	ETAT_MAT		en_service		ht_max	ID_POSTE	id_mat	id_mat_KTL	libelle	date_delivrance_refus	
id_poste	id_poste	nom_com	ht_nacelle	ETAT_N		etat_eolie		ht_nacelle	ID_ZDE	id_parc	nb_mats_au	nom_commune	date_delivrance_rejet	
id_zde	id_zde	nom_parc	id_au	ETAT_PC		etat_mat		i_mat_anc	N_PARCEL	id_poste	nb_mats_de	nom_eolienne	date_depot_demande_autorisation	
n_com_mat	n_parcel	numero	id_mat	ETAT_T		etat_proce		i_parc_anc	NUMERO	id_zde	numero_par	periode_allumage_desc	date_depot_demande_autorisation_repowering	
n_dpt_mat	nom_parc	pu_nominal	id_parc	HT_MAT		exploitant		i_pc_anc	PRECIS_POS	n_parcel	puissanc_1	periode_allumage_lib	date_maj	
n_parcel	numero	srce_geom	id_poste	HT_MAX		gardesol		id_au	PU_NOMINAL	numero	puissance_	puissance	date_mise_en_service	
numero	precis_pos	sys_coord	n_parcel	HT_NACELLE		ht_mat		id_mat	SRCE_GEOM	precis_pos		reference_modele	date_repowering	
precis_pos	srce_geom	type_eol	numero	ID_AU		ht_max		id_parc	SYS_COORD	pu_nominal		type_feu_desc	etat_parc	
pu_nominal	sys_coord	x_pc	pu_nominal	ID_DREAL		ht_nacelle		id_poste	X_PC	srce_geom		type_feu_lib	id_parc]
srce_geom	x_pc	y_pc	srce_geom	ID_ICPE		id_eolienn		id_zde	Y_PC	sys_coord		x_aerogenerateur	nom_commune	
sys_coord	у_рс		sys_coord	ID_IGN		id_map		insee_com		x_pc		y_aerogenerateur	nom_exploitant	
valid_gest			x_pc	ID_MAT		id_mat		n_commune		у_рс			nom_usuel	
x_pc			у_рс	ID_PARC		id_parc		n_mat_diff					repowering	
у_рс				INSEE_COM		id_pc		n_parc					repowering_notable	
				N_PARCEL		id_sre		n_parcel					repowering_substantiel	
				NOM_COM		n_parcel		numero]				siret	
				NOM_PARC		nom_commun		obs					statut_admin	
				PU_NOMINAL		nom_parc		precis_pos					statut_parc	
				RACCORDE		operateur		pu_nominal					type	
				REFUSE		precis_pos		srce_geom]					
				SRCE_GEOM		puissanc_2		sys_coord						
				×		secteur		x_icpe						
				Y		srce_geom		xcoord						
						sys_coord		y_icpe				1		
						type_proce		ycoord						
						x_pc						1		
						x_rgf93						1		
						у_рс								
						y_rgf93						1		