



HAL
open science

Utilisation de grandes quantités de données au service de la modélisation pour la transition énergétique

Robin Girard

► **To cite this version:**

Robin Girard. Utilisation de grandes quantités de données au service de la modélisation pour la transition énergétique. Sciences de l'ingénieur [physics]. Université Grenoble - Alpes, 2020. tel-04028236

HAL Id: tel-04028236

<https://minesparis-psl.hal.science/tel-04028236>

Submitted on 25 Mar 2023

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.

Document de préparation à l’habilitation à diriger
des recherches.

Utilisation de grandes quantités de données au
service de la modélisation pour la transition
énergétique

Robin Girard
MINES-ParisTech /PSL Research University
Département Énergétique et Procédés / PERSEE / Groupe ERSEI

Soutenue le 12/02/2020.

Composition du jury :

Hervé Gueguen – Professeur des universités – Centrale Supélec-Paris-Saclay (rapporteur)
Manuel Matos – Professeur – Université de Porto, Portugal (rapporteur)
Benoit Robyns – Directeur de recherche – Université Lille 1(rapporteur)
Nouredine Hadjsaid – Professeur des universités – Grenoble INP
Georges Kariniotakis – Directeur de recherche – MINES ParisTech
Marc Petit – Professeur des universités – Centrale Supélec-Paris-Saclay

Introduction

Après une thèse en mathématiques appliquées au laboratoire Jean Kuntzmann de l'université de Grenoble, j'ai pris le virage de l'énergétique et je suis arrivé directement à MINES ParisTech au centre qui s'appelle maintenant le centre Procédés, Énergies Renouvelables et Systèmes Énergétiques (PER-SEE) centre de recherche commun de MINES ParisTech et d'ARMINES, à Sophia Antipolis. Dès lors j'ai évolué progressivement en tant que chercheur dans quelques-uns des multiples aspects du monde de l'énergie au travers de projets, d'encadrement de thèses et de stages, en tant qu'enseignant prenant des responsabilités au sein de l'école dans les enseignements existant ou montant de nouveaux cours, mais également en tant que citoyen en intervenant dans des projets et des conférences pour le compte de la transition énergétique.

L'objet de ce document est de proposer une synthèse de mes activités de recherche depuis 2008 à MINES ParisTech et de conclure par quelques perspectives. Ces travaux s'inscrivent dans un contexte que je présente dans la Section 1 et qui s'articule autour des trois éléments importants que sont la transition énergétique, la libéralisation et l'évolution des bases de données. La synthèse, développée dans la Section 2, est structurée en 3 sous-thématiques : la prédiction à court terme, la modélisation du système électrique et du marché de l'électricité, et la modélisation prospective du territoire et de la charge à partir de grandes bases de données. Pour finir, j'évoque quelques pistes pour ma recherche à venir dans la Section 3. Cette synthèse est l'occasion d'évoquer mes publications (conférences et articles), mais aussi certaines de mes activités d'enseignement, mes activités de rayonnement scientifique, les thèses et les stages que j'ai encadrés et les projets de recherche auxquels j'ai contribué. À chacun de ces thèmes est dédié une section contenant une liste exhaustive.

Table des matières

1	Contexte	6
1.1	Transition énergétique	6
1.2	Libéralisation	11
1.3	Place de la donnée dans l'analyse des systèmes énergétiques .	11
2	Description de la recherche effectuée	13
2.1	Prédiction à court terme	13
2.1.1	Structure des modèles de prédiction et particularités de la prédiction de production photovoltaïque (PV) . .	15
2.1.2	Prédictions probabilistes : quantiles et scénarios	16
2.1.3	Prédiction de consommation locale	19
2.1.4	Modèles statistiques spatio-temporels	22
2.1.5	Prédiction de rampes de production	27
2.2	Modélisation et simulation du système électrique et du marché de l'électricité pour la gestion et la planification	31
2.2.1	Optimisation de la flexibilité pour l'équilibre offre-demande du système centralisé	33
2.2.2	Optimal Power Flow et réseau de distribution	38
2.2.3	Modélisation et simulation du marché de l'électricité .	45
2.3	Énergie et territoire, modélisation prospective à partir de grandes bases de données	51
2.3.1	Des bases de données pour l'analyse et l'optimisation de la transition énergétique dans les territoires	53
2.3.2	Modélisation de la consommation électrique	58
2.3.3	Modélisation et simulation de la production renouve- lable et des erreurs de prédiction à court terme	68
2.3.4	Sensibilité des estimations de gisements territoriaux pour la transition et des analyses de cycle de vie. . . .	71
3	Projet de recherche	78
4	Contrats de recherche et rayonnement scientifique	82
4.1	Liste des contrats de recherche et description de ma participation	82
4.2	Rayonnement scientifique	84
5	Enseignement et encadrement	85
5.1	Thèses encadrées et publications associées	85
5.2	Stages : encadrements et tutorats	86

5.3 Enseignements	88
6 Bibliographie personnelle	89
6.1 Articles publiés dans des revues scientifiques à comité de lecture	89
6.2 Articles soumis dans des revues à comité de lecture	92
6.3 Conférences	92
7 Bibliographie	98
8 HDR Robin Girard - Articles publiés dans des revues scientifiques à comité de lecture	100
CV	102

1 Contexte

1.1 Transition énergétique

La transition énergétique en France est souvent résumée par les points de la Figure 1, même si aujourd’hui l’objectif des 50 % de production nucléaire a été repoussé à 2030-2035. Dans ce cadre, en France ou dans d’autres pays dans le monde, le système électrique est appelé à un renouvellement profond qui soulève un certain nombre de questions. Cela concerne sa structure, son organisation, la consommation et la production d’électricité. Pour les pays qui entreprennent aujourd’hui une telle transition, les motivations sont nombreuses. En première ligne, il y a bien le changement climatique, mais aussi la raréfaction progressive des ressources fossiles, le besoin d’accroître l’indépendance énergétique, le coût et l’impact du nucléaire (risque et déchets). En France, cette transition est aussi liée au vieillissement du parc nucléaire actuel, à sa taille et au fait qu’il ait été construit en 15 ans seulement. Dans tous les cas, le système électrique est appelé à jouer un rôle important et les défis autour de cette transition, qui méritent un traitement quantitatif et scientifique, sont nombreux.

Dans l’exemple du système français, la loi de transition énergétique se traduit à la fois au niveau de la production et de la consommation. Concernant la production à partir de sources renouvelables, son augmentation repose en partie sur le vecteur électricité avec le développement de l’éolien et du photovoltaïque, mais le nucléaire n’est pas écarté. Concernant la consommation, c’est d’abord la maîtrise de la demande, en particulier celle du chauffage, qui doit jouer un rôle important via des rénovations de bâtiments. Mais l’électrification du transport, avec le véhicule électrique, et du chauffage, via l’utilisation de Pompes à Chaleur (PAC), vont jouer un rôle important et faire évoluer les caractéristiques de la consommation d’électricité. L’ensemble de ces mesures doit se faire sous le contrôle de l’évolution des émissions de gaz à effet de serre dont le détail en France pour 2016 est donné dans la Figure 1.

Nécessité d’une modélisation avancée dans l’évaluation des coûts.

Dans la Figure 1 on peut voir que les émissions de CO₂ du secteur de l’électricité en France sont assez faibles (de l’ordre de 50 *MT* de *CO_{2eq}*) du fait de l’importance du parc nucléaire. Aussi, si le rôle du développement des EnR électriques dans la baisse des émissions de GES dans le monde est incontestable, il faut observer que ces moyens de production ne feront pas vraiment évoluer un mix électrique français déjà relativement performant sur

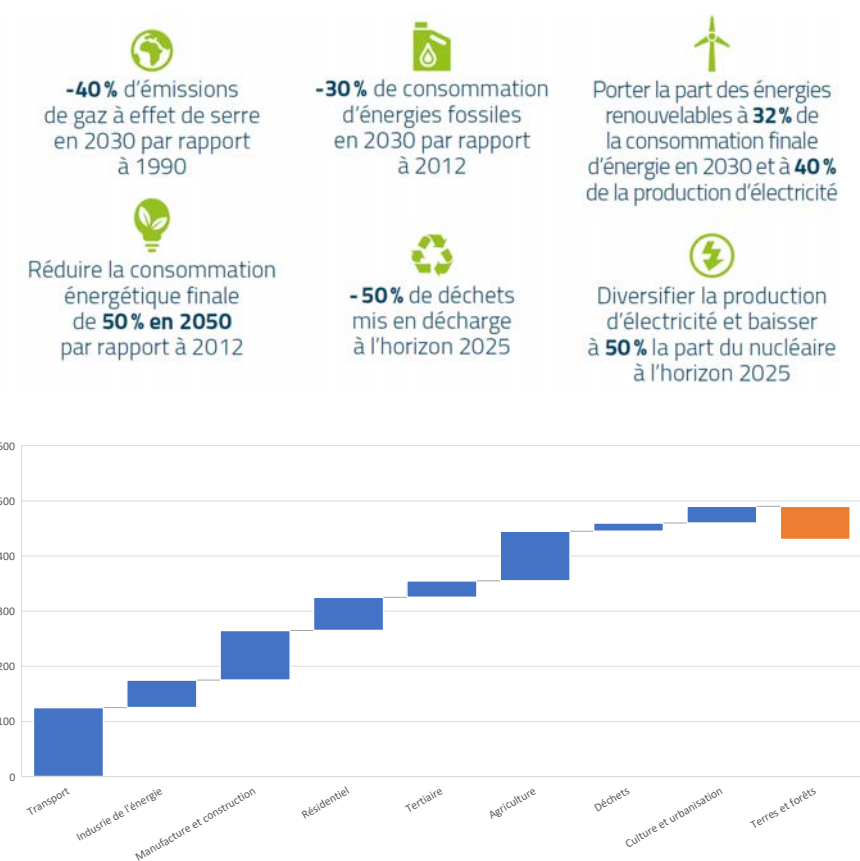


FIGURE 1 – Haut : **La transition énergétique en France** est menée dans le cadre d'une loi : la loi de transition énergétique pour la croissance verte dont les objectifs clés sont rappelés ici.

Bas : **Émissions de CO2 de différents secteurs en France en 2016.** Il faut noter ici que les émissions de l'industrie de l'énergie sont 80 % liées à la production d'électricité et que les émissions du secteur résidentiel sont 84 % dues au chauffage. Une grande partie du tertiaire l'est aussi. Sources <https://www.statistiques.developpement-durable.gouv.fr>

le plan des émissions de GES. Sur l'importance que prendront ces moyens de production par rapport au nucléaire dans le futur, les différents scénarios prospectifs français (comme ceux de l'ADEME¹ qui complètent la stratégie nationale bas carbone) restent assez ouverts. Dans le monde, le GIEC suggère également un développement des énergies renouvelables et du nucléaire. En France, la question soulève un certain nombre de problèmes complexes -le nucléaire reste un élément important de la stratégie industrielle française, mais implique certains risques-, mais la difficulté du choix entre ces énergies est qu'à ces enjeux variés et difficilement quantifiables s'ajoute un problème technico-économique lié à la comparaison entre le coût de production d'un système renouvelable reposant sur des moyens de production intermittents et celui d'un système plus nucléaire reposant sur une puissance pilotable.

En effet, le coût actualisé (Levelized Cost of Electricity - LCOE) des moyens de production renouvelables a beaucoup baissé ces dernières années. La production d'énergie renouvelable est devenue compétitive avec les moyens de production conventionnels² mais ces coûts ne sont qu'une partie du coût du système électrique et il s'agit aussi d'être capable de quantifier les coûts liés aux contraintes d'équilibre offre demande sur le système électrique, ainsi que ceux liés au réseau de transport et de distribution. Ces coûts ne reflètent pas non plus les impacts environnementaux des filières ainsi que les risques futurs d'épuisement de ressources minérales.

Plusieurs difficultés sont en effet posées par cette **production renouvelable**. Sans faire un inventaire exhaustif de celles-ci on peut relever les plus importantes. D'un côté, la production renouvelable est **variable et non complètement prédictible**, ce qui pose des problèmes à plusieurs échelles d'espace et de temps. La Figure 2 fait la synthèse des différentes modalités

1. Voir <https://www.ademe.fr/actualisation-scenario-energie-climat-ademe-2035-2050>

2. Le coût de production du PV et de l'éolien est de plus en plus bas et en 2017/2018 en France on a pu voir des projets éoliens et PVs se monter pour des tarifs d'achat autour de 50–60 euros/MWh seulement, voir <https://www.ccomptes.fr/sites/default/files/2018-04/20180418-rapport-soutien-energies-renouvelables.pdf> tableau 3 p53 . Au Portugal, l'été 2019, Akuo Energy s'est engagé à produire de l'énergie PV pour moins de 20 euros/MWh et d'autres projets dans le monde ont vu le jour à des prix similaires. Dans le même temps le coût du nucléaire en France est similaire à celui du PV et de l'éolien : le coût de production des centrales rénovées est estimé à 65–70 euros/MWh par la Cour des comptes (apport annuel de 2016). Enfin, pour ce qui est de la nouvelle génération de centrales EPR, son coût aujourd'hui est en Europe autour de 110 euros/MWh. Ces chiffres reflètent le coût d'une tête de série et pas une situation stable , le coût futur est difficile à estimer et très sensible au taux d'actualisation ou à la durée de construction. Pour le renouvelable ils résultent d'une baisse considérable, alors qu'ils sont plutôt à la hausse pour le nucléaire (aujourd'hui la production nucléaire du parc amorti de EDF est vendue à 42 euros/MWh), ce qui indique que les rapports se sont inversés ces dix dernières années.

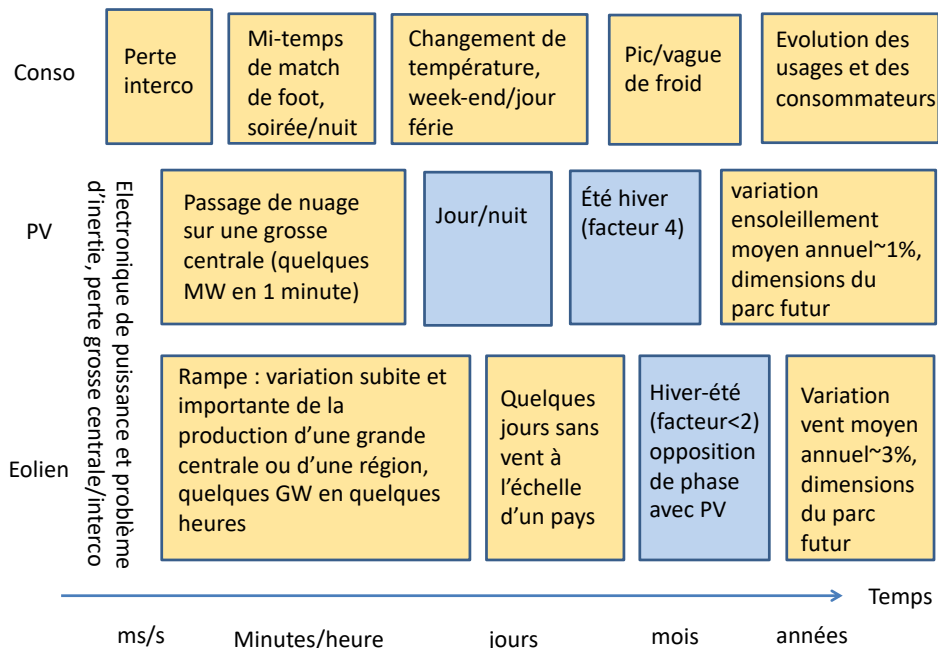


FIGURE 2 – Quelques exemples de **variations** de la consommation et de la production (PV/éolienne) à différentes échelles de temps. En bleu, les variations qui n'impliquent pas vraiment de problème de prédictibilité.

de ce problème. De l'autre, la production renouvelable est **décentralisée**. Ceci implique d'une part qu'elle est connectée majoritairement au réseau de distribution d'électricité dont la gestion et le dimensionnement doivent être repensés. D'autre part, la planification énergétique demande une articulation à des échelles territoriales très variées et la contrainte d'utilisation du territoire peut être forte.

Dans le même temps, les évolutions de la **consommation** impliquent de nouvelles distributions des charges sur le réseau électrique et une évolution de la variabilité à différentes échelles de temps. La réglementation thermique de 2020 favorisera le chauffage électrique au détriment du gaz. Les pompes à chaleur (PAC) se développent pour répondre à cette nouvelle demande mais leur efficacité en période de grand froid reste une question ouverte. Elles peuvent impliquer localement des appels de puissance importants et l'évolution de la thermosensibilité de la consommation électrique sur le moyen/long

terme reste difficile à évaluer. Les véhicules électriques vont eux aussi poser des problèmes nouveaux sur le réseau, en termes d'appel de puissance et de variabilité.

Pour autant, la consommation, tout comme la production renouvelable, ne doivent pas être regardées comme de simples contraintes, car elles sont aussi, et à plus d'un titre, un levier pour la transition énergétique. En effet, la consommation électrique qui était jusqu'à maintenant une donnée du problème que l'on cherchait à estimer à l'avance est appelée à prendre de plus en plus de place dans les processus de décisions (planification et gestion) du système énergétique. Sur le plan de la planification, cela inclut la maîtrise de la thermosensibilité de la consommation via un contrôle de l'efficacité des bâtiments (rénovation, réglementations thermiques) et un suivi des modes de chauffage utilisés, le contrôle de l'avènement des nouveaux usages comme le véhicule électrique. Sur le plan de la gestion, cela inclut le pilotage d'une partie de la demande comme l'eau chaude sanitaire ou le véhicule électrique. De même, la production renouvelable peut être modulée, dans la limite de ce que permet le vent ou le soleil (et donc plutôt à la baisse), afin de lever une contrainte réseau ou pour participer à l'équilibrage de l'offre et de la demande, son implantation sur le territoire peut être coordonnée pour en limiter les impacts comme cela a été fait par exemple dans le cadre des S3REnR.

Enfin, dans cette transition énergétique, les différents vecteurs que sont l'électricité, le gaz et la chaleur doivent être pensés dans leurs complémentarités, ce qui implique des couplages nouveaux et des optimisations plus complexes. Ici, comme dans le système électrique, ces couplages concernent un vaste spectre d'échelles de temps et d'espace. On pense par exemple à la production de gaz (méthane/dihydrogène) pour absorber le surplus de production photovoltaïque en été, offrant ainsi une flexibilité saisonnière au système centralisé, ou pour soulager le réseau de distribution en levant des contraintes posées par une production locale. La production de gaz vert et sa place vis-à-vis de l'utilisation de biomasse posent un grand nombre de questions en termes de développement durable et d'utilisation du territoire. En hiver, la micro-cogénération (chaleur/électricité à partir de gaz) est parfois présentée comme une solution pour maîtriser la thermosensibilité électrique. Cette transition énergétique est donc à l'origine de nouveaux problèmes d'optimisation et de statistique pour la gestion et la planification qui sont sans doute plus complexes que ceux que nous avons connus jusqu'à maintenant. En outre ces problèmes font appel à des champs disciplinaires très variés (optimisation, statistiques, énergétique, géographie, économie, etc.) qu'il s'agit

d'utiliser conjointement et de manière équilibrée.

1.2 Libéralisation

Ces changements en cours se font dans un environnement réglementaire transformé par la volonté de la Commission européenne de libéraliser le secteur de l'énergie. En France, ceci s'est en particulier traduit par le démantèlement du monopole EDF-GDF, avec la création, en ce qui concerne l'électricité, d'un gestionnaire du réseau de transport d'électricité RTE, d'un gestionnaire de la majeure partie du réseau de distribution d'électricité ENEDIS, d'un producteur fournisseur EDF, et l'apparition progressive de producteurs et de fournisseurs alternatifs interagissant sur un marché de l'électricité. Aujourd'hui, on ne peut pas dire que cet environnement soit complètement stabilisé et un certain nombre d'ajustements importants sont nécessaires, notamment pour maintenir la qualité d'approvisionnement tout en s'assurant de l'efficacité des investissements dans le système électrique tant sur le plan de la production que du réseau ou même de l'efficacité énergétique.

Plusieurs évolutions récentes ou en cours dans les mécanismes de marché transforment la manière dont se développe le système électrique. Parmi celles-ci, certaines concernent particulièrement les producteurs renouvelables, comme la sortie progressive des tarifs d'achat, mais d'autres concernent de manière plus générale des groupes de consommateurs et de producteurs qui peuvent être regroupés dans une même sous-partie du réseau (de distribution), on parle alors de micro-réseau (micro-grid noté MG), ou disséminés à différents endroits du réseau de transport, on parle alors de centrales virtuelles (virtual power plant noté VPP). Les VPP vont être autorisés à participer au marché de la réserve primaire, éventuellement avec des offres asymétriques, les MG sont amenés à jouer un rôle important dans le cadre de réglementations autorisant l'autoconsommation collective. Mais cette réglementation remet en cause un principe important du service public de l'électricité : la péréquation tarifaire.

La modélisation du système électrique actuel ne peut pas passer à côté de ces nouvelles formes d'organisation économique. Les enjeux scientifiques et sociaux apportés par la libéralisation sont à ce titre assez nombreux.

1.3 Place de la donnée dans l'analyse des systèmes énergétiques

La transition énergétique est concomitante d'une révolution numérique qui a débuté au début des années quatre-vingt-dix : augmentation des puis-

sances de calculs et de stockage de données (que l'on quantifie souvent via la loi de Moore), mise à disposition libre et simple de grandes quantités de données (on parle de *l'open data*). Le monde de l'énergie est un champ de l'industrie qui n'échappe pas à ce développement des technologies de l'information. L'augmentation de la quantité d'informations observées ainsi que du nombre d'appareils commandés est une partie importante de la modernisation à venir du système électrique. Les données correspondantes proviendront d'une quantité accrue d'observations de la production, de la consommation, du réseau, mais aussi d'autres informations sur les consommateurs, et plus généralement sur le territoire. Cette modernisation doit permettre d'accroître sa capacité à intégrer au mieux de nouvelles sources de productions renouvelables et des usages électriques en constante mutation, en ce qui concerne la phase de gestion (via plus d'observation et de contrôle des productions et consommations), mais aussi la phase de planification (via une simulation prospective impliquant une description spatiale et temporelle du territoire plus fine).

Comme bien souvent avec l'accroissement de la quantité de données observées, leur utilisation soulève plus de questions que de solutions. L'exploitation de ces données pour l'amélioration de l'opération et de la planification du système passe par l'élaboration de modèles statistiques et de méthodes d'optimisation adaptés à la taille des problèmes et des données, tirant parti de structures sous-jacentes de plus en plus complexes. La quantification des bénéfices de cette complexité est toujours une question difficile et importante. La quantification des impacts de ces nouvelles technologies ne l'est pas moins, il s'agit notamment de savoir évaluer les consommations en énergie et en matières premières de ces éléments dits "intelligents". L'analyse de cycle de vie est pour ce faire un outil pertinent.

2 Description de la recherche effectuée

Dans cette section, je présente mon travail de recherche jusqu'à aujourd'hui en excluant celui de ma thèse sur l'analyse statistique de données en grande dimension, avec une application à l'analyse d'images hyperspectrales. Celui-ci a donné lieu à trois publications qui ne sont pas dans le domaine de l'énergie, mais qui reflètent ma première expérience avec des données en grande dimension [ArticleP4, ArticleP3, ArticleP2].

Cette Section est structurée en 3 sous-thématiques : la prédiction à court terme, la modélisation du système électrique et du marché de l'électricité, et la modélisation prospective du territoire et de la charge à partir de grandes bases de données.

J'ai pris soin de mentionner toutes mes publications en limitant l'utilisation d'autres références que l'on retrouvera dans les articles en annexe. J'ai replacé cette recherche dans un contexte socio-économique dans l'introduction et j'ai cherché, tout au long du texte, à mentionner le lien entre (i) ma recherche académique qui se traduit le plus souvent par des publications et des encadrements de thèse, (ii) ma recherche partenariale qui se traduit par des projets de recherche en partenariat avec d'autres laboratoires de recherche et des industriels et (iii) mes enseignements qui se traduisent par des cours ou des encadrements de projets étudiants. Les contrats de recherche, les publications (conférences ou articles), mes activités d'enseignement et de rayonnement scientifique sont listés dans des Sections dédiées regroupées à la fin de ce document.

2.1 Prédiction à court terme

La production éolienne, la production photovoltaïque et la consommation électrique sont par nature variables. Il est nécessaire de comprendre ces variations pour améliorer les stratégies de gestion du réseau électrique. Ainsi, une solution pour permettre d'accroître la capacité du système électrique à s'adapter -sa flexibilité- à ces nouveaux moyens de production renouvelables intermittents est l'amélioration des prévisions de production, de l'échelle nationale à une échelle plus fine pour des horizons allant de quelques minutes à quelques mois. La prédiction de la production, plusieurs années à l'avance, est aussi un problème en soi, on parle dans ce cas d'estimation et de caractérisation de la ressource. C'est un enjeu pour la planification des investissements dans le système électrique et non pour son opération. Cette section est dédiée à la prédiction à court terme et nous traitons le problème long terme dans la Section 2.2.

J'ai travaillé sur la prédiction à court terme de l'**éolien** dès mon arrivée à l'école des mines, en septembre 2008, dans le cadre de trois différents projets de recherche. J'ai été *project manager* du projet européen FP7 SafeWind [Proj9], j'ai contribué au projet FP6 Anemos.plus [Proj3], et au projet ADEME ENSEOLE [Proj14]. Dans tous ces projets, j'ai contribué par des travaux et la rédaction de rapports. Le projet SafeWind a permis le financement de la thèse de Arthur Bossavy [These2] autour de la prédiction des variations de la production éolienne. C'est la première thèse que j'ai co-encadré. J'ai aussi encadré des stages sur ce sujet [Stages5, Stages6, Stages8].

Dans ces projets, j'ai développé des collaborations avec des statisticiens, des chercheurs de la communauté du système électrique, des opérateurs de systèmes électriques, mais aussi des météorologues. D'une manière générale, ces sujets de recherche et les projets associés ont alimenté le reste de ma recherche, ainsi qu'une partie des cours que j'ai montés autour de la météorologie pour l'éolien [Ens9] et surtout sur le fonctionnement et l'évolution du système électrique dans un contexte de transition énergétique [Ens7, Ens11].

J'ai débuté mon travail sur la prédiction à court terme de la **production PV** dans le cadre du projet européen More Micro-grids [Proj15] et de stages [Stages4, Stages9]. Ce travail s'est prolongé dans le cadre des projets Nice Grid [Proj4] et Berrien [Proj11] financés par l'ADEME, et dans le cadre du projet [Proj13] en partenariat avec un développeur de centrales PV. Ces projets ont permis de financer la thèse de Ghislain Agoua [These11]. Aujourd'hui, nous poursuivons nos travaux sur cette thématique avec Compagnie Nationale du Rhône (CNR) qui finance la thèse CIFRE en cours [These6] sur la prédiction des productions éoliennes et PV à partir de grandes quantités de données et de modèles spatio-temporels. Ces sujets ont également inspiré quelques activités d'enseignement, essentiellement des TP [Ens3, Ens4].

Enfin, la prédiction à court terme de la **consommation locale** est un sujet de recherche plus récent, sur lequel j'ai travaillé principalement dans le cadre d'une partie de la thèse d'Alexis Gerosier [These5] et du projet européen Sensible [Proj12].

Nous présentons ici notre contribution à ce domaine en rappelant quelques éléments sur la structure des modèles de prévision à court terme et les particularités de la prédiction de la production photovoltaïque dans la Sous-section 2.1.1, puis sur les prévisions probabilistes (quantiles et scénarios) dans la Sous-section 2.1.2. Ces deux premières sous-sections sont l'occasion d'évoquer une partie de mes travaux. Elles sont suivies de trois Sous-sections traitant de sujets plus originaux : tout d'abord sur les prédictions à court terme de consommation locale dans la Sous-section 2.1.3, puis dans la Sous-

section 2.1.4 sur le développement de modèles spatio-temporels et enfin sur la caractérisation et la prédiction de rampes dans la Sous-section 2.1.5.

2.1.1 Structure des modèles de prédiction et particularités de la prédiction de production photovoltaïque (PV)

La prédiction de la production éolienne et photovoltaïque, ainsi que de la consommation peuvent rentrer dans un cadre commun pour lequel je vais donner quelques notations. Dans tous les cas, pour un horizon h donné (ou pour un ensemble d’horizons), on cherche à trouver une relation entre la variable aléatoire à prédire $P_{t+h|t} = \mathbb{E}[P_{t+h}|t]$ donnant la puissance (consommée ou produite) à l’instant $t + h$ sachant toutes les informations connues à l’instant t , et ces informations connues à l’instant t . Ces informations peuvent être de la puissance passée mesurée, des mesures distantes M_t (images satellites, observations météorologiques, mesures de centrales voisines, etc.) ou des prévisions météorologiques issues d’un modèle physique NWP_t . La relation recherchée est donnée par une fonction de régression f , et la prédiction faite est de la forme :

$$\hat{P}_{t+h|t} = f(P_t, P_{t-1}, \dots, P_{t-l}, M_t, NWP_t, h).$$

La fonction de régression f peut être obtenue par une démarche qui est purement statistique ou purement physique, mais elle est aussi souvent un mélange des deux.

La prédiction à court terme de la production PV est un domaine de recherche plus récent que celui correspondant dans le monde de l’éolien. Nous avons proposé un état de l’art de la prédiction PV dans [Confs2]. Les différences fondamentales avec le domaine de la prédiction de production éolienne sont que les variables météorologiques qui influencent la production sont maintenant l’irradiation et dans une moindre mesure les variables influant la température des panneaux (température extérieure et vent). La relation entre la production et l’irradiation directe reçue et projetée sur le plan du panneau est relativement linéaire, ce qui implique une étape de projection de l’irradiation prédite dans les modèles. Cette relation linéaire cache surtout une composante déterministe liée à la course du soleil qu’il est important d’intégrer dans la modélisation. En effet, cette composante n’est pas une simple composante périodique journalière puisqu’elle contient aussi des variations saisonnières. Il est donc utile de travailler avec une série temporelle dans laquelle cette composante a été calculée et retirée. Si P_t est la production, nous cherchons donc à travailler avec une série du type

$$u_t = P_t/g(P_t^{sim})$$

où P_t^{sim} est la production simulée en ciel clair et g une fonction bien choisie. Le choix de la fonction g , ou plus généralement de la transformation à appliquer à la production P_t , est souvent appelé stationnarisation, même si le terme n’a pas nécessairement ici rigoureusement le sens que lui donnent les statisticiens. Dans la littérature, la production ciel clair est obtenue par l’application d’un modèle purement physique ou statistique. Les modèles physiques donnent la course du soleil et incluent éventuellement les effets liés à la traversée de l’atmosphère. Les modèles statistiques sont construits à partir d’un quantile extrême appliqué à de longues séries de production observées indicées par l’heure de la journée et le jour de l’année, ou à partir de données sélectionnées par un expert, ou sur la base d’un seuil sur un indice de clarté.

Nos premiers travaux sur ce sujet ont été menés dans le cadre du stage [Stages4]. Puis, dans le cadre du stage [Stages9], nous avons proposé une modélisation au pas de temps de 5 secondes. Partant du constat que les méthodes existantes n’améliorent pas toujours les prédictions et impliquent un certain nombre d’irrégularités dans séries normalisées, notamment le matin et le soir, nous avons proposé dans [ArticleP18] une nouvelle méthode de stationnarisation. La Figure 3 présente les améliorations obtenues. Ces méthodes sont importantes pour la modélisation spatio-temporelle.

Un certain nombre de nos travaux ont été appuyés par des mises en œuvre de nos algorithmes dans un contexte opérationnel. C’était par exemple un des objectifs du projet SafeWind pour l’éolien [Proj9], et des projets Sensible [Proj12] et NiceGrid [Proj4] pour la production PV et la consommation. Ces travaux ont donné lieu à des publications de conférence, pour la prédiction de consommation [Confs40] et PV [Confs17]. Dans ces problèmes, les tailles variées des données disponibles et la qualité des données qui remontent en temps réel est souvent un sujet important. Dans [Confs17] nous avons étudié l’influence de la profondeur de l’échantillon d’apprentissage sur les performances de prédiction à court terme de la production PV. Les résultats de l’expérimentation du projet Nice Grid faite à partir de prédictions probabilistes (i.e. prédictions de quantiles, voir paragraphe suivant) ont pu être présentées dans [Confs35] et dans [Confs21, Confs11].

2.1.2 Prédictions probabilistes : quantiles et scénarios

Dans un certain nombre de cas d’application de la prédiction de la production ou de la consommation, il est utile de savoir prédire non seulement

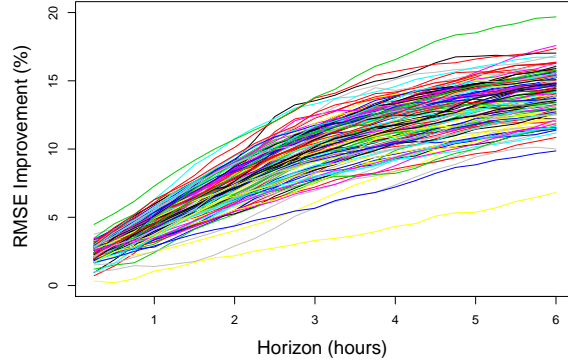


FIGURE 3 – **Amélioration de la prédiction photovoltaïque (PV) obtenue grâce à la stationnarisation des données.** Chaque ligne représente une centrale PV dans le centre de la France. Le pas de temps utilisé est de 15 minutes, et les périodes considérées pour le test et pour l’apprentissage sont d’un an.

la valeur d’une puissance à venir, mais aussi l’incertitude autour de la valeur prédite. En effet, dans certaines configurations météorologiques, par exemple pour le cas de l’éolien si le vent est fort, on peut dire à l’avance que les erreurs de prédiction sont susceptibles d’être plus importantes qu’en moyenne. Ainsi, les méthodes probabilistes visent à la prédiction de certaines caractéristiques du processus stochastique $P_{t+h|t}$. La plupart des approches existantes concernent la prédiction des distributions marginales (i.e. distribution de $P_{t+h|t}$ à h fixé) à partir de méthodes de régression de quantile ou par une modélisation de toute la distribution. Nous avons proposé un état de l’art pour le cas du PV intégrant ces méthodes [Confs2]. Toujours dans le cas du PV et dans le cadre du projet NiceGrid [Proj4] nous avons proposé une nouvelle méthode paramétrique de régression de quantile. Cette méthode a été présentée à la conférence [Confs16] et sa description en détail figure dans le rapport de projet NiceGrid [Proj4] D6.5. Nous avons proposé un exemple d’utilisation de prévision de quantile pour l’optimisation d’une batterie couplée à un système PV dans [Confs20].

Certaines approches visent à modéliser la structure temporelle de dépendance entre les prévisions pour différents horizons c’est-à-dire la distribution du vecteur $(P_{t+h|t})_h$, h parcourant ici l’ensemble des horizons temporels considérés. Le processus sous-jacent peut aussi être spatio-temporel dans le cas d’un ensemble de centrales ou de sites mesurés. Une manière de parvenir

à la description du processus sous-jacent est évidemment d'en proposer une modélisation paramétrique sous forme de séries temporelles ou de vecteurs aléatoires, par exemple en utilisant un modèle gaussien. Dans le domaine de l'énergie, ce type de description est plutôt rare et l'on préfère en général décrire le processus en donnant un grand nombre de réalisations appelées scénarios. Les scénarios sont donc le résultat de la modélisation et doivent être considérés comme des réalisations indépendantes du processus stochastique correspondant. Nous les noterons ici $(P_{t+h|t}^\omega)_{h,\omega}$, ω parcourant un ensemble Ω de scénarios. Les prédictions de ce genre sont plus complexes et plus difficiles à appréhender que les prédictions des lois marginales faites par les méthodes de quantiles. Elles sont cependant utiles dans les applications qui nous concernent. Par exemple, la prédiction de la structure de dépendance temporelle est importante pour l'optimisation de l'utilisation d'un stockage ou son dimensionnement. En effet, si les erreurs de prédiction de la production pour différents horizons successifs sont corrélées, la capacité d'un stockage qui serait utilisé pour gérer celles-ci est beaucoup plus sollicitée que dans le cas où elles ne sont pas corrélées et où le stockage se vide et se remplit à une fréquence élevée. Le même type de besoins s'applique à la gestion du réseau ou plus généralement du système électrique pour la corrélation spatiale entre les erreurs.

Nous avons proposé une méthode spécifique pour la prédiction de scénarios de consommation de véhicules électriques (nous parlons de ces travaux dans la Sous-section suivante) dans [Confs46, ArticleP22].

Concernant la prédiction à court terme de scénarios, nous avons proposé des méthodes d'évaluation des scénarios qui ne soient pas simplement des évaluations des distributions marginales (pour h fixé) dans [ArticleP5]. Deux approches y sont proposées et nous illustrons par des exemples simples leurs avantages et leurs inconvénients. La première approche proposée est une généralisation du *energy score* qui est traditionnellement définie par :

$$ES_t = \frac{1}{|\Omega|} \sum_{\omega} \|P_{t+h|t} - P_{t+h|t}^\omega\|_2 - \frac{1}{2|\Omega|^2} \sum_{\omega} \sum_{\omega'} \|P_{t+h|t}^\omega - P_{t+h|t}^{\omega'}\|_2.$$

La généralisation proposée repose sur l'utilisation de normes de type Sobolev (norme L_p sur les variations du processus) au lieu de normes L_2 . Nous montrons empiriquement que dans le cas où les erreurs forment un processus gaussien avec une fonction de covariance du type $\exp((h_1 - h_2)^2/\gamma^2)$ l'évaluation proposée est en fait une manière indirecte de vérifier que la valeur de γ utilisée pour la génération des scénarios reflète bien la régularité des

observations.

La seconde approche proposée, dite approche par « événement », est en lien avec nos travaux sur les événements extrêmes présentés à la Sous-section 2.1.5. Il s’agit de définir un événement (par exemple une variation importante et soudaine de la production éolienne), de prédire sa probabilité d’occurrence pour chaque horizon h à partir de la fréquence d’apparition de cet événement sur les différents scénarios. L’évaluation des scénarios prédits se ramène alors à l’évaluation d’une prédiction probabiliste d’événement pour laquelle il existe déjà un grand nombre d’indicateurs. Nous proposons ici d’utiliser le score de Brier, qui est classique en météorologie. Ce type d’approche d’évaluation a l’avantage d’être adapté à certaines spécificités des processus non stationnaires. Dans [ArticleP5], nous montrons l’intérêt d’utiliser cette méthode d’évaluation plutôt que celle reposant sur le *energy score* en faisant varier la définition de l’événement à l’origine du critère. Nous comparons ainsi les performances de prédiction de scénarios de production éolienne, obtenues avec un modèle de copules gaussien et un modèle utilisant des ensembles météorologiques de ECMWF. Nous montrons que le choix de l’événement peut changer considérablement les résultats, mais que d’une manière générale, les ensembles météorologiques apportent une information supplémentaire qui n’est pas comprise dans le modèle de copule gaussien.

2.1.3 Prédiction de consommation locale

Un nombre important de mesures de la consommation électrique est désormais disponible grâce à une nouvelle génération de compteurs, comme le linky en France, ou à des *box* vendues par des entreprises privées pour assurer le suivi de la consommation. Dans le même temps, un certain nombre d’évolutions rendent pertinente une gestion de l’énergie à une échelle plus fine, c’est le cas notamment de nouveaux cadres réglementaires comme celui de l’auto-consommation collective ou du développement de productions décentralisées. Dans ce contexte, les différents acteurs des réseaux électriques cherchent à raffiner leur modèle de prévision à court terme de la consommation électrique : des modèles de prévision à l’échelle d’un quartier plutôt qu’à l’échelle d’une région. La résolution spatiale et temporelle de ces nouveaux *smart meters* installés au niveau des bâtiments est plus fine que celle à laquelle nous étions confrontés par le passé et le traitement de ce volume important de données requière de nouvelles méthodes de prédiction à court terme. Par ailleurs, la qualité et la disponibilité de ces données sont loin d’être parfaites et les méthodes développées doivent pouvoir intégrer un

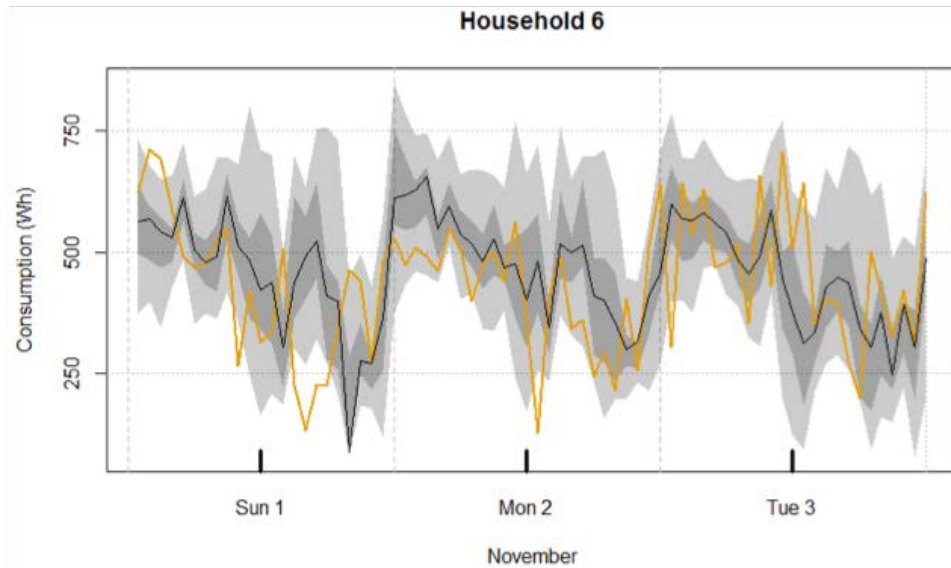


FIGURE 4 – Exemple de prédiction un jour à l’avance de la consommation d’un foyer. La ligne jaune est la consommation finalement observée, la ligne noire la prédiction du quantile à 50 %, l’intervalle en gris sombre correspondant à l’intervalle entre les quantiles à 30 et 70 % et l’intervalle en gris clair entre les quantiles) 10 et 90 %.

fonctionnement dégradé adapté à ces contraintes opérationnelles.

Dans [Confs40], nous avons développé une méthode de prédiction de quantiles reposant sur une décomposition additive de la dépendance des quantiles futurs à trois variables explicatives : (i) la consommation observée à la même heure que celle pour laquelle la prédiction est faite, mais le jour précédent, (ii) pour cette même heure, la médiane des valeurs observées pendant toute la semaine précédente et (iii) une prédiction météorologique de la température. Un enjeu auquel nous avons aussi répondu ici était de pouvoir adapter le modèle à l’absence temporaire de certaines variables explicatives afin de faire face à une plus grande variété de situations opérationnelles. Nous avons évalué ce modèle à partir de différents critères et montré l’amélioration par rapport à un modèle de persistance pour les critères déterministes. Dans [Confs39], nous avons montré l’intérêt d’utiliser cette méthode dans un contexte de gestion prédictive. Un exemple de prédiction est donné par la Figure 4.

Dans [ArticleP20], nous nous sommes appuyés sur un cas de démonstra-

tion dans la ville de Evora, en partenariat avec le gestionnaire du réseau de distribution espagnol EDP Distribution, pour énumérer les cas de figure apparaissant dans la pratique. Ce cas d'étude nous a amenés à généraliser l'approche de [Confs40] en proposant un plus grand nombre de modèles et une articulation optimisée entre leurs utilisations. Le déploiement de l'algorithme dans des conditions opérationnelles (dans le cadre du projet Sensible) est également discuté dans [ArticleP20]. Nous avons montré qu'il était intéressant de combiner plusieurs modèles adaptés à des situations différentes vis-à-vis de la disponibilité des données. Nous avons aussi observé que certaines situations opérationnelles peuvent rendre difficile l'évaluation des résultats si les données manquantes sont trop nombreuses.

Dans [ArticleP22], nous avons proposé une méthode de génération de scénarios de consommation de véhicules électriques (VE) dans un contexte de prévision court terme. Nous observons ici une série temporelle de consommation de 46 bornes de recharge individuelles de véhicules électriques. Les données sont issues du projet Américain Pecan Street et sont en accès libre. Pour chaque borne, nous observons au pas de temps d'une minute la consommation agrégée de tous les véhicules en charge. Cette méthode ne repose pas à proprement parler sur le schéma général présenté à la Sous-section 2.1.1, mais sur une modélisation de la charge d'un VE à partir de quelques caractéristiques élémentaires : le début de charge, la puissance de charge et la durée de charge. Les charges sont supposées être ici de simples fonctions créneau. La modélisation proposée est appliquée dans un contexte de prédiction à court terme grâce à plusieurs étapes. (i) La première étape consiste en une détection de blocks de charge par une méthode de filtrage et d'estimation des caractéristiques correspondantes. (ii) La deuxième étape est une étape d'apprentissage statistique de la distribution bi-variée du couple (durée, heure de début de charge) à partir d'une méthode à noyau, couplée à une méthode de clustering des VE qui permet de définir quatre grandes catégories de charge de VE. Ces deux étapes sont suffisantes pour la simulation de scénarios long terme de consommation de VE. (iii) La troisième étape vise à la simulation de ces caractéristiques par rapport à des informations connues en un jour à l'avance.

Cette méthode est utilisée pour prédire des scénarios de la consommation agrégée des 42 véhicules un jour à l'avance. Étant donné qu'elle part d'un modèle véhicule par véhicule, il s'agit ici d'une méthode de type *bottom up*. Un exemple de prédiction est donné dans la Figure 5. Cette méthode de prédiction à court terme de scénario est comparée à une méthode de persistance et une méthode statistique avancée (Gradient Tree Boosting Model, GBT)

Mesure d'erreur	Persistance	GTB	<i>Bottom up</i>
MAE (kW)	6.24	4.86	4.87
CRPS (kW)	3.63	3.59	

TABLE 1 – Comparaison de la méthode de simulation *bottom up* par rapport à la persistance et à la méthode GTB. Les critères utilisés sont la Mean Absolute Error (MAE, critère évaluant la qualité de la prédiction de la médiane) et le Continuous Ranked Probability Score (CRPS, un score évaluant l'ensemble de la distribution prédite).

décrite dans l'article, les résultats de la comparaison sont donnés dans la Table ??.

2.1.4 Modèles statistiques spatio-temporels

Contexte et problématique. Une solution pour permettre d'améliorer la prédictibilité pour des horizons de production allant de quelques minutes à quelques heures est la mise à jour rapide de prévisions de production, de l'échelle nationale à une échelle plus fine. En effet, ces prévisions sont aujourd'hui mises à jour toutes les heures à l'échelle d'un pays, d'une région ou d'une centrale. Comme nous l'avons vu dans la Sous-section 2.1.1, les modèles existants reposent principalement sur la combinaison, via un modèle statistique, de deux types de données. D'un côté les mesures de production des centrales obtenues à partir des systèmes SCADA et actualisées tous les quarts d'heure, de l'autre les prévisions météorologiques de vent ou d'irradiation obtenues à partir de modèles numériques de prévision météorologiques (NWP). Ces prévisions météorologiques sont mises à jour en général toutes les 6 heures, parfois toutes les 3 heures. Aujourd'hui il existe des produits proposant des mises à jour horaires. La Figure 6 illustre l'importance relative de ces deux types de données en fonction de l'horizon de prédiction dans le cas de la production éolienne. On peut aussi y voir qu'à l'échelle d'un pays ou d'une région, les données de production éolienne reflètent la situation météorologique avec une granularité spatiale et temporelle supérieure à celle que l'on peut appréhender à partir de modèles météorologiques et plus adaptée aux échelles temporelles des variations de la production éolienne. Le même constat peut-être fait à propos de la production photovoltaïque.

Ce travail sur la modélisation spatio-temporelle a débuté dans le cadre du projet européen Safewind [Proj9] pour le cas de l'éolien et s'est poursuivi dans le projet ANR WinPower [Proj2]. Il a donné lieu à une collaboration

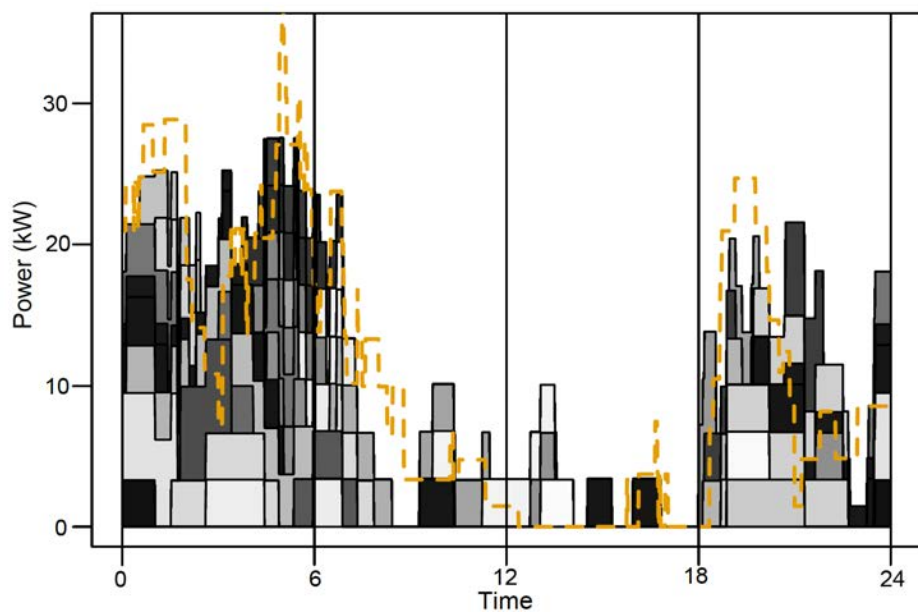


FIGURE 5 – Exemple de prédiction *bottom up* un jour à l'avance de la consommation résultante de la charge des 42 véhicules électriques. La ligne orange est la consommation finalement observée.

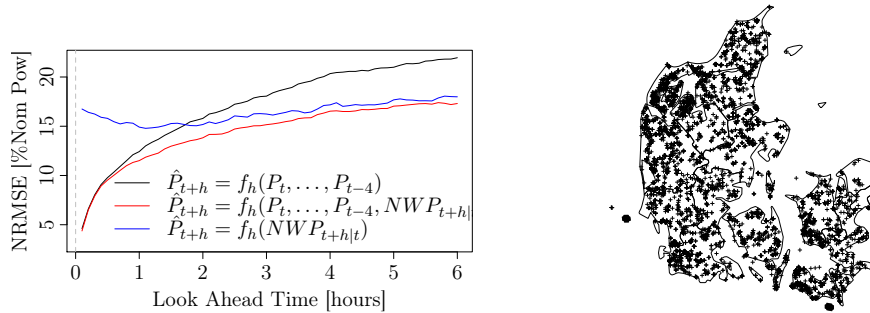


FIGURE 6 – À gauche horizons temporels pour lesquels on peut espérer un bénéfice à l’utilisation de modèles spatio-temporels. Sont représentées : les erreurs de prévision de la production sur une ferme française, obtenues à partir de trois modèles de prévisions (non spatio-temporels) utilisant des prévisions météorologiques (NWP) ou pas. Les mises à jour de ces prévisions se font toutes les six heures (en même temps que l’arrivée des NWPs). À droite est représenté potentiel d’amélioration en termes de réseau de capteurs : plus de 6000 turbines au Danemark, aujourd’hui plus de 50 % des fermes correspondantes fournissent au gestionnaire Energinet des informations de production mises à jour toutes les quinze minutes.

avec Denis Allard de l’INRA sur l’analyse des structures spatio-temporelles des erreurs de prédiction dont je décris le contenu dans le paragraphe suivant. La majeure partie de mon travail sur la modélisation spatio-temporelle s’est concentré plus directement sur la prédiction à court terme de la production éolienne puis surtout de la production photovoltaïque. Les deux derniers paragraphes de cette Sous-section sont dédiés à ces travaux qui se sont effectués essentiellement dans le cadre de deux thèses : celle de Ghislain Agoua [These11] soutenue en 2017 et celle de Kevin Bellinger [These6] qui est commencée depuis 2018.

Description de déplacements de structure. J’ai proposé deux indicateurs statistiques spatio-temporels pour mesurer des déplacements de structure dans [ArticleP7] et j’ai montré l’intérêt de ces indicateurs pour quantifier les déplacements de structures spatio-temporelles dans les erreurs de prédiction à court terme de la production éolienne au Danemark. Les processus étudiés ici sont des erreurs de prédiction qui peuvent être considérées comme des processus aléatoires

$$\epsilon_{t+h|t}^x = P_{t+h} - \hat{P}_{t+h|t}$$

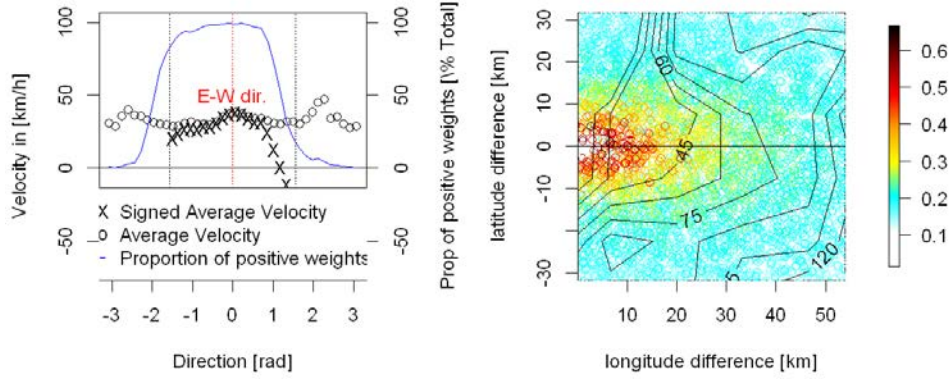


FIGURE 7 – Vitesse et direction de propagation des erreurs de prédiction de production éolienne au Danemark selon la première approche pour définir ces vitesses. Cette approche repose sur une étude des maxima locaux de fonctions spatio-temporelles de corrélations. À gauche vitesse moyenne de propagation estimée $V(\theta)$ (voir Équation (5) dans [ArticleP7] pour la définition). À droite : chaque point représente une paire de centrale $s(x_1, x_2)$, a pour coordonnées $x_1 - x_2$, la couleur donne le maximum de corrélation $\rho_{x_1, x_2}(h)$ et les lignes de niveau approchent le temps $h^*(p_1, p_2)$ pour lequel ce maximum est atteint.

indexés par une localisation x , un instant de prédiction t et un horizon de prédiction h .

Le premier indicateur statistique repose sur la définition d’un temps de propagation entre un site x_1 et un site x_2 via une analyse des maxima locaux de la fonction

$$\rho_{x_1, x_2} : h \rightarrow \text{corr}(\epsilon_{\cdot, +15|}^{x_1}, \epsilon_{\cdot, +15+h|}^{x_2}).$$

Les temps et directions de propagation obtenus à partir de cette définition sont représentés sur la Figure 7.

Avec cette définition, et pour les données de production éolienne utilisées au Danemark, on obtient une propagation ouest-est avec une vitesse de 23.66 km/h . Dans le cas où la relation entre deux sites distants n’est pas bien décrite par une relation linéaire, l’utilisation peut ne pas être suffisante. Dans ce cas nous avons proposé une autre quantification de l’interdépendance qui repose sur une décomposition de la variance de prédiction d’un

site en fonction d'un autre.

Le second indicateur statistique proposé repose sur l'utilisation d'un modèle spatio-temporel de propagation d'onde planeaire pour approcher à chaque instant t le processus spatio-temporel indexé en h et $x : \epsilon_{t+h|t}^x$. Ce modèle d'onde planeaire contient une composante paramétrique donnant une vitesse de propagation. À chaque instant t on peut donc estimer et définir une vitesse \hat{v}_t de propagation. Nous sommes partis de l'hypothèse d'une propagation ouest-est restée sur un modèle simple. On peut là aussi calculer une vitesse moyenne de propagation : $1/T \sum_{t=1}^T \hat{v}_t$ qui est ici de l'ordre de 30 km/h . Nous avons également proposé une étude statistique de la relation entre la vitesse du vent et la vitesse de propagation (advection). Cette relation n'est pas une simple relation linéaire. En d'autres termes on identifie bien une propagation à une vitesse donnée et plus la vitesse du vent est élevée, plus la vitesse de propagation est grande, mais la relation entre la vitesse de propagation et la vitesse du vent n'a pas une explication simple et directe.

Modèles de prédiction spatio-temporelle et procédure d'estimation. Une partie de ma recherche s'est portée sur des **modèles statistiques spatio-temporels de prédiction à court terme** de la production éolienne, ou PV. La première innovation que nous avons apportée avec ces modèles était d'utiliser les centrales de production réparties sur le territoire comme un réseau de capteurs. Puis, dans le cas de la production PV, nous avons rajouté des données satellites et des prédictions issues de modèles météorologiques. L'enjeu n'a pas tant été ici l'innovation statistique (les modèles utilisés sont des adaptations mineures techniques de modélisation statistiques qui existent depuis quelques années déjà) que de montrer l'intérêt de ces grandes quantités de données pour améliorer la prédiction à court terme. Les modèles développés sont des modèles linéaires de la forme

$$\hat{P}_{t+h|t}^x = \beta_{h,x}^0(Z_{t,x}) + \sum_{l=0}^L \sum_{y \in \mathcal{X}} \beta_{h,x}^{l,y}(Z_{t,x}) P_{t-l}^y + g(\text{Satellite}, \text{NWP}), \quad (1)$$

où P_t^x est la production d'un site x donné parmi l'ensemble des sites considérés \mathcal{X} à l'instant t , $\hat{P}_{t+h|t}^x$ est la prédiction effectuée à l'instant t pour l'horizon h , $l = 0, \dots, L$ sont des décalages temporels, $\beta_{h,x} = (\beta_{h,x}^0, (\beta_{h,x}^{l,y})_{l,y})$ est le paramètre du modèle pour l'horizon h éventuellement fonction d'une ou plusieurs covariables $Z_{t,x}$ disponible à l'instant t pour la localisation x . La partie du modèle $g(\text{Satellite}, \text{NWP})$ intégrant les images satellites et les

prédictions météorologiques n'a été développée que pour le PV. Elle est discutée à la fin de ce paragraphe.

Plusieurs méthodes d'estimation du paramètre β_h sont possibles. Si β_h ne dépend pas d'une covariable et tant que le nombre d'observations de l'échantillon d'apprentissage est suffisamment grand devant le nombre de sites une méthode des moindres carrés est suffisante. Lorsque le nombre de paramètres devient trop grand devant le nombre d'observations, nous avons choisi de restreindre la complexité du modèle en pénalisant la norme l_1 du paramètre β_h (modèle de type LASSO). Dans le cas où $\beta_{h,x}$ dépend d'une covariable $Z_{t,x}$, nous estimons $\beta_{h,x}(z)$ en nous ramenant au cas précédent par une pondération plus importante des dates de l'échantillon d'apprentissage pour lesquelles $Z_{t,x} - z$ est suffisamment petit. Nous utilisons des poids de lissage exponentiels et un paramètre de lissage γ qui est choisi par validation croisée. Cette procédure d'estimation locale réduit le nombre d'observations effectivement utilisées pour l'estimation de $\beta_{h,x}(z)$ et, dans ce cas, l'utilisation d'un modèle de type LASSO est préférable. J'ai contribué à la mise en œuvre de ce type de méthode pour l'éolien dans [ArticleP7, Confs10, Confs8, Confs19] et pour le photovoltaïque dans le cadre de la thèse de Ghislain Agoua [These11] et de celle de Kevin Bellinger [These6] qui ont donné lieu à des publications dans un cadre déterministe, [ArticleP18, Confs30] mais aussi dans un cadre probabiliste [ArticleP17]. Je vais maintenant présenter certains résultats obtenus dans ces travaux.

Évaluation de l'apport de ces modèles. Nous avons montré que les modèles présentés dans le paragraphe précédent permettent une amélioration substantielle de la prédiction à court terme dans le cas de la production éolienne et dans le cas de la production PV. Les améliorations obtenues dans le cas du PV sont représentées dans la Figure 8. Des résultats similaires ont été obtenus dans le cas de l'éolien et sont présentés dans les articles et rapport mentionnés dans le paragraphe précédent.

2.1.5 Prédiction de rampes de production

Les situations de fortes et brusques variations de production éolienne, que l'on appelle **les rampes**, ont été identifiées par un certain nombre d'opérateurs comme des situations délicates. Nous avons travaillé à la caractérisation de ces situations et à la prédiction de l'incertitude temporelle relative à celles-ci principalement dans le cadre de la thèse d'Arthur Bossavy [These2].

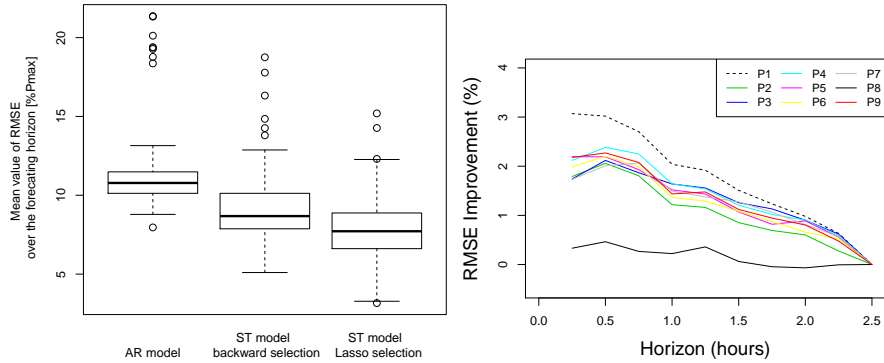


FIGURE 8 – **Amélioration de la prédiction photovoltaïque (PV) par des modèles spatio-temporels.** À gauche : boxplot des erreurs avec et sans modèle spatio-temporel et intérêt de l'utilisation d'une procédure de type LASSO. À droite amélioration supplémentaire obtenue en utilisant une covariable météorologique dans le modèle spatio-temporel (ici la vitesse du vent). Chaque ligne représente une centrale PV dans le centre de la France. Le pas de temps utilisé est de 15 minutes et les périodes considérées pour le test et pour l'apprentissage sont de 1 an. Issue de [ArticleP18].

Caractérisation mathématique de ces phénomènes. Dans [Confs15, ArticleP10], nous avons proposé plusieurs plusieurs caractérisations possibles de ces événements en suivant une démarche empruntée au domaine du traitement du signal [ext1]. Afin de pouvoir comparer les différentes approches nous avons proposé un modèle de rampe qui est décrit en détail dans [ArticleP10] et qui peut être synthétisé par la Figure 9.

Les approches utilisées pour la caractérisation de ces événements sont des approches reposant sur un filtrage et un seuillage à différentes échelles. La phase de filtrage à différentes échelles temporelles permet de mesurer différents niveaux et différents types de variations. Nous avons comparé plusieurs filtres, ainsi que différentes méthodes pour combiner les signaux filtrés aux différentes échelles. Les résultats sont mesurés par la capacité d'une méthode à prédire une variation et, lorsque la variation est bien prédite, nous regardons aussi la capacité de la méthode de prédiction à estimer correctement l'instant d'apparition de la rampe. Nous avons analysé la sensibilité des résultats aux différents paramètres du modèle. Notre travail a permis d'écarter complètement la définition de rampe proposée dans [ext6] et utilisée dans la littérature en montrant sa trop grande sensibilité au bruit. Nous avons montré que la combinaison des informations comprises à toutes les échelles est

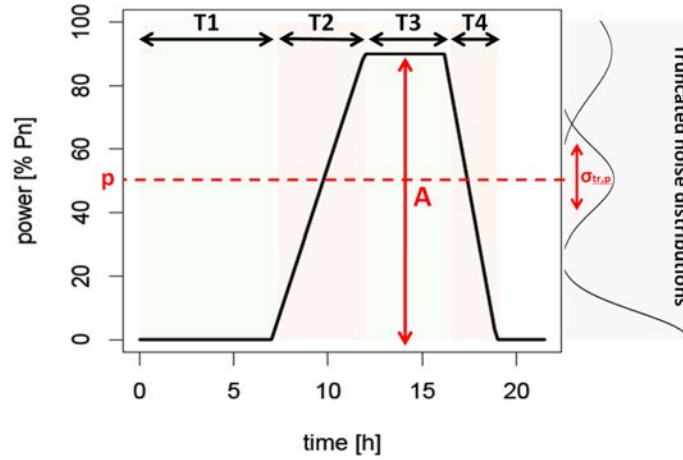


FIGURE 9 – **Modèle de rampe** proposé dans [ArticleP10] pour étudier les avantages et inconvénients des différentes approches de caractérisation des rampes. C’est un modèle pour un processus comportant 4 régimes différents

efficace même si les approches correspondantes ne sont pas toujours optimales. Ce type de démarche était proposée dans la littérature du traitement du signal [ext5] et de l’éolien [ext12]. Nous avons montré que dans un cas plus stationnaire (un cas où la fréquence d’apparition des rampes reste régulière), une sélection d’échelles telle que celle proposée dans [ext3] permet alors d’améliorer les performances de détection.

Prédiction de l’incertitude temporelle autour de ces phénomènes.

Nous avons proposé dans une série de conférences [Confs1, Confs6, Confs4, Confs3] puis finalement dans l’article [ArticleP6] une méthode pour prédire l’incertitude temporelle autour de l’avènement des rampes. Cette méthode repose sur l’utilisation des ensembles météorologiques de ECMWF qui sont des scénarios générés à partir de perturbations des conditions initiales d’un modèle météorologique. Si l’on évalue la proportion de ces prédictions d’ensemble (i.e. la proportion de ces scénarios) qui prédisent une rampe dans un intervalle de temps donné, on ne peut malheureusement pas espérer que celle-ci puisse être directement utilisée comme une probabilité. En effet les ensembles météorologiques ne sont pas nécessairement bien calibrés, pas pour cette utilisation en tout état de cause. C’est pourquoi nous avons proposé une méthode de calibration reposant sur une régression logistique, permettant de faire le lien entre la probabilité d’avoir effectivement une rampe et le nombre de scénarios dans lesquels la rampe est détectée.

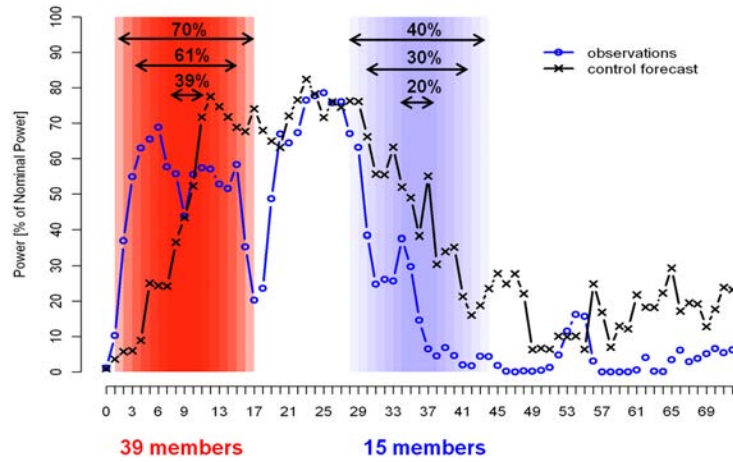


FIGURE 10 – **Prédiction de l'incertitude temporelle** autour d'une rampe, résultat du modèle de prédiction proposé dans [ArticleP6]. Ici le modèle de prédiction utilisé utilise en entrée les 50 prédictions météorologiques fournies par les « prédictions d'ensemble » du Centre européen de prédiction (ECMWF).

Cette méthode n'est bien évidemment pas conditionnée à l'utilisation des ensembles de ECMWF et il est possible de l'utiliser avec des ensembles purement statistiques (souvent appelés les « poor man ensemble » voir e.g. [ext7]).

2.2 Modélisation et simulation du système électrique et du marché de l'électricité pour la gestion et la planification

Les évolutions à venir du système électrique (évolution de la demande et des usages de l'électricité, évolution des moyens de production) requièrent des méthodes de modélisation, de simulation et d'optimisation adaptées vis-à-vis des enjeux de planification et de gestion. Ces modèles doivent permettre de tenir compte des fluctuations de la production à l'échelle nationale et continentale, mais aussi des effets de ces nouvelles productions connectés dans le réseau de distribution. La diversité grandissante des échelles spatiales et temporelles impliquées et l'articulation nécessaire de celles-ci impliquent des difficultés nouvelles. Par ailleurs, dans un contexte de libéralisation les éléments de ce système interagissent par le biais de règles de marché dont il est important de comprendre et analyser le fonctionnement. Dans ce contexte, la multiplicité des acteurs et la complexité de leurs interactions s'ajoutent à celle du système électrique.

L'optimisation mathématique est au cœur des modèles correspondants. Un enjeu important vis-à-vis de son utilisation est de trouver un cadre de modélisation adapté à la réalité que l'on cherche à modéliser, et offrant des méthodes de résolution efficaces. Les modélisations correspondantes doivent donc répondre à des problèmes d'optimisation complexes intégrant un grand nombre de variables à un grand nombre d'échelles. Cette complexité impose souvent des non-linéarités et des non-convexités, mais dans le même temps la dimensionnalité des problèmes et la singularité de leurs organisations rendent nécessaire la mise en œuvre de modèles exploitant des structures géométriques simples.

J'ai débuté mon activité de recherche dans cette vaste thématique en 2011 dans le cadre du projet européen SafeWind [Proj9] par une modélisation des **marchés de l'électricité**. Il s'agissait alors de comprendre l'impact des erreurs de prédiction à court terme sur le revenu et les choix d'investissements de producteurs éoliens. Ce travail s'est effectué dans le cadre d'un stage [Stages7] puis de la thèse de Fiona Foucault [These4].

Partant d'un questionnaire sur le rôle du stockage à l'**échelle d'un quartier** dans un échange avec le Centre Scientifique et Technique du Bâtiment (CSTB), je me suis également tourné 2011 vers des questions de planification du réseau de distribution dans le cadre de la thèse de Yassine Abdelouadoud [These1] (thèse financée par le CSTB). J'ai pu poursuivre ces

travaux dans le cadre du projet européen Sensible [Proj12] qui a notamment permis de financer la thèse de Etta Grover-Silva [These3] sur ce sujet et un stage [Stages16]. Un projet Carnot, le projet SmartCity [Proj10] est venu compléter ces financements en 2015. Ces travaux ont été l’occasion de plusieurs projets d’étudiants proposés dans le cadre de l’enseignement MIG du cycle IC de MINES ParisTech [Ens1].

Dans le cadre de la thèse [These1] et de différents sujets du MIG [Ens1] autour de 2013, j’ai approché le thème de l’impact de la variabilité des renouvelables à l’échelle du **système centralisé** sur les problèmes de gestion et de dimensionnement. Cela a notamment motivé ma coordination de la réponse à un appel à projets de l’ADEME « vers un mix 100 % renouvelable ». Nous avons finalement remporté et contribué à mener cette étude coordonnée par Artelys en partenariat avec Énergie Demain pour le compte de l’ADEME[Proj5]. En parallèle, j’ai poursuivi cette activité dans le cadre du projet [Proj2] et plus récemment dans le cadre d’une thèse CIFRE avec RTE [These9]. Sur ce sujet, je participe ponctuellement à la tâche 25 de l’Agence Internationale de l’Énergie (AIE) [Ray6]. Cette thématique a beau n’être pas très développée dans mes publications elle n’en est pas moins un sujet qui a nourri beaucoup de mes enseignements et en particulier une série de cours sur le fonctionnement et la planification du système électrique [Ens7, Ens9, Ens11]. C’est aussi une thématique que j’ai retrouvé assez souvent dans mes tutorats d’élèves de master [Stages1, Stages2] ou d’élèves de l’option Machine Énergie du cycle IC [Stages3].

Ces travaux sur la simulation et l’optimisation du système électrique se sont progressivement rapprochés de besoins industriels sur l’évaluation des revenus du stockage et de la production renouvelable sur le marché de l’électricité. Cela a relancé récemment mes travaux sur la modélisation des marchés de l’électricité. L’objectif y est de modéliser la dynamique des prix sur le marché de l’électricité. Les travaux sur le stockage ont été menés dans le cadre projet Berrien [Proj11] et ceux en lien avec des producteurs ont été mené dans le cadre d’un projet avec le producteur Coruscant [Proj13]. Ce sujet est maintenant traité dans la thèse de Valentin Mahler [These10]. Ici encore différents sujets du MIG [Ens1] m’ont poussé à avancer sur ce thème, souvent en amont de la création de projets financés. Le MIG sur le stockage en 2013 a par exemple joué un rôle moteur dans la création du projet Berrien [Proj11]. C’est ici également un sujet récurrent ces dernières années dans les tutorats de stage.

Nous commençons la présentation de nos travaux les plus récents sur la modélisation des besoins de flexibilité dans le système électrique centralisé

pour l'équilibre offre-demande. Puis, dans la Sous-section 2.2.2 nous présentons nos travaux relatifs à l'optimisation dans le réseau de distribution. Enfin, la partie concernant le marché de l'électricité (valeur de la prédictibilité et modélisation des prix) est présentée dans la Sous-section 2.2.3.

Il est à noter que l'utilisation des outils que nous avons développés et que nous décrivons dans cette Section est conditionnée à une modélisation de la consommation et de la production future à différentes échelles d'espace et de temps. Ces sujets sont traités dans la Section 2.3.

2.2.1 Optimisation de la flexibilité pour l'équilibre offre-demande du système centralisé

Les problèmes d'optimisation qui permettent de faire la planification ou l'opération du système électrique sont de plus en plus difficiles à résoudre du fait de la grande dimensionnalité nouvelle qu'implique un traitement satisfaisant des variations du PV et de l'éolien à plusieurs échelles d'espace et de temps. C'est notamment l'optimisation et le dimensionnement des flexibilités nécessaires aux différentes échelles qui posent problème. Mais le stockage en particulier impose aux problèmes d'optimisation des contraintes difficiles à gérer.

Définition et détermination des besoins de Flexibilité dans le système. L'utilisation du terme de flexibilité dans l'étude du système électrique n'est pas tout à fait nouvelle, mais elle peut avoir plusieurs sens selon le contexte et son utilisation pour désigner une adaptabilité à la variabilité et à la faible prédictibilité de ses charges sur le réseau est assez récente (dans les publications scientifiques, son utilisation remonte à 2010). Dans le cadre de la Thèse de Thomas Heggarty [These9] en partenariat avec RTE nous travaillons autour des différentes définitions de la Flexibilité et sur son estimation. La flexibilité comprise comme capacité du système à s'adapter aux variations concerne donc une variété de problèmes (dont une partie a été mentionnée dans l'introduction de ce document Figure 2) selon l'échéance qui sont rassemblés Figure 11.

Dans [ArticleP23], nous avons proposé une méthode statistique pour estimer le besoin de Flexibilité. Cette méthode repose sur une décomposition sur trois échelles temporelles -jour, semaine, année- des signaux de consommation et de production renouvelable variable (PV et éolien), ou directement de demande nette. Un besoin de flexibilité en puissance ainsi qu'un besoin

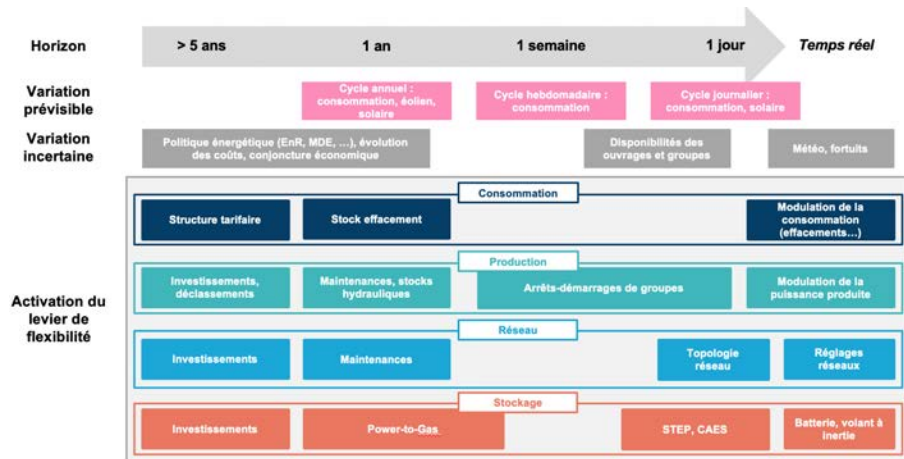


FIGURE 11 – **Besoins et leviers de flexibilité** dans le système électrique à différentes échelles d'espace et de temps.

en énergie sont ensuite calculés pour chaque échelle de temps. Un exemple de représentation est donné Figure 12.

Nous proposons ensuite une analyse de sensibilité de l'évolution de ce besoin à l'échelle européenne en fonction de l'évolution de plusieurs paramètres : (i) la pénétration du PV, (ii) la pénétration de l'éolien, (iii) la pénétration du chauffage électrique, (iv) la pénétration de la climatisation, et (v) le niveau d'inter-connection en Europe. L'analyse de sensibilité est menée à partir d'une technique statistique que j'ai souvent utilisée dans mes différents travaux : les indices de Sobol. Je reviens sur ce thème dans la Section 2.3.

Cette méthode a également été appliquée aux séries de production réalisées des moyens pilotables du système français dans le but de voir quels moyens répondaient à quels besoins. Cette démarche permet de dépasser certains clichés qui nous font ranger par exemple le nucléaire dans les moyens non pilotables alors que le nucléaire (via son cycle de maintenance) permet de répondre à un besoin majeur de flexibilité à l'échelle annuelle imposée par la consommation de chauffage. C'est ce que l'on peut observer sur la deuxième partie de la Figure 12. On peut questionner ce besoin dans la mesure où le chauffage électrique s'est développé en France en réponse à une surcapacité nucléaire. Cependant, le développement du chauffage électrique répond aujourd'hui à une exigence de décarbonation du système énergétique.

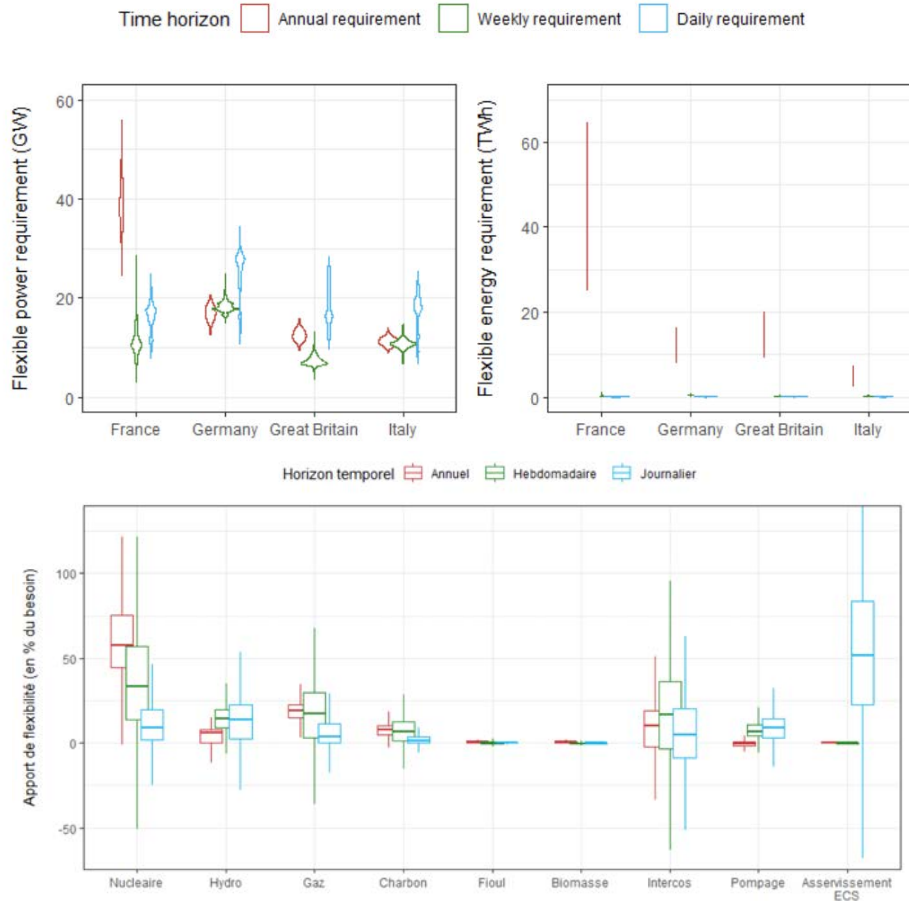


FIGURE 12 – **Besoins de flexibilité** (en haut) en puissance et en énergie estimés à différentes échelles de temps sur la demande (sans tenir compte de la production renouvelable variable) pour différents pays d'Europe en 2017. Les besoins de flexibilité sont ici représentés sous forme de distribution statistique.

Couverture du besoin en énergie (en bas) en France en 2015 estimé par la même méthode que celle utilisée pour définir le besoin, mais appliquée aux séries temporelles de production observées.

Les perspectives de ce travail sont dans l'estimation d'un lien entre ces besoins de flexibilité estimés et la manière dont ils sont couverts lors de l'optimisation de capacité du système. Ce type de lien doit permettre à terme à la fois d'accélérer la résolution du problème d'optimisation de capacité (avec contraintes d'opération) dans un système fortement renouvelable, et de rendre plus compréhensive la description des besoins de flexibilité.

Dans le cadre du cours d'une semaine que j'ai monté sur l'évolution du système électrique dans le contexte de la transition énergétique [Ens11], j'ai mis en place une plateforme pédagogique qui interface d'une manière intuitive le logiciel R avec le logiciel d'optimisation AMPL afin que les étudiants puissent faire l'expérience de poser et résoudre simplement des problèmes d'opération et d'investissement dans le système électrique Européen et y comprendre certains enjeux de la transition énergétique vis-à-vis du système électrique.

Optimisation rapide pour la gestion d'un stockage. Une source de flexibilité importante dans le système électrique repose sur le stockage électrique et la gestion de la demande. Sur le plan de l'optimisation de la gestion du système, les contraintes correspondantes sont a minima des contraintes en puissance (les plus souvent deux contraintes : une pour le stockage et une pour le déstockage) et des contraintes en énergie (ici aussi deux contraintes : imposer au stockage de n'être jamais moins que vide ou plus que plein). Ces dernières contraintes imposent un couplage temporel qui peut être très profond si le stockage est grand et qui rend la résolution du problème de gestion du système difficile. Ce couplage est encore plus problématique lorsqu'il s'agit de problèmes d'investissement (avec contraintes de gestion) et les algorithmes correspondants intègrent le plus souvent le stockage de manière séparée.

Nous avons travaillé dans deux directions autour de ce problème. Dans la thèse de Yassine Abdelouadoud [These1] (chapitre 2), nous avons tout d'abord proposé de décomposer le problème à l'échelle nationale sur une année au pas de temps horaire en 52 problèmes hebdomadaires au pas de temps horaire couplé à un problème maître couvrant l'année au pas de temps hebdomadaire. Ainsi les stockages journaliers et les cycles journaliers des Stations de Transfert d'Énergie par Pompage (STEP) sont gérés par les 52 problèmes alors que les stockages saisonniers et la maintenance des centrales nucléaires sont gérés par le problème annuel. Dans [These1] un troisième couplage est proposé avec les réseaux de distribution dont nous parlons dans la Sous-section 2.2.2. La Figure 13 illustre de manière schématique le fon-

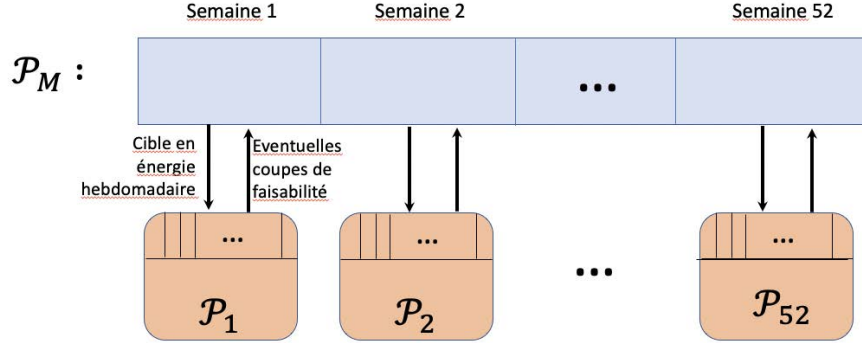


FIGURE 13 – **Décomposition du problème de gestion** à deux échelles de temps. Le problème maître à l'échelle annuelle envoie des objectifs en termes de production d'énergie par moyen de production et les 52 problèmes de second niveau peuvent être résolus en parallèle et font éventuellement remonter des coupes imposantes par exemple le bridage d'une production éolienne si ses variations au pas de temps horaire sont trop grandes.

tionnement de cet algorithme.

En parallèle, sur la même question, nous avons proposé un nouvel outil de **programmation dynamique** [Confs29, Confs37] qui permet de résoudre très rapidement un problème de gestion du stockage (seul interagissant avec le système via un marché) formulé de la manière suivante :

$$\begin{aligned}
 \min \quad & \sum_{i=1}^n C_i(x_i) \\
 \text{s.t.} \quad & \begin{cases} lb_i \leq x_i \leq ub_i & i = 1, \dots, n \\ lbC_i \leq \sum_{j=1}^i x_j \leq ubC_i & i = 1, \dots, n \end{cases}
 \end{aligned} \tag{2}$$

où lbC, ubC, lb et ub sont des vecteurs de \mathbb{R}^n , et C_i ($i = 1 \dots, n$) sont des fonctions de coût. Dans le cas où ces fonctions de coûts sont convexes et linéaires par morceau, ou quadratiques par morceau, nous avons implémenté un algorithme qui permet de résoudre ce problème en un temps quasi linéaire en fonction du nombre n de pas de temps. La manière d'utiliser la programmation dynamique ici n'est pas standard, elle est décrite dans [Confs29], elle permet d'obtenir une solution exacte et s'inspire des techniques de transformée de Legendre rapide. Nous avons illustré ce résultat par une comparaison empirique avec l'utilisation du solveur commercial CPLEX les résultats sont présentés Table 2. L'outil correspondant a été mis à disposition en open source dans un package du logiciel R [Ray9].

TABLE 2 – Comparaison empirique entre l’algorithme de programmation dynamique proposé dans [Confs29] et l’utilisation d’un solveur linéaire de l’état de l’art (CPLEX). Ici les fonctions C_i sont linéaires. Certaines valeurs n’ont pas pu être obtenues suite à des problèmes de mémoire elles sont notées « NC ».

n	100	1000	5000	10^4	10^5	10^6
dyn.prog. [ms]	0.34	3.06	28.8	61.5	649.2	6285
CPLEX [ms]	4.31	724	63579	NC	NC	NC

2.2.2 Optimal Power Flow et réseau de distribution

En ce qui concerne le réseau de distribution, mes travaux ont commencé un peu avant 2012 avec la thèse de Yassine Abdelouadoud [These1] via une recherche sur des problèmes d’optimisation : les *Optimal Power Flow* (OPF). Ceux-ci permettent de modéliser de nombreux cas de figure dans le réseau de distribution, tant sur le plan de la gestion que de la planification. La thèse de Etta Grover Silva [These3] est venue compléter cette activité de recherche en y ajoutant des éléments de modélisation permettant de tenir compte : (i) d’incertitudes (*stochastic Optimal Power Flow*), (ii) des particularités de la basse tension. Elle a aussi transformé le modèle d’opération en un modèle de planification pour le positionnement et le dimensionnement de stockage.

Formulation de base. Nous avons contribué au domaine de l’**optimisation d’un réseau de distribution actif** à partir d’un *Optimal Power Flow* (OPF) et d’un nouvel algorithme d’optimisation pour résoudre ce problème. Comme nous l’avons montré par un cas d’application dans [Confs13], l’OPF permet d’optimiser les puissances actives et réactives dans le réseau tout en tenant compte des contraintes de tension et d’intensité. Le problème d’optimisation correspondant est assez classique, on y minimise une fonction de coût impliquant l’intensité dans les lignes ou les puissances actives et réactives dans ces lignes sous des contraintes sur les intensités et les niveaux de tension. La relation entre la tension, les intensités dans les lignes et les puissances injectées/soutirées est donnée par un système d’équations dit de *load flow*. La fonction de coût peut impliquer des pertes, le coût de l’électricité achetée moins celui de l’électricité vendue, la quantité de PV coupé, une distance à un objectif donné par le système centralisé...

Pour expliquer notre apport sur ce sujet nous allons rappeler les équations correspondantes. Un point de départ pour obtenir les équations de *load flow*

dans un réseau radial peut être trouvé dans [ext10]. Les équations résultent de l'application des lois de Kirchoff en centrant le calcul sur le flux dans chaque branche. Le modèle correspondant est dit *branch flow* BFM, et donne la relation entre les puissances actives et réactives (notées P et Q) dans chaque branche, l'intensité dans chaque branche et les tensions en chaque nœud. Chaque branche est modélisée par une impédance $z = r + jx$. La puissance P entrant au nœud i dans la branche se dirigeant vers le nœud j est notée $P_{i,j}$, le même type de notation est utilisé pour Q , z , r , x et I . Le réseau étant supposé radial chaque nœud i a au plus un père qui est noté $f(i)$ et un ensemble de nœuds fils notés $s(i)$. Une fois ces notations fixées le BFM s'écrit de la manière suivante.

$$P_{f(j),j} - r_{f(j),j} I_{f(j),j}^2 - \sum_{k \in s(j)} P_{j,k} - P_j^{load} + P_j^{st} + P_j^{pv} = 0 \quad (3a)$$

$$Q_{f(j),j} - x_{f(j),j} I_{f(j),j}^2 - \sum_{k \in s(j)} Q_{j,k} - Q_j^{load} + Q_j^{st} + Q_j^{pv} = 0 \quad (3b)$$

$$U_{f(j)} - U_j - 2(r_{f(j),j} P_{f(j),j} + x_{f(j),j} Q_{f(j),j}) + z_{f(j),j}^2 I_{f(j),j}^2 = 0 \quad (3c)$$

avec :

$$I_{f(j),j}^2 = \frac{P_{f(j),j}^2 + Q_{f(j),j}^2}{U_{f(j)}^2} \quad (4a)$$

$$U_j = V_j^2 \quad (4b)$$

et les contraintes suivantes sur l'intensité et la tension

$$I_{f(j),j}^2 \leq (I_{f(j),j}^{max})^2 \quad (5a)$$

$$V_j^{min} \leq V_j \leq V_j^{max} \quad (5b)$$

$$V_1 \in V_{OLTC} \quad (5c)$$

la tension du nœud source peut être réglable dans un ensemble V_{OLTC} qui permet de modéliser les possibles changements de prise sur les transformateurs.

Les injections du stockage et de la production photovoltaïque peuvent

être contraintes par les puissances apparentes des onduleurs :

$$P_j^{st2} + Q_j^{st2} \leq S_j^{max,pv2} \quad (6a)$$

$$P_j^{pv2} + Q_j^{pv2} \leq S_j^{max,st2} \quad (6b)$$

$$P_j^{pv} \in [0, P_j^{pv,sp}] \quad (6c)$$

$$P_j^{st} \in \begin{cases} [P_j^{st,sp}, S_j^{max,st}] & \text{if } P_j^{st,sp} < 0 \\ [-S_j^{max,st}, P_j^{st,sp}] & \text{if } P_j^{st,sp} > 0 \end{cases} \quad (6d)$$

Dans ces équations, la non-convexité est concentrée dans l'Équation (4a) qui est une contrainte d'égalité sur une forme quadratique. Plusieurs travaux (voir par exemple [ext10]) ont suggéré une relaxation de cette contrainte sous la forme de l'inégalité (7) qui permet alors d'avoir un ensemble de contraintes s'écrivant sous la forme d'un cône de second ordre. Cette formulation donne accès à une gamme d'outils efficaces dédiés à l'optimisation « semi-définie positive » et reposant sur des techniques de type points intérieurs.

$$I_{f(j),j}^2 U_{f(j)} \geq P_{f(j),j}^2 + Q_{f(j),j}^2 \quad (7)$$

Ajout de coupes. Depuis quelques années, un certain nombre de travaux se sont penchés sur ce type de relaxation pour résoudre des OPF dans le réseau de distribution. Tous ces travaux ont établi un ensemble de conditions suffisantes, relatives à la fonction de coût où à l'ensemble des injections, qui permettent d'assurer l'exactitude de la relaxation, nous pouvons citer par exemple [ext10, ext8, ext13, ext9, ext11, ext14].

Dans le cadre de la thèse de Yassine Abdelouadoud [These1] et dans l'article [ArticleP11] nous avons montré que la relaxation définie par l'Équation 7 n'était pas exacte pour des cas où l'on cherche à minimiser une fonction de coût intégrant la minimisation des pertes et ajoutée à une autre fonction qui tend à lui être opposée. Cela peut être le cas par exemple lorsque la production renouvelable est assez importante pour que sa coupure volontaire puisse contribuer à la diminution des pertes en ligne mais que cela corresponde à une perte économique. Les chances d'une telle configuration sont faibles aujourd'hui mais nous avons montré que dans un contexte de forte pénétration renouvelable elle est très probable. Il était donc nécessaire d'apporter une solution à l'inexactitude de cette relaxation et ce que nous avons fait dans [Confs24, ArticleP11] en proposant l'ajout itératif de coupes permettant de rapprocher les solutions du problème relaxé pour lesquelles la contrainte 7 n'est pas une égalité d'une solution au problème initial non relaxé.

Dans [ArticleP11], nous montrons mathématiquement que les coupes que nous avons proposées permettent à chaque étape de réduire le domaine à explorer et sans écarter la solution du problème non relaxé. En d’autres termes, la coupe ajoutée à une étape de l’algorithme permet de forcer la nouvelle solution du problème relaxé à se rapprocher de la contrainte d’égalité 4a. Nous illustrons dans [ArticleP11] la rapidité de la convergence de l’algorithme de coupes proposé.

Extensions du problème. Les Équations précédentes ne sont données que pour un seul pas de temps. Or, dans le cas où l’on utilise un stockage ou pour modéliser la gestion de la demande, il est nécessaire d’introduire des contraintes de stock qui impliquent un couplage entre les différents pas de temps qui peut être introduit par l’ajout de contraintes linéaires (égalités et inégalités). Le principe de l’algorithme et ses propriétés mathématiques s’étendent à ce cas de figure **multitemporel**. Nous avons montré l’intérêt de cette extension dans le cadre de la thèse de Etta Grover-Silva [These3] qui a fait suite à la thèse de Yassine Abdelouadoud et nous avons publié ce travail dans [Confs38, Confs41].

La thèse de Etta Grover-Silva nous a permis d’explorer d’autres extensions. Dans [ArticleP19], nous proposons une extension au cas **stochastique** (multi-temporel et multi-scénarios) du problème de gestion auquel nous avons également proposé l’ajout du pilotage de la demande de chauffages sous des **contraintes de confort thermique**. Des prévisions de scénarios sont générées un jour à l’avance pour la demande non contrôlable, pour la consommation d’eau chaude, pour la production PV et pour la température. L’algorithme de résolution est ici sensiblement le même et l’on peut penser qu’une exploitation de la structure par block du problème pourrait permettre de mettre en place un algorithme plus rapide utilisant du calcul parallèle et une décomposition du problème.

Nous avons proposé une adaptation de l’OPF au dimensionnement et aux positionnements de stockages distribués dans le réseau de distribution dans [These3, ArticleP14, Confs38, Confs41]. Sur le plan de la modélisation, il ne s’agit ici que d’ajouter le coût d’investissement du stockage dans la fonction de coût et d’intégrer la capacité (en nombre d’heures) et la puissance (en MW) dans les variables à optimiser. La résolution du problème d’optimisation est-elle plus difficile pour au moins deux raisons? La première est que le problème de planification impose de travailler avec une période d’évaluation

suffisamment longue et représentative pour refléter la diversité des besoins à la fois vis-à-vis de la profondeur des contraintes (qui implique un besoin de capacité), de leur intensité (qui implique un besoin de puissance) et de leur localisation. La deuxième raison est qu'à cette forte dimensionnalité s'ajoute une non-convexité imposée par la forme du coût du stockage qui est proportionnel à l'énergie de stockage installée et donc au produit entre la capacité (le nombre d'heures maximum de charge) et la puissance. Nous avons proposé une technique de résolution qui repose sur une alternance entre l'évaluation des besoins indépendamment de leur localisation et une évaluation de la localisation des besoins.

L'importance du déséquilibre entre les phases dans le réseau de basse tension a également été étudiée dans la thèse de Etta Grover-Silva [These3] (chapitre 2). Nous avons montré qu'il était suffisant d'effectuer un OPF en deux temps. Dans le premier temps, nous proposons une résolution avec un modèle de phases équilibrées (celui discuté dans les paragraphes précédents) mais en prenant une marge sur les contraintes de tension et d'intensité. Ceci constitue à la fois une relaxation (le fait de négliger le déséquilibre entre les phases) et un renforcement des contraintes (marge prise sur la tension et l'intensité). L'enjeu est de vérifier que le renforcement des contraintes permet de compenser l'effet de la relaxation. Dans le second temps, nous appliquons à la solution de l'OPF un simple *load flow* avec des équations intégrant une modélisation plus sérieuse du déséquilibre pour vérifier que la marge choisie à la première étape est suffisante. Une adaptation des marges est finalement possible, mais rarement nécessaire.

Couplage avec un système centralisé : décomposition, coordination et réduction de dimension. Dans les travaux présentés sur le réseau de distribution reposant sur les OPF, la fonction objectif externalise les objectifs et les contraintes du système centralisé auquel le réseau de distribution est connecté, par exemple en utilisant des prix de marché. Or, dans le cas où un grand nombre d'actifs du système sont connectés au réseau de distribution d'électricité, ils participent également à l'équilibre du système centralisé et tirent leurs revenus de celui-ci. Aussi, il n'est pas possible de les inclure dans la gestion du système centralisé sans considérer les contraintes du réseau de distribution et il n'est pas non plus possible de considérer qu'ils n'influencent pas la formation des prix sur le système centralisé.

Dans le cadre de la thèse de Yassine Abdelouadoud [These1] nous avons proposé un algorithme de gestion coordonné du système centralisé et d'un

ensemble de réseaux de distribution actifs. C'est l'objet du deuxième chapitre de cette thèse qui n'a pour l'instant été publié dans aucun journal. Dans l'algorithme proposé, le problème d'optimisation centralisé, problème primaire, est résolu sans tenir compte des contraintes du réseau de distribution. La solution est ensuite intégrée dans la fonction objectif de l'algorithme de simulation du réseau de distribution. On cherche ainsi à minimiser la distance à la solution obtenue par le système centralisé tout en tenant compte des contraintes du réseau de distribution. Si ce dernier problème n'est pas faisable ou si la distance obtenue n'est pas nulle, alors le problème de gestion du réseau de distribution retourne des contraintes supplémentaires sur les puissances actives commandables. Nous illustrons le fonctionnement de cet algorithme en simulant le système français. Pour ce système la décomposition temporelle du problème de gestion du système centralisé que nous avons décrite au début de cette Section est utilisée. Les contraintes de congestion dans le réseau de transport ne sont pas modélisées. Un ensemble de 3 réseaux typiques représentant les réseaux urbains, semi-urbains et ruraux est proposé pour modéliser le réseau de distribution français.

Sur la Figure 14, on peut voir l'évolution au fur et à mesure de ces itérations, de la puissance PV qui doit être ajustée à la baisse lors de la réalisation de l'OPF dans l'ensemble des réseaux de distribution.

L'évaluation du caractère contraignant des actifs contrôlables du réseau de distribution doit être faite sur tous les pas de temps de l'année. Ces 8760 OPF sont une charge de calcul qu'il est utile de chercher à réduire. Nous avons proposé dans le chapitre 4 de la thèse de Yassine Abdelouadoud [These1] [Confs14] et dans [Confs14] un critère de criticité rapide à calculer reposant sur une linéarisation des équations de power flow. Il permet une estimation rapide du niveau de criticité de chaque pas de temps vis-à-vis des contraintes de tension et d'intensité. L'algorithme proposé repose sur un tri des 8760 niveaux de criticité observés par ordre croissant pour un réseau de distribution donné. Des optimal power flow sont ensuite appliqués pas de temps par pas de temps jusqu'à ce que plusieurs pas de temps de suite se révèlent être non critiques (i.e. sans contraintes actives de tension et d'intensité). Dans la mesure où cette méthode permet d'éviter de vérifier complètement les contraintes de tension et d'intensité par un optimal power flow complet, elle constitue une relaxation dont nous discutons dans [Confs14] l'inexactitude.

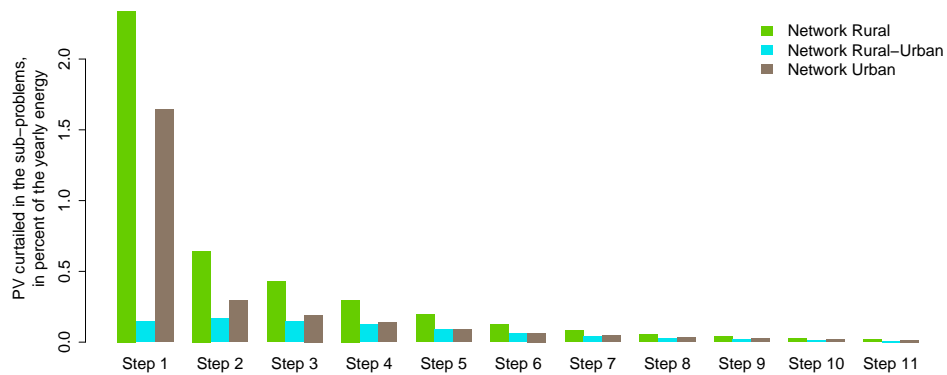


FIGURE 14 – **Convergence vers zéro** des différences entre les objectifs de production PV donnés par le système centralisé et la solution faisable dans l'ensemble des réseaux de distribution. Le résultat est représenté au fur et à mesure des itérations dans l'algorithme de couplage entre le système centralisé et le réseau de distribution. En quelques étapes, la modification des contraintes des actifs envoyée par les systèmes décentralisés au système centralisé permet une intégration satisfaisante des contraintes du réseau de distribution.

2.2.3 Modélisation et simulation du marché de l'électricité

Mon activité de recherche autour du marché de l'électricité a commencé en 2011 dans un lien étroit avec celle sur la prédiction à court terme. Dans le cadre du projet SafeWind [Proj9], j'ai encadré le stage de Kevin Laquaine [Stages7] qui a donné lieu à une publication [ArticleP8] sur la valeur de la prédictibilité dans les décisions d'investissement à partir d'analyses statistiques d'historiques de marchés. Ce travail s'est poursuivi par la thèse de Fiona Foucault [These4] dans le cadre du projet ANR WinPower [Proj2] qui a donné lieu aux publications [Confs18, Confs26, Confs27, Confs29, ArticleS1] et qui repose sur une méthodologie de modélisation du marché à partir de problèmes d'optimisation multi-niveau. Ces travaux sont présentés dans le deuxième paragraphe ci-dessous.

Suite à ce travail et face à l'importance grandissante pour diverses applications de la modélisation de la dynamique des prix futurs, j'ai tourné ma recherche vers la modélisation des prix de l'électricité à partir d'approches combinant de l'apprentissage statistique avec les principes de base de la physique du système électrique et du fonctionnement du marché de l'électricité. Ces travaux sont menés aujourd'hui dans le cadre de la thèse de Valentin Mahler [These10], ils sont récents et ont donné lieu à la publication [Confs49]. Ils sont décrits dans le paragraphe ci-dessous.

Simulation de la dynamique des prix dans un mix électrique futur.

Une des conséquences de la libéralisation du système électrique est que la gestion des actifs du système n'est plus entreprise par un acteur centralisé, mais par un grand nombre d'acteurs qui communiquent via des plateformes d'échange. Leurs revenus sont donc déterminés par l'évolution des prix dans ces plateformes d'échange et la compréhension de l'évolution de ces revenus, l'optimisation si possible des stratégies d'investissement vis-à-vis de ces revenus sont des problèmes importants.

Plusieurs évolutions aujourd'hui renforcent l'importance de ce problème. Tout d'abord là **sortie des tarifs d'achat réglementés** pour l'éolien et le photovoltaïque, via un mécanisme de complément de revenu, introduit de nouvelles sources d'incertitude sur les revenus du producteur. Les agrégateurs qui vendent leur production, ainsi que les producteurs eux-mêmes qui signent des contrats avec les agrégateurs doivent comprendre ces incertitudes et si possible maîtriser les risques correspondants. Dans un futur un peu moins proche, si l'augmentation de la part dans le mix électrique de l'électricité renouvelable se poursuit on peut penser que le **stockage de l'électricité**

deviendra primordial pour la maîtrise des coûts du système. Dans ce cas on peut imaginer qu'il tirera une partie de sa rentabilité en arbitrant des différences de prix sur le marché de l'électricité : achetant l'électricité à bas prix pour la revendre lors des pointes. Dans ces deux cas de figure, le stockage et la production renouvelable, les revenus sous-jacents ne dépendent pas seulement de la valeur moyenne du prix de l'électricité mais aussi et surtout de sa dynamique.

Les évolutions réglementaires, et plus généralement l'évolution du mix électrique lui-même, déterminent cette dynamique des prix. En ce qui concerne les marchés, il faut noter plusieurs changements. Tout d'abord, les volumes qui y transitent augmentent assez nettement en France et dans tous les pays du monde. Ensuite les règles de fonctionnement évoluent avec l'apparition par exemple sur le marché EPEX-spot de la possibilité de vendre des blocs temporels, ou encore dans certains pays la prise en compte des contraintes de congestion via des marchés à prix locaux ou nodaux. Enfin de nouveaux marchés, comme le marché de capacité en France, changent les possibilités de revenu qu'un producteur peut espérer obtenir en participant sur les marchés.

Aujourd'hui l'effet des renouvelables sur la dynamique des prix est déjà observable en France, mais aussi à plus forte raison dans d'autres pays, comme l'Allemagne ou l'Espagne. Par exemple, on constate bien souvent que la forte intégration de production photovoltaïque ou éolienne dans le mix énergétique abaisse le niveau des prix (car ce sont des productions de coûts marginaux faibles), ce qui impacte les décisions d'investissement. C'est une des raisons de la mise en place du marché de capacité. L'apparition de prix négatifs, bien que reflétant les contraintes temporelles de gestion des moyens conventionnels, est aussi liée à la priorité des productions éoliennes et PV sur le marché. Dans le même temps, un certain nombre de décisions concernant le système électrique sont susceptibles d'affecter durablement les prix de marché. C'est le cas de la programmation pluriannuelle de l'énergie (la PPE) qui indique entre autres le développement futur des EnR et du nucléaire en France. Mais il existe d'autres facteurs importants comme la durée des carénages de centrales nucléaires qui sont prévus pour les années à venir, la date de mise en service de Flamanville, l'évolution des volumes de l'ARENH (aujourd'hui EDF est obligé de vendre 100 *TWh* par an d'électricité nucléaire à un prix fixé par l'état), l'évolution des interconnexions, l'évolution des mixes électriques voisins, l'évolution de la sensibilité de la consommation à la température qui est liée à l'efficacité énergétique.

Dans le contexte actuel de transition énergétique vers des moyens de production d'électricité renouvelable et du fait de la libéralisation du marché de l'électricité français, il est nécessaire de mieux modéliser le comportement

des marchés spot de l'électricité et la manière dont ces derniers peuvent réagir à des changements du mix énergétique. Dans le même temps, un grand nombre de données sont disponibles pour mieux comprendre et modéliser les prix de l'électricité dans plusieurs types de systèmes. En effet, on dispose aujourd'hui de plusieurs années de prix de marché dans différents pays avec des caractéristiques variées et dans certains marchés on peut même obtenir les courbes d'offre. Aujourd'hui ces données sont soit utilisées directement pour établir un *business model* sur la base de données passées sans tenir compte de l'effet de changement de mix électrique ou du couplage entre les marchés, soit le mix électrique et la formation des prix sur le marché sont modélisés avec des outils classiques de simulation du système électrique centralisé et alors la grande quantité de données observées n'est pas utilisée.

J'ai commencé à travailler sur ce sujet dans le cadre de travaux personnels, puis cela est devenu un sujet récurrent dans les MIG [Ens1] (sur les micro STEP en 2013, le power 2 gaz en 2015, et la sortie des tarifs d'achat en 2017) et plus récemment j'ai pu financer un stage [Stages20] par les projets Berrien [Proj11] (pour évaluer la valeur futur des micro-STEP) et Coruscant [Proj13] (pour évaluer les conséquences pour un producteur renouvelable de la sortie des tarifs d'achat réglementés). L'étude bibliographique réalisée dans le stage [Stages20] a révélé que les outils mathématiques existants se basent principalement soit sur (i) une modélisation purement statistique, soit sur (ii) une modélisation explicite du système électrique ne tirant pas partie des données disponibles. L'objectif scientifique de ces travaux en cours et à venir est de proposer un modèle permettant une synergie entre les deux voies existantes dans la littérature intégrant (i) un modèle statistique permettant de tirer parti des grandes bases de données d'historique de fonctionnement des systèmes électriques et des marchés européens et (ii) des paramètres physiques décrivant le réseau et les contraintes de fonctionnement sous-jacentes.

Aujourd'hui ces travaux se poursuivent dans le cadre de la thèse de Valentin Mahler [These10]. Nous avons proposé dans ces travaux (ceux du stage et dans la première année de thèse) un modèle pour simuler la dynamique des prix de l'électricité sur le marché spot qui repose à la fois sur une modélisation explicite du système électrique et du marché de l'électricité et sur une modélisation statistique des offres de prix faites par les producteurs. La combinaison de ces deux techniques de modélisation permet d'avoir un modèle qui peut en même temps être utilisé dans un cadre prospectif et s'ajuster à une réalité plus complexe pour laquelle un grand nombre d'observations sont disponibles. Des premiers résultats prometteurs ont été présentés à la conférence [Confs49].

Valeur de la prédictibilité dans la décision d’investissement. La participation sur le marché de l’électricité des producteurs renouvelables implique une nouvelle constitution du revenu. Nous avons déjà évoqué le rôle important que jouent les prix sur le marché *day ahead* ainsi que la dynamique de ces prix et sa relation avec celle de la production. Les erreurs de prédiction faites sur ces marchés, lors de la déclaration des offres la veille pour le lendemain, impliquent aussi des coûts supplémentaires. Ainsi le revenu du producteur peut être décomposé en un revenu lié à sa production et une perte de revenu liée à ses erreurs de prédiction. Nous avons étudié l’importance dans la pratique de ces deux facteurs dans le cadre du projet SafeWind [Proj9] et du stage de Kevin Laquaine [Stages7] qu’il a permis de financer.

Les travaux de ce stage ont donné lieu à une publication [ArticleP8] dans laquelle nous proposons une analyse quantitative des contributions de la prédictibilité et du facteur de charge dans les revenus de producteurs éoliens participant au marché de l’électricité. Cette analyse repose sur l’utilisation d’historiques de données de marché de NordPool et d’un jeu de données fourni par Energinet dans le cadre du projet européen SafeWind et se composant de plus de 200 fermes éoliennes observées sur presque 4 ans au pas de temps de 15 minutes. Ces dernières données étaient très originales à l’époque et ont aussi servi au travail sur la modélisation spatio-temporelle déjà évoquée. Les résultats obtenus ici sont synthétisés dans la Figure 15.

Pour ce cas d’étude spécifique, nous avons montré que le revenu annuel Rev_x (en [euros/MW/an]) d’un producteur éolien se trouvant en x est très bien approché par une fonction simple impliquant son facteur de charge CF_x et sa prédictibilité $NMAE_x$ (mesurée par la Normalized Mean Absolute Error NMAE) par la relation :

$$Rev_x = 3960CF_x - 448.52*(NMAE_x - 9.753) + 121*(NMAE_x - 9.753)^2 - 2173 + \epsilon_x \quad (8)$$

où ϵ_x est un terme résiduel dont nous avons montré empiriquement qu’il ne contribuait pas à plus de 0.2 % de la variance de Rev_x . Dans cette équation, nous avons également montré que la contribution $r^2(CF, Rev)$ de CF_x à la variance de Rev_x est de plus de 99.7 %. Avec ces ordres de grandeur et étant donné la corrélation positive qu’il existe entre $NMAE_x$ et CF_x il n’est définitivement pas possible de trouver un cas pour lequel il est aujourd’hui intéressant de considérer la prédictibilité dans la phase d’investissement d’un producteur éolien.

Dans le cadre de la thèse de Fiona Foucault [These4], nous avons cherché à poursuivre ces travaux de deux manières. D’une part à partir d’une approche

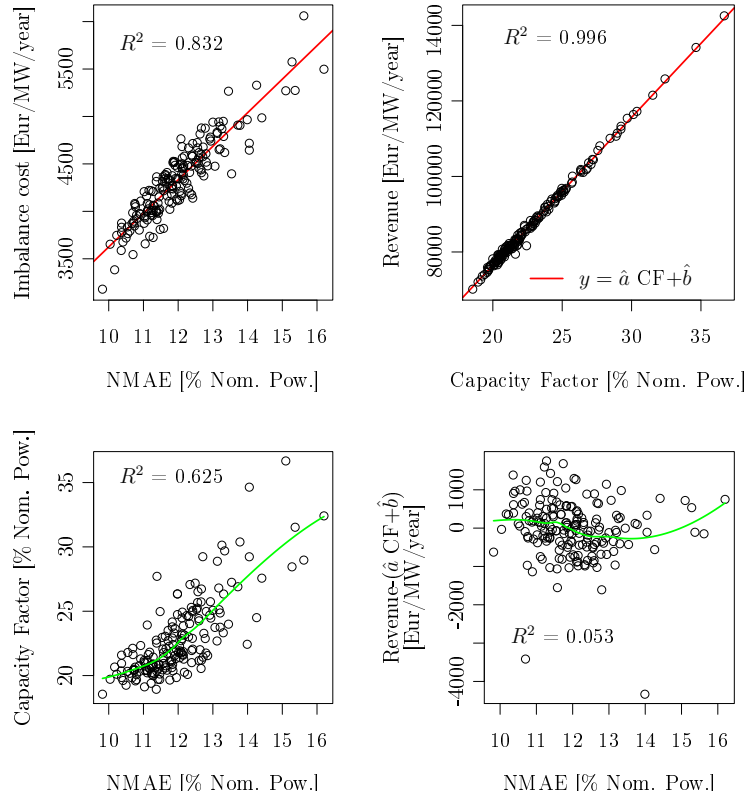


FIGURE 15 – **Valeur de la prédictibilité au Danemark** dans la phase d’investissement d’un producteur éolien. Ces 4 graphiques illustrent les résultats obtenus lors d’une analyse des données historiques de prix et de production éolienne au Danemark sur une période de 2 ans (2008 – 2010). Chaque point représente une centrale éolienne pour laquelle des erreurs de prédiction ont été simulées, de la production a été observée et des revenus ont été calculés à partir de données de marché historiques.

statistique basée sur des historiques et la mise en place de modèles simplifiés, et d'autre part à partir d'une approche de simulation d'un marché à prix locaux et à deux niveaux de marchés (*day-ahead* et *real time*). Nous allons décrire la première approche, avant d'évoquer rapidement la seconde.

Bien que l'article [ArticleP8] discute la généralité de ces résultats par quelques heuristiques, on peut penser que ceux-ci sont propres au cas de NordPool. Dans un article de conférence, [Confs18] nous avons étendu ces résultats en analysant d'autres marchés européens et le marché américain PJM. Nous avons montré que dans ces marchés la part de variance expliquée $r^2(Rev, NMAE)$ par la prédictibilité dans la relation statistique entre le revenu, le facteur de charge et la prédictibilité (relation du type de celle décrite par l'Équation 8) était plus petite de 0.1 %.

Dans un article soumis [ArticleS1] qui correspond à un des chapitres de la thèse de Fiona Foucault, nous avons cherché à établir quels étaient les déterminants de cette part de variance expliquée. Nous avons déterminé formellement une relation approchée liant le revenu à des déterminants liés à la production (prédictibilité et facteur de charge) et des déterminants liés au marché (prix moyen sur le marché *day ahead* et prix moyen de règlement des écarts). Cette relation s'applique dans un cadre où le revenu du producteur est donné par une expression de la forme

$$Revenue = \overline{\lambda_{DA}}CF - \widehat{\Delta\lambda} \times NMAE + \epsilon. \quad (9)$$

où $\overline{\lambda_{DA}}$ est le prix moyen sur le marché *day ahead*, $\widehat{\Delta\lambda}$ est un terme lié au prix moyen de règlement des écarts et ϵ un terme suffisamment petit pour être négligé. Il faut noter que cette relation n'est pas vraie en général surtout si la corrélation entre les prix et la production est trop importante. Pour autant, dans le cas des marchés étudiés cette relation est vérifiée. Nous montrons en particulier que c'est le cas aujourd'hui de la plupart des marchés avec un *dual pricing* et nous étudions aussi la formulation alternative dans le cas du *single pricing*. Dans ce cadre nous définissons

$$\alpha = \frac{cov(CF, NMAE)}{\sigma_{CF}\sigma_{NMAE}}, \quad Y = \frac{\sigma_{NMAE}}{\sigma_{CF}}, \quad X = \frac{\widehat{\Delta\lambda}}{\lambda} \quad (10)$$

où $cov(x, y)$ est la covariance entre le vecteur x et le vecteur y et σ_x l'écart type de x . À partir de cela, nous avons montré que de manière approchée $r^2(Rev, NMAE)$ est donnée par l'expression

$$r^2 = 1 - \frac{(1 - \alpha^2)X^2Y^2}{1 + X^2Y^2 - 2XY\alpha} \quad (11)$$

Cette formule permet en outre de séparer l'effet des caractéristiques X du marché de celles Y et α du producteur. Nous montrons dans le cas du Danemark que la corrélation α entre le facteur de charge et la prédictibilité est de 0.77 et que Y est de l'ordre de 0.35. Elle permet surtout un calcul rapide de $r^2(Rev, NMAE)$ dans le cadre d'une analyse de marché préliminaire, elle donne un indicateur de surveillance de marché. Étant donné que le dénominateur de l'équation 11 est petit nous avons finalement, pour $\alpha = 0.77$ et $Y = 0.35$:

$$1 - r^2(Rev, NMAE) \approx (1 - \alpha^2)X^2Y^2 \approx 0.05X^2.$$

La seconde approche sur laquelle nous avons travaillé dans le cadre de la thèse de Fiona Foucault a impliqué la modélisation couplée de trois éléments : (i) le problème d'investissement du producteur (voulant maximiser ses bénéfices) (ii) un modèle de marché *day ahead* nodal permettant de simuler les prix et l'influence du choix des investisseurs sur ceux-ci (iii) un modèle de marché temps réel permettant de simuler l'impact des erreurs de prédiction sur le revenu en tenant compte de l'impact du choix de l'investisseur sur ceux-ci. Le problème d'optimisation correspondant est un problème stochastique à 3 niveaux et nous avons proposé une méthode pour le résoudre qui repose sur l'utilisation de coupes de Benders. Nous avons exploité cet outil de simulation pour analyser l'impact futur des erreurs de prédiction dans la décision d'investissement et la conclusion encore ici était que l'impact reste faible. Cette méthode ainsi que son utilisation pour analyser la valeur de la prédictibilité dans l'investissement sont décrites dans la thèse de Fiona Foucault et dans l'article de conférence [Confs27] mais n'ont à ce jour pas donné lieu à une publication dans un journal.

2.3 Énergie et territoire, modélisation prospective à partir de grandes bases de données

La compréhension et la modélisation des enjeux de la transition énergétique requièrent une connaissance et une modélisation fine des territoires, de la diversité des systèmes énergétiques qui y sont déployés, et de la variété des conditions dans lesquelles ils sont exploités.

Les caractéristiques importantes de ces systèmes énergétiques sont par définition celles qui ont un impact significatif sur leur faisabilité, leurs coûts, ou leurs impacts environnementaux. L'enjeu vis-à-vis de la démarche prospective que nous adoptons ici est à la fois de les identifier et de modéliser leurs impacts en tenant compte de la complexité des phénomènes sous-jacents et de la quantité de données disponibles. Il faut intégrer une complexité assez

grande pour ne pas manquer des interactions ou des effets importants et en même temps assez faible pour ne pas dépendre d'une information trop peu fiable. Par exemple, nous avons vu que la variabilité de la production éolienne, en plus de son facteur de charge, est nécessaire à la quantification du coût du système électrique. Aussi, sa compréhension à partir de déterminants météorologiques et techniques est-elle essentielle à toute analyse, tant sur le plan économique qu'environnemental. Dans la modélisation de cette variabilité à l'échelle d'un territoire, comme une région de France, on ne peut négliger l'évolution du ratio entre la taille des pales et la puissance de la machine qui impacte la courbe de puissance et le facteur de charge, mais l'on aurait tort de vouloir imposer à un modèle la connaissance passée et future du profil aérodynamique des pales de toutes les machines du territoire étudié.

Dans tous les domaines discutés dans cette section se retrouve un enjeu d'équilibre entre biais et variance dans la modélisation et l'on peut pêcher autant par un manque d'ambition vis-à-vis des données existantes que par une surparamétrisation au mieux inutile. Dans le même temps, les modélisations que nous proposons doivent intégrer à la fois une composante physique afin de pouvoir évaluer l'impact d'évolutions technologiques et une composante statistique qui permettent de tirer parti des grandes quantités de données existantes pour gommer les simplifications opérées par la modélisation, et donner un crédit plus important à notre analyse de la complexité. En outre, la qualité d'un modèle ou de l'analyse d'un système complexe ne vient jamais selon nous de sa seule complexité, mais plutôt de notre capacité à en montrer l'intérêt. Ainsi notre démarche a toujours été de partir de modèles aussi simples que possible pour obtenir des ordres de grandeur, puis de raffiner ces modèles en démontrant progressivement la valeur ajoutée du raffinement (ajout de nouvelles données ou ajout d'une complexité nouvelle dans la modélisation). Cette combinaison méthodologique entre l'analyse physique des systèmes énergétiques complexes et l'analyse de grandes bases de données est un élément important de ma démarche. Il résulte en partie de l'expertise que j'ai acquise avant d'arriver à l'école de mines dans le domaine de l'analyse de données en grande dimension et de l'informatique. On peut parler d'économétrie de l'énergie, mais ce domaine de recherche nous allons le voir est aussi par certains aspects liés à la géographie.

Les travaux développés dans cette Section concernent différents types de problèmes autour des EnR, du stockage, sur le dimensionnement du réseau électrique ou encore sur l'efficacité énergétique. La liste des projets correspondants permettra d'en structurer la présentation.

Dans le cadre du projet MOSAIC [Proj1], j'ai mis en place un partenariat depuis 2013 avec le gestionnaire de réseau ENEDIS pour la modélisation de la consommation et de la production à échelle d'un quartier en tirant parti entre autres de la grande quantité de données que possède ENEDIS sur ses clients. Ces travaux sont décrits dans la Sous-section 2.3.2.

Dans le cadre du **projet Smart Reserve** [Proj7] à l'échelle de la France, puis dans la cadre du projet REstable [Proj16] (j'ai coordonné le WP 1 de ce projet dédié aux analyses prospectives) à l'échelle de l'Europe, j'ai développé la mise en place d'outils pour la simulation de la production EnR et des erreurs de prévision correspondantes. Ce travail à l'échelle de l'Europe a été effectué en grande partie par une étudiante postdoc Evgeniya Ishkina. Ces travaux (et une partie des travaux du projet suivant) sont présentés dans la Sous-section 2.3.3.

Dans le cadre du **projet « vers un mix 100 % renouvelable »** [Proj5] j'ai coordonné l'étude du gisement national l'estimation du potentiel français de production renouvelable par une cartographie du territoire, ce travail a été effectué en grande partie par l'étudiant en postdoc Arthur Bossavy. Ces travaux et ceux des deux projets suivants sont présentés dans la Sous-section 2.3.4.

Dans le cadre du **projet Shape** [Proj6] avec l'ADEME et en partenariat avec le CSTB, nous mettons en place des méthodes pour accompagner les collectivités dans leurs processus de décision pour la transition énergétique. Ce projet est récent, mais une thèse est en cours [These7].

Dans le cadre du **projet Berrien** [Proj11], j'ai permis à Antoine Rogeau (actuellement en thèse pour le projet Shape) d'effectuer un stage pour estimer le potentiel géolocalisé du stockage micro-hydraulique en France à partir de systèmes d'information géographique (SIG).

2.3.1 Des bases de données pour l'analyse et l'optimisation de la transition énergétique dans les territoires

Un parcours rapide de bases de données existantes pour la modélisation de la transition énergétique. En France, de nombreux acteurs gèrent des bases de données, aux origines et applications multiples. On peut bien entendu faire une distinction entre les acteurs publics et privés, mais aussi entre les données qui sont en Open Data ou en accès privatif (propriétaire ou accessible mais payant). Ces acteurs font parfois partie du monde de l'énergie, auquel cas leurs données ont un intérêt certain pour la transition énergétique, mais d'autres données provenant d'instituts plus « généraux »

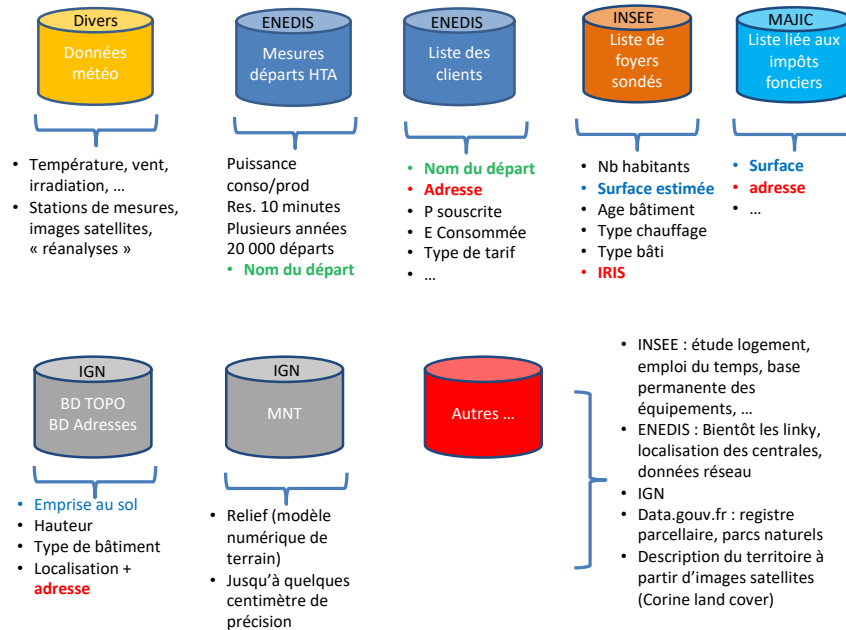
s'avèrent également précieuses comme les données de cadastre ou les données de recensement. Nous présentons une partie des bases de données utilisées dans la Figure 16.

L'Institut géographique national (IGN) est l'acteur cartographique national principal et fournit des services au grand public, avec quelques cartes disponibles en accès libre, mais surtout aux secteurs privés et académiques. Les données sont proposées dans des formats vectorisés et utilisables sur des moyens informatiques, et comprennent autant des cartes altimétriques, hydrographiques, la modélisation du parc bâti ou encore des routes, des réseaux d'électricité, etc.

L'institut National des Statistiques et des Études Économiques (INSEE) est l'organisme public en charge du recensement de la population française, et de toutes les études statistiques qui y sont liées. L'ensemble des statistiques et études sont disponibles en accès libre à une échelle spatiale fine, définie par un découpage géographique appelé maille IRIS (de l'ordre de 1000 – 2000 habitants). On peut mentionner ici que le recensement est une source importante, mais également diverses études en accès restreint comme l'étude logement ou l'étude emploi du temps.

Le gouvernement possède quant à lui aussi de nombreuses données liées au domaine de l'énergie. Certaines sont proposées en accès libre sur le site de données « open data » du gouvernement, comme le cadastre français, rendu public courant 2017, ou encore une base détaillée sur les logements sociaux, décrivant le parc français constitué de près de 5 millions de logements. D'autres données sont en accès restreint, notamment proposées aux collectivités, comme les données de la DGFIP (imposition) sur les bâtiments, appelées MAJIC. D'autres organismes peuvent y accéder, sous réserve de s'affranchir d'un coût d'acquisition.

D'autres acteurs possèdent des données pouvant être utilisées dans le domaine énergétique. Les grands acteurs de l'énergie, tels que les fournisseurs d'énergie, les distributeurs et les transporteurs d'électricité et de gaz, possèdent des données propriétaires (qui leur sont propres) bien entendu liées à leur domaine d'expertise. Ainsi, ENEDIS possède des données de consommation de ses clients mesurées aujourd'hui à l'échelle des 20000 départs HTA au pas de temps de 10 minutes ou client par client avec des relevés de compteur bisannuels. Ces données vont devenir de plus en plus précises et volumineuses avec le déploiement du Linky. Les données client dépassent les énergies mesurées et concernent, client par client, un certain nombre d'informations comme la puissance souscrite, le type de contrat, la puissance de raccordement, le poste électrique de raccordement, etc. Toutes ces données ne sont pas publiques, mais elles sont une base essentielle dans notre projet



1

FIGURE 16 – **Synthèse des bases de données** disponibles en France pour l'analyse de la transition énergétique, la modélisation des gisements techniques de production et de stockage (cf. Section 2.3.4), la simulation de la production et de la consommation (cf. Section 2.3), la prédiction de la production à court terme (cf. Section 2.1) ou à long terme (cf. Sous-section 2.3.3).

MOSAIC [Proj1] que nous discutons essentiellement dans la Sous-Section 2.3.2.

De même, GRDF possède des données de consommation de Gaz, le nom et la localisation de ses clients, etc. Suite à la LTECV, GRDF et ENEDIS proposent leur propre site de données *OpenData*, où l'on peut trouver des données agrégées de consommation à une échelle cohérente avec l'INSEE, à savoir la maille IRIS, mais aussi à la maille régionale, la maille départementale, communale ou EPCI. D'autres données comme la consommation à l'échelle nationale, des profils de consommateurs ou des localisations de bornes électriques pour véhicules électriques et hybrides sont également disponibles.

Enfin, les données permettant de caractériser dans l'espace et dans le

temps l'évolution d'une situation météorologique sont essentielles. Pour caractériser la production solaire, éolienne ou hydraulique d'un côté, mais aussi pour caractériser la consommation électrique qui dépend beaucoup de la température. Ces données peuvent être issues d'observations de la terre comme des stations météorologiques ou des images satellites. Elles peuvent aussi croiser des observations de la terre et des modèles météorologiques, on parle alors de « réanalyses ». Ces données sont souvent produites et gérées par des organismes météorologiques nationaux comme Météo France, DWD, ou la NASA, ou internationaux comme ECMWF, mais un certain nombre de PME se sont aussi spécialisées ces dernières années dans la gestion et la production de telles données et de plus en plus le terme de « météorologie de l'énergie » est utilisé pour les caractériser. On notera ici l'importance des données météorologiques également dans la gestion du système électrique avec les prévisions à court terme *Numerical Weather Prediction* (NWP) issues des modèles météorologiques que nous avons largement utilisés dans nos travaux sur la prédiction à court terme.

Le problème de l'association de bases de données. Ces bases de données permettent une caractérisation plus fine des facteurs qui impactent notre consommation d'électricité. On peut ici parler d'économétrie de l'énergie tant beaucoup de méthodes développées dans le domaine de l'économétrie peuvent s'avérer utiles. La dimension géographique de ces données est également importante.

Les bases de données disponibles offrent des informations de natures et de sources variées qui sont parfois redondantes, parfois complémentaires. Leur utilisation simultanée est donc un enjeu important à la fois dans un but de consolidation et dans un but d'accroissement de l'information utilisée. Il y a malheureusement un certain nombre de difficultés dans ces tâches la principale étant que les périmètres considérés dans les différentes bases de données ne sont pas forcément les mêmes et il n'existe pas toujours une manière simple de pouvoir plonger une base dans une autre.

Par exemple la base de données du recensement produite par l'INSEE offre des renseignements pour chaque foyer français ou plus exactement pour chaque foyer sondé mais les données sont extrapolées à chaque foyer, cependant que pour des raisons de confidentialité évidentes chacun de ces foyers n'est pas associé à une localisation géographique exacte mais via son appartenance à un IRIS. Les informations du recensement sont malgré tout

utiles à la caractérisation de la consommation électrique de chaque client ne serait-ce que par exemple parce qu'elles permettent de connaître le mode de chauffage utilisé. Dans le même temps ENEDIS possède pour caractériser ses clients une liste avec des adresses permettant une géolocalisation assez précise, des consommations mesurées, des informations concernant le type d'abonnement et la puissance souscrite. L'enrichissement de cette base de données avec l'information contenue dans la base du recensement est donc intéressant mais difficile. Nous avons proposé dans le cadre du travail de thèse de Thibaut Barbier [These8] et dans l'article de conférence [Confs34] une méthode permettant cet enrichissement que nous appelons aussi *matching*. Cette méthode repose sur la formulation d'un problème d'optimisation et la mise en œuvre d'un algorithme de résolution adapté.

Un autre exemple de problème de *matching* est donné lorsque l'on cherche à enrichir la base de données des clients de ENEDIS avec la base MAJIC ou lorsque l'on cherche à enrichir la base de données géo-référencée des bâtiments de l'IGN (BD TOPO) avec MAJIC. Ce problème a été étudié dans la cadre de la thèse de Antoine Rogeau [These7], dans l'article de conférence [Confs47] et dans le travail pour le projet Shape [Proj6].

Optimisation de la rénovation et du changement de mode de chauffage pour la transition énergétique. Les politiques s'engagent aujourd'hui dans la transition énergétique à différentes échelles. À l'échelle nationale par des lois comme celle que nous avons mentionnée en introduction, mais également à l'échelle des collectivités territoriales, dans les régions avec les SRCAE ou à des échelles plus fines comme avec les territoires à énergie positive. Les objectifs correspondants sont le plus souvent donnés à une échelle macroscopique comme une réduction de la consommation d'énergie, une production locale d'énergie renouvelable, ou une réduction d'émissions de GES mais les stratégies de mise en œuvre ne sont pas précisées. La connaissance du territoire permise par les bases de données mentionnées aux paragraphes précédents doit permettre d'envisager une description plus précise des actions concrètes à mettre en œuvre pour atteindre ces objectifs.

Dans la cadre de la thèse de Antoine Rogeau [These7], effectuée en partenariat avec le CSTB dans le projet Shape [Proj6], nous proposons des techniques d'optimisation pour accompagnement des collectivités dans la mise en œuvre de la transition énergétique. Dans l'article de conférence [Confs50] et l'article soumis, [ArticleP1] nous avons proposé une méthode numérique

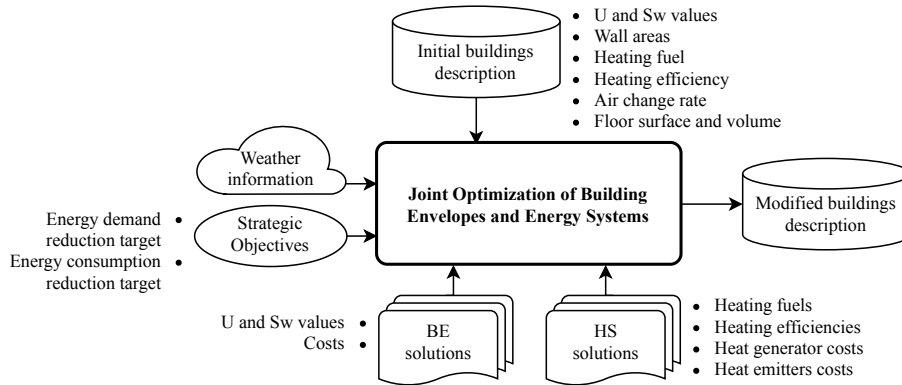


FIGURE 17 – **Structure de l’outil d’optimisation** de la rénovation et du changement de mode de chauffage dans un territoire développé dans [ArticleP1].

permettant d’optimiser un ensemble de stratégies de rénovation et de changement de mode de chauffage dans le but d’atteindre des objectifs multiples à l’échelle du territoire (e.g. réduction de consommation d’énergie, réduction d’émission de GES). L’objectif à l’échelle du territoire implique un couplage entre tous les bâtiments dans le problème d’optimisation. L’ensemble des données utilisées et la structure de l’outil d’optimisation sont représentés à la Figure 17. Le travail en cours ici est de mettre en œuvre une méthode permettant de générer des scénarios alternatifs variés et proches du scénario optimal pour proposer aux décideurs une panoplie de trajectoires variées, mais toutes efficaces.

2.3.2 Modélisation de la consommation électrique

La transition énergétique et la volonté d’accroître le contrôle en temps réel du réseau de distribution obligent à repenser les règles de dimensionnement du réseau de distribution d’électricité. Jusqu’à maintenant ce réseau était dimensionné pour faire face à quelques événements extrêmes liés à la perte d’un ouvrage ou à une consommation forte liée à une situation de grand froid. La transition énergétique appelle ici à un changement de paradigme puisque les nouveaux usages à venir (comme le véhicule électrique), l’efficacité énergétique et la production renouvelable injectée dans le réseau de distribution transforment la nature même des flux dans le réseau. Ainsi, les événements dimensionnant pour le réseau de distribution peuvent être

par exemple liés à une forte production locale ou à l'utilisation du véhicule électrique et non plus à une période de grand froid.

Les possibilités nouvelles de contrôle des puissances actives et réactives injectées et soutirées doivent permettre de soulager le réseau dans les heures les plus complexes et seront, elles aussi, à l'origine de changements dans les règles de dimensionnement du réseau. La mesure des économies sous-jacentes, qui sont le plus souvent des économies d'investissement reste un enjeu important aujourd'hui. Ce contexte amène de nouveaux problèmes d'optimisation sur lesquels nous avons travaillé et qui sont discutés dans la Sous-section 2.2.

Je présente ici nos travaux sur la simulation de la charge (production et consommation). Dans ce contexte j'ai monté le projet MOSAIC qui est une collaboration directe avec ENEDIS sur la simulation de la charge dans le réseau de distribution. Cette collaboration dure depuis maintenant 2013 et je coordonne ce projet qui a pour l'instant a donné lieu à de nombreux stages, un postdoc et une thèse soutenue [These8]. Ce projet est aussi alimenté par les thèses [These5] et [These7]. Le montage d'une suite est en cours pour 2020.

Le projet MOSAIC repose à la fois sur une modélisation physique et sur l'analyse statistique de grandes quantités de données. La modélisation physique est sommaire, mais concerne chaque client et chaque appareil pour un très grand nombre de clients (potentiellement tous les clients de milliers de départements HTA). Il ne s'agit pas ici d'avoir une modélisation la plus fine possible, mais plutôt un modèle comportant des caractéristiques physiques essentielles et auxquelles il soit possible d'accéder par un simple accès à des bases de données déjà disponibles. Les données disponibles sont ici :

- les données ENEDIS : l'ensemble des données clients (table donnant des informations client par client), des mesures de départ HTA faites par ENEDIS depuis plusieurs années à un pas de temps de dix minutes. Prochainement les données linky seront intégrées à ce projet.
- des données décrivant les bâtiments (BD TOPO IGN), la BD parcellaire, le recensement et MAJIC.
- des données décrivant les habitants : les locaux français (base MAJIC obtenue à partir des impôts locaux), des données de recensement (INSEE).
- des données météorologiques (réanalyses, stations météorologiques, images satellites, simulations, etc.).
- des données décrivant les entreprises françaises (base SIRENE de l'INSEE).

L'outil développé a été intégré au système informatique de ENEDIS qui

le déploiement dans ses antennes régionales pour permettre l'analyse de différents projets novateurs de développement du réseau. Il est aussi utilisé par notre groupe de recherche pour des analyses prospectives par exemple dans l'étude de la répliquabilité de certains démonstrateurs *smart grid*. [These8, Confs34, Confs25, ArticleP13, Confs31][These5] [Confs42, Confs45]. Ce travail s'organise autour de plusieurs enjeux scientifiques parmi lesquels les plus importants sont :

1. Évaluer la valeur de l'information apportée par les différentes bases de données existantes pour la caractérisation de la consommation.
2. Fournir à ENEDIS un ensemble de modèles validés sur des données observées pour simuler l'évolution de la charge dans le cadre de projets de transition énergétique territoriaux effectués en partenariat avec les collectivités.
3. Accompagner ENEDIS vers la redéfinition de ses règles de dimensionnement.

Simulation dynamique par une méthode statistique directe. Dans le cadre de la thèse [These5] et dans l'article, [ArticleP13] nous avons proposé une méthode statistique permettant de reconstruire la dynamique de la consommation électrique de catégories de clients à partir de l'observation d'un grand nombre de départs HTA sur lesquels seule la somme de la consommation associée à ces catégories est mesurée. Ces travaux constituent une méthode de référence purement statistique pour l'exercice consistant à simuler un départ HTA sur lequel la dynamique de la consommation n'est pas observée à partir de la description des clients desservis.

La modélisation s'inspire du domaine de la séparation de sources. La charge au niveau d'un départ HTA f à un temps t $d^f(t)$ la charge non-thermosensible mesurée sur ce départ est supposée être l'agrégation de différents profils $d_1(t), \dots, d_K(t)$ associés à K catégories de consommateurs avec des poids p_1^f, \dots, p_K^f :

$$d^f(t) = \sum_{k=1}^K p_k^f d_k(t) + \varepsilon^f(t), \quad (12)$$

où $\varepsilon^f(t)$ est un terme résiduel. On observe $d^f(t)$ dans un grand nombre de départs $f = 1, \dots, F$. Sur chacun de ces départs, les proportions des différentes catégories sont obtenues à partir de la base de données client de ENEDIS, mais l'on peut aussi utiliser d'autres bases pour créer différentes

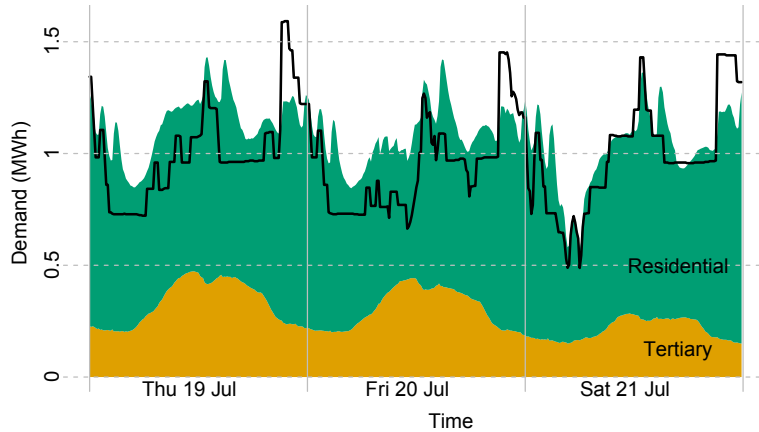


FIGURE 18 – **Prédiction d'un nouveau départ HTA** dans la ville de Lyon à partir de deux catégories : tertiaire et résidentiel. L'observation de la consommation pour ces trois jours est donnée par le trait noir.

catégories. Nous avons proposé un algorithme d'optimisation permettant de retrouver les fonctions $d_k(t)$ en les décomposant dans une base bien choisie et en cherchant les coefficients de la décomposition qui permettent de minimiser une erreur quadratique.

Un exemple de résultat sur un départ de Lyon (qui n'était pas dans la base d'apprentissage) est donné dans la Figure 25 et les résultats moyens obtenus et détaillés dans l'article sont rappelés Table 3.

Nous montrons que cette méthode peut être utilisée pour plusieurs objectifs. Tout d'abord pour comparer l'intérêt qu'il y a dans différentes catégorisations de clients. En effet, deux catégorisations données vont conduire à des profils de catégories différents et à des performances de prédiction différentes. La comparaison que l'on fait alors permet d'obtenir des réponses partielles au premier enjeu scientifique mentionné dans le paragraphe précédent (évaluer la valeur de sources d'information existantes).

Ensuite, la méthode proposée permet de séparer l'observation de la consommation au niveau d'un départ HTA en consommations de plusieurs catégories, on peut par exemple séparer la consommation observée au niveau d'un départ donné en consommation résidentielle et consommation tertiaire.

Region	Year	2 categories	8 categories	9 categories	12 categories
BLOIS	2010	24.36 (2.32)	23.90 (2.75)	24.04 (2.99)	26.78 (2.91)
	2011	23.87 (1.62)	22.79 (1.42)	22.91 (1.16)	24.78 (2.09)
	2012	22.84 (1.26)	22.54 (1.17)	22.09 (1.24)	24.09 (2.17)
	2013	22.34 (2.06)	22.31 (1.98)	21.32 (1.96)	23.34 (2.04)
LYON	2010	19.05 (2.39)	19.42 (2.71)	18.29 (2.24)	19.23 (1.94)
	2011	19.28 (1.24)	18.06 (1.55)	18.56 (1.20)	18.46 (1.42)
	2012	19.07 (1.35)	18.21 (1.72)	18.30 (1.34)	19.00 (1.86)
	2013	18.06 (1.03)	17.92 (2.03)	17.49 (1.12)	18.68 (1.91)
RENNES	2010	22.57 (0.96)	21.67 (1.23)	21.59 (1.04)	22.70 (1.57)
	2011	22.62 (1.22)	21.54 (1.48)	21.57 (1.06)	22.10 (1.08)
	2012	22.75 (1.11)	22.96 (0.99)	22.39 (0.98)	22.61 (0.84)
	2013	24.94 (1.03)	23.99 (1.37)	24.14 (1.26)	24.08 (1.37)

TABLE 3 – RMSE (en %) de notre modèle pour 3 zones différentes sur 4 années considérées et ce pour différentes catégorisations. La simulation est effectuée 100 fois et la RMSE moyenne est donnée avec un écart type entre parenthèses. Les meilleurs résultats obtenus sur les 4 catégorisations sont fournis en caractères gras.

En outre comme nous l’avons déjà dit cette méthode purement statistique pourra nous servir de référence lors de l’évaluation des méthodes de simulation mixtes (physiques et statistiques) proposées dans le paragraphe suivant et qui constituent le simulateur MOSAIC.

Enfin, nous illustrons l’intérêt de cet outil pour la planification de l’évolution d’un poste source d’un quartier qui s’agrandirait par rapport aux méthodes usuelles qui n’utilisent aujourd’hui qu’une homothétie de la mesure au niveau du départ HTA. Cette dernière application est une réponse assez simple à l’enjeu de redéfinition des règles de dimensionnement.

Les perspectives d’utilisation de cette méthode sont nombreuses, mais la première d’entre elles est d’appliquer cet algorithme sur des zones plus vastes et en intégrant des catégories de clients nouvelles issues par exemple de la base de données MAJIC qui n’étaient pas disponibles à l’époque où ce travail a été mené.

La méthode physique et statistique proposée : MOSAIC. La méthode statistique présentée dans le paragraphe précédent, si elle peut servir de référence pour le problème qui est le nôtre, n’en reste pas moins difficile à exploiter dans l’exercice de la planification dans un contexte de transition énergétique. En effet, dans ce contexte où les différents usages de l’électricité

se transforment (apparition du véhicule électrique, rénovation thermique des bâtiments, efficacité énergétique, production renouvelable, stockage, etc.), les évolutions de la pointe de consommation sur un nouveau départ HTA ou sur un départ HTA existant ne peuvent plus être obtenues par une simple homothétie des consommations passées ou par une méthode purement statistique. Cependant, il n'est pas non plus souhaitable d'en rester à une méthode purement physique qui ne tirerait pas parti de la grande quantité de données disponibles. C'est ce compromis que nous avons cherché à faire en mettant en œuvre la méthode MOSAIC.

Nous donnons ici dans les grandes lignes la méthode de simulation MOSAIC, mais le but est plutôt d'en montrer la philosophie que d'en montrer tous les détails. Nous renvoyons le lecteur à la thèse de Thibaut Barbier et aux articles correspondants pour plus d'informations. Elle repose sur une modélisation simple de chaque appareil et de chaque client. La méthode utilisée pour choisir les appareils est décrite dans la thèse de Thibaut Barbier. La simulation est stochastique et se décompose en 3 temps :

1. la simulation statique : choix des appareils possédés par chaque client ;
2. la simulation dynamique : simulation du lancement des appareils ;
3. l'évaluation de l'énergie consommée sur une grille temporelle (non nécessairement régulière).

La simulation statique repose sur l'utilisation de taux de possession d'appareils en fonction des caractéristiques du client (type de tertiaire, résidentiel, nombre de personnes, surface du logement, etc.).

Concernant la simulation dynamique. On peut distinguer plusieurs types de modèles :

- les modèles par blocs de puissance ;
- les modèles par blocs de puissance périodiques ;
- les modèles de climatisation et de chauffage électrique.

Un exemple de modèle par bloc de puissance est donné par la machine à laver et est illustré par la Figure 19. Les usages suivants ont été choisis pour être modélisés de cette façon : machine à laver, sèche-linge, lave-vaisselle, ordinateur, éclairage, piscine, four, plaques de cuisson, chauffe-eau électrique et véhicules électriques.

De façon plus formelle, la consommation des appareils modélisés par blocs de puissance constante d'un client j s'exprime à un instant t par :

$$P^j(t) = \sum_{a=1}^{A^j} \sum_{n=1}^{N_a^j} \sum_{m=1}^{M_a} P_{a,m}^j \mathbb{1}_{d_{a,n,m}}(t - T_{a,n,m}) \quad (13)$$

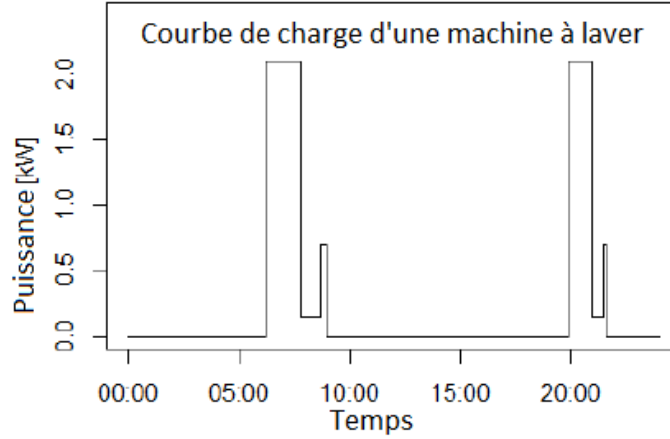


FIGURE 19 – Tracé de la courbe de charge d’une machine à laver modélisée par blocs de puissance constante, lancée deux fois dans la même journée

Avec A^j le nombre d’appareils du client j , M_a le nombre de blocs modélisant l’appareil a , N_a^j son nombre de lancements. On a ici trois variables dépendant de l’appareil, du client et du bloc de puissance considéré : la puissance du bloc $P_{a,m}^j$, la durée du bloc $d_{a,m}^j$, et la date de lancement du bloc $T_{a,m}^j$. La simulation de ces variables est faite à partir de tirages aléatoires en tenant compte de caractéristiques des clients et dans le but d’obtenir à une échelle suffisamment agrégée des courbes cibles de consommation par appareil. Ces courbes cibles sont obtenues dans un premier temps par des campagnes de mesure ou des études mais l’objectif est ici à terme d’intégrer ces courbes cibles, comme des paramètres dans le simulateur qui seront estimés à partir de l’observation d’un grand nombre de départs. Ils peuvent donc être considérés comme un degré de liberté dans le simulateur. Plus de détails sont donnés dans la thèse de Thibaut Barbier et dans les articles donnés au début de cette section. Les profils par blocs périodiques relèvent d’une modélisation similaire mais avec des blocs se répétant périodiquement, cela concerne le réfrigérateur et le congélateur. Les paramètres qui déterminent ce modèle sont la puissance, la durée, la période et le rapport cyclique.

En ce qui concerne le chauffage et la climatisation, nous utilisons un modèle mono-zone de type RC.

$$\frac{dT_{in}}{dt}(t) + \frac{1}{RC}[T_{in}(t) - T_{out}(t)] = \frac{1}{C}[C_{op}(t) \times P(t) + S_{ves}E(t)] \quad (14)$$

où R ($^{\circ}C/W$) représente la résistance thermique modélisant les déperdi-

tions entre l'intérieur et l'extérieur du bâtiment. C est la capacité thermique ($J/^\circ C$), modélisant la chaleur stockée dans le bâtiment. P est la puissance électrique (W) consommée par le système de chauffage, qu'il s'agisse de convecteurs ou d'une pompe à chaleur. Cop est le coefficient de performance du moyen de chauffage, dans le cas d'une PAC, il dépend de la température extérieure et de la température du fluide. E est l'irradiation solaire sur plan horizontal orienté sud (W/m^2). S_{ves} la surface de vitrage équivalent sud (m^2). T_{in} est la température intérieure ($^\circ C$) moyenne du bâtiment, T_{out} la température extérieure ($^\circ C$).

La résistance thermique équivalente est liée aux caractéristiques géométriques du bâtiment par la relation :

$$\frac{1}{R} = U_{mur} \times S_{mur} + U_{pb} \times S_{pb} + U_{ph} \times S_{ph} + U_{baie} \times S_{baie} + Ara \times V \quad (15)$$

où S représente la surface et U un coefficient de transmission thermique et où la convention d'indice suivante a été utilisée : mur – murs sans baies, pb – plancher bas, ph – plancher haut, baie – parois vitrées, et ra – renouvellement air. Le choix de U est ici crucial dans la modélisation, nous avons utilisé dans un premier temps des valeurs qui dépendent de l'âge de construction du bâtiment et qui sont issues de la littérature du bâtiment mais un des enjeux de MOSAIC ici est d'avoir une méthode automatique d'estimation de cette matrice de valeurs pour U en fonction de l'élément considéré et de l'âge du bâtiment. Une modélisation plus fine de U a été obtenue à partir de la relation statistique observée entre les valeurs de consommation annuelle thermo-sensible données par l'open Data ENEDIS à la maille IRIS sur toute la France et les surfaces de logements obtenues à partir du recensement et de l'étude logement de l'INSEE. Les valeurs obtenues sont présentées Figure 20

La géométrie du bâtiment et plus généralement les données caractérisant les clients (e.g. nombre de personnes, consommation en énergie annuelle) sont obtenues en croisant les bases de données suivantes : la BD TOPO, MAJIC, la base du recensement de l'INSEE et la base client de ENEDIS. Un exemple de simulation est donné Figure 21.

La simulation du véhicule électrique a fait l'objet de travaux particuliers dédiés à la fois à la prédiction à court terme et à la simulation long terme [ArticleP22]. Nous avons déjà donné une description des modèles développés dans la section sur la prédiction à court terme.

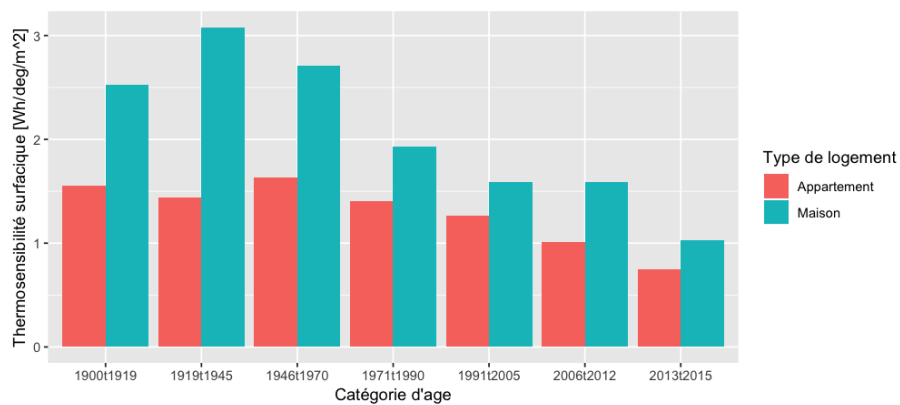


FIGURE 20 – Consommation thermosensible surfacique des résidences principales chauffées à l’électricité estimée par catégorie et date de construction de logement. Ces estimations sont faites à partir d’un côté des Open Data ENEDIS donnant pour tous les IRIS de France la consommation thermosensible résidentielle de 2017, et de l’autre les données du recensement 2015 de l’INSEE qui permet (combiné à l’enquête logement) d’obtenir les surfaces chauffées à l’électricité par type de logement. Certains IRIS ont été écartés (données manquantes ou différences trop importantes entre 2017 et 2015).

Consommation électrique moyenne de 500 clients résidentiels

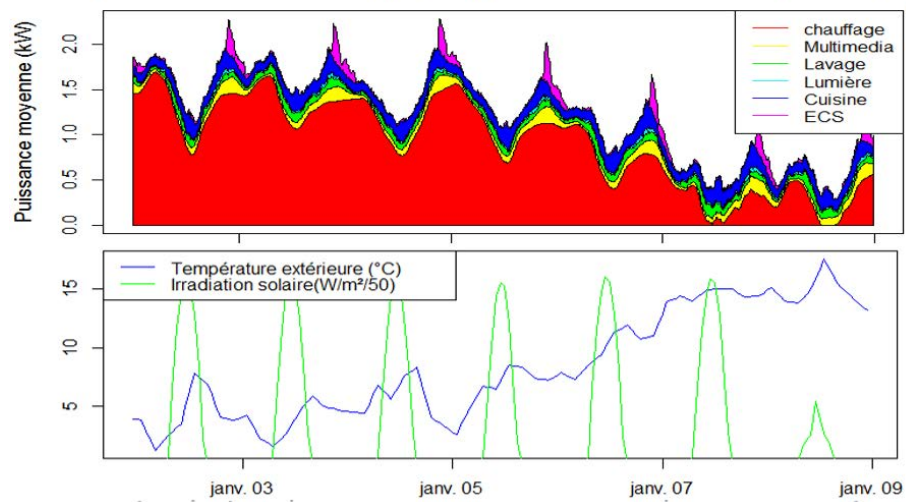


FIGURE 21 – Exemple de résultats de simulation de 500 clients résidentiels. Dans le cadre supérieur est représentée la consommation électrique ramenée par client, décomposée selon les grandes catégories d’usages. Dans le cadre inférieur figurent les conditions météorologiques

2.3.3 Modélisation et simulation de la production renouvelable et des erreurs de prédiction à court terme

Une autre partie de mon travail a concerné la modélisation et la simulation de la **production renouvelable et des erreurs de prédiction** à court terme de cette production, toujours dans une démarche prospective pour l’analyse des systèmes électriques. Ce travail a été entrepris par un stage dans le cadre du projet WinPower [Proj2] puis poursuivi par un premier postdoc de Arthur Bossavy dans le cadre du projet [Proj5] et un second (de Evgenyia Ichikina) dans le cadre du projet [Proj16]. Le travail sur la simulation des erreurs de prédiction de la production éolienne et PV était au cœur du projet SmartReserve [Proj7] dont j’étais le coordinateur et s’est poursuivi dans le WP 1 du projet européen REStable [Proj16].

Il est à noter que ma connaissance de la météorologie et de la prédiction que j’ai développée dès mon arrivée aux mines dans le cadre du projet SafeWind, a joué un rôle important dans cette recherche. Une des contributions du projet SafeWind pour ce problème notamment été de sensibiliser les météorologues du centre européen de prédiction (European Center for Medium Range Weather Forecast, ECMWF) à la problématique de la prédiction de production éolienne, cela s’est traduit par un article de conférence [Confs9] mais aussi par la décision d’ECMWF d’archiver les vitesses de vent à 100 *m*. Les nouvelles analyses de ECMWF (ERA5) sorties en 2017 contiennent d’ailleurs ces vitesses de vent (alors que ni ERA40 ni MERRA, les deux autres bases de données de réanalyse ne les contenaient).

Simulation de la production renouvelable. Dans ce problème la production n’est pas toujours observée et simulée à l’échelle d’une centrale, par exemple si l’on cherche à modéliser la production éolienne d’une région française il faut tenir compte de la diversité d’implantations de centrales. La variabilité de la production correspondante est moins importante mais les facteurs expliquant cette production sont plus nombreux puisqu’un grand nombre de localisations géographiques sont à prendre en compte. Finalement on peut distinguer trois éléments importants dans nos méthodes de simulation de la production renouvelable :

- la simulation de vent ou d’irradiation solaire ;
- la simulation des implantations de centrale ;
- la transformation de vent ou d’irradiation en production renouvelable.

La simulation du vent repose sur l’utilisation de réanalyses météorologiques et l’irradiation solaire sur l’utilisation des données SODA qui sont produites par le centre OIE de l’école des mines. Ces dernières données sont

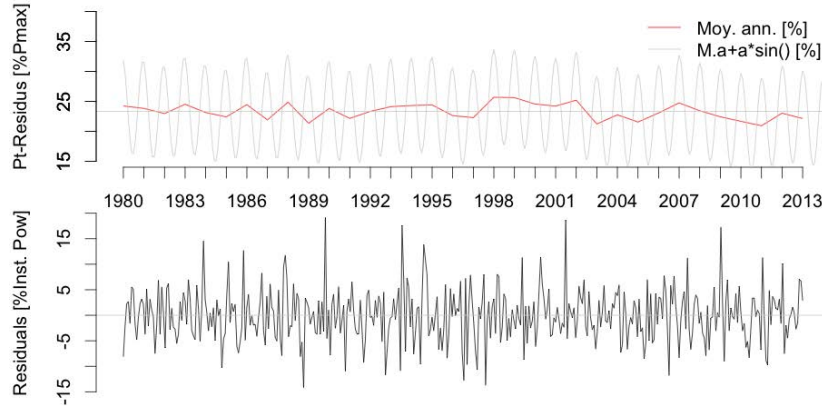


FIGURE 22 – Exemple d’une simulation de la production éolienne sur la France obtenue à partir des réanalyses de vent de la NASA (MERRA) jusqu’à 1980, calculée au pas de temps horaire puis ramené au pas de temps mensuel. Le modèle de simulation a été calibré à l’échelle régionale sur la production éolienne de 2013. Les simulations sont décomposées en trois composantes : une variation d’une année sur l’autre, une variation saisonnière sinusoïdale et un résidu ne présentant pas de régularité particulière.

obtenues en combinant des images satellites, des réanalyses et des modèles de transfert radiatif. Ces données de réanalyse couvrent souvent des périodes assez longues et peuvent être remplacées par des modèles de changement climatique. Elles peuvent donc servir à comprendre d’un côté certaines régularités de la production d’un pays ou d’une région et de l’autre la distribution des événements extrêmes. Nous donnons Figure 22 un exemple de décomposition long terme de la production éolienne obtenue à partir d’un modèle sur la France avec des données de vent de la NASA (MERRA).

La simulation de l’implantation des centrales est un enjeu important, nous avons observé dans le projet [Proj5] que les résultats peuvent être assez variables selon les choix de localisation. Une première méthode utilisée consiste à rassembler des données de localisation actuelles. Cette méthode a été utilisée dans [Confs36, Confs28]. Une deuxième méthode plus prospective réside dans l’optimisation des localisations futures ou dans leur simulation. Nous avons proposé une méthode reposant sur l’optimisation du facteur de charge et d’un terme aléatoire. Le terme aléatoire permet de modéliser le fait que les centrales ne sont pas toujours construites dans les zones les plus

favorables sur le plan de la production d'énergie.

La simulation de la production repose ensuite sur des modèles statistiques ou physiques qui permettent de transformer les données météorologiques en données de puissance. Dans le projet WinPower [Proj2] et dans [Confs23] nous avons discuté ce type de modèle. Un enjeu ici est d'avoir un modèle suffisamment explicite sur le plan de la physique pour permettre de modéliser l'impact d'évolution technologique. Par exemple pour la production PV il faut être capable de modéliser l'impact de différents systèmes de *tracking*. Il existe aujourd'hui des centrales PV équipées de systèmes de *tracking* qui permettent d'incliner les panneaux à l'horizontale lorsqu'il y a des nuages pour mieux récupérer le rayonnement diffus majoritaire dans ces cas. Pour la production éolienne, on parle beaucoup aujourd'hui de l'augmentation du rapport entre la longueur des pales et la puissance installée. Cette augmentation modifie la courbe de puissance et permet des meilleurs facteurs de charge et il faut donc des modèles intégrant ce type de phénomène. Les modèles proposés et utilisés dans [Confs36, Confs28] permettent ce type d'analyse. En même temps que cette modélisation physique, il s'agit d'avoir des modèles ayant des degrés de liberté ajustables sur des données observées. Le caractère statistique des modèles permet de pallier leurs imperfections et est aussi un gage de qualité dans les simulations. Les données observées sont des données régionales en France ou des données nationales dans plusieurs pays européens. Les modèles correspondants ont été décrits dans les rapports des projets [Proj16] et [Proj5] ainsi que dans [Confs36] et [ArticleP12].

Simulation des erreurs de prédiction à court terme de la production renouvelable variable

Dans un contexte prospectif, les besoins de réserve tertiaire sont susceptibles d'évoluer avec le volume des erreurs de prédiction à court terme du PV et de l'éolien. On peut essayer d'estimer les niveaux d'erreur futurs en supposant qu'elles seront similaires dans leur forme à celles qui sont faites aujourd'hui tout en grandissant proportionnellement au facteur de charge. Mais cette approche comporte plusieurs limitations :

- elle ne permet pas de simuler plus d'années météorologiques que celles pour lesquelles ces erreurs ont été observées ;
- elle ne permet pas de modéliser l'effet de foisonnement qui peut apparaître dans différentes configurations futures de localisation des centrales.

Dans le cadre Projet SmartReserve [Proj7] à l'échelle de la France puis

dans le cadre du projet REStable [Proj16] à l'échelle européenne nous avons proposé une méthode pour simuler les erreurs de prédiction de la production. À l'échelle de la France et dans la continuité des travaux présentés dans le paragraphe précédent, nous avons utilisé ces modèles de simulation dans le projet prospectif « vers un mix 100% renouvelable » [Proj5]. Ces travaux ont donné lieu à deux articles de conférence [Confs33, Confs32].

2.3.4 Sensibilité des estimations de gisements territoriaux pour la transition et des analyses de cycle de vie.

La recherche de modèles simplifiés des systèmes complexes est une tâche statistique relativement classique qui s'apparente à la régression, et la recherche de paramètres déterminants est dans ce cadre une sélection de variables. La méthode d'analyse généralisée de la variance introduite par Sobol [ext4] et Saltelli [ext2] permet d'estimer la part de la variance d'une variable Y expliquée par un ensemble de variables $X = (X_1, \dots, X_p)$. Elle est donc un outil intéressant pour l'analyse de sensibilité et la hiérarchisation des paramètres d'entrée d'un système. Les paramètres d'entrées peuvent être aussi nombreux qu'on le souhaite et permettent de refléter :

- l'existence d'une incertitude sur une valeur (comme le niveau d'acceptabilité dans la détermination du potentiel éolien) ;
- l'existence d'un levier significatif (la distance minimale à un radar est un levier pour le potentiel éolien, l'amélioration de la performance lors de la fabrication des panneaux a été un levier important pour le PV ces dernières années) ;
- ou d'une variabilité intrinsèque de la variable considérée (c'est le cas par exemple du facteur de charge du PV ou de l'éolien qui sont liés aux conditions météorologiques moyennes qui peuvent varier d'une année sur l'autre ou d'un site à l'autre).

L'observation de la variabilité de la sortie étudiée en fonction de ces paramètres a bien un intérêt mais nous utilisons des techniques de Sobol pour estimer la part de variance de la sortie expliquée par chacune des entrées. On peut alors mettre en avant les facteurs qui déterminent le plus la variabilité de la sortie.

Un avantage important de cette méthode est qu'elle ne repose presque aucune hypothèse concernant la relation qui lie Y aux variables explicatives. Elle est donc tout indiquée pour l'analyse de systèmes complexes pour lesquels un consensus de modélisation est difficile à trouver. Nous avons déployé cette approche dans un certain nombre de domaines, comme l'analyse des

facteurs influençant la variabilité du système électrique [ArticleP23] (dans le cadre de la thèse de Thomas Heggarty, nous avons déjà discuté ce travail), mais aussi l'estimation des impacts environnementaux des systèmes énergétiques [ArticleP9, Confs12, Confs22, Confs43, ArticleP15], et l'estimation de gisements de production renouvelable [Confs28, Confs36, ArticleP12], ou de stockage micro-hydraulique [ArticleP16]. Par opposition aux méthodes « locales » l'analyse globale de la sensibilité n'impose pas de restreindre pas l'analyse à un domaine de valeurs de X . Cependant nous avons montré dans [ArticleP15] qu'elle impose de faire des hypothèses sur la distribution statistique de X qui peuvent avoir un impact sur les résultats.

Analyse de sensibilité dans l'analyse de cycle de vie des systèmes énergétiques. Dans le domaine de l'énergie en particulier ces analyses de cycle de vie sont essentielles et peuvent concerner plusieurs types d'indicateurs. Les objectifs de baisse d'**émissions de CO_2** fixés par les gouvernements dans le cadre des programmes de lutte contre le changement climatique en appellent à des évaluations fines des émissions, et toute énergie même renouvelable ou nucléaire, a à ce titre un impact qu'il faut savoir évaluer. Un autre indicateur important est le **temps de retour énergétique**, car sa baisse prévue dans le passage du pétrole aux énergies renouvelables inquiète certains économistes et historiens. Enfin la croissance du secteur de l'énergie à l'échelle planétaire, et en particulier celui du secteur des énergies renouvelables, contribue à l'augmentation de la demande mondiale de **matériaux** et l'évaluation de cette contribution en fonction de la rareté est alors nécessaire. Pour ces trois familles d'indicateurs importantes, les méthodes d'analyse de cycle de vie (ACV) forment un ensemble d'outils pertinents qui reposent sur l'utilisation de grandes bases de données (comme la base eco-invent) contenant des informations sur les impacts (e.g. consommation d'énergie, de matières premières, émissions de CO_2) d'un ensemble de processus et sur leurs interdépendances.

Une particularité des ACV est leur l'ambition de faire la lumière sur tous les impacts d'un système dans sa plus grande complexité en considérant celui-ci « du berceau à la tombe ». Toutes ces évaluations reposent sur des bases de données décrivant les systèmes énergétiques et les territoires de manière aussi précise que possible. Elles sont également des outils utilisés directement par les décideurs, les citoyens et bientôt les consommateurs dans leur ensemble et à ce titre elles ont tendance à tendre dans leur utilisation ou leur formulation, vers un objet très simple. Cette simplification est le plus

souvent utile et pratique, elle permet une large diffusion de l'impact, mais elle peut aussi être instrumentée pour discrétiser un système et alors c'est toute la puissance du crédit que l'on donne à l'outil qui s'en trouve diminué. En outre les descriptions disponibles dans les bases de données sont loin d'être parfaites et il faut s'adapter entre autres à la variabilité des conditions de mise en œuvre ainsi qu'aux lacunes et aux incertitudes qu'elles contiennent.

Face à cette difficulté scientifique, j'ai travaillé à proposer des méthodes qui ne permettent pas seulement une estimation -des potentiels, émissions, temps de retour énergétique ou utilisation de matières premières- mais aussi à une évaluation et une analyse des incertitudes sous-jacente et de leurs origines. Cette analyse de sensibilité repose sur le calcul d'indices de Sobol. Mes travaux sur ce sujet ont pour l'instant été le résultat d'une collaboration suivie avec Isabelle Blanc qui était dans mon centre de recherche jusqu'en 2012 et qui dirige des recherches sur le thème des ACV. J'ai travaillé sur une méthodologie générale pour analyser les sources d'incertitudes dans une analyse de cycle de vie qui a été appliquée aux systèmes photovoltaïques et éoliens ainsi qu'à la géothermie [ArticleP9, Confs12, Confs22, Confs43]. J'ai contribué à cette activité par ma connaissance des indices de Sobol et des systèmes énergétiques, essentiellement dans le cadre des thèses de Pierre-Yves Padey et Martino Lacirignola qui ont donné une importance centrale aux indices de Sobol. Dans cette dernière thèse, j'ai proposé l'idée d'une analyse de sensibilité qui a été exploitée et a donné lieu à la publication [ArticleP15]. J'ai monté un cours pour une journée sur le sujet de la variabilité, de l'incertitude et de l'analyse de sensibilité dans les ACV et j'ai donné ce cours dans une série de Cours Doctoral Excellence (CDE1) organisé dans le cadre du réseau EcoSD [Ray7], piloté par Isabelle Blanc à MINES ParisTech.

Les impacts environnementaux des systèmes éoliens et photovoltaïques dépendent d'un grand nombre de facteurs qui relèvent évidemment de la localisation, mais aussi des technologies utilisées. Certains de ces facteurs sont entachés d'incertitudes ou de variabilité.

J'ai travaillé sur une méthodologie générale pour analyser les sources d'incertitudes dans une analyse de cycle de vie qui a été appliquée aux systèmes photovoltaïques et éoliens ainsi qu'à la géothermie [ArticleP9, Confs12, Confs22, Confs43]. Ces analyses de sensibilité de type Sobol ont l'avantage de nécessiter peu d'hypothèses, mais j'ai contribué à une méthodologie nouvelle pour permettre de tenir compte de la sensibilité des résultats à ces hypothèses [ArticleP15]. Dans le cas du PV les données évoluent très vite, car les procédés de fabrication des panneaux ont considérablement évolué ces dernières années. Dans tous les cas un paramètre important est donné par le facteur d'émission du pays dans lequel les panneaux sont construits et

le facteur de charge.

Dans le cas de la **production éolienne**, nous avons montré dans [ArticleP9, Confs12] que la durée de vie du système LT exprimée en années et son facteur de charge FC en pourcentage expliquent presque 85 % de la variabilité des émissions de GES Em à travers la relation

$$Emission \left[\frac{gCO_{2eq}}{kWh} \right] = \frac{5253 - 5.2LT}{FC \times LT}$$

Dans le cas de la **géothermie profonde**, nous avons montré dans [ArticleP15] qu'un modèle réduit à 5 paramètres était suffisant pour expliquer plus de 70 % de la variance des émissions de GES via la relation :

$$Emission \left[\frac{gCO_{2eq}}{kWh} \right] = \frac{Nw \cdot (\omega_1 \cdot z + \omega_2) + LT \cdot (\omega_3 \cdot f + \omega_4 \cdot P_{ORC})}{(P_{ORC} - f \cdot \omega_5) \cdot LT}$$

avec Nw le nombre de puits (un ou deux), LT la durée de vie de l'installation en années, f le flux en $[kg/s]$, P_{ORC} la capacité installée en kW ,

$$w_1 = 120,70 [gCO_{2eq}/(m \cdot h/y)], \quad w_2 = 5161,87 [gCO_{2eq}/(h/y)],$$

$$w_3 = 61,82 [gCO_{2eq} \cdot s/(kg \cdot h)], \quad w_4 = 6,42 [gCO_{2eq}/(hWh)]$$

et

$$w_5 = 6,10 [kW \cdot s/kg].$$

Détermination du gisement éolien, PV et micro-hydraulique en France. Une difficulté importante posée aujourd'hui par la production renouvelable est l'utilisation du territoire qu'elle impose. Cette difficulté se traduit par une acceptabilité mitigée de ces modes de production, voire parfois par une franche opposition. J'ai proposé des méthodes pour l'évaluation du potentiel technique qui est lié à la fois aux conditions météorologiques et à la surface disponible. Ces méthodes croisent des contraintes techniques, des contraintes environnementales, des contraintes réglementaires et reposent encore une fois sur l'utilisation de modèles mathématiques et de grandes bases de données géographiques. Elles sont elles aussi associées à une analyse de sensibilité.

Nous avons débuté notre travail sur l'estimation du gisement en 2013 avec le projet « vers un mix 100 % renouvelable » [Proj5, Confs36]. Dans ce projet, avec Arthur Bossavy (alors en postdoc), nous avons proposé une méthodologie pour l'estimation du gisement solaire et éolien. Cette méthodologie repose sur l'utilisation de plusieurs SIG et se décompose en 3 temps :

- Application de masques permettant de retirer toutes les zones interdites pour la construction (e.g. pour l'éolien distance aux bâtiments, aux radars, altitude ou pente trop élevée, etc.);
- Estimation du facteur de charge et filtrage des zones où celui-ci est trop faible;
- Application d'un facteur d'acceptabilité et d'un ratio d'utilisation de surface par MW installé sur les surfaces restantes.

Une synthèse de la sensibilité des résultats aux paramètres de facteur de charge et au ratio d'acceptabilité est présenté Figure 23.

Dans [ArticleP12] nous avons proposé une extension de cette méthodologie intégrant une analyse de sensibilité à partir d'indices de Sobol et nous avons illustré son fonctionnement sur deux régions françaises.

Ce type de méthodes a pu être appliqué également au stockage micro-STEP dans l'article [ArticleP16]. Ces travaux ont été effectués dans le cadre du projet ADEME Berrien et du stage d'Antoine Rogeau. La méthodologie proposée est ici un peu plus complexe, elle repose sur les points suivants.

1. Détection des réservoirs (dépressions détectées à partir d'un modèle numérique de terrain et bassins existants disponibles via l'IGN);
2. Création des connexions de réservoirs en considérant de faibles limites structurales (hauteur de chute et distance);
3. Application de différents niveaux de contraintes techniques et environnementales sur les bassins et les connexions (acceptabilité structurale, environnementale, urbaine, occupation du sol, etc.);
4. Sélection des meilleurs sites par optimisation de l'appariement des sites lorsque des réservoirs sont interconnectés, critère énergétique (voir Section 2.4);
5. Analyse de sensibilité (Bouclant sur les étapes 3 et 4).

Une synthèse de la méthodologie est présentée en Figure 25. Un exemple de localisation est donné par la Figure 25. Les résultats hors analyse de sensibilité sont présentés dans la Table 4 et donnent un gisement en France du stockage micro-STEP assez significatif (une ou quelques dizaines de GW).

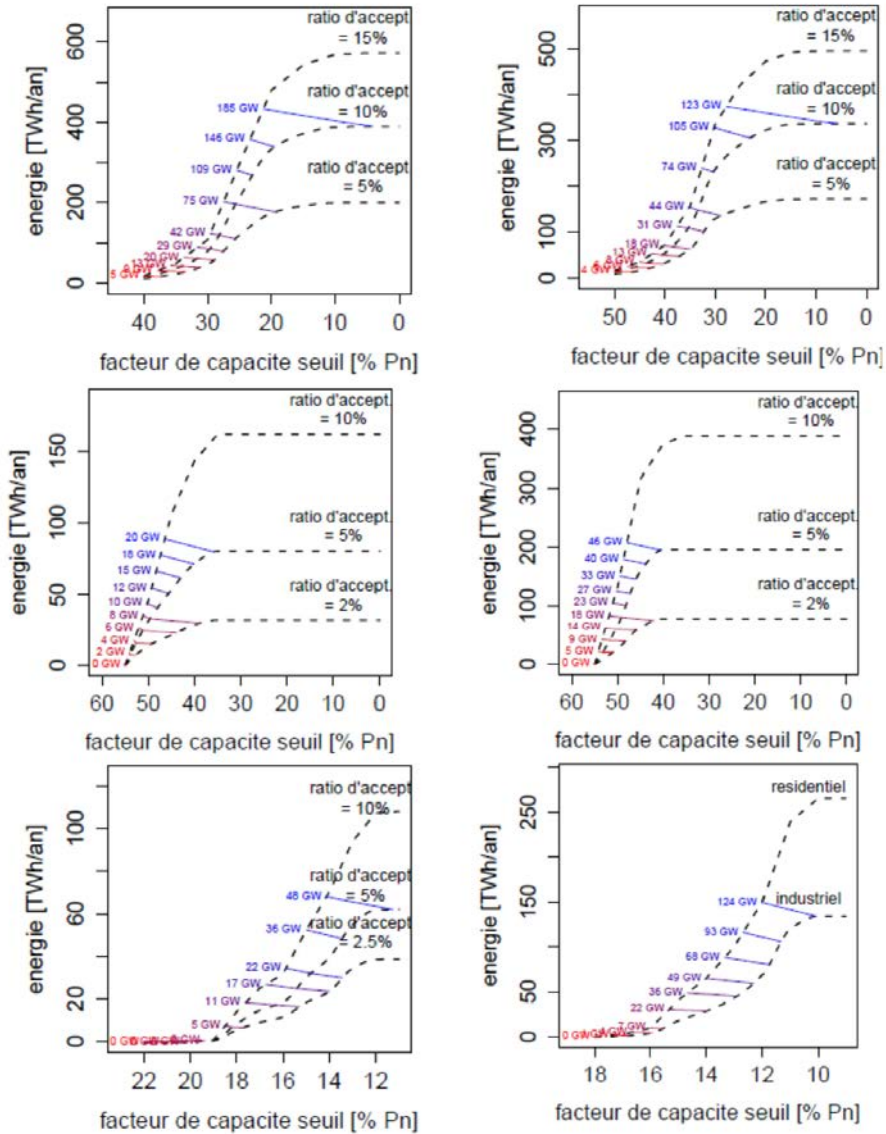


FIGURE 23 – Synthèse des résultats obtenus pour le gisement éolien (en mer et sur terre) et pour le gisement PV France. Sensibilité des gisements renouvelables nationaux aux valeurs de ratio d'acceptabilité et de facteur de charge seuil. Les résultats montrés ici concernent de haut en bas et de gauche à droite, les différentes filières renouvelables : éolien terrestre (turbines anciennes et nouvelle génération), l'éolien en mer (turbines non flottantes et flottantes), le PV au sol et sur toitures (résidentielles et industrielles).

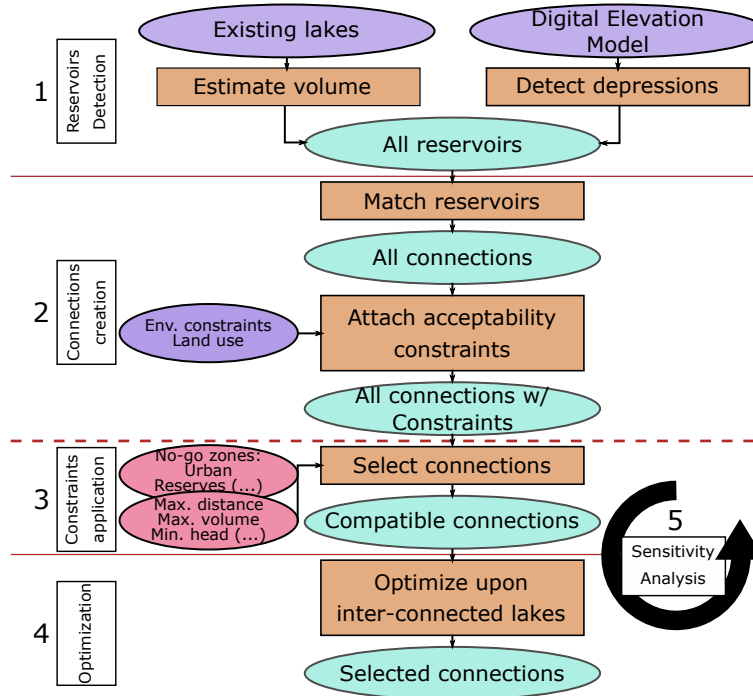


FIGURE 24 – Synthèse de la méthodologie en 5 parties proposée pour l’estimation du gisement de stockage micro-STEP et appliquée en France.

TABLE 4 – Résultats obtenus pour différents types d’optimisation des interconnexions : T_1 connexion entre deux lacs existants, T_2 connexion entre un lac existant et une dépression de terrain, T_3 connexion entre deux dépressions de terrain. Pour les codes couleur : rose : Type 1, orange : Type 2 et vert : Type 3

Conne- c- Na- ture	Pri- ority	Energy bef. optim. (GWh)	Final Energy (GWh)	Proportions of built connections
All	No	78.42	33.56	
All	Yes	78.42	32.93	
T1 T2	Yes	51.18	21.76	
T1	/	30.35	14.35	

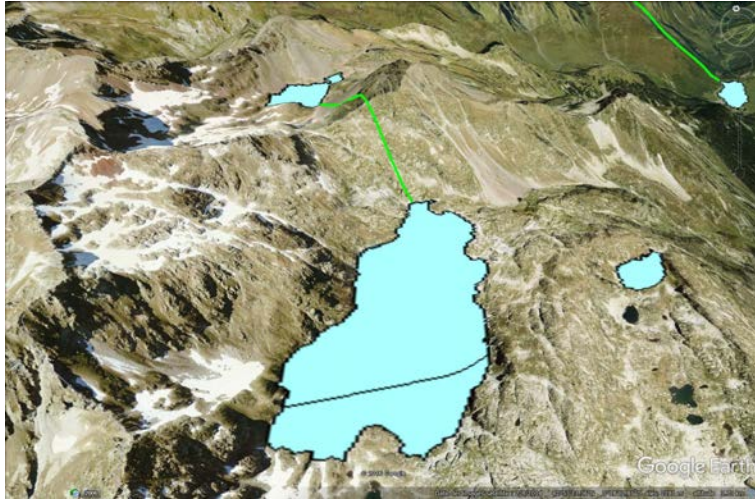


FIGURE 25 – **Exemple de localisation d’une micro-STEP** avec remplissage des deux cuvettes identifiées et une matérialisation du raccordement proposé

3 Projet de recherche

Pour mes recherches à venir, j’ai un projet qui peut se décliner selon la même structure que ma recherche jusqu’à maintenant tant la plupart des sujets évoqués ne sont pas fermés et soulèvent de nouvelles questions encore aujourd’hui.

Concernant la prédiction de production, l’intégration d’une grande quantité de données aux méthodes spatio-temporelles actuelles est une voie prometteuse. On pourrait par exemple combiner des données de nature différente, comme des données de différents types de radars, des images satellites, des mesures sur site de production ou des observations météorologiques. Les modèles correspondants pourraient, quant à eux, inclure des contraintes ou des équations modélisant de manière simplifiée la dynamique de l’atmosphère, rapprochant ainsi le domaine de la statistique spatio-temporelle de celui de l’assimilation rapide et de la météorologie simplifiée. Dans le même temps, la consolidation des données de production avec ces mêmes méthodes est un sujet important. En effet, la qualité des données de production est variable et la démultiplication des sites impose de savoir traiter automatiquement les valeurs aberrantes. Or, ces valeurs aberrantes peuvent beaucoup dégrader les phases d’estimation et de prédiction. Dans ce sens, une pers-

pective intéressante est d'utiliser des mesures sur des sous-systèmes pour détecter des pannes dans les historiques et les écarter de l'échantillon d'apprentissage ou les corriger. Dans un deuxième temps, une correction dynamique pour améliorer les prédictions serait intéressante. Dans le cas du PV et de l'éolien, les sous-systèmes peuvent être des transformateurs instrumentés ou même des mesures obtenues au niveau des onduleurs.

Concernant la planification du réseau de distribution et de transport, incluant des contraintes de gestion, les perspectives existantes sont nombreuses. Elles touchent à la prise en compte des incertitudes avec des algorithmes stochastiques, le couplage avec les contraintes des autres systèmes connectés, comme le système électrique centralisé connecté en amont, mais aussi les systèmes connectés en aval, comme les bâtiments, les systèmes et infrastructures de transport, les systèmes de stockage ou les réseaux de gaz et de chaleur. La prise en compte de ces aspects nécessitera des travaux afin d'améliorer les performances des algorithmes utilisés. Plusieurs pistes me semblent être pertinentes, comme la réduction de scénarios, la décomposition-coordination multi-échelle des problèmes d'optimisation (et des scénarios dans le cas stochastique) ou la mise au point de techniques de réduction de dimension (relaxation automatique guidée par des critères de criticité). La résolution de ces problèmes d'optimisation complexe n'est pas ici le seul problème. En effet, dans un contexte de libéralisation et de participation des acteurs locaux à l'équilibre du réseau principalement de nouveaux modèles de simulation intégrant les différents acteurs et leurs interactions sont nécessaires. La question de la réglementation, la discussion de l'efficacité des mécanismes de marchés sera forcément un sujet d'intérêt. L'adaptation des outils d'optimisation que nous avons développés via des modèles de marché (du type de ceux développés dans le cadre de la thèse de Fiona Foucault) ou des modèles type multi-agent sont des pistes de recherche importantes pour une analyse plus réaliste vis-à-vis du morcellement des acteurs. Sur le plan de l'optimisation les problèmes de gestion du système électrique nécessitent également des approches multi-échelles et stochastique. En outre ces problèmes intègrent souvent des non-linéarités et des chaînes de décision complexes. Des approches permettant de résoudre ces problèmes en continu, au fur et à mesure des remontés d'observation, de la mise à jour des prédictions ou des informations de disponibilité, sont nécessaires.

La modélisation de la consommation électrique et de la production va considérablement changer dans les années à venir avec l'arrivée des compteurs intelligents et le déploiement des énergies renouvelables. La quantité de données correspondante sera bien plus grande et l'on peut penser que les travaux entamés sur l'utilisation des données de mesures disponibles aujour-

d’aujourd’hui au niveau des 20000 départs HTA en France ne sont qu’un préliminaire à des questions autrement plus nombreuses et complexes. Ces données vont permettre de répondre à des enjeux importants, telles la caractérisation de l’efficacité de la rénovation thermique du bâtiment ou la segmentation et la modélisation de la consommation électrique du secteur tertiaire. Le véhicule électrique et son interaction avec le réseau sont aussi des sujets qui sont loin d’être épuisés. On peut penser par exemple aux perspectives offertes par l’optimisation des stratégies de charge d’une flotte de véhicules autonomes, ou plus simplement au pilotage national de la charge via des signaux tarifaires pour participer à l’équilibre offre-demande. Dans tous ces nouveaux problèmes, les algorithmes que nous avons déjà développés devront être adaptés à l’augmentation de la complexité des données et des modèles. Dans le contexte de la modélisation de la production renouvelable, plusieurs enjeux semblent importants, comme l’intégration d’une modélisation plus fine des centrales installées (par la modélisation dans le temps des puissances installées, l’intégration continue des hauteurs de mat, tailles de pales et puissances de machines, etc.) ou la prise en compte des impacts du changement climatique.

Sur le plan de la modélisation de l’acceptabilité, il est nécessaire d’explorer la nature pluri-disciplinaire de ce problème en collaborant des psychologues, ainsi que des sociologues. Cependant, une piste importante à plus court terme est de travailler sur l’impact sur le champ visuel ou auditif de l’éolien de manière quantitative et en comparaison avec les autres infrastructures existantes (routes, bâtiments, etc.). En marge de ces études, je pense que la sensibilité de l’analyse de sensibilité de type Sobol aux hypothèses de travail devrait être plus approfondie tant sur le plan pratique, avec des applications particulières, que sur le plan théorique.

L’analyse des impacts économiques et environnementaux est loin d’être un sujet clos vis-à-vis du système électrique. Les systèmes étudiés jusqu’à maintenant dans la majorité des ACV étaient des systèmes de production simples, et peu de travaux concernent l’impact environnemental du système électrique dans son ensemble (en tenant compte donc des besoins de flexibilité) et l’étude des paramètres importants. L’évaluation des ressources matérielles nécessaires à la transition énergétique est un sujet qui fait couler beaucoup d’encre aujourd’hui dans la presse grand public, et sur lequel les experts sont partagés, la mise en place d’un méta-modèle ici serait pertinente. Dans une perspective de forte pénétration renouvelable, la biomasse peut jouer un rôle important permettant la génération de chaleur, de gaz vert, de biocarburant. Mais l’analyse environnementale de tous ces usages concurrents de la biomasse en tenant compte de ce qu’elle apporte à chaque

système énergétique, et de l'utilisation de la surface au sol disponible qui se fait parfois au détriment d'autres surfaces (agricoles ou forestières) est un enjeu important sur lequel j'aimerais me pencher.

On peut évidemment déceler dans cet ensemble de projets quelques idées communes comme la décomposition multi-échelle qui, bien que n'étant pas une idée neuve (qui était d'ailleurs un outil important dans ma thèse), n'en reste pas moins un concept puissant qui pourrait être fécond dans à peu près tous les domaines mentionnés. Finalement, ce sont aussi les interactions entre ces différentes thématiques qui seront dans les années à venir une source intéressante de recherches. Les domaines sur lesquels j'ai bâti mon expertise ne s'intègrent pas dans un champ de recherche homogène et si l'on se contente de les lister ils peuvent paraître éloignés les uns des autres : OPF dans le réseau de distribution, analyse de la variabilité dans l'ACV, estimation de gisement à partir de systèmes d'information géographique, optimisation des rénovations d'un parc de bâtiments, planification des flexibilités d'un système électrique européen fortement renouvelable, simulation de la courbe de charge électrique locale, prédiction à court terme. Dans le même temps, il est évident que les interactions entre ces domaines offrent des perspectives riches et nombreuses qui font un des intérêts de mon projet de recherche.

4 Contrats de recherche et rayonnement scientifique

4.1 Liste des contrats de recherche et description de ma participation

- [Proj1] MOSAIC. *MOdélisation Simulation et Analyse des Flexibilités de Consommation*. Coordinateur des différents projets de 2013 à maintenant, écriture de la totalité des contrats. ENEDIS. 2013-2020.
- [Proj2] WINPOWER. *Intégration d'énergie éolienne par un réseau de courant continu à haute tension*. Responsable scientifique de la contribution ARMINES, WP leader. Participation au montage. ANR. 2011-2015.
- [Proj3] ANEMOS.PLUS. *Advanced Tools for the Management of Electricity Grids with Large-Scale Wind Generation*. Participation technique. EU-FP6. 2007-2011.
- [Proj4] NICEGRID. *Un quartier solaire intelligent*. Responsable de la contribution de recherche scientifique sur la prédiction de la production PV. Contribution à la description technique de notre travail lors du montage du projet. ADEME (AMI). 2012-2015.
- [Proj5] Vers un mix électrique 100 POURCENT RENOUEVELABLE. Montage du projet en partenariat avec Artelys qui a ensuite coordonné le projet. Responsable de la contribution de ARMINES. Etude réalisée pour l'ADEME. 2013-2015.
- [Proj6] SHAPE. *Simulations Physiques pour l'Aide à la Planification Energétique*. Responsable scientifique et coordinateur du projet. Participation au montage. ADEME. 2017-2020.
- [Proj7] SMARTRESERVE. *Evaluation de l'évolution de la prédictibilité liée à une intégration massive du photovoltaïque et de l'éolien au niveau national*. Responsable scientifique et coordinateur du projet. Coordination du montage. ADEME. 2013-2015.
- [Proj8] Carnot 2014. *Analyse et exploitation de grandes bases de données pour la reconstitution de la demande en énergie*. Responsable scientifique et coordinateur du projet. Coordination du montage. Carnot. 2014-2015.

- [Proj9] SAFEWIND. *Multi-scale data assimilation, advanced wind modeling and forecasting with emphasis to extreme weather situations for a secure large-scale wind power integration*. Project manager, responsable scientifique de la contribution ARMINES. Projet déjà monté avant mon arrivée aux mines. EU - FP7. 2008-2012.
- [Proj10] SMARTCITY. *développement d'une approche globale d'intégration des systèmes de stockage d'énergie dans l'espace urbain*. Responsable scientifique. Participation au montage. Carnot. 2015-2016.
- [Proj11] Micro STEP de BERRIEN. *Micro-STEP installée sur une carrière en fin d'exploitation et alimentée par des énergies renouvelables*. Contribution au montage du projet sous la direction de l'entreprise UNITE. Responsable scientifique de la recherche ARMINES, WP leader. ADEME (AMI). 2014-2018.
- [Proj12] SENSIBLE. *Storage-Enabled Sustainable Energy for Buildings and Communities*. Contribution technique (prédiction de production et planification stockage), participation au montage. EU- H2020. 2014-2018.
- [Proj13] CORUSCANT. *Amélioration de la prévision de production photovoltaïque et simulation des revenus d'un producteur renouvelable sur le marché de l'électricité. Projet en partenariat direct avec l'industriel Coruscant*. Coordinateur du projet. Coordination du montage du projet sous la direction. Coruscant. 2014-2019.
- [Proj14] ENSEOLE. *Prédiction de la production éolienne*. Contribution à la production scientifique ARMINES, organisation de réunions. Projet déjà monté avant mon arrivée aux mines. ADEME. 2008-2009.
- [Proj15] More MICROGRIDS. Responsable de la contribution de ARMINES sur la prédiction de la production PV, (projet déjà monté avant mon arrivée aux mines). EU-FP5. 2008-2009.
- [Proj16] RESTABLE. *Improvement of Renewables-based System Services*. Contribution scientifique, WP leader. Contribution à la description technique de notre travail lors du montage du projet. EU - ERA-Net Smart Grids. 2016-2019.

4.2 Rayonnement scientifique

- [Ray1] Reviewing pour de nombreux journaux internationaux de rang A : Wind Energy, IEEE Transactions on Power Systems, IEEE Signal Processing Magazine, International Journal of Forecasting, Quarterly Journal of the Royal Meteorological Society, Journal of Renewable and Sustainable Energy, Journal of Machine Learning Research, Applied Energy.
- [Ray2] Conférences grand public. Essentiellement dans le cadre de la transition énergétique. Participation au réseau de conférences citoyennes dans les Alpes Maritimes (Biot 2018, Villeneuve Loubet 2019), Journée Alternatibat 2018 à Biot, Pint of Science 2018 à Nice.
- [Ray3] Intervention à l'Institut des Hautes Etudes pour la Science et la Technologie (IHEST). Atelier "Autoconsommation d'électricité solaire PV". Novembre 2017.
- [Ray4] Rapporteur de la thèse de Amanda Lenzi. Danmarks Tekniske Universitet, DTU. Juin 2017.
- [Ray5] Membre du comité scientifique de l'OFATE et Conférence. Membre du comité scientifique de l'Office Franco-Allemand pour la Transition énergétique (OFATE). Participation à une table ronde à Berlin sur le couplage des différents secteurs énergétiques. Novembre 2017.
- [Ray6] Participation à la tâche 25 de l'Agence Internationale de l'Energie. Tâche sur l'intégration de l'éolien au système électrique. Depuis 2015.
- [Ray7] Analyse de sensibilité pour l'ACV. Cours Doctoral Excellence (CDE1) organisé dans le cadre du réseau EcoSD. Cours de 3 heures et participation à l'organisation des TP. 2013-2015.
- [Ray8] Mise en place d'un blog sur la transition énergétique. Juin 2019 <https://www.energy-alternatives.eu/>.
- [Ray9] Création d'un package R pour la programmation dynamique ConConPiWiFun (Optimisation with Continuous Convex Piecewise (Linear and Quadratic) Functions). 2015.

5 Enseignement et encadrement

5.1 Thèses encadrées et publications associées

- [These1] Seddik Yassine ABDELOUADOUD. “Intégration des énergies renouvelables au réseau de distribution d’électricité”. encadrement 70%, directeur François Neirac. Soutenue en 2014, maintenant au CSTB. Article : [ArticleP11], conférences : [Confs13, Confs24, Confs14]. Thèse de doct. MINES ParisTech.
- [These2] Arthur BOSSAVY. “Caractérisation et prédiction probabiliste des variations brusques et importantes de la production éolienne”. encadrement 70%, directeur Didier Mayer. Soutenue en 2012. Maintenant à Artelys. Publications : Articles [ArticleP10, ArticleP6] puis en post-doc sous ma direction [ArticleP12]. Conférences (6 pendant la thèse dont la moitié comme premier auteur et les autres en post-doc) : [Confs1, Confs3, Confs4, Confs15, Confs19, Confs6, Confs17, Confs32, Confs16, Confs35, Confs36, Confs33, Confs21, Confs20]. Thèse de doct. MINES ParisTech.
- [These3] Grover Silva ETTA. “Optimisation de l’intégration de la production renouvelable dans un système électrique intelligent”. encadrement 70%, directeur Georges Kariniotakis. Soutenue en 2017, maintenant à écoCO2. Articles : [ArticleP14, ArticleP19], conférences : [Confs38] et [Confs41]. Thèse de doct. MINES ParisTech.
- [These4] Fiona FOUCAULT. “Optimisation de l’implantation de centrales éoliennes dans l’environnement d’un marché à prix locaux”. encadrement 70%, directeur Georges Kariniotakis. Soutenue en dec 2016. Maintenant à la mairie de Paris. Conférences (3 comme premier auteur) : [Confs26, Confs27, Confs18, Confs29]. Thèse de doct. MINES ParisTech.
- [These5] Alexis GÉROSIER. “Prédiction à court terme de la consommation électrique d’un quartier”. encadrement 70%, directeur Georges Kariniotakis. Soutenue en 2019, maintenant chez Raiden. Articles [ArticleP13, ArticleP20, ArticleP22], conférences (2 comme premier auteur) : [Confs40, Confs46, Confs39]. Thèse de doct. MINES ParisTech.

- [These6] Bellinger KÉVIN. “Prévision à court-terme de la production des centrales virtuelles à base d’énergie renouvelables météo-dépendantes.” encadrement 70%, directeur Georges Kariniotakis. Soutenance prévue en 2021. Conférence : [Confs48]. Thèse de doct. MINES ParisTech, thèse CIFRE avec CNR.
- [These7] Antoine ROGEAU. “Méthodes d’aide à la décision pour la mise en oeuvre locale de la transition énergétiques”. encadrement 70%, directeur Georges Kariniotakis. Soutenance prévue en 2019. Article (en stage) : [ArticleP16], un article soumis. Conférences [Confs45, Confs47, Confs50]. Thèse de doct. MINES ParisTech.
- [These8] Barbier THIBAUT. “Modélisation de la consommation électrique à partir de grandes masses de données pour la simulation des alternatives énergétiques du futur”. encadrement 70%, directeur François Neirac. Soutenue en 2017, maintenant à MINES ParisTech et Optimum Tracker. Article : [ArticleP13], conférences (2 comme premier auteur) : [Confs45, Confs42, Confs31, Confs25, Confs34]. Thèse de doct. MINES ParisTech.
- [These9] Heggarty THOMAS. “Optimisation technico-économique du mix de flexibilités”. encadrement 70%, directeur Georges Kariniotakis. Soutenance prévue en 2020. Article : [ArticleP23]. Thèse de doct. MINES ParisTech, thèse CIFRE avec RTE.
- [These10] Mahler VALENTIN. “Modélisation de l’évolution des prix sur le marché de l’électricité dans un contexte de forte pénétration des énergies renouvelables”. encadrement 70%, directeur Georges Kariniotakis. Soutenance prévue 2021. Conférence : [Confs49]. Thèse de doct. MINES ParisTech.
- [These11] Agoua XWÉGNON GHISLAIN. “Méthodes spatio-temporelles pour la prédiction à court terme de la production photovoltaïque”. encadrement 70% directeur Georges Kariniotakis. Soutenue en 2017, maintenant à EDF R&D. Articles : [ArticleP18, ArticleP17], conférences (1 comme premier auteur) : [Confs30, Confs41]. Thèse de doct. MINES ParisTech.

5.2 Stages : encadrements et tutorats

- [Stages1] TUTORAT. Master Eurec. Suivi de 4 étudiants. Depuis 2009.
- [Stages2] TUTORAT. Master ALEF. Suivi de 12 étudiants. depuis 2010.

- [Stages3] TUTORAT. Option Machine Energie de MINES ParisTech. Suivi de 16 étudiants. depuis 2012.
- [Stages4] Moussa ALFRED (2009). Encadrement. Prévion de la production de centrales PV et de l'incertitude associée. Université Paul Sabatier.
- [Stages5] Arthur BOSSAVY (2009). Encadrement. Prévion de l'incertitude de la production d'une centrale éolienne dans le cas de rampes. Université Paul Sabatier.
- [Stages6] Pauline SURGOT (2010). Encadrement. Classification automatique de situations météorologiques pour la prédiction de la production des centrales renouvelables. Université de Nantes.
- [Stages7] Kevin LAQUAINE (2011). Encadrement. Intégration de la prédictibilité dans l'estimation de la ressource éolienne. Université de Rennes.
- [Stages8] Graziella MORANDI (2011). Encadrement. Détection et classification automatiques de valeurs aberrantes et particulières de production de fermes éoliennes et de centrales PV. ENSAI.
- [Stages9] Paul PUGET (2011). Encadrement. Prévion de la production PV à très court terme. ENSMSE.
- [Stages10] Thibaut BARBIER (2012). Encadrement. Modélisation dynamique des usages résidentiels de l'électricité ENS Rennes.
- [Stages11] Vincent BARBESANT (2013). Encadrement. Programmation dynamique appliquée à la gestion d'un stockage. Ecole Polytechnique.
- [Stages12] Liakopoulou KATERINA (2013). Encadrement. Simulation d'un marché à prix locaux. TU Delft.
- [Stages13] Akter SHOUROV (2013). Encadrement. Evaluation des besoins de réserve vis à vis de l'intégration d'énergie PV. Master REST - Polytechnique.
- [Stages14] Thibaut BARBIER (2014). Encadrement. Simulation de la courbe de charge électrique à partir de méthode *bottom up*. ENS Rennes.
- [Stages15] Semlali HICHAM (2014). Encadrement. Methodes de validation de scénarios de consommation électrique pour la modélisation prospective des smart grids. ENSMSE.

- [Stages16] Chen LE (2014). co-Encadrement avec Etta Grover-Silva. Optimisation stochastique pour la gestion de stockage dans le réseau de distribution. Master ICARE.
- [Stages17] Yuan ZHAO (2014). Encadrement. Implémentation C++ d'une méthode de simulation *bottom up* de la consommation électrique. Master ICARE.
- [Stages18] Rogeau ANTOINE (2015). Encadrement. Estimation du potentiel de centrales Micro-STEP en France. Mines de Nantes.
- [Stages19] Parrain TRISTAN (2016). Encadrement. Programmation dynamique pour l'opération du système électrique. Université du Mans.
- [Stages20] Billeau SÉBASTIEN (2017). Encadrement. Modélisation des prix du marché *spot* de lélectricité. MINES ParisTech.

5.3 Enseignements

- [Ens1] MIG ALEF. (Métier de l'Ingénieur Généraliste, Alternatives Energétiques du Futur). Encadrement de 12 étudiants de première année du cycle ingénieur civil pendant 3 semaines à plein temps. Organisation des visites, des cours, du travail et de la présentation des résultats. Création d'un nouveau sujet chaque année sur les alternatives énergétiques du futur en partenariat avec un industriel. Exemples de sujets traités : Energie des mers, micro-STEP, thermo-sensibilité en PACA, power to gaz, fin du tarif d'achat pour les EnR, dimensionner le réseau de Paris-Saclay, les smart-grid dans la région de Nice, ... Depuis 2011.
- [Ens] UE14. Unité d'enseignement en première année à MINES ParisTech depuis septembre 2019. Participation à la coordination, mise en place des 27 mini-projets et de la session poster associée.
- [Ens2] PESTO Energie. Module de formation de 6 semaines plein temps à destination du corps des mines sur l'énergie. Participation depuis 2012 (cours, proposition et suivi de mini-projets, visites), puis depuis 2015 coordination et participation.
- [Ens3] TP Analyse de données de production PV. Post Master EU-REC énergies renouvelables. Volume : 12 heures par an. 2010-2015.

- [Ens4] TP prédiction à court terme de production PV. Ecole des Mines de Saint Etienne. Volume : 12 heures par an. 2010-2013.
- [Ens5] Matlab/Simulink. Introduction à Matlab et SIMULINK (TP), Post Master EUREC énergies renouvelables. Volume : 12 heures par an. 2008-2012.
- [Ens6] Refonte du cycle IC. Participation à la refonte du cycle ingénieur civil (IC) des mines. 2018-2019.
- [Ens7] Cours Master Eurec. Intégration des EnR électriques au système. Volume : 12 heures par an. 2012-2016.
- [Ens8] Cours Master PSL. Intégration des EnR électriques au système. Volume : 12 heures par an. 2012-2016.
- [Ens9] Cours Master REST. Système électrique, météorologie pour l'éolien, prédiction à court terme. Volume : 12 heures par an. 2012-2016.
- [Ens10] TP Analyse de données pour l'énergie. Master Eurec. Volume : 16 heures par an. 2009-2014.
- [Ens11] Evolutions du système électrique dans le contexte de la transition énergétique. MINES ParisTECH. Volume : 36 heures par an. 2018-2019.

6 Bibliographie personnelle

6.1 Articles publiés dans des revues scientifiques à comité de lecture

- [ArticleP1] Antoine ROGEAU et al. "A Joint Optimization of Building-envelope and Heating-system Retrofits at Territory-scale for Enhanced Decision Aiding". *Applied Energy* (Accepted January 2020).
- [ArticleP2] Robin GIRARD. "Segmentation d'images hyper-spectrales". *Traitement du signal* 23 (2006), p. 277–288.
- [ArticleP3] Robin GIRARD. "Plugin procedure in segmentation and application to hyperspectral image segmentation". *Electronic journal of statistics* 4 (2010), p. 655–676.

- [ArticleP4] Robin GIRARD. “Fast rate of convergence in high-dimensional linear discriminant analysis”. *Journal of Nonparametric Statistics* 23.1 (2011), p. 165–183.
- [ArticleP5] Pierre PINSON et Robin GIRARD. “Evaluating the quality of scenarios of short-term wind power generation”. *Applied Energy* 96 (2012), p. 12–20.
- [ArticleP6] Arthur BOSSAVY, Robin GIRARD et Georges KARINIOTAKIS. “Forecasting ramps of wind power production with numerical weather prediction ensembles”. *Wind Energy* 16.1 (2013), p. 51–63.
- [ArticleP7] Robin GIRARD et Denis ALLARD. “Spatio-temporal propagation of wind power prediction errors”. *Wind Energy* 16.7 (2013), p. 999–1012.
- [ArticleP8] Robin GIRARD, K. LAQUAINE et Georges KARINIOTAKIS. “Assessment of wind power predictability as a decision factor in the investment phase of wind farms”. *Applied Energy* 101 (2013), p. 609–617.
- [ArticleP9] Pierryves PADEY et al. “From LCAs to Simplified Models : A Generic Methodology Applied to Wind Power Electricity”. *Environmental Science and Technology* 47.3 (2013), p. 1231–1238.
- [ArticleP10] Arthur BOSSAVY, Robin GIRARD et Georges KARINIOTAKIS. “An edge model for the evaluation of wind power ramps characterization approaches”. *Wind Energy* 18.7 (2014), p. 1169–1184.
- [ArticleP11] Seddik ABDELOUADOUD et al. “Optimal power flow of a distribution system based on increasingly tight cutting planes added to a second order cone relaxation”. *International Journal of Electrical Power and Energy Systems* 69 (2015), p. 9–17.
- [ArticleP12] Arthur BOSSAVY, Robin GIRARD et Georges KARINIOTAKIS. “Sensitivity analysis in the technical potential assessment of onshore wind and ground solar photovoltaic power resources at regional scale”. *Applied Energy* 182 (2016), p. 145–153.

- [ArticleP13] Alexis GEROSSIER, Thibaut BARBIER et Robin GIRARD. “A novel method for decomposing electricity feeder load into elementary profiles from customer information”. *Applied Energy* 203 (2017), p. 752–760.
- [ArticleP14] Etta GROVER-SILVA, Robin GIRARD et Georges KARINIOTAKIS. “Optimal sizing and placement of distribution grid connected battery systems through an SOCP optimal power flow algorithm”. *Applied Energy* (2017).
- [ArticleP15] Martino LACIRIGNOLA et al. “LCA of emerging technologies : addressing high uncertainty on inputs’ variability when performing global sensitivity analysis”. *Science of the Total Environment* 578 (2017), p. 268–280.
- [ArticleP16] Antoine ROGEAU, Robin GIRARD et Georges KARINIOTAKIS. “A generic GIS-based method for small Pumped Hydro Energy Storage (PHES) potential evaluation at large scale”. *Applied Energy* 197 (2017), p. 241–253. ISSN : 0306-2619.
- [ArticleP17] Xwégnon Ghislain AGOUA, Robin GIRARD et Georges KARINIOTAKIS. “Probabilistic Model for Spatio-Temporal Photovoltaic Power Forecasting”. *IEEE Transactions on Sustainable Energy* (2018), p. 11. ISSN : 1949-3029.
- [ArticleP18] Xwégnon Ghislain AGOUA, Robin GIRARD et Georges KARINIOTAKIS. “Short-Term Spatio-Temporal Forecasting of Photovoltaic Power Production”. *IEEE Transactions on Sustainable Energy* 9.2 (2018), p. 538–546.
- [ArticleP19] Grover-Silva ETTA et al. “A stochastic optimal power flow for scheduling flexible resources in microgrids operation”. *Applied Energy* 229 (2018), p. 201–208. ISSN : 0306-2619.
- [ArticleP20] Alexis GEROSSIER et al. “Robust Day-Ahead Forecasting of Household Electricity Demand and Operational Challenges”. *Energies* 11.12 (2018), p. 3503.
- [ArticleP21] Andrea MICHIORRI et al. “Storage sizing for grid connected hybrid wind and storage power plants taking into account forecast errors autocorrelation”. *Renewable Energy* 117 (2018), p. 380–392.
- [ArticleP22] Alexis GEROSSIER, Robin GIRARD et Georges KARINIOTAKIS. “Modeling and Forecasting Electric Vehicle Consumption Profiles”. *Energies* 12.7, 1341 (2019), p. 1–14.

- [ArticleP23] Thomas HEGGARTY et al. “Multi-temporal assessment of power system flexibility requirement”. *Applied Energy* 238 (2019), p. 1327–1336.

6.2 Articles soumis dans des revues à comité de lecture

- [ArticleS1] Robin GIRARD et al. “Measures of market incentives toward predictability in the investment phase of wind power production”. *Applied Energy* (Submitted October 2019).

6.3 Conférences

- [Confs1] Arthur BOSSAVY, Robin GIRARD et Georges KARINIOTAKIS. “Forecasting Uncertainty Related to Ramps of Wind Power Production”. *Proceedings of European Wind Energy Conference and Exhibition, EWEC*. 2010.
- [Confs2] Bella ESPINAR et al. “Photovoltaic Forecasting : A state of the art”. *5th European PV-Hybrid and Mini-Grid Conference*. 2010, p. 250–255.
- [Confs3] Arthur BOSSAVY, Robin GIRARD et Georges KARINIOTAKIS. “A Probabilistic Approach to Forecast Ramps of Wind Power Production using Ensemble”. *11th EMS Annual Meeting - 11th European Conference on Applications of Meteorology (ECAM)*. 2011.
- [Confs4] Arthur BOSSAVY, Robin GIRARD et Georges KARINIOTAKIS. “Forecasting Ramps of Wind Power Production”. *IEEE 16th ISAP Conference : Intelligent Systems Application to Power Systems*. 2011.
- [Confs5] Gilney DAMM et Robin GIRARD. “WINPOWER Project - Wind energy integration by DC network”. *EWEA OFFSHORE - European Wind Energy Association annual event*. 2011.
- [Confs6] Robin GIRARD, Arthur BOSSAVY et Georges KARINIOTAKIS. “Advanced ramp forecasting for different temporal scales”. *EWEA - European Wind Energy Association annual event*. 2011.
- [Confs7] Robin GIRARD et Georges KARINIOTAKIS. “Analysing spatio-temporal patterns of wind power production for off shore wind power scenario generation”. *EWEA OFFSHORE - European Wind Energy Association annual event*. 2011.

- [Confs8] Georges KARINIOTAKIS et Robin GIRARD. “Towards Improved Wind Power Forecasting Technology with Focus on Extremes. The SafeWind Project”. *Annual EWEA 2011 Conference*. 2011.
- [Confs9] Pierre PINSON et al. “Benefits from the increased contribution of the European weather forecasting community to wind power prediction”. *EWEA - European Wind Energy Association annual event*. 2011.
- [Confs10] Georges KARINIOTAKIS et al. “Advances In Short-Term Wind Power Forecasting With Focus on 'Extreme' Situations - SafeWind”. *EWEA - European Wind Energy Association annual event*. 2012.
- [Confs11] Andrea MICHIORRI et al. “A local energy management system for solar integration and improved security of supply : The Nice Grid project”. *3rd IEEE PES International Conference and Exhibition on Innovative Smart Grid Technologies (ISGT Europe)*. 2012.
- [Confs12] Pierryves PADEY et al. “Understanding LCA results variability : developing global sensitivity analysis with Sobol indices. A first application to photovoltaic systems”. *International Symposium on Life Cycle assessment and Construction Civil engineering and buildings*. 2012.
- [Confs13] Seddik Yassine ABDELOUADOUD et al. “Planning-oriented yearly simulation of energy storage operation in distribution system for profit maximization, voltage regulation and reserve provisioning”. *22nd International Conference and Exhibition on Electricity Distribution - CIRED*. 2013.
- [Confs14] Yassine Seddik ABDELOUADOUD et al. “A criticality criterion to decrease the computational burden in multistage distribution system optimal power flow”. *PowerTech*. 2013.
- [Confs15] Arthur BOSSAVY, Robin GIRARD et Georges KARINIOTAKIS. “A Novel Methodology for comparison of different wind power ramp characterization approaches”. *EWEA - European Wind Energy Association annual event*. 2013.

- [Confs16] Arthur BOSSAVY et al. “A model to make probabilistic forecasts of PV production from a plan-oriented parameterization of solar irradiance”. *2nd International Conference Energy & Meteorology - Weather & Climate for the Energy Industry, ICEM*. 2013.
- [Confs17] Arthur BOSSAVY et al. “The impact of available data history on the performance of photovoltaic generation forecasting models”. *22nd International Conference on Electricity Distribution - CIRED*. 2013.
- [Confs18] Fiona FOUCAULT, Robin GIRARD et Georges KARINIOTAKIS. “The impact of electricity market schemes on predictability being a decision factor in the wind farm investment phase”. *EWEA - European Wind Energy Association annual event*. Austria, fév. 2013, p. 1–8.
- [Confs19] Robin GIRARD, Arthur BOSSAVY et Georges KARINIOTAKIS. “Advanced spatio-temporal wind power forecasting with distributed wind farms as sensors”. *EWEA - European Wind Energy Association annual event*. 2013.
- [Confs20] Andrea MICHIORRI et al. “Impact of PV forecasts uncertainty on batteries management in microgrids”. *PowerTech - Towards carbon free society through smarter grids*. 2013.
- [Confs21] Andrea MICHIORRI et al. “The value of schedule update frequency on distributed energy storage performance in renewable energy integration”. *The 4th IEEE/PES European Innovative Smart Grid Technologies (ISGT), International Conference and Exhibition*. 2013.
- [Confs22] Pierryves PADEY et al. “Global Sensitivity Analysis : a tool to analyse LCA variability of energy systems”. *SETAC Europe 23rd Annual Meeting*. 2013.
- [Confs23] Pierre PINSON, Robin GIRARD et Patrick MCSHARRY. “Stochastic power generation”. *DTU International Energy Report : Energy storage options for future sustainable energy systems*. Sous la dir. de Leif Sonderberg Petersen HANS HVIDTFELDT LARSEN. Technical University of Denmark, 2013.

- [Confs24] Seddik Yassine ABDELOUADOUD et al. “Iterative linear cuts strenghtening the second-order cone relaxation of the distribution system optimal power flow problem”. *IEEE PES T&D Conference*. 2014.
- [Confs25] Thibaut BARBIER et al. “A novel approach for electric load curve holistic modelling and simulation”. *Proceedings of Med-Power - The 9th Mediterranean Conference on Power Generation, Transmission Distribution and Energy Conversion*. 2014.
- [Confs26] Fiona FOUCAULT, Robin GIRARD et Georges KARINIOTAKIS. “A robust investment strategy for generation capacity in an uncertain demand and renewable penetration environment”. *11th International Conference on the European Energy Market (EEM)*. 2014.
- [Confs27] Fiona FOUCAULT, Robin GIRARD et Georges KARINIOTAKIS. “Wind farm strategic investment considering forecast errors penalties in a nodal prices market”. *EWEA - European Wind Energy Association annual event*. 2014.
- [Confs28] Robin GIRARD. “Vers un scénario énergétique 100% EnR (Energies renouvelables) pour la France”. *Journée Futuring Cities – Ville et Energie Durable, Institut Mines Telecom*. 2014.
- [Confs29] Robin GIRARD et al. “Fast dynamic programming with application to storage planning”. *IEEE PES T&D Conference and Exposition*. 2014.
- [Confs30] Xwégnon Ghislain AGOUA, Robin GIRARD et Georges KARINIOTAKIS. “Spatio-temporal models for photovoltaic power short-term forecasting”. *Solar Integration workshop*. 2015.
- [Confs31] Aurélie FERRAGE et al. “Using simulated predictive load curves to improve DSO’s network development planning methods integrating smart grids functionalities”. *23rd International Conference on Electricity Distribution - CIRED 2015*. 2015.
- [Confs32] Robin GIRARD et al. “Evaluation of the level of prediction errors and sub-hourly variability of PV and wind generation in a future with a large amount of renewables”. *23rd International Conference on Electricity Distribution - CIRED*. 2015.

- [Confs33] Robin GIRARD et al. “Evaluation of the level of prediction errors of wind and PV in a future with a large amount of renewables”. *EWEA Technology Workshop, Wind Power Forecasting*. 2015.
- [Confs34] Thibaut BARBIER et al. “An optimisation algorithm for matching large scale databases on customers for improved characterisation of electricity consumption”. *Proceedings of Med-Power - The 10th Mediterranean Conference on Power Generation, Transmission, Distribution and Energy Conversion*. 2016.
- [Confs35] Alexis BOCQUET et al. “Assessment of probabilistic PV production forecasts performance in an operational context”. *Proceedings of the 6th Solar Integration Workshop - International Workshop on Integration of Solar Power into Power Systems*. 2016.
- [Confs36] Anne-Laure DUBILLY et al. *A 100% renewable electricity mix Analyses and optimisations Testing the boundaries of renewable energy-based electricity development in metropolitan France by 2050*. Rapp. tech. Ademe, 2016.
- [Confs37] Robin GIRARD. “Deux contributions pour l’optimisation d’un système électrique fortement renouvelable”. *Journées MAS 2016 - Phénomènes complexes et Hétérogènes*. 2016.
- [Confs38] Etta GROVER-SILVA, Robin GIRARD et Georges KARINIOTAKIS. “Multi-temporal Optimal Power Flow for Assessing the Renewable Generation Hosting Capacity of an Active Distribution System”. *IEEE/PES Transmission and Distribution Conference and Exposition (T&D)*. 2016.
- [Confs39] Carlos Adrian CORREA FLOREZ et al. “Residential Electrical and Thermal Storage Optimisation in a Market Environment”. *CIREN 2017 - 24th International Conference on Electricity Distribution*. Glasgow, United Kingdom, juin 2017.
- [Confs40] Alexis GEROSIER et al. “Probabilistic Day-Ahead Forecasting of Household Electricity Demand”. *CIREN 2017 - 24th International Conference on Electricity Distribution*. Glasgow, United Kingdom, juin 2017.

- [Confs41] Etta GROVER-SILVA et al. “A stochastic multi-temporal optimal power flow approach for the management of grid connected storage”. *CIREC (Congres International des Réseaux Electriques de Distribution)*. 2017.
- [Confs42] Nicolas KONG et al. “Long Term Forecast of Local Electrical Demand and Evaluation of Future Impacts on the Electricity Distribution Network”. *CIREC 2017*. Glasgow, United Kingdom, juin 2017.
- [Confs43] Paula PEREZ-LOPEZ et al. “Global sensitivity analysis in LCA of emerging technologies : Accounting for inputs variability”. *SETAC Europe 2017 - 27th Annual Meeting*. SETAC. Brussels, Belgium, mai 2017.
- [Confs44] Nicolas RETIÈRE et al. “Fractal Grid – Towards the Future Smart Grid”. *CIREC 2017 - 24th International Conference on Electricity Distribution*. Glasgow, United Kingdom, juin 2017, p. 1236.
- [Confs45] Antoine ROGEAU et al. “Evolution of electrical distribution grid sizing considering self-consumption of local renewable production”. *CIREC 2017 - 24th International Conference on Electricity Distribution*. Glasgow, United Kingdom, juin 2017.
- [Confs46] Alexis GEROSIER, Robin GIRARD et George KARINIOTAKIS. “Modeling Electric Vehicle Consumption Profiles for Short-Term Forecasting and Long-Term Simulation”. *11th Mediterranean Conference on Power Generation, Transmission, Distribution and Energy Conversion, MEDPOWER 2018*. Dubrovnik (Cavtat), Croatia, nov. 2018.
- [Confs47] Antoine ROGEAU et al. “An address matching algorithm for household-scale databases for enhancing electricity demand characterisation”. *Proceedings of MedPower - The 12th Mediterranean Conference on Power Generation, Transmission, Distribution and Energy Conversion*. 2018.
- [Confs48] Kevin BELLINGUER et al. “Short-term PV generation forecasting using multiple heterogeneous sources of data”. *PVSEC 2019 - European PV Solar Energy Conference*. 2019.

- [Confs49] Valentin MAHLER et al. “Simulation of day-ahead electricity market prices using a statistically calibrated structural model”. *16th International Conference on the European Energy Market (EEM)*. Ljubljana, Slovenia, sept. 2019.
- [Confs50] Antoine ROGEAU et al. “Retrofitting a building stock : modeling and optimization for decision aiding at territory scale”. *Journal of Physics : Conference Series*. Lausanne, Switzerland : IOP Publishing, 2019.

7 Bibliographie

- [ext1] J. CANNY. “A computational approach to edge detection”. *Pattern Analysis and Machine Intelligence, IEEE Transactions on PAMI-8.6* (1986), p. 679–698.
- [ext2] Toshimitsu HOMMA et Andrea SALTELLI. “Importance measures in global sensitivity analysis of nonlinear models”. *Reliability Engineering & System Safety* 52.1 (1996), p. 1–17. ISSN : 0951-8320.
- [ext3] T. LINDBERG. “Edge detection and ridge detection with automatic scale selection”. *International Journal of Computer Vision* 30(2) (1998), p. 117–154.
- [ext4] I.M SOBOL. “Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates”. *Mathematics and Computers in Simulation* 55 (2001). The Second {IMACS} Seminar on Monte Carlo Methods, p. 271–280. ISSN : 0378-4754.
- [ext5] P. BAO, L. ZHANG et X. WU. “Canny edge detection enhancement by scale multiplication”. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 27.9 (2005), p. 1485–1490.
- [ext6] N. J. CUTLER et al. “Detecting, categorizing and forecasting large ramps in wind farm power output using meteorological observations and WPPT”. *Wind Energy* 10.5 (2007), p. 453–470.
- [ext7] P. PINSON et al. “From probabilistic forecasts to statistical scenarios of short-term wind power production”. *Wind Energy* 12(1) (2009), p. 51–62.

- [ext8] Subhonmesh BOSE et al. “Optimal power flow over tree networks”. *2011 49th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. 2011, p. 1342–1348.
- [ext9] Baosen ZHANG et D. TSE. “Geometry of feasible injection region of power networks”. *2011 49th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. 2011, p. 1508–1515.
- [ext10] M. FARIVAR et al. “Optimal inverter VAR control in distribution systems with high PV penetration”. *IEEE Power and Energy Society General Meeting*. 2012, p. 1–7.
- [ext11] Na LI, Lijun CHEN et S.H. LOW. “Exact convex relaxation of OPF for radial networks using branch flow model”. *Smart Grid Communications (SmartGridComm), 2012 IEEE Third International Conference on*. 2012, p. 7–12.
- [ext12] C. GALLEGRO et al. “A wavelet-based approach for large wind power ramp characterisation”. *Wind Energy* 16.2 (2013), p. 257–278. ISSN : 1099-1824.
- [ext13] Lingwen GAN et al. “Optimal Power Flow in Distribution Networks”. *Proc. 52nd IEEE Conference on Decision and Control*. 2013.
- [ext14] Javad LAVAEI, David TSE et Baosen ZHANG. “Geometry of Power Flows and Optimization in Distribution Networks”. *IEEE Transactions on Power Systems* (2014), p. 572–583.
- [ext15] Welsch MANUEL et al. “Incorporating flexibility requirements into long-term energy system models – A case study on high levels of renewable electricity penetration in Ireland”. *Applied Energy* 135 (2014), p. 600–615. ISSN : 0306-2619.
- [ext16] Poncelet KRIS et al. “Impact of the level of temporal and operational detail in energy-system planning models”. *Applied Energy* 162 (2016), p. 631–643. ISSN : 0306-2619.
- [ext17] Hans-Kristian RINGKJØB, Peter M. HAUGAN et Ida Marie SOLBREKKE. “A review of modelling tools for energy and electricity systems with large shares of variable renewables”. *Renewable and Sustainable Energy Reviews* 96 (2018), p. 440–459. ISSN : 1364-0321.

8 HDR Robin Girard - Articles publiés dans des revues scientifiques à comité de lecture

- [ArticleP1] Antoine ROGEAU et al. “A Joint Optimization of Building-envelope and Heating-system Retrofits at Territory-scale for Enhanced Decision Aiding”. *Applied Energy* (Accepted January 2020).
- [ArticleP2] Robin GIRARD. “Segmentation d’images hyper-spectrales”. *Traitement du signal* 23 (2006), p. 277–288.
- [ArticleP3] Robin GIRARD. “Plugin procedure in segmentation and application to hyperspectral image segmentation”. *Electronic journal of statistics* 4 (2010), p. 655–676.
- [ArticleP4] Robin GIRARD. “Fast rate of convergence in high-dimensional linear discriminant analysis”. *Journal of Nonparametric Statistics* 23.1 (2011), p. 165–183.
- [ArticleP5] Pierre PINSON et Robin GIRARD. “Evaluating the quality of scenarios of short-term wind power generation”. *Applied Energy* 96 (2012), p. 12–20.
- [ArticleP6] Arthur BOSSAVY, Robin GIRARD et Georges KARINIOTAKIS. “Forecasting ramps of wind power production with numerical weather prediction ensembles”. *Wind Energy* 16.1 (2013), p. 51–63.
- [ArticleP7] Robin GIRARD et Denis ALLARD. “Spatio-temporal propagation of wind power prediction errors”. *Wind Energy* 16.7 (2013), p. 999–1012.
- [ArticleP8] Robin GIRARD, K. LAQUAINE et Georges KARINIOTAKIS. “Assessment of wind power predictability as a decision factor in the investment phase of wind farms”. *Applied Energy* 101 (2013), p. 609–617.
- [ArticleP9] Pierryves PADEY et al. “From LCAs to Simplified Models : A Generic Methodology Applied to Wind Power Electricity”. *Environmental Science and Technology* 47.3 (2013), p. 1231–1238.
- [ArticleP10] Arthur BOSSAVY, Robin GIRARD et Georges KARINIOTAKIS. “An edge model for the evaluation of wind power ramps characterization approaches”. *Wind Energy* 18.7 (2014), p. 1169–1184.
- [ArticleP11] Seddik ABDELOUADOUD et al. “Optimal power flow of a distribution system based on increasingly tight cutting planes added to a second order cone relaxation”. *International Journal of Electrical Power and Energy Systems* 69 (2015), p. 9–17.
- [ArticleP12] Arthur BOSSAVY, Robin GIRARD et Georges KARINIOTAKIS. “Sensitivity analysis in the technical potential assessment of onshore wind and ground solar photovoltaic power resources at regional scale”. *Applied Energy* 182 (2016), p. 145–153.

- [ArticleP13] Alexis GEROSSIER, Thibaut BARBIER et Robin GIRARD. “A novel method for decomposing electricity feeder load into elementary profiles from customer information”. *Applied Energy* 203 (2017), p. 752–760.
- [ArticleP14] Etta GROVER-SILVA, Robin GIRARD et Georges KARINIOTAKIS. “Optimal sizing and placement of distribution grid connected battery systems through an SOCP optimal power flow algorithm”. *Applied Energy* (2017).
- [ArticleP15] Martino LACIRIGNOLA et al. “LCA of emerging technologies : addressing high uncertainty on inputs’ variability when performing global sensitivity analysis”. *Science of the Total Environment* 578 (2017), p. 268–280.
- [ArticleP16] Antoine ROGEAU, Robin GIRARD et Georges KARINIOTAKIS. “A generic GIS-based method for small Pumped Hydro Energy Storage (PHES) potential evaluation at large scale”. *Applied Energy* 197 (2017), p. 241–253. ISSN : 0306-2619.
- [ArticleP17] Xwégnon Ghislain AGOUA, Robin GIRARD et Georges KARINIOTAKIS. “Probabilistic Model for Spatio-Temporal Photovoltaic Power Forecasting”. *IEEE Transactions on Sustainable Energy* (2018), p. 11. ISSN : 1949-3029.
- [ArticleP18] Xwégnon Ghislain AGOUA, Robin GIRARD et Georges KARINIOTAKIS. “Short-Term Spatio-Temporal Forecasting of Photovoltaic Power Production”. *IEEE Transactions on Sustainable Energy* 9.2 (2018), p. 538–546.
- [ArticleP19] Grover-Silva ETTA et al. “A stochastic optimal power flow for scheduling flexible resources in microgrids operation”. *Applied Energy* 229 (2018), p. 201–208. ISSN : 0306-2619.
- [ArticleP20] Alexis GEROSSIER et al. “Robust Day-Ahead Forecasting of Household Electricity Demand and Operational Challenges”. *Energies* 11.12 (2018), p. 3503.
- [ArticleP21] Andrea MICHIORRI et al. “Storage sizing for grid connected hybrid wind and storage power plants taking into account forecast errors autocorrelation”. *Renewable Energy* 117 (2018), p. 380–392.
- [ArticleP22] Alexis GEROSSIER, Robin GIRARD et Georges KARINIOTAKIS. “Modeling and Forecasting Electric Vehicle Consumption Profiles”. *Energies* 12.7, 1341 (2019), p. 1–14.
- [ArticleP23] Thomas HEGGARTY et al. “Multi-temporal assessment of power system flexibility requirement”. *Applied Energy* 238 (2019), p. 1327–1336.

Robin GIRARD
1 rue Claude Daunesse, MINES ParisTech PERSEE
06560 Sophia Antipolis

(+33)662507809
robin.girard@mines-paristech.fr

Né le 12 Juillet 1980
Nationalité française
Permis B, véhicule

Curriculum Vitae

Situation administrative

- Aout 2012** – CHARGÉ DE RECHERCHE , Centre PERSEE, Groupe Energies renouvelables et systèmes électriques intelligents, [MINES ParisTech](#), Sophia Antipolis, France.
Présent
- Sept 2008** – ASSISTANT DE RECHERCHE , Centre Energetique et Procédés, Groupe Energies renouvelables et réseaux. [MINES ParisTech](#), Sophia Antipolis, France.
Aout 2012
- Oct 2007** – ATTACHÉ TEMPORAIRE D'ENSEIGNEMENT ET DE RECHERCHE, [Université Joseph Fourier](#), Grenoble, France.
Aout 2008
- Oct 2004** – DOCTORANT EN MATHÉMATIQUES APPLIQUÉES, (Bourse région Rhones Alpes), [Université Joseph Fourier](#), Grenoble, France.
Sept 2007
- .
- Oct 2004** – MONITEUR D'ENSEIGNEMENT SUPÉRIEUR EN MATHÉMATIQUES APPLIQUÉES, [Université Pierre Mendès France](#), Grenoble, France.
Sept 2007
- .

Cursus et titres universitaires

- Oct 2004** – DOCTORAT DE MATHÉMATIQUES APPLIQUÉES, [LJK](#), Grenoble. Titre : Réduction de dimension en statistique et application à la segmentation d'images hyperspectrales. Sous la direction de Anestis Antoninadis (Professeur UJF) et Sophie Lambert-Lacroix (Professeur UJF), France.
Juin 2008
- 2003 – 2004** DIPLÔME MASTER RECHERCHE EN MATHÉMATIQUES APPLIQUÉES DE L'UNIVERSITÉ JOSEPH FOURIER, Grenoble, France.
- 2001 – 2004** [DIPLÔME D'INGÉNIEUR DE L'ECOLE NATIONALE SUPÉRIEUR D'INFORMATIQUE ET DE MATHÉMATIQUES APPLIQUÉES DE GRENOBLE](#), Grenoble, France.

Compétences, Divers

- Programmation Java, C, C++, Fortran, Pascal, Delphi, Ada, Python.
Software R, S, Matlab et Simulink, Scilab, Maple, keynote, latex, Eclipse, SVN, MS-Project, Word, Excel, Powerpoint, Open-Office, RapidWeather.
Data Format GRIB et Netcdf, outils nco, ncl, et GRIB_API, ainsi que interfaçage avec R et Matlab. Utilisation du système [MARS](#) du centre Européen de prédiction. Oracle.
Système Mac, Unix et Linux (écriture de script shell), Windows.
Langues Anglais : Lu, parlé, Ecrit.