



**HAL**  
open science

# Analyses probabilistes pour l'étude des réseaux électriques de distribution

Fallilou Diop

► **To cite this version:**

Fallilou Diop. Analyses probabilistes pour l'étude des réseaux électriques de distribution. Autre. Université Paris Saclay (COMUE), 2018. Français. NNT : 2018SACLC055 . tel-01936560

**HAL Id: tel-01936560**

**<https://theses.hal.science/tel-01936560>**

Submitted on 27 Nov 2018

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

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



# Analyses probabilistes pour l'étude des réseaux électriques de distribution

Thèse de doctorat de l'Université Paris-Saclay

École doctorale n° 575 Electrical, Optical, Bio :  
Physics and Engineering (EOBE)  
Spécialité de doctorat : Génie Électrique

Thèse présentée et soutenue à Gif-sur-Yvette, le 25 juin 2018, par

**Fallilou DIOP**

Composition du jury :

Demba DIALLO Professeur Université Paris Sud	Président
Raphaël CAIRE Maître de Conférences HDR INP Grenoble	Rapporteur
Bruno FRANCOIS Professeur Centrale Lille	Rapporteur
Corinne ALONSO Professeur INP Toulouse	Examineur
Marc PETIT Enseignant Chercheur HDR CentraleSupélec	Directeur de Thèse
Martin HENNEBEL Enseignant Chercheur CentraleSupélec	Encadrant
Mohamed ZAYANI Ingénieur Chercheur	Invité



# Remerciements

Je tiens tout d'abord à remercier l'IRT SYSTEM X d'avoir financé et co-encadré ces travaux de recherche. Dans ce cadre, je remercie particulièrement mes encadrants Mme Amira BEN HAMIDA et M Mohamed HAYKEL ZAYANI pour leurs conseils, leur soutien et leurs encouragements tout au long de cette thèse.

Je voudrais également remercier grandement mon directeur de thèse, M Marc PETIT, enseignant chercheur à CentraleSupélec, pour toute son aide. Je suis ravi d'avoir travaillé en sa compagnie car outre son appui scientifique, il a toujours été là pour me soutenir et me conseiller au cours de l'élaboration de cette thèse.

Je tiens à remercier M Martin HENNEBEL, enseignant chercheur à CentraleSupélec, qui m'a encadré tout au long de cette thèse et qui m'a fait partager ses brillantes idées. Qu'il soit aussi remercié pour sa gentillesse, sa disponibilité permanente, son aide précieuse pour la rédaction du manuscrit et pour les nombreux encouragements qu'il m'a prodigués.

M Bruno FRANCOIS, Professeur à l'Ecole Centrale Lille, et M Raphaël CAIRE, Maître de Conférences à l'Institut National Polytechnique de Grenoble, m'ont fait l'honneur d'être rapporteurs de ma thèse, ils ont pris le temps de m'écouter et de discuter avec moi. Leurs remarques m'ont permis d'envisager mon travail sous un autre angle. Pour tout cela je les remercie. Je tiens à remercier M Demba DIALLO, professeur Université Paris Sud, pour avoir accepté de présider mon jury de thèse et pour sa participation scientifique. Je remercie également Mme Corinne ALONSO, professeur à l'Institut National Polytechnique de Toulouse, pour l'honneur qu'il me fait d'être dans mon jury de thèse.

Je tiens à remercier particulièrement tous les membres de l'équipe SCE de l'IRT SYSTEMX, pour toutes nos discussions et leurs conseils qui m'ont accompagné tout au long de mes travaux de thèse.

A mon père Serigne DIOP, ma mère Khady TOURE,  
à ma chère épouse SODA FALL DIOP,  
à mes soeurs et frères Fatou Kiné DIOP, Sokhna DIARA DIOP, Cheikh Ahmadou  
Khadim DIOP, Fatim TOURE,  
à Aly Fall, Abdoulaye FALL, Sylveria DA SILVA,  
à mes neveux Mohamed Fadel FALL, Assane FALL, Isma FALL, Fallilou SARR,  
ainsi qu'à l'ensemble des mes amis et proches.

# Table des matières

<b>Remerciements</b>	<b>3</b>
<b>1 Introduction Générale</b>	<b>11</b>
1.1 Contexte global de la thèse . . . . .	11
1.2 Description générale du système électrique . . . . .	12
1.2.1 Schéma historique . . . . .	13
Production . . . . .	13
Le réseau de transport . . . . .	16
Distribution . . . . .	17
1.2.2 Un nouveau paradigme du système électrique . . . . .	20
Production décentralisée . . . . .	20
Véhicules électriques . . . . .	22
1.3 Problématiques . . . . .	23
1.3.1 Impacts liés à l'intégration de la production photovoltaïque . . . . .	24
1.3.2 Le déséquilibre . . . . .	24
1.3.3 Tension . . . . .	24
1.3.4 Impact sur les régulateurs . . . . .	25
1.3.5 Surintensité . . . . .	27
1.3.6 Protection . . . . .	27
1.3.7 Harmoniques . . . . .	27
1.3.8 Îlotage . . . . .	28
1.4 Plan du mémoire . . . . .	28
<b>2 Modélisation probabiliste dans l'étude des réseaux électriques de distribution</b>	<b>31</b>
2.1 Introduction . . . . .	31
2.2 Théorie des probabilités . . . . .	32
2.2.1 Notion de variables aléatoires . . . . .	32
2.2.2 Notion de densité de probabilité et de fonction de répartition . . . . .	33
Notion de fonction de répartition . . . . .	34
2.2.3 Notion de moments et d'estimateur . . . . .	34
Moments . . . . .	34
Estimation des moments . . . . .	37
2.3 Modélisation probabiliste de la production et de la consommation d'électricité . . . . .	38
2.3.1 État de l'art . . . . .	38

	Modélisation de la production . . . . .	38
	Modélisation de la consommation . . . . .	40
2.3.2	Modèle production PV . . . . .	41
	Estimation de la densité de probabilité . . . . .	41
	Modèle de la nébulosité basé sur le clustering . . . . .	43
	Échantillonnage . . . . .	46
2.3.3	Modèle de la consommation . . . . .	47
	Consommation des véhicules électriques . . . . .	48
	Consommation interne . . . . .	49
2.4	Modélisation des composants du réseau . . . . .	54
2.4.1	Les transformateurs . . . . .	55
2.4.2	Les lignes et câbles . . . . .	57
2.4.3	Les charges . . . . .	61
2.4.4	La production décentralisée . . . . .	62
2.5	Conclusion . . . . .	63
<b>3</b>	<b>Méthodes de Calcul</b>	<b>65</b>
3.1	Introduction . . . . .	65
3.2	Le Load flow déterministe . . . . .	66
3.2.1	La méthode Forward backward sweep . . . . .	67
	Exemple sur une branche monophasée . . . . .	67
3.2.2	Choix de l'outil de simulation . . . . .	70
	GridLab-D . . . . .	70
	OpenDSS . . . . .	71
3.3	Méthodes Load flow probabilistes . . . . .	72
3.3.1	État de l'art . . . . .	73
	Méthodes de simulation Monte Carlo . . . . .	73
	Les méthodes analytiques . . . . .	74
	Les méthodes approximées . . . . .	75
	Les autres méthodes . . . . .	76
3.3.2	Méthode Monte Carlo . . . . .	78
	Estimation de la probabilité de dépassement des seuils de fiabilité . . . . .	79
3.3.3	Méthode d'estimation ponctuelle (PEM) . . . . .	80
	Principes . . . . .	80
	Méthode Hong à une variable . . . . .	82
	Méthode Hong à plusieurs variables d'entrée . . . . .	85
	Application au calcul de load flow probabiliste . . . . .	87
3.4	Conclusion . . . . .	89
<b>4</b>	<b>Cas d'études</b>	<b>93</b>
4.1	Introduction . . . . .	93
4.2	Réseau Test IEEE 4 noeuds . . . . .	94
4.2.1	Le réseau . . . . .	94
	Application des méthodes LFP en mode instantané . . . . .	96
	Résultats . . . . .	96
4.3	Réduction des incertitudes . . . . .	98

4.3.1	Principe de la réduction . . . . .	99
	Réduction par clustering . . . . .	99
	Réduction par PEM . . . . .	99
4.3.2	Cas du paramètre nébulosité . . . . .	100
4.3.3	Cas du paramètre consommation interne . . . . .	102
4.3.4	Cas du paramètre distance parcourue VE . . . . .	106
4.3.5	Conclusion . . . . .	107
4.4	Comparaison entre méthode MC et méthode pseudo MC appliquées à toutes les variables . . . . .	109
4.4.1	Modèles des paramètres pour la méthode MC . . . . .	109
4.4.2	Méthode pseudo MC . . . . .	110
4.4.3	Conclusion . . . . .	114
4.5	Réseau Test de grande taille . . . . .	114
4.6	Résultats . . . . .	116
4.7	Applications . . . . .	117
<b>5</b>	<b>Conclusion Générale</b>	<b>121</b>
<b>6</b>	<b>Annexes</b>	<b>133</b>
6.1	Annexe 1 : Matrices généralisées des quadripoles du réseau test IEEE 4 noeuds . . . . .	133
6.2	Annexe 2 : Principe du FBS appliqué au réseau IEEE 4 noeuds . . . . .	134





# Liste des nomenclatures

$X$	Variable aléatoire
$f$	Densité de probabilité d'une variable aléatoire
$F$	fonction de répartition d'une variable aléatoire
$Mo_k$	Moment d'origine d'ordre $k$
$P_{pv}$	Puissance photovoltaïque produite à l'instant $t$
$c(t)$	Indice de nébulosité à l'instant $t$
$I(t)$	Irradiance solaire à l'instant $t$
$S$	Surface de l'installation PV
$\eta$	Rendement panneaux PV
$P_{consomsee}(t)$	Puissance totale soutirée en un instant $t$
$P_{ev}(t)$	Puissance soutirée par la recharge des VE
$P_{exo}(t)$	Puissance soutirée par les autres équipements
$P_{rech}$	Puissance de recharge d'un VE
$L$	Profil journalière de consommation domestique
$kc$	nombre de cluster
$ic$	Indice d'un cluster
$jc$	Indice d'un profil dans un cluster
$M$	Variable aléatoire discrète qui modélise le numéro de cluster
$mc$	Valeur de $M$
$V$	Tension électrique
$I$	Courant électrique
$Z$	Impédance d'une ligne
$n$	Nombre de paramètres aléatoires en entrée d'un problème de LFP
$m$	Nombre de points d'approximation PEM
$N$	Nombre de scénarios aléatoires d'une méthode Monte Carlo
$X$	Paramètre aléatoire d'entrée d'un load flow
$i$	Indice d'un paramètre aléatoire d'entrée du load flow
$U$	Grandeur électrique résultant d'un load flow
$h$	fonction non linéaire du load flow
$r$	Ordre d'un moment statistique
$k$	Indice d'un point d'approximation PEM pour une variable donnée
$l$	Indice d'un paramètre aléatoire d'entrée du load flow



# Chapitre 1

## Introduction Générale

### 1.1 Contexte global de la thèse

Cette thèse a été financée par l'IRT systemX dans le cadre du projet SCE (smart city energy analytic). L'IRT travaille en collaboration avec des industriels, des académiques et des strat-up sur des problématiques de recherche en lien avec les domaines de la mobilité, de l'énergie et des télécommunications.

Ledit projet SCE se donne pour objectif de fournir des études d'aide à la décision pour une gestion intelligente du système électrique. Ces outils sont basés sur une valorisation des données de production et de consommation d'électricité mais aussi météorologiques comme la nébulosité et financière notamment les prix d'électricité. Un des use cases de ce projet est le développement d'un outil de simulation dédié à modéliser le déploiement des panneaux photovoltaïques domestiques sur un horizon de 20 ans. Le niveau de déploiement est déterminé en fonction de trois critères :

- Un critère économique qui inclut le coût d'investissement initial et la rentabilité de l'installation,
- Un critère politique qui permet de prendre en compte les politiques incitatifs de l'Etat notamment la mise en place de tarif de rachat garantie,
- Un critère social qui symbolise la catégorie socio professionnelle des agents simulés et l'impact du voisinage.

Afin de compléter ce modèle, la prise en compte des impacts de la production PV sur les contraintes techniques du réseau était nécessaire. C'est dans ce cadre qu'interviennent

les travaux de cette thèse qui visent à développer des méthodes de quantification de l'effet de la production PV sur le réseau de distribution en prenant en compte l'effet intermittente de ce type de production.

## 1.2 Description générale du système électrique

Par système électrique (SE) on entend l'ensemble de la chaîne d'alimentation en électricité. Il regroupe les centrales de production, les sites de consommation et les réseaux qui acheminent l'électricité des centrales de production vers les sites de consommation. Le stockage de l'électricité en très grande quantité étant très coûteux, le SE actuel repose sur un équilibre en temps réel entre la production et la consommation.

Dans ce cadre, le schéma historique du système électrique repose sur une organisation unidirectionnelle partant de la production vers les consommateurs. Dans ce schéma, la production est centralisée au niveau des centrales de production, réparties sur le territoire, dont les puissances vont de quelques mégawatts au millier de mégawatts. La production est groupée et faite en grande quantité dans le but de favoriser l'effet d'échelle. La plupart des centrales sont connectées sur un réseau dit de transport qui couvre l'ensemble du territoire et permet d'acheminer l'électricité de ces centrales de production aux zones de consommation éloignées. Pour des raisons technico-économiques, le transport est effectué sous très haute tension triphasée (de 63 à 90 kV en France)). Or la majeure partie des équipements électriques dans les foyers (éclairage, télévision, réfrigérateur,...) est monophasée et requière une alimentation basse tension monophasée. Le réseau de distribution assure l'interface entre le réseau de transport et les consommateurs de faible puissance alimentés en HTA et basse tension. La tension est successivement abaissée au niveau des postes sources puis au niveau des postes de distribution pour alimenter les usagers raccordés.

La libéralisation du marché de l'électricité, la volonté de trouver une alternative aux sources d'énergie fossile, les politiques incitatives d'installation et de revente de production d'origine renouvelable (tarifs de rachat garantis), le développement des véhicules

électriques et du stockage ont fait émerger un nouveau paradigme du système électrique. La production y est décentralisée avec des unités de production sur les différents niveaux, que ce soit sur le réseau de transport (comme dans le schéma historique), ou sur les réseaux de distribution moyenne et basse tension (avec des panneaux photovoltaïques sur les toits des maisons par exemple). A cela s'ajoute le développement du stockage qui offre plus de flexibilité au système électrique, notamment avec les batteries des véhicules électriques. Parce que ce nouveau paradigme offre plus de services et met en jeu plus d'acteurs, il présente aussi plus d'incertitudes sur les comportements des éléments qui le constituent. Par exemple, la puissance injectée sur le réseau par les unités de production décentralisées ou les besoins de recharge des véhicules électriques comportent des niveaux d'incertitude dont il faut tenir compte dans la gestion des réseaux.

Dans ce qui suit, après avoir détaillé le fonctionnement du schéma historique et ce nouveau paradigme, nous évoquerons les problématiques induites par ce nouveau paradigme sur la gestion des réseaux de distribution basse tension.

### 1.2.1 Schéma historique

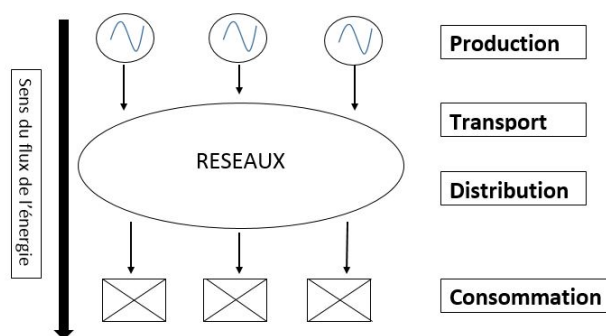


FIGURE 1.1 – Schéma historique du système électrique

#### Production

Dans l'organisation historique du système électrique, la production de l'électricité est centralisée dans des sites de production dont les capacités s'échelonnent de quelques dizaines à plusieurs centaines de mégawatts (MW). Ces niveaux de puissance se justifiaient

par l'amélioration du rendement obtenue avec l'augmentation de la taille des machines tournantes utilisées pour produire l'électricité. [1].

Classiquement, la production de l'électricité reposait sur le principe de la transformation d'énergie primaire en énergie électrique. Les sources primaires peuvent être réparties en deux grandes catégories selon leur origine :

- Sources primaires fossiles : fioul, charbon, gaz, uranium.
- Sources primaires renouvelables : eau, soleil, vent, biomasse, géothermie.

Pour les centrale à combustible, la chaleur dégagée par la combustion produit de la vapeur qui fait tourner une turbine qui entraîne à son tour un alternateur pour produire de l'électricité. Pour sources de production renouvelables, biomasse mise à part, le principe de fonctionnement est différent : dans le cas des centrales hydrauliques et éoliennes, la source primaire en occurrence l'eau (le vent) entraîne directement la turbine. Dans les centrales solaires, la production de l'électricité est basée sur une transformation directe de la lumière solaire en énergie électrique grâce à l'effet photovoltaïque. L'utilisation d'un type de centrale donné est déterminé par son coût de production et l'ordre de grandeur de puissance qu'elle peut fournir.

TABLE 1.1 – Ordres de grandeur d'unités de production en fonction de la technologie utilisée

Type	Puissance
Centrale nucléaire	900 à 1650 MW
1 Éolienne	2 à 5 MW
Centrale hydro-électrique	quelques kW à 1800 MW
Centrale thermique à flamme	120 à 700 MW
10 $m^2$ de panneau PV	1kWc

Le tableau 1.2.1 montre le coût de production de quelques sources d'énergie en France.

Sur les centrales conventionnelles fonctionnant à base de combustibles (gaz, fioul, charbon, biomasse, nucléaire etc...), il est possible de prévoir, moduler et contrôler la puissance produite en agissant sur la quantité de combustible à brûler. Par contre, dans le cas des centrales éoliennes et photovoltaïques, les sources primaires (vent, soleil) dépendent des conditions climatiques, et leur production ne peut pas être contrôlée au delà de ce qu'ils peuvent produire. De ce fait, il est impossible de prévoir avec exactitude la puissance électrique produite par ces centrales puisqu'elle dépend de l'intermittence

TABLE 1.2 – Coût de production des différentes sources d'énergies (source [2])

(en euros par MWh)	UFE/DGEC (dernières données actualisées pour 2011)	Énergies 2050 (rapport Février 2012)	Cours des comptes (Janvier 2012 pour le nucléaire, Juillet 2013 pour les renouvelables)	Tarif d'achat au 1 <sup>er</sup> Septembre 2013
Nucléaire	42,3	56	49,5	
Hydroélectricité	55		43-188	60,7 durant 20 ans + primes pour les petites installations
CCGT gaz	61	69		
Éolien terrestre	65	73	62-102	82 durant 10 ans puis 28 à 82 durant 5 ans selon les sites
Charbon	66	67		
Fioul	86			
Éolien off-shore	143	102	87-116	130 durant 10 ans, puis 30 à 130 durant 10 ans selon les sites
Photovoltaïque	217	150	114 -547	110 à 200 selon les sites

des phénomènes météorologiques (vitesse du vent, passage de nuage etc...). Ce caractère intermittent des productions éolienne et photovoltaïque rend difficile leur intégration sur les réseaux de transport et de distribution électriques.

La principale caractéristique de l'électricité est qu'elle n'est pas stockable. Pour palier ce phénomène, le système physique constitué des réseaux de transport et de distribution doit assurer la connexion entre les producteurs et les consommateurs et maintenir en temps réel l'équilibre entre les puissances injectées et soutirées.



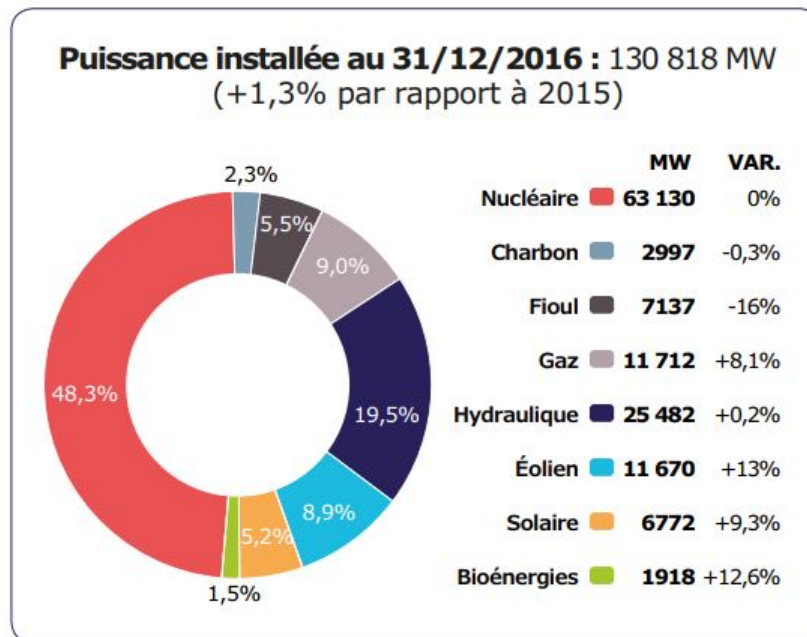


FIGURE 1.2 – Mix énergétique France en 2016 (source RTE)

## Le réseau de transport

Le réseau de transport achemine l'électricité produite, sur une très grande distance des sites de production aux postes sources du réseau de distribution. Il est constitué de lignes exploitées en Haute Tension B (HTB), de sorte à réduire le courant et minimiser les pertes Joule, ainsi les chutes de tension sur les longues distances [1]. Pour augmenter la capacité, la puissance est acheminée en régime triphasé alternatif. Selon les caractéristiques des lignes utilisées (isolant, section du conducteur, etc.) et les conditions d'exploitation (tension et courant admissibles, chute de tension admissible, possibilité de report en cas d'incident), une portion de ligne du réseau de transport permet d'acheminer des dizaines voire des centaines de mégawatts sur plusieurs centaines de kilomètres. Le tableau 1.2.1 donne une idée de la puissance transmissible en fonction de la tension et de la distance. Les centrales de production n'étant pas toujours à proximité des zones de consommation, et avec les quantités d'énergie produite mises en jeu, le réseau de transport joue un rôle stratégique dans l'exploitation des systèmes électriques. Il permet l'acheminement de grandes quantités d'énergie sur de longues distances, dans des conditions techniques et économiques favorables.

TABLE 1.3 – Puissance transmissible et distance ([1])

Tension(kV)	Puissance(MW)	Distance(km)
63	20	
150	80	100
225	200	200
400	700	400
	1200	100
750	2500	200
	1000	600

Le réseau de transport fonctionne suivant des critères de tension et de fréquence bien définis. La fréquence nominale, qui est par exemple de  $50Hz$  en France et  $60Hz$  aux États-Unis, est proportionnelle à la vitesse de rotation des alternateurs des centrales de production. Ces valeurs de fréquence sont le résultat d'un compromis entre la taille des machines (inversement proportionnelle à la fréquence) et les pertes dans les machines et les conducteurs (proportionnelles à la fréquence)[1]. Tout déséquilibre entre la production et la consommation dans le système se traduit par un décalage de la fréquence par rapport à sa valeur nominale.

Pour des raisons de sécurité d'approvisionnement, le réseau de transport a une structure maillée. Ceci permet de garantir la continuité d'approvisionnement en cas de perte d'un ouvrage important. Le respect de la règle dite du  $N - 1$  permet d'assurer la continuité d'alimentation en cas de perte d'un élément du réseau de transport. Concrètement, cela veut dire que chaque nœud du réseau est alimenté aux moins par deux lignes. Ces mesures s'expliquent par l'importance stratégique du réseau de transport et que la perte d'une portion de ligne sans redondance pourrait impacter des dizaines voire des centaines de milliers de consommateurs.

## Distribution

Le réseau de distribution permet l'alimentation des consommateurs finaux en basse tension sur l'ensemble du territoire. De ce fait, il est beaucoup plus dense que le réseau de transport. Par exemple, le réseau de transport de France métropolitaine est constitué de lignes d'une longueur cumulée dépassant les 100 000 km alors que les réseaux de distribution cumulent plus d'un million de km [1], la moitié étant exploitée en HTA, le

reste en BT. Leur topologie est essentiellement radiale avec des longueurs de lignes en BT de l'ordre de la centaine de mètres.

Le réseau de distribution est alimenté sur sa partie moyenne tension (HTA) par des postes sources alimentés depuis le réseau de transport (HTB). Ces postes sont équipés de transformateurs HTB/HTA pour abaisser la tension à 20 kV entre phases [3] et d'autres équipements de mesure et de contrôle-commande, de sectionneurs, de disjoncteurs. Sur chaque poste source, sont raccordés des départs moyenne tension dont le nombre varie de quelques unités à plusieurs dizaines [3]. Ces départs alimentent les consommateurs raccordés en moyenne tension et les postes de distribution basse tension (BT) équipés de transformateurs HTA/BT.

En milieu rural, le poste de distribution, constitué d'un transformateur sur poteau ou en cabine simplifiée peut alimenter un ou deux départs basse tension (BT) et desservir une dizaine de clients domestiques. En zone urbaine, les postes en cabine ou en immeuble peuvent alimenter jusqu'à huit départs (voire 16 dans le cas de postes à deux transformateurs) et desservir jusqu'à 300 logements [3]. Le niveau de tension des départ BT est de 230/400V et les branchements sont dimensionnés pour  $12kVA$  en monophasé. Pour des puissances supérieures, le raccordement se fait en triphasé jusqu'à la limite de  $250kVA$ .

L'essentiel de la consommation raccordée au réseau de distribution basse tension est constituée de charges monophasées, qui vont engendrer du déséquilibre entre les phases. Afin de créer un conducteur de neutre, et d'éviter les surtension dues au déséquilibre, les secondaires des transformateurs de distribution sont connectés en étoile [4]. Historiquement, le déséquilibre entre phases est limité par le foisonnement des charges connectées sur ces trois phases, ce qui permet dans certains cas de dimensionner le conducteur de neutre avec une section inférieure à celle des trois phases.

Le gestionnaire du réseau de distribution (tout comme celui du réseau de transport) est tenu de respecter un certain nombre de contraintes pour garantir une qualité de service fixée par le législateur. Le respect de ces contraintes permet, d'une part, d'assurer la sécurité d'approvisionnement et d'autre part de garantir les conditions de bon fonctionnement des équipements du réseau (lignes, transformateurs, contrôleurs de tension) et des équipements électriques des utilisateurs raccordés au réseau. Tous ces équipements ont été conçus et dimensionnés pour fonctionner avec des niveaux de tension, de courant

et de puissance bien définis.

Parmi ces contraintes, on peut citer d'abord le maintien global de la tension dans la plage réglementaire. Pour le réseau BT, le critère d'évaluation de la tenue globale de tension est le pourcentage d'utilisateurs dits mal connectés, c'est à dire le pourcentage d'utilisateurs du réseau dont le point de connexion a subi au moins une fois dans l'année une tension inférieure à 90% ou supérieure à 110% de la tension nominale admise en ce point pendant une durée supérieure à 10 minutes [5]. La qualité de service est considérée comme non respectée si ce pourcentage calculé pour l'ensemble des réseaux dans un département dépasse 3% en France métropolitaine et 5% dans les DOM TOM [5].

A côté de la tenue de tension au niveau global, le respect de la tenue de tension au niveau d'un point particulier du réseau exige que la tension mesurée en ce point de connexion n'excède pas les limites citées plus haut  $[0.9V_n \text{ } 1.1V_n]$  pendant une durée de plus de 10 minutes. En plus, dans le cas du réseau de distribution BT, un critère de gradient de tension doit être respecté : la chute de tension observée à ce point ne doit pas dépasser 2% si une charge monophasée d'un kilowatt ( $1kW$ ) y était raccordée [5].

Le rôle du gestionnaire de réseau est avant tout d'assurer la sécurité et la continuité de l'approvisionnement en électricité. Cette continuité de l'alimentation peut être mise en danger par des événements météorologiques extrêmes, des défaillances des composants du système ou un déséquilibre entre la demande et la production disponible. Le critère de continuité de l'approvisionnement est quantifié par le pourcentage d'utilisateurs mal alimentés. Un utilisateur est réputé mal alimenté s'il a subi au cours d'une année :

- Une durée totale de coupures longues strictement supérieure à 13 heures,
- Un nombre de coupures longues (supérieures à 3 minutes) strictement supérieur à 6,
- Un nombre de coupures brèves (1 seconde à 3 minutes) strictement supérieur à 35

L'indice de continuité globale de l'alimentation est supposé non respecté si le pourcentage d'utilisateurs mal connecté dépasse 5% [5].

## 1.2.2 Un nouveau paradigme du système électrique

Le schéma historique basé sur une centralisation des moyens de production et un transit unidirectionnel du flux de puissances se transforme aujourd'hui sous l'effet de la libéralisation du secteur de l'électricité, de la volonté de développer des sources d'énergies renouvelables et de l'utilisation de nouveaux usages comme les véhicules électriques. L'organisation historique du système électrique (cf figure 1.1) tend vers une architecture caractérisée par des unités de productions de petite puissance, de plus en plus décentralisées 1.3.

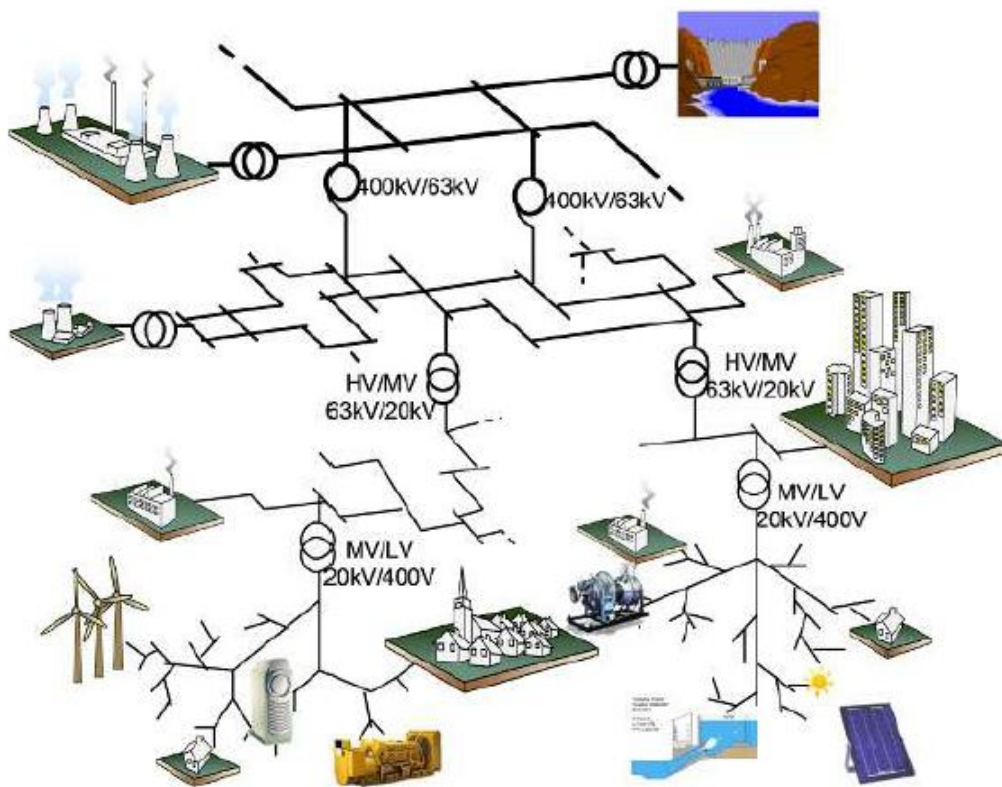


FIGURE 1.3 – Nouvelle organisation du système électrique (source [6])

### Production décentralisée

La libéralisation des marchés, qui a amené à la séparation des activités de production, de transport et de distribution, ainsi les politiques incitant à la production et à la vente d'électricité d'origine renouvelable par la mise en place de tarifs de rachat garantis ont permis de développer une nouvelle catégorie de producteurs d'électricité. Des particuliers

peuvent ainsi injecter leur production sur le réseau en bénéficiant d'un prix de revente attractif.

Ces unités de production peuvent être des panneaux photovoltaïques de quelques kilowatts sur le toit d'une maison ou encore une ferme de quelques dizaines de kilowatts de panneaux solaires, d'éoliennes ou une mini centrale de cogénération. Ces nouveaux moyens de production décentralisés, de puissances unitaires faibles, sont raccordés aux réseaux de distribution où ils modifient la répartition des flux. On constate ainsi une circulation bidirectionnelle de l'énergie produite, qui peut remonter vers les postes de distribution et les postes sources. L'intégration de cette énergie, dont certains moyens de production présentent un caractère intermittent, amène des complications dans la gestion des réseaux.

Du fait de leur dépendance directe aux conditions météorologiques (soleil, présence de

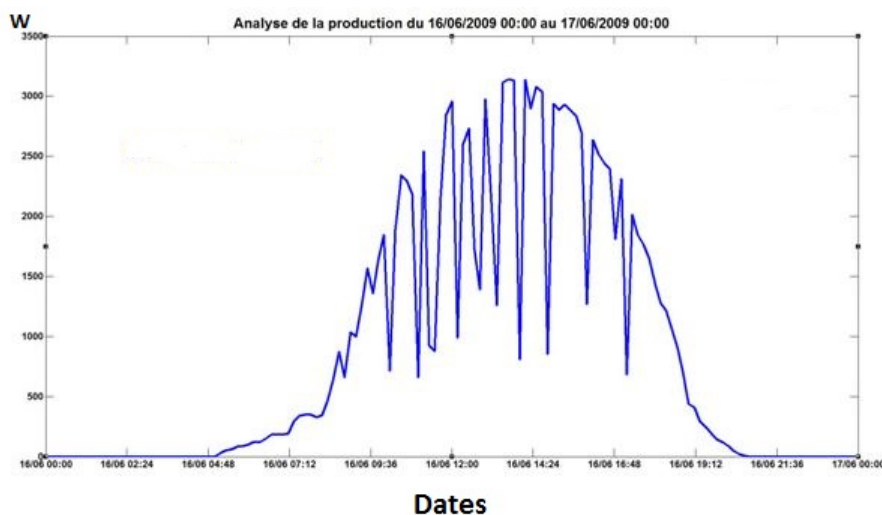


FIGURE 1.4 – Profil de puissance générée par des panneaux photovoltaïques lors d'une journée ensoleillée avec nuages (source [7])

nuages, vitesse du vent) qui peuvent être changeantes d'une journée à une autre ou d'une heure à une autre dans la même journée (cf image 1.4), l'énergie produite par des panneaux solaires ou un parc éolien présente un caractère aléatoire. Pour être en mesure d'assurer les conditions de sécurité et de continuité d'approvisionnement, les gestionnaires de réseaux de distribution (GRD) et de transport (GRT) doivent prendre en compte ces incertitudes production dans la planification et la gestion des réseaux.

## Véhicules électriques

La volonté de réduire la dépendance aux sources d'énergie fossile s'est traduite dans le domaine de la mobilité par le développement des véhicules électrifiés (VEs). L'utilisation des VEs limite la consommation de produits dérivés du pétrole (essence, gasoil) et participe ainsi à la réduction des émissions de gaz à effet de serre (à condition que le mix énergétique pour produire de l'électricité soit décarbonné). Malgré l'existence de certains verrous économiques et techniques (coût et taille des batteries, durée de la recharge...) qui limitent le taux de pénétration des VEs (par exemple en France 22 000 VEs immatriculés en 2016 sur plus de 2 millions de véhicules immatriculés chaque année), leur développement est en croissance (cf figure 1.5). Les réseaux de distribution doivent

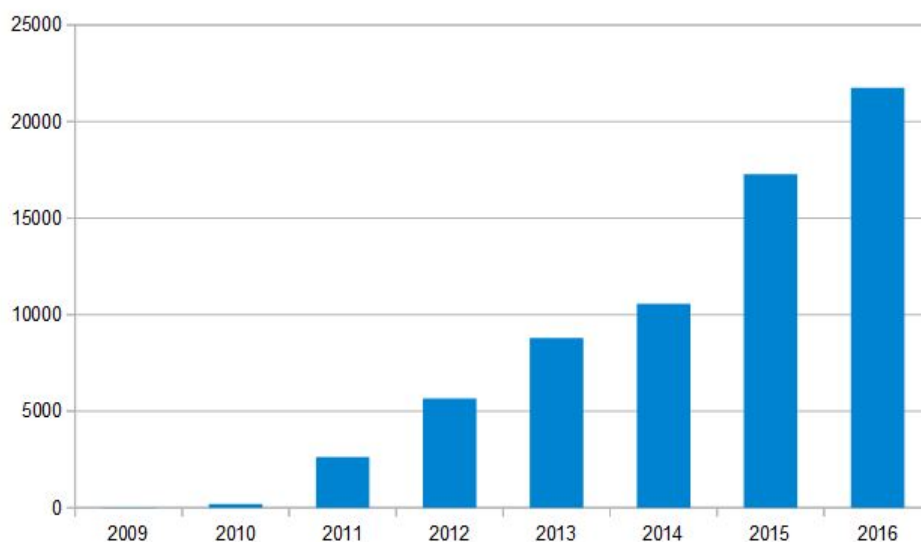


FIGURE 1.5 – Nombre de véhicules électriques immatriculés en France par année (source [8])

être en mesure d'alimenter ces nouveaux types de consommation, via le raccordement de bornes de recharges dont les puissances peuvent varier de  $3.3kVA$  à  $43kVA$  [9].

Les demandes de recharge de ces véhicules peuvent se situer à toute heure de la journée selon les habitudes de mobilité de chacun. Elles constituent ainsi des blocs de consommation d'énergie de durées et d'amplitudes variables et situées à des instants aléatoires dans la journée. Ceci rajoute ainsi de l'incertitude à la consommation d'électricité qui était déjà fluctuante.

La cohabitation entre les éléments du schéma historique et les éléments issus du nou-

TABLE 1.4 – Capacités des batteries de quelques modèles de VE (source [10])

Voitures	Capacité batterie
Tesla Model S 75	75 kWh
Tesla Model S 90D	90 kWh
Tesla Model S 100D	100 kWh
Tesla Model X 75D	75 kWh
Tesla Model X 100D	100 kWh
Renault ZOE	22 kWh
Renault ZOE Z.E. 40	41 kWh
Nissan LEAF	24 kWh
Nissan LEAF 2017	30 kWh
Smart ForTwo Electric Drive	17,6 kWh
Opel Ampera-E	60 kWh
Volkswagen e-Golf	35,8 kWh
BMW i3 22	22 kWh
BMW i3 33	33 kWh

veau paradigme (production décentralisée, véhicules électriques, stockage, micro réseau...) fait émerger des problématiques dans la gestion du système électrique. Cette cohabitation, qui est nécessaire puisque les composants du système existant ne peuvent être remplacés d'un seul coup, aura des impacts sur les réseaux électriques et sur le fonctionnement global du système. Notre étude va se porter sur les effets de la production décentralisée, à l'échelle domestique, et des véhicules électriques sur le fonctionnement du réseau de distribution BT. Le réseau de distribution BT est le composant du système existant qui est en interaction directe avec le consommateur final.

### 1.3 Problématiques

Dans ce nouveau contexte, la garantie des conditions de fiabilité du réseau (sécurité et continuité d'approvisionnement, contraintes techniques...) passe par une estimation préalable des impacts de la production décentralisée, et plus particulièrement celle des panneaux PV et de la consommation des VEs. Pour plus de fiabilité, le modèle d'estimation de ces impacts devra prendre en compte les incertitudes sur la production et la consommation raccordées au réseau.



### 1.3.1 Impacts liés à l'intégration de la production photovoltaïque

L'injection de puissance sur un réseau de distribution BT peut causer des problèmes d'augmentation locale de la tension, de surintensité, de déséquilibres de tension, de courant etc...[11]. L'organisation historique du système électrique montre que le réseau de distribution n'a pas été conçu pour y raccorder de la production. La production de panneaux PV, sur les toits des maisons, injectée sur ce réseau crée un flux de puissance inverse qui peut créer des contraintes difficilement gérables avec les modes d'exploitation actuels du réseau de distribution [12].

### 1.3.2 Le déséquilibre

Le raccordement de consommateurs monophasés, avec différents niveaux de puissances, sur le réseaux de distribution crée du déséquilibre sur ce réseau. Ce déséquilibre peut être réduit grâce à une bonne répartition des charges afin que les puissances circulant sur différentes phases soient proches et donc que le réseau soit presque équilibré. Le raccordement de panneaux PV domestiques pouvant s'effectuer en monophasé, cela constitue une source supplémentaire de déséquilibre de tension comme de courant si leur répartition n'est pas équilibrée [13],[14]. Le déséquilibre de courant a pour conséquence l'augmentation du courant dans le neutre qui peut atteindre des proportions supérieures au courant nominal admis dans le conducteur [15], [16], [17]. Le niveau de déséquilibre en tension peut lui être quantifié par un indice de déséquilibre  $T_{des}$  égal au rapport entre les composantes homopolaire et directe des tensions triphasées [18] (cf équation 1.1) .

$$T_{des} = \frac{V_{homopolaire}}{V_{directe}} \times 100 \quad (1.1)$$

### 1.3.3 Tension

L'injection de puissance produite par la production décentralisée fait varier la circulation du flux de puissance dans le réseau, ce qui impacte le profil de tension [19],[20] et [21]. En effet, dans une branche d'un réseau à courant alternatif, la différence de valeur

efficace des tensions aux deux extrémités peut être approximée par la formule suivante :

$$\Delta V \approx \frac{P \times R + Q \times X}{V} \quad (1.2)$$

Avec  $P$  la puissance active consommée,  $Q$  la puissance réactive,  $R$  la résistance série de branche,  $X$  la réactance série de branche et  $V$  la moyenne des tensions aux deux extrémités. Selon la norme IEEE 1547, un système de production photovoltaïque doit fonctionner avec un facteur de puissance unitaire, c'est à dire qu'il n'absorbe ni n'injecte de puissance réactive sur le réseau. Le problème de surtension due à l'injection de puissance active est moins sensible dans le réseau de transport puisque le rapport  $X/R$  est élevé. Cependant, dans le cas des réseaux de distribution, la section des conducteurs est moins grande, ce qui se traduit par des résistances plus élevée, et l'écartement des conducteurs étant plus large, les réactances de ligne ont tendance à être plus faibles. Le rapport  $X/R$  y est proche de l'unité, et par conséquent les transits de puissance active ont le même poids dans la chute de tension que les transits de puissance réactive. En l'absence de production décentralisée sur un réseau de distribution, l'amplitude de tension décroît de façon monotone le long du réseau de la source vers les terminaux, s'il n'existe pas de système de régulation. L'installation de sources de production décentralisées comme des panneaux PV par exemple, peut être considérée du point de vue du réseau comme une réduction de la puissance active consommée. Ainsi la chute de tension est atténuée. Cependant, quand la production de PV est plus grande que la consommation locale, le consommateur avec un panneau PV est considéré comme un producteur qui fournit de l'énergie. De ce fait, la tension locale va être supérieure à celle de ses voisins. Ce cas de figure peut se présenter même si la capacité de l'installation est plus petite que la consommation maximale locale, puisque le profil de production PV et le profil de consommation dans une journée ne coïncident généralement pas.

### 1.3.4 Impact sur les régulateurs

La consommation dans le réseau varie tout le temps et la structure du réseau n'est pas non plus fixe. De ce fait, pour continuer à assurer une qualité de service suffisante, des régulateurs sont utilisés. Dans le réseau de distribution, le but principal est de garder la

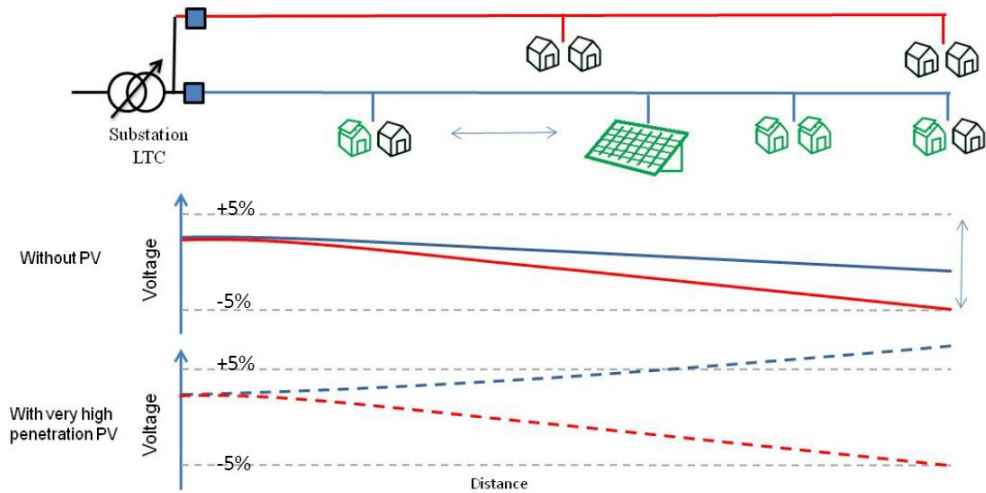


FIGURE 1.6 – Exemple d’une fonction de masse

tension dans une plage normée. Trois moyens de régulation de tension sont généralement utilisés : le changement de prise du transformateur (LTC pour Load tap changer), que l’on peut régler en charge sur certaines installations (OLTC pour On load tap changer) ou bien hors charge pour les autres (Off load tap changer), notamment utilisés en BT. Le régulateur de tension (SVR pour step voltage regulator) et les bancs de capacité sont les autres moyens utilisés pour la régulation de la tension. Les LTC et le branchement de condensateur se font aujourd’hui au niveau des postes source. Quand le réglage de ces derniers n’assure pas un profil de tension respectant la norme dans le réseau, Il est nécessaire de prendre des mesures correctrices.

L’intermittence de la production pose un problème lorsque le taux de pénétration de PV est élevé. La variation de l’irradiance conduit à une variation de puissance produite du PV et donc de la tension sur le réseau. Cette fluctuation de tension peut entraîner, sur une courte période, des déclenchements répétés des systèmes de régulations, ce qui se répercute sur leur durée de vie. En général, les GRD (Gestionnaires du réseau de distribution) définissent un seuil de déclenchement par journée pour ces appareils. L’étude [22] montre sur un cas réel, que l’intégration de PV augmente le nombre de déclenchement de banc de condensateur. Ce dernier dépasse le nombre de déclenchement limite défini par le gestionnaire.

### 1.3.5 Surintensité

La surintensité peut causer des problèmes thermiques dus à l'effet Joule. Avec un taux limité de pénétration de PV, l'électricité produite et celle consommée localement se compensent et le transit depuis la source est réduit. Cependant, une injection de puissance et de courant de PV, peut créer un flux de puissance inverse. L'excès de production peut aboutir à une surcharge de la capacité de la ligne ou du transformateur de distribution local. Une surintensité fait augmenter la température de la ligne et des transformateurs et réduit la durée de vie de l'appareil. Comme nous l'avons vu précédemment 1.3.2, le risque de surintensité dans le conducteur de neutre est augmenté lorsque l'on raccorde de la production décentralisée monophasée, qui accentue le déséquilibre entre les phases.

### 1.3.6 Protection

L'intégration de nouvelles sources de production contribue au courant de défaut et influence le comportement transitoire après le défaut. La contribution du courant injecté par un système de PV en cas de défaut peut aboutir à un phénomène d'aveuglement des dispositifs de protection. [23]. En outre, si le relais est conçu pour un réseau radial, le courant de défaut qui vient du PV pourrait faire déclencher ce relais à cause de la non-sensibilité au sens. Il est donc important d'étudier la contribution de courant de défaut depuis le panneau photovoltaïque. En ce qui concerne le courant de défaut, un point particulier est que le photovoltaïque est raccordé au réseau par un onduleur. Il n'y a pas de modèle standard pour le courant de défaut depuis le PV, le comportement variant selon son algorithme de contrôle et le matériel utilisé. L'intégration de PV pourrait nécessiter une reconfiguration pour les dispositifs de protection. Notre étude portant sur le régime stationnaire, nous n'abordons pas cette problématique.

### 1.3.7 Harmoniques

Le système de production PV est associé à de l'électronique de puissance (onduleur notamment) qui crée des harmoniques dans le réseau. Néanmoins, les systèmes photovoltaïques modernes profitent de la technologie de transistor qui normalement produit peu d'harmoniques mais néanmoins pose certains problèmes évoqués par [24].

### 1.3.8 Îlotage

Quand le réseau de distribution est coupé en amont du consommateur et que les productions décentralisées restent connectées, la zone du réseau qui contient les PV est isolée du reste du réseau public. Cette configuration est appelée l'îlotage. Ce phénomène n'est pas toléré et le système de PV doit se déconnecter du réseau en ce moment-là. Cette fonctionnalité est exigée par la norme IEEE 1547 et nous ne l'abordons pas dans notre étude pour les mêmes raisons que la protection.

## 1.4 Plan du mémoire

Le but de cette thèse est d'étudier des modèles probabilistes pour estimer les impacts de la production photovoltaïque et des véhicules électriques sur le réseau de distribution basse tension. Le chapitre 2 pose les bases théoriques de la modélisation probabiliste en s'appuyant sur la théorie des probabilités et celle de la modélisation des réseaux électriques de distribution en régime permanent. La deuxième partie de ce chapitre évoque les modèles probabilistes de consommation électrique et de production photovoltaïque développés dans le cadre de cette thèse.

Dans le chapitre 3, sont évoquées les méthodes de calcul de répartition de puissance (Load flow) utilisées dans l'étude du comportement des réseaux électriques en régime permanent en fonction des puissances injectées et soutirées en chaque point. Les méthodes classiques de load flow déterministe ne permettent pas de prendre en compte la diversité des scénarios de production et de consommation possibles. Pour cela, des méthodes probabilistes sont étudiées permettant ainsi de mieux prendre en compte l'aspect aléatoire de la production et de la consommation d'électricité. A ce titre, trois méthodes seront étudiées : la première, plus intuitive, est la méthode de Monte Carlo basée sur une répétition de calcul de load flow déterministes avec des scénarios aléatoires de puissances produites et consommées en chaque point du réseau. La deuxième méthode, dite méthode PEM (Point Estimate method), est basée sur une approximation des densités de probabilité des paramètres d'entrées en quelques points caractéristiques. Enfin la dernière méthode, dite pseudo Monte Carlo, est basée sur la répartition des valeurs possibles des paramètres d'entrée en quelques clusters afin de réduire les scénarios possibles.

Le chapitre 4 présente une application de ces méthodes probabilistes étudiées pour estimer la tenue de tension et le courant de neutre, d'abord sur un réseau simple à 4 nœuds, puis sur un réseau test d'EPRI de grande taille pour tester leurs limites. Enfin, le chapitre 5 est une conclusion générale pour souligner les apports et les perspectives de cette thèse.



# Chapitre 2

## Modélisation probabiliste dans l'étude des réseaux électriques de distribution

### 2.1 Introduction

Dans le but d'étudier l'impact de la production décentralisée intermittente et des consommations d'électricité sur les contraintes techniques du réseau, une phase préalable de modélisation du comportement des différentes entités mises en jeu est nécessaire. Dans ce cadre, ce chapitre traite de la modélisation des composants du réseau de distribution d'électricité dans leur fonctionnement en régime permanent ainsi que celle des unités de production et des charges qui y sont raccordées. Le caractère aléatoire de la production et de la consommation d'électricité transitant sur le réseau en un instant donné justifie l'utilisation de modèles probabilistes afin de mieux représenter l'incertitude sur ces éléments.

Ce chapitre sera introduit par un rappel sur quelques notions de théorie des probabilités qui seront utiles pour l'élaboration des méthodes de quantification d'incertitude sur les réseaux, puis nous exposerons les travaux dans la littérature portant, à notre connaissance, sur ce sujet, ainsi que les modèles probabilistes de production photovoltaïque et de consommation d'électricité développés et utilisés dans cette thèse . Enfin, les modèles



électriques des composants du réseau (transformateurs, lignes, charges...) utilisés pour l'étude du système en régime permanent seront décrits.

## 2.2 Théorie des probabilités

### 2.2.1 Notion de variables aléatoires

Une variable aléatoire est une variable dont la valeur est le résultat d'une expérience aléatoire. Par exemple, considérons l'expérience aléatoire consistant à tirer une boule dans un panier contenant 2 boules, une noire et une blanche. Soit  $X$  un nombre prenant les valeurs 0 ou 1 selon la couleur de la boule tirée à chaque expérience.

$$X = \begin{cases} 0 & \text{si la boule tirée est noire} \\ 1 & \text{si la boule tirée est blanche} \end{cases}$$

$X$  est une variable aléatoire symbolisant la couleur de la boule tirée à chaque expérience.

La valeur d'une variable aléatoire ne peut être connue avant la réalisation de l'expérience aléatoire dont elle est le résultat. Néanmoins, il est possible de connaître à l'avance la probabilité d'occurrence des valeurs possibles. Par exemple, dans l'expérience précédente,  $X$  ne peut prendre que deux valeurs (0 ou 1) avec des probabilités  $P(X = 0) = \frac{1}{2}$  et  $P(X = 1) = \frac{1}{2}$ .

On distingue les variables aléatoires entre celles de nature discrète de celles de nature continue. Une variable aléatoire discrète ne peut prendre qu'un nombre limité de valeurs dans l'intervalle de valeurs possibles de l'expérience aléatoire. Par exemple, la variable qui représente le numéro de face issu d'une expérience de lancement d'un dé est discrète. Une variable aléatoire est dite continue si elle peut prendre une infinité de valeurs dans l'intervalle des résultats possibles de l'expérience aléatoire. Par exemple les variables aléatoires représentant la durée de vie d'une batterie, la valeur d'une tension ou encore l'énergie journalière consommée par un foyer sont de nature continue.

## 2.2.2 Notion de densité de probabilité et de fonction de répartition

La fonction de probabilité, également appelée fonction de masse ou distribution de probabilité [25], d'une variable aléatoire discrète est une fonction  $f$  qui représente toutes les valeurs  $x$  possibles de la variable  $X$  et leur probabilité. la figure 2.1 représente la fonc-

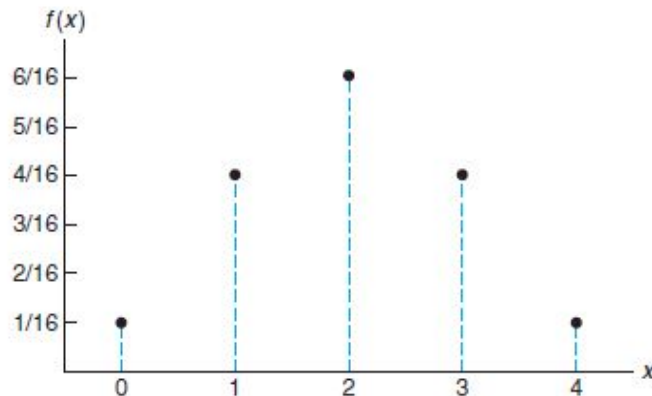


FIGURE 2.1 – Exemple d'une fonction de masse ( source [25])

tion de probabilité d'une variable aléatoire discrète  $X = 0, 1, 2, 3, 4$  dont les probabilités des valeurs sont données par :

$$\left\{ \begin{array}{l} P(X = 0) = 1/16 \\ P(X = 1) = 1/4 \\ P(X = 2) = 3/8 \\ P(X = 3) = 1/4 \\ P(X = 4) = 1/16 \end{array} \right.$$

Si  $X$  est une variable aléatoire discrète, la fonction  $f$  est une densité de probabilité de la variable  $X$  si elle respecte les conditions ci-dessous :

$$\left\{ \begin{array}{l} f(x) \geq 0 \\ \sum_x f(x) = 1 \\ P(X = x) = f(x) \end{array} \right.$$

Dans le cas d'une variable aléatoire continue, on parle de densité de probabilité (cf

figure 2.2). Les conditions pour qu'une fonction  $f$  soit la densité de probabilité de la variable  $X$  sont :

$$\begin{cases} f(x) & \geq 0 \\ \int_{-\infty}^{+\infty} f(x) dx & = 1 \\ P(a \leq X \leq b) & = \int_a^b f(x) dx \end{cases}$$

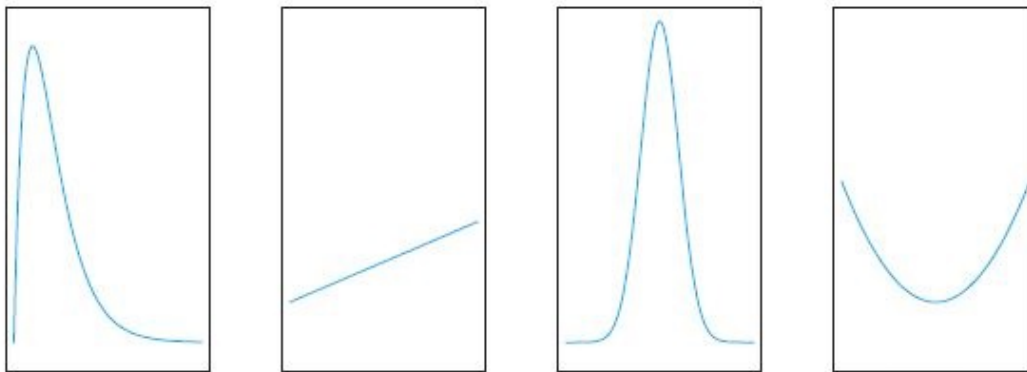


FIGURE 2.2 – Exemple d'une fonction de densité de probabilité d'une variable aléatoire continue (source [25])

### Notion de fonction de répartition

La fonction de répartition  $F(x)$  d'une variable aléatoire  $X$  de densité de probabilité  $f(x)$  est définie par [25]. Elle est définie comme étant la probabilité d'avoir une valeur de la variable  $X$  inférieure à une valeur donnée  $x$ .

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(x) dx \quad (2.1)$$

### 2.2.3 Notion de moments et d'estimateur

#### Moments

Le moment d'une variable aléatoire est un indicateur de la dispersion de cette variable. On distingue les moments à l'origine, les moments centrés, et les moments centrés réduits. Soit  $X$  une variable aléatoire continue de densité de probabilité  $f$ . Le moment à l'origine

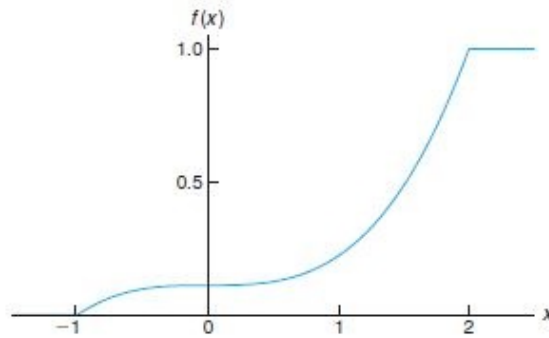


FIGURE 2.3 – Exemple d’une fonction de répartition d’une variable aléatoire continue (source [25])

d’ordre  $k$  de  $X$  noté  $Mo_k$  est défini par l’équation 2.2 :

$$Mo_k = E[X^k] = \int_{-\infty}^{+\infty} x^k f(x) dx \quad (2.2)$$

Le moment d’origine d’ordre zéro est toujours égal à 1, alors que celui d’ordre 1, est l’espérance de la distribution de probabilité de la variable  $X$  notée  $\mu_X$  (cf équation 2.3).

$$\mu_X = E[X] = \int_{-\infty}^{+\infty} x f(x) dx \quad (2.3)$$

Les moments centrés définissent la disparité des valeurs de la variable aléatoire  $X$  autour de son espérance. Le moment centré d’ordre  $k$  de  $X$  est donné par l’équation 2.4 :

$$E[(X - E[X])^k] = \int_{-\infty}^{+\infty} (x - \mu_X)^k f(x) dx \quad (2.4)$$

Le moment centré d’ordre 0 est égal à 1, celui d’ordre 1 est nul, et celui d’ordre 2 est

connu sous le nom de variance, qui est le carré de l'écart-type donné par l'équation 2.5 :

$$\sigma_X^2 = E[(X - E[X])^2] = \int_{-\infty}^{+\infty} (x - \mu_X)^2 f(x) dx \quad (2.5)$$

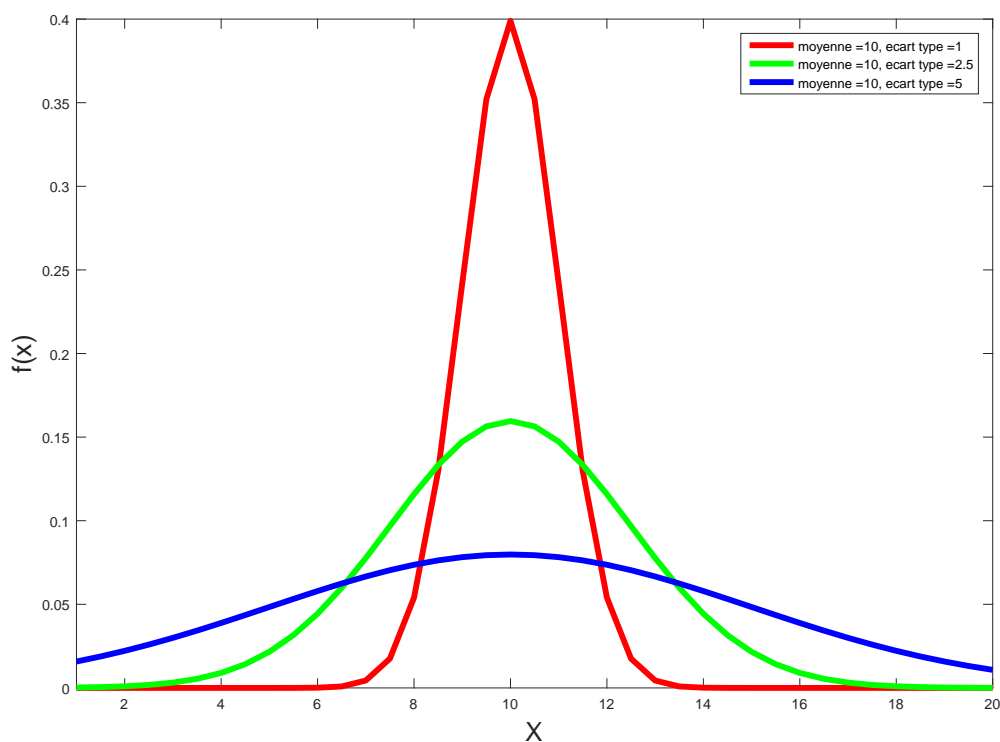


FIGURE 2.4 – Exemples de distributions avec différentes valeurs d'écart-type

Le moment centré réduit d'ordre  $k$  est le rapport du moment centré d'ordre  $k$  par l'écart type élevé à la puissance  $k$ .

$$E\left[\left(\frac{X - E[X]}{\sigma}\right)^k\right] = \frac{1}{\sigma^k} \int_{-\infty}^{+\infty} (x - \mu_X)^k f(x) dx \quad (2.6)$$

Le moment centré réduit d'ordre 1 est nul (cf équations 2.4 et 2.6), celui d'ordre 2 est égal à 1. Ceux d'ordre 3 et 4 renseignent sur les aspects de formes de la distribution de

probabilité. Ils sont connus respectivement sous les noms de coefficient d'asymétrie (ou Skewness) et d'aplatissement (Kurtosis) définis par les équations 2.7 et 2.8. Le coefficient de Skewness définit l'aspect plus ou moins symétrique de la distribution par rapport à son espérance. Le Kurtosis révèle l'aspect pointu ou plat du sommet de la distribution.

$$\lambda_{X,3} = E\left[\left(\frac{X - E[X]}{\sigma}\right)^3\right] = \frac{1}{\sigma^3} \int_{-\infty}^{+\infty} (x - \mu_X)^3 f(x) dx \quad (2.7)$$

$$\lambda_{X,4} = E\left[\left(\frac{X - E[X]}{\sigma}\right)^4\right] = \frac{1}{\sigma^4} \int_{-\infty}^{+\infty} (x - \mu_X)^4 f(x) dx \quad (2.8)$$

### Estimation des moments

Considérons une variable aléatoire  $X$  résultante d'une expérience d'ingénierie (électrique, mécanique...). Cette expérience ne peut être répétée infiniment. De ce fait, on ne dispose que d'un nombre fini de valeurs du paramètre  $X$ . Dans un tel cas, comment conclure sur les caractéristiques (espérance, écart type...) de la population de  $X$  en ne disposant que d'un échantillon fini de valeurs de  $X$  appelé population ?

L'idée est de calculer un point qui permette d'approximer la valeur d'un paramètre d'une population statistique. Par exemple, la moyenne ou la médiane d'un échantillon de valeurs d'une variable aléatoire  $X$ , de taille  $n$ , peut être utilisée pour estimer l'espérance de  $X$  [25]. L'estimateur ainsi obtenu ne permet pas d'avoir une valeur précise de grandeur estimée sans une certaine erreur.

Soit un estimateur  $\hat{\Theta}$  dont la valeur  $\hat{\theta}$  est une estimation d'un paramètre inconnu  $\theta$  de la distribution de la variable  $X$ . L'estimateur est dit non biaisé si la moyenne de ses estimations est égal au paramètre  $\theta$  estimé [25].

$$\mu_{\hat{\Theta}} = E(\hat{\Theta}) = \theta \quad (2.9)$$

Soit une variable aléatoire  $X$ , de moyenne  $\mu_X$  et d'écart type  $\sigma_X$  inconnus. On dispose d'un échantillon de taille  $N$  de la variable  $X$ . En appliquant le principe de l'équation 2.9,

on peut démontrer que  $\bar{X}$  (cf équation 2.10) et  $S^2$  (cf équation 2.11) sont respectivement des estimateurs non biaisés de l'espérance et de la variance ( $\sigma_X^2$ ) de  $X$ .

$$\bar{X} = \frac{1}{N} \sum_{i=1}^{i=N} X_i \quad (2.10)$$

$$S^2 = \frac{1}{n-1} \sum_{i=1}^N (X_i - \bar{X})^2 \quad (2.11)$$

Bien que  $S^2$  soit un estimateur sans biais de  $\sigma^2$ ,  $S$  n'est pas un estimateur non biaisé de  $\sigma$ . Il existe un biais qui peut être quand même négligé dans le cas d'un nombre d'échantillons  $N$  élevé [25].

Il peut arriver qu'un paramètre  $\theta$  d'une variable aléatoire ait plusieurs estimateurs non biaisés  $\hat{\Theta}_i$ . Dans ce cas, l'estimateur le plus efficace est celui à plus faible variance [26].

## 2.3 Modélisation probabiliste de la production et de la consommation d'électricité

### 2.3.1 État de l'art

Les modèles probabilistes de consommation et de production d'électricité dans la littérature sont historiquement basés sur une approximation de la distribution de puissance active par une loi de probabilité qui est généralement de type normale pour la consommation, Weibull pour la production éolienne et bêta pour la production photovoltaïque [27]. La partie qui suit présente une description des différentes variantes de ces modèles probabilistes de production et de consommation d'électricité trouvées dans la littérature.

#### Modélisation de la production

Alors que la puissance active des générateurs conventionnels est généralement modélisée par une distribution de Bernoulli [28], on retrouve plus de diversités dans les modèles

probabilistes de production d'électricité d'origine renouvelable. Les modèles probabilistes des unités de production d'électricité sont de plus en plus utilisés avec l'avènement de la production renouvelable intermittente. C'est dans ce cadre que les travaux décrits dans [29] ont développé un modèle stochastique de la production photovoltaïque en supposant que la puissance PV produite à chaque instant suit une loi bêta.

Au lieu de considérer la puissance de sortie comme paramètre aléatoire, les travaux décrits dans [30] proposent un modèle qui calcule de la puissance photovoltaïque produite  $P$  en fonction de l'irradiance  $r$ , de la surface des panneaux  $A$ , du rendement  $\eta$  et d'une erreur sur la prévision de la température  $\Delta T$  pondéré par une constante de température  $k$  fournie par le constructeur (cf équation 2.12).

$$P = rA\eta(1 - k\Delta T) \quad (2.12)$$

L'aspect aléatoire est ici caractérisé par l'irradiance  $r$ , modélisée par une loi bêta et l'erreur sur la prévision de température modélisée par une loi normale centrée. La loi de probabilité de la puissance  $P$  est alors obtenue par le produit de la loi de probabilité de l'irradiance et de celle de l'erreur sur la température, avec l'hypothèse que ces deux variables sont indépendantes. Même si cette approche considère plus de paramètres dans le modèle, elle omet la nébulosité qui est un facteur aussi déterminant dans la production du panneau PV. Les auteurs dans [27] proposent une densité de probabilité de la puissance de PV produite en fonction de la densité de probabilité du rapport entre l'irradiance mesurée au sol et l'irradiance solaire maximale théorique.

Ces modèles nécessitent l'acquisition de données météorologiques (irradiance, température, nébulosité...) sur le site d'étude. En l'absence de données, les travaux de [31] permettent néanmoins une modélisation stochastique de la production basée sur une représentation des passages de nuages. Il s'agit d'un modèle de variation spatio-temporelle de la couverture nuageuse qui ne nécessite pas de données d'historique d'irradiance. Il est basé sur une génération de profils aléatoires d'ombre nuageuse déterminés à partir d'informations de distribution de taille de nuages. Toujours dans ce cadre, les travaux



dans [32] ont proposé un modèle d'ombre de nuages au sol à l'aide de fractales et de l'algorithme "midpoint displacement".

En outre, les travaux dans [33] ont proposé un modèle de génération d'ombres nuageuses au sol calibré grâce à des données de production mesurées. Une ombre nuageuse est modélisée par 2 ou 5 cercles concentriques selon que l'étude porte sur des petites installations PV domestiques ou bien des installations de grande taille de plusieurs centaines de kW à quelques MW. La puissance de PV produite est considérée égale à sa puissance théorique en ciel clair si l'ensemble du panneau se trouve en dehors du cercle d'extrémité. Sinon, la puissance produite est un pourcentage de la puissance théorique maximale installée, pourcentage qui varie de 80% à 30% au fur et à mesure que la position du panneau est proche du centre de l'ombre nuageuse. Les diamètres des cercles modélisant les nuages, le taux de couverture nuageuse de la zone et la densité de l'ombre sont calibrés en comparant la production simulée à celle mesurée sur des sites de production réels.

Les travaux dans [34] proposent un modèle stochastique de la puissance produite par les panneaux PV en fonction des conditions météorologiques, en distinguant trois situations : journée avec ciel clair, journée avec ciel nuageux et journée pluvieuse. Il propose une méthode de calcul pour chaque type de jour mais ne propose pas une méthode pour déterminer la probabilité d'occurrence d'un type de jour donné.

## **Modélisation de la consommation**

Les travaux dans [35] et [27] proposent une modélisation de la puissance consommée par une loi normale estimée à partir de distribution de mesures de consommation. La corrélation entre les puissances consommées pour chaque pas de temps peut être prise en compte en utilisant des distributions normales jointes ou un modèle d'auto régression (AR) comme le montre les travaux dans [36].

Les auteurs dans [37] trouvent que l'utilisation d'une loi normale pour modéliser les comportements probabilistes de la consommation est très simplificatrice. En effet, la distribution des valeurs de puissances consommées à une heure donnée ne suit pas toujours parfaitement une loi normale. Ils proposent un modèle basé sur la loi  $\chi^2$ . Une variable est

dite distribuée suivant une loi  $\chi^2$  à  $k$  degrés de liberté si elle est la somme des carrés de  $k$  variables aléatoires indépendantes distribuées suivant une loi normale centrée réduite. La forme de cette distribution dépend de la valeur de  $k$ . Plus  $k$  est grand, plus elle tend vers une loi normale. En ajustant la valeur de  $k$ , il est ainsi possible de mieux approximer une distribution de puissance consommée à une heure donnée [37].

Dans [38], un modèle probabiliste de charge basé sur le clustering est proposé. Des données de consommations journalières à l'horizon d'une année sont réparties en quatre saisons, chaque saison étant représentée par 2 clusters (jour ouvré et week-end). Chacun des huit clusters est ainsi représenté par un profil type de consommation défini par le centre du cluster. La différence entre le profil type d'un cluster et les autres profils de ce cluster est calculée puis approximée par une densité de probabilité de Weibull. Ainsi, chaque cluster est défini par un profil type (centres) et la densité de probabilité de l'erreur autour de ce profil type.

### 2.3.2 Modèle production PV

Deux modèles de production PV ont été développés dans le cadre de cette thèse. Ils sont appliqués sur des données de nébulosité fournies par l'entreprise REUNIWATT dans le cadre du projet SCE. Ces données sont issues de prévisions de la nébulosité sur le site de Carpentras (France). Les prévisions sont faites sur un horizon de trois heures et sur une durée allant du 1er Janvier 2011 au 31 Décembre 2013.

#### Estimation de la densité de probabilité

La puissance de sortie d'une installation PV de surface  $S$ , à un instant  $t$  est donnée par l'équation 2.13

$$P_{pv}(t) = (1 - c(t)) \times I(t) \times S \times \eta \quad (2.13)$$

où  $I(t)$  désigne l'irradiance (en  $W/m^2$ ) à l'instant  $t$ .  $S$  et  $\eta$  désignent respectivement la surface totale des panneaux installés en ( $m^2$ ) et leur rendement. L'aspect aléatoire de la production PV est modélisé par le paramètre aléatoire  $c(t)$  qui désigne l'indice de nébulosité en un instant  $t$ . Le modèle de la production photovoltaïque peut être vu comme une combinaison d'un facteur déterministe - la puissance théorique installée - et d'un

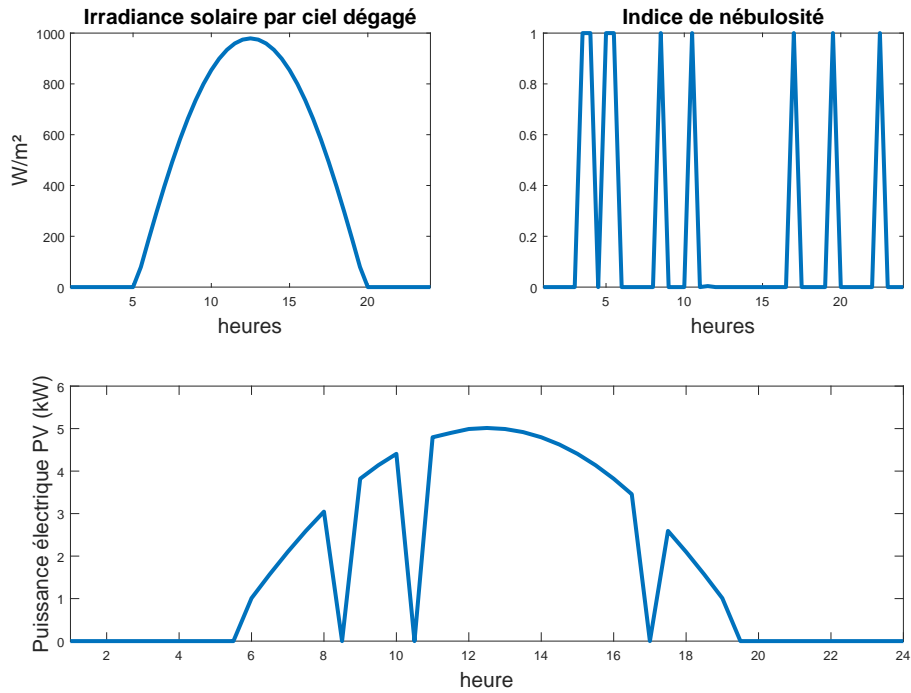


FIGURE 2.5 – Puissance électrique générée par un profil d’irradiance et de nébulosité donné

facteur probabiliste permettant de modéliser l’incertitude de la production (figure 2.5). La puissance théorique installée est fonction de la surface des panneaux, du rendement et de l’irradiance pour une inclinaison donnée. Le facteur probabiliste noté  $c(t)$  dans l’équation 2.13 est un paramètre aléatoire continu, compris entre 0 et 1, qui modélise l’indice de nébulosité. Pour un ciel clair,  $c$  est proche de 0, ce qui implique une production réelle proche de la puissance théorique maximale productible à cet instant. Par contre un ciel couvert est traduit par une valeur de  $c$  proche de 1, d’où une production réelle proche de zero. La loi de probabilité de l’indice de nébulosité a été estimée à partir de la distribution de données de nébulosité mesurées sur une année. Les données sont réparties par saison puis regroupées par heure de la journée. La forme de la distribution, pour chaque pas horaire, suit une loi bêta dont les paramètres varient légèrement d’une heure à une autre.

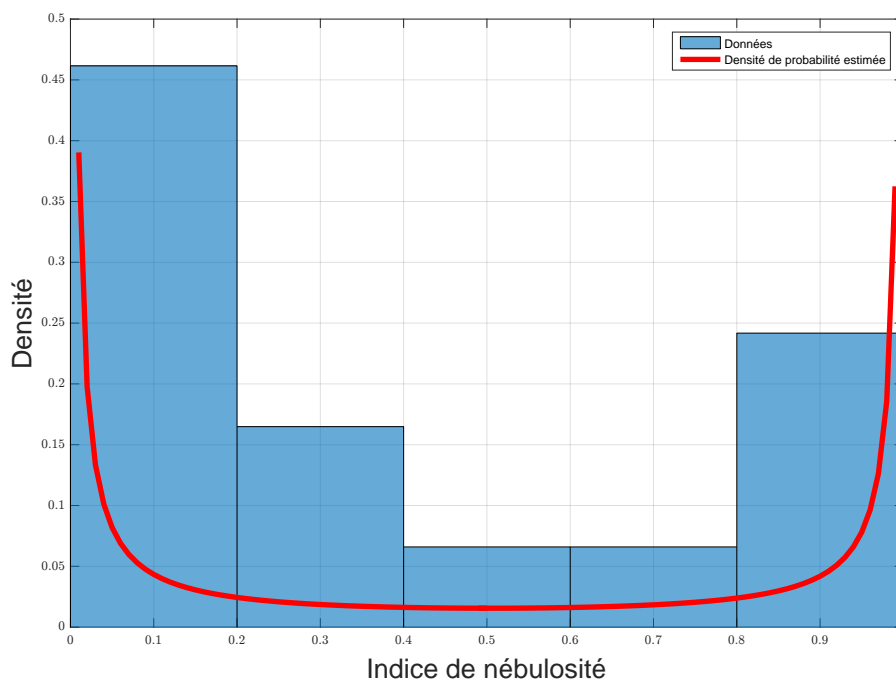


FIGURE 2.6 – Distribution de probabilité de l'indice de nébulosité

### Modèle de la nébulosité basé sur le clustering

Le modèle de nébulosité présenté ci dessus, qui consiste à approximer la distribution des mesures de nébulosité par une loi bêta ne permet pas de prendre en compte la corrélation entre les pas de temps consécutifs. Pour pallier ce problème, un nouveau modèle sera étudié. D'abord, les données de nébulosité seront réparties en fonction des saisons de l'année. Pour chaque saison, une technique de clustering sera appliquée aux différents profils de nébulosité journaliers de la saison, ceci afin de regrouper les jours ayant le même profil de nébulosité. Les figures 2.8, 2.7, 2.9 et 2.10 montrent les quatre saisons de l'année avec les profils types de nébulosité des différents clusters et leur distribution, c'est à dire le nombre de jours observés dans chaque cluster.

La saison d'été (cf figure 2.7) est caractérisée principalement par quatre types de jours selon le profil de nébulosité. Le profil le plus représentatif, que l'on retrouve sur plus de 60% des jours de la saison est le profil de nébulosité proche de 0 sur toute la journée (cluster 1). On constate que les jours de ciel clair sont les plus présents dans cette saison. Les trois autres profils types de nébulosité représentés par les clusters 2, 3 et 4 indiquent

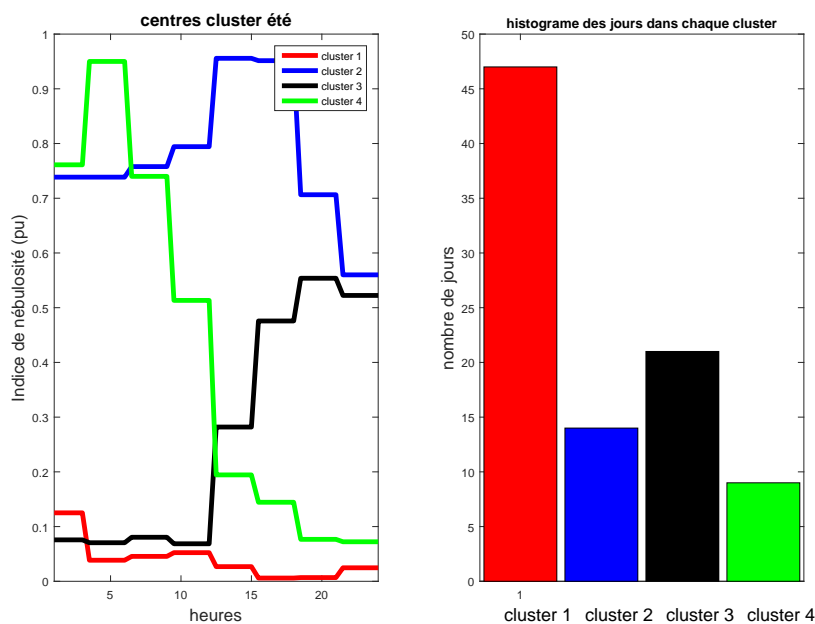


FIGURE 2.7 – Types de profils journaliers de la nébulosité en été et leur distribution

TABLE 2.1 – Récapitulatif des probabilités d’occurrence des profils type de nébulosité selon la saison

Type de jour	Été	Hiver	Printemps	Automne
Ensoleillée	62%	15.56%	34.07%	38.78%
Nuageuse	10%	47.78%	17.58%	37.18%
Ensoleillée matin	16%	16.67%	25.27%	12.9%
Nuageuse matin	9%	20%	23.08%	17.2%

les journées avec un ciel couvert toute la journée (10% des jours étudiés), les journées de transition avec ciel clair la matinée et présence de nuages l’après midi (16% des jours étudiés) et les journées avec ciel couvert le matin s’éclaircissant l’après midi (9% des jours étudiés).

La saison d’hiver (cf figure 2.8) est aussi caractérisée par les quatre types de jours que l’on retrouve en été. Les jours les plus représentatifs sont ceux du cluster 2, avec plus de 45%. Ces jours présentent un profil de nébulosité proche de 1 toute la journée, c’est à dire une journée avec un ciel pratiquement couvert. Les journées avec un profil ensoleillé sont celles du cluster 1 et présentent 15.56% de l’échantillon étudié. Les clusters 3 et 4 représentent respectivement les journées ensoleillées le matin puis nuageuses l’après midi et vice et versa.

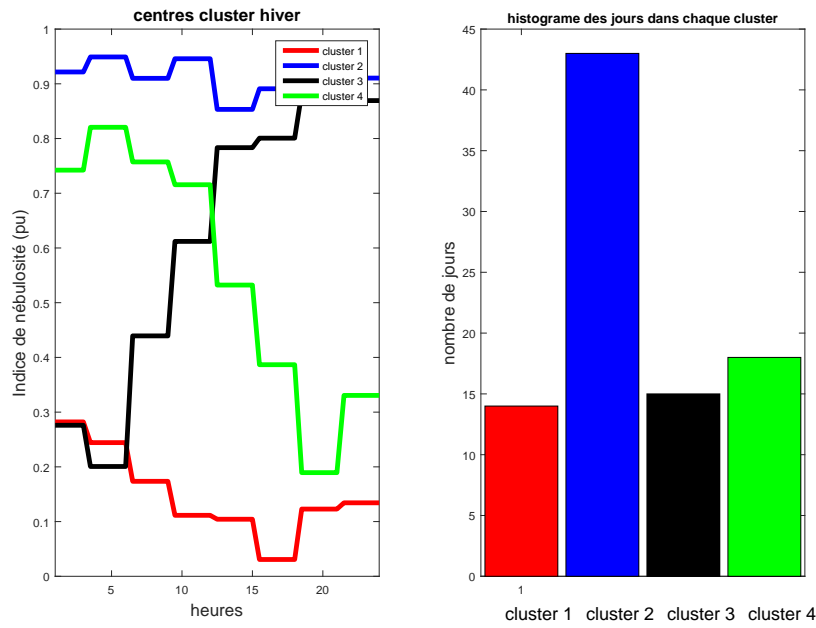


FIGURE 2.8 – Types de profils journaliers de la nébulosité en hiver et leur distribution

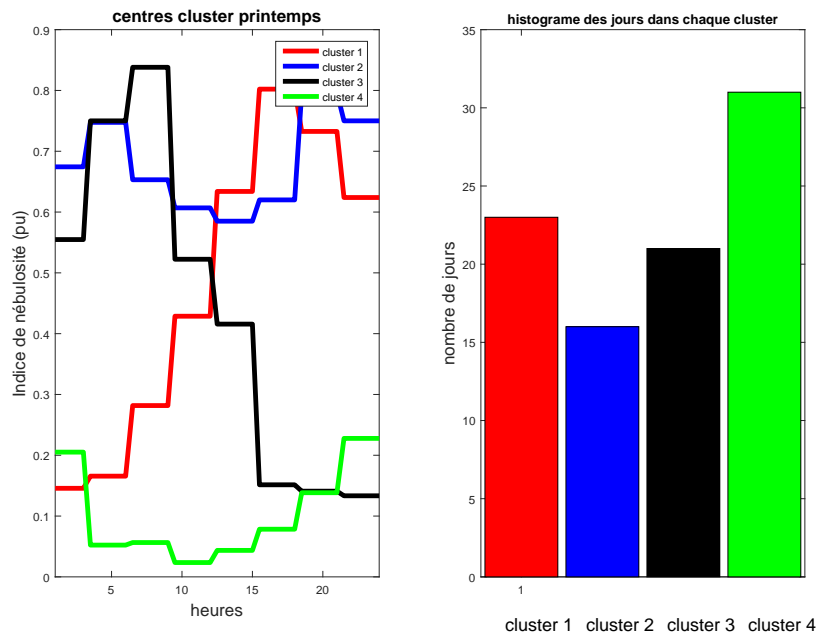


FIGURE 2.9 – Types de profils journaliers de la nébulosité au printemps et leur distribution

Au printemps (cf figure 2.9), on retrouve les 4 profils types (ensoleillé, nuageux, ensoleillé/nuageux et nuageux/ensoleillé) avec des proportions plus équilibrées et valant

respectivement 34, 17, 25 et 23%.

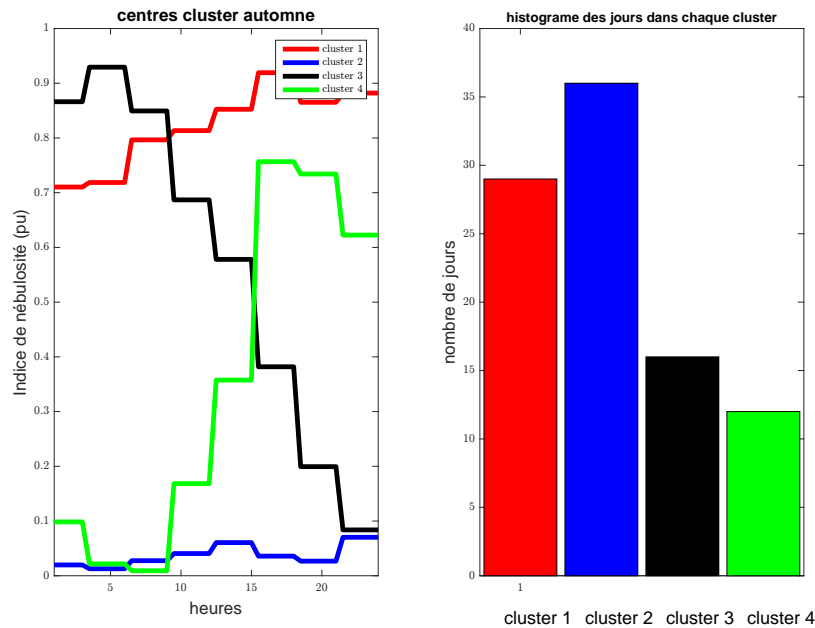


FIGURE 2.10 – Types de profils journaliers de la nébulosité en automne et leur distribution

Les journées de la saison d’automne (cf figure 2.10) sont aussi réparties selon les quatre profils types de nébulosité évoqués dans les autres saisons. Par contre, durant cette saison, il y a pratiquement autant de chance d’avoir une journée nuageuse (31%) qu’une journée ensoleillée (38%). Il en est de même pour les types de journée au cours desquelles la nébulosité s’inverse entre le matin et l’après midi.

## Échantillonnage

Soit  $T$  la variable aléatoire discrète qui indique le type de jour selon le profil de nébulosité. Elle est définie par :

- $T = 1$  si le jour échantillonné appartient au cluster 1 de la saison
- $T = 2$  si le jour échantillonné appartient au cluster 2 de la saison
- $T = 3$  si le jour échantillonné appartient au cluster 3 de la saison
- $T = 4$  si le jour échantillonné appartient au cluster 4 de la saison

On désigne respectivement par  $p_1$ ,  $p_2$ ,  $p_3$  et  $p_4$  les probabilités qu’un jour donné d’une

saison donnée soit dans les clusters 1, 2, 3 ou 4. Donc, on peut écrire :

$$Pr(T = 1) = p_1 \quad (2.14)$$

$$Pr(T = 2) = p_2 \quad (2.15)$$

$$Pr(T = 3) = p_3 \quad (2.16)$$

$$Pr(T = 4) = p_4 \quad (2.17)$$

$$(2.18)$$

Les probabilités  $p_i$ , pour  $i \in \{1, 2, 3, 4\}$ , sont obtenues en calculant la proportion de jours dans le cluster  $i$  sur le nombre total de jours de la saison. Pour obtenir  $N$  échantillons de la variable aléatoire  $T$ , on se base sur la méthode d'échantillonnage proposée dans [39]. On échantillonne  $N$  valeurs d'une variable aléatoire  $U$  uniformément distribuée dans l'intervalle  $[0, 1]$ . Ensuite l'intervalle  $[0, 1]$  est subdivisé en un nombre de sous intervalles égal au nombre de valeurs possibles de la variable  $T$ , en l'occurrence 4 dans notre cas, de telle sorte que la largeur du  $i^{ieme}$  intervalle soit égale à  $Pr(T = i) = p_i$ , avec  $i \in \{1, 2, 3, 4\}$ . A chaque fois que la valeur échantillonnée de  $U$  se trouve dans un sous intervalle  $i$ , on en déduit que  $T$  est égal à  $i$ .

### 2.3.3 Modèle de la consommation

Vue d'un exploitant de réseau de distribution, la consommation électrique domestique présente, elle aussi, un caractère aléatoire que nous cherchons à représenter. Pour ce faire, nous distinguons la consommation électrique liée à la recharge des véhicules électrifiés (VE) de la consommation purement domestique, pour deux raisons : d'une part la consommation des VEs a actuellement un potentiel de croissance important, et d'autre part son placement temporel est une variable permettant d'optimiser le système. La puissance consommée en un instant  $t$  par un foyer, s'exprime alors :

$$P_{consomee}(t) = P_{ev}(t) + P_{exo}(t) \quad (2.19)$$

Avec la consommation du véhicule électrique (VE) notée  $P_{ev}$  et la consommation totale de tous les autres équipements de la maison notée  $P_{exo}$ .



## Consommation des véhicules électriques

Le modèle probabiliste des consommations de VE est basé sur une génération aléatoire des distances parcourues. La distance permet d'estimer la quantité d'énergie nécessaire qui, répartie sur un intervalle de temps, donne le profil de recharge.

Les résultats de l'Enquête Nationale Transports Déplacements (ENTD 2008) réalisée en France ont permis d'obtenir la distribution des distances parcourues par les véhicules individuels. Cette distribution a été approchée par une loi log-normal de moyenne  $\mu = 30km$  et d'écart-type  $\sigma = 0.85$  (cf figure 2.11).

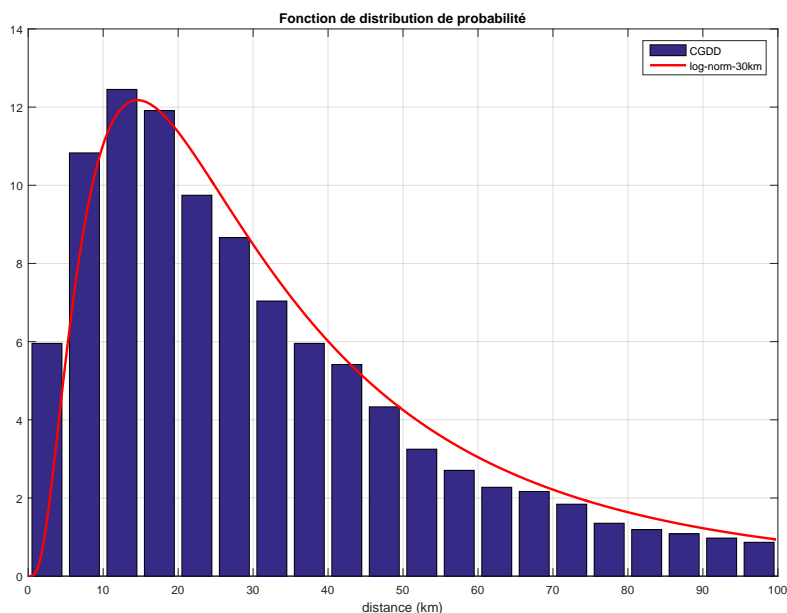


FIGURE 2.11 – Distribution de probabilité des distances parcourues par des VE

Les distances sont générées aléatoirement à partir de la distribution de probabilité présentée à la figure 2.11. L'énergie de recharge nécessaire est calculée sur une base de consommation de  $20kWh/100km$  [40] en supposant que le niveau de capacité de la batterie du véhicule est tel que la recharge peut se faire uniquement à domicile. Pour rappel, les véhicules électriques de série les plus récents ont des capacités de batteries de  $40kWh$ .

La puissance de recharge  $P_{rech}$ , l'énergie nécessaire  $E$  et l'intervalle de temps  $[t1 t2]$

durant lequel la recharge devra être effectuée sont liés par la relation 2.20.

$$E = \int_{t1}^{t2} P_{rech}(t) dt \quad (2.20)$$

## Consommation interne

### Problématique de l'échantillonnage pas à pas

Les modèles probabilistes classiques de puissance consommée proposés dans la littérature consistent à définir, pour chaque instant horaire, une loi de probabilité pour la puissance active. ces modèles restent applicables dans le cas d'une étude en mode instantané. Sur une étude en série temporelle, par exemple à l'horizon d'une journée ils présentent des limites. Supposons que l'on discrétise la journée en vingt-quatre pas de temps d'une heure. Puis, pour chaque heure la puissance active consommée soit représentée par une loi de probabilité connue et propre à cette heure. Donc pour un foyer donné, nous disposons de vingt-quatre lois de probabilité représentant le caractère aléatoire de la puissance active. Si les vingt-quatre valeurs de puissance sont échantillonnées indépendamment à partir des distributions horaires, le profil de consommation résultant risque d'être complètement irréaliste.

Pour pallier ce problème, une des alternatives développées est basée sur une technique d'échantillonnage qui impose une contrainte sur la valeur échantillonnée à chaque instant. Le principe consiste à définir, pour chaque pas horaire, des valeurs limites (une maximale et une minimale) de sorte que la puissance échantillonnée reste sur une plage. Autrement dit, cela revient à définir une enveloppe, comme celle sur la figure 2.12, qui respecte la forme d'un profil type de consommation résidentielle. Avec une telle technique, les profils aléatoires échantillonnés auront une allure et des niveaux de valeurs réalistes.

Les bornes supérieures et inférieures de cette enveloppe représentent les variations maximales de la puissance consommée à une heure donnée. Il est nécessaire d'avoir une connaissance sur les habitudes de consommation des usagers résidentiels pour pouvoir définir les enveloppes. La partie qui suit présente comment sont définies les enveloppes des profils pour chaque usager.

### Le clustering pour définir les enveloppes des profils

En présence de plusieurs centaines d'usagers aux habitudes de consommation dif-

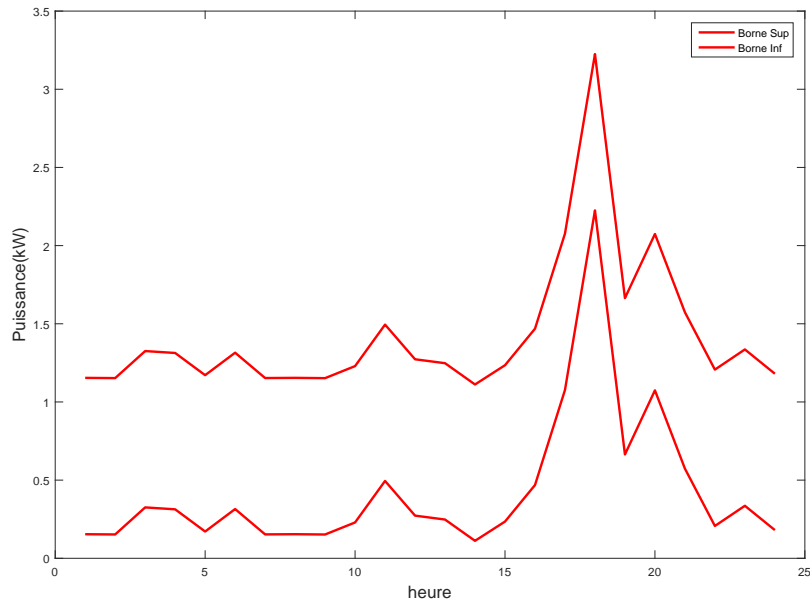


FIGURE 2.12 – Exemple d’enveloppe de profils de consommation résidentiels

férentes, il est particulièrement difficile, en simulation, de considérer une enveloppe de profils types propre à chacun d’entre eux. Le clustering permet, à partir de données de consommation de sites existants, d’identifier des groupes homogènes de consommateurs avec les mêmes habitudes. Cela permet de réduire le nombre d’enveloppes types à considérer de plusieurs centaines à une dizaine.

### Les données

Les données utilisées dans cette étude sont des consommations électriques de 6445 bâtiments recueillies en Irlande. Ces compteurs sont liés à différents types de bâtiments dont des foyers, des PME et autres. Nous nous sommes focalisés sur les usagers de type foyer. Parmi ces usagers, nous y trouvons différentes catégories socio-professionnelles selon l’âge, le statut d’emploi, le montant des revenus, les types d’appareils dans la maison, le statut familial, le nombre d’occupants, la surface de maison,...

Les mesures, effectuées par pas de trente minutes, couvrent une période de dix-huit mois. Nous avons travaillé avec les profils journaliers moyennés de 3487 foyers. Ces profils ont été calculés sur la consommation du mois de Novembre, mois durant lequel il n’y a quasiment pas de phénomène particuliers comme des congés par exemple.

## La méthode de clustering utilisée : Le K-means

L'algorithme *K – moyennes* ou *k – means* (en anglais) consiste à partitionner des données en  $kc$  groupes appelés *clusters*,  $kc$  étant un nombre entier naturel. Le critère utilisé pour réaliser cette partition consiste à minimiser la somme des carrés des distances entre les points d'un cluster et la moyenne de ce cluster (équation 2.21) [41].

Supposons  $n$  profils  $(L_1, L_2, L_3, \dots, L_n)$  représentés dans notre cas par les 3487 vecteurs de profils de consommation journalière. On cherche à partitionner ces  $n$  profils en  $kc$  ensembles  $S = \{S_1, S_2, \dots, S_{kc}\}$  avec  $kc \leq n$  en minimisant la distance entre les points à l'intérieur de chaque ensemble.

$$\underset{S}{\operatorname{argmin}} \sum_{ic=1}^{kc} \sum_{L_{jc} \in S_{ic}}^{\max} \|L_{jc} - \mu_{ic}\|^2 \quad (2.21)$$

$\mu_{ic}$  étant la moyenne des points à l'intérieur du cluster  $S_{ic}$ .

Le nombre optimal de cluster  $kc$  est déterminé en fonction de l'inertie. Pour un nombre de cluster  $kc$ , on calcul la somme des distances entre chaque point d'un cluster et son centre. La somme totale de ces distances représente l'inertie (équation 2.22). La figure 2.13 montre la variation de cette inertie en fonction du nombre de clusters considéré. Le nombre optimal de cluster  $kc$  correspond à celui à partir duquel la variation de l'inertie est quasi linéaire. Dans notre cas, le nombre optimal de cluster est  $Kc = 5$ .

$$\operatorname{Inertie} = \sum_{ic=1}^{kc} \sum_{L_{jc} \in S_{ic}}^{\max} |L_{jc} - C_{ic}| \quad (2.22)$$

## Résultats du clustering et échantillonnage des profils aléatoires

La figure 2.15 montre la répartition des profils dans les 5 clusters. Pour chaque cluster, les bornes supérieure et inférieure de l'enveloppe des profils (figure 2.14) sont déterminées respectivement, à tout instant, par la puissance maximale et minimale parmi les éléments du cluster. Soient  $kc$  clusters et  $n_{ic}$  le nombre d'éléments  $V$  dans chaque cluster  $ic$  où  $ic \in \{1, 2, \dots, kc\}$ , alors les bornes inférieure et supérieure de l'enveloppe de chaque cluster

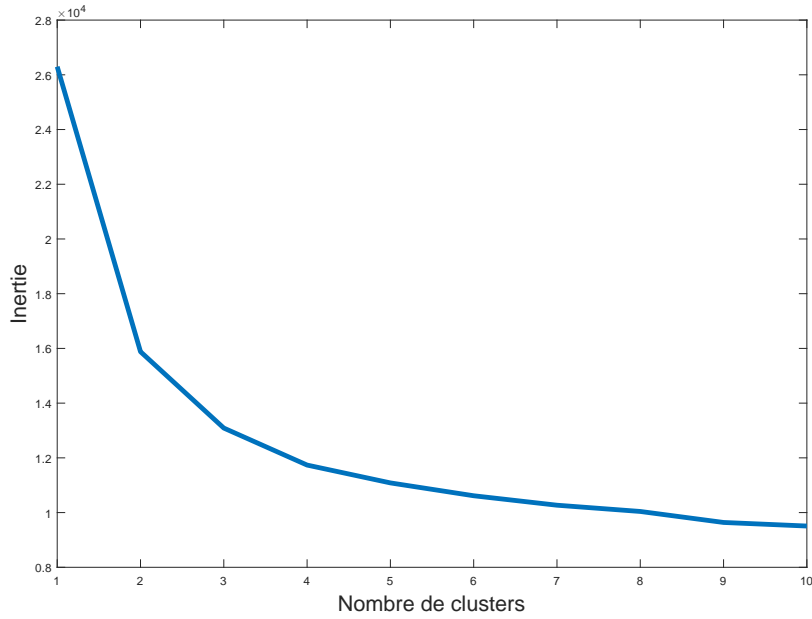


FIGURE 2.13 – Variation de l’inertie en fonction du nombre de cluster  $k$

$ic$  pour chaque instant  $h$  sont données respectivement par les équations 2.23 et 2.24 :

$$Borne_{inf}^{ic}(h) = \min(L_1(h), L_2(h), \dots, L_{n_{ic}}(h)) \quad (2.23)$$

$$Borne_{sup}^{ic}(h) = \max(L_1(h), L_2(h), \dots, L_{n_{ic}}(h)) \quad (2.24)$$

L’échantillonnage des profils à l’intérieur d’une enveloppe se fait suivant l’équation 2.25 :

$$P(h) = Borne_{inf}(h) + r \times (Borne_{sup}(h) - Borne_{inf}(h)) \forall h \quad (2.25)$$

La puissance horaire  $P(h)$  est échantillonnée en fonction des bornes limites de l’enveloppe et d’un nombre aléatoire  $r$  uniformément distribué sur l’intervalle  $[0, 1]$ . Avec cette approche, les profils échantillonnés respectent les limites d’un profil de consommation réelle (figure 2.16). Par contre, ils sont assujettis à de fortes variations qui ne sont pas représentatives d’une consommation électrique domestique, car cette méthode ne prend pas en compte la corrélation entre les puissances dans le temps.

La deuxième méthode développée consiste à considérer le centre d’un cluster comme le

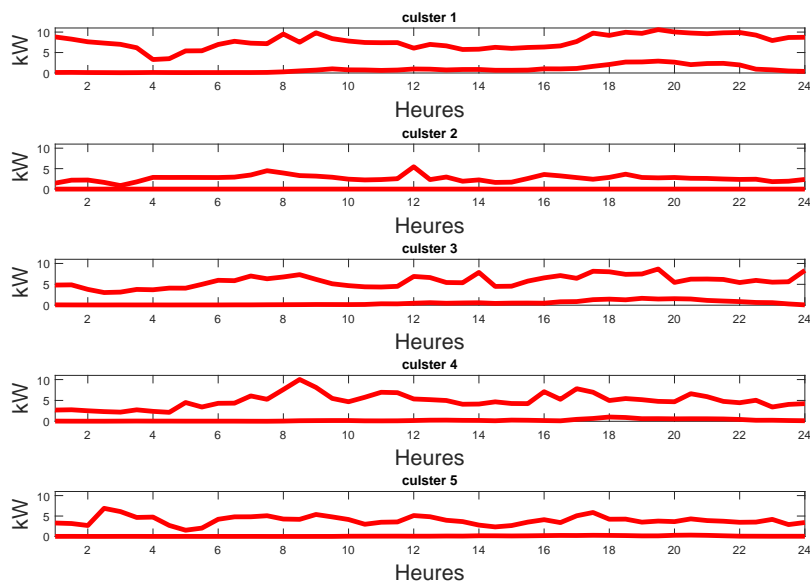


FIGURE 2.14 – Enveloppes des clusters

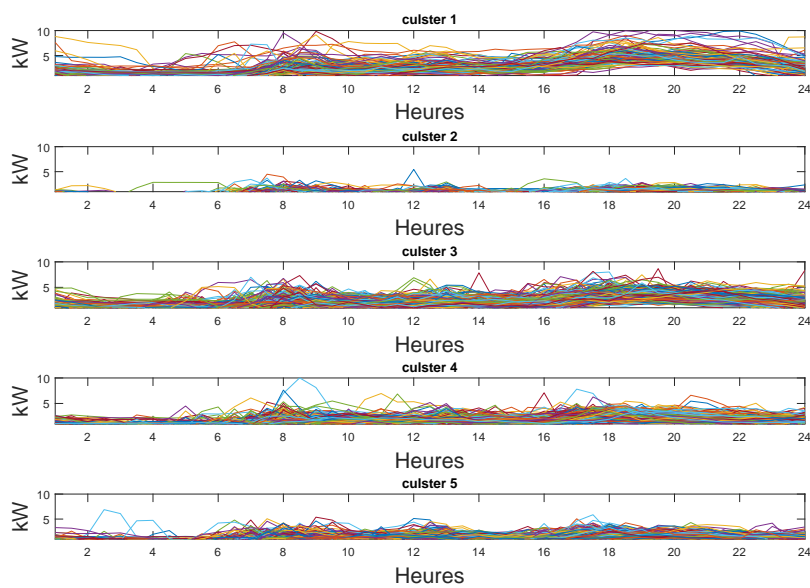


FIGURE 2.15 – Répartition des profils de consommation journalière dans les clusters

profil type de l'ensemble des foyers appartenant à ce cluster 2.17. Le centre est représentatif des profils de consommation dans un cluster donné et ne présente pas un caractère

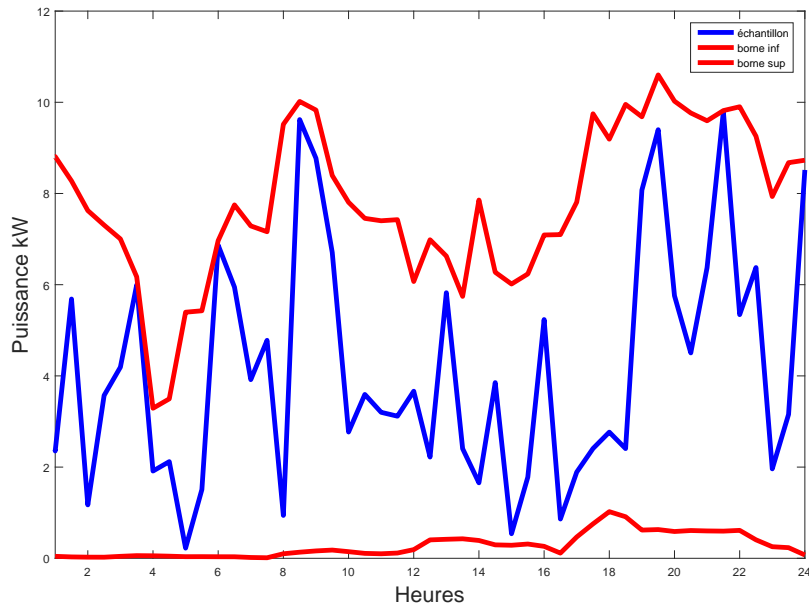


FIGURE 2.16 – Exemple de profil aléatoire échantillonné suivant une enveloppe

oscillatoire contrairement à la méthode précédente.

L'aspect aléatoire de la consommation est modélisé par une variable aléatoire discrète  $M$  qui prend la valeur  $mc$  quand le profil échantillonné est le centre du cluster  $mc$ . Autrement dit,  $M$  définit l'appartenance d'un foyer à l'un des cinq clusters. La probabilité d'appartenance à un cluster donné est déterminée par le rapport entre la taille du cluster et la taille totale des clusters (figure 2.18). Le même principe d'échantillonnage que celui présenté dans le cas du profil type de nébulosité sera appliqué.

## 2.4 Modélisation des composants du réseau

Le réseau de distribution basse tension est déséquilibré. De ce fait, l'étude d'un tel réseau nécessite, pour chaque élément, la modélisation des trois phases. Les schémas équilibrés monophasés classiques ne peuvent être utilisés dans ce cas. Cette partie présente la modélisation des équipements du réseau en mode déséquilibré pour une étude en régime permanent.

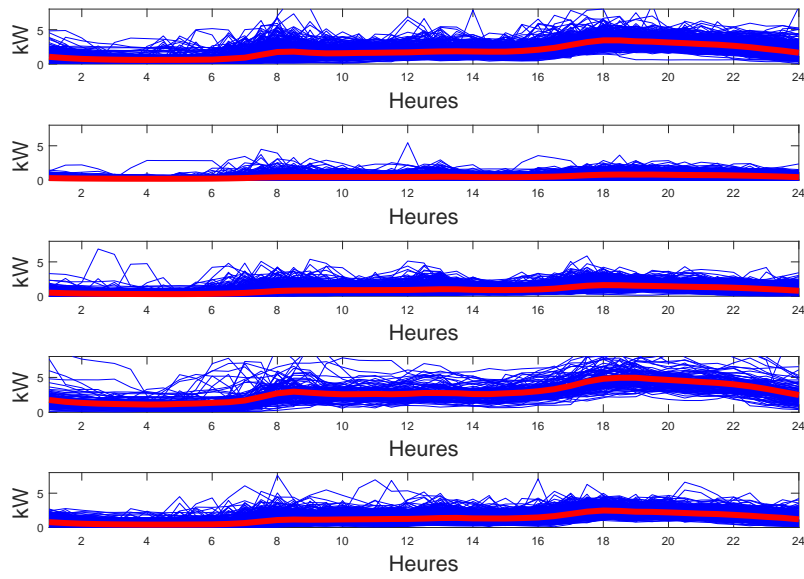


FIGURE 2.17 – Clusters avec leur centres

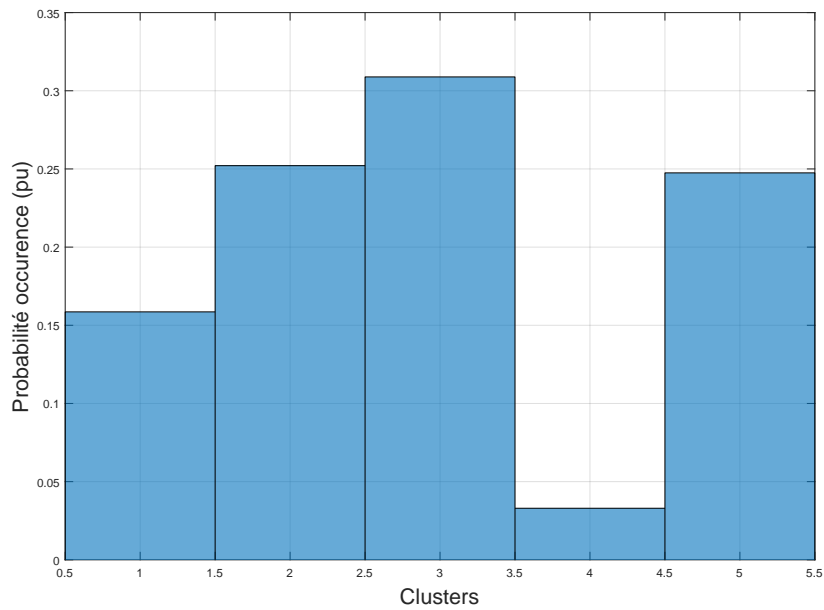


FIGURE 2.18 – Distribution des profils étudiés dans les différents clusters

### 2.4.1 Les transformateurs

On retrouve généralement deux types de transformateurs sur le réseau de distribution :



- Ceux que l'on retrouve au niveau des postes sources. Ils permettent de faire le lien entre le réseau de répartition (90, 63 ou 45 kV) et le réseau de distribution moyenne tension [1].
- Les transformateurs HTA/BT des postes de distribution qui sont chargés d'abaisser la tension jusqu'au niveau d'alimentation des consommateurs basse tension.

Ces transformateurs triphasés sont connectés à des charges de natures différentes qui peuvent parfois être une combinaison de charges triphasées (c'est le cas des moteurs triphasés) et monophasées (essentiellement les consommateurs domestiques). De ce fait, il existe différentes manières de coupler le transformateur au réseau. Le type de couplage utilisé est un paramètre essentiel à prendre en compte dans l'étude des réseaux électriques. On distingue généralement :

- **Le couplage Triangle - Étoile avec neutre à la terre** : Il est généralement utilisé au niveau des postes de distribution alimentant des départs composés de 4 lignes. Le couplage étoile au secondaire permet aussi d'alimenter des charges monophasées réparties entre les trois phases comme c'est le cas en France au niveau des transformateurs HTA/BT.
- **Le couplage Étoile - Triangle** : Généralement utilisé pour l'alimentation de combinaisons de charges triphasées.
- **Le couplage Étoile - Étoile avec neutre à la terre** : Utilisé pour l'alimentation de charges monophasées et triphasées dans les systèmes de distribution à 4 conducteurs (1 neutre et 3 phases).
- **Le couplage Triangle - Triangle** : Utilisé dans les systèmes de distribution à trois conducteurs pour l'alimentation des charges triphasées .
- **Le couplage Étoile ouvert - Triangle ouvert** : Il est souvent utilisé lorsque l'on est en présence de charges triphasées.

La figure 2.19 offre une représentation générale des courants et tensions aux cotés primaires et secondaire d'un transformateur triphasé. L'idée est de définir une relation généralisée entre les grandeurs électriques (courants et tensions) du côté secondaire et celles du côté primaire. Cette relation est définie par les matrices généralisées. Elles permettent d'exprimer les tensions et courants du primaire en fonction de ceux du secondaire comme le montrent les équations 2.26 et 2.27 [4].

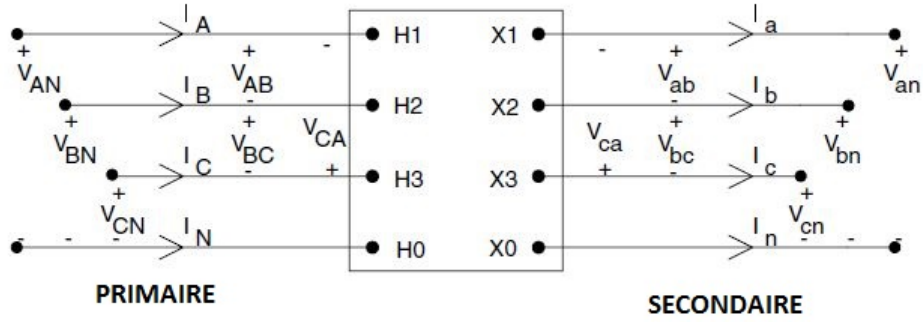


FIGURE 2.19 – Représentation des tensions et courants aux cotés primaire et secondaire du transformateur (source [4])

$$\begin{bmatrix} V_{AN} \\ V_{BN} \\ V_{CN} \end{bmatrix} = \begin{bmatrix} a_t \\ b_t \end{bmatrix} \begin{bmatrix} V_{an} \\ V_{bn} \\ V_{cn} \end{bmatrix} + \begin{bmatrix} I_a \\ I_b \\ I_c \end{bmatrix} \quad (2.26)$$

$$\begin{bmatrix} I_A \\ I_B \\ I_C \end{bmatrix} = \begin{bmatrix} c_t \\ d_t \end{bmatrix} \begin{bmatrix} V_{an} \\ V_{bn} \\ V_{cn} \end{bmatrix} + \begin{bmatrix} I_a \\ I_b \\ I_c \end{bmatrix} \quad (2.27)$$

Les termes  $V_{AN}$ ,  $V_{BN}$  et  $V_{CN}$  représentent respectivement les tensions entre ligne et neutre des phases A, B, et C du côté primaire tandis que  $V_{an}$ ,  $V_{bn}$  et  $V_{cn}$  font référence aux tensions phase-neutre au secondaire du transformateur. De même  $I_A$ ,  $I_B$ ,  $I_C$  sont les courants des phases du côté primaire et  $I_a$ ,  $I_b$ , et  $I_c$  sont ceux des phases du côté secondaire.  $[a_t]$ ,  $[b_t]$ ,  $[c_t]$  et  $[d_t]$  sont les matrices généralisées du transformateur.

## 2.4.2 Les lignes et câbles

En fonction de la nature des études et des données disponibles, on distingue généralement trois types de modèles de lignes aériennes et souterraines des réseaux de distribution : le modèle en Pi, le modèle modifié et le modèle approximé.

## 1. Le modèle en Pi

Ce modèle considère les matrices d'impédances  $[Z_{abc}]$  et d'admittances de shunt  $[Y_{abc}]$  des phases. Ces dernières sont des matrices  $(3 \times 3)$  renfermant les informations sur les caractéristiques et les distances entre les conducteurs. Pour chaque phase,

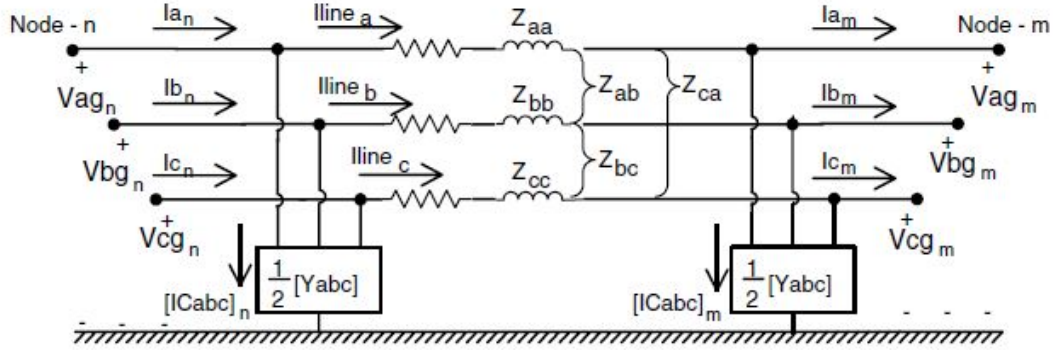


FIGURE 2.20 – Modèle en Pi d'une ligne de distribution (source [4])

la tension phase-neutre à l'extrémité  $n$  de la ligne est exprimée en fonction de la tension et du courant à l'extrémité  $m$  par l'intermédiaire des matrices généralisées (équation 2.28).

$$[V_{abc}]_n = [a] \cdot [V_{abc}]_m + [b] \cdot [I_{abc}]_m \quad (2.28)$$

avec :

$$\begin{aligned} [a] &= [U] + \frac{1}{2} \cdot [Z_{abc}] \cdot [Y_{abc}] \\ [b] &= [Z_{abc}] \end{aligned} \quad (2.29)$$

Sachant que  $[U]$  est une matrice identité de  $(3 \times 3)$ .

De même, le courant des phases à l'extrémité  $n$  est exprimé en fonction des tensions et courants des phases de l'extrémité  $m$  par :

$$[I_{abc}]_n = [c] \cdot [V_{abc}]_m + [d] \cdot [I_{abc}]_m \quad (2.30)$$

avec :

$$\begin{aligned} [c] &= [Y_{abc}] + \frac{1}{4} \cdot [Y_{abc}] \cdot [Z_{abc}] \cdot [Y_{abc}] \\ [d] &= [U] + \frac{1}{2} \cdot [Z_{abc}] \cdot [Y_{abc}] \end{aligned} \quad (2.31)$$

## 2. Le modèle modifié

Dans ce modèle, l'admittance de shunt est négligée (cf figure 2.21). Les matrices généralisées deviennent :

$$[a] = [U] \quad (2.32)$$

$$[b] = [Z_{abc}] \quad (2.33)$$

$$[c] = [0] \quad (2.34)$$

$$[d] = [U] \quad (2.35)$$

La valeur de l'admittance shunt est très faible dans plusieurs situations [4]. Né-

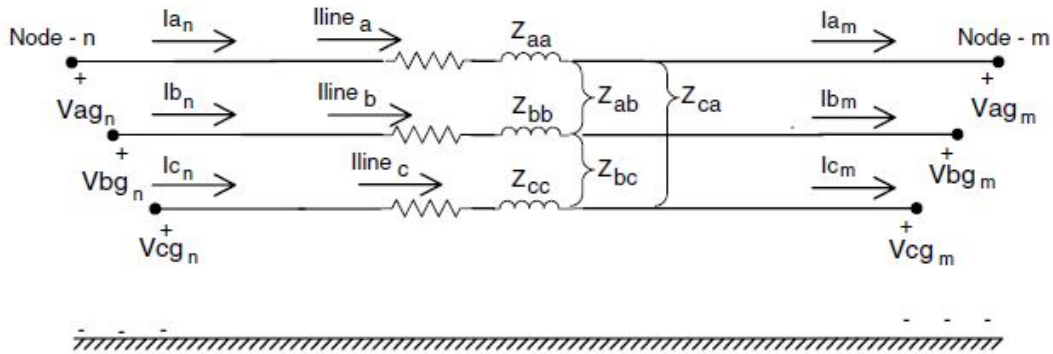


FIGURE 2.21 – Modèle modifié d'une ligne de distribution(source[4])

gliger cette admittance permet de simplifier les calculs tout en gardant un faible niveau d'erreur. Cependant, il n'est pas recommandé de négliger l'admittance dans les cas suivants : un long réseau rural, des lignes d'alimentation d'éclairage et des câbles souterrains.

### 3. Le modèle approximé

Il ignore l'admittance de shunt et suppose que les seuls paramètres connus sont les séquences directes et homopolaires des impédances de la ligne.

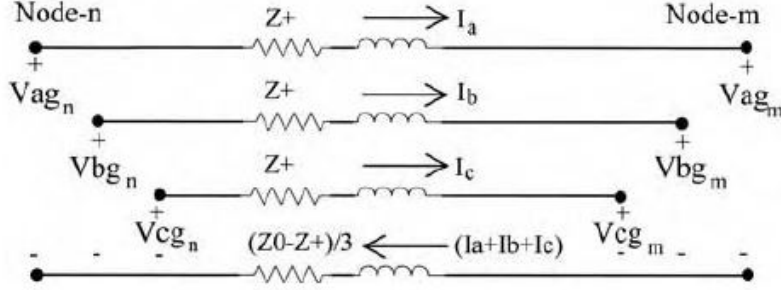


FIGURE 2.22 – Modèle approximé d'une ligne de distribution (source [4])

À partir des composantes symétriques homopolaire  $Z_0$  et directe  $Z_+$ , la matrice des séquences est définie :

$$Z_{seq} = \begin{bmatrix} Z_0 & 0 & 0 \\ 0 & Z_+ & 0 \\ 0 & 0 & Z_+ \end{bmatrix} \quad (2.36)$$

Une transformation inverse permet d'obtenir la matrice approximée des phases :

$$Z_{approx} = [A_s] \cdot [Z_{seq}] \cdot [A_s]^{-1} \quad (2.37)$$

$$Z_{approx} = \frac{1}{3} \begin{bmatrix} (2Z_+ + Z_0) & (Z_0 - Z_+) & (Z_0 - Z_+) \\ (Z_0 - Z_+) & (2Z_+ + Z_0) & (Z_0 - Z_+) \\ (Z_0 - Z_+) & (Z_0 - Z_+) & (2Z_+ + Z_0) \end{bmatrix} \quad (2.38)$$

La tension à l'extrémité  $n$  est calculée en fonction des tensions et courants à l'extrémité  $m$  :

$$[V_{abc}]_n = [V_{abc}]_m + [Z_{approx}][I_{abc}]_m \quad (2.39)$$

De là on déduit :  $[a] = 1$  et  $[b] = Z_{approx}$ .

### 2.4.3 Les charges

Les charges sont spécifiées par la puissance complexe consommée. Le calcul du courant consommé nécessite la connaissance de la tension aux bornes de la charge. C'est pour cela que les techniques de load flow itératives initialisent cette tension en la supposant égale à la tension nominale du réseau. Les charges sur le réseau de distribution sont monophasées, biphasées ou triphasées. Elles peuvent être branchées en étoile ou en triangle. On distingue généralement quatre méthodes pour les modéliser :

- Puissances active et réactive constantes (PQ)
- Courant constant (I)
- Impédance constant (Z)
- Combinaisons des trois (ZIP)

Dans le cadre de notre étude, les charges sont des consommateurs domestiques. Elles sont branchées en étoile et modélisées en PQ car les données disponibles sont celles des puissances actives, le facteur de puissance étant fixé arbitrairement à 0,93.

#### Charges connectées en étoile

La figure 2.23 illustre le branchement en étoile d'une charge. On définit la charge par les modules et phases des puissances complexes sur les phases a, b et c (équations 2.40, 2.41, 2.42) et ceux des tensions phase-neutre :  $|V_{an}| \angle \delta_a$ ,  $|V_{bn}| \angle \delta_b$  et  $|V_{cn}| \angle \delta_c$  initialement choisies.

$$|S_a| \angle \theta_a = P_a + jQ_a \quad (2.40)$$

$$|S_b| \angle \theta_b = P_b + jQ_b \quad (2.41)$$

$$|S_c| \angle \theta_c = P_c + jQ_c \quad (2.42)$$

Le modèle PQ consiste à exprimer les courants des phases de la charge en fonction

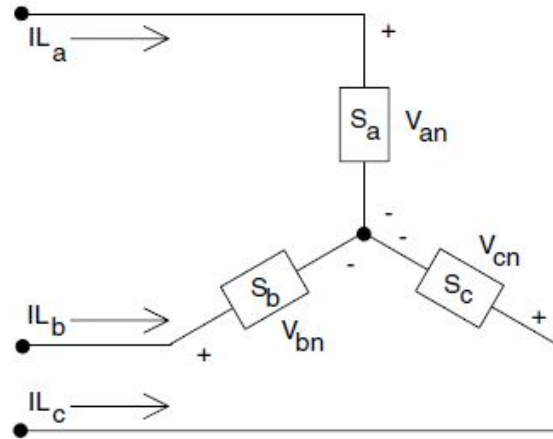


FIGURE 2.23 – Modèle de charge en étoile (source [4])

des puissances et des tensions. Ce qui donne :

$$\begin{aligned}
 Ich_a &= \left(\frac{S_a}{V_{an}}\right)^* = \frac{|S_a|}{|V_{an}|} \angle(\delta_a - \theta_a) = |Ich_a| \angle\varphi_a \\
 Ich_b &= \left(\frac{S_b}{V_{bn}}\right)^* = \frac{|S_b|}{|V_{bn}|} \angle(\delta_b - \theta_b) = |Ich_b| \angle\varphi_b \\
 Ich_c &= \left(\frac{S_c}{V_{cn}}\right)^* = \frac{|S_c|}{|V_{cn}|} \angle(\delta_c - \theta_c) = |Ich_c| \angle\varphi_c
 \end{aligned} \tag{2.43}$$

Dans le cas d'une charge monophasée ou biphasée, on applique les mêmes équations précédentes en supposant un courant nul sur la phase manquante.

#### 2.4.4 La production décentralisée

Un nœud avec de la production décentralisée peut être modélisé en mode PQ ou PV.

##### 1. Modèle PQ

Dans ce modèle on suppose que la puissance active et la puissance réactive injectées par le générateur sont constantes. De ce fait, l'unité de production est considérée comme une charge de puissance négative dans les calculs de répartition des flux.

##### 2. Modèle PV

Ce modèle suppose que le générateur fournit une puissance active constante. La puissance réactive dépend de la tension désirée au nœud d'injection.

## 2.5 Conclusion

Afin de modéliser l'incertitude de la production d'électricité d'origine photovoltaïque et de la consommation, deux méthodes ont été développées dans le cadre de cette thèse :

1. Une méthode classique basée sur l'estimation d'une loi de probabilité à partir de mesures, pratique pour les études en mode instantané.
2. L'autre basée sur l'utilisation du clustering dans le but de définir des profils types pondérés avec une certaine probabilité.

L'aspect aléatoire de la production est modélisé par l'indice de nébulosité qui est une des principales causes de l'intermittence de la production PV. L'indice de nébulosité, sur une plage horaire donnée, suit une loi bêta bornée entre 0 et 1. Cependant cette manière de modéliser le comportement de la nébulosité ne prend pas en compte la corrélation entre deux pas de temps successifs et les échantillons aléatoires tirés à partir de la loi bêta sont très oscillatoires. Ils reflètent soit un ciel complètement couvert soit un ciel complètement dégagé, et ne représentent que très rarement le cas d'une situation intermédiaire.

La deuxième méthode est basée sur une répartition des profils de nébulosité mesurés, d'une saison donnée, en quatre clusters afin de dégager des journées types de nébulosité. Ainsi les quatre types de journée qui se dégagent représentent une journée à ciel clair, une journée à ciel couvert, une journée avec la matinée claire et l'après midi nuageuse et vice et versa. Chaque type de journée est défini par une probabilité d'occurrence qui varie d'une saison à une autre. Cette méthode permet de réduire l'incertitude sur les données d'origine à quatre profils types pondérés par leur probabilité d'occurrence.

Pour mettre en exergue la composante VE qui est un élément essentiel dans la transition énergétique, la consommation est modélisée en distinguant celle du véhicule électrique des autres équipements électriques domestiques. La distance parcourue qui détermine l'énergie nécessaire, et l'instant de début de recharge sont considérés comme des paramètres aléatoires dans le modèle du profil de recharge des véhicules électrifiés. Ils sont modélisés respectivement par une loi lognormal et une loi uniforme. La puissance des bornes de recharges est supposée identique vu que nous sommes dans le cadre de recharge à domicile pour tous les véhicules.



Le consommation des autres équipements est modélisée dans un premier temps par un nombre aléatoire uniformément distribué entre deux bornes. Mais cette approche aboutit à des profils de consommation journaliers très oscillatoires. Pour pallier cet effet, une deuxième méthode basée sur le clustering est développée.

# Chapitre 3

## Méthodes de Calcul

### 3.1 Introduction

L'objectif de ce chapitre est de présenter les méthodes de calcul utilisées pour l'étude des réseaux de distribution. Les études en régime permanent sont faites à l'aide de calculs de répartition des puissances (Load Flow). Le Load flow permet de déterminer l'amplitude et le déphasage des tensions et courants sur chaque nœud et branche du réseau en fonction des puissances injectées et soutirées en ces nœuds et des paramètres physiques du réseau. L'approche classique suppose une connaissance déterministe des valeurs des données d'entrées (puissances actives, réactives,...) en chaque instant, de ce fait nous la désignerons par Load Flow Déterministe (LFD) dans la suite. Cette méthode, par son approche déterministe, simplifie la résolution du problème d'étude des réseaux car elle néglige le caractère incertain de certains paramètres comme la production intermittente d'un panneau photovoltaïque ou les instants de demande de recharge d'un véhicule électrique. En conséquence, elle ne permet pas de modéliser tous les états possibles.

Or, avec une pénétration croissante des sources de production intermittentes (cf figure 3.1) et un besoin de puissance pour la recharge des véhicules électriques également en évolution croissante, le réseau de distribution est soumis à chaque instant à une production et une consommation d'électricité à caractère incertain.

Dans le but de prendre en compte l'intermittence de la production et la variabilité de la demande dans le modèle de calcul de load flow, des méthodes de Load Flow probabilistes

Figure: 15

## Solar PV Global Capacity and Annual Additions, 2006-2016

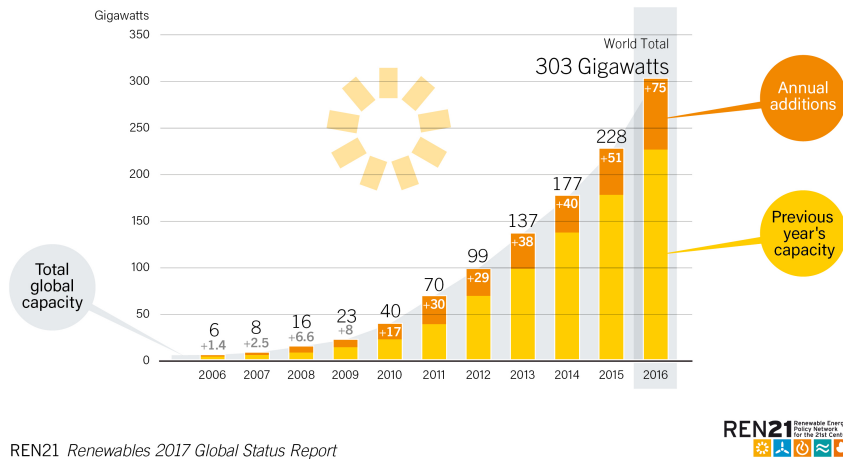


FIGURE 3.1 – Capacité globale d'énergie photovoltaïque installée entre 2006 et 2017

(LFP) sont utilisées. Ces dernières offrent la possibilité de considérer plusieurs scénarios de production et de consommation en modélisant les puissances injectées et soutirées par des variables aléatoires. Le résultat du calcul de Load flow est ainsi une variable aléatoire qui peut être représentée par une densité de probabilité ou une fonction de répartition.

La première partie de ce chapitre est consacrée à la présentation d'une méthode de load flow déterministe largement utilisée dans l'étude des réseaux électriques à structure radiale. La seconde partie évoquera les méthodes de load flow probabilistes en détaillant celles étudiées dans le cadre de cette thèse.

## 3.2 Le Load flow déterministe

Dans le cas d'un réseau de transport, des méthodes itératives de calcul de load flow sont utilisées telles que les méthodes de Newton Raphson et Gauss Siedel. L'application de ces méthodes sur les réseaux de distribution n'aboutit pas facilement à la convergence du fait du caractère radial des ces derniers [4]. Ceci a motivé le développement de méthodes de calcul, mieux adaptées aux réseaux radiales, comme par exemple celle dite de balayage

amont-aval connue en anglais sous l'appellation Forward Backward Sweep (FBS).

### 3.2.1 La méthode Forward backward sweep

#### Exemple sur une branche monophasée

Dans le but d'illustrer son principe de fonctionnement, la méthode FBS est appliquée ici sur une branche monophasée d'un réseau (figure 3.2). La branche considérée est définie par sa tension source  $V_s$  et les impédances des tronçons de lignes  $Z_{ab}$ . Les charges, souvent non linéaires dans les réseaux de distribution, sont représentées par leur puissance complexe monophasée  $S$ . La méthode FBS est basée sur un calcul itératif des tensions

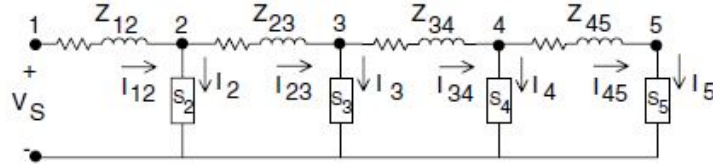


FIGURE 3.2 – Exemple d'une branche monophasée d'un réseau de distribution (source [4])

complexes en chaque nœud et des courants complexes dans chaque tronçon de ligne, en partant du nœud à l'extrémité (nœud 5) vers la source (nœud 1). Initialement, la valeur de la tension de départ au nœud 5 est fixée arbitrairement. La connaissance de la tension du nœud 5 ainsi que de la puissance complexe de la charge 5 permet d'écrire :

$$I_5 = \left(\frac{S_5}{V_5}\right)^* \quad (3.1)$$

Le courant  $I_{45}$  dans la portion de ligne 4-5, étant le même que le courant  $I_5$ , l'application de la loi des maille donne :

$$V_4 = V_5 + Z_{45} * I_{45} \quad (3.2)$$

De là, on déduit le courant alimentant la charge connectée au nœud 4 :

$$I_4 = \left(\frac{S_4}{V_4}\right)^* \quad (3.3)$$

L'application de la loi des nœuds permet d'obtenir le courant dans le tronçon de ligne reliant les nœuds 3 et 4,

$$I_{34} = I_4 + I_{45} \quad (3.4)$$

puis la loi des mailles pour calculer  $V_3$ .

$$V_3 = Z_{34} * I_{34} + V_4 \quad (3.5)$$

Puis le courant aux bornes de la charge 3 en utilisant toujours la puissance complexe

$$I_3 = \left(\frac{S_3}{V_3}\right)^* \quad (3.6)$$

Le courant dans la ligne entre les nœuds 2 et 3

$$I_{23} = I_3 + I_{34} \quad (3.7)$$

ce qui donne la tension au nœud 2

$$V_2 = Z_{23} * I_{23} + V_3 \quad (3.8)$$

On en déduit le courant aux bornes de la charge connectée au nœud 2

$$I_2 = \left(\frac{S_2}{V_2}\right)^* \quad (3.9)$$

Le courant dans la portion de ligne entre les nœuds 1 et 2 ainsi que la tension aux bornes du nœud 1 sont donnés par :

$$I_{12} = I_2 + I_{23} \quad (3.10)$$

$$V_1 = Z_{12} * I_{12} + V_2 \quad (3.11)$$

$$(3.12)$$

La détermination de la tension  $V_1$  en partant du nœud 5 constitue la phase appelée **Forward Sweep**. A l'issue de cette phase, la tension  $V_1$  ainsi obtenue est comparée à la tension de source spécifiée  $V_S$  dans le but de vérifier la convergence de la solution.

$$\epsilon = \frac{||V_S| - |V_1||}{V_S} \quad (3.13)$$

La non convergence se traduit par une erreur  $\epsilon$  supérieure au seuil de tolérance  $\delta$  généralement fixé à 0.0001 [4]. Dans ce cas, on procède à la phase appelée **Backward Sweep**. Cette phase consiste à calculer les tensions des nœuds en partant cette fois ci du nœud 1 vers le nœud 5 tout en utilisant courants calculés dans la phase **Forward Sweep**. En initialisation de cette phase, on suppose logiquement la tension au nœud 1 égale à la tension source  $V_S$  déjà spécifiée au départ. En appliquant la loi des mailles de la source vers l'extrémité, les tensions des nœuds sont données par :

$$V_1 = V_S \quad (3.14)$$

$$V_2 = V_1 - Z_{12} \times I_{12} \quad (3.15)$$

$$V_3 = V_2 - Z_{23} \times I_{23} \quad (3.16)$$

$$V_4 = V_3 - Z_{34} \times I_{34} \quad (3.17)$$

$$V_5 = V_4 - Z_{45} \times I_{45} \quad (3.18)$$

$$(3.19)$$

La phase **Forward Sweep** peut recommencer cette fois ci avec la tension  $V_5$  déterminée à l'issue de la phase **Backward Sweep** et non avec celle admise initialement. Les phases **Forward-Backward Sweep** sont ainsi répétées jusqu'à ce que l'on ait une représentation la plus exacte possible des niveaux de tensions et de courants du réseau, c'est à dire un niveau d'erreur  $\epsilon$  inférieur au seuil de tolérance  $\delta$ .

La méthode FBS est très utilisée dans l'étude des réseaux à structure arborescente. Elle offre la possibilité d'étudier des systèmes d'une, deux ou trois phases [42]. Contrairement à la méthode de Newton Raphson, elle ne nécessite pas l'utilisation d'un modèle composite du réseau, la mise à jour de l'état d'un nœud se faisant seulement au niveau local avec ses voisins directs. Ceci allège la complexité de mise en œuvre des calculs. Pour ces raisons,

la méthode FBS est largement utilisée dans l'étude des systèmes de grande taille ou dans des processus nécessitant une répétition du calcul de load flow [43].

Parmi les autres méthodes d'étude des réseaux électriques, on peut citer la méthode dite Newton Raphson. Elle est basée sur une centralisation des calculs qui repose sur l'utilisation du principe de gradient descendant. Elle implique un modèle composite complet du réseau et l'utilisation de l'inverse de la matrice de gradient de toutes les équations de load flow du système. Elle est plus complexe que la méthode FBS et requiert plus d'itérations, donc de temps de simulation.

### 3.2.2 Choix de l'outil de simulation

Les critères de choix de l'outil de simulation sont basés sur le caractère gratuit et open source de l'outil, son habilité à être interfacée avec d'autres outils de simulation et sa capacité à modéliser les réseaux de distribution triphasés déséquilibrés. Dans ce cadre le logiciel **GridLab-D** développé par la Pacific Northwest National Laboraotry et **OpenDss** développé par EPRI (Electric Power Research Institute) ont été étudiés.

Les travaux dans [44] encadrés dans le cadre de cette thèse ont proposé une étude comparative 3.1 entre les deux logiciels d'étude de réseau sus mentionnés.

#### GridLab-D

GridLab-D est multi-plateforme. Il fonctionne sous Windows, Linux et MacOS. Ses avantages résident dans le fait qu'il permet d'intégrer les modèles des équipements électriques internes des foyers ainsi que les caractéristiques physico-thermiques des maisons. Ce qui peut être utile dans le cas des études de gestion de demande et d'effacement. De même il offre la possibilité d'intégrer directement dans la simulation les paramètres externes tels que la température, l'humidité et l'ensoleillement. Cet aspect peut servir à estimer la consommation et la production électrique sans passer par une étape de traitement de données par un autre outil. Cependant, Gridlab-D ne permet pas de prendre en compte un modèle d'un générateur de production sous la forme d'un nœud PV (P : puissance active, V : tension) et les modèles internes des équipements électriques sont peu modifiables. L'absence de nœud de type PV empêche la modélisation de certains types

TABLE 3.1 – Comparaison entre GridLab-D et OpenDSS

	GridLab-D	OpenDSS
Vitesse de simulation	-	+
IHM	Non	Oui
Granularité des modèles	Possibilité de modéliser les équipements internes d'un foyer	Se limite aux modèles des équipements du réseau
Simulation en série temporel	Oui	Oui
Modification des modèles	Moins flexible, des configurations sont fixes et définis à l'intérieur	Supporte des configurations spécifiques (e.x. transformateur-YY)
Couplage avec d'autres outils	Requête HTTP ou manipuler en utilisant le module interne via TCP/IP	Manipuler via l'interface de COM (Component Object Model) ou DLL (Dynamic Link Library)
Support de type de nœuds	PQ, $V\delta$	PV, PQ, $V\delta$
Plateforme	Windows, Linux, MacOS	Windows
Communauté des utilisateurs	Active	Active
Popularité	-	+

de production. Le faible choix sur la configuration des modèles électriques de certains équipements limite les possibilités d'utilisation de GridLab-D.

## OpenDSS

OpenDSS est aussi un simulateur gratuit, développé depuis 1997 et dédié à la simulation du réseau de distribution. En 2008, EPRI (Electric Power Research Institut) a publié ce logiciel sous une licence ouverte (Open Source). Contrairement à GridLab-D, OpenDSS est muni d'une IHM (Interface Homme Machine), ce qui rend son usage plus facile. En plus, il fournit une interface COM (Component Object Model), ce qui permet de coupler avec presque tous les langages (C++, Matlab, Python, VBA, etc..).

Le choix s'est porté sur OpenDSS car ce logiciel offre plus de choix et de possibilités sur le plan des modèles électriques des composants du réseau. De la comparaison qualitative faite dans les travaux de [44], on peut déduire que GridLab-D est mieux approprié pour



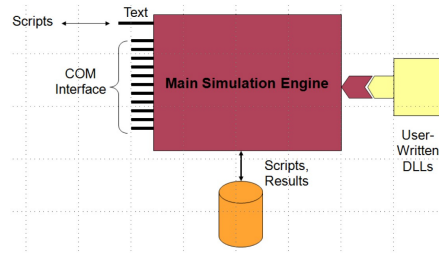


FIGURE 3.3 – Les structures du logiciel OpenDSS [44]

l'étude des Smart Grid à l'échelle de la maison que pour l'étude du réseau en amont du compteur de la maison.

### 3.3 Méthodes Load flow probabilistes

Une méthode de Load Flow Probabiliste (LFP) est une technique qui permet d'étudier les systèmes électriques en prenant en compte l'environnement incertain dans lequel ils évoluent. Dans un calcul de load flow déterministe (présenté ci dessus), les paramètres du réseau ainsi que les données d'entrée (puissances, impédances,...) sont spécifiés par des valeurs fixes sans prendre en compte leurs éventuelles variations, donc leur caractère incertain est négligé.

Dans le but de modéliser cette incertitude, le calcul de load flow probabiliste considère les paramètres d'entrée comme des variables aléatoires continues définies par leur densité de probabilité et leurs moments statistiques. De ce fait, les résultats du calcul de load flow (tension, courant,...) sont de facto des variables aléatoires continues avec leurs densités de probabilité et leurs moments.

Les applications des méthodes probabilistes pour l'étude des réseaux électriques sont diverses et variées. Le load flow probabiliste est généralement appliqué dans les études en phase de planification d'un système électrique [45]. Lors de cette phase une estimation des indices de fiabilité du réseau en fonction des différents scénario de production et de demande permettrait un dimensionnement adéquat des composants du réseau. Pour les mêmes raisons, il est aussi utilisé pour des études d'expansion de réseau [46], [47].

Avec l'avènement des sources de production intermittentes, la question de l'incertitude

est plus que jamais présente dans l'étude des réseaux électriques. Dans ce cadre, les méthodes de load flow probabilistes sont utilisées dans les études d'impact de la position, de la puissance et du taux de pénétration des sources de production photovoltaïques sur les réseaux de transport [48] mais aussi de distribution moyenne et basse tensions [49], [50].

Pour ses missions de contrôle et de diagnostic, les gestionnaires des réseaux peuvent utiliser les techniques probabilistes pour la détection de défauts [51], la prévision des pertes sur le réseau [52], ou bien l'optimisation de la capacité d'une source de production intermittente [53].

La partie qui suit va présenter les principales techniques de load flow probabilistes utilisées dans la littérature.

### 3.3.1 État de l'art

Les différentes méthodes probabilistes de calcul de répartition de puissance peuvent être regroupées en quatre catégories, à savoir les méthodes de simulation de type Monte Carlo, les méthodes analytiques, les méthodes approximées et les méthodes hybrides [54].

#### Méthodes de simulation Monte Carlo

La méthode de simulation Monte Carlo (MC) [55], [46], [56] également appelée méthode numérique consiste à répéter un calcul de Load flow déterministe avec différents scénarios de production et de consommation. Pour chaque scénario, les valeurs des paramètres d'entrée sont tirées aléatoirement à partir de densités de probabilité.

La méthode de Monte Carlo utilise les équations non linéaires de load flow présentées dans la section précédente, de ce fait elle est considérée comme la référence de toutes les autres méthodes en terme de précision [57]. D'un autre côté, elle nécessite un très grand nombre de simulations de load flow déterministe pour arriver à un tel niveau de précision, ce qui implique un long temps de simulation pour les réseaux de grande taille. Par contre cette précision n'est pas sensible à la taille du réseau et la méthode de Monte Carlo permet aussi de prendre en compte les changements de configuration du réseau et

la corrélation entre les paramètres d'entrée [58].

Pour réduire le nombre de simulations nécessaires sans perdre en précision, des méthodes basées sur des techniques de réduction de variance [59] telles que l'échantillonnage dirigé et l'échantillonnage LHS (Latin hypercube sampling) [60] sont utilisées.

### **Les méthodes analytiques**

Les premières méthodes analytiques étaient basées sur des techniques de convolution pour estimer les densités de probabilité des résultats de load flow en fonction de celles des paramètres aléatoires d'entrées [61], [62]. La convolution permet de calculer analytiquement la distribution de probabilité du résultat de load flow en fonction des distributions de paramètres d'entrée qui étaient généralement modélisées par une loi normale pour les charges et une loi binomiale pour les unités de production [63], [64]. La convolution suppose une linéarisation des équations de load flow, une totale indépendance ou une corrélation linéaire entre les variables d'entrées, et une configuration constante du réseau [57].

Mais la performance des méthodes basées sur la convolution est affectée par le niveau de dispersion des paramètres d'entrées [65]. Dans ce cadre la convolution classique est remplacée par la FFT (Transformée de Fourier Rapide)[66] qui est utilisée pour améliorer la précision et gagner en temps de simulation.

Les calculs de convolution sont complexes [67]. Les méthodes des cumulants permettent de s'affranchir de cette complexité en effectuant des opérations arithmétiques sur les caractéristiques numériques des variables aléatoires comme les moments et les cumulants [68]. Les travaux dans [69],[70] proposent des variantes pour prendre en compte le changement de configuration du réseau.

Les méthodes analytiques permettent d'estimer des moments statiques. De ce fait elles sont combinées avec des techniques de reconstruction de densité de probabilité comme celles dite de "Maximum Entropy (ME)" [71] ou de Gram-Charlier (GC) [67] pour estimer les densités des paramètres aléatoires de sortie.

Les méthodes analytiques permettent de réduire le temps de simulation par rapport

à la méthode de simulation Monte Carlo [66]. Par contre, la linéarisation des équations de load flow qu'elles impliquent diminue la précision des estimations[57]. Dans ce sens, des méthodes telles que la multi linéarisation [46] et la PLF quadratique [72],[73] sont utilisées pour améliorer la précision. Mais malgré cette amélioration, ces méthodes restent moins précises que la méthode MC.

### **Les méthodes approximées**

Les premières méthodes dites approximées, sont basées sur les séries de Taylor tronquées [74], la méthode de Kaplan [75] et la méthode dite de source d'incertitude commune [76]. La méthode des séries de Taylor tronquées repose sur les calculs des dérivés de la fonction liant les variables aléatoires d'entrée et de sortie. La méthode Kaplan propose une discrétisation des densités de probabilité continues des variables d'entrée. Même si elle est simple à mettre en œuvre, elle ne permet pas de prendre en compte des variables d'entrée corrélées et nécessite un très grand nombre de pas de simulation pour seulement une vingtaine de paramètres aléatoire d'entrée [77]. La méthode dite de source commune fonctionne avec l'hypothèse selon laquelle une variable d'entrée  $X$  est une fonction de plusieurs variables aléatoires dépendantes de distribution normale ou lognormale. De ce fait son utilisation est limitée.

Plus récemment on retrouve aussi dans la littérature la méthode FOSM (First Order Second Moment - Premier ordre second moment) qui consiste à estimer les moments d'ordre premier et second des variables de sortie, en fonction du développement en série de Taylor au premier ordre de la fonction reliant variables de sortie et d'entrée[78].

La méthode PEM (Point Estimate Method) [79],[80], s'inscrit dans le même principe. Elle est basée sur une approximation des densités de probabilité des paramètres d'entrée par des points pondérés avec leurs probabilités respectives. Ces probabilités, aussi appelées poids, déterminent l'influence d'un point donné sur le résultat de sortie. Chaque variable aléatoire d'entrée peut être représentée par  $m$  points. Les positions des points et leurs poids, pour une variable d'entrée donnée, sont exprimés en fonction des moments d'ordre  $2m - 1$  de cette variable. Pour  $n$  variables, on obtient ainsi  $m * n$  points. Le Load

flow déterministe est calculé alors  $m*n$  fois, le résultat intermédiaire pour un point donné est pondéré par son poids. Le moment d'ordre  $j$  du paramètre de sortie est obtenu cumulativement en sommant les résultats intermédiaires élevés à la puissance  $j$  et pondérés par leurs poids respectifs.

L'avantage de la méthode PEM est qu'elle ne nécessite pas d'avoir une représentation exacte de la densité de probabilité des paramètres d'entrée [81] du fait qu'elle requiert seulement les moments d'ordre premier. L'autre avantage est que l'utilisation des équations non linéaires de load flow déterministe rend sa précision meilleure que celle des méthodes analytiques dans une certaine mesure [68] pour un temps de calcul raisonnable [82].

Cependant des études citées dans [58] ont montré certaines limites de la méthode PEM. Pour certaines distributions bornées et non gaussiennes, comme les distributions lognormale ou exponentielle les points peuvent être localisées en dehors de leur domaine si l'écart-type est grand. La méthode PEM ne permet pas non plus de prendre en compte les changements de la configuration du système électrique au cours de la simulation.

Les méthodes approximées offrent un compromis entre la précision des grandeurs électriques estimées et le temps de simulation [83]. À l'instar des méthodes analytiques, les méthodes approximées utilisent les caractéristiques numériques (moments et cumulants) des paramètres aléatoires d'entrée. Mais les deux catégories de méthodes diffèrent sur la manière de calculer les caractéristiques des paramètres de sortie, puisque les méthodes dites approximées sont basées sur l'utilisation des équations non linéaires de Load Flow. Elles ont ce dernier aspect en commun avec les méthodes de simulation de Monte Carlo.

D'autres méthodes ont été utilisées, dans la littérature, pour l'étude probabiliste des réseaux électriques.

### **Les autres méthodes**

Les travaux dans [84] présentent une technique de calcul basée sur la méthode Taguchi. Elle étudie l'effet de plusieurs variables d'entrée possibles sur la ou les variables de sortie. Les variables aléatoires d'entrée sont considérées comme des facteurs avec différents niveaux. A l'image de la méthode PEM, le nombre de niveaux dépend de la précision

recherchée. Les niveaux sont des points équidistants de part et d'autre de la moyenne de la distribution du paramètre aléatoire. Le load flow déterministe est calculé en utilisant les niveaux sélectionnés sur une base qualitative. Il est démontré que le nombre de simulations requises est inférieure à celui de la méthode PEM, en plus d'une amélioration de la précision des moyennes et écart types estimés pour une distribution normale [84].

Une extension probabiliste de la méthode Forward Backward Sweep est proposée dans [43]. Elle estime les moments des paramètres de sortie en fonction de ceux des paramètres d'entrée sans linéarisation des équations de load flow. Comme un FBS classique, la phase de backward sweep (balayage amont) permet de calculer les moments des courants dans les branches en partant de l'extrémité vers la source. De la source à l'extrémité, la phase de forward sweep (balayage vers l'aval) calcule les moments des tensions des nœuds et des charges. Dans la version déterministe, les courants sont calculés par la relation algébrique entre puissance et tension. Dans le cas probabiliste, il n'est pas possible d'avoir une valeur exacte. Une approximation est alors faite en utilisant un gradient d'erreur calculé dans la phase Forward Sweep. Selon l'auteur, la méthode fournit des résultats satisfaisants mais ses estimations ne dépassent pas les moments d'ordre 2.

Une méthode proposant une combinaison de la méthode analytique des cumulants avec des simulations de load flow déterministe (MC) est présentée dans [85]. Elle permet, dans une même étude donnée, de prendre en compte des variables aléatoires discrètes comme une unité de production au fioul et des variables aléatoires continues comme la production d'une éolienne. Les distributions des réponses (résultats) du système aux variables d'entrée continues sont estimées par la méthode des cumulants. Celles correspondant aux variables d'entrée discrètes sont estimées par la technique de multiple load flow déterministes (MC). La réponse du système aux deux types de variables est alors obtenue par la convolution des deux réponses précédentes.

Notre étude s'est portée sur deux de ces méthodes, à savoir la méthode de Monte Carlo qui peut être considérée comme référence en terme de précision [86] et la méthode PEM pour le compromis qu'elle offre entre temps de simulation et précision des estimations

[82].

### 3.3.2 Méthode Monte Carlo

La méthode MC est une méthode numérique permettant d'estimer une valeur en utilisant des procédés aléatoires. Elle consiste à effectuer un nombre  $N$  fois le calcul de load flow déterministe en considérant différents scénarios de production et de consommation

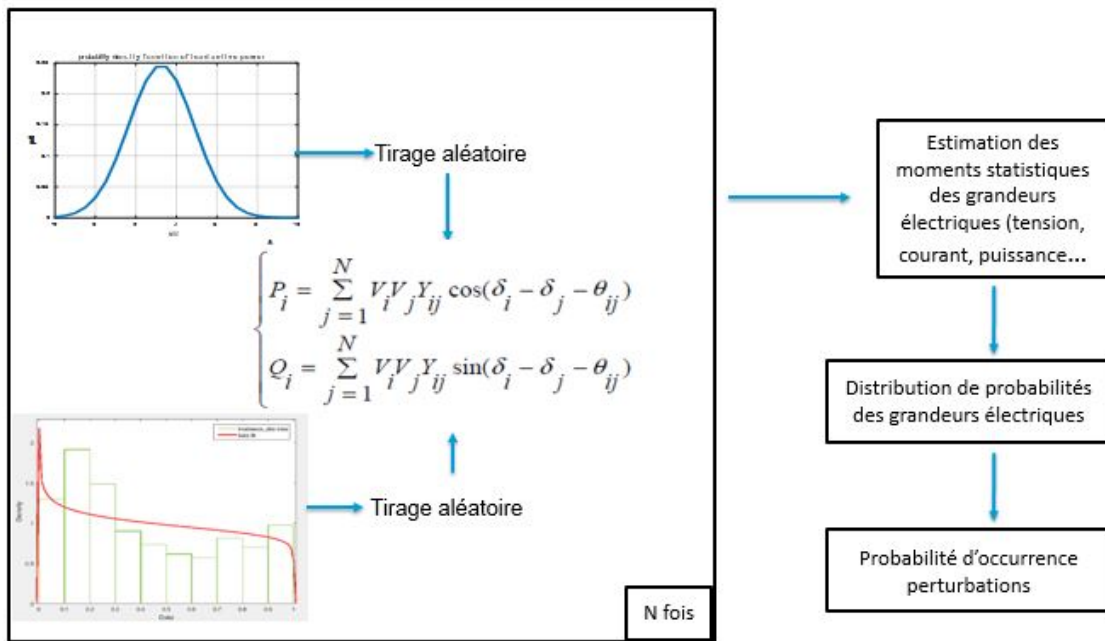


FIGURE 3.4 – Schéma de principe de la méthode de Monte Carlo

puissances actives des charges sont générées aléatoirement à partir de leurs densités de probabilité respectives. Ces paramètres d'entrée sont échantillonnés indépendamment et identiquement distribués, les résultats correspondants sont aussi indépendants et identiquement distribués.

A l'issue des  $N$  scénarios, la moyenne et l'écart type des grandeurs observées sont estimés par les équations 3.20, 3.21 [87].

$$\mu_X \approx \frac{1}{N} \sum_{i=1}^N X_i \quad (3.20)$$

$$\sigma_X^2 \approx \frac{1}{N-1} \sum_{i=1}^N X_i^2 - \mu_X^2 \quad (3.21)$$

Selon le théorème de la limite centrale des grands nombres [39] :

$$\lim_{N \rightarrow +\infty} Pr \left| \frac{\bar{\mu} - \mu}{\sigma \sqrt{N}} \leq x \right| = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} dx \quad (3.22)$$

avec  $Pr$  une expression de probabilité,  $N$  le nombre de scénarios simulés,  $\bar{\mu}$  et  $\mu$  respectivement les espérances échantillonnées et réelles du système résultant et  $\sigma$  l'écart-type réel du système résultant. Pour un certain niveau de confiance  $\alpha$ , alors l'erreur sur la moyenne estimée peut s'écrire :

$$\lim_{N \rightarrow +\infty} Pr |\bar{\mu} - \mu| \leq \frac{x_\alpha \sigma}{\sqrt{N}} \approx \frac{2}{\sqrt{2\pi}} \int_0^{x_\alpha} e^{-\frac{x^2}{2}} dx \quad (3.23)$$

Ceci montre que l'erreur d'estimation est inversement proportionnelle à la racine carrée du nombre de scénarios considérés. Plus  $N$  est grand, plus on tend vers la convergence. L'amélioration de la précision d'une décimale impliquerait une augmentation du nombre de simulations nécessaires par 100 [58]. En pratique, la convergence est déterminée en fonction de la variance du système résultant. On considère que le problème converge lorsque la variation de la variance entre deux itérations successives est inférieure à un seuil donné [88].

Les lois de probabilité des grandeurs résultantes sont estimées à partir de la distribution des résultats observés à l'issue de tous les scénarios en utilisant des méthodes telle que celle dite du maximum de vraisemblance.

### **Estimation de la probabilité de dépassement des seuils de fiabilité**

Soit  $X_{lim}$  la valeur limite d'un paramètre électrique (tension, courant de neutre, facteur de déséquilibre...) au delà de laquelle, les conditions de bon fonctionnement du réseau ne sont plus respectées et  $p$  la probabilité de dépassement de cette valeur. Cette probabilité est calculée par l'équation 3.24 où  $P(X \leq X_{lim})$  est l'image de la fonction de répartition



du paramètre  $X$  en  $X_{lim}$ .

$$p = P(X \geq X_{lim}) = 1 - P(X \leq X_{lim}) \quad (3.24)$$

### 3.3.3 Méthode d'estimation ponctuelle (PEM)

#### Principes

La méthode d'estimation ponctuelle (ou PEM pour Point Estimate Method) permet d'estimer les moments d'une variable aléatoire  $U$  qui est une fonction d'une ou de plusieurs autres variables aléatoires  $(X_1, X_2, \dots, X_i, \dots, X_n)$ .

$$U = h(X_1, X_2, \dots, X_i, \dots, X_n) \quad (3.25)$$

Contrairement à la méthode de Monte Carlo, elle ne nécessite pas de tirage d'échantillons aléatoires. Elle est dite méthode approximée car la loi de probabilité  $f_{X_i}$  du paramètre d'entrée  $X_i$  est réduite à une concentration de points associée à une pondération respective à chaque point. Ces points approximatifs sont estimés en fonction de la moyenne (équation 3.26), de l'écart type (équation 3.27), du coefficient d'asymétrie (Skewness) (équation 3.28) et du coefficient d'aplatissement (Kurtosis) (équation 3.29) de chaque paramètre d'entrée.

$$\mu_X = E[X] = \int_{-\infty}^{+\infty} x f_X(x) dx \quad (3.26)$$

$$\sigma_X^2 = E[(X - E[X])^2] = \int_{-\infty}^{+\infty} (x - \mu_X)^2 f_X(x) dx \quad (3.27)$$

$$\lambda_{X,3} = E\left[\left(\frac{X - E[X]}{\sigma}\right)^3\right] = \frac{1}{\sigma^3} \int_{-\infty}^{+\infty} (x - \mu_X)^3 f_X(x) dx \quad (3.28)$$

$$\lambda_{X,4} = E\left[\left(\frac{X - E[X]}{\sigma}\right)^4\right] = \frac{1}{\sigma^4} \int_{-\infty}^{+\infty} (x - \mu_X)^4 f_X(x) dx \quad (3.29)$$

Dans les travaux de Krishna et al [89], on retrouve un récapitulatif des variantes de la méthode PEM développées à ce jour (cf tableau 3.2). La méthode PEM développée par

TABLE 3.2 – Comparaison qualitative des méthodes PEM

Auteur de la méthode	Nombre de simulation	Efficacité sur réseau de grande taille	Variables corrélés	Variables asymétriques
Rosenblueth [90]	$2^n$	Très basse	Oui	Oui
Li [91]	$O(n^3)$	Basse	Oui	Oui
Harr [92]	$2n$	Haute	Oui	Non
Hong [93]	$Kn$ ou $Kn + 1$	Haute	Non	Oui

Rosenblueth concentre la loi de probabilité du paramètre aléatoire d'entrée  $X$  en deux points équidistants de sa moyenne, de sorte à couvrir les trois premiers moments centrés de  $X$ , pour estimer ceux de  $U$ . Cette méthode est relativement simple à mettre en œuvre tout en garantissant un niveau de précision acceptable lorsque l'on est en présence d'une seule variable aléatoire d'entrée  $X$ . Dans le cas de  $n$  variables d'entrée, cette méthode nécessiterait  $2^n$  concentrations localisées en  $2^n$  points pour remplacer la densité de probabilité jointe des variables d'entrée.

Les travaux de Li [91] ont montré que le fait d'utiliser  $2^n$  concentrations lorsque l'on est en présence de plusieurs variables d'entrée n'est pas efficace pour les problèmes de grande taille. Il a développé une variante qui réduit le nombre de points à  $2n$  mais ne couvre pas plus loin que les moments centrés d'ordre 2.

Quant à Harr [92], il a développé lui aussi une méthode PEM utilisant  $2n$  points, couvrant aussi les deux premiers moments centrés en introduisant la prise en compte des variables corrélées.

Les méthodes de Li et Harr proposent une approximation en deux points, mais ne prennent pas en compte les moments d'ordre supérieur à 2 comme le skewness ou le kurtosis. Les travaux de Morales [83] ont montré que le niveau de précision obtenu avec cette approche décroît considérablement avec la taille du problème.

La méthode de Hong [93] propose une estimation des moments du paramètre de sortie

$U$  en utilisant  $K * n$  concentrations couvrant les premiers  $(2K - 1)$  moments de chaque variable aléatoire d'entrée. Cette dernière est utilisée dans nos travaux parce qu'elle offre un meilleur compromis entre temps de simulation et précision.

### Méthode Hong à une variable

Dans un premier temps, considérons une variable aléatoire  $Z$  fonction d'une seule variable aléatoire  $X$  de densité de probabilité  $f_X(x)$ . Les deux étant liées par la relation :

$$U = h(X) \quad (3.30)$$

L'objectif est d'estimer les moments premiers de la variable aléatoire  $U$  en fonction de ceux du paramètre aléatoire d'entrée  $X$ . On désigne par  $\mu_X$ ,  $\sigma_X$ ,  $\lambda_{X,3}$  et  $\lambda_{X,4}$  la moyenne, l'écart type, le coefficient de skewness et celui de kurtosis du paramètre  $X$  (cf 3.26, 3.27, 3.28 et 3.29). Notons par  $M_r^c(X)$  le moment centré d'ordre  $r$  de  $X$  donné par l'équation 3.31.

$$M_r^c(X) = \int_{-\infty}^{+\infty} (x - \mu_X)^r f_X(x) dx \quad (3.31)$$

En posant :

$$\lambda_{X,r} = \frac{M_r^c(X)}{\sigma_X^r} \quad (3.32)$$

on retrouve les coefficients de skewness et de kurtosis pour  $r = 3$  et  $r = 4$ , sachant que

$$\lambda_{X,1} = 0 \quad (3.33)$$

et

$$\lambda_{X,2} = 1 \quad (3.34)$$

Le développement en séries de Taylor de la fonction  $h$  au point  $\mu_X$  donne [93] :

$$h(x) = h(\mu_X) + \sum_{k=1}^{\infty} \frac{1}{k!} h^{(k)}(\mu_X) (x - \mu_X)^k \quad (3.35)$$

avec  $h^{(k)}$  la  $k^{ieme}$  dérivée de la fonction  $h$ . Notons par  $\mu_U$  la moyenne du paramètre  $U$ .

Elle peut être calculée par l'espérance de l'équation 3.30.

$$\mu_U = E(U) = E(h(X)) \quad (3.36)$$

Si  $X$  est une variable aléatoire continue de loi de probabilité  $f_X$ , alors on peut écrire : [94]

$$E(g(X)) = \int_{-\infty}^{+\infty} g(x)f_X(x) dx \quad (3.37)$$

En appliquant l'équation 3.37, l'espérance de  $U$  s'écrit :

$$E(h(X)) = \int_{-\infty}^{+\infty} h(x)f_X(x) dx \quad (3.38)$$

En remplaçant  $h(x)$  par son développement en série de Taylor et après simplification, on obtient :

$$\mu_U = h(\mu_X) + \sum_{k=1}^{\infty} \frac{1}{k!} h^{(k)}(\mu_X) \lambda_{X,k} \sigma_X^k \quad (3.39)$$

On choisit deux points (appelés concentrations) sur la densité de probabilité de  $X$  positionnés en  $x_1$  et  $x_2$  définis par l'équation 3.40 avec  $\xi_k$  une constante à déterminer pour chaque position  $k$ .

$$x_k = \mu_X + \xi_k \sigma_X, k = 1, 2 \quad (3.40)$$

Soit  $w_k$  le poids d'un point  $x_k, k = 1, 2$  qui détermine l'influence de ce dernier sur le résultat final. L'idée est d'approximer l'espérance de  $U$  par  $w_1 h(x_1) + w_2 h(x_2)$  [93]. Le calcul de  $w_1 \times h(x_1) + w_2 \times h(x_2)$  grâce à l'expression en série de Taylor de  $h$  3.35 est donné par l'équation 3.41 :

$$w_1 h(x_1) + w_2 h(x_2) = h(\mu_X)(w_1 + w_2) + \sum_{k=1}^{\infty} \frac{1}{k!} h^{(k)}(\mu_X) (w_1 \xi_1^k + w_2 \xi_2^k) \sigma_X^k \quad (3.41)$$

Par identification des équations 3.39 et 3.41, on peut écrire :

$$w_1 \xi_1^k + w_2 \xi_2^k = \lambda_{X,k} \quad (3.42)$$

Nous sommes en présence de quatre inconnues,  $w_1, \xi_1, w_2$  et  $\xi_2$ . Un système à quatre équations peut être obtenu à la condition de prendre en compte les moments centrés

d'ordre 0, 1, 2 et 3 [93]. L'équation 3.42 peut être réécrite :

$$\sum_{j=1}^K w_j \xi_j^k = \lambda_{X,k} \quad (3.43)$$

avec  $k = 0, 1, \dots, 2K - 1$ . Pour rappel  $K$  désigne le nombre de points utilisés pour remplacer le paramètre d'entrée  $X$ . Ici  $K = 2$ , ce qui implique un système de quatre équations à quatre inconnues :

$$\begin{cases} w_1 + w_2 & = 1 \\ w_1 \xi_1 + w_2 \xi_2 & = 0 \\ w_1 \xi_1^2 + w_2 \xi_2^2 & = 1 \\ w_1 \xi_1^3 + w_2 \xi_2^3 & = \lambda_{X,3} \end{cases}$$

La résolution de ce système, détaillée en annexe, donne :

$$\xi_1 = \frac{\lambda_{X,3}}{2} + \sqrt{1 + \left(\frac{\lambda_{X,3}}{2}\right)^2} \quad (3.44)$$

$$\xi_2 = \frac{\lambda_{X,3}}{2} - \sqrt{1 + \left(\frac{\lambda_{X,3}}{2}\right)^2} \quad (3.45)$$

$$w_1 = \frac{-\xi_2}{\xi_1 - \xi_2} \quad (3.46)$$

$$w_2 = \frac{\xi_1}{\xi_1 - \xi_2} \quad (3.47)$$

Revenons à l'équation 3.41, sachant que  $w_1 + w_2 = 1$  et en combinant avec l'équation 3.43 [93], on peut écrire :

$$w_1 h(x_1) + w_2 h(x_2) = h(\mu_X) + \sum_{k=1}^3 \frac{1}{k!} h^{(i)}(\mu_X) \lambda_{X,k} \sigma_X^k + \sum_{k=4}^{\infty} \frac{1}{k!} h^{(i)}(\mu_X) (w_1 \xi_1^k + w_2 \xi_2^k) \sigma_X^k \quad (3.48)$$

D'où

$$h(\mu_X) + \sum_{i=k}^3 \frac{1}{k!} h^{(i)}(\mu_X) \lambda_{X,k} \sigma_X^k = w_1 h(x_1) + w_2 h(x_2) - \sum_{k=4}^{\infty} \frac{1}{k!} h^{(k)}(\mu_X) (w_1 \xi_1^k + w_2 \xi_2^k) \sigma_X^k \quad (3.49)$$

En substituant l'équation 3.49 dans l'équation 3.39, on obtient :

$$\mu_U = w_1 h(x_1) + w_2 h(x_2) + \sum_{k=4}^{\infty} \frac{1}{k!} h^{(i)}(\mu_X) (\lambda_{X,k} - (w_1 \xi_1^k + w_2 \xi_2^k)) \sigma_X^k \quad (3.50)$$

Ainsi l'espérance de  $U$  peut être obtenue selon une approximation au troisième ordre en négligeant les dérivées  $h^{(k)}$  pour  $k = 4, 5, \dots$  [93]. Il s'agit de la somme des images par  $h$  des points  $x_k$  pondérés par le poids  $w_k$  de chaque point.

$$\mu_U \cong w_1 h(x_1) + w_2 h(x_2) \quad (3.51)$$

De la même manière, le moment central d'ordre 2 de  $U$  peut être approximé par [93] :

$$E(U^2) \cong w_1 h(x_1)^2 + w_2 h(x_2)^2 \quad (3.52)$$

Cette approche peut être généralisée à l'utilisation de  $m$  points, pour représenter la densité de probabilité de la variable d'entrée  $X$ , localisés en  $x_k = \mu_X + \xi_k \sigma_X$ ,  $k = 1, 2, \dots, m$  pondérés avec  $w_k$ . L'utilisation d'un grand nombre de point permet de couvrir un ordre plus élevé de moment de la variable  $X$ , augmentant ainsi la précision de la méthode selon [54]. Pour  $m$  points, on couvre les  $2m - 1$  premiers moments. Dans ce cas, l'espérance du paramètre résultant  $U$  est estimée par [93] :

$$\mu_U \cong \sum_{j=1}^m w_j h(x_j) \quad (3.53)$$

### Méthode Hong à plusieurs variables d'entrée

Considérons cette fois ci une variable aléatoire  $U$  fonction de  $n$  variables aléatoires  $X_{l \text{avecl}} = 1, 2, \dots, n$  comme le montre l'expression 3.54 :

$$Z = h(X_1, X_2, \dots, X_n) \quad (3.54)$$

Cette fois ci, les points auxquels la fonction  $h$  sera évaluée sont définis non pas par la simple expression  $x_{l,k}$  de chaque variable mais par un vecteur  $\Gamma$  constitué de l'approximation  $x_{l,k}$  d'une variable donnée et des valeurs moyennes de toutes les autres variables (cf équation 3.55) . Comme dans le cas d'une seule variable d'entrée, chaque vecteur est pondéré par un poids  $w_{X_{l,k}}$  avec  $l = 1, 2, \dots, n$  et  $k = 1, 2, \dots, m$ .

Autrement dit, à chaque variable d'entrée  $X_l$ , on associe  $m$  vecteurs à  $n$  éléments et

$m$  poids ( $m$  nombre de points  $k$  à considérer et  $n$  nombre de variables d'entrée). Les éléments du vecteur  $k$  associé à la variable d'entrée  $X_l$  sont les valeurs moyennes des autres variables et la valeur de l'approximation  $x_{X_l,k}$  de la variable  $X_l$ .

$$\Gamma_{X_l,k} = (\mu_{X_1}, \mu_{X_2}, \dots, x_{X_l,k}, \dots, \mu_{X_n}) \quad (3.55)$$

Sachant que le point  $x_{l,k}$  est défini par :

$$x_{l,k} = \mu_{X_l} + \xi_{X_l,k} \sigma_{X_l} \quad (3.56)$$

La présence de plusieurs variables amène à l'utilisation du développement en série de Taylor multi variables de la fonction  $h$  autour des moyennes des paramètres aléatoires  $X_l$  [93]. Par une démarche similaire au cas d'une seule variable, il est possible d'établir  $(2m - 1)$  équations pour chaque variable  $X_l$  respectant la condition [93] :

$$\sum_{k=1}^m w_{X_l,k} (\xi_{X_l,k})^j = \lambda_{X_l,j} \quad (3.57)$$

avec  $j = 1, 2, \dots, 2m - 1$ ,  $l = 1, 2, \dots, m$  et  $\lambda_{X_l,j} = M_j^c(X_l)/\sigma^j$ . En plus nous savons que  $\sum_{l=1}^n \sum_{k=1}^m w_{X_l,k} = 1$ , ce qui permet d'écrire :

$$\sum_{k=1}^m w_{X_l,k} = \frac{1}{n} \quad (3.58)$$

avec  $l = 1, 2, \dots, n$ .

Les équations 3.57 et 3.58 fournissent un système de  $2m$  équations à  $2m$  inconnues pour chaque variable. La résolution de cette équation pour  $m = 3$ , par exemple, donne pour chaque point  $k$  d'une variable  $X_l$ , sa constante interne  $\xi_{X_l,k}$  et son poids  $w_{X_l,k}$  exprimés en fonction des coefficients de skewness (asymétrie)  $\lambda_{X_l,3}$  et de Kurtosis (apla-

tissement)  $\lambda_{Xl,4}$  du variable  $X_l$ .

$$\xi_{Xl,k} = \frac{\lambda_{Xl,3}}{2} + (-1)^{3-k} \sqrt{\lambda_{Xl,4} - \frac{3(\lambda_{Xl,3})^2}{4}} \quad k = 1, 2 \quad (3.59)$$

$$\xi_{Xl,3} = 0 \quad (3.60)$$

$$w_{Xl,k} = \frac{(-1)^{3-k}}{\xi_{Xl,k}(\xi_{Xl,1} - \xi_{Xl,2})}, \quad k = 1, 2 \quad (3.61)$$

$$w_{Xl,3} = \frac{1}{n} - \frac{1}{\lambda_{Xl,4} - (\lambda_{Xl,3})^2} \quad (3.62)$$

On remarque que pour  $m = 3$ ,  $\xi_{Xl,3} = 0$ , ce qui implique que  $x_{l,3} = \mu_{Xl}$ ,  $l = 1, 2 \dots n$ . De ce fait,  $n$  des  $3n$  concentrations seront localisées au même point  $(\mu_{X1}, \mu_{X2}, \dots, \mu_{Xn})$ . En posant  $w_0$ , comme la somme de leur pondération, on aura :

$$w_0 = \sum_{l=1}^n w_{l,3} = 1 - \frac{1}{\lambda_{Xl,4} - (\lambda_{Xl,3})^2} \quad (3.63)$$

L' estimation du moment d'ordre  $j$  de la variable résultante  $Z$ .

$$E(U^j) \cong w_0 (h(\mu_{X1}, \mu_{X2}, \dots, \mu_{Xn}))^j + \sum_{l=1}^n \sum_{i=1}^m w_{Xl,i} (h(\mu_{X1}, \mu_{X2}, \dots, x_{l,i}, \dots, \mu_{Xn}))^j \quad (3.64)$$

### Application au calcul de load flow probabiliste

Sous des conditions de configuration de réseau et des paramètres de lignes constants, le problème de load flow peut être modélisé par une fonction non linéaire  $h_s$  reliant les puissances injectées et soutirées aux grandeurs électriques de sortie (tensions, courants, ...) désignées par  $U_s$  comme le montre l'équation 3.65. Du fait de leur caractère incertain, les puissances injectées/soutirées sont considérées comme des paramètres aléatoires qu'on notera par  $X$ . Le problème de load flow peut être considéré alors comme des grandeurs électriques aléatoires (tension, courant..)  $U_s$  fonction de plusieurs variables aléatoires d'entrée  $X_i$ . De ce fait la méthode Hong à plusieurs variables d'entrée est applicable dans



ce cas.

$$U_s = h_s(X_1, X_2, X_3, \dots, X_i, \dots, X_n) \quad (3.65)$$

L'application de la méthode Hong à  $n$  variables implique une représentation de chaque variable d'entrée  $X_i$  par  $m$  points  $P_{X_i,k}$  à la probabilité  $w(X_i, k)$  avec  $k = 1, 2, \dots, m$ . La position du point d'approximation  $k$  d'une variable  $X_i$  est donnée par l'équation 3.66.

$$P_{X_i,k} = \mu_{X_i} + \xi_{X_i,k} \times \sigma_{X_i} \quad (3.66)$$

Pour couvrir les  $2m - 1$  premiers moments, le poids  $w_{X_i,k}$  du point  $P_{X_i,k}$  ainsi que la constante  $\xi_{X_i,k}$  doivent satisfaire les conditions définies par le système d'équations ci dessous :

$$\begin{cases} \sum_{k=1}^m w_{X_i,k} & = 1 \\ \sum_{k=1}^m w_{X_i,k} (\xi_{X_i,k})^j & = M_{X_i,j}^c \end{cases}$$

dont  $j = 1, 2, \dots, 2m - 1$  et  $M_{X_i,j}^c$  est le moment centré d'ordre  $j$  du paramètre  $X_i$ .

L'augmentation du nombre  $m$  de points d'approximation se traduit par l'utilisation d'informations supplémentaires de la loi de probabilité (comme le coefficient d'aplatissement dont on n'aurait pas besoin si  $K$  était égal à deux), ce qui améliorerait sans doute la précision du résultat. Mais il en résulterait aussi proportionnellement une croissance du nombre d'évaluation nécessaire du load flow déterministe. Rappelons que ce dernier est égal à  $(m * n)$ .

Pour un problème portant sur un réseau de distribution couvrant une grande zone géographique avec des unités de production photovoltaïque domestiques, on aurait plusieurs milliers de paramètres d'entrées à considérer. Le nombre de simulations déterministes requis peut avoisiner la dizaine de millier. Il faut arriver à un nombre de points  $m$  qui fournirait une précision acceptable tout en minimisant le nombre de calculs déterministes requis. Les conclusions des travaux de Morales [83] indiquent une faible précision pour une méthode PEM à 2 points sur un réseau IEEE 118 nœuds par rapport à une méthode PEM à 3 points. Pour ces raisons, nous avons porté notre étude en choisissant  $m = 3$ . Les poids et constantes internes des points de chaque variable sont alors donnés par les équations 3.62.

Une fois ces inconnues déterminées, le load-flow déterministe est calculé  $m * n$  fois. Pour chaque calcul, le paramètre considéré est remplacé par ses  $m$  points à tour de rôle et les autres paramètres par leur valeur moyenne comme l'indique l'équation 3.67. Cette opération est répétée jusqu'à ce que tous les  $n$  paramètres aléatoires d'entrée soient considérés.

$$Us(Xi, k) = h(\mu_{X1}, \mu_{X2}, \mu_{X3}, \dots, P_{Xi,k}, \dots, \mu_{Xn}) \quad (3.67)$$

L'étude porte sur le réseau de distribution caractérisé par l'existence d'une multitude de sites à basse et moyenne consommation essentiellement constitués de foyers domestiques. Ce qui se traduit dans le modèle par un grand nombre  $n$  de variables aléatoires d'entrée et en conséquence beaucoup de calculs de load flow déterministes. Les travaux dans [43] ont montré que la méthode FBS est efficace dans ces cas de figure, raison pour laquelle nous l'avons utilisée pour évaluer les calculs de load flow déterministes. Les moments théoriques d'ordre  $r$  des variables résultantes sont estimés cumulativement suivant l'équation 3.68. Dans notre cas, le résultat de load flow  $Us$  désigne soit la tension en  $pu$ , soit les parties réelles et imaginaires des courants de ligne, soit les parties réelles et imaginaires des puissances, soit le module du courant de neutre.

$$E(Us^r) \simeq \sum_{l=1}^m \sum_{k=1}^m w_{Xi,k} (Us(Xi, k))^r \quad (3.68)$$

L'espérance de la variable de sortie est alors estimée par  $\mu_{Us} = E(Us)$  et son écart-type par  $\sigma_{Us} = \sqrt{E(Us^2) - E(Us)^2}$ .

### 3.4 Conclusion

La quantification de l'impact de la production décentralisée et des consommations sur les contraintes techniques du réseau se fait grâce à un calcul de load flow déterministe. Le load flow permet de calculer les amplitudes et phases des courants et tensions en chaque point du réseau en fonction des flux de puissances injectées et soutirées en un instant donné. Parmi les méthodes déterministes de calcul de load flow, la méthode FBS est mieux appropriée pour les réseaux à structure radiale tels que les réseaux de distribution

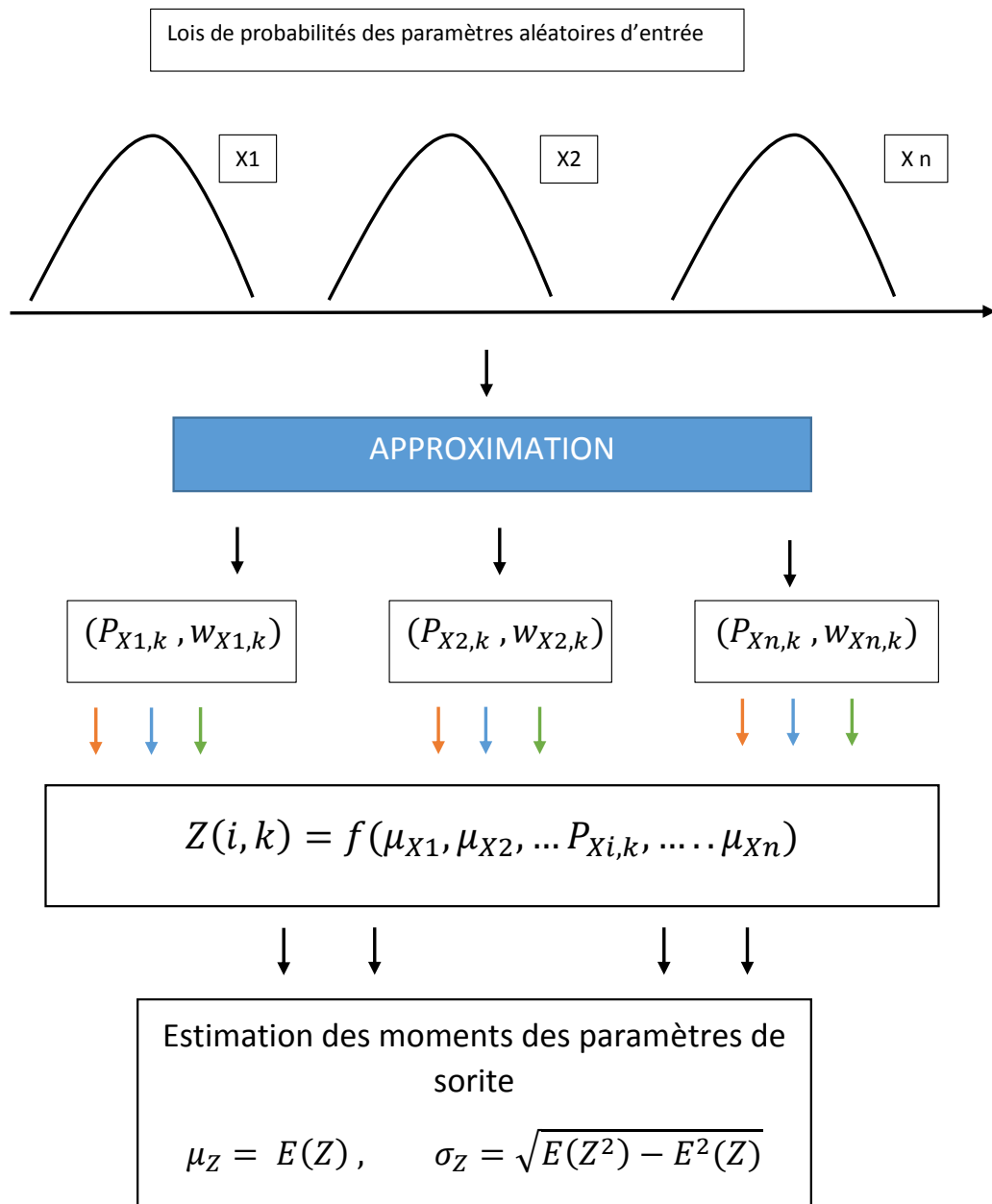
d'électricité.

Dans le but de prendre en compte l'incertitude de la production de certains types de sources telles que l'éolien ou le photovoltaïque ainsi que celle de la consommation, des méthodes de load flow probabilistes ont été développées. Ces méthodes peuvent être réparties en quatre catégories :

1. La méthode Monte Carlo qui consiste à calculer plusieurs milliers de fois le load flow déterministe en considérant des scénarios de production et de consommation aléatoires. Elles sont efficaces en termes de précision mais nécessitent beaucoup de temps de simulation.
2. Les méthodes analytiques basées sur une linéarisation des équations de load flow et sur des opérations arithmétiques, expriment les moments ou cumulants des paramètres de sortie du load flow en fonction de ceux des paramètres aléatoires d'entrée. Elles requièrent moins de temps de simulation mais sont moins précises que la méthode Monte Carlo.
3. Les méthodes approximées, telle que la méthode PEM qui est basée sur une approximation de la densité de probabilité des paramètres aléatoires d'entrée en quelques points représentatifs des moments d'ordre un, deux, trois et même parfois quatre. Chaque point est pondéré par un poids. Le load flow déterministe est calculé en remplaçant chaque paramètre d'entrée par ses points d'estimations à tour de rôle. Elle nécessite moins de temps de simulation que la méthode MC et offre une meilleure précision que les méthodes analytiques.
4. Les autres méthodes qui sont des méthodes rencontrées dans notre étude bibliographique, sont généralement basées sur une combinaison des méthodes citées ci dessus.

Notre étude s'est portée sur la méthode Monte Carlo, parce qu'elle est considérée comme référence en termes de précision, et la méthode PEM pour le compromis temps de calcul/précision qu'elle offre. Ces deux méthodes seront appliquées pour estimer la tension et le courant de neutre dans deux réseaux de tailles différentes avec différents types de modèles probabilistes de production et de consommation.

FIGURE 3.5 – Principe de l'estimation par la méthode PEM





# Chapitre 4

## Cas d'études

### 4.1 Introduction

La consommation raccordée sur le réseau de distribution basse tension est essentiellement constituée de charges monophasées telles que les foyers résidentiels ou l'éclairage public. Ce qui implique un couplage en étoile au secondaire des transformateurs de ce réseau, et l'existence d'un quatrième fil (le neutre) dont la section peut être inférieure à celle des trois autres phases. La différence des puissances des charges connectées sur les phases crée du déséquilibre auquel s'ajoute (ou se compense) celui créé par la différence des puissances PV monophasées installées sur les toits des bâtiments et raccordées au réseau. Ce déséquilibre fait croître le courant dans le conducteur de neutre qui peut atteindre des valeurs supérieures à la capacité physique de ce dernier.

L'utilisation de câbles souterrains en zone urbaine implique un accès difficile aux conducteurs pour d'éventuelles actions de maintenance et de renouvellement. De ce fait, en phase de planification, il est nécessaire d'avoir un bon dimensionnement des conducteurs. Ceci en estimant au préalable l'intensité à supporter selon les cas de production et de consommation dans la zone à connecter.

Le but de ce chapitre est l'application des méthodes de load flow probabilistes étudiées (méthode Monte Carlo et méthode PEM) pour l'estimation des critères de fiabilité comme le courant dans le neutre et la tension d'un réseau de distribution moyenne tension et par la même occasion, de proposer un outil de quantification de l'impact du déséquilibre de puissance de PV installée sur le courant de neutre.

Dans un premier temps, les méthodes d'approximations PEM et clustering sont appliquées aux variables d'entrée prises une à une dans le but d'évaluer leurs performances et limites en fonction du type de variable en considérant la méthode MC comme référence. Ensuite elles sont appliquées à l'ensemble des variables dans le cadre d'une étude comparative plus générale.

Deux réseaux tests sont utilisés dans ce chapitre :

- Un réseau de petit taille (4 nœuds) pour développer les études préalables et comparer les méthodes
- Un réseau de grande taille (900 nœuds) pour développer l'outil de quantification introduit ci dessus sur un système réaliste et vérifier la sensibilité des méthodes à la taille du réseau.

## 4.2 Réseau Test IEEE 4 noeuds

### 4.2.1 Le réseau

Dans un premier temps un petit réseau de distribution moyenne tension est utilisé dans le but de tester les méthodes de calcul présentées dans le chapitre 3. Les données du réseau sont fournies par l'organisme IEEE et repris dans les travaux de [4]. Il s'agit

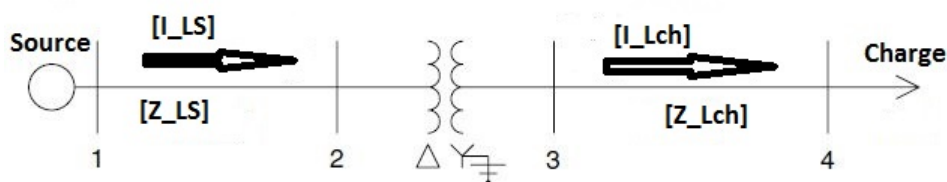


FIGURE 4.1 – Zone d'étude

d'un réseau de distribution moyenne tension triphasé alimenté par une source de tension triphasée équilibrée caractérisée par une tension entre phase  $U_S$  de  $12,7kV$ .

La ligne entre les noeuds 1 et 2 désignée, par  $L_{12}$ , est une ligne aérienne triphasée à trois fils connectés au primaire de trois transformateurs monophasés couplés en triangle. Elle est longue de 2000 pieds soit environ 600 m. On note respectivement par  $[Z_{L_{12}}]$  et  $[I_{L_{12}}]$  la matrice d'impédances de cette ligne et le vecteur des courants la traversant. Les valeurs des impédances sont données en Ohms dans la matrice suivante :

$$[[Z_{L_{12}}]] = \begin{bmatrix} 0.14 + j0.53 & 0.03 + j0.32 & 0.03 + j0.27 \\ 0.03 + j0.32 & 0.14 + j0.53 & 0.03 + j0.29 \\ 0.03 + j0.27 & 0.03 + j0.29 & 0.14 + j0.53 \end{bmatrix} \quad (4.1)$$

Le second segment de ligne désigné par  $L_{34}$  reliant les noeuds 3 et 4 est une ligne triphasée aérienne à quatre fils cette fois, alimentant un groupe de charges déséquilibrées et connectées en étoile. Elle est longue de 2500 pieds soit environ 800 m. On note par  $[Z_{L_{34}}]$  et  $[I_{L_{34}}]$  respectivement la matrice d'impédance de cette ligne et le vecteur des courants la traversant. La matrice d'impédance de cette ligne, réduite par réduction de Kron et donnée aussi en Ohms, vaut :

$$[Z_{L_{34}}] = \begin{bmatrix} 0.19 + j0.50 & 0.06 + j0.23 & 0.05 + j0.17 \\ 0.06 + j0.23 & 0.19 + j0.48 & 0.06 + j0.19 \\ 0.05 + j0.17 & 0.06 + j0.19 & 0.19 + j0.49 \end{bmatrix} \quad (4.2)$$

Le conducteur de neutre, en aluminium renforcé avec un brin d'acier pour un diamètre total d'environ de 14 mm. Ce qui implique une section d'environ de  $160 \text{ mm}^2$  [95] et donc un courant nominal autour de 320 A. Les conducteurs des phases des deux portions de ligne sont également faits en aluminium renforcé avec des brins d'acier pour un diamètre total de 18 mm, donc une section d'environ  $262 \text{ mm}^2$ , d'où un courant nominal fixé autour de 524 A.

Le groupe transformateur, couplé en mode Triangle-Etoile avec neutre à la terre, est composé de trois transformateurs monophasés de mêmes caractéristiques, à savoir une puissance de 2000 kVA chacun, une tension de 12,47 kV au primaire pour 2,4 KV au



secondaire et une impédance (en pu) de  $z_t = (1, 0 + j6, 0)\%$ .

Ce réseau alimente une charge triphasée déséquilibrée connectée au nœud 4, branchée en étoile et définie par son modèle PQ donné par :

$S_a = 750$  kVA avec un facteur de puissance de 0,85

$S_b = 1000$  kVA avec un facteur de puissance de 0,90

$S_c = 1250$  kVA avec un facteur de puissance de 0,95

Les résultats de l'application de la méthode déterministe FBS peuvent être consulter en Annexe 3 et permettent d'évaluer la répartition des courants et tensions sur les lignes et nœuds du réseau en un instant donné sachant que toute la production et la consommation sont raccordées au nœud 4 du réseau.

### **Application des méthodes LFP en mode instantané**

Les méthodes PEM et MC présentées en chapitre 3 sont appliquées. Les puissances actives  $P_{4a}$ ,  $P_{4b}$ ,  $P_{4c}$  des charges raccordées aux phase a, b, et c du nœud 4 sont modélisées par des lois normales de moyennes égales aux valeurs nominales des charges fournies par les données du réseau ci-dessus et d'écart type égaux à 250 kW chacun. On suppose que le facteur de puissance de la charge est connu et fixé égal aux valeurs par défaut pour chaque phase.

Les puissances injectées par les trois panneaux photovoltaïques  $P_{4a_{pv}}$ ,  $P_{4b_{pv}}$ ,  $P_{4c_{pv}}$  suivent le modèle présenté au chapitre 2 en distinguant la nébulosité, probabiliste et modélisée par une loi de probabilité bêta, de la puissance théorique installée qui est déterministe. Les panneaux photovoltaïques, étant dans la même zone, sont soumis au même paramètre aléatoire de nébulosité. Ce problème comporte alors 4 paramètres aléatoires.

### **Résultats**

Les valeurs moyennes de la tension au nœud 4 ainsi que les courants dans les phases et le neutre de la ligne connectée à la charge sont présentées dans le tableau 4.2.1 récapitulant les résultats obtenus par la méthode PEM, la méthode MC et la méthode de load flow déterministe calculée avec les valeurs moyennes des variables aléatoires d'entrée.

TABLE 4.1 – Comparaison des résultats des load flow déterministe, Monte Carlo et PEM appliqués au réseau 4 nœuds

	Phases	LFD	$\mu_{PEM}$	$\sigma_{PEM}$	$\mu_{MC}$	$\sigma_{MC}$
$V4(\text{pu})$	a	0,97	0,95	0,03	0,95	0,03
	b	0,94	0,93	0,03	0,93	0,03
	c	0,92	0,92	0,02	0,92	0,029
$ I4 (A)$	a	216,57	240,66	127,04	242,43	123,5
	b	249,13	280,15	150,60	285,6	147,36
	c	424,44	430,14	152,56	431,41	152,31
$arg(I4)(^\circ)$	a	-82,46	-97,34	31,6	-83,77	32,53
	b	158,41	66,06	132,59	92,74	112,02
	c	60,97	58,91	2,88	58,05	4,57
$ I_{neutre} (A)$		244,67	283,97	140,19	288,05	139,92
$arg(I_{neutre})(^\circ)$		88,42	84,74	56,71	68,34	63,95

Les moyennes et écarts-types des modules et phases des grandeurs estimés par les méthodes de MC et PEM sont dans le même ordre de grandeur pour un temps de simulation de 28 secondes pour la première contre 1.95 seconde pour la seconde. Cette première comparaison est basée sur un cas simple en mode statique avec des puissances de charge modélisées par des lois normales et un paramètre de nébulosité modélisé par une loi bêta de paramètres  $a = 0.0037$  et  $b = 0.0117$ . Dans une telle configuration la méthode PEM et la méthode MC aboutissent à des résultats quasi similaires.

Dans cette partie, une étude en mode statique sur un seul instant horaire donné a été faite. Dans les deux sections suivantes, l'étude en série temporelle à l'horizon d'une journée sera menée. Ceci sera l'occasion de tester les modèles de production et de consommation en séries temporelles développés au chapitre 2. La première section ci dessous traite de l'efficacité d'une méthode de LFP donnée en fonction du type de variable. Les méthodes utilisées sont la méthode PEM et une méthode basée sur le clustering qui se basent toutes les deux sur le principe de la réduction de l'incertitude.

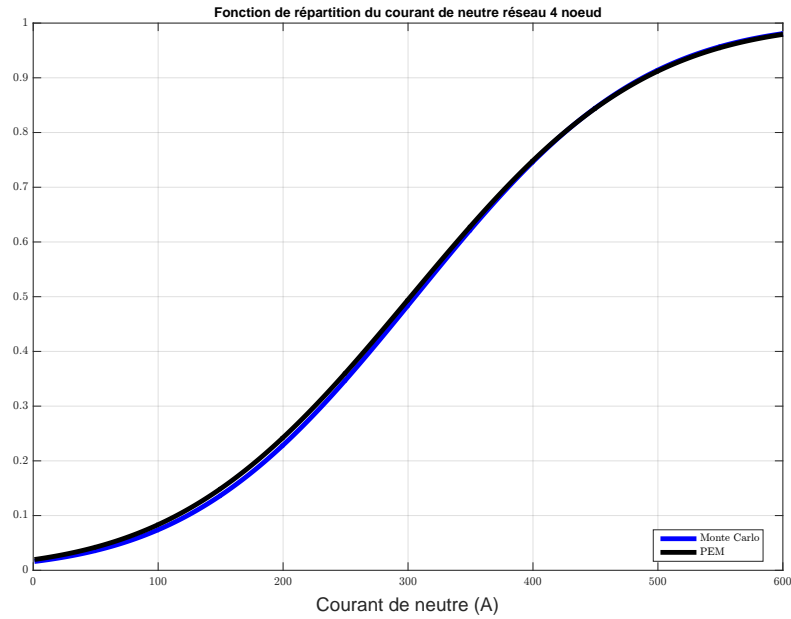


FIGURE 4.2 – Fonction de répartition des valeurs de courant de neutre dans la ligne  $L_{34}$  du réseau 4 nœuds

### 4.3 Réduction des incertitudes

Dans le but de vérifier l'intérêt et la pertinence d'utilisation des méthodes d'approximation de LFP, qu'elles soient basées sur la méthode PEM ou bien sur le clustering (pseudo MC), l'étude qui va suivre va comparer deux méthodes d'approximation en prenant la méthode MC comme référence. On considère les paramètres aléatoires d'entrée listés ci dessous :

- La nébulosité.
- La distance parcourue par les véhicules électriques.
- L'instant de début de recharge des VEs.
- Le profil de consommation des foyers.

Les lois de probabilité modélisant les paramètres de distances parcourues, d'instant de début de recharge et de profil de consommation interne sont indépendantes du numéro de phase. Dans une optique simplificatrice, on suppose que tout véhicule raccordé sur une phase donnée admet la même loi de probabilité de distance parcourue. La même hypothèse est considérée pour la consommation interne et l'instant de début de recharge.

Le principe de la réduction d'incertitude consiste à approximer une variable par ses points PEM, dans un premier temps, puis par les centres de ses clusters dans un second temps, tandis que les autres variables sont échantillonnées aléatoirement à partir de leur distribution de probabilité comme dans une méthode Monte Carlo. Ainsi on peut tester l'applicabilité d'une méthode (PEM ou clustering) en fonction du type de distribution et de la variable aléatoire considérée.

L'étude est appliquée au réseau test IEEE 4 nœuds pour des raisons de simplicité et de réduction de temps de calcul. Les cas de simulation considérés sont :

- Cas 1 (cas témoin) : Simulation Monte Carlo appliquée à toutes les variables. Ce cas sert de référence de comparaison.
- Cas 2 : Modèle réduit par PEM appliqué à une variable, simulation MC appliquée aux autres variables.
- Cas 3 : Modèle réduit par clustering appliqué à une variable, simulation MC appliquée aux autres variables.

### 4.3.1 Principe de la réduction

#### Réduction par clustering

Les scénarios possibles d'un paramètre aléatoire, par exemple la nébulosité, sont regroupés en clusters qui sont eux définis par leur profil type de nébulosité et leur probabilité d'occurrence. On est ainsi en présence de quatre profils type de nébulosité et donc de quatre profils type de production photovoltaïque (cf figure 4.3). Pour chacun de ces profils, on cherche à estimer la tension et le courant dans le conducteur de neutre du réseau test en tirant aléatoirement les valeurs des autres variables aléatoires (distances, instants de recharge, consommations internes). Le niveau de l'impact de chaque profil type sur le résultat final (courant de neutre ou tension) est pris en compte en pondérant le résultat correspondant à ce profil par sa probabilité d'occurrence.

#### Réduction par PEM

La réduction par méthode PEM consiste à appliquer la méthode PEM vue au chapitre 2 sur une variable à la fois. On calcule la tension et le courant de neutre du réseau test par

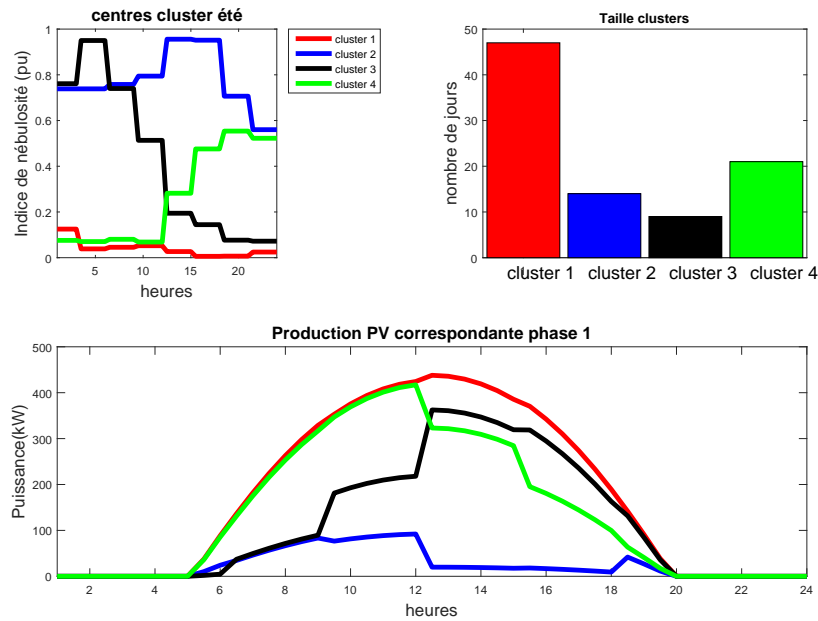


FIGURE 4.3 – Scénarios de production photovoltaïque en fonction du profil type de nébulosité

une méthode de load flow probabiliste en considérant un à un les points d'approximation du paramètre considéré. Le résultat de chaque calcul est pondéré par le poids du point d'approximation correspondant. La somme de tous les résultats pondérés par les poids respectifs des points utilisés en entrée donne le résultat final.

### 4.3.2 Cas du paramètre nébulosité

L'application de la réduction PEM à la distribution de nébulosité conduit, pour chaque pas de temps, à trois points qui impliquent trois scénarios de production PV possibles (cf figure 4.4). En regardant de plus près les points pour un pas de temps donné, à 14h par exemple, la forme de loi bêta implique une concentration de points, à très fort poids (plus de 70%), proches de zéro, une concentration de points proches de un et un troisième point situé sur la moyenne avec un poids quasi nul.

La réduction par clustering aboutit aux quatre clusters de la figure 4.3 qui correspondent, pour une saison d'été, à une journée ensoleillée, une journée nuageuse et une journée avec matinée nuageuse et après midi ensoleillée et vice et versa.

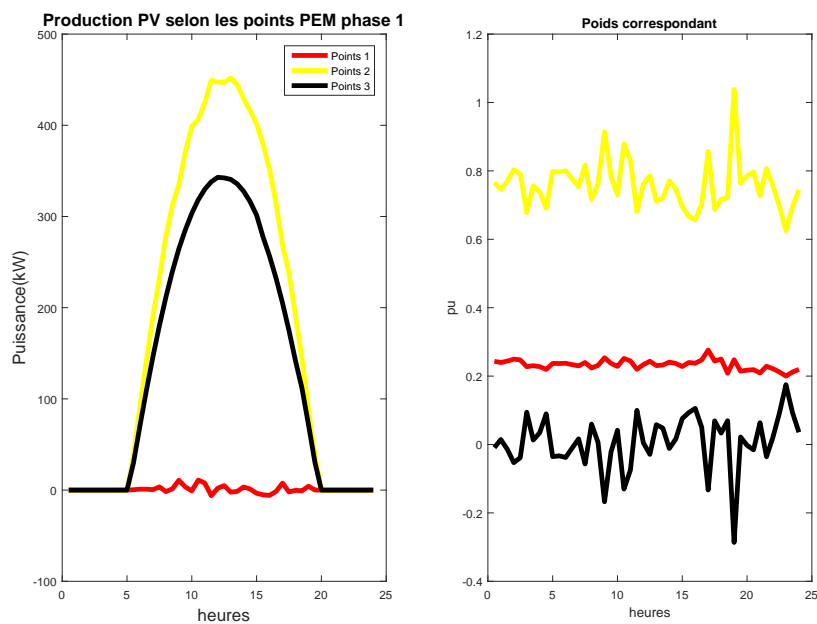


FIGURE 4.4 – Scénarios de production PV en fonction des points approximés par la méthode PEM

TABLE 4.2 – Points d’approximation PEM de la nébulosité à 14h

Points	0,992	0,003	0,239
Poids	0,240	0,770	-0,011

Le résultat (cf tableau 4.4) montre que pour certains instants horaires, les estimations de la méthode PEM sont aberrantes car les points d’approximation calculés sont en dehors des bornes de la loi bêta. Prenons par exemple, l’instant horaire correspondant à 16h. Les moments de la loi bêta caractérisant la nébulosité à cet instant sont donnés par le tableau 4.3 avec  $\mu$ ,  $\sigma$ ,  $\lambda_3$ ,  $\lambda_4$  désignant respectivement la moyenne, l’écart type, le coefficient de skewness (ou d’asymétrie) et de kurtosis (ou d’aplatissement)

TABLE 4.3 – Moments statistique de loi bêta de la nébulosité à 18h

$\mu$	$\sigma$	$\lambda_3$	$\lambda_4$
0.2554	0.4331	1.1219	2.3555

Le calcul des constantes  $\xi$  pour les 3 points à partir des équations 3.62 donne :

$$\xi = \begin{bmatrix} 1,74 \\ -0,62 \\ 0 \end{bmatrix} \quad (4.3)$$

Rappelons que chaque point  $k$  d'une variable aléatoire d'entrée  $pl$  est donné par :

$$P_{pl,k} = \mu_{pl} + \sigma_{pl}\xi_{pl,k} \quad (4.4)$$

En appliquant l'équation 4.4 à notre cas où la variable  $pl$  représenterait la nébulosité et  $k$  un nombre entier, et vues les données ci dessus, on obtient :

TABLE 4.4 – Concentration de points pour la nébulosité à 18h

	$k = 1$	$k = 2$	$k = 3$
$P_{pl,k}$	1,01	-0,01	0,25
Poids	0,24	0,67	-0,66

La conséquence est qu'il en résulte des valeurs de points de nébulosité qui sont en dehors des bornes fixées [0 1]. Ceci implique des puissances de production photovoltaïque qui sont soit négatives, soit supérieures à la puissance maximale théorique installée, ce qui est physiquement impossible. Pour cette raison, nous ne représenterons que les résultats obtenus par les méthode MC et clustering 4.5 et 4.6.

Pour une distribution bêta, l'approximation par clustering est plus appropriée que celle par méthode PEM.

### 4.3.3 Cas du paramètre consommation interne

Dans le cas de la réduction PEM, l'aspect aléatoire de la consommation interne est modélisé par un nombre aléatoire discret  $R_{int}$  qui détermine l'indice du cluster auquel appartient le profil type de consommation interne (cf chapitre 2). Une telle distribution peut être approximée, par les trois points du tableau 4.5 avec leur poids respectif.

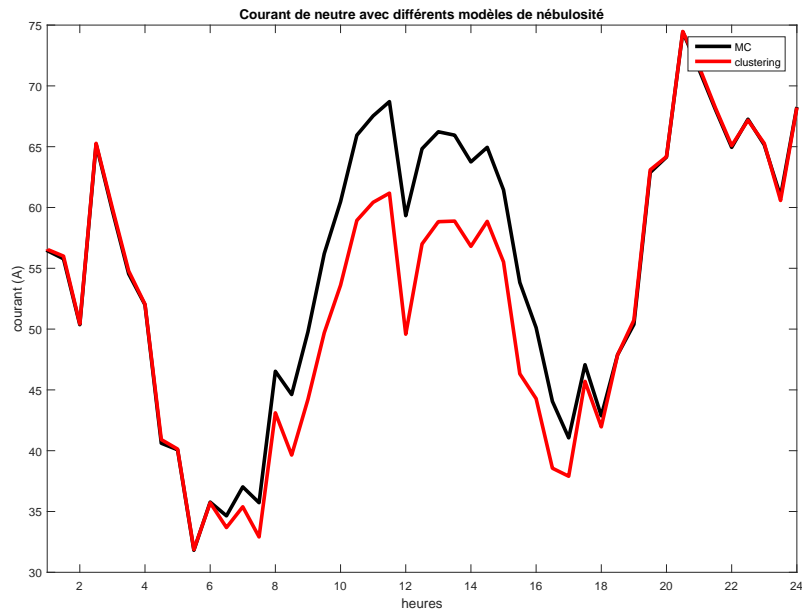


FIGURE 4.5 – Courant dans le neutre selon différents modèles de nébulosité

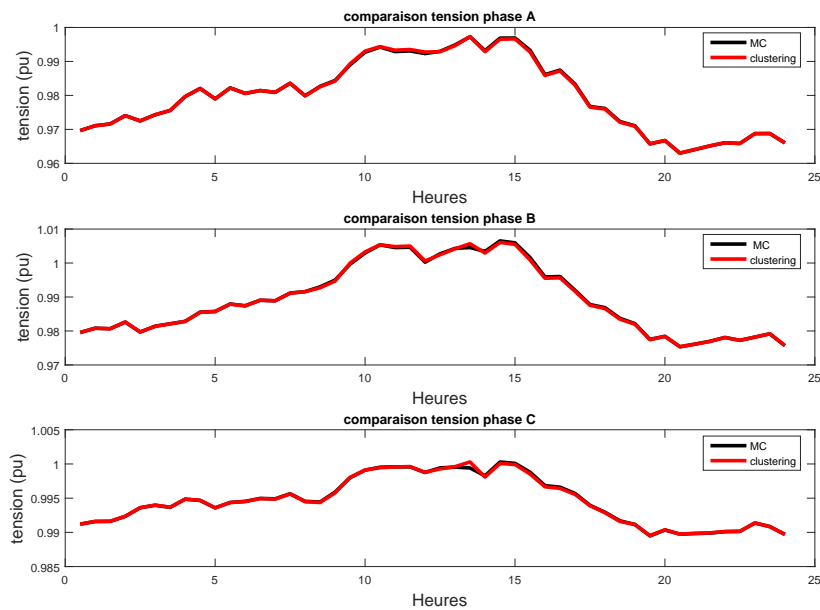


FIGURE 4.6 – Tensions sur chaque phase selon différents modèles de nébulosité

La nature discrète de la distribution oblige à arrondir les points calculés aux valeurs entières les plus proches qui sont 5, 1, et 3.



TABLE 4.5 – Points d’approximation PEM de la variable consommation interne

Points	4,9	1,02	3,13
Poids	0,31	0,26	0,42

L’application de la méthode de clustering K-means aux données de consommation étudiées aboutit aux cinq clusters (cf chapitre 2) dont les profils types et distributions sont représentés sur la figure 4.7.

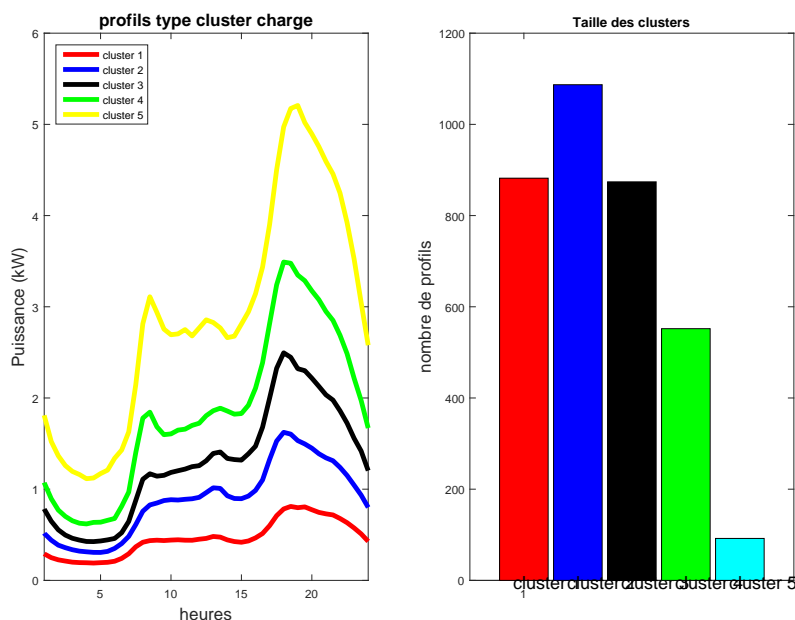


FIGURE 4.7 – Profils types et tailles des clusters de consommation

Les profils types des clusters montrent des habitudes de consommation similaires mais de niveaux de puissances différents. La répartition est surtout basée sur le niveau d’énergie consommée.

Les figures 4.8 et 4.9 montrent que les résultats obtenus pour chacune des méthodes sont proches de ceux obtenus par la méthode de Monte Carlo. Donc pour ce type de distribution, il n’y a pas de différences significatives entre la réduction par clustering et celle par méthode PEM.

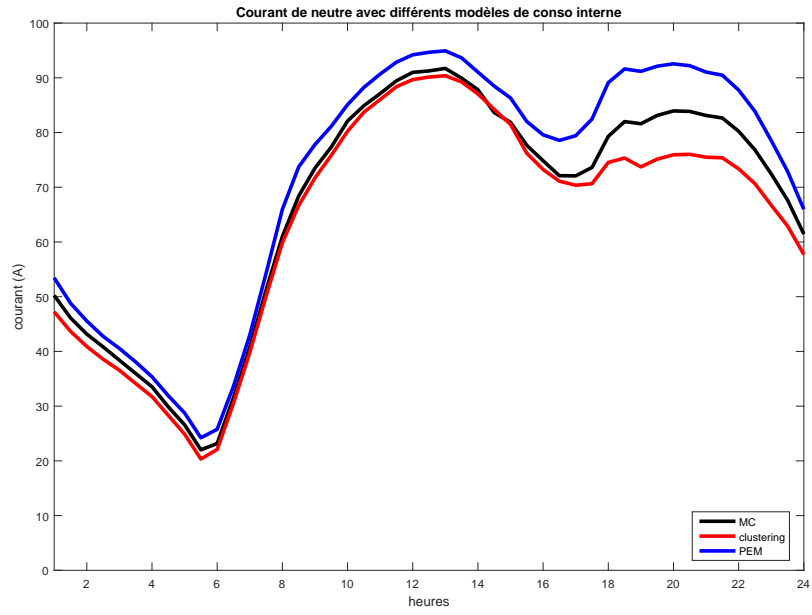


FIGURE 4.8 – Courant dans le neutre en fonction du modèle de consommation interne

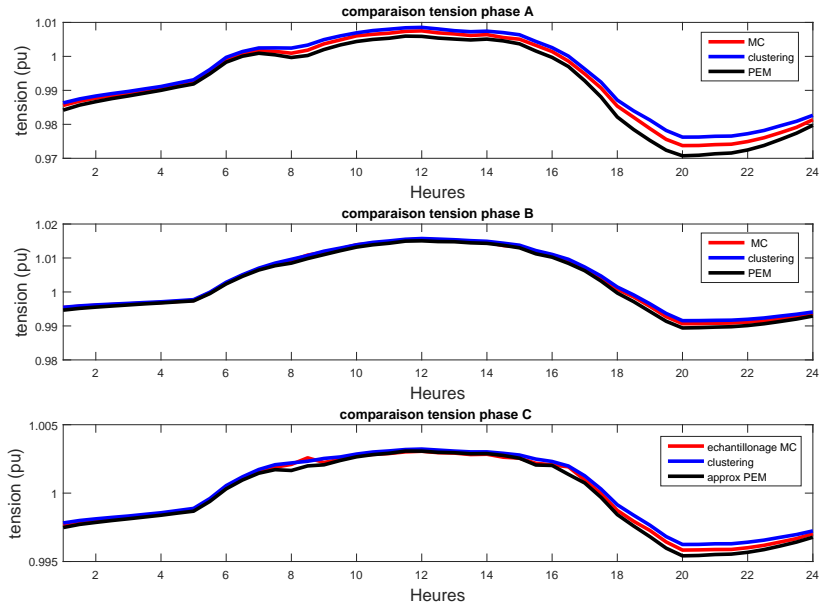


FIGURE 4.9 – Tensions en fonction du modèle de consommation interne

### 4.3.4 Cas du paramètre distance parcourue VE

La réduction de l'incertitude par méthode PEM appliquée au paramètre distance parcourue conduit aux points du tableau 4.6 avec leurs poids respectifs. On note une forte concentration des valeurs autour de la moyenne.

TABLE 4.6 – Points d'approximation PEM de la variable Distance

Points	35,61	27,67	30
Poids	0,02	0,04	0,94

La réduction par clustering du paramètre distance a consisté à appliquer un algorithme de répartition à un échantillon de 10000 valeurs de distances obtenues à partir de la distribution de probabilité lognormale (moyenne 30 km et écart type 0.85). Le nombre de clusters est choisi en faisant un compromis entre la taille du problème (puisqu'il faudra tester pour chaque cluster) et une segmentation raisonnable en fonction de la distribution des données.

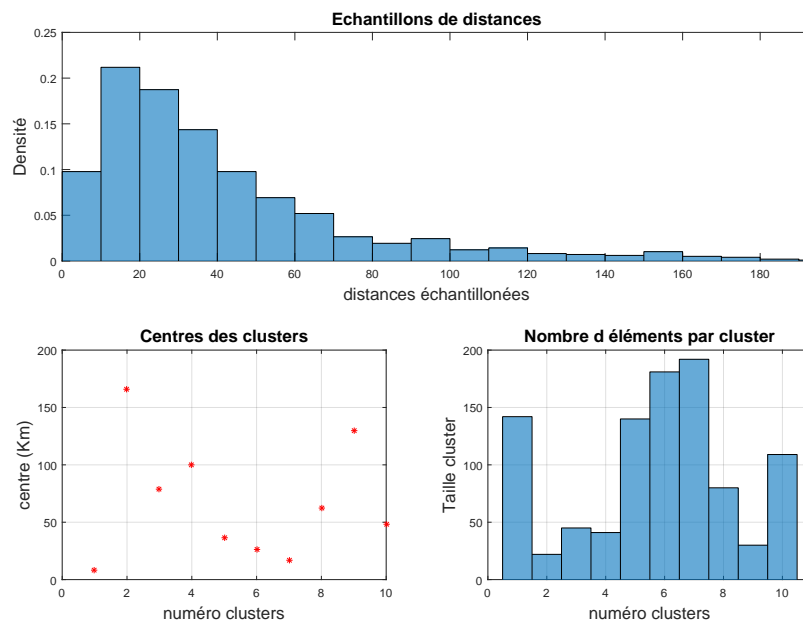


FIGURE 4.10 – clustering des données de distances

Ici nous avons choisi arbitrairement un nombre de clusters égal à 10. La figure 4.10 représente les échantillons de distances, générés selon une loi lognormale, et sur lesquels le clustering est appliqué. Sur la même figure sont aussi représentées les distances types qui constituent les centres des clusters ainsi que la distribution des tailles. Les clusters 7, 6, 5, et 1, de centres respectifs situés autour de la moyenne de la distribution lognormale, à savoir 16,45 km, 36,5 km, 21,3 km et 8,6 km, sont les plus représentatifs de la série avec des probabilités d'occurrence de 19,5%, 18,4%, 14,2% et 14,5%. Ceci s'explique par le fait que les parcours quotidiens des usagers particuliers (qui sont la base de cette étude) se font majoritairement sur des moyennes à faibles distances. Suivent les clusters 10, 8, 3 et 4, de centres situés entre 48 et 100 km (48 km, 62 km, 78 km et 99 km) et de probabilités d'occurrence 11,1%, 8,1%, 4,6% et 4,2%. Enfin les clusters 9 et 2, représentant les distances supérieures à 100 km, ont les probabilités d'occurrence les plus rares situées respectivement autour de 3% et 2%.

L'application des méthodes d'approximation PEM et clustering sur l'unique paramètre de distance parcourue aboutit aux résultats présentés sur les figures 4.11 et 4.12. Le clustering est mieux approprié par ce qu'il prend en compte aussi les longues distances de parcours même si leur fréquence d'occurrence est très faible. Au contraire, les approximations de la méthode PEM sont limitées seulement aux valeurs proches de la moyenne (cf tableau 4.6). Les distances qui sont au delà de 40 km ne sont pas prises en compte.

### 4.3.5 Conclusion

Les types de distribution utilisés pour tester la pertinence des méthodes de réduction sont choisis de sorte à couvrir plusieurs formes : Une forme de distribution convexe centrée autour de sa moyenne (distribution normale), une distribution convexe asymétrique (distribution lognormale), une distribution concave (distribution bêta) et une loi discrète. Le tableau 4.7 permet d'avoir un récapitulatif du comportement des méthodes d'approximation en fonction de la nature de distribution. La méthode PEM est efficace pour des distributions de type concave mais fournit des résultats aberrants pour la distribution de type bêta et peut fausser le modèle dans le cas d'un paramètre aléatoire discret. Au

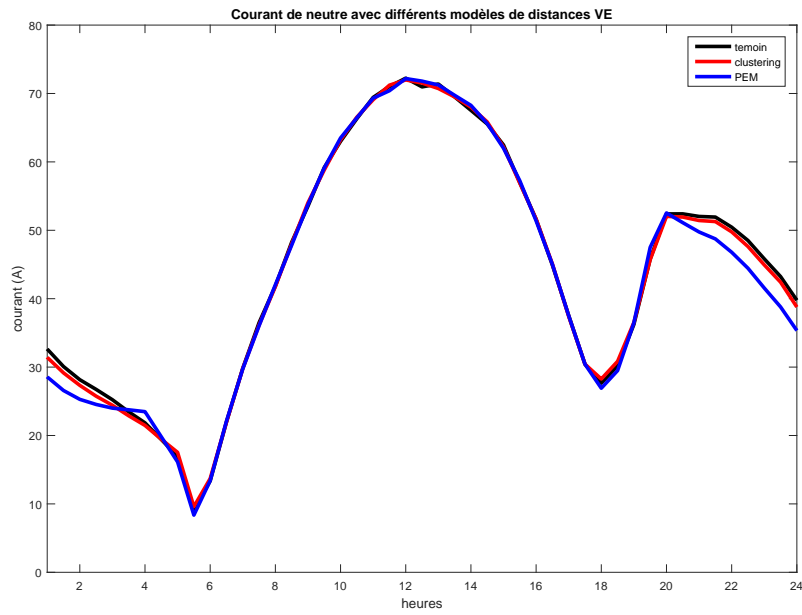


FIGURE 4.11 – Courant dans le neutre selon modèles du paramètre distance

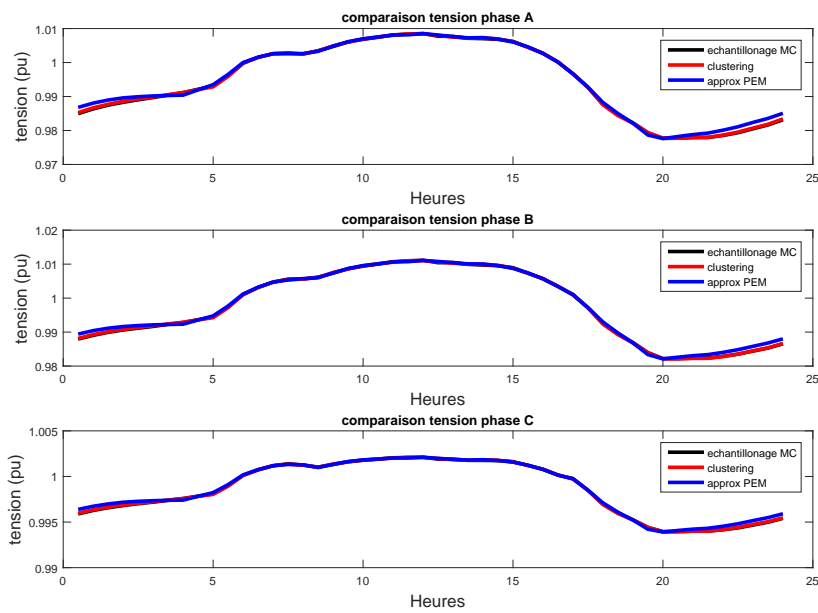


FIGURE 4.12 – Tensions selon les modèles du paramètre distance

contraire, la méthode basée sur le clustering a un domaine d'applicabilité plus large.

TABLE 4.7 – Récapitulatif approximation PEM et clustering en fonction de type de distribution

Distribution	Normale	Bêta	Discrète	Lognormale
PEM	=	-	=	$\approx$
Clustering	=	+	=	$\approx$

Il vient d’être démontré dans cette partie que les performances de la méthode PEM sont limitées pour certains type de distribution de variable aléatoire ce qui n’est pas le cas pour la méthode basée sur le clustering, appelée également méthode pseudo MC, qui sera utilisée dans la suite des travaux.

## 4.4 Comparaison entre méthode MC et méthode pseudo MC appliquées à toutes les variables

Les méthodes MC et pseudo MC sont appliquées au réseau test 4 nœuds dans le but d’une étude comparative plus globale.

### 4.4.1 Modèles des paramètres pour la méthode MC

Le caractère aléatoire de la production photovoltaïque est toujours modélisé par l’indice de nébulosité. Pour chaque pas de temps, la distribution des données de nébulosité est approximée par une loi bêta. Avec une distribution toutes les demies heures, on est en présence de 48 lois bêta de moyennes comprises entre 0.2 et 0.3. La surface totale de panneaux installés sur chaque phase est choisie arbitrairement égale à 4000, 5000 et 6000 $m^2$  de manière à générer du déséquilibre. La figure 4.13 montre un exemple d’échantillon de profil de puissance produite sur chaque phase du nœud 4. Comme expliqué au chapitre 2, la demande de chaque maison est modélisée en distinguant la consommation des véhicules électriques (VE) de celle des autres équipements domestiques. Pour éviter de saturer le transformateur de 2000 $kVA$ , on suppose que 200 maisons sont connectées à la phase  $a$ , 250 à la phase  $b$  et 300 à la phase  $c$ . Chaque maison dispose d’un VE et d’une unité de production PV raccordée au réseau.

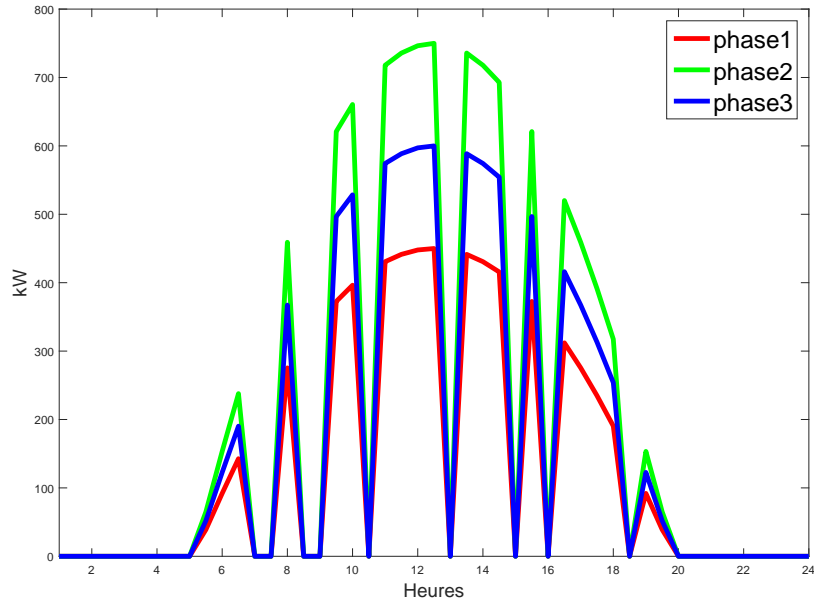


FIGURE 4.13 – Exemple Production photovoltaïque totale sur chaque phase

Le modèle de recharge des VE est le même que celui présenté au chapitre 2 avec une distance de parcours suivant une loi lognormale et l'heure de début de recharge qui est un nombre entier uniformément distribué entre le début de la plage horaire de recharge et l'instant maximal de début de recharge, c'est à dire l'instant au plus tard à partir duquel, la voiture doit commencer sa recharge pour avoir le niveau d'énergie nécessaire à son trajet et ce avant 8h du matin.

La figure 4.14 montre les profils de recharge des VE au niveau de chaque phase du nœud 4 en supposant une distance parcourue égale à la valeur moyenne de 30km et un instant de début de recharge choisi aléatoirement suivant une loi uniforme.

Le profil de consommation interne est modélisé par une variable aléatoire discrète (comme vue au chapitre 2) qui détermine le profil type d'un cluster parmi les cinq clusters de consommation interne (cf figure 4.7).

#### 4.4.2 Méthode pseudo MC

La méthode pseudo MC ou méthode de LFP basée sur le clustering consiste à répartir les données de chaque paramètre d'entrée en clusters comme vue précédemment. Ensuite

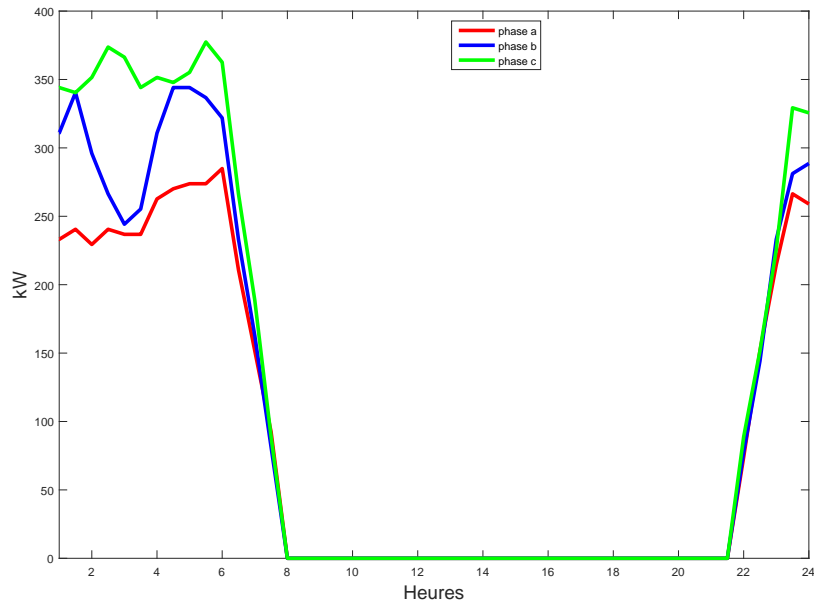


FIGURE 4.14 – Recharge totale des VE sur chaque phase

une pseudo méthode de Monte Carlo est appliquée en considérant, non pas des échantillons de valeurs aléatoires, mais les centres des clusters de chaque paramètre jusqu'à couvrir l'ensemble des combinaisons possibles de centres des clusters sur les différents paramètres. Chaque combinaison est pondérée par le produit normalisé des probabilités d'occurrence des clusters la constituant, de sorte à avoir la somme de toutes les probabilités des combinaisons égale à 1. Le résultat final est la somme pondérée des résultats des combinaisons. On l'appelle méthode pseudo MC ou quasi déterministe puisque l'incertitude sur les paramètres d'entrée est réduite à un nombre fini de clusters définis par leurs centres.

Les données utilisées impliquent que les profils de nébulosité peuvent être répartis en 4 clusters pondérés par leur probabilité d'occurrence, les consommations internes par 5 clusters, et 10 clusters pour les distances parcourues par les VEs. Le paramètre instant de début de recharge sera tiré aléatoirement pour chaque combinaison. Le total des combinaisons sera égal à  $4 \times 5 \times 10 = 200$ . Cette méthode pseudo MC nécessite alors 200 calculs de load flows déterministes, ce qui réduit considérablement le temps de simulation par rapport à une méthode de MC classique.



Appliquée sur le réseau test à 4 nœuds, cette méthode ne nécessite que 5 secondes de temps de simulation pour un résultat quasi similaire à la méthode Monte Carlo en termes de valeurs moyennes (cf figure 4.15). Cependant, la valeur moyenne seule ne four-

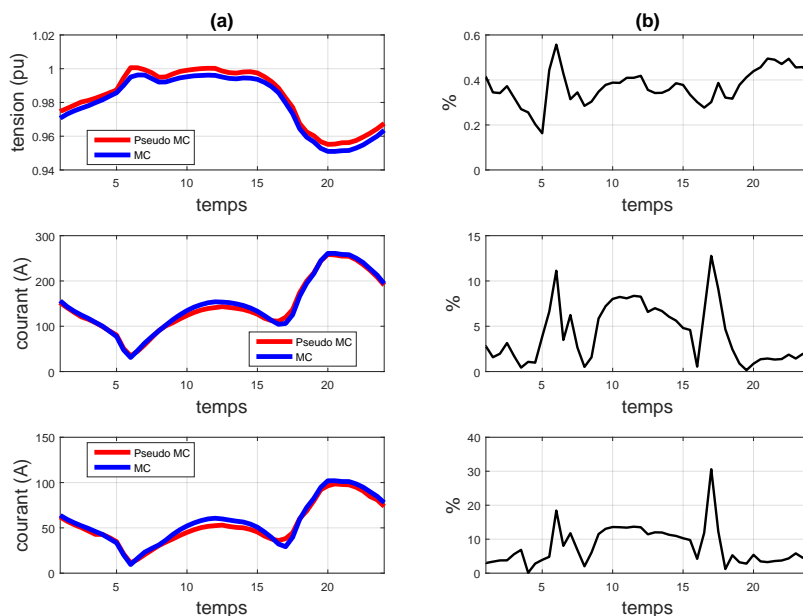


FIGURE 4.15 – Comparaison des valeurs moyennes estimées par les méthodes MC et Pseudo MC

(a) : Valeurs moyennes de la tension au nœud 4, du courant de phase A et du courant de neutre de la ligne  $L_{34}$  estimées par méthode MC et par la méthode pseudo MC

(b) : Erreur relative sur les estimations des valeurs moyennes

nit pas assez d'information sur les possibles variations du courant de neutre en fonction des scénarios. Elle ne permet pas, par exemple, d'avoir une idée précise du risque potentiel d'atteindre le seuil critique de valeur du courant de neutre. Ce genre d'information nécessite de déterminer la densité de probabilité et la fonction de répartition.

La figure 4.16 montre des valeurs extrêmes de courant plus importantes sur la distribution obtenue par méthode pseudo MC que celle obtenue par méthode MC. Ceci s'explique par le fait que la répartition en cluster des valeurs possibles des paramètres d'entrée permet de mieux prendre en compte les valeurs extrêmes même si leur probabilité d'occurrence est faible.

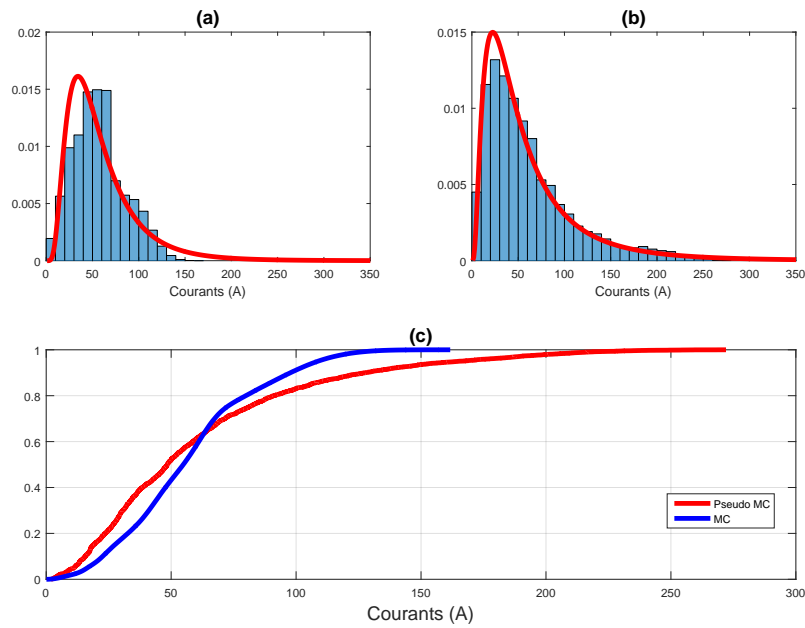


FIGURE 4.16 – Distribution du courant de neutre sur l'ensemble des scénarios de MC et pseudo MC appliqués au réseau 4 noeuds sur une journée

(a) : Estimation densité de probabilité en fonction des résultats des scénarios de la méthode MC

(b) : Estimation densité de probabilité en fonction des résultats des scénarios de la méthode pseudo MC

(c) : Comparaison des fonctions de répartition empiriques des méthode MC et pseudo MC

En plus du courant de neutre, la tension constitue aussi un paramètre de la qualité de l'alimentation électrique fournie par le réseau. Les distributions de tensions estimées par la méthode MC (cf figure 4.17) ne correspondent à aucune des lois de probabilité classiques. De ce fait une technique d'estimation par noyau est utilisée pour déterminer la loi de probabilité correspondante. L'estimation par noyau, qui est une généralisation de celle par histogrammes, consiste à estimer la densité de chaque point d'observation par une gaussienne centrée en ce point et de largeur définie par un paramètre de lissage  $h$ . L'estimateur final, qui est une loi appelée Kernel, est formé par la moyenne des gaussiennes en chaque point.

Les distributions de tensions estimées par méthode pseudo MC sont approximées par des lois de Weibull (cf figure 4.17). Les différences entre les deux histogrammes s'expliquent par le nombre d'échantillon observé qui est plus élevé dans le cas de la méthode MC.

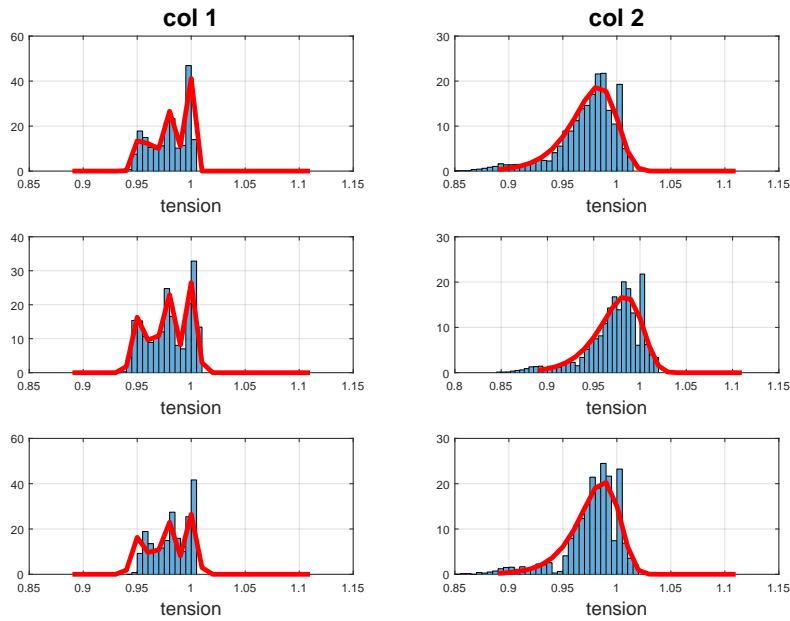


FIGURE 4.17 – Distribution des tensions sur les trois phases du nœud 4 du réseau 4 nœuds :

col 1 : Distributions et densités de probabilités des tensions sur les phases A, B et C estimées par la méthode MC

col 2 : Distributions et densités de probabilités des tensions sur les phases A, B et C estimées par la méthode pseudo MC

### 4.4.3 Conclusion

La méthode pseudo MC, appliquée à l'ensemble des variables, aboutit à des résultats du même ordre de grandeur que la méthode MC pour un temps de simulation 60 fois moins important (dans le cas du réseau test IEEE 4 nœuds). Dans la partie qui suit, ces résultats seront confrontés en appliquant les deux méthodes sur un réseau test de plus grande taille.

## 4.5 Réseau Test de grande taille

Cette partie a un double objectif qui consiste à :

- Tester le comportement de la méthode pseudo Monte Carlo, introduite précédemment, sur un réseau de grande taille. Les estimations seront comparées à celles obtenues par la méthode de Monte Carlo qui sert de référence. La méthode PEM

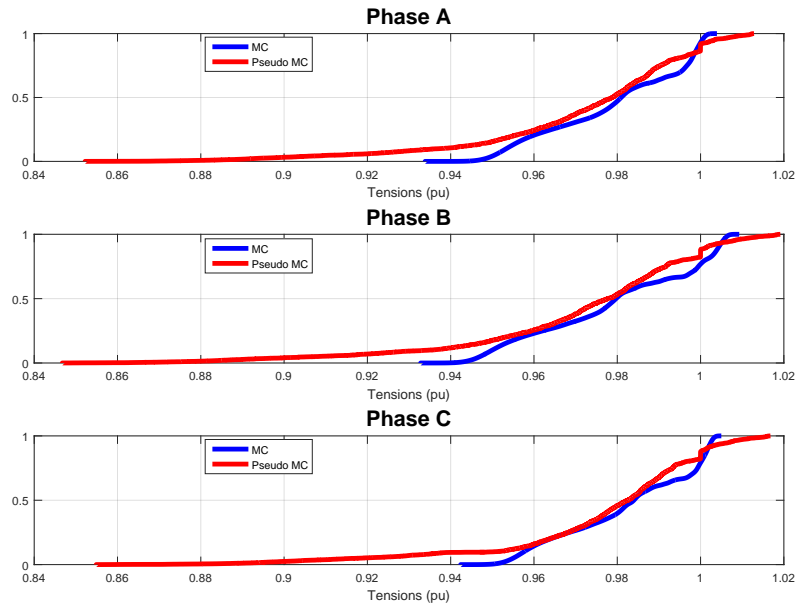


FIGURE 4.18 – Fonctions de répartition empiriques des tensions sur les trois phases du nœud 4 estimées par méthode MC et pseudo MC

n'est pas appliquée puisque les résultats précédents ont montré ses limites sur une distribution bêta.

- Appliquer une des méthodes LFP pour quantifier l'impact du déséquilibre de production PV sur la probabilité de dépassement des limites du courant de neutre.

Le réseau test *ckt5* de l'EPRI est utilisé. Ce réseau couvre une zone géographique de  $205\text{km}^2$  environ. Le réseau *ckt5* est de topologie américaine, c'est à dire qu'il est composé d'une grande partie de moyenne tension ( $7.2\text{kV}$ ). Des portions à basse tension sont alimentées via des transformateurs monophasés  $7.2\text{kV}/240\text{V}$ , de puissances de 15, 25, 37.5, et 50 kVA. Chaque transformateur alimente entre 2 et 6 maisons selon sa puissance nominale.

Le réseau *ckt5* dans sa totalité est trop vaste pour notre étude. De ce fait, une zone de  $6\text{km}^2$  est définie sur ce réseau comme l'indique la figure 4.19. Cette zone correspond à 49 portions de réseaux BT alimentés par 147 ( $49 \times 3$ ) transformateurs monophasés  $7.2\text{kV}/240\text{V}$  à 25 kVA. Chaque phase d'une portion alimente alors 3 à 5 consommateurs basse tension monophasés.

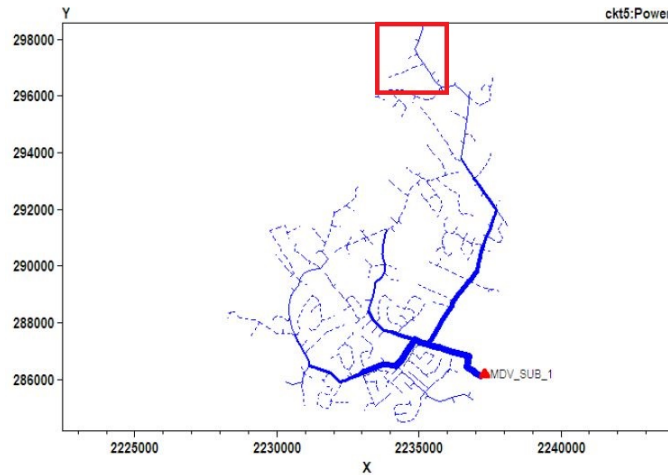


FIGURE 4.19 – Zone d'étude

## 4.6 Résultats

La figure 4.20 montre une comparaison entre les valeurs moyennes du courant de neutre estimées par méthode MC et pseudo MC appliquées au réseau ckt5. La ligne triphasée sélectionnée pour montrer les résultats alimente une dizaine de maisons dont quatre sont connectées sur la phase A, quatre sur la phase B et trois sur la phase C. Les surfaces totales de panneaux installés sur chaque phase sont respectivement 59, 53 et 42 mètres carrés. Chaque maison est équipée d'une borne de recharge de véhicule électrique de puissance égale à 3.7 kW et d'autres équipements internes dont le profil de consommation correspond à un des profils type des clusters définis en chapitre 2.

La comparaison des valeurs moyennes estimées (cf figure 4.20) confirme les conclusions préalablement observées sur le réseau de petite taille. La méthode de pseudo clustering, pour un temps de simulation largement inférieur à la méthode Monte Carlo, fournit des estimations de valeurs moyennes dont la précision est proche de celle de la méthode MC.

La distribution des résultats (cf figure 4.21) de courant de neutre observés sur l'ensemble des scénarios MC et pseudo MC sur toute une journée type d'été permet d'estimer de manière empirique le risque global de dépassement de courant de neutre auquel on peut s'attendre. La partie du réseau observée étant en MT, elle est alors surdimensionnée par rapport au nombre de charges et de productions modélisées. Il faut noter qu'ici l'objectif, dans un premier temps, est de comparer le comportement des deux méthodes en termes

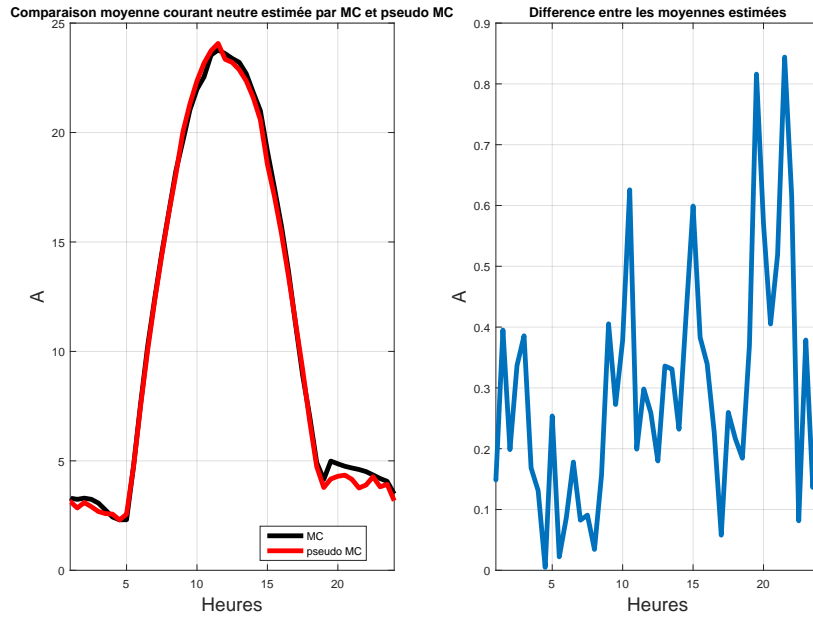


FIGURE 4.20 – Comparaison des valeurs moyennes estimées par les méthodes MC et Pseudo MC

de densité de probabilité et de fonction de répartition indépendamment de la limite de courant de neutre fixée. En supposant cette limite égale à 25 A, le risque de dépassement observé est de l'ordre de 3% pour avec la méthode MC et 9% avec la méthode pseudo MC. Ce risque est plus élevé avec la méthode pseudo MC puisqu'elle permet de mieux considérer les scénarios de production et de consommation extrêmes.

La partie du réseau observée étant en MT, elle est alors surdimensionnée par rapport au nombre de charges et de productions modélisées. Il faut noter qu'ici l'objectif, dans un premier temps, est de comparer le comportement des deux méthodes en termes de densité de probabilité et de fonction de répartition indépendamment de la limite de courant de neutre fixée. En supposant cette limite égale à 25 A

## 4.7 Applications

Comme expliqué en introduction, avec le raccordement de production PV domestique, le réseau de distribution d'électricité se trouve dans un environnement pour lequel il n'a pas été dimensionné ou structuré au départ. L'un des intérêts des méthodes de LFP

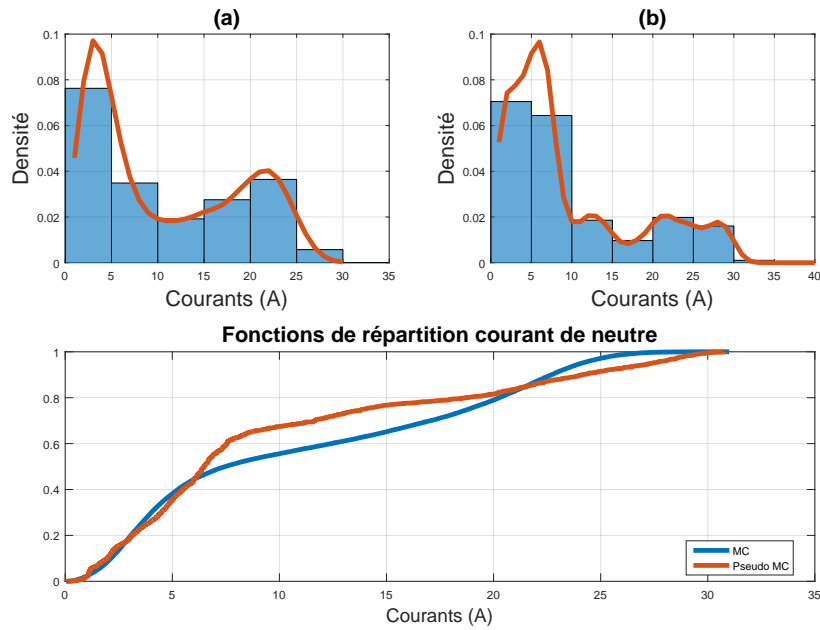


FIGURE 4.21 – Densités de probabilité et fonctions de répartition du courant de neutre  
(a) : Estimation par méthode MC.  
(b) : Estimation par méthode pseudo MC

est de fournir au gestionnaire de réseau des outils permettant de quantifier le taux de pénétration maximal de PV qui risquerait de causer des défaillances sur les critères de fonctionnement du réseau, comme par exemple la tension ou le courant de neutre. Dans cette partie, sera présenté un cas d'étude permettant de quantifier l'impact de la production PV sur le courant dans le neutre d'un réseau de distribution.

L'objectif ici est d'appliquer une méthode LFP dans le but d'estimer la probabilité de dépassement du courant de neutre en fonction du déséquilibre de production photovoltaïque installée entre les phases d'un noeud en moyenne tension.

Sur un noeud donné de notre zone d'étude, on définit par  $P_T$  la puissance totale théorique installée sur toutes les trois phases. Cette puissance est calculée en fonction des surfaces et des inclinaisons des panneaux sur chaque phase. La proportion de puissance installée sur chacune des phases a, b, et c est définie par  $k_i$ , comme l'indique l'équation 4.5.

$$k_i = \frac{P_i}{P_T}, i = 1, 2, 3 \quad (4.5)$$

Pour une puissance  $P_T$  fixe, en variant les  $k_i$ , on agit sur les différences de puissances installées sur chaque phase, ce qui permet de réduire ou d'augmenter le niveau de déséquilibre. L'indice de déséquilibre  $T_{un}$  est défini comme étant la différence maximale de production installée entre deux des trois phases 4.6.

$$T_{un} = \max(|P_1 - P_2|, |P_2 - P_3|, |P_3 - P_1|) \quad (4.6)$$

Pour un gain de temps de simulation, la méthode pseudo Monte Carlo est appliquée pour différents indices de déséquilibre  $T_{un}$ . L'évolution, sur une journée, des probabilités de dépassement du courant de neutre en fonction du taux de déséquilibre  $T_{un}$  est représentée par la figure 4.22, où le niveau de déséquilibre est donné en pourcentage de la puissance totale de PV installée sur les trois phases. Plus ce pourcentage est grand, plus la production totale est concentrée sur une seule phase. La limite du courant de neutre est fixée égale au courant nominal du conducteur de neutre qui est de l'ordre de 300 A. Pour arriver à ce niveau, la puissance des charges connectées est augmentée de sorte à faire fonctionner le réseau aux environs de sa charge nominale.

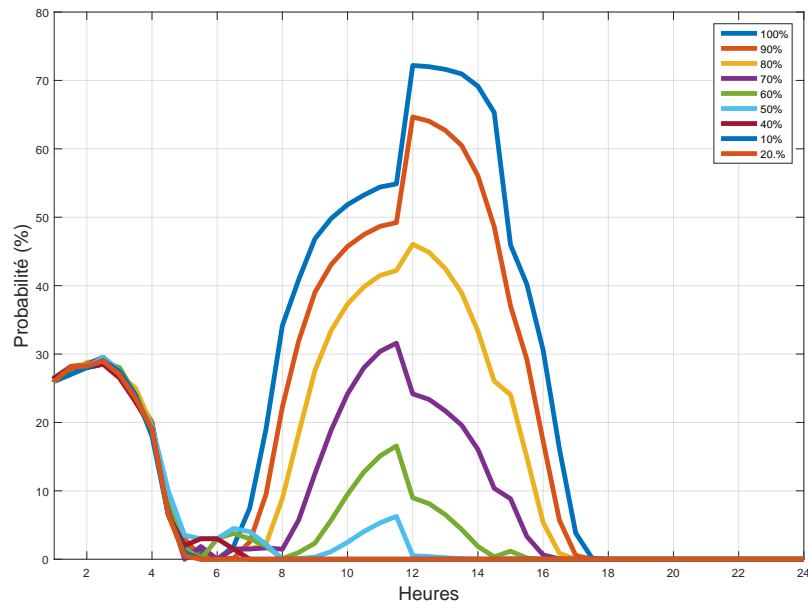


FIGURE 4.22 – Probabilités de dépassement limites du courant de neutre en fonction de différents niveaux de déséquilibre estimées par méthode Pseudo MC



Les résultats montrent une forte corrélation entre le déséquilibre de production et le courant de neutre. On note aussi une probabilité de dépassement importante aux heures de recharge des VEs. Pour une puissance totale maximale de 154 KW installée sur l'ensemble des trois phases du nœud sélectionné, la probabilité de dépassement de la limite du courant de neutre reste quasiment nulle pour un niveau de déséquilibre de 20% de la puissance totale installée. Cette probabilité atteint les 30%, au heures de production PV maximale, lorsque le déséquilibre est aux alentours de 70% de la puissance maximale installée. Le risque de dépassement du courant de neutre commence à atteindre un niveau critique (60%) lorsque le déséquilibre atteint environ 85% de la puissance totale installée.

# Chapitre 5

## Conclusion Générale

Le but de cette thèse est d'étudier des modèles probabilistes pour estimer les impacts de la production photovoltaïque et des véhicules électriques sur les réseaux de distribution moyenne et basse tension.

Les mutations observées sur le système électrique (production décentralisée, véhicules électriques, stockage, micro réseau...) font émerger des problématiques d'ordre économiques et techniques dans la gestion de ce dernier. Parmi ces problématiques, l'impact de la production décentralisée d'origine photovoltaïque et des VEs sur les contraintes techniques du réseau de distribution notamment la tension et le courant de neutre. L'incertitude sur la puissance PV produite et sur l'utilisation des VEs implique la nécessité de développer des modèles probabilistes de consommation et de production d'électricité. L'apport de cette thèse dans ce cadre peut être résumé par :

1. Un modèle probabiliste de la production PV basé sur le modèle de l'indice de nébulosité représentatif du phénomène des passages de nuages, qui est la principale cause de l'intermittence de ce type de production. La nébulosité est considérée comme un paramètre aléatoire continu, de valeur comprise entre 0 (ciel complètement dégagé) et 1 (ciel complètement couvert) et définie par une loi de probabilité bêta. Ce modèle est simple d'utilisation mais les échantillons aléatoires tirés à partir de la loi bêta reflètent soit un ciel complètement couvert soit un ciel complètement dégagé, donc ne permettent pas de représenter les cas de situation intermédiaire ou de transition.

2. Un deuxième modèle probabiliste de nébulosité basé cette fois ci sur un paramètre aléatoire discret qui définit le profil type de nébulosité d'un jour donné. Les données de nébulosité sont réparties en saisons. Pour chaque saison, le clustering permet de regrouper les jours aux profils de nébulosité similaires. La moyenne de ces profils est considérée comme le profil type T du groupe de jours. Ainsi la probabilité d'occurrence d'un profil T est définie par la taille de son cluster par rapport au nombre de jour total de la saison considérée. Ce modèle offre plus de possibilités de scénarios de profils de nébulosité tout en gardant une corrélation entre les points des instants horaires successifs.
3. Un modèle de recharge de véhicule électrique basé sur la distance parcourue et l'instant de début de recharge, considérés comme paramètres aléatoires de lois respectivement lognormal et uniforme. La distance est convertie en énergie requise par le biais de la constante  $0.2kWh/km$ . Sous l'hypothèse d'une recharge à domicile uniquement, l'instant de début de recharge est uniformément distribué entre 18h et 06h du matin et la puissance de recharge est constante et supposée la même sur toute les bornes.
4. Un modèle de consommation des autres équipements de la maison qui repose sur le même principe que le deuxième modèle de nébulosité décrit plus haut. L'incertitude est réduite à quelques profils type de clusters pondérés par leur probabilité d'occurrence représentative de la taille d'un cluster donné par rapport à la taille totale de tous les clusters.

Les techniques de calculs de load flow déterministes, utilisées généralement, dans l'étude des réseaux ne permettent pas de prendre en compte les incertitudes sur la production et la consommation raccordées sur le réseau. De ce fait des techniques de load flow probabilistes ont été étudiées dans cette thèse.

1. La méthode Monte Carlo (MC) qui consiste à calculer plusieurs milliers de fois le load flow déterministe en considérant des scénarios de production et de consommation aléatoires. Elle est considérée comme référence en termes de précision mais nécessite beaucoup de temps de simulation.
2. La méthode PEM qui est basée sur une approximation de la densité de probabilité

des paramètres aléatoires d'entrée en quelques points représentatifs des moments d'ordre un, deux, trois et parfois quatre. Chaque point est pondéré par un poids. Le load flow déterministe est calculé en remplaçant chaque paramètre d'entrée par ses points d'estimations. Elle est présentée comme une méthode qui nécessite moins de temps de simulation que la méthode MC pour un niveau de précision acceptable.

En plus de ces méthodes existantes, nous avons développé une troisième méthode, appelée méthode pseudo MC, qui repose aussi sur une répétition du calcul de load flow déterministe. Cette fois ci, l'incertitude sur les paramètres aléatoires d'entrée est réduite à quelques clusters définis par leur centre et une probabilité d'occurrence. Autrement dit, on suppose que chaque paramètre ne peut prendre que les valeurs des centres de ses clusters avec une probabilité respective. Ainsi l'on dispose d'un nombre fini de valeurs possibles (nombre de clusters) pour un paramètre donné. Le load flow déterministe est calculé en considérant toutes les combinaisons possibles des valeurs des différents paramètres. Ceci réduit considérablement le nombre de scénario de calcul par rapport à une méthode Monte Carlo avec tirage de valeurs aléatoires à partir d'une loi.

Les méthodes PEM et pseudo MC sont des méthodes d'approximation (ou de réduction) pour s'affranchir du temps de calcul élevé de la méthode MC. L'un des apports de cette thèse est de comparer la pertinence de ces méthodes, en termes de précision par rapport à la méthode MC, selon différents types de lois de probabilité pour les paramètres d'entrée. On constate, comme le rappelle le tableau 5.1, que les performances de la méthode PEM sont limitées par rapport à celle de la méthode pseudo MC dans le cas de distribution de type convexe ou discrète.

TABLE 5.1 – Récapitulatif approximation PEM et clustering en fonction de type de distribution

Distribution	Normale	Bêta	Discrète	Lognormale
PEM	=	-	-	≈
Clustering	=	+	+	≈

La méthode pseudo MC proposée dans cette thèse, appliquée au réseau test de 4

nœuds, aboutit à des résultats (en termes de moyennes et densité de probabilité) de niveaux de précision acceptables par rapport à la méthode Monte Carlo pour un temps de simulation de 5 secondes contre 300 secondes pour MC. Appliquée à un réseau de taille plus réaliste, la méthode pseudo MC aboutit à des probabilités de dépassement des valeurs limites plus importantes que celles estimées par la méthode MC. La modélisation par clustering implique des probabilités d'occurrence des situations extrêmes plus élevées que dans le cas d'une modélisation par densité de probabilité

Une application de load flow probabiliste étudiée dans cette thèse porte sur l'estimation de la probabilité de dépassement des limites du courant dans le fil de neutre en fonction du déséquilibre de puissance installée entre les phases d'un bus. Ceci dans le but d'estimer le niveau de déséquilibre maximal admissible pour maintenir le bon niveau de courant dans le neutre. Les résultats montrent une forte corrélation entre le déséquilibre de PV et le courant dans le neutre et que le risque de dépassement commence à être considérable (60%) lorsque le déséquilibre atteint environ 85% de la somme des puissances installées sur les phases.

Le réseau *ckt5* utilisé pour tester les méthodes a une topologie d'un réseau de distribution urbain étalé. Une des perspectives de ces travaux pourrait être l'application des outils présentés sur des données d'un réseau urbain de topologie plus dense et un réseau rural, afin de confronter les résultats obtenus.

En outre, les probabilités de violation de risque estimées grâce à ces méthodes peuvent servir comme données d'entrée à un problème d'optimisation de consommation et de production. Ceci, dans le but de définir des stratégies de planification d'utilisation des charges et de contrôle de la production injectée pour limiter les contraintes techniques sur le réseau.

# Bibliographie

- [1] Jean-Claude Sabonnadière and Nouredine Hadjsaid. *Lignes et réseaux électriques 1 : Lignes d'énergie électrique*. Hermès-Lavoisier, 2007.
- [2] Connaissance des énergies. Coûts de production de l'électricité en france. <https://www.connaissancedesenergies.org/fiche-pedagogique/couts-de-production-de-l-electricite-en-france>, 2013.
- [3] ESR. Description physique du réseau public de distribution (rpd) d'électricité, 2015.
- [4] William H Kersting. *Distribution system modeling and analysis*. CRC press, 2012.
- [5] Ministère de l'énergie. Arrêté du 24 décembre 2007 pris en application du décret num 2007-1826 du 24 décembre 2007 relatif aux niveaux de qualité et aux prescriptions techniques en matière de qualité des réseaux publics de distribution et de transport d'électricité. <https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000017754536>, 2007.
- [6] Jean-Claude Sabonnadière and Nouredine Hadjsaid. Des réseaux électriques aux « smartgrids ». *Encyclopédie de l'énergie*, 2015.
- [7] SMART GRIDS CRE. L'intégration des enr. <http://www.smartgrids-cre.fr/index.php?p=integrationenr-hei-l2ep>, 2011.
- [8] Fiche auto. Evolution du nombre de voitures Électriques vendues en france. <http://www.fiches-auto.fr/articles-auto/l-auto-en-chiffres/s-1941-evolution-du-nombre-de-voitures-electriques-vendues-en-france.php>, 2017.
- [9] avere France. Les puissance de recharge des véhicules électriques. [http://www.avere-france.org/Site/Article/?article\\_id=5883](http://www.avere-france.org/Site/Article/?article_id=5883), 2017.
- [10] Lesvoitresselectrique.com. Comparatif autonomie des voitures Électriques. <https://www.les-voitures-electriques.com/guide/autonomie-voitures-electriques/>, 2017.
- [11] Miroslav M Begovic, Insu Kim, Damir Novosel, Julio Romero Agüero, and Ajeet Rohatgi. Integration of photovoltaic distributed generation in the power distribution grid. In *System Science (HICSS), 2012 45th Hawaii International Conference on*, pages 1977–1986. IEEE, 2012.

- [12] Jaroslaw Krata, Tapan Kumar Saha, and Ruifeng Yan. Large scale photovoltaic system and its impact on distribution network in transient cloud conditions. In *Power & Energy Society General Meeting, 2015 IEEE*, pages 1–5. IEEE, 2015.
- [13] Johan Desmet, Isabel Sweertvaegher, Greet Vanalme, Kurt Stockman, and Ronnie Belmans. Analysis of the neutral conductor current in a three phase supplied network with nonlinear single phase loads. In *Electric Machines and Drives Conference, 2001. IEMDC 2001. IEEE International*, pages 448–453. IEEE, 2001.
- [14] Dobrivoje P Stojanovic, Lidija M Korunovic, and Aleksandar Jovic. Measurement and analysis of neutral conductor current in low voltage distribution network. In *EUROCON 2009, EUROCON'09. IEEE*, pages 1481–1486. IEEE, 2009.
- [15] Makarand Sudhakar Ballal, Hiralal Murlidhar Suryawanshi, Mahesh K Mishra, and Bhupesh Nemichand Chaudhari. Interturn faults detection of transformers by diagnosis of neutral current. *IEEE Transactions on Power Delivery*, 31(3) :1096–1105, 2016.
- [16] Thomas M Gruz. A survey of neutral currents in three-phase computer power systems. *IEEE Transactions on Industry Applications*, 26(4) :719–725, 1990.
- [17] Md Jan E Alam, Kashem M Muttaqi, and Danny Sutanto. Alleviation of neutral-to-ground potential rise under unbalanced allocation of rooftop pv using distributed energy storage. *IEEE Transactions on Sustainable Energy*, 6(3) :889–898, 2015.
- [18] P Pillay and M Manyage. Definitions of voltage unbalance. *IEEE Power Engineering Review*, 21(5) :50–51, 2001.
- [19] Ruifeng Yan, Stephanie Roediger, and Tapan Kumar Saha. Impact of photovoltaic power fluctuations by moving clouds on network voltage : A case study of an urban network. In *Power Engineering Conference (AUPEC), 2011 21st Australasian Universities*, pages 1–6. IEEE, 2011.
- [20] Yujun He. Optimisation du plan de tension d’un réseau de distribution à l’aide de la production distribuée, 2015.
- [21] B Wojszczyk and M Brandao. High penetration of distributed generation and its impact on electric grid performance-utility perspective. In *2011 IEEE PES Innovative Smart Grid Technologies*. 2011.
- [22] MACPHERSON James VASILIC Slavko STEWART, Emma. Analysis of high-penetration levels of photovoltaics into the distribution grid on oahu, hawaii. 2013.
- [23] WOYAK Jeremy COSTYK David SEGUIN, Rich. High-penetration pv integration handbook for distribution engineers, 2016.
- [24] JW Smith, R Dugan, M Rylander, and T Key. Advanced distribution planning tools for high penetration pv deployment. In *Power and Energy Society General Meeting, 2012 IEEE*, pages 1–7. IEEE, 2012.

- [25] Ronald E Walpole, Raymond H Myers, Sharon L Myers, and E Ye Keying. *Probability & Statistics for Engineers & Scientists, MyStatLab*. Pearson Higher Ed, 2016.
- [26] Saeed Ghahramani. *Fundamentals of Probability : With Stochastic Processes*. CRC Press, 2015.
- [27] S Conti and S Raiti. Probabilistic load flow for distribution networks with photovoltaic generators part 1 : Theoretical concepts and models. In *Clean Electrical Power, 2007. ICCEP'07. International Conference on*, pages 132–136. IEEE, 2007.
- [28] B Rajanarayan Prusty and Debashisha Jena. Modeling of correlated photovoltaic generations and load demands in probabilistic load flow. In *India Conference (INDICON), 2015 Annual IEEE*, pages 1–6. IEEE, 2015.
- [29] SH Karaki, RB Chedid, and R Ramadan. Probabilistic performance assessment of autonomous solar-wind energy conversion systems. *IEEE Transactions on Energy Conversion*, 14(3) :766–772, 1999.
- [30] Miao Fan, Vijay Vittal, Gerald Thomas Heydt, and Raja Ayyanar. Probabilistic power flow studies for transmission systems with photovoltaic generation using cumulants. *IEEE Transactions on Power Systems*, 27(4) :2251–2261, 2012.
- [31] Ward T Jewell and Timothy D Unruh. Limits on cloud-induced fluctuation in photovoltaic generation. *IEEE Transactions on energy conversion*, 5(1) :8–14, 1990.
- [32] Chengrui Cai and Dionysios C Aliprantis. Cumulus cloud shadow model for analysis of power systems with photovoltaics. *IEEE Transactions on Power Systems*, 28(4) : 4496–4506, 2013.
- [33] Mehdi Ebad and William Mack Grady. A cloud shadow model for analysis of solar photovoltaic power variability in high-penetration pv distribution networks. In *Power and Energy Society General Meeting (PESGM), 2016*, pages 1–5. IEEE, 2016.
- [34] Munira Batool, Syed M Islam, and Farhad Shahnia. Stochastic modeling of the output power of photovoltaic generators in various weather conditions. In *Power Engineering Conference (AUPEC), 2016 Australasian Universities*, pages 1–5. IEEE, 2016.
- [35] M Shukla and G Radman. Optimal power flow using probabilistic load model. In *System Theory, 2005. SSST'05. Proceedings of the Thirty-Seventh Southeastern Symposium on*, pages 439–442. IEEE, 2005.
- [36] Peiyuan Chen, Birgitte Bak-Jensen, and Zhe Chen. Probabilistic load models for simulating the impact of load management. In *Power & Energy Society General Meeting, 2009. PES'09. IEEE*, pages 1–8. IEEE, 2009.
- [37] Tianshu Zhang, Xiaohui Song, Xiaoli Meng, Jie Yu, and Xiaoyi Chen. A probabilistic load model based on chi-square method for distribution network. 2013.



- [38] Mohammed E Nassar and MM A Salama. A novel probabilistic load model and probabilistic power flow. In *Electrical and Computer Engineering (CCECE), 2015 IEEE 28th Canadian Conference on*, pages 881–886. IEEE, 2015.
- [39] Peter Olofsson and Mikael Andersson. *Probability, statistics, and stochastic processes*. John Wiley & Sons, 2012.
- [40] Nicolas MEUNIER. La vérité sur la consommation des véhicules électriques @ONLINE, Aug. 2012. URL [https://www.challenges.fr/automobile/dossiers/la-verite-sur-la-consommation-des-voitures-electriques\\_2623](https://www.challenges.fr/automobile/dossiers/la-verite-sur-la-consommation-des-voitures-electriques_2623).
- [41] Feteh Nassim Melzi, Mohamed Haykel Zayani, Amira Ben Hamida, Allou Same, and Latifa Oukhellou. Identifying daily electric consumption patterns from smart meter data by means of clustering algorithms. In *Machine Learning and Applications (ICMLA), 2015 IEEE 14th International Conference on*, pages 1136–1141. IEEE, 2015.
- [42] S Sunisith and K Meena. Backward/forward sweep based distribution load flow method. *International Electrical Engineering Journal (IEEJ)*, 5(9), 2014.
- [43] Eduard Janecek and Daniel Georgiev. Probabilistic extension of the backward/forward load flow analysis method. *IEEE Transactions on Power Systems*, 27(2) :695–704, 2012.
- [44] Dexian Li. Stage de fin d’étude :intégration des contraintes électriques dans un modèle de déploiement de panneaux photovoltaïques dans un réseau électrique de distribution, 2016.
- [45] Mohammad Javad Mirzaei, Ahad Kazemi, and Omid Homaei. A probabilistic approach to determine optimal capacity and location of electric vehicles parking lots in distribution networks. *IEEE Transactions on Industrial Informatics*, 12(5) :1963–1972, 2016.
- [46] AM Leite Da Silva and VL Arienti. Probabilistic load flow by a multilinear simulation algorithm. In *IEE Proceedings C (Generation, Transmission and Distribution)*, volume 137, pages 276–282. IET, 1990.
- [47] AM Leite Da Silva, JW Marangon Lima, SMP Ribeiro, VL Arienti, M Th Schilling, X Vieira Filho, and SM Soares. Operational and expansion planning of brazilian systems based on probabilistic load flow. In *Probabilistic Methods Applied to Electric Power Systems, 1991., Third International Conference on*, pages 197–202. IET, 1991.
- [48] Walid El-Khattam, YG Hegazy, and MMA Salama. Investigating distributed generation systems performance using monte carlo simulation. *IEEE Transactions on Power Systems*, 21(2) :524–532, 2006.
- [49] S Conti, S Raiti, and C Di Gregorio. Probabilistic load flow for distribution networks with photovoltaic generators part 2 : Application to a case study. In *Clean Electrical Power, 2007. ICCEP’07. International Conference on*, pages 137–141. IEEE, 2007.

- [50] E Ciapessoni, D Cirio, and A Pitto. Effect of renewable and load uncertainties on the assessment of power system operational risk. In *Probabilistic Methods Applied to Power Systems (PMAPS), 2014 International Conference on*, pages 1–6. IEEE, 2014.
- [51] Hao-Tian Zhang, Kang Chang, Huiling Zhang, and Loi Lei Lai. A novel probabilistic approach for evaluating fault ride-through capability of wind generation. In *Machine Learning and Cybernetics (ICMLC), 2016 International Conference on*, pages 135–140. IEEE, 2016.
- [52] Matej Rejc and Miloš Pantoš. Probabilistic short-term transmission-loss forecasting based on a two-point estimate method. In *Energy Market (EEM), 2010 7th International Conference on the European*, pages 1–6. IEEE, 2010.
- [53] Qiaoyan Bian, Qian Xu, Liying Sun, Leiqi Zhang, Hao Wu, and Huanhai Xin. Grid-connected wind power capacity optimization based on the principle of maximum entropy. In *PES General Meeting/ Conference & Exposition, 2014 IEEE*, pages 1–5. IEEE, 2014.
- [54] Juan M Morales, Luis Baringo, Antonio J Conejo, and Roberto Mínguez. Probabilistic power flow with correlated wind sources. *IET generation, transmission & distribution*, 4(5) :641–651, 2010.
- [55] P Jorgensen, JS Christensen, and JO Tande. Probabilistic load flow calculation using monte carlo techniques for distribution network with wind turbines. In *Harmonics and Quality of Power Proceedings, 1998. Proceedings. 8th International Conference On*, volume 2, pages 1146–1151. IEEE, 1998.
- [56] Damien Fleming Frame, GW Ault, and Sikai Huang. The uncertainties of probabilistic lv network analysis. In *Power and Energy Society General Meeting, 2012 IEEE*, pages 1–8. IEEE, 2012.
- [57] Peiyuan Chen, Zhe Chen, and Birgitte Bak-Jensen. Probabilistic load flow : A review. In *Electric Utility Deregulation and Restructuring and Power Technologies, 2008. DRPT 2008. Third International Conference on*, pages 1586–1591. IEEE, 2008.
- [58] Gan Li. *Stochastic analysis and optimization of power system steady-state with wind farms and electric vehicles*. PhD thesis, University of Birmingham, 2012.
- [59] Bie Zhaohong and Wang Xifan. Studies on variance reduction technique of monte carlo simulation in composite system reliability evaluation. *Electric Power Systems Research*, 63(1) :59–64, 2002.
- [60] Panida Jirutitijaroen and Chanan Singh. Comparison of simulation methods for power system reliability indexes and their distributions. *IEEE Transactions on Power Systems*, 23(2) :486–493, 2008.
- [61] RN Allan and MRG Al-Shakarchi. Probabilistic ac load flow. In *Proceedings of the Institution of Electrical Engineers*, volume 123, pages 531–536. IET, 1976.

- [62] Barbara Borkowska. Probabilistic load flow. *IEEE Transactions on Power Apparatus and Systems*, (3) :752–759, 1974.
- [63] RN Allan, B Borkowska, and CH Grigg. Probabilistic analysis of power flows. In *Proceedings of the Institution of Electrical Engineers*, volume 121, pages 1551–1556. IET, 1974.
- [64] RN Allan and MRG Al-Shakarchi. Probabilistic techniques in ac load-flow analysis. In *Proceedings of the Institution of Electrical Engineers*, volume 124, pages 154–160. IET, 1977.
- [65] RN Allan and AM Leite Da Silva. Probabilistic load flow using multilinearisations. In *IEE Proceedings C (Generation, Transmission and Distribution)*, volume 128, pages 280–287. IET, 1981.
- [66] RN Allan, AM Leite Da Silva, and RC Burchett. Evaluation methods and accuracy in probabilistic load flow solutions. *IEEE Transactions on Power Apparatus and Systems*, (5) :2539–2546, 1981.
- [67] Pei Zhang and Stephen T Lee. Probabilistic load flow computation using the method of combined cumulants and gram-charlier expansion. *IEEE transactions on power systems*, 19(1) :676–682, 2004.
- [68] Gan Li and Xiao-Ping Zhang. Comparison between two probabilistic load flow methods for reliability assessment. In *Power & Energy Society General Meeting, 2009. PES'09. IEEE*, pages 1–7. IEEE, 2009.
- [69] Zechun Hu and Xifan Wang. A probabilistic load flow method considering branch outages. *IEEE Transactions on Power Systems*, 21(2) :507–514, 2006.
- [70] Alexander C Melhorn and Aleksandar Dimitrovski. Three-phase probabilistic load flow in radial and meshed distribution networks. *IET Generation, Transmission & Distribution*, 9(16) :2743–2750, 2015.
- [71] Trevor Williams and Curran Crawford. Probabilistic load flow modeling comparing maximum entropy and gram-charlier probability density function reconstructions. *IEEE Transactions on Power Systems*, 28(1) :272–280, 2013.
- [72] M Brucoli, F Torelli, and R Napoli. Quadratic probabilistic load flow with linearly modelled dispatch. *International Journal of Electrical Power & Energy Systems*, 7 (3) :138–146, 1985.
- [73] Li Xiaoming, Chen Xiaohui, Yin Xianggen, Xiang Tiejuan, and Liu Huagang. The algorithm of probabilistic load flow retaining nonlinearity. In *Power System Technology, 2002. Proceedings. PowerCon 2002. International Conference on*, volume 4, pages 2111–2115. IEEE, 2002.
- [74] KS Li. Point-estimate method for calculating statistical moments. *Journal of Engineering Mechanics*, 118(7) :1506–1511, 1992.

- [75] Stanley Kaplan. On the method of discrete probability distributions in risk and reliability calculations—application to seismic risk assessment. *Risk Analysis*, 1(3) : 189–196, 1981.
- [76] Qin Zhang. A general method dealing with correlations in uncertainty propagation in fault trees. *Reliability Engineering & System Safety*, 26(3) :231–247, 1989.
- [77] Robert E Kurth and David C Cox. Discrete probability distributions for probabilistic fracture mechanics. *Risk Analysis*, 5(3) :235–240, 1985.
- [78] Marcelino Madrigal, K Ponnambalam, and VH Quintana. Probabilistic optimal power flow. In *Electrical and Computer Engineering, 1998. IEEE Canadian Conference on*, volume 1, pages 385–388. IEEE, 1998.
- [79] Chun-Lien Su. Probabilistic load-flow computation using point estimate method. *IEEE Transactions on Power Systems*, 20(4) :1843–1851, 2005.
- [80] Yuanbing Xiao, Haozhong Cheng, Ting Lei, Shenxi Zhang, Weijie Luan, and Di Yang. Research on probabilistic load flow method based on point estimation and dimension reduction by dominated variables. In *Electric Utility Deregulation and Restructuring and Power Technologies (DRPT), 2015 5th International Conference on*, pages 599–603. IEEE, 2015.
- [81] Iman Kaffashan and Turaj Amraee. Probabilistic undervoltage load shedding using point estimate method. *IET Generation, Transmission & Distribution*, 9(15) :2234–2244, 2015.
- [82] Angela Russo, P Varilone, and P Caramia. Point estimate schemes for probabilistic harmonic power flow. In *Harmonics and Quality of Power (ICHQP), 2014 IEEE 16th International Conference on*, pages 19–23. IEEE, 2014.
- [83] Juan M Morales and Juan Perez-Ruiz. Point estimate schemes to solve the probabilistic power flow. *IEEE Transactions on Power Systems*, 22(4) :1594–1601, 2007.
- [84] Ying-Yi Hong, Faa-Jeng Lin, and Ti-Hsuan Yu. Taguchi method-based probabilistic load flow studies considering uncertain renewables and loads. *IET Renewable Power Generation*, 10(2) :221–227, 2016.
- [85] Xuexia Zhang, Zhiqi Guo, and Weirong Chen. Probabilistic power flow method considering continuous and discrete variables. *Energies*, 10(5) :590, 2017.
- [86] Defu Cai, Xiaoping Li, Kunpeng Zhou, Junhui Xin, and Kan Cao. Probabilistic load flow algorithms considering correlation between input random variables : A review. In *Industrial Electronics and Applications (ICIEA), 2015 IEEE 10th Conference on*, pages 1139–1144. IEEE, 2015.
- [87] Celine Baranger Julien Mathiaud. Methode de monte-carlo, 2012/2013.
- [88] Reuven Y Rubinstein and Dirk P Kroese. *Simulation and the Monte Carlo method*, volume 10. John Wiley & Sons, 2016.

- [89] Attoti Bharath Krishna, Nikhil Gupta, KR Niazi, and Anil Swarnkar. Probabilistic power flow in radial distribution systems using point estimate methods. In *Advanced Computing and Communication Systems (ICACCS), 2017 4th International Conference on*, pages 1–6. IEEE, 2017.
- [90] Emilio Rosenblueth. Two-point estimates in probabilities. *Applied Mathematical Modelling*, 5(5) :329–335, 1981.
- [91] KS Li. Point-estimate method for calculating statistical moments. *Journal of Engineering Mechanics*, 118(7) :1506–1511, 1992.
- [92] Milton E Harr. Probabilistic estimates for multivariate analyses. *Applied Mathematical Modelling*, 13(5) :313–318, 1989.
- [93] HP Hong. An efficient point estimate method for probabilistic analysis. *Reliability Engineering & System Safety*, 59(3) :261–267, 1998.
- [94] Ronald E Walpole, Raymond H Myers, Sharon L Myers, and E Ye Keying. *Probability & Statistics for Engineers & Scientists, MyStatLab*. Pearson Higher Ed, 2016.
- [95] Priority Wire and INC Cable. Acsr – aluminum conductor steel reinforced. <http://www.prioritywire.com/specs/ACSR.pdf>, 2013.

# Chapitre 6

## Annexes

### 6.1 Annexe 1 : Matrices généralisées des quadripoles du réseau test IEEE 4 noeuds

TABLE 6.1 – Matrices généralisées  $L_{12}$

$[a]$	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$
$[b]$	$\begin{pmatrix} 0.14 + 0.53i & 0.03 + 0.32i & 0.03 + 0.27i \\ 0.03 + 0.32i & 0.14 + 0.53i & 0.03 + 0.29i \\ 0.03 + 0.27i & 0.33 + 0.i & 0.14 + 0.53i \end{pmatrix}$
$[c]$	0
$[d]$	$[a]$
$[A]$	$[a]^{-1}$
$[B]$	$[a]^{-1} * [b]$

TABLE 6.2 – Matrices généralisées  $L_{34}$

$[a]$	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$
$[b]$	$\begin{pmatrix} 0.19 + 0.50i & 0.06 + 0.23i & 0.05 + 0.17i \\ 0.06 + 0.23i & 0.19 + 0.48i & 0.06 + 0.19i \\ 0.05 + 0.17i & 0.06 + 0.19i & 0.19 + 0.49i \end{pmatrix}$
$[c]$	0
$[d]$	$[a]$
$[A]$	$[a]^{-1}$
$[B]$	$[a]^{-1} * [b]$

TABLE 6.3 – Matrices généralisées de l'ensemble transformateur

[a]	$\begin{pmatrix} 0 & -3.4 & -1.7 \\ -1.7 & 0 & -3.4 \\ -3.4 & -1.7 & 0 \end{pmatrix}$
[b]	$\begin{pmatrix} 0 & -0.09 - 0.59i & -0.04 - 0.29i \\ -0.04 - 0.29i & 0 & -0.09 - 0.59i \\ -0.09 - 0.59i & -0.04 - 0.29i & 0 \end{pmatrix}$
[c]	$0$
[d]	$\begin{pmatrix} 0.19 & -0.19 & 0 \\ 0 & 0.19 & -0.19 \\ -0.19 & 0 & 0.19 \end{pmatrix}$
[A]	$\begin{pmatrix} 0.19 & 0 & -0.19 \\ -0.19 & 0.19 & 0 \\ 0 & -0.19 & 0.19 \end{pmatrix}$
[B]	$\begin{pmatrix} 0.02 + 0.17i & 0 & 0 \\ 0 & 0.02 + 0.17i & 0 \\ 0 & 0 & 0.02 + 0.17i \end{pmatrix}$

## 6.2 Annexe 2 : Principe du FBS appliqué au réseau IEEE 4 noeuds

Dans un premier temps les matrices généralisées (présentées au chapitre 2) des composants du système sont déterminées. L'exemple de réseau considéré ci dessus est composé de trois quadripoles séries à savoir le segment de ligne  $L_{12}$ , le transformateur et le segment de ligne  $L_{34}$ .

On a vu au chapitre 2 que les matrices généralisées [a], [b], [c], [d], [A] et [B] d'un composant série triphasé étaient liées aux vecteurs de tension et courant en entrée et à ceux de sortie du composant par la relation :

$$[V_{abc}]_s = [a] \cdot [V_{abc}]_{ch} + [b] \cdot [I_{abc}]_{ch} \quad (6.1)$$

$$[I_{abc}]_s = [c] \cdot [V_{abc}]_{ch} + [d] \cdot [I_{abc}]_{ch} \quad (6.2)$$

Par  $[V_{abc}]_s$  et  $[I_{abc}]_s$ , sont désignés respectivement la tension et le courant complexes des phases a, b, et c de l'extrémité du composant située du coté de la source. Et par  $[V_{abc}]_{ch}$  et  $[I_{abc}]_{ch}$ , tension et courant complexes des phases a, b, et c de l'extrémité du composant situé du coté de la charge.

L'équation exprimant les tensions de l'extrémité coté charge en fonction de celle coté source est comme suit :

$$[V_{abc}]_{ch} = [A] \cdot [V_{abc}]_s - [B] \cdot [I_{abc}]_{ch} \quad (6.3)$$

Les expressions des matrices généralisée vues au chapitre 2, appliquées à notre modèle de réseau ci dessus (avec les admittances de shunt des lignes  $L_{12}$  et  $L_{34}$  négligées), donneraient les résultats présentés dans annexe 2.

Initialement, la valeur de la tension au nœud 4 est choisie égale à la valeur nominale de  $2.4kV$  avec un déphasage de 30 degrés entre les phases. En appliquant la méthode FBS présentée au chapitre 2, lors de la première phase de Forward Sweep, on peut calculer : Le courant à la borne de la charge qui vaut alors :

$$I4 = \begin{bmatrix} 312.5\angle-61.7^\circ \\ 416.66\angle-175.8^\circ \\ 520.83\angle71.8^\circ \end{bmatrix} (A) \quad (6.4)$$

La tension en  $kV$  et le courant en  $A$  calculés à la sortie du transformateur et donc à l'entrée de la ligne  $L_{34}$  sont :

$$V3 = \begin{bmatrix} 2.47\angle-29.52^\circ \\ 2.53\angle-148.3^\circ \\ 2.50\angle94.05^\circ \end{bmatrix} (kV) \quad (6.5)$$

$$I3 = \begin{bmatrix} 312.5\angle-61.7^\circ \\ 416.66\angle-175.8^\circ \\ 520.83\angle71.8^\circ \end{bmatrix} (A) \quad (6.6)$$

De là, la tension et courant à l'entrée du transformateur (extrémité ligne  $LS$  coté source) sont calculés :

$$V2 = \begin{bmatrix} 7.95\angle3.2^\circ \\ 7.34\angle-113.4^\circ \\ 7.64\angle120.4^\circ \end{bmatrix} (kV) \quad (6.7)$$

$$I2 = \begin{bmatrix} 118.23\angle-23.5^\circ \\ 150.30\angle-137.7^\circ \\ 148.25\angle88.8^\circ \end{bmatrix} (A) \quad (6.8)$$

Et enfin la tension et le courant du nœud connecté à la source :

$$V1 = \begin{bmatrix} 7.98\angle3.3^\circ \\ 7.37\angle-113.2^\circ \\ 7.7\angle121.06^\circ \end{bmatrix} (kV) \quad (6.9)$$

$$I1 = \begin{bmatrix} 118.23\angle-23.5^\circ \\ 150.30\angle-137.7^\circ \\ 148.25\angle88.8^\circ \end{bmatrix} (A) \quad (6.10)$$

Cette étape marque la fin de la première phase Forward. La tension en phase au nœud 1 notée  $U1$  est comparée à celle de la source. Les composantes 1, 2 et 3 de la tension  $U1$  sont exprimées en fonction des tension simples des phases a, b, et c autrement dit les composantes 1, 2 et 3 du vecteur  $[V1]$  calculé ci dessous :

$$U1_1 = V1_a - V1_b \quad (6.11)$$

$$U1_2 = V1_b - V1_c \quad (6.12)$$

$$U1_3 = V1_c - V1_a \quad (6.13)$$

$$(6.14)$$



Da là, il est possible de définir une matrice  $[D]$  qui permet le passage de tensions simples aux tensions composées d'un bus donnée :

$$[D] = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ -1 & 0 & 1 \end{bmatrix} \quad (6.15)$$

$$[U1] = [D].[V1] = \begin{bmatrix} 13.06\angle 33.6^\circ \\ 13.4\angle -85.4^\circ \\ 13.43\angle 152.8^\circ \end{bmatrix} \quad (6.16)$$

L'erreur sur les modules des tensions  $[U1]$  et  $[US]$  est de :

$$\epsilon = \frac{|[US] - [U1]|}{12.47} = \begin{bmatrix} 0.0809 \\ 0.1126 \\ 0.0926 \end{bmatrix} pu \quad (6.17)$$

L'erreur  $\epsilon$  étant supérieure au seuil de tolérance  $\delta$  fixé à  $0.0001pu$ , la phase "Backward sweep" est entamée en supposant que la tension phase neutre au noeud 1  $[V1]$  est égale à la tension phase neutre de la source  $VS$ .

L'erreur  $\epsilon$  est devenue inférieure à  $\delta$  après 4 itérations correspondant à :

$$\epsilon = \begin{bmatrix} 0.26.10^{-4} \\ 0.62.10^{-4} \\ 0.36.10^{-4} \end{bmatrix} pu \quad (6.18)$$

Les courants dans les phases et le neutre de la ligne  $L_{34}$  neutre ainsi que la tension au noeud 4 calculée en  $pu$  par rapport à la tension nominale au secondaire du transformateur, c'est à dire  $2400V$  sont données :

$$V_{4f} = \begin{bmatrix} 0.94 \\ 0.91 \\ 0.92 \end{bmatrix} (pu) \quad (6.19)$$

$$I_{4f} = \begin{bmatrix} 331.21\angle -63.7^\circ \\ 454.31\angle -179.3^\circ \\ 565.81\angle 64.6^\circ \end{bmatrix} (A) \quad (6.20)$$

$$I_{nf} = 218.91\angle 107.35^\circ (A) \quad (6.21)$$

Des installations de production photovoltaïques sont raccordées au noeud 4 du réseau. Elles sont représentés par leur puissance active produite (en  $kW$ ). De ce fait, un modèle PQ est plus approprié qui considère cette production comme une charge négative à facteur de puissance égal à 1. En guise d'exemple, on suppose une production sur chacune des trois phases définie en fonction de la surface, de leur rendement, de l'irradiance et d'un indice de nébulosité :

$$S_{4PV} = \begin{bmatrix} 429.82\angle 0^\circ \\ 716.37\angle 0^\circ \\ 429.82\angle 0^\circ \end{bmatrix} (KVA) \quad (6.22)$$

Avec une production photovoltaïque modélisé par un nœud PQ, la charge totale sur les trois phases du nœud 4  $S_{4_{tot}}$  qui sera considérée dans la suite des calculs, est considérée comme une somme de puissance consommée  $S_4$  et d'une puissance négative  $(-1) * S_{4_{PV}}$  représentant la production photovoltaïque.

$$S_{4_{tot}} = S_4 - S_{4_{PV}} \quad (6.23)$$

L'application de la méthode FBS aboutit aux résultats ci dessous après 3 itérations pour un même seuil  $\delta$  de 0.0001 avec une erreur  $\epsilon$  de :

$$\epsilon = \begin{bmatrix} 0.18.10^{-4} \\ 0.25.10^{-4} \\ 0.12.10^{-4} \end{bmatrix} pu \quad (6.24)$$

Les modules des courants dans les phases et le neutre la ligne  $L_{34}$  et la tension en  $pu$  au nœud 4 sont :

$$I_{4_{fd}} = \begin{bmatrix} 200.27 \angle -89^\circ \\ 232.98 \angle 153.9^\circ \\ 274.96 \angle 48.7^\circ \end{bmatrix} (A) \quad (6.25)$$

$$I_{n_{fd}} = 111.28 \angle 102.85^\circ \quad (6.26)$$

$$V_{4_{fd}} = \begin{bmatrix} 0.95 \\ 0.94 \\ 0.94 \end{bmatrix} (pu) \quad (6.27)$$

**Titre :** Analyses probabilistes pour l'étude des réseaux électriques de distribution

**Résumé :** Les mutations observées sur le système électrique (production décentralisée, véhicules électriques, stockage, micro réseau...) font émerger des problématiques d'ordres économiques et techniques dans la gestion de ce dernier. Parmi eux, l'impact sur les niveaux de tension et de courant de neutre des réseaux de distribution. Le but de cette thèse est d'étudier des modèles probabilistes pour estimer ces impacts. L'incertitude sur la puissance PV produite et sur l'utilisation des VEs implique la nécessité de développer des modèles probabilistes de consommation et de production d'électricité. Deux modèles différents de production et de consommation ont été étudiés : L'un basé sur l'approximation de données historiques par une densité de probabilité, l'autre reposant sur la réparti-

tion des données en groupes définis par un profil type et une probabilité d'occurrence. Des techniques de load flow probabilistes ont été étudiées dans cette thèse pour prendre en compte l'effet intermittente de la production PV et l'incertitude sur la consommation. Une technique basée sur la méthode de simulation Monte Carlo, une deuxième basée sur l'approximation PEM et une dernière basée sur l'utilisation du clustering appelée méthode pseudo Monte Carlo. Après avoir comparé la pertinence des méthodes sur deux réseaux test, la méthode pseudo Monte Carlo est appliquée, pour son gain en temps de simulation et son adaptabilité, dans un cas d'application qui porte sur l'estimation de la probabilité de dépassement des limites du courant de neutre en fonction du déséquilibre de production PV installée.

**Title :** Probabilistic load flow computation for unbalanced distribution grids with distributed generation

**Abstract :** The current changes on the electrical system bring out economic and technical issues in the management of the latter. Among these issues, the impact of distributed generation and VEs on the technical constraints of the distribution network. The aim of this thesis is to study probabilistic models to estimate the impacts of photovoltaic production and electrical vehicles on medium and low voltage distribution networks. Two different probabilistic models of

production and consumption were studied : one based on the fitting of historical data by one probability density function, the other one based on the data clustered in groups defined by a standard profile and a probability of occurrence. Three probabilistic load flow techniques have been studied in this thesis. The first is based on the Monte Carlo simulation method, the second is based on the PEM approximation method and the last, based on the use of clustering, is called pseudo Monte Carlo method.

