

Filière Energie et techniques environnementales

Orientation Smart Grid

Travail de bachelor Diplôme 2020

Guillaume Dubuis

*Prédiction d'événements extrêmes dans les
réseaux électriques à l'aide de machine learning*

- Professeur
Philippe Jacquod
- Expert
Matthias Bucher
- Date de la remise du rapport
14.08.2020

Ce rapport est l'original remis par l'étudiant.
Il n'a pas été corrigé et peut donc contenir des inexactitudes ou des erreurs.

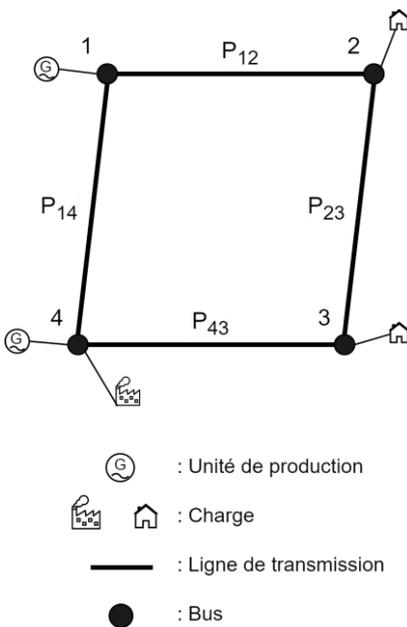
SYND	ETE	TEVI
X	X	X

Filière / Studiengang ETE	Année académique / Studienjahr 2019/20	No TD / Nr. DA SG/2020/89
Mandant / Auftraggeber <input checked="" type="checkbox"/> HES—SO Valais <input checked="" type="checkbox"/> Industrie <input type="checkbox"/> Etablissement partenaire <i>Partnerinstitution</i>	Etudiant / Student Guillaume Dubuis Professeur / Dozent Philippe Jacquod	Lieu d'exécution / Ausführungsort <input checked="" type="checkbox"/> HES—SO Valais <input type="checkbox"/> Industrie <input type="checkbox"/> Etablissement partenaire <i>Partnerinstitution</i>
Travail confidentiel / vertrauliche Arbeit <input type="checkbox"/> oui / ja ¹ <input checked="" type="checkbox"/> non / nein	Expert / Experte (données complètes) Matthias Bucher, Swissgrid	

Titre / Titel Prédiction d'événements extrêmes dans les réseaux électriques à l'aide de machine learning
Description / Beschreibung Avec la transition énergétique, les réseaux électriques sont appelés à fonctionner dans des régimes opérationnels très différents, par contre leur sécurité de fonctionnement est toujours évaluée selon les mêmes protocoles "n-1" depuis des décennies. Le but de ce travail de diplôme est dans un premier temps d'investiguer numériquement des modèles simples de réseaux électriques dans des modes de fonctionnement normaux ou potentiellement problématiques. Dans un deuxième temps on construira un algorithme de machine learning capable d'identifier rapidement les situations potentiellement problématiques. Le but est de construire un code numérique capable de prédire l'apparition de modes de fonctionnement dangereux du point de vue de la sécurité d'approvisionnement avant que ceux-ci ne se manifestent, ou au moins en temps réel.
Objectifs / Ziele — Comprendre le fonctionnement des réseaux électriques haute tension, — Comprendre la modélisation numérique de ces réseaux, — Comprendre et être capable de coder et d'utiliser un algorithme prédictif de machine learning.

Signature ou visa / Unterschrift oder Visum Responsable de l'orientation / filière <i>Leiter der Vertiefungsrichtung / Studiengang:</i>  ¹ Etudiant / Student : 	Délais / Termine Attribution du thème / Ausgabe des Auftrags: 25.05.2020 Présentation intermédiaire / Zwischenpräsentation Semaine / Woche 26 (22.06 – 26.06.2020) Remise du rapport / Abgabe des Schlussberichts: 14.08.2020, 12:00 Exposition / Ausstellung der Diplomarbeiten: 28.08.2020 (si autorisé / falls genehmigt) Défense orale / Mündliche Verfechtung: Semaine / Woche 36 (31.08 – 04.09.2020)
---	--

¹ Par sa signature, l'étudiant-e s'engage à respecter strictement la directive DI.1.2.02.07 liée au travail de diplôme.
Durch seine Unterschrift verpflichtet sich der/die Student/in, sich an die Richtlinie DI.1.2.02.07 der Diplomarbeit zu halten.



Prédiction d'événements extrêmes dans les réseaux électriques à l'aide de machine learning

Diplômant **Guillaume Dubuis**

Objectif du projet

Ce travail étudie l'intégration d'algorithmes de machine learning dans l'évaluation de la sécurité des réseaux de transport électrique.

Méthodes | Expériences | Résultats

Dans le domaine des réseaux électriques, le critère de sécurité est défini par une exploitation des lignes dans leurs limites opérationnelles. Les gestionnaires de réseau de transport (GRT) utilisent les calculs de « load flow » afin de déterminer les flux de puissance des lignes. La première étape de ce travail, a été le développement d'un algorithme de load flow basé sur la méthode itérative de Newton-Raphson.

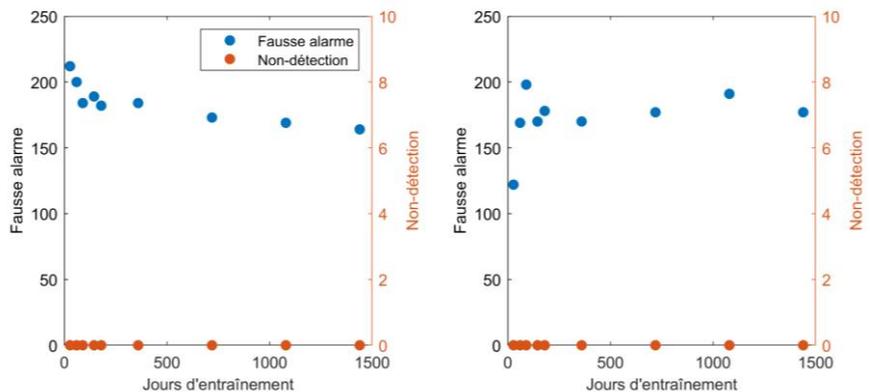
Après validation, cet algorithme a permis de générer les jeux de données nécessaires à l'entraînement et l'évaluation des modèles de machine learning. Ces derniers ont pour but de classifier les situations en deux catégories : saine ou à risque. Les performances sont examinées selon les deux types d'erreurs possibles : les non-détections et les fausses alarmes. Selon ces critères, deux modèles se distinguent : le réseau de neurones et l'estimateur à vecteur de support.

Travail de diplôme
| édition 2020 |

Filière
*Energie et techniques
environnementales*

Domaine d'application
Smart Grid

Professeur responsable
*Philippe Jacquod
philippe.jacquod@hevs.ch*



Évolution des performances du réseau de neurones (gauche) et de l'estimateur à vecteur de support (droite)

Ces résultats montrent que les algorithmes plus évolués de machine learning peuvent apporter une aide supplémentaire dans la gestion des réseaux électriques.

Table des matières

Liste des figures	II
Liste des tableaux	III
Glossaire	IV
1 Introduction	1
1.1 Objectifs	1
1.2 Réseau électrique.....	1
1.2.1 Structure du réseau	2
1.2.2 Sécurité N –1	3
1.2.3 Calculs de load flow	5
1.2.4 Stratégie proposée.....	5
1.3 Machine learning.....	6
2 Méthodologie.....	7
2.1 Solution au load flow par Newton – Raphson.....	7
2.1.1 Validation de l’algorithme.....	10
2.2 Création des jeux de données	10
2.2.1 Cas de test IEEE	10
2.2.2 Limite thermique	11
2.2.3 Génération des jeux de données.....	11
2.2.4 Congestion des jeux de données	13
2.3 Développement du machine learning	14
2.3.1 Les estimateurs	14
2.3.2 Optimisation des hyperparamètres.....	17
2.3.3 Matrice de confusion	19
3 Résultats	21
3.1 Load flow implémenté vs MATPOWER.....	21
3.2 Performances des estimateurs.....	21
4 Conclusions	25
Références.....	26
Annexes.....	27

Liste des figures

Figure 1: Illustration des sept niveaux du réseau électrique [3]	2
Figure 2: Illustration du fléchissement sur une ligne électrique.....	3
Figure 3: Schéma unifilaire d'un réseau électrique à cinq bus	4
Figure 4: Schéma unifilaire d'un réseau symétrique à trois bus	8
Figure 5: Diagramme séquentiel de la méthode Newton–Raphson, i représente le nombre d'itérations [7]	9
Figure 6 : Illustration de la topologie du cas de test IEEE 57–bus, les 50 charges (en bleu), les 6 générateurs (en vert) et le « slack bus » (en rouge).....	10
Figure 7 : Quart d'heure de congestion en fonction de la charge de ligne pour les jeux de test des deux scénarios	13
Figure 8: Quart d'heure de congestion pour le pire cas des calculs en $N - 1$, en fonction de la charge de ligne pour les deux scénarios	14
Figure 9: Classification de 3 étiquettes dans un espace à deux dimensions ($k = 15$) [14] .	15
Figure 10: SVC avec deux classes séparées par un hyperplan (trait continu) avec trois échantillons sur les limites (traitillé), appelés « vecteurs de support» [14].....	16
Figure 11: Réseau de neurones (MPLC) à une couche avec une sortie [14]	16
Figure 12: Forêt d'arbre décisionnel avec 600 arbres [14]	17
Figure 13: Principe de validation croisée avec 5 découpes du jeu d'entraînement initial (gris), les jeux de validation (bleu) et les jeux d'entraînement (vert)	18
Figure 14 : Illustration d'une matrice de confusion.....	20
Figure 15 : Fausses alarmes (bleu) et non-détections (rouge) en fonction du nombre de jours d'entraînement pour les différents modèles sur le scénario alourdi, les caractéristiques utilisées sont les injections	21
Figure 16: Fausses alarmes (bleu) et non-détections (rouge) en fonction du nombre de jours d'entraînement pour les différents modèles sur le scénario alourdi, les caractéristiques utilisées sont les productions	23
Figure 17: Fausses alarmes (bleu) et non-détections (rouge) en fonction du nombre de jours d'entraînement pour les différents modèles sur le scénario normal, les caractéristiques utilisées sont les injections	23
Figure 18: Fausses alarmes (bleu) et non-détections (rouge) en fonction du nombre de jours d'entraînement pour les différents modèles sur le scénario normal, les caractéristiques utilisées sont les productions	24

Liste des tableaux

Tableau 1: Type de bus pour les calculs de load flow [7]	8
Tableau 2: Caractéristiques des jeux de données selon les scénarios	13
Tableau 3: Valeurs des hyperparamètres retenus après la validation croisée pour chaque modèle sur les deux scénarios.....	19

Glossaire

- Caractéristique** De l'anglais « feature », elle désigne les entrées des modèles d'apprentissage automatique
- Estimateur** Autre mot signifiant modèle d'apprentissage automatique
- Étiquette** De l'anglais « label », elle désigne les sorties des modèles d'apprentissage automatique
- GRT** Gestionnaire de réseau de transport
- HT, MT, BT** Abréviation pour « Haute, Moyenne ou Basse Tension »
- IEEE** Institute of Electrical and Electronics Engineers
- kNN** Modèle d'apprentissage automatique des k plus proches voisins ou « k-Nearest Neighbors »
- NN** Modèle d'apprentissage automatique à réseau de neurones ou « Neural Network »
- Nœud** Dans un système électrique, un nœud est l'endroit où les lignes électriques sont connectées ensemble. Si deux ou plusieurs lignes sont connectées ensemble, ils sont sur le même nœud. Les nœuds sont aussi appelés bus.
- RF** Modèle d'apprentissage automatique à forêts d'arbres décisionnels « Random Forest »
- SVC** Modèle d'apprentissage automatique à vecteurs de support ou « Support Vector Classification »
- Topologie** Dans un système électrique, la topologie représente la manière dont les nœuds sont interconnectés les uns avec les autres.

1 Introduction

La loi sur l'énergie entrée en vigueur en janvier 2018 comporte trois orientations. La première orientation vise à réduire la consommation énergétique de la Suisse. La deuxième et la troisième orientation tendent au remplacement des centrales nucléaires et des centrales à énergies fossiles par des énergies renouvelables. La production annuelle d'électricité à partir des nouvelles énergies renouvelables doit atteindre au moins 11,4 térawattheures d'ici 2035 [1]. Ces énergies renouvelables ont pour principale caractéristique d'être intermittentes ce qui accroît les problématiques de gestion et de sécurité du réseau électrique. Avec l'avancée des nouvelles technologies de l'information et de la communication, le secteur de l'énergie est en pleine mutation. Les acteurs du marché de l'électricité y voient une occasion de se moderniser afin d'augmenter leur fiabilité, leur efficacité et leur sécurité. À l'heure actuelle, la gestion des réseaux est basée sur les capacités et l'expérience des opérateurs. Dans le contexte de la transition énergétique, les gestionnaires envisagent le développement d'intelligences artificielles, en particulier, des algorithmes d'apprentissage automatique de l'anglais « machine learning » qui aideront à la détection et à la résolution de situations problématiques.

1.1 Objectifs

Dans un premier temps, ce travail va étudier les numérisations des réseaux de transports. L'objectif est de développer un algorithme de « load flow » afin de déterminer le niveau de sécurité d'un réseau. Puis, dans un second temps, il s'agit d'étudier la capacité des modèles de machine learning à évaluer eux aussi ce niveau de sécurité. L'objectif final est de détecter les événements extrêmes sur le réseau électrique avant qu'ils ne se produisent, ou au moins en temps réel.

Ce rapport est structuré en trois parties. Le premier chapitre « Introduction » porte sur la présentation du réseau électrique et de l'apprentissage automatique. Le deuxième « Méthodologie » développe la numérisation du réseau de transport et la mise en œuvre de l'apprentissage automatique. Finalement, le troisième présente les résultats et les conclusions.

1.2 Réseau électrique

L'objectif de ce sous-chapitre est de présenter le réseau électrique dans son ensemble afin d'en comprendre son fonctionnement ainsi que son système de gestion. L'énergie électrique peut être transformée en de nombreuses autres formes d'énergie tels que l'énergie mécanique (moteurs électriques), l'énergie thermique (appareil de chauffage) ou l'énergie lumineuse. Le principal avantage de l'énergie électrique est sa facilité de transport. Le réseau électrique sert justement à transporter cette dernière des producteurs aux consommateurs. Depuis la mise en service de la première ligne électrique suisse en 1884, le réseau électrique n'a pas cessé de s'agrandir.

L'augmentation des clients et de leur consommation pousse à la création de nouvelles centrales de production et de lignes de transport. L'énergie électrique correspond à un quart de la consommation énergétique totale de la Suisse [2].

1.2.1 Structure du réseau

Le réseau électrique suisse combine les centrales de production, un réseau de transport, des réseaux de distribution et des consommateurs. Il s'agit d'un système PTDU (Production, Transport, Distribution, Utilisation). Ce dernier, illustré par la Figure 1, est articulé sur sept niveaux : quatre de tension et trois de transformation. Cette infrastructure historique a été dimensionnée pour un flux de puissance unidirectionnel, c.-à-d. le transport de l'énergie des producteurs aux consommateurs.

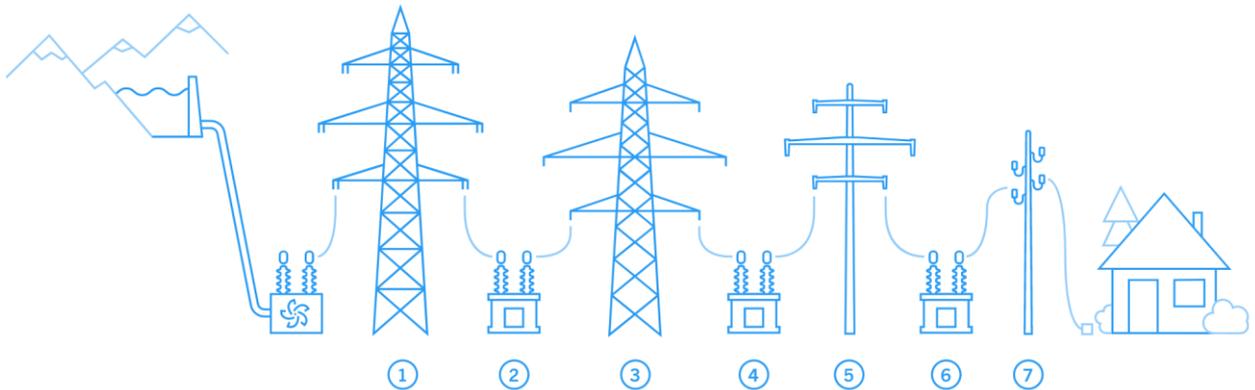


Figure 1: Illustration des sept niveaux du réseau électrique [3]

Le niveau 1 correspond au réseau de transport. Il s'agit de lignes majoritairement aériennes de très haute tension (THT, 220 et 380 kilovolts). Les connexions internationales sont raccordées à ce niveau. Ce réseau relie les principales installations de production électrique au réseau de niveau de tension inférieur (3 selon Figure 1) et permet le transit d'une grande quantité d'énergie. Swissgrid est l'unique propriétaire et gestionnaire du réseau de transport (GRT) en Suisse. Il régule la fréquence et la tension du niveau 1 (service système).

Le niveau 3 correspond aux réseaux de distribution suprarégionaux. Il s'agit de lignes à haute tension (HT, 36 et 170 kilovolts). Ce réseau a le même objectif que celui de transport avec un seuil de tension inférieur dû aux contraintes technico-économiques. Par conséquent, les quantités d'énergie transmises et les distances de transmission sont moindres. Le niveau 3 n'est géré par Swissgrid. Par exemple, le réseau suprarégional valaisan appartient à plusieurs sociétés tandis que sa surveillance a été confiée aux FMV (Forces Motrices Valaisannes).

Le niveau 5 correspond au réseau de distribution régional. Il s'agit de lignes à moyenne tension (MT, 1 à 36 kilovolts). Les gestionnaires de ce réseau de distribution (GRD) sont plus de 600 en Suisse et gèrent aussi le niveau 7 qui correspond aux réseaux de distribution locaux. Les lignes du niveau 7 sont en basse tension (BT, inférieure à 1 kilovolt, majoritairement 400 volts) et sont le plus souvent enterrées.

Pour chacun des niveaux de tension, le réseau électrique comporte une topologie différente. Le réseau HT et THT présente une topologie maillée, avec plusieurs liaisons nœud à nœud, pour maximiser la sécurité d'approvisionnement. Pour les réseaux MT et BT, les topologies radiales et

arborescentes sont privilégiées, car moins coûteuses. En effet, plus on descend en niveau de tension, plus il y a de kilomètres de lignes.

1.2.2 Sécurité N –1

Les flux de puissance dans les lignes induisent eux-mêmes du courant qui va chauffer les lignes électriques. Il s'agit de l'effet Joule et est la principale cause de pertes sur le réseau de transport. Ce phénomène est dû à la résistance de la ligne électrique qui s'oppose à la circulation du courant. Comme la ligne chauffe, le conducteur se dilate et se rapproche du sol, phénomène appelé « fléchissement ». Lorsque la puissance augmente, le fléchissement s'accroît également. Ces situations dangereuses peuvent amener à des courts-circuits qui sont néfastes pour l'infrastructure et pour les obstacles environnants.

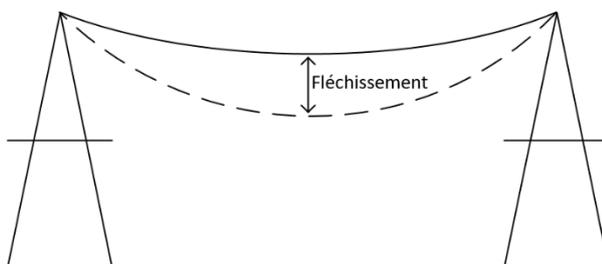


Figure 2: Illustration du fléchissement sur une ligne électrique

Afin d'éviter les coupures, les lignes possèdent une « limite thermique » qui se traduit par la puissance maximale de la ligne. Lorsque cette limite est atteinte au-delà d'un certain temps, une protection agit automatiquement sur la ligne électrique et la déconnecte du réseau. Le principal inconvénient de ces protections est que la ligne est déconnectée et augmente souvent le flux de puissance électrique dans certaines autres lignes. Ces dernières pourraient elles-mêmes surchauffer et être automatiquement déconnectées les unes après les autres. Ce phénomène est appelé une « surcharge en cascade » et peut conduire à une panne d'électricité locale voir un « black-out ». C.-à-d., une panne de courant à large échelle. En 2003, l'Italie a perdu son réseau à la suite de la perte d'une ligne internationale la reliant à la Suisse. Sa consommation a été reportée sur d'autres lignes, ce qui a entraîné leur déclenchement pour cause de surcharge.

Il est possible de dépasser la limite thermique pendant un certain temps. Celui-ci dépend du pourcentage de surcharge calculé selon l'équation (6). Une faible surcharge laisse plus de temps aux dispatchers afin de résoudre le problème qu'une surcharge conséquente. Par exemple, une charge de 110 % est soutenable pendant environ 15 minutes [4].

Après une coupure à large échelle, le redémarrage se fait par séquence et commence autour des sites de production capables de démarrer sans courant électrique (Générateur au pétrole, batterie, aimants permanents). En effet, il faut exciter les machines électriques, assurer les sécurités et les séquences de connexion au réseau. Le réseau s'étend alors depuis ces sites. Par conséquent, les coupures de courant peuvent durer plusieurs heures voir des jours. Afin d'éviter de tels événements

dramatiques, le GRT veille à ce que le flux dans chaque ligne soit toujours inférieur à la limite thermique. Ceci est souvent appelé sécurité « N ». Il s'agit de la première étape, mais elle n'est pas suffisante. Les conditions atmosphériques, comme la foudre et le givre, les machines de chantier et les oiseaux peuvent également être source de défauts (en plus des déclenchements dû à la limite thermique). Les réseaux de transport couvrent souvent de vastes zones et sont composés de plusieurs centaines voire de milliers de lignes. Ce grand nombre de lignes augmente la probabilité qu'un défaut provoque le déclenchement de l'une d'elles. Les GRT exploitent alors souvent leur réseau électrique en utilisant le critère de sécurité « N-1 ». Qu'est-ce qui se passe lorsqu'une composante du réseau disparaît ? Bien que la consommation et la production restent identiques, le réseau est alors différent et les charges des lignes sont déviées ailleurs. Aucune ligne ne doit dépasser sa limite thermique. L'opération est répétée pour chaque ligne. Si aucune situation dangereuse n'est détectée, le réseau est sûr à N-1. Au contraire, lorsqu'un régime potentiellement dangereux est détecté, une action curative est mise en place afin de rendre le réseau sûr à N-1. Par exemple, il est possible de demander d'augmenter la production dans une certaine centrale et de la diminuer dans une autre.

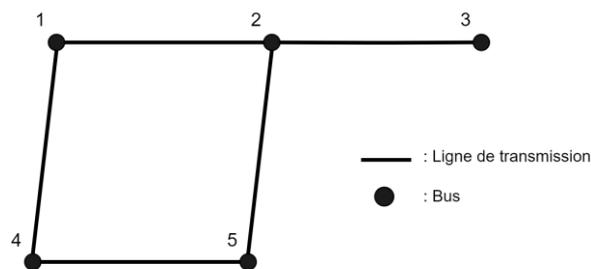


Figure 3: Schéma unifilaire d'un réseau électrique à cinq bus

Certaines liaisons connectent uniquement un bus en fin de ligne (bus N° 3 Figure 3). Ce bus est alors alimenté par une seule ligne. Lorsque celle-ci est supprimée, le bus n'est plus alimenté. Ce type de ligne n'est pas considéré dans les calculs en N-1 des réseaux de transport. Bien que ce dernier soit maillé, il y aura toujours quelques nœuds périphériques. Sur le réseau européen, il y a entre 3 500 à 4 000 nœuds dont une centaine sont dans ce cas (ordre de grandeur approximatif). La redondance peut parfois venir du niveau inférieur de tension (p. ex. MT).

L'Annexe 1 issue de la statistique de l'électricité 2018 [2] montre que la Suisse importe davantage d'électricité pendant les mois d'hiver. Ces derniers sont ainsi les plus risqués, du point de vue des puissances transmises par les lignes en suisse. D'autres facteurs influencent ce risque. Dans les moments de fortes production et consommation, les lignes sont chargées, mais cela signifie aussi que beaucoup de centrales sont en production et qu'elles peuvent absorber les erreurs plus aisément que lorsque moins de générateurs sont en services. Finalement les jours d'été, où une faible consommation est entièrement couverte par des énergies renouvelables (photovoltaïque, éolienne), sont potentiellement des journées dangereuses. En effet, peu de moyens de production peuvent absorber les pics et creux de consommation. Ce domaine s'appelle la dynamique des réseaux et ne sera pas pris en compte dans ce travail.

Actuellement, les actions curatives aux problématiques en $N-1$ sont trouvées par les collaborateurs « dispatcher » expérimentés de Swissgrid. Les solutions sont essentiellement basées sur la mémoire des événements passés. Ce qui implique de former des successeurs avant que les collaborateurs mentionnés ne quittent l'entreprise. Cette démarche a très bien fonctionné jusqu'à présent. Néanmoins, une approche systématique s'adapterait probablement mieux aux évolutions futures du réseau. Ces dix dernières années, le réseau électrique a véritablement été intégré au niveau européen, du point de vue physique (interconnexion) et des marchés (libéralisation). Ce travail se concentre sur l'intégration physique.

1.2.3 Calculs de load flow

La technique de détermination des tensions de bus dans un réseau est généralement appelée « load flow ». Lorsque l'amplitude de la tension et l'angle de tension sur tous les bus sont connus, l'état du système est complètement déterminé et toutes les propriétés du système peuvent être calculées (par exemple la puissance transmise par les lignes et les pertes). Le point de départ d'un problème de load flow est un schéma unifilaire du réseau électrique (par exemple Figure 3), à partir duquel les données d'entrée des algorithmes peuvent être obtenues.

Le nombre de nœuds dans les réseaux est si élevé que les calculs sont trop complexes pour être faits à la main. C'est pourquoi des méthodes numériques sont utilisées. Principalement, des méthodes itératives pour la résolution d'équations algébriques non linéaires. Elles se nomment :

- Gauss – Seidel
- Newton – Raphson
- Fast – decoupled

La méthode retenue dans ce travail est celle de Newton – Raphson. L'expérience a montré que Newton – Raphson converge dans de nombreux cas où Gauss – Seidel diverge pour des calculs de load flow. De plus, le nombre d'itérations nécessaires à la convergence est indépendant du nombre de nœuds N étudiés pour Newton – Raphson, mais augmente avec N pour Gauss – Seidel. La plupart des problèmes de flux de puissance de Newton – Raphson convergent en moins de 10 itérations [5].

1.2.4 Stratégie proposée

Les calculs de load flow permettent au GRT d'évaluer sécurité (N et $N-1$) des réseaux. Au vu du nombre élevé de lignes, ces calculs en $N-1$ nécessitent une grande quantité d'opérations. Il en résulte un temps de calcul important. De plus, les réseaux électriques vont évoluer, car les grosses productions centralisées vont progressivement réduire leurs productions au profit de petites productions décentralisées (photovoltaïque, mini-hydraulique, éolienne). Ces nouveaux moyens de production, étant moins contrôlés, impliquent d'augmenter la fréquence des calculs $N-1$. Afin de s'assurer en amont qu'il n'est pas nécessaire de mobiliser du contrôle (des réserves) le cas échéant.

De plus, les gestionnaires de réseau européen comme EDF pour la France, différentes entités différentes pour l'Allemagne (Amprion, TransnetBW, TenneT TSO et 50Hertz), TERNA pour l'Italie, centralisent leurs résultats afin de détecter les problèmes. Il revient aux gestionnaires concernés de résoudre le problème. Il est possible que la suppression d'une ligne en Suisse crée une difficulté en Allemagne. Cette problématique est difficilement détectable à l'heure actuelle. Ainsi, les calculs N-1 au niveau européen seraient très intéressants dans la mesure où ils seraient effectués assez rapidement afin de permettre au GRT de réagir dans un délai acceptable.

L'enjeu est alors de réduire le temps d'évaluation de la sécurité. A moins d'augmenter la puissance de calcul, il est peu probable qu'une amélioration de calcul de load flow par la méthode itérative apporte un gain de temps significatif. Une solution « plus intelligente » serait d'estimer la sécurité à l'aide de machine learning. Ces algorithmes novateurs sont généralement plus efficaces et l'état du réseau serait ainsi déterminé plus rapidement. Lorsque ces derniers détectent une situation à risque, les calculs de load flow traditionnels sont effectués. Cela permet de réduire la fréquence de ceux-ci. Cette stratégie nécessite des algorithmes de machine learning fiable. Ils doivent détecter toutes les situations à risque sans émettre trop de fausses alarmes, au risque de ne pas améliorer la stratégie actuelle.

1.3 Machine learning

Le machine learning est l'art de « Donner à une machine la capacité d'apprendre sans la programmer de façon explicite »¹. En pratique, un modèle mathématique est développé à partir de données expérimentales. Les trois techniques sont l'apprentissage supervisé, l'apprentissage non supervisé et l'apprentissage par renforcement. La machine reçoit des données caractérisées par des variables d'entrées X et annotées d'une variable de sortie Y . Dans le vocabulaire du machine learning, les variables X sont appelées « Features » (caractéristiques) et la variable Y est nommée « label » (étiquette). L'objectif est que la machine apprenne à prédire la valeur de sorties Y en fonction des caractéristiques d'entrée qu'on lui donne. Tout d'abord, le type de modèle est spécifié. Il en existe différents types : des modèles linéaires, polynomiaux, à arbre de décision, des réseaux de neurones, etc. Ensuite, les hyperparamètres du modèle sont choisis, par exemple, le nombre de branches dans l'arbre de décision ou le nombre de neurones dans le réseau. Le modèle est entraîné avec les caractéristiques d'entrées ce qui constitue les données d'apprentissage ou « training set ». Un algorithme d'optimisation trouve les hyperparamètres qui fournissent les meilleures performances sur le training set. Il s'agit par exemple de trouver la pente et l'ordonnée à l'origine pour une régression linéaire. C'est ce que l'on appelle la phase d'entraînement. Le processus est alors terminé et le modèle est prêt à être utilisé. Lorsque la machine reçoit de nouvelles données sans étiquette, elle va utiliser le modèle pour prédire la valeur de l'étiquette. Cette méthode permet de résoudre des problèmes de régression quand la sortie est une variable continue (quantitative) et des problèmes de classification quand la sortie est une variable discrète (qualitative).

¹ Arthur Lee Samuel (1901–1990) pionnier du machine learning

Cette technologie est déjà utilisée au sein de Swissgrid au travers de la « predictive maintenance », c.-à-d. que l'algorithme aide à la gestion des travaux de maintenance. Par exemple, dans une centrale électrique la maintenance se fait à l'aide de capteurs sur la turbine. De plus, elle est ouverte périodiquement afin de détecter les éventuels défauts (inspection de la conduite forcée). Les endroits potentiellement problématiques sont ainsi localisés. Concernant le réseau de transport, il est difficilement envisageable d'aller vérifier tous les kilomètres de lignes. Swissgrid connaît les régimes d'utilisation de la ligne. Les constructeurs ont fourni les plages de fonctionnement. Un algorithme de machine learning indique lorsqu'il est nécessaire d'aller vérifier sur place. Ce délai dépend de l'utilisation et de la météo. Une surcharge n'a pas le même effet en hiver qu'en été, car la température extérieure influence les caractéristiques de la ligne.

2 Méthodologie

Ce chapitre présente les principaux développements de ce travail. Dans un premier temps, la méthode de calculs des load flow est décrite. Puis, le type de réseau électrique ainsi que les jeux de données sont introduits. Finalement, l'implémentation des modèles de machine learning est présentée.

2.1 Solution au load flow par Newton – Raphson

Le load flow permet de déterminer l'amplitude de la tension et l'angle de tension sur tous les bus d'un réseau donné. Pour ce faire, il faut spécifier la topologie du réseau ainsi que les puissances actives et réactives injectées. Les données d'entrée sont constituées de données de bus et de données de ligne de transport. Les quatre variables suivantes sont associées à chaque bus : amplitude de tension, angle de phase, puissance réelle nette et puissance réactive fournie au bus. Sur chaque bus, deux de ces variables sont spécifiées comme données d'entrée, et les deux autres sont des inconnues à calculer par l'algorithme [6].

Dans les réseaux, il y a trois types de bus (voir Tableau 1). Le bus de référence ou « slack bus » est unique dans le réseau. Sa tension est fixée à 1 pu (unité réduite ou « per unit ») et son angle à 0. Les bus de production « PV bus », dont la puissance et la tension sont connues, s'avèrent minoritaires dans les réseaux. Les bus de charge « PQ bus », dont la puissance active et la puissance réactive sont connues, constituent la majeure partie des nœuds. Les données d'entrée pour chaque ligne de transport comprennent l'impédance des circuits équivalents π en pu ainsi que les deux numéros de bus auxquels la ligne est connectée.

Type de bus	Description	Est une variable connue ?			
		U_i	δ_i	P_i	Q_i
Slack bus	Unique dans le réseau	Oui	Oui	Non	Non
PV bus	Bus de production	Oui	Non	Oui	Non
PQ bus	Bus de charge	Non	Non	Oui	Oui

Tableau 1: Type de bus pour les calculs de load flow [7]

Un réseau symétrique avec N bus est considéré. La tension et l'angle au bus numéro i sont notés U_i et δ_i . La Figure 4 montre un schéma avec N égal à 3. Toutes les variables sont exprimées en pu.

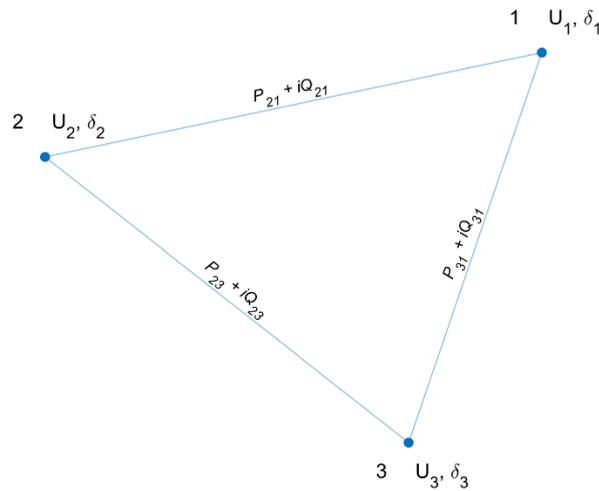


Figure 4: Schéma unifilaire d'un réseau symétrique à trois bus

À partir des lois de Kirchhoff (loi des nœuds et loi des mailles), la puissance active injectée au bus numéro i notée P_i peut être définie comme,

$$P_i = U_i \sum_{j=1}^N Y_{ij} U_j \cos(\theta_{ij} + \delta_j - \delta_i) \quad i = 1, 2, \dots, N, \quad (1)$$

où Y_{ik} et θ_{ij} sont respectivement le module et l'angle de l'admittance entre le bus numéro i et j . Alors que Q_i est la puissance réactive donnée par

$$Q_i = -U_i \sum_{j=1}^N Y_{ij} U_j \sin(\theta_{ij} + \delta_j - \delta_i) \quad i = 1, 2, \dots, N. \quad (2)$$

Le système d'équations, composé des équations (1) et (2), est non linéaire et nécessite une résolution numérique comme la méthode itérative de Newton–Raphson. Le but est de déterminer les tensions sur tous les bus d'un système. À partir d'une estimation initiale $x^{(0)}$ (estimation des inconnues du Tableau 1), une séquence d'estimations progressivement meilleures est effectuée ($x^{(1)}$, $x^{(2)}$, ...) qui converge, sauf exception, vers la solution [8] [9]. Le diagramme séquentiel de la Figure 5 décrit les différentes étapes de résolution. Les résidus Δg représentent la différence entre les puissances cibles (les puissances connues du Tableau 1) et les puissances calculées (à l'itération i). La matrice jacobienne notée $JAC^{(i)}$ permet de déterminer l'amélioration $\Delta x^{(i)}$ à apporter aux estimations afin de se rapprocher de la solution. Elle est composée des dérivées partielles (au point $x^{(i)}$) du système d'équations. Lorsque les résidus Δg sont inférieurs au critère de convergence ε , les calculs ont convergés vers la solution.

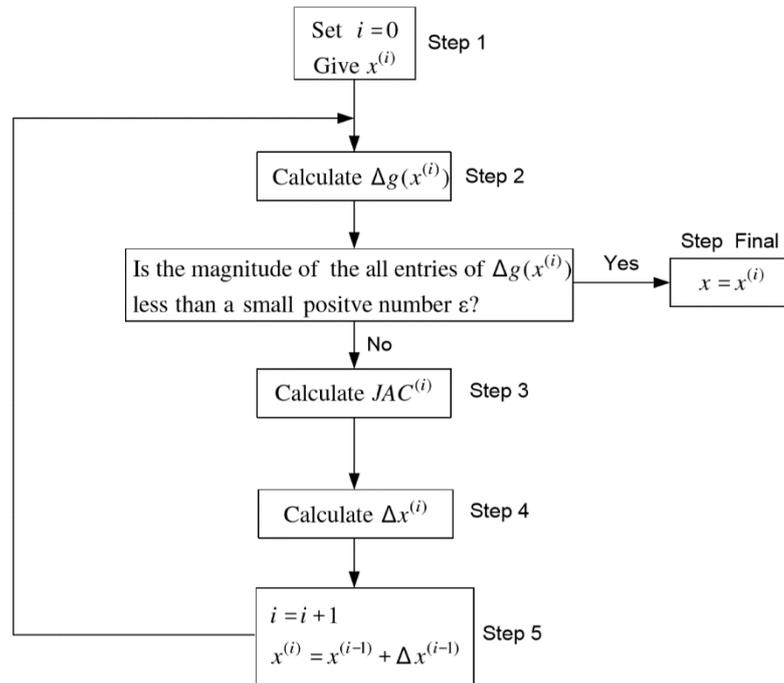


Figure 5: Diagramme séquentiel de la méthode Newton–Raphson, i représente le nombre d'itérations [7]

Une fois que les amplitudes de tensions et les angles sont connus, les puissances actives P_{ij} et les puissances réactives Q_{ij} entre le nœud i et le nœud j sont déterminées à l'aide des équations (3) et (4). Dans ces équations, g_{ij} et b_{ij} sont respectivement la partie réelle et la partie imaginaire de l'admittance de la ligne.

$$P_{ij} = g_{ij}U_i^2 - U_iU_j(g_{ij} \cos(\theta_{ij}) + b_{ij}\sin(\theta_{ij})) \quad (3)$$

$$Q_{ij} = -b_{ij}U_i^2 - U_iU_j[g_{ij} \sin(\theta_{ij}) + b_{ij}\cos(\theta_{ij})] \quad (4)$$

Davantage d'informations sur la méthode, la matrice jacobienne et les équations utilisées peuvent être trouvées dans les sources [6, 7, 8, 9]. L'algorithme développé est disponible sur Github (D. Guillaume, «Prédiction d'événements extrêmes sur les réseaux électriques à l'aide de machine learning,» disponible: <https://github.com/Guillaume-cpu/machineLearning-reseauElectrique>). Il a été développé sur MATLAB du fait de la facilité du calcul matriciel sur ce logiciel.

2.1.1 Validation de l'algorithme

Afin de valider l'algorithme implémenté, les résultats obtenus seront comparés avec ceux obtenus avec MATPOWER pour le cas d'IEEE « case57 ». MATPOWER est un ensemble de fichiers MATLAB pour résoudre les load flow et les problèmes de flux de puissance optimale aussi appelé « optimal power flow ». Il est conçu comme un outil de simulation simple d'emploi et ajustable pour chercheurs et éducateurs [10].

2.2 Création des jeux de données

2.2.1 Cas de test IEEE

Dans ce travail, le cas de test IEEE 57-bus a été choisi pour valider l'algorithme de « load flow » et en tant qu'exemple de topologie dans les scénarios de l'apprentissage automatique. Il comporte plus de 50 bus ce qui est intéressant pour de premiers essais. L'IEEE de l'anglais « Institute of Electrical and Electronics Engineers » est une association professionnelle mondiale reconnue dans l'étude des réseaux électriques.

Le cas de test IEEE 57-bus représente une simple approximation du système électrique américain tel qu'il était au début des années 1960. Les données ont été fournies par M. Iraj Dabbagchi et converties en format commun de données IEEE par M. Rich Christie à l'Université de Washington en août 1993 [11]. Il comporte les données de bus de charge, les données de production et les données de lignes.

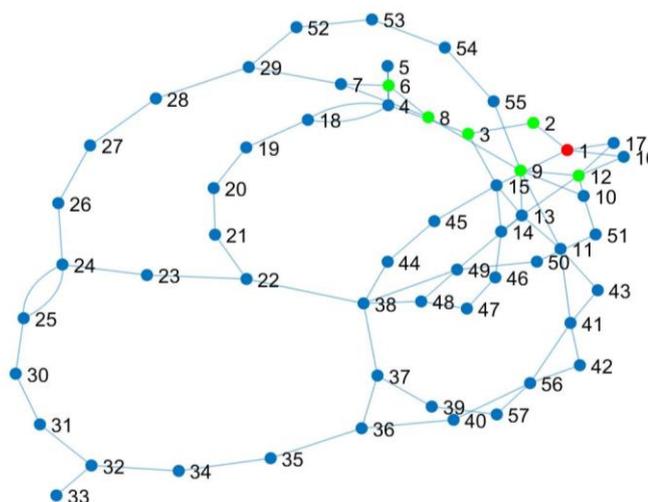


Figure 6 : Illustration de la topologie du cas de test IEEE 57-bus, les 50 charges (en bleu), les 6 générateurs (en vert) et le « slack bus » (en rouge)

Le IEEE 57–bus est constitué de 80 lignes, 50 bus de charge et de 7 bus de production. Ces derniers ne sont pas répartis sur l’ensemble du réseau. Cette répartition ne correspondant pas à un cas de réseau de transport standard (les congestions en N –1 peuvent en être affectées). La ligne reliant le bus N° 32 au bus N° 33 est la seule alimentant ce nœud. Lorsque celle–ci est supprimée pour la sécurité N –1 le nœud n’est plus connecté. C’est pourquoi ce nœud et cette ligne ont été supprimés du réseau dans le cadre de ce travail.

Les résultats de cette étude sont spécifiques à ce réseau. Pour une application au niveau européen, il serait nécessaire de se baser sur la topologie de ce réseau. L’objectif ici est de valider la méthode de détection à l’aide du machine learning.

2.2.2 Limite thermique

Les limites thermiques des lignes ne sont pas fournies pour le réseau IEEE 57–bus. La puissance thermique limite P_{ij}^{th} est approximée par

$$P_{ij}^{th} \cong U_i U_j b_{ij} \sin(\Delta\theta_{max}) \cong U_i^2 b_{ij} 0.3, \quad (5)$$

où U_i et U_j sont les tensions aux bornes de la ligne et b_{ij} est la susceptance de la ligne. Alors la surcharge Q_{ij}^{th} de la ligne est défini par

$$Q_{ij}^{th} = \frac{|P_{ij}|}{P_{ij}^{th}}. \quad (6)$$

Selon l’équation (6), la puissance de la ligne est à son maximum lorsque la différence d’angle entre les tensions est de 90° , car le sinus est égal à un. Cet état n’est pas fonctionnel et un tel réseau serait instable. Lorsque l’angle atteint environ 25° , le courant dans la ligne dépasse la limite thermique et la situation ne peut pas perdurer. D’un point de vue théorique, il est possible d’obtenir des courants plus élevés, mais physiquement les contraintes techniques ne permettent pas une différence d’angle supérieur à 30° (fléchissement § 1.2.2).

2.2.3 Génération des jeux de données

En apprentissage supervisé, deux types de données sont nécessaires : les données d’entraînement qui servent à optimiser les hyperparamètres des modèles. Et les données de test qui sont utilisées afin de connaître les erreurs finales et d’estimer les performances futures des modèles. Ainsi, il est important que ce jeu de données soit le plus représentatif possible de la réalité. Les données de test doivent être différentes des données d’entraînement afin d’éviter que les résultats ne soient biaisés. En effet, les estimateurs ont bien plus de facilité avec des données déjà rencontrées.

Concernant la phase de test, le choix des courbes de charges pour chaque nœud et des courbes de production pour les nœuds PU est primordial afin de correspondre au mieux à des situations réelles. Des profils journaliers sont créés à partir de données globales du réseau de transport suisse fournies par Swissgrid [12]. Les productions et consommations échantillonnées au quart d’heure sont

disponibles pour l'année 2009 jusqu'à 2019. Une période de 12 semaines est sélectionnée pour chaque année. Les courbes de charges journalières sont normalisées par rapport à leur puissance maximale. Une fois normalisés, les 84 profils sont répartis sur les différents nœuds. La répartition se fait selon le jour de la semaine. Ainsi, dans notre jeu de test, les profils de charges du lundi correspondent à un profil observé un lundi sur le réseau suisse et ainsi de suite pour les autres jours de la semaine. Le processus est répété sur les années suivantes, ce qui permet de simuler 15 jours, soit deux semaines et un lundi. Ainsi, il est assuré qu'il y aura des pics et des creux de consommations représentatifs dans le jeu. Ensuite les séries temporelles de chaque nœud sont multipliées par un facteur de densité de population. Certains nœuds ont une consommation plus élevée, car la population est plus importante à ses environs. Ce facteur a été choisi aléatoirement entre 1 et 10 pour chaque nœud et reste le même en tout temps et pour tous les jeux de données.

La puissance réactive est calculée à l'aide de la puissance active. Dans le réseau de transport, le rapport entre la puissance active et réactive d'un nœud est généralement compris entre 0,1 et 0,2. Un quotient a été attribué au hasard pour chaque bus dans cette plage. Il reste lui aussi le même en tout temps et pour tous les jeux de données.

Concernant la phase d'entraînement, le jeu de données a été créé de manière aléatoire. Néanmoins, le facteur de densité de population a été respecté. Ce jeu de données n'est plus une série temporelle (une situation n'a pas de lien avec la précédente). Il n'y a de ce fait plus de notions de pics ou de creux de consommation. Les performances des estimateurs sont liées au nombre de situations étudiées durant la phase d'entraînement. Ainsi, la génération aléatoire permet d'atteindre aisément un grand nombre de cas.

Lorsque l'algorithme de « load flow » a été validé, il est utilisé afin de fournir les étiquettes aux jeux de données précédemment décrits. Il y a deux types d'étiquettes : la sécurité N et la sécurité N-1. Ce critère de sécurité est évalué grâce à l'algorithme de load flow. Lorsque la surcharge Q_{ij}^{th} sur une des lignes est plus grande que 1, la situation est étiquetée à risque. Dans le cas contraire, elle est saine. Comme les estimateurs fonctionnent avec des valeurs numériques, les situations saines en N et en N-1 sont étiquetées avec un 0 et les situations à risque avec un 1.

2.2.4 Congestion des jeux de données

Ainsi deux scénarios ont été créés. Chaque scénario dispose d'un jeu d'entraînement et d'un jeu de test. Le premier scénario est dit « normal » et le second « alourdi ». Ce qui les différencie c'est l'occurrence des situations risquées comme le montre le Tableau 2.

Scénario	Jeu	Situation	Situation risquée	Occurrence des situations risquées en N [%]	Série
Normal	Entraînement	140 160 (4 ans)	7 090	5,3	Aléatoire
	Test	1 440 (2 sem.)	67	4,6	Temporelle
Alourdi	Entraînement	140 160	43 107	31	Aléatoire
	Test	1 440	363	25	Temporelle

Tableau 2: Caractéristiques des jeux de données selon les scénarios

La Figure 7 montre les histogrammes de congestion pour les flux de puissance des jeux de tests sur le réseau 57-bus. Ils montrent le flux normalisé avec la limite thermique pour les 80 lignes. Pour rappel, la phase d'entraînement est constituée de 15 jours avec un échantillonnage au quart d'heure. L'allure des congestions en N correspond à des situations réelles.

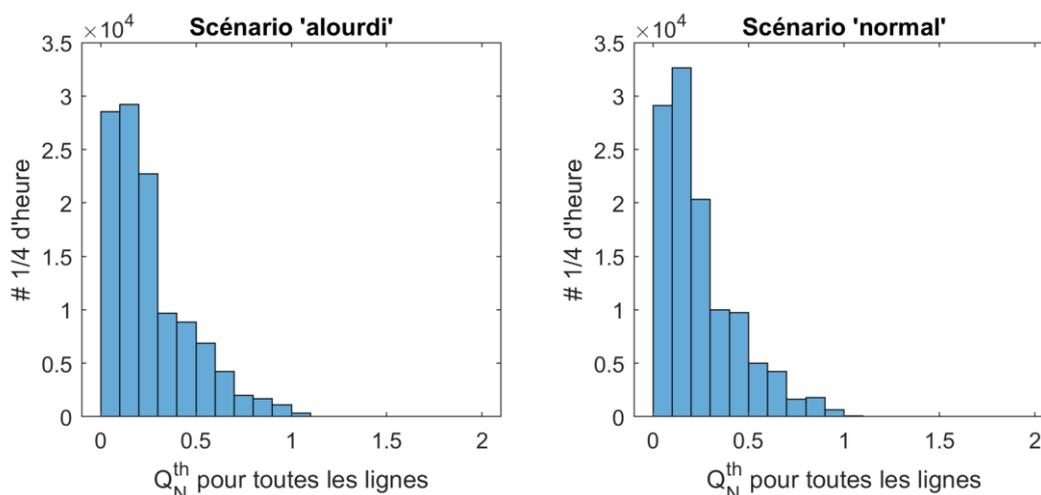


Figure 7 : Quart d'heure de congestion en fonction de la charge de ligne pour les jeux de test des deux scénarios

La congestion en N – 1 a été effectuée. Chaque ligne du réseau 57-bus est coupée successivement et un calcul de flux de puissance avec le réseau résultant est effectué. Le pire cas est défini comme le flux le plus élevé consécutif à la coupure de n'importe quelle ligne pour un quart d'heure donné. La Figure 8 montre les histogrammes pour le pire cas de congestions N–1 pour les deux scénarios en fonction de la charge sur les lignes. Les deux scénarios sont fortement congestionnés. Sur les

15 jours de test, le scénario normal présente 88 % de situation à risque alors que le taux monte à 96 % pour le scénario alourdi. Dans les deux scénarios, le pire flux dépasse souvent le triple de la limite thermique. Ceci peut être attribué à la répartition non homogène des bus de production sur le réseau 57-bus (Figure 6).

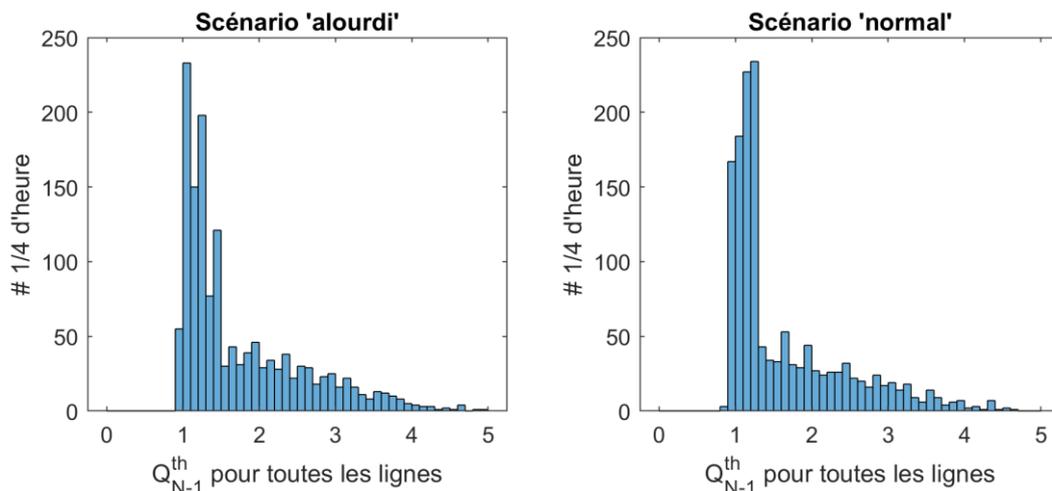


Figure 8: Quart d'heure de congestion pour le pire cas des calculs en $N - 1$, en fonction de la charge de ligne pour les deux scénarios

2.3 Développement du machine learning

L'objectif est de développer des modèles de machine learning capable d'évaluer le niveau de sécurité des réseaux électriques. Plus précisément, les modèles estiment le niveau de sécurité du réseau à un instant t avec les informations disponibles à ce moment. Ce qui revient à remplacer les algorithmes traditionnels de load flow. Pour chaque estimateur, trois types d'entrées sont expérimentés : les injections (charges et productions au temps t), uniquement les charges, puis uniquement les productions. La topologie du réseau étudié comporte 6 nœuds de production (le slack bus n'est pas compté). Les modèles avec les caractéristiques de production effectuent des estimations avec 12 % de l'état du système connu.

La principale problématique est que les situations à risque sont peu fréquentes sur les réseaux électriques. C'est pourquoi, deux scénarios ont été générés, un scénario alourdi en congestions et un scénario normal. De plus, les tests sont effectués en variant le nombre de données observées par les modèles durant la phase d'entraînement.

2.3.1 Les estimateurs

La première interrogation s'est située dans les types d'estimateurs à utiliser. Plusieurs modèles sont envisageables comme les arbres de décisions, les réseaux de neurones. Le choix s'est porté sur les quatre modèles suivants (parmi les plus utilisés): les k plus proches voisins (kNN), les machines à vecteurs de support (SVC), les réseaux de neurones (NN), et les forêts d'arbres décisionnels (RF). Ce choix couvre les modèles basés sur des distances dans un espace à n dimensions (kNN et SVC) et des modèles basés sur des fonctions de combinaisons avec des poids (NN et RF).

Les modèles sont développés sur python à l'aide de la librairie open source scikit-learn [13]. Celle-ci permet la mise en œuvre des algorithmes d'apprentissage automatique. Un grand nombre de modèles y est implémenté avec une architecture orientée objet dans laquelle chaque modèle dispose de sa propre classe. Il suffit de créer une instance de celle-ci et de préciser les hyperparamètres. Ceux-ci sont primordiaux, car les performances en dépendent directement. Sur le site web de scikit-learn, des tutoriels en anglais décrivent chaque modèle. Après l'initialisation, le modèle est entraîné, évalué et utilisé avec trois méthodes communes à toutes les classes de la librairie.

kNN – Le modèle des k plus proches voisins ou « k -Nearest Neighbors », est la méthode la plus utilisée pour la classification. L'algorithme est parmi les plus simples, car aucun modèle mathématique ne doit être développé. Les données d'apprentissage sont simplement stockées. Pour déterminer l'étiquette associée à une caractéristique, cette méthode prend en compte les k plus proches voisins dans l'échantillon d'apprentissage de la nouvelle caractéristique X . Le choix du nombre de voisins est très dépendant des données.

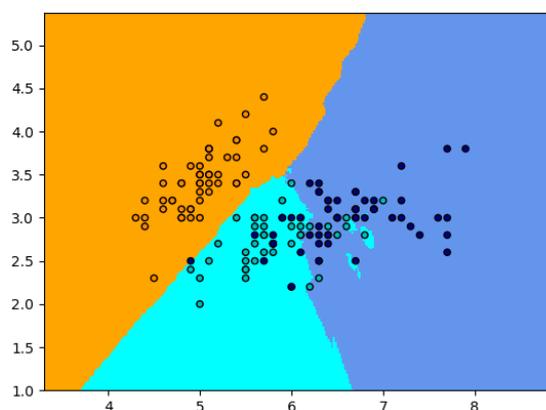


Figure 9: Classification de 3 étiquettes dans un espace à deux dimensions ($k = 15$) [14]

Sur la Figure 9, une nouvelle caractéristique se trouvant dans la zone orange sera étiquetée orange. En effet, la majorité de ses 15 plus proches voisins ont eux aussi l'étiquette orange. Ceci est vrai pour les deux autres couleurs.

SVC – Les modèles SVC ou « Support Vectors Classification » sont une des grandes familles de modèle en apprentissage supervisé. Ils ont pour avantage de travailler avec des données de grandes dimensions et nécessitent peu d'hyperparamètres. Ils sont utilisés pour de la classification et pour de la régression. Le principe consiste à séparer les différentes données à l'aide d'un hyperplan. Les vecteurs de support situés proches de la séparation permettent de calculer l'hyperplan optimal pour la classification.

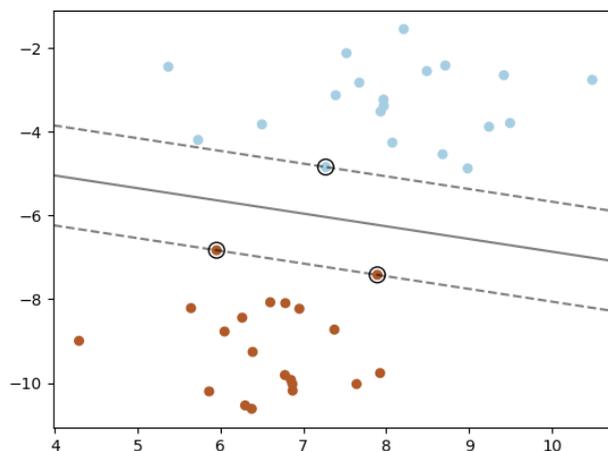


Figure 10: SVC avec deux classes séparées par un hyperplan (trait continu) avec trois échantillons sur les limites (traitillé), appelés « vecteurs de support » [14]

NN – Les réseaux des neurones ou « Neural Network » imitent le fonctionnement du cerveau. Un neurone reçoit une ou plusieurs informations en entrées et en renvoie une en sortie. Les entrées ont des poids différents et la sortie est déterminée selon la fonction du neurone. Une couche est formée de plusieurs neurones en parallèle. Dans un réseau, plusieurs couches de neurones sont mises en série. Une fonction d'optimisation détermine le poids des entrées de chaque neurone produisant le meilleur résultat.

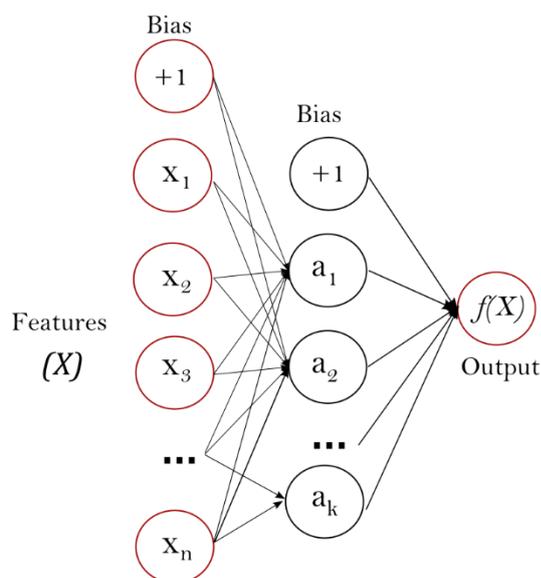


Figure 11: Réseau de neurones (MPLC) à une couche avec une sortie [14]

RF– Les forêts d’arbres décisionnels ou « Random Forest » combinent la simplicité des arbres de décision et la flexibilité ainsi que la généralisation des « ensembles méthodes ». Chaque arbre est entraîné sur des données aléatoires parmi le jeu d’entraînement. Une prédiction est effectuée en moyennant la prédiction de chaque arbre de la forêt.

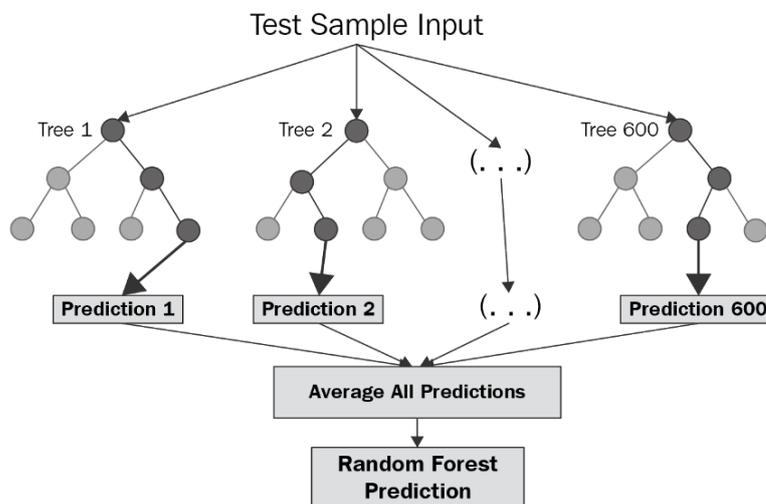


Figure 12: Forêt d’arbre décisionnel avec 600 arbres [14]

2.3.2 Optimisation des hyperparamètres

Les hyperparamètres d’un modèle influencent ses performances et définissent sa sensibilité. Lors d’apprentissage supervisé, il est important de ne pas être en surapprentissage ou « overfitting ». Ce qui signifie qu’un modèle se spécifie sur les données d’entraînement et perd tout sens de généralisation. Ceci s’observe par un taux d’erreur sur les données d’entraînement faible et un taux d’erreur sur les données de test élevé. Plus la différence entre les taux d’erreur ($T_{\text{train}} < T_{\text{test}}$) est grande plus on est en overfitting. Au contraire, un phénomène de sous-apprentissage ou « underfitting » peut aussi être observé. Le modèle est beaucoup trop généraliste et ne fait plus attention aux particularités des données d’entraînement. Ainsi le meilleur modèle est celui dont le taux d’erreur est le plus faible et le plus semblable possible entre les données d’entraînement et de test. Il n’y a alors ni overfitting ni underfitting. Par exemple, un estimateur kNN est généralement en overfitting avec un hyperparamètre k égal à un. Le modèle est très sensible aux spécificités des données d’entraînement. Alors qu’un nombre de voisins élevé aura tendance à classer une nouvelle entrée dans la classe le plus représentée pendant la phase d’entraînement.

Afin d’améliorer les performances des estimateurs, il est nécessaire d’optimiser les hyperparamètres. La méthode consiste à les modifier les uns après les autres, puis d’effectuer la phase d’entraînement et d’enregistrer le score obtenu sur les données de test. Les hyperparamètres qui ont eu le meilleur score sont retenus. Mais attention, cette méthode force les estimateurs à se spécifier sur les données de test. Elles ne pourront plus être utilisées pour l’évaluation finale de l’estimateur, car ces données auront déjà été vues. En effet, le modèle les a indirectement vues, car il est optimisé avec elle. Pour cette raison, une troisième section est prélevée dans le jeu d’entraînement, le jeu

de validation. Il permet de chercher les réglages du modèle qui donne les meilleures performances tout en gardant de côté le jeu de test. En résumé, pour comparer deux modèles avec des réglages différents, ils sont entraînés sur le jeu d'entraînement. Puis, l'estimateur avec le meilleur score sur le jeu de validation est maintenu. Finalement, ses performances futures sont évaluées à l'aide du jeu de test.

La démarche décrite précédemment ne garantit pas que la manière d'extraire le jeu de validation du jeu d'entraînement est optimale. Il est possible que le choix d'une autre portion pour l'échantillon de validation change l'estimateur retenu. Pour s'assurer la sélection du meilleur modèle, il faut procéder à une validation croisée ou « cross validation ». Elle consiste à entraîner et à valider le modèle sur plusieurs découpes possibles du jeu d'entraînement voir Figure 13. Par exemple, en le découpant en cinq parties, l'entraînement peut être fait sur les quatre premières et la validation sur la dernière partie. Puis l'opération est répétée pour toutes les quatre autres configurations possibles. La performance finale est la moyenne des cinq scores obtenus. Cela garantit de choisir le modèle qui a eu, en moyenne, les meilleures performances.

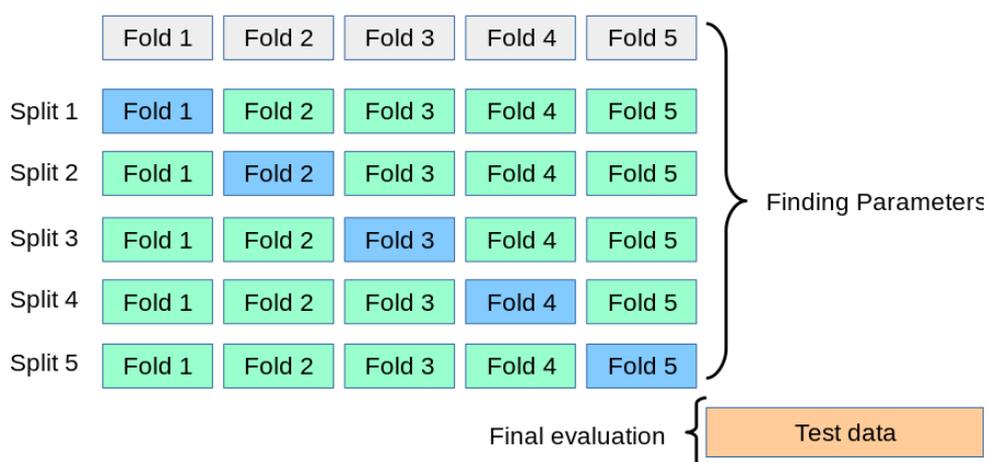


Figure 13: Principe de validation croisée avec 5 découpes du jeu d'entraînement initial (gris), les jeux de validation (bleu) et les jeux d'entraînement (vert)

Dans ce travail, les hyperparamètres des modèles évalués ont été réglés à l'aide de la première année des données dans les jeux d'entraînement. Les hyperparamètres retenus ainsi que la plage testée sont présentés dans le Tableau 3.

Estimateurs	Hyperparamètres	Paramètres à essayer	Paramètres retenus selon le type d'entrée		
			Alourdi / Normal		
			Injections	Charges	Productions
kNN	Nombre de voisins	1 à 200	41 / 4	51 / 11	171 / 31
SVC	Paramètre de régularisation C	0.01 à 600	9.6 / 502	1.6 / 7,6	0.6 / 0,01
NN	Nombre de Neurons dans une couche	50 à 150	115 / 120	105 / 110	100 / 100
RF	Nombre d'arbres	10 à 200	167 / 65	95 / 82	93 / 50

Tableau 3: Valeurs des hyperparamètres retenus après la validation croisée pour chaque modèle sur les deux scénarios

2.3.3 Matrice de confusion

Afin d'évaluer les performances des algorithmes de classification, le taux d'erreur n'est pas suffisant. Les situations sont classifiées selon deux classes : saines ou risquées. Deux types d'erreurs sont alors possibles : les situations faussement estimées en saines, appelées fausses alarmes, et les situations risquées non classées comme telles et estimées à tort comme saines, appelées non-détections. En matière de sécurité des réseaux, la non-détection est plus grave qu'une fausse alarme. Ainsi, la matrice de confusion permet de les distinguer. Chaque colonne de la matrice représente la classe estimée et chaque ligne représente la classe réelle.

Dans la Figure 14, sur les 1'380 situations saines réelles, l'estimateur a prédit que 180 étaient risquées (fausse alarme), et sur les 67 situations risquées, il a prédit que 12 étaient saines (non-détection). Toutes les prédictions correctes sont situées dans la diagonale du tableau (flèche en bleu), il est donc facile d'inspecter visuellement le tableau pour les erreurs de prédiction, car elles seront représentées par des valeurs en dehors de celle-ci. La somme de chaque ligne donne le nombre de cas dans les classes respectives.

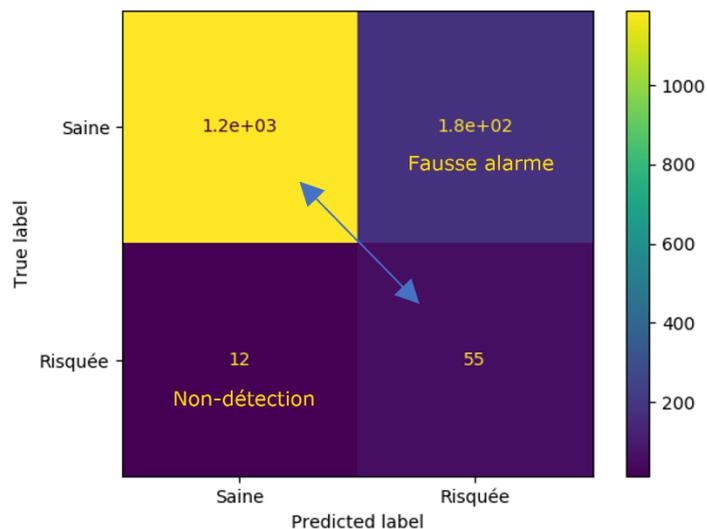


Figure 14 : Illustration d'une matrice de confusion

Le premier critère d'évaluation est un nombre de non-détections égale à zéro. Une situation dangereuse doit être signalée dans tous les cas. Le deuxième critère est un ratio de fausse alarme Q_{fa} le plus faible possible. Ce ratio, défini par l'équation (7), est calculé avec le nombre de fausses alarmes E_{FA} et le nombre total de situations estimées sur la période de test M_{test} .

$$Q_{fa} = \frac{E_{FA}}{M_{test}} \quad (7)$$

3 Résultats

3.1 Load flow implémenté vs MATPOWER

Les tensions sur les bus et les angles obtenus à l’aide de l’algorithme implémenté et à l’aide de MATPOWER sont identiques jusqu’à huit chiffres après la virgule (cas de test IEEE 57–bus). Ces résultats sont obtenus sous certaines conditions :

- La différence entre la puissance active consommée et la puissance active produite doit être faible. Il en va de même pour la puissance réactive.
- Les valeurs initiales de Newton–Raphson doivent être semblables.
- Les capacités « shunt » sur les nœuds ainsi que les réglages en charge ou « tap changing » des transformateurs doivent être négligés.

Au préalable, un autre test a été effectué sur le cas IEEE 30–bus. Sous les mêmes conditions, les résultats sont identiques jusqu’à 7 chiffres après la virgule. L’algorithme implémenté permet résoudre les équations de load flow avec succès.

3.2 Performances des estimateurs

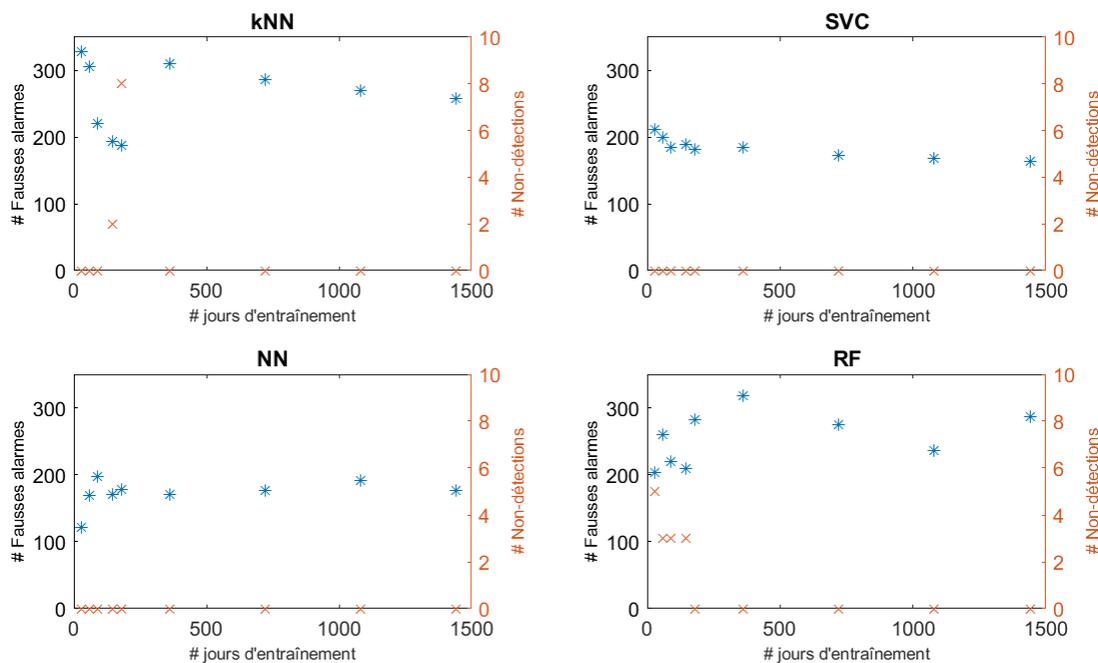


Figure 15 : Fausses alarmes (bleu) et non-détections (rouge) en fonction du nombre de jours d’entraînement pour les différents modèles sur le scénario alourdi, les caractéristiques utilisées sont les injections

La phase d’optimisation des hyperparamètres de chaque estimateur a été réalisée sur la première année du jeu d’entraînement correspondant. Les hyperparamètres sont ensuite constants pour tous les tests. Ensuite, les modèles ont été entraînés sur une période croissante de 28 jours à environ 4 ans. Finalement, chaque estimateur est évalué sur le jeu de test du scénario pour chaque période

d'entraînement. Sur les Figure 15 et Figure 16, les performances sont représentées pour le scénario alourdi. Les fausses alarmes et les non-détections issues de la matrice de confusion sont représentées en fonction du nombre de jours du jeu d'entraînement. Pour rappel, le scénario alourdi signifie que le taux de situation risquée est élevé pour les jeux d'entraînement et de test.

Le premier critère de performance est de signaler toutes les situations à risque. Ce critère est éliminatoire et se traduit par zéro non-détection. Tous les modèles avec les injections en caractéristique d'entrée (Figure 15) respectent ce critère à partir de 360 jours d'entraînement.

Le deuxième critère de performance est un nombre de fausses alarmes faible. Les meilleurs modèles sont alors le SVC et le NN pour les injections en entrées (Figure 15). Respectivement, ils ont aux mieux 164 et 170 fausses alarmes. Le quotient de fausses alarmes (voir équation (7)) est alors de moins de 12 %. Pour le modèle NN, les fausses alarmes se stabilisent dans cette zone alors que pour le modèle SVC, une légère diminution se profile avec l'augmentation des jours d'entraînement. Il est alors probable que des données supplémentaires continuent à l'améliorer.

De manière générale, l'augmentation du nombre de jours d'entraînement accroît les performances. Le modèle a plus de chance de reconnaître les particularités et de s'améliorer. Néanmoins, davantage de données d'entraînement n'améliorent pas infiniment le modèle. C'est une information utile lorsque la récolte de données est coûteuse.

Les résultats obtenus avec les caractéristiques de charges (Annexe 2) sont aussi performants qu'avec les injections. En effet, l'estimation avec les charges est faite avec 89 % de l'état du réseau connu (50 charges pour 56 injections au total). Des estimations fiables peuvent être obtenues sans connaître l'état du réseau complet.

La Figure 16 représente les résultats obtenus avec les productions en entrées. La sécurité du réseau est évaluée avec seulement 12 % de l'état du réseau connu. Les performances sont inférieures à celles obtenues à l'aide des caractéristiques d'injections. Le premier critère n'est jamais respecté. Néanmoins, le modèle NN se distingue avec peu de situations risquées non signalées. Avec 1 440 jours d'entraînement, il y a 13 non-détections sur 363 situations risquées soit 3,5 %. Le ratio de fausse alarme s'élève alors à 16,5 %.

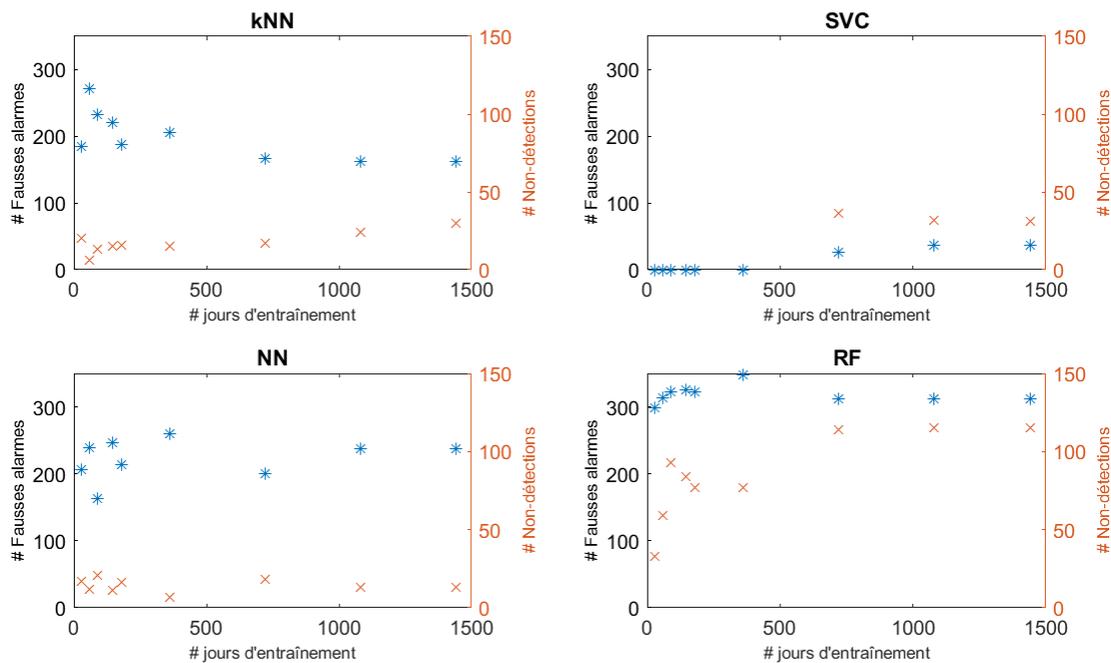


Figure 16: Fausses alarmes (bleu) et non-détections (rouge) en fonction du nombre de jours d'entraînement pour les différents modèles sur le scénario alourdi, les caractéristiques utilisées sont les productions

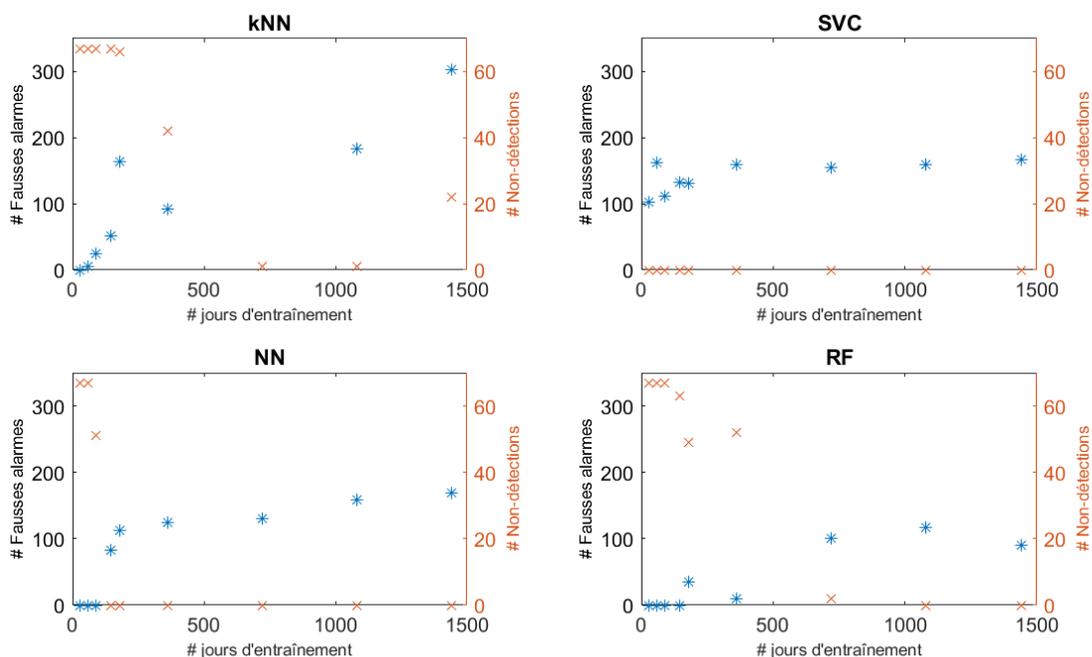


Figure 17: Fausses alarmes (bleu) et non-détections (rouge) en fonction du nombre de jours d'entraînement pour les différents modèles sur le scénario normal, les caractéristiques utilisées sont les injections

La Figure 17 représente les performances des estimateurs sur le scénario normal. Ce dernier scénario signifie que le taux de situation risquée est d'environ 5 % pour les jeux d'entraînement et de

test. Il est plus représentatif des réseaux existants que le scénario alourdi. Comme pour le scénario alourdi (Figure 16), le premier critère est respecté pour les modèles SVC, NN et RF. Le deuxième critère met l'estimateur RF en évidence. Avec 1 440 jours d'entraînement, il détecte toutes les situations à risque et émet 90 fausses alarmes. Le ratio de fausses alarmes Q_{fa} est de 6 %. Cette performance est meilleure que le modèle NN sur le scénario alourdi ($Q_{fa}=12\%$, Figure 16).

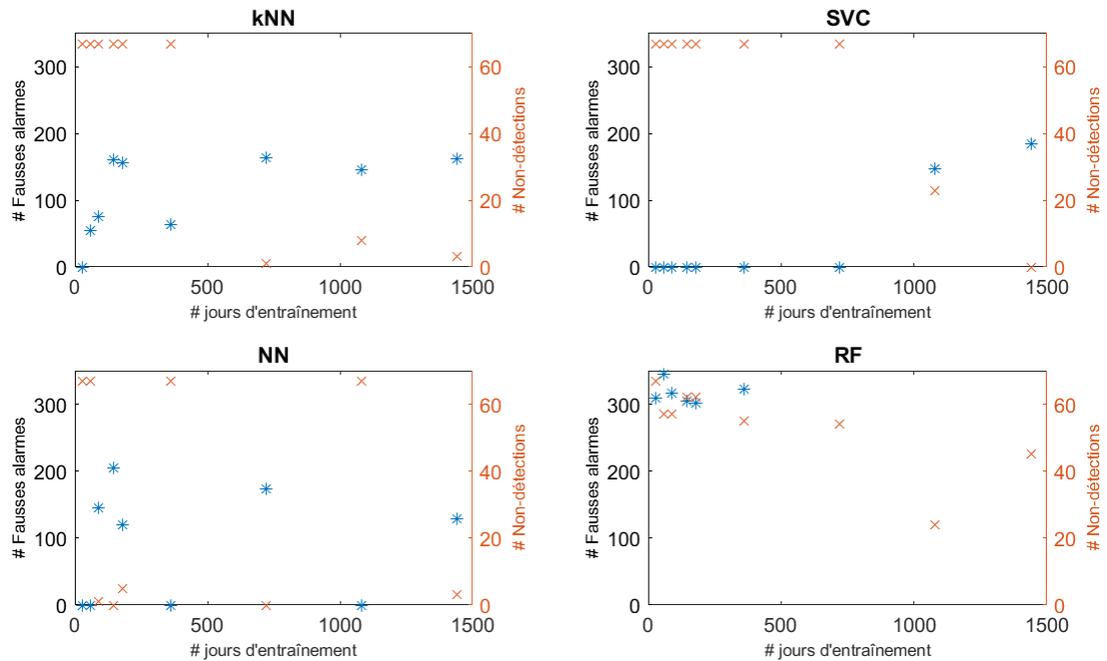


Figure 18: Fausses alarmes (bleu) et non-détections (rouge) en fonction du nombre de jours d'entraînement pour les différents modèles sur le scénario normal, les caractéristiques utilisées sont les productions

La Figure 18 représente les performances des estimateurs sur le scénario normal avec les productions en entrées. La faible fréquence des situations à risques impacte les résultats. Cependant, les modèles SVC et NN se distinguent une nouvelle fois (comme sur la Figure 16). Pour 1'440 jours d'entraînement le SVC est meilleur, car il détecte toutes les situations à risque avec 200 fausses alarmes.

4 Conclusions

Un algorithme de load flow basé sur la méthode de Newton–Raphson a été implémenté. Il permet de déterminer les flux de puissance de chaque ligne afin d'évaluer la sécurité en N . Une comparaison avec les résultats obtenus par MATPOWER a permis de valider l'algorithme. Ensuite, des jeux de données ont été générés afin d'entraîner et d'évaluer quatre estimateurs d'apprentissage automatique. La phase d'entraînement comporte des données aléatoires contrairement à la phase de test qui elle est constituée par une série temporelle basée sur la consommation globale de la Suisse. Les modèles ont pour but de classifier pour chaque situation le niveau de sécurité. Soit une situation est jugée saine, soit elle est à risque. Les performances sont examinées selon les deux types d'erreurs possibles. Les non-détections et les fausses alarmes. Selon ces critères, deux modèles se distinguent, le réseau de neurones et l'estimateur à vecteur de support. Ils détectent toutes les situations à risque et émettent 12 % de fausses alarmes sur la période de test sur un scénario congestionné. Le modèle de forêt d'arbres décisionnels a obtenu les meilleures performances sur le scénario normal avec 0 non-détection et 6 % de fausses alarmes. Ces résultats montrent que les estimateurs de machine learning peuvent apporter une aide supplémentaire dans la gestion des réseaux électriques.

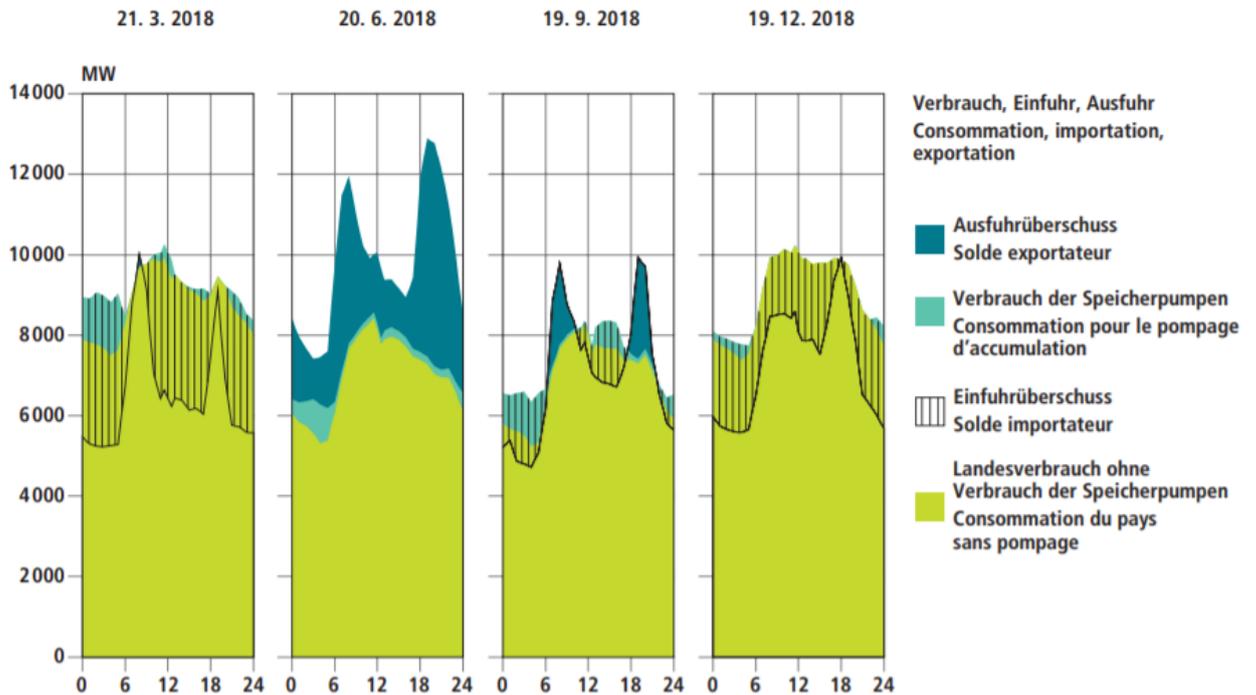
Cette étude a des perspectives pour de multiples points. Le modèle à réseau de neurones (NN) a obtenu de bonnes performances. Une optimisation plus avancée des hyperparamètres pourrait permettre d'améliorer ses résultats. Par exemple, la prochaine étape serait d'évaluer la sécurité en $N-1$ en non plus N . Puis, il serait intéressant d'ajouter une nouvelle catégorie de classification. Il s'agirait alors d'indiquer 3 niveaux de sécurité : « vert » pour un réseau sûr à $N-1$, « Orange » pour une sécurité $N-1$ qui entraînerait une exploitation d'une ligne à plus de 100 % et « Rouge » pour une sécurité $N-1$ qui entraînerait une charge à plus de 110 %. Selon les résultats, une estimation quantitative est envisageable. Elle indiquerait le taux de surcharge de ligne qui est en cause lorsque la sécurité $N-1$ n'est pas respectée. Le but ici est toujours d'évaluer la sécurité du réseau plus vite que ne le font les algorithmes de load flow. Il faudrait quantifier le gain de temps réalisé pour l'étude d'un réseau à 400 nœuds (~ le nombre de nœuds du réseau suisse).

Une autre possibilité de développer davantage ce travail serait la prédiction de la sécurité du réseau dans les heures voire les jours suivant la prédiction. Par exemple, l'estimateur prédit les événements extrêmes dans les 6 heures qui suivent. Puis, des connaissances spécifiques du réseau pourraient être ajoutées aux estimateurs afin d'améliorer leurs performances. Et finalement, les modèles devraient être évalués sur des données historiques d'exploitation de réseau afin de garantir leur efficacité.

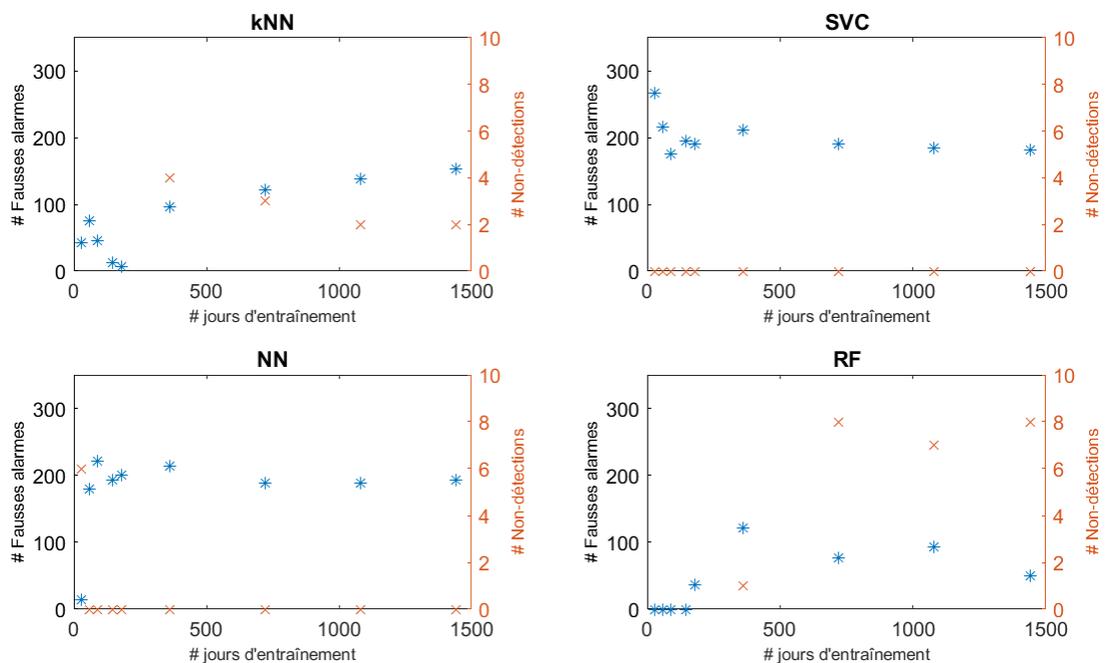
Références

- [1] Pronovo SA, «Facts SRI,» [En ligne]. Available: <https://pronovo.ch/fr/sri-facts-sri/>. [Accès le 06 06 2020].
- [2] OFEN, «Statistique suisse de l'électricité,» 2018.
- [3] Swissgrid SA, «Niveau de réseau,» [En ligne]. Available: <https://www.swissgrid.ch/fr/home/operation/power-grid/grid-levels.html>. [Accès le 17 juin 2020].
- [4] A. Berizzi, «The Italian 2003 blackout,» IEEE Power Engineering Society General Meeting, Denver, 2004.
- [5] Hart et Tinney, «Power Flow Solution by Newton's Method,» chez *IEEE Transactions on power apparatus and systems*, 1967, p. 1449.
- [6] J. D. Glover, M. S. Sarma et T. J. Overbye, «The power-flow problem,» chez *Power System Analysis and Design*, 5 éd., 2012, pp. 325-331.
- [7] L. Soder et M. Ghandhari, «Newton-Raphson method,» chez *Static Analysis of Power Systems*, Royal Institute of Technology, 2010, pp. 87-91.
- [8] J. D. Glover, M. S. Sarma et T. J. Overbye, «ITERATIVE SOLUTIONS TO NONLINEAR ALGEBRAIC EQUATIONS: NEWTON-RAPHSON,» chez *Power System Analysis and Design*, 5 éd., 2012, pp. 321-325.
- [9] J. D. Glover, M. S. Sarma et T. J. Overbye, «POWER-FLOW SOLUTION BY NEWTON-RAPHSON,» chez *Power System Analysis and Design*, 5 éd., 2012, pp. 334-343.
- [10] R. D. Zimmerman, C. E. Murillo-Sánchez et R. J. Thomas, «MATPOWER: Steady-State Operations, Planning, and Analysis Tools for Power Systems Research and Education,» chez *IEEE Transactions on Power Systems*, 2011, pp. 12-19.
- [11] UW Power Systems Test Case Archive, «IEEE 57-Bus Power Flow Test Case,» [En ligne]. Available: <https://eGRIDdata.org/dataset/ieee-57-bus-power-flow-test-case>. [Accès le 20 06 2020].
- [12] Swissgrid SA, «Données de réseau,» [En ligne]. Available: <https://www.swissgrid.ch/fr/home/operation/grid-data.html>. [Accès le 29 juin 2020].
- [13] Pedregosa et al., «Scikit-learn: Machine Learning in Python,» *Journal of Machine Learning Research*, vol. 12, pp. 2825-2830, 2011.
- [14] scikit-learn, «Supervised learning,» [En ligne]. Available: https://scikit-learn.org/stable/supervised_learning.html. [Accès le 13 juillet 2020].
- [15] D. Guillaume, «Prédiction d'événements extrêmes sur les réseaux électriques à l'aide de machine learning,» 2020. [En ligne]. Available: <https://github.com/Guillaume-cpu/machineLearning-reseauElectrique>.

Annexes



Annexe 1 : Diagramme de la consommation pour le 3e mercredi du mois issue de la figure 17 de la statistique de l'électricité 2018 [2]



Annexe 2: Fausses alarmes (bleu) et non-détections (rouge) en fonction du nombre de jours d'entraînement pour les différents modèles sur le scénario alourdi, les caractéristiques utilisées sont les charges