Etude de l'intégration de la prévision de la ressource solaire dans un système de gestion énergétique hybride (solaire / combustible fossile) pour des sites industriels isolés sans stockage

Fuqiang Zhuang

PSL UNIVERSITÉ PARIS

THÈSE DE DOCTORAT
DE L'UNIVERSITÉ PSL

Préparée à MINES Paris

Etude de l'intégration de la prévision de la ressource solaire dans un système de gestion énergétique hybride (solaire / combustible fossile) pour des sites industriels isolés sans stockage

Soutenue par

Fuqiang ZHUANG

Le 13 décembre 2022

Ecole doctorale n° 621

ISMME (Ingénierie des Systèmes, Matériaux, Mécanique, Energétique)

Spécialité

Énergétique et Génie des Procédés

Composition du jury :

Prof. Richard, PEREZ University at Albany - SUNY	*Rapporteur*
Prof. Cristina, CORNARO Université de Rome	*Rapporteure*
Prof. Philippe, LAURET Université de la Réunion	*Président du jury* *Examinateur*
Dr. Elke, LORENZ Fraunhofer ISE	*Examinatrice*
Dr. Sylvain, CROS Ecole Polytechnique	*Examinateur*
Dr. Rodrigo, AMARO E SILVA MINES Paris, PSL	*Examinateur*
Dr. Yves-Marie, SAINT-DRENAN MINES Paris, PSL	*Examinateur*
Prof. Philippe, BLANC MINES Paris, PSL	*Directeur de thèse*

MINES PARIS | PSL

REMERCIEMENT

La récolte est enfin venue après trois années d'aventure, marquant la fin de cette thèse. Tout au long de ce parcours, j'ai connu des moments confus, fatigants, stressants, réussis, mais surtout des moments joyeux, grâce au soutien et à l'encouragement des personnes qui m'ont aidé. Cette expérience m'a permis de devenir plus sceptique et curieux professionnellement, mais aussi plus objective et optimiste personnellement.

Je tiens tout d'abord à exprimer ma gratitude envers mes encadrants : Philippe Blanc, mon directeur de thèse, pour sa bienveillance, sa patience et ses idées brillantes sur la recherche, évidement il y a aussi ses blagues inoubliables; Yves-Marie Saint-Drenan, mon maître de thèse, pour sa patience, sa disponibilité, et ses conseils en recherche et rédaction; Benoît Gschwind, mon maître de thèse, pour son enthousiasme et son aide en programmation; Rodrigo Amaro e Silva, mon maître de thèse, pour sa réactivité et ses conseils utiles sur la présentation. Je voudrais également remercier Hadrien Verbois, qui n'était pas officiellement encadrant, mais qui a fourni un travail équivalent, pour sa bienveillance, son aide dans la lecture et ses suggestions sur mon manuscrit. Grâce à vous tous, j'ai pu éviter de nombreuses erreurs et vous remercie d'avoir pris le temps de m'accompagner dans les détails de ma thèse.

Je tiens également à remercier tous les membres du jury pour leur temps et leurs efforts consacrés à la lecture de ma thèse. Les commentaires constructifs des rapporteurs et les suggestions intéressantes des autres membres du jury m'ont permis d'améliorer la qualité de mon manuscrit.

Je suis également reconnaissant envers l'équipe d'encadrement industriel chez SPIE, qui m'a accompagné tout au long de la réalisation de cette thèse. Merci à mon maître de thèse industriel, Philippe Salvado, pour ses explications et conseils avisés sur les sujets liés au système d'énergie, en particulier pour sa patience et sa disponibilité. Merci à Pierre Saint-Girons, qui a proposé ma candidature pour cette thèse CIFRE et qui a toujours fait preuve de patience et de bienveillance. Merci à Jérôme Viscaino pour ses explications techniques et ses expériences qui m'ont permis de me connecter aux réalités opérationnelles. Je tiens à remercier tous mes collègues de SPIE qui m'ont aidé et encouragé dans la réalisation de cette thèse, tels que Laurent, Laure, Antoine, Bashkim, Stéphane, Bianca, etc. Je tiens à remercier particulièrement l'initiateur de ce projet de thèse, Jean-Philippe Mangione, pour sa bienveillance, sa patience ainsi que ses précieux conseils sur la vie professionnelle et personnelle. Sans lui, je n'aurais pas pu effectuer mon alternance d'ingénieur ni réaliser cette thèse passionnante.

Je remercie Thomas Carrière pour sa gentillesse et sa patience en me donnant des cours de prévision et répondant à mes questions. Je tiens également à remercier chaleureusement l'équipe de Solaïs - Sébastien, Christophe et Daniel - pour leur bienveillance et leur réactivité. Votre contribution a été précieuse pour mener à bien ce projet de thèse.

Le Centre de recherche O.I.E de l'école des Mines Paris m'a accueilli avec une grande amitié et j'en remercie tous ses membres. Je remercie Thierry en particulier, qui m'a prodigué des conseils éclairés tout au long de mon parcours de thèse, notamment en me rappelant que cette épreuve est un marathon, et qu'il est essentiel de respecter son propre rythme et d'avancer petit à petit. Je suis

également reconnaissant envers Lionel, qui a toujours pris le temps de m'expliquer des nouvelles choses et de m'aider à améliorer mon niveau de français. Mes sincères remerciements vont également à Paula, Isabelle, Mireille, Thibaut, Mélanie et Raphaël pour leur soutien scientifique et leur bonne humeur, ainsi qu'à Sandra pour son assistance administrative. Je tiens également à remercier les autres membres d'O.I.E., d'avoir fait de mon séjour parmi eux une expérience aussi enrichissante et mémorable. Les moments de convivialité partagés avec Romain, Mélodie, Benoit T, Ghada, Joris, Xuemei, Joanna, Mathilde, Mehefa, Alejandra, Jeremie, Sara, Gabriel, Lu, Vadim, etc. m'ont permis de vivre des moments de détente et de partage inoubliables.

Enfin, je ne saurais oublier de remercier ma famille, mes parents, mes frères et sœurs, pour leur confiance et leur soutien sans faille tout au long de mon parcours. Et je tiens tout particulièrement à remercier ma copine Mingkong, qui m'a apporté un soutien inestimable en relisant mes travaux tard le soir, en écoutant mes présentations à presque endorme, en me donnant la force nécessaire dans les moments de découragement, et en illuminant ma vie de sa présence et de sa joie de vivre au quotidien.

RESUME ETENDU

Le contexte de cette thèse repose sur le constat que près de trois quarts des émissions de gaz à effet de serre dans le monde sont liées à la production d'énergie, comme montré par la Figure 1. En se penchant de plus près sur la consommation de ressources énergétiques, les ressources fossiles continuent de jouer un rôle crucial. Cette réalité nous rappelle notamment qu'il existe encore de nombreux sites industriels, en particulier des sites isolés avec une forte demande d'énergie tels que celui de notre cas d'étude : une mine d'extraction d'or isolée qui utilisent encore des énergies fossiles comme ressources principales.

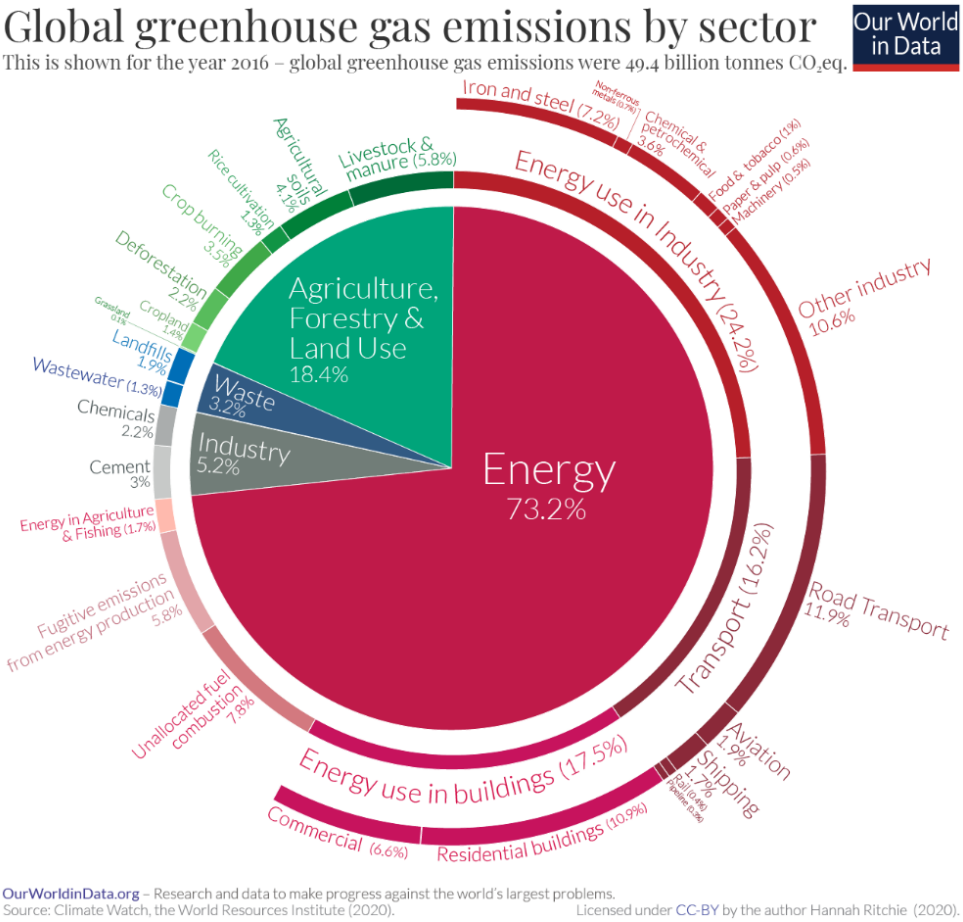


Figure 1 : émission de Gaz à Effet de Serre par secteur (*Climate Watch,* 2020).

Plus précisément, le site de notre cas d'étude se trouve au Mali, en Afrique de l'Ouest, et est alimenté par une centrale électrique de 60 MW, composée de 27 groupes électrogènes diesel présentant différentes caractéristiques. Ce site d'extraction d'or a été construit et mis en service il y a plusieurs années.

Avec la prise de conscience croissante de l'importance sur la protection de l'environnement, ces industries sont soumises à des contraintes de plus en plus importantes, telles que le prix du fioul et les taxes élevées, ainsi que des politiques plus strictes en matière d'émissions de gaz à effet de serre.

L'isolement de ce site industriel rend de plus difficile l'approvisionnement du fioul, ce qui entraîne une augmentation des prix.

Afin de répondre à l'augmentation des besoins de puissance, de réduire l'impact environnemental et les coûts de production d'énergie, l'opérateur du site recherche des alternatives aux énergies fossiles. L'énergie photovoltaïque (PV) est une option intéressante en raison de son empreinte carbone réduite. En effet, produire 1 kWh d'électricité avec du diesel émet environ 20 fois plus de CO2 équivalent qu'avec de l'énergie PV, selon des études d'Analyse du Cycle de Vie (ACV). L'intégration d'énergie PV dans ce système d'énergie joue un rôle important sur la réduction de l'impact environnemental. L'opérateur du site a donc décidé d'installer une centrale PV afin de construire un système d'énergie hybride sans stockage, comme montré par la Figure 2.

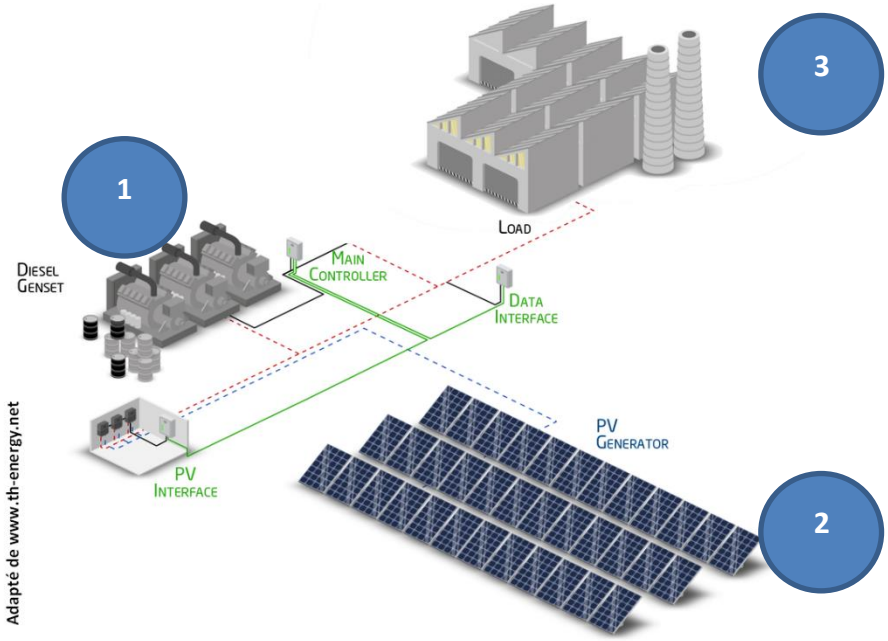


Figure 2 : schéma illustratif du système d'énergie hybride isolé, sans stockage. La zone (1) représente les générateurs diesel, la zone (2) représente le champ photovoltaïque et la zone (3) représente la partie industrielle consommatrice de l'énergie provenant du système hybride (1) + (2).

La mise en place d'un système hybride d'énergie est une solution prometteuse, cependant, elle est confrontée à des contraintes liées à trois types de variabilités :

- La panne potentielle de différents composants du système d'énergie ;
- La variabilité de la demande en énergie ;
- La variabilité solaire causée par les changements saisonniers, diurnes, ainsi que les conditions météorologiques telles que les tempêtes de sable, la vapeur d'eau dans l'air, ainsi que la couverture nuageuse.

Afin de pallier ces variabilités, des solutions de stockage telles que des batteries, des stations de transfert d'énergie par pompage (STEP), des volants d'inertie et des stations d'air comprimé sont proposées. Cependant, ces solutions ont des inconvénients tels que des coûts d'investissement élevés, des besoins en matières premières, des impacts environnementaux, et des contraintes géographiques.

Dans cette thèse, la solution envisagée est l'utilisation d'un système de gestion de puissance, appelé Power Management System (PMS), combiné à une planification –ou *dispatching*– intelligent des

groupes électrogènes diesel et à une réserve tournante dimensionnée avec l'aide de la prévision solaire probabiliste.

Le *dispatching* intelligent repose sur la notion d'*Unit Commitment*, une question mathématique classique qui vise à trouver, à chaque instant, la meilleure combinaison de moyens de production d'électricité avec un coût de production minimal ou un bénéfice maximal en cas de vente d'électricité, comme montré au Figure 3. Ce *dispatching* est basé sur la prévision solaire probabiliste, utilisant comme élément principal la médiane, qui représente le percentile de 50% de la prévision pour le *dispatching*.

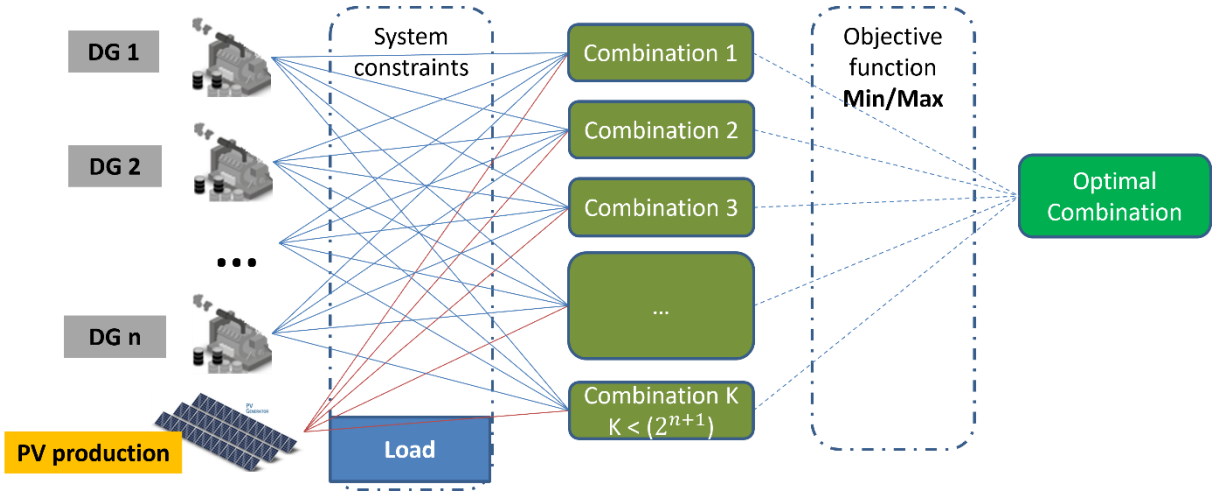


Figure 3 : diagramme d'explication du principe de l'*Unit Commitment* (DG : *Diesel Generator*).

Toutefois, dans la réalité, il est impossible d'avoir une prévision parfaite, ce qui peut entraîner des variations dans la production. Pour faire face à cette variabilité, nous utilisons la réserve tournante, *spinning reserve* (SR), des groupes électrogènes diesel pouvant servir à compenser toute ou partie de la variabilité résiduelle mal anticipée. Cette réserve est divisée en deux parties distinctes : la réserve tournante positive et la réserve tournante négative. La réserve tournante positive (SR+) est la capacité de réserve des groupes électrogènes correspondant à la part posititive de la différence entre la puissance nominale et la puissance effectivement produite. La réserve tournante négative (SR-), quant à elle, correspond à la différence positive entre sa puissance courante de sortie et sa capacité minimale de la plage de fonctionnement, permet de ralentir les groupes diesel en produisant moins de puissance.

Comme le montre la Figure 4, en cas de sous-production, c'est-à-dire dans le cas d'une prévision inférieure à la production effective, nous pouvons utiliser la réserve tournante positive pour produire plus de puissance et maintenir l'équilibre entre la production et la demande. Si cela n'est pas suffisant, il faut alors envisager le délestage de la charge, qui est une solution pouvant provoquer des dégâts, donc très coûteux et fortement déconseillée. En revanche, en cas de surproduction, nous pouvons ralentir les groupes diesel pour produire moins de puissance et si nécessaire, procéder à l'écrêtage de la production PV via une commande aux onduleurs.

En comparant ces deux situations de variation, nous constatons que la surproduction est plus facile à gérer et a un impact moins important. Dans la simulation de cette thèse, nous considérons uniquement le cas de sous-production en dimensionnant la réserve tournante positive en fonction de l'incertitude sous la médiane de la prévision probabiliste.

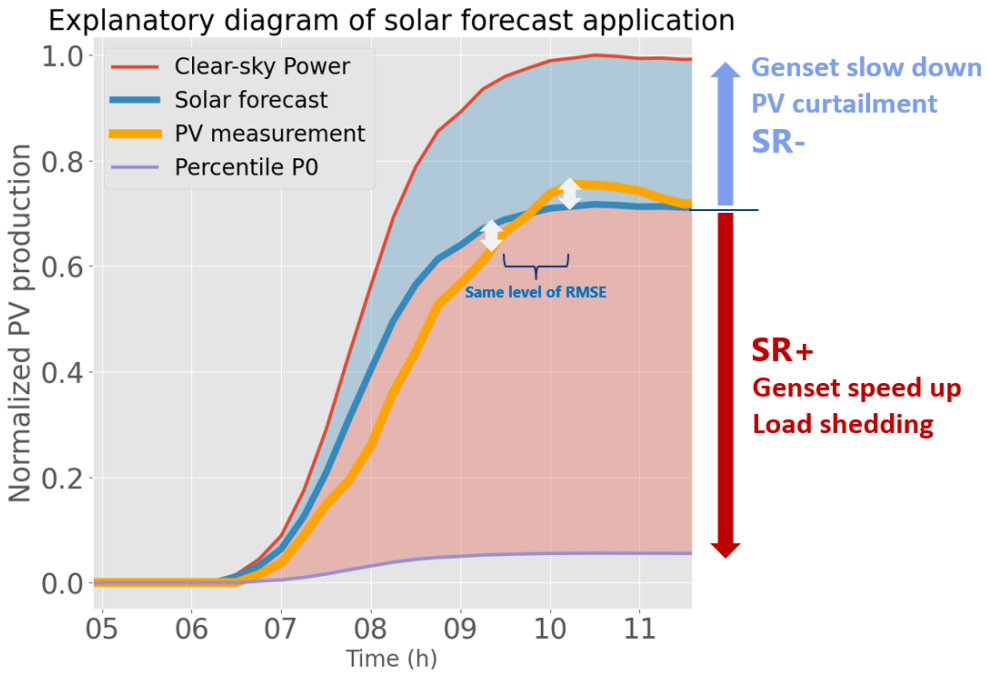


Figure 4 : schéma de la prévision probabiliste et du dimensionnement de la réserve tournante (La courbe rouge est la production en condition ciel clair, la courbe bleue est le niveau médian de la prévision probabiliste, la courbe violette est le niveau le plus bas de la prévision probabiliste, la courbe jaune est la production réelle de PV).

Compte tenu de ces éléments, la qualité de la prévision est cruciale pour le fonctionnement optimal du système hybride diesel / solaire. Dans le cadre de cette thèse, nous nous sommes penchés sur la relation entre la qualité de la prévision probabiliste et la performance du système hybride en tant que question scientifique principale. Pour mieux comprendre cette question, nous avons identifié trois sous-questions (SQ) :

- Quelle est la meilleure métrique pour évaluer la performance du système hybride ?
- Comment mettre en place une plateforme pour simuler le comportement du système hybride de manière correcte ?
- Quels sont les avantages d'utiliser un système hybride avec prévision probabiliste par rapport à un système sans énergie photovoltaïque ?

SQ1 : Quelle est la meilleure métrique pratique pour évaluer la performance du système hybride ?

Tout d'abord, pour répondre à la première sous-question, nous avons élaboré une méthode basée sur le coût du système d'énergie : le coût final du système est donc l'indicateur de la performance du système. Cette approche prend en compte les critères classiques tels que la stabilité du réseau, la performance économique du système, ainsi que la performance de la prévision. Le coût du système est divisé en trois parties :

1) Le coût des groupes électrogènes,
    a. ☒Coût d'investissement - CAPEX (*Capital Expenditure*)
    b. ☑Coût d'opération et maintenance - OPEX (*Operational Expenditure*)
2) Le coût du système PV
    a. ☑ Coût actualisé de l'énergie PV (Levelized cost of PV energy - *LCOE*)
3) Le coût de pénalité en cas de déséquilibre entre la production et la demande.
    a. ☑Coût d'excès d'énergie -> Ecrêtage du PV dans le cas d'une réserve tournante négative insuffisante
    b. ☑Coût du déficit d'énergie -> Délestage de la charge

Seule la partie OPEX (*Operational Expenditure*) est prise en compte pour le coût des génératrices, tandis que le CAPEX (*Capital Expenditure*) n'est pas considéré car les groupes électrogènes préexistaient après l'intégration de la partie PV. Au niveau du coût de la partie solaire on a choisi d'utiliser le coût conventionnel de PV à 0.04 $/kWh LCOE (*Levelized Cost Of Energy*) pour la simulation. Enfin, pour des actions pénalisantes comme l'écrêtage de PV en cas de surproduction et délestage de la charge en cas de déficit, il y a un coût associé s'ajoutant sur le coût du système à la fin.

Pour simplifier la simulation, certaines hypothèses ont été faites. Premièrement, les pannes potentielles des différents composants du système d'énergie, telles que les pannes de génératrices diesel, les dysfonctionnements des trackers et les pannes des onduleurs, n'ont pas été considérées. En effet, ces machines ont des états de santé différents et il est difficile de réévaluer la probabilité de panne de chacune d'entre elles. En outre, si l'on prenait en compte les pannes potentielles, le principe de sécurité de niveau "n-1" pour le dimensionnement de la réserve tournante devrait être appliqué, ce qui nécessiterait une capacité de réserve tournante importante pour couvrir toutes les variations solaires en l'absence de panne.

Deuxièmement, l'incertitude de la demande n'a pas été prise en compte dans la simulation, car le profil de demande réel utilisé dans l'étude est relativement stable.

Troisièmement, on considère que si l'équilibre entre la production et la demande est respecté en échelle de 15-min, l'équilibre transitoire est aussi respecté. Cette hypothèse, très forte, est faite car la simulation dynamique demande beaucoup plus de ressource de calcul et n'est pas l'objectif de cette thèse. Les équilibres à 15 min assurées par les outils de planification sont donc uniquement des conditions *sine qua non*.

Dans notre simulation, nous avons exploré trois méthodes distinctes de prévision. La première méthode, appelée CH-PeEn (*Complete History-Persistence Ensemble*), utilise toutes les données historiques disponibles sur plusieurs années pour générer des profils typiques de distribution de la production. Pour éviter les perturbations causées par les saisons de pluie qui induisent une forte

fluctuation du niveau d'éclairement, comme montré par la Figure 5, nous avons généré donc un profil typique pour chaque mois plutôt qu'un seul profil pour toute l'année.

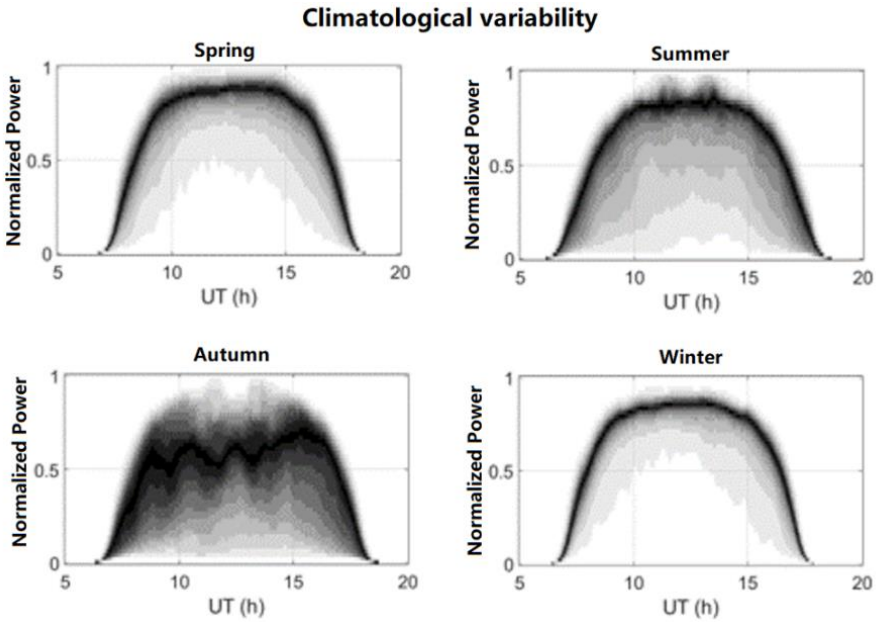


Figure 5 : profils typiques de la dispersion de production photovoltaïque (normalisée par la puissance maximale), pour différentes saisons du cas d'études au Mali (l'axe horizontal est le temps universel d'une journée, l'axe vertical la puissance normalisée de la production PV. Chaque courbe dans ces figures représente une journée de base des données, de plus la couleur est foncée, de plus la probabilité est élevée).

La deuxième méthode, appelée MCM (*Markov Chain Mixture*), comme montré avec la Figure 6, le principe de cette méthode consiste à classer la base de données en plusieurs classes, à générer une matrice de transition entre ces classes, et à déduire la distribution de la prévision pour l'état futur en fonction de l'état de prédiction et de cette matrice de transition. Pour une prévision plus longue, on peut générer des matrices de transition en ordre n en multipliant n fois la matrice de transition d'ordre 1.

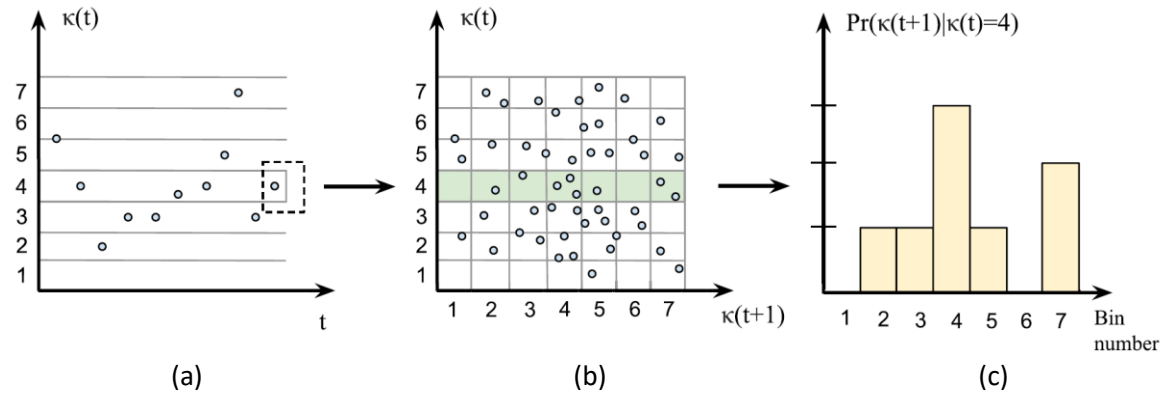


Figure 6 : diagramme d'explication sur la méthode de prévision MCM (Munkhammar et al., 2019). (a)Les éléments de la série temporelle d'indice de ciel clair sont catégorisés suviant plusieurs (ici, 7) classes équiréparties ; (b) : définition de la matrice de transition (ici de taille 7 x 7) entre différentes classes ; (c) : déduction de la distribution de la probabilité de transition pour le temps suivant à partir d'une classe donnée pour le temps présent.)

Enfin, la dernière méthode utilisée est la méthode de prévision parfaite, qui consiste à utiliser la série temporelle passée comme prédiction. A noter que cette méthode n'existe évidemment pas pour l'opération : elle est utilisée pour faire des analyses et simulations. Nous avons examiné ces trois

méthodes différentes pour évaluer leur efficacité respective dans la prévision de la production d'énergie solaire. La méthode de prévision parfaite permet d'évaluer la meilleure intégration possible du PV dans le système. A l'inverse, la méthode CH-PeEn permet d'évaluer le minimum que l'on puisse faire pour cette intégration. La méthode MCM est une méthode opérationnelle simple que l'on peut alors situer entre les deux méthodes « extrêmes » précédentes.

SQ2 : Comment correctement implémenter un simulateur numérique du comportement du système hybride ?

Pour représenter le comportement du système hybride, nous avons considéré six groupes électrogènes diesel avec différentes caractéristiques, d'une puissance totale d'environ 60 MWc. Ces six groupes diesel sont considérés au lieu des 27 réellement utilisés, car certains groupes sont identiques et sont allumés ou éteints simultanément : regrouper ces groupes diesel permet ainsi d'avoir une meilleure visibilité pour l'analyse.

Ces groupes fonctionnent au fioul lourd et ont une plage de fonctionnement définie entre 40% et 100% de leur puissance nominale. Les données d'entrée sont la production photovoltaïque réelle et la mesure de la demande réelle sur un an. La réserve tournante dans cette simulation est dimensionnée en fonction de l'incertitude de la prévision indexée sur la différence entre la médiane et un faible percentile et d'un tampon supplémentaire égal à 10% de la capacité de l'installation photovoltaïque. Ce tampon permet d'absorber les erreurs de prévision solaire, principalement au début et à la fin de la journée.

Afin de réaliser la simulation, nous avons mis en place un simulateur numérique en trois étapes. Tout d'abord, comme montré avec la Figure 7, la prévision de la production photovoltaïque (la partie jaune) et de la demande (en couleur bleue) est effectuée. Ensuite, la demande résiduelle (la partie noire) est comblée en utilisant la production diesel, qui correspond à la différence entre la demande et la production photovoltaïque.

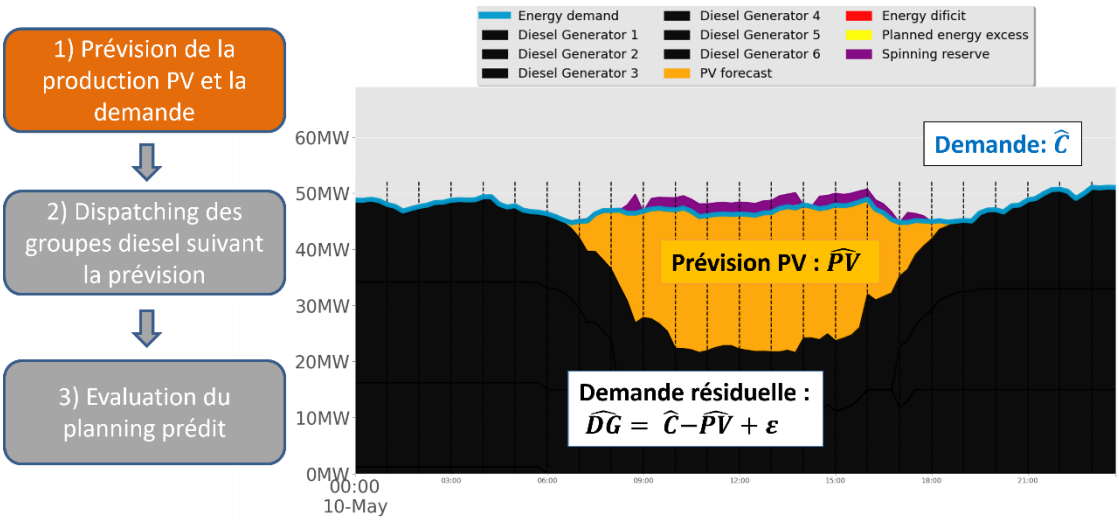


Figure 7 : étape 1 de la méthode de simulation – Prévision de la production PV et de la demande.

Dans la deuxième étape, nous procédons au *dispatching* des groupes diesel en utilisant les informations de prévision. Une fois que la demande et la production photovoltaïque ont été prévues, comme montré Figure 8, l'optimiseur génère une plannificationde *dispatch* des groupes diesel (la

partie grise) en utilisant le principe de l'*Unit Commitment*, tel qu'expliqué précédemment. Ce simulateur permet ainsi de gérer efficacement la distribution de l'énergie.

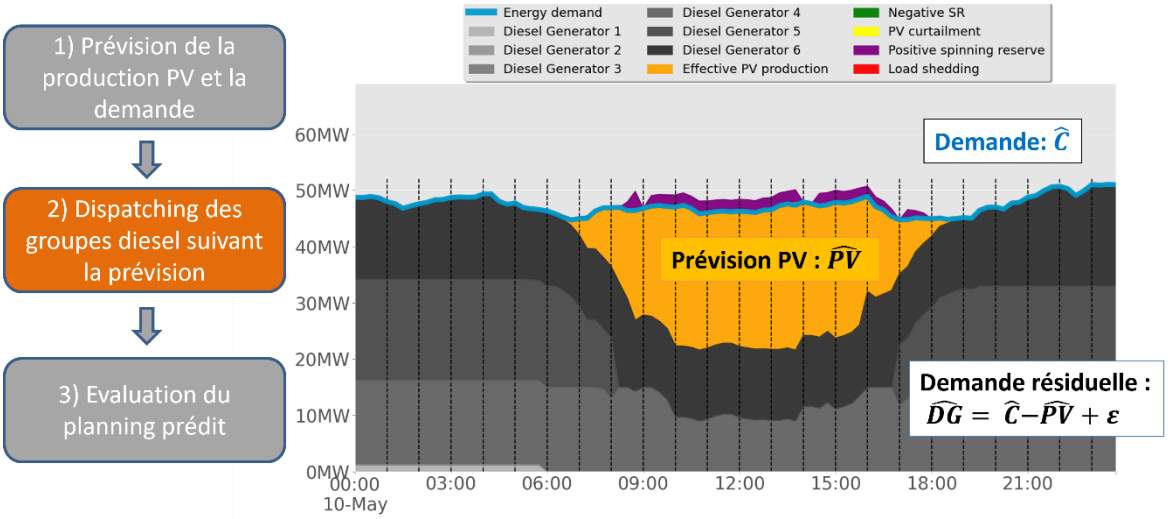


Figure 8 : étape 2 de la méthode de simulation – Optimisation du *dispatching* des générateurs diesel.

La dernière étape est d'évaluer ce planning basé sur la prévision avec les données effectives *a posteriori*. Comme montré par la Figure 9, la somme de la production doit être égale à la demande. En cas de déséquilibre entre la production et la demande, un coût de pénalité est ajouté au coût final du système par exemple pour de l'écrêtage du PV en cas de surproduction ou encore pour le délestage de la charge en cas de sous-production.

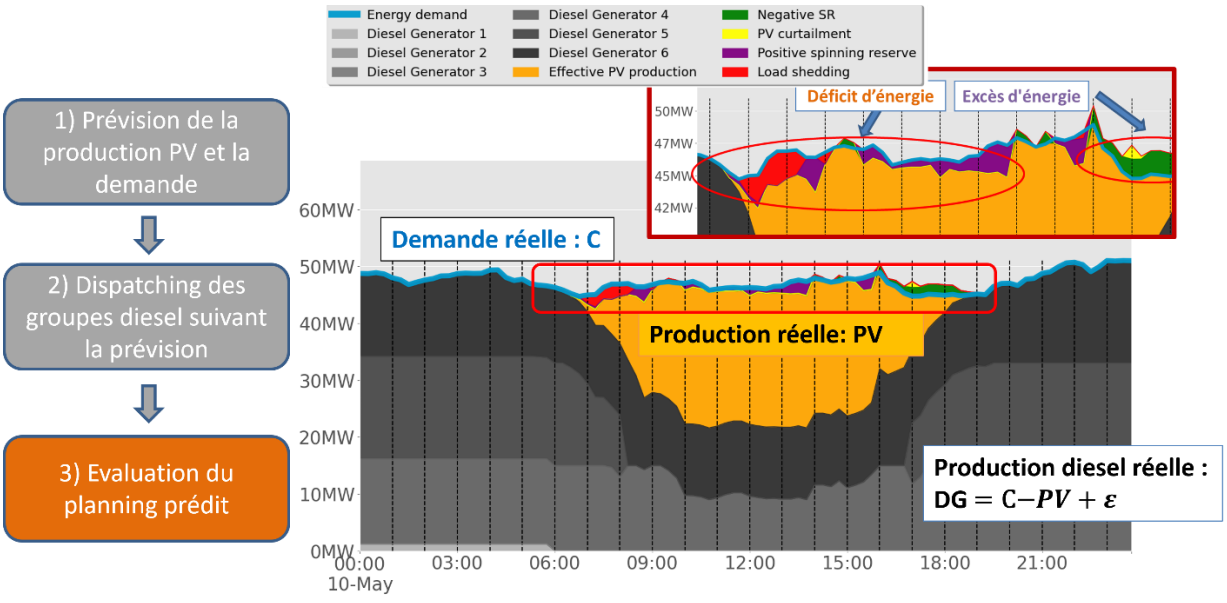


Figure 9 : étape 3 de la méthode de simulation – Evaluation du planning prédit.

Un horizon de prévision plus long permet de mieux organiser le dispatching des groupes diesel. Cependant, une prévision avec un horizon trop long perd un peu de sa performance en termes de précision. Cependant, le dispatching est généré en fonction de l'horizon de prévision. Si l'horizon de prévision est très court et que la prévision ne peut pas prédire des informations plus lointaines, les groupes diesel risquent de changer fréquemment leur état, ce qui réduirait significativement leur

durée de vie. Nous avons donc trouvé un compromis entre la qualité de la prévision et la durée de vie des groupes diesel en introduisant la fréquence de mise à jour. Ce paramètre permet d'utiliser un horizon de prévision plus long pour avoir plus d'informations et mieux organiser le *dispatching*, ainsi que de changer l'ordre plus rapidement pour renouveler la prévision avec une meilleure précision.

Le simulateur développé dans le cadre de la thèse est hautement configurable, comme montré par le Tableau 1 comprenant le taux de pénétration photovoltaïque en capacité, la méthode de prévision utilisée, la quantité de réserve tournante positive, l'horizon de prévision, la fréquence de mise à jour des ordres de dispatching des groupes électrogènes, ainsi que les types de journée. En dehors des paramètres listés, les informations plus concrètes du système d'énergie comme les caractéristiques des machines ou encore le prix du fioul sont tous configurables.

PV rate in capacity	Forecast method	Uncertainty range used for energy buffer sizing		Forecast Lead Time	Update Time	Type of day
0%	CH-Persistence Ensemble (CH-PeEn) forecast	P50-P0	With extra buffer = 10% PV capacity	1-hour	30-min	Variable (25-05-2021)
25%		P50-P5		2-hour	60-min	Clear (10-05-2021)
50%		P50-P10		3-hour		
75%	Markov chain based forecast	P50-P15		6-hour	120-min	Annual data (333 days)
100%		P50-P20		12-hour	180-min	
125%		P50-P25				
150%	Perfect forecast					
175%						
200%						

Tableau 1 : détails des paramètres configurables de la plateforme simulation.

SQ3 : Quel sera l'avantage d'utiliser un système hybride avec prévision probabiliste par rapport à un cas sans énergie PV ?

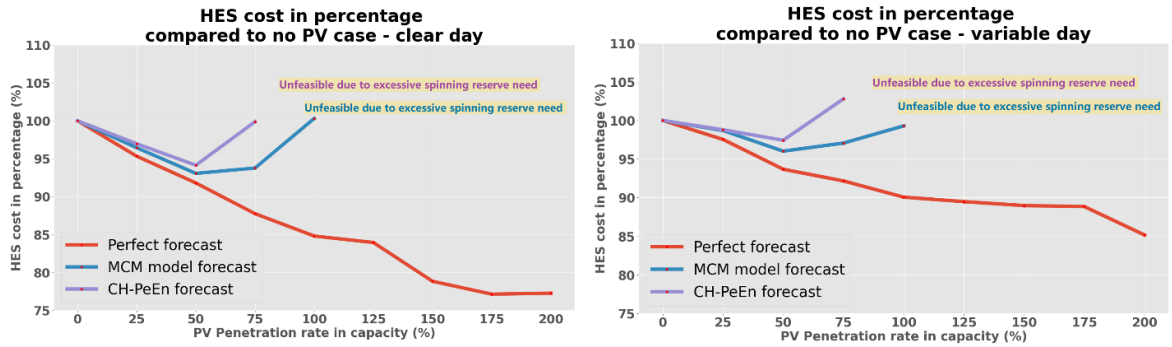
Afin de répondre à la question posée, nous avons utilisé notre simulateur. Un test de sensibilité des paramètres est nécessaire pour bien comprendre le résultat. Dans cette thèse, nous avons donc réalisé plusieurs scénarios avec différentes configurations.

Scénario 1 – L'influence des différents taux de pénétration PV sur la performance économique :

Le premier scénario de notre étude a consisté à évaluer l'impact de différents taux de pénétration photovoltaïque (PV) sur la performance économique du système. Nous avons fait varier le taux de pénétration PV et testé toutes les méthodes de prévision avec une quantité de réserve tournante dimensionnée à partir de la marge d'incertitude entre le percentile 0 (P0, valeur minimale) et la médiane P50 de la prévision, ainsi qu'un tampon supplémentaire. L'horizon de prévision a été fixé à une heure et le temps de mise à jour à 30 minutes. Nous avons également testé différents types de journées, telles que des journées de ciel clair et des journées variables par la présence d'une couverture nuageuse.

Nos résultats pour le ciel clair (cf. Figure 10a) ont montré que l'augmentation du taux de pénétration PV permettait de réduire le coût du système jusqu'à une limite minimale dépendant de la qualité de méthode de prévision. Ce minimum est atteint à cause de l'augmentation du coût du système due à l'écrêtage du PV à mesure que le taux de pénétration PV augmente. Pour les méthodes de prévision moins précises, leur marge d'incertitude devient très importante et entraîne une forte sollicitation de

la réserve tournante, ce qui rend la situation infaisable en termes d'optimisation quand le taux de pénétration PV devient trop important.



(a) (b)

Figure 10 : résultat de la simulation sur différents taux de pénétrations PV pour un jour clair (a) et variable (b) (HES : Hybrid Energy System) (scenario 1).

Pour la journée variable (cf. Figure 10b), le schéma général est similaire à celui de la journée avec ciel clair, mais le gain économique est moins élevé car il y a moins de ressources solaires. En conséquence, le taux de pénétration PV peut être augmenté davantage en raison de cette limitation. Sur ces résultats, trois conclusions peuvent être tirées :

- L'intégration de l'énergie PV dans un système d'énergie peut réduire le LCOE (Levelized cost of energy) avant la limite,
- Un taux de pénétration PV élevé n'est pas forcément avantageux au niveau du gain économique,
- Le gain économique est fortement dépendant sur la qualité de la prévision.

Scénario 2 – L'influence des différentes quantités de la réserve tournante sur la performance économique.

Le scénario 2 de notre étude vise à évaluer l'impact de la quantité de réserve tournante positive sur la performance du système. Nous avons fait varier principalement la quantité de réserve tournante et fixé le taux de pénétration à 50%, ce qui correspond à la valeur la plus proche de la situation réelle du cas d'étude initial. L'horizon de prévision et le temps de mise à jour ont été fixés à 1 heure et 30 minutes, respectivement. Nous avons testé toutes les méthodes de prévision et tous les types de journée dans ce scénario.

Les résultats de ce scénario, comme montré par la Figure 11, indiquent que des niveaux de réserve tournante plus faibles permettent d'atteindre des taux de pénétration PV plus élevés, mais la réserve pour stabilité du système diminue également. Nous avons observé que, pour un niveau donné de réserve tournante journalière, différents taux de stabilité peuvent être obtenus. Ainsi, l'organisation de la réserve tournante est importante pour le coût global du système. Dans ce contexte, l'utilisation de la prévision solaire s'avère être central pour améliorer l'hybridation diesel / solaire.

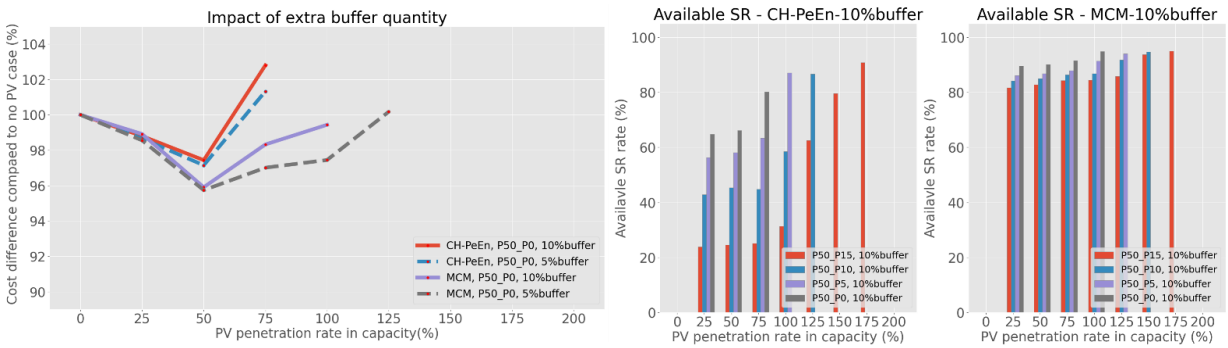


Figure 11 : résultats sur la simulation en explorant l'influence de la quantité de la réserve tournante (scénario 2).

Scénario 3 – L'influence des différents horizons de prévision sur la performance économique :

Le scénario 3 de notre étude évalue l'impact de différents horizons de prévision sur la performance économique du système hybride. Nous avons uniquement fait varier l'horizon de prévision de 1 à 3 heures, tandis que les autres paramètres étaient identiques à ceux du scénario 1.

Comme montré avec la Figure 12, les résultats obtenus indiquent qu'en général, un horizon de prévision plus long peut améliorer le gain économique du système, surtout pour les méthodes de prévision de haute qualité. Toutefois, nous avons remarqué que le gain économique est plus élevé pour les journées de ciel clair que pour les journées variables, car il y a une différence en termes de ressources solaires. Pour les méthodes de prévision qui ne sont pas conçues pour les horizons de prévision plus longs, comme la méthode MCM, l'augmentation de l'horizon de prévision peut avoir un effet négatif, entraînant une perte de précision.

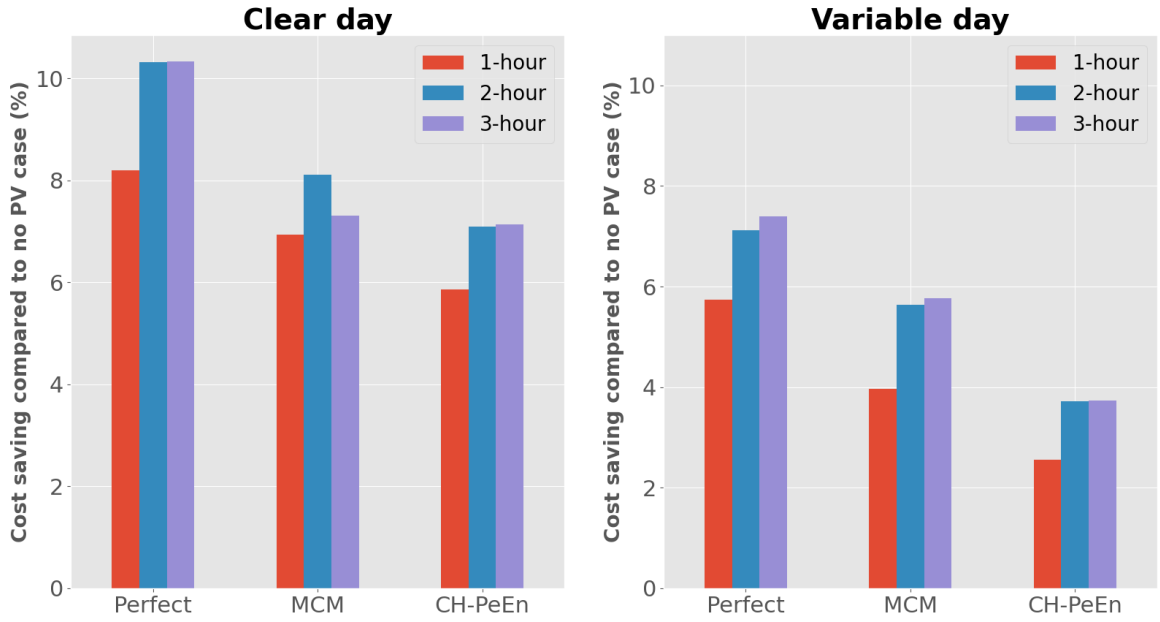


Figure 12 : résultats de la simulation sur l'influence des différents horizons de prévision (scénario 3).

Scénario 4 – L'influence des différentes fréquences de mettre à jour l'ordre de dispatching sur la performance économique :

Enfin, nous nous intéressons sur l'influence de différents temps de mise à jour de dispatching sur la performance du système. Dans le scénario 4, on fait varier le temps de mise jour de 30 à 180-min. La reste des paramètres restent les mêmes comme le scénario 3.

Le résultat montré au Figure 13, indiques qu'en général, un temps de mise à jour plus court permet d'améliorer le gain économique du système. Ainsi, le manque de ressources solaires rend le gain d'économie moins important en journée variable. Aussi noter que le capex des groupes électrogènes diesel n'est pas considéré dans cette simulation, sinon, le gain avec un temps de mise à jour court peut être réduit en raison de perdre la durée de vie.

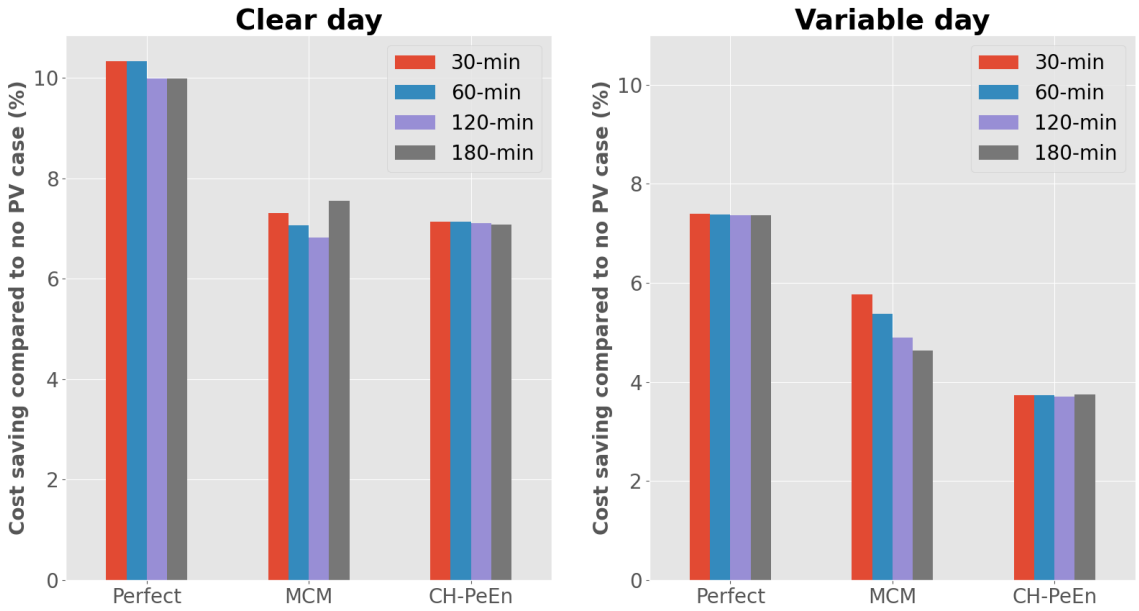


Figure 13 Résultat de la simulation sur l'influence des différentes fréquences de mise à jour

Scénario 5 – Le gain d'économique en utilisant différentes méthodes de prévision :

Après avoir compris l'impact des différents paramètres sur les performances du système, nous avons exploré, dans le scénario 5, les avantages d'utiliser l'hybridation diesel / solaire en combinaison avec une prévision à court terme, par rapport à un cas sans énergie solaire. Dans ce scénario, nous avons conservé la plupart des éléments du scénario 3, en fixant l'horizon de prévision à 3 heures pour toutes les méthodes de prévision, avec une extension de 12 heures pour la prévision parfaite. La différence principale est l'utilisation de données sur un an plutôt que sur deux jours (ciel et variable), ce qui rend de nos résultats plus fiables.

Comme montré par la Figure 14, pour un taux de pénétration de 50% en termes de capacité, l'utilisation de la méthode de CH-PeEn nous a permis d'obtenir un gain économique annuel d'environ 2,8% par rapport au cas sans énergie solaire. En examinant les coûts du système de manière plus détaillée, nous constatons que le coût de la consommation du fioul a été réduit par rapport à la référence, mais le coût associé à l'utilisation de l'énergie solaire a augmenté, notamment le coût de maintien de la réserve tournante qui est dimensionnée en fonction de la marge d'incertitude trop importante. L'utilisation d'un modèle plus avancé, tel que le MCM, nous a permis d'obtenir un gain

économique de 5,2%. Cette méthode permet de réduire les coûts liés à la réserve tournante, car elle dispose d'une marge d'incertitude plus faible. En cas de prévision parfaite sur un horizon de 3-h, le gain économique peut atteindre 8,5%. Si nous étendons l'horizon de prévision à 12-h, le gain peut augmenter jusqu'à 11%, car l'optimiseur peut mieux organiser le dispatching des groupes électrogènes diesel. Ainsi, nous pouvons conclure que l'utilisation des méthodes de prévision peut offrir une marge de gain économique allant de 2,8% à 11% par rapport au cas sans énergie solaire. La méthode très simple MCM permet d'atteindre presque une valeur médiane entre ces deux extrêmes.

Il est important de noter que cette marge de gain n'est pas générique et ne s'applique qu'à cette configuration de simulation. L'objectif principal du simulateur développé dans le cadre de cette thèse est de fournir aux utilisateurs la possibilité de configurer librement tous les paramètres de la simulation, afin qu'ils puissent obtenir une réponse adaptée à leur propre configuration.

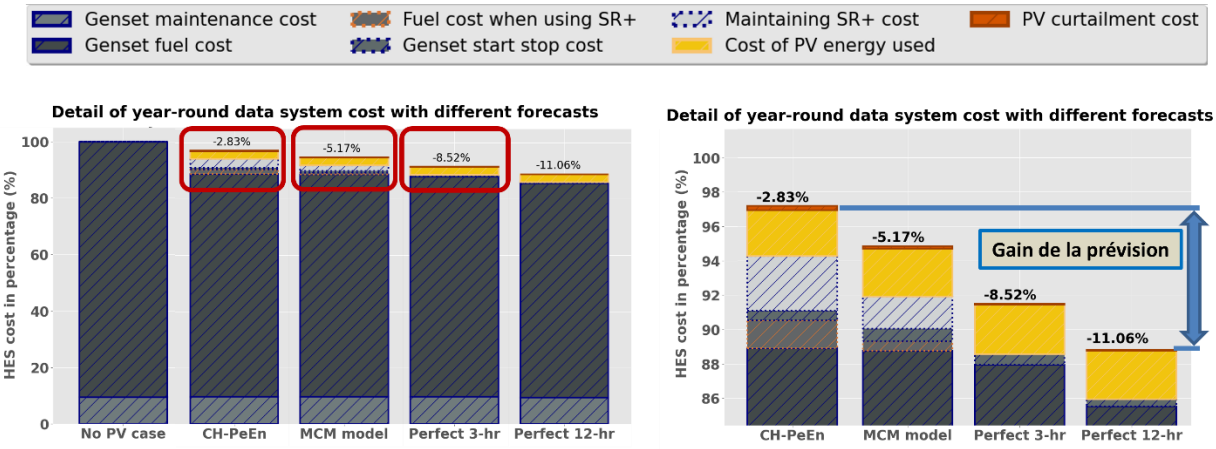


Figure 14 : résultats de la simulation en explorant le gain d'utilisation de la prévision entre la version minimale CH-PeEN et la version parfaite, sans erreur sur 12 h d'horizon.

Question Principale : Quelle est la relation entre la qualité de la prévision probabiliste et la performance du système hybride ?

Après avoir compris le fonctionnement du simulateur ainsi que les sensibilités des différents paramètres, nous avons exploré la question scientifique principale de notre étude. Dans un premier temps, nous nous sommes intéressés à la performance statistique des méthodes de prévision en utilisant des métriques courantes telles que le RMSE (*Root Mean Square Error*), le MAE (*Mean Absolute Error*), le nCRPS (normalized *Continuous Ranked Probability Score*) et le PINAW (*Prediction Interval Normalized Average Width*) pour obtenir des résultats journaliers tout au long de l'année à comparer aux gains financiers journaliers du système hybride apportés par les différentes méthodes de prévision testées. Ces résultats ont montré que la méthode MCM était plus performante que la méthode CH-PeEn en raison d'une distribution d'erreur plus centrée et plus petite. Comme montré la Figure 15, la distribution de CRPS pour le CH-PeEn a deux modes principaux très distincts, ce qui montre que sa performance en ciel clair et variable est très différente, et la distribution générale est plus grand que cellle du CRPS du modèle MCM.

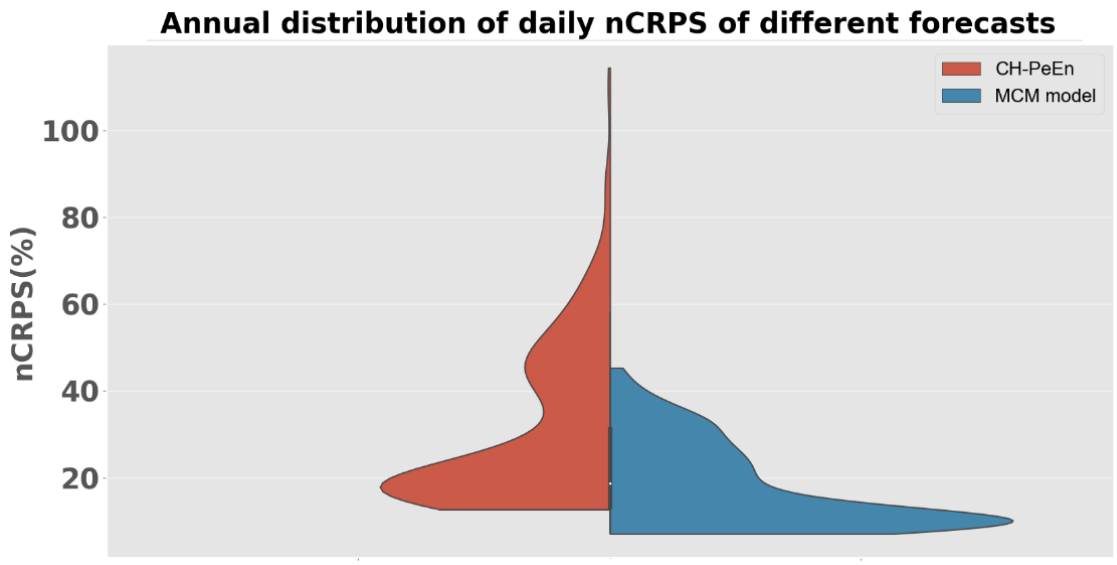


Figure 15 : résultats sur le nCRPS des méthodes prévisions utilisées.

En ce qui concerne la performance économique, nous avons constaté que la prévision de meilleure qualité (Modèle MCM) se comportait mieux dans toutes les conditions météorologiques, y compris lors de ciel clair ou variable. Toutefois, la méthode CH-PeEn ne présentait de bonnes performances que lors des journées ciel clair. Il est important de souligner que le coût du système est l'indicateur de la performance de celui-ci. Cependant, deux coûts du système ne sont pas suffisants pour déterminer la véritable performance pratique, car le coût du système dépend aussi de plusieurs éléments tels que le taux de pénétration PV, l'état du ciel, etc. C'est pourquoi nous avons introduit un nouvel indicateur appelé « *Skill Score of cost* » (SSc), inspiré du « *Skill Score* » utilisé fréquemment dans le domaine de la prévision. Ce SSc est défini comme étant le rapport entre la différence de performance entre le cas considéré et le cas parfait par rapport au cas sans PV.

$$SSc = \frac{Coût_{cas\ sans\ PV} - Coût_{cas\ parfait}}{Coût_{cas\ sans\ PV} - Coût_{cas\ parfait,\ réf}} < 100\%$$

Le SSc utilise le cas sans énergie PV et le cas parfait comme références, ce qui permet d'assurer la bonne marge de performance pour tous les cas considérés.

Les résultats obtenus présentés Figure 16 permettent de constater que l'utilisation du modèle de prévision MCM assure la performance du système dans différentes conditions météorologiques, alors que la méthode CH-PeEn est nettement moins efficace en cas de conditions météorologiques variables. On peut en conclure que l'indicateur SSc est approprié pour évaluer la vraie performance des méthodes de prévision sur la performance économique du système.

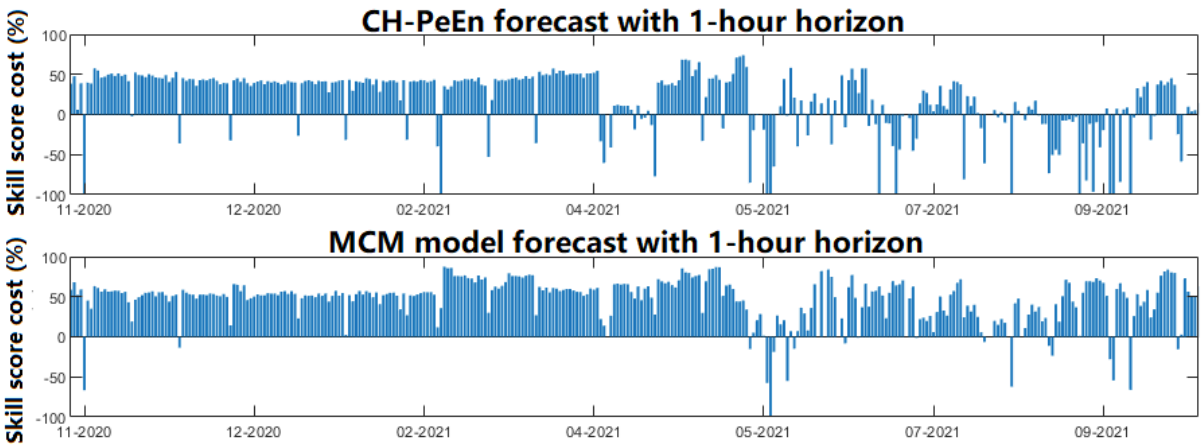


Figure 16 : résultat de la performance économique en utilisant les méthodes de prévisions CH-PeEN et MCM.

Les métriques statistiques classiques comme le nCRPS, le PINAW, etc. ont été utilisées pour évaluer la performance statistique des méthodes de prévision utilisées dans le système.

La Figure 17 montre la relation entre le nCRPS et le SSc pour les deux méthodes de prévision CH-PeEN et MCM. Ces résultats indiquent qu'il peut y avoir de nombreuses possibilités en termes de SSc pour un même niveau d'erreur statistique. Autrement dit, les métriques statistiques comme le nCRPS ne suffisent pas à représenter la performance économique du système. La relation entre ces deux types de performance n'est pas linéaire et la corrélation associée n'est pas claire. Cette conclusion est valable pour toutes les autres métriques qui ont été testées, telles que RMSE, MAE, PINAW et PICP, etc.

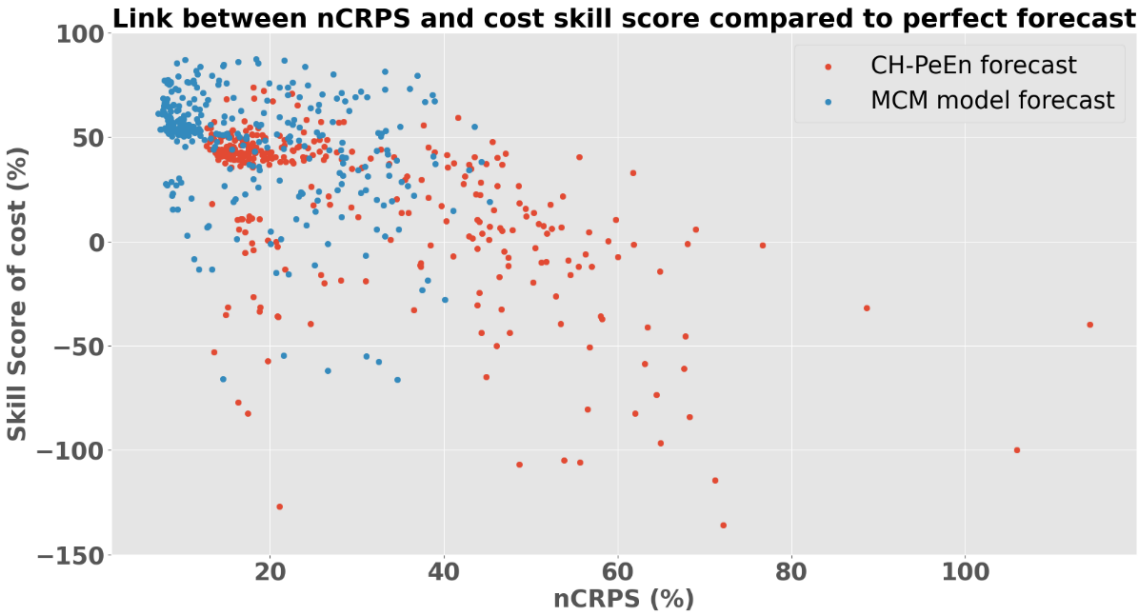


Figure 17 : relation entre le nCRPS et le Skill Score des coûts (SSc) avec les deux différentes méthodes de prévision CH-PeEN et MCM.

L'utilisation du simulateur développé dans cette thèse apparait comme une meilleure option pour évaluer la performance de la prévision dans une application réelle comparée aux métriques classiquement utilisées.

En conclusion, plusieurs points ont été tirées de cette thèse. Tout d'abord, le concept de système hybride sans stockage utilisant la prévision solaire est prouvé comme étant faisable. Deuxièmement, les métriques classiques ne permettent pas directement de représenter la valeur ajoutée de l'utilisation de la prévision solaire sur un tel système hybride diesel / solaire. Enfin, le simulateur développé dans le cadre de cette thèse configurable afin de s'adapter à un grand nombre de cas est un moyen efficace d'évaluer la performance des méthodes de prévision dans une application réelle pour un lieu donné, compte tenu de la variabilité intra-journalière de sa ressource solaire. Ainsi, l'utilisation de différentes méthodes de prévision avec cette méthodologie permet de dériver la marge de gain économique potentiel : la prévision parfaite fournit la borne supérieure et la méthode CH-PeEN, basique, fournit la borne inférieure du gain économique d'un certain niveau de pénétration PV dans le système hybride diesel / solaire.

Malgré tous les efforts déployés dans cette thèse, il y a des limites aux résultats de la simulation en raison des hypothèses faites et de la limite des données disponibles. Par conséquent, en termes de perspective, une simulation étendue avec plus de contraintes, des applications du système hybride automatisé avec la précondition que les groupes diesel pilotables par des automates sont envisageables. Enfin, l'étude de prédimensionnement avec le concept de surdimensionnement du PV contrôlé par son écrêtage, avec et sans système de stockage d'énergie pour des profils de demande quasi constants est intéressante à examiner. Le simulateur tel que développé dans le cadre de la thèse permettrait de mener ces études, moyennant quelques améliorations.

ABSTRACT

Driven by a decreasing trend in the cost of photovoltaic (PV) and serious environmental concerns, the solar energy sector has grown considerably over the last decades and is expected to keep developing fast in the future.

Nevertheless, solar resource variability raises some difficulties for the maintain of energy supply/demand balance, especially in an off-grid context with a large share of fluctuating solar energy.

The energy system studied in this work is an isolated site (i.e. it is not connected to the local or national grid). The best way to make use of solar energy in this context is therefore to build a Hybrid Energy System (HES). Unlike the grid-connected HES that have a redundant capacity to secure the power supply/demand balance, an off-grid HES has to maintain the frequency and voltage by itself.

Energy storage systems (ESS) are usually implemented as an energy buffer to mitigate the solar variability. However, the benefits stemming from the use of storage systems are accompanied by some adverse effects: additional investment costs, higher environmental impacts, maintenance issues, etc. Hence, to bypass these shortcomings, an alternative to ESS is explored in this work. We investigate to which extent the need for storage system can be avoided by using a dedicated Power Managing System (PMS) with short-term probabilistic solar forecasting, providing PV power forecasting along with uncertainty.

In the literature, there are lots of work study about solar forecasting and algorithms for Unit commitment (UC) problem, however, there is no particular work study on this kind of concrete problem for an insular storage-less HES. This concept could eventually allow isolated grid operators to avoid the huge initial investment and bypass the constraint of installing an ESS. Therefore, this thesis aims to find out the feasibility of this concept and the gain of using a PMS with short-term probabilistic forecasting compared to a no PV case.

To be able to evaluate the HES economic performance, we implemented a simulation platform based on a cost-based approach, which evaluates the energy system performance in cost, and the final system cost is used as the performance indicator. This simulation platform has several configurable parameters for different scenarios, such as the PV penetration rate, the generator dispatch update time, the forecast horizon, different sizing approaches for spinning reserve, different forecast methods, and eventually the hardware setup as well as the fuel cost.

In our simulation, the dynamic dispatching of genset is driven by the probabilistic forecast method, which means that the overall effective system cost is dependent on the forecast method quality and the solar variability. Therefore, we explore the relationship between the statistical performance of forecast and the final system economic performance. In order to reflect the influence of forecast quality in system economic performance, we introduce a system cost skill score using the no PV and perfect cases as the reference, where the denominator is the cost difference between no PV and perfect case, and the numerator is the cost difference between the no PV and forecast model case.

The obtained results show that the classic statistical metrics are able to give general distinguishing information between different forecast methods, but they are not able to correctly reflect the practical economic performance.

Hence, using a cost-based simulation tool maybe is a direct and intuitive option to evaluate the influence of forecast method quality in final system economic performance. Accompanied by the cost skill score, this setup is very suitable to assess the system performance of a HES using solar forecast information.

TABLE OF CONTENTS

ABBREVIATION

PMS	Power Management System
RES	Renewable Energy Sources
HES	Hybrid Energy System
ESS	Energy Storage System
LCOE	Levelized Cost of energy
SR	Spinning Reserve
PV	Photovoltaic
GHG	Greenhouse Gas
UC	Unit Commitment
GHI	Global Horizontal Irradiation
GTI	Global Tilted Irradiance
DNI	Direct Normal Irradiance
DHI	Diffuse Horizontal Irradiance
CMV	Cloud Motion Vector
IEA	International Energy Agency
IRENA	International Renewable Energy Agency
NWP	Numerical Weather Prediction
GFS	Global Forecast System
ECMWF	European Centre for Medium-Range Weather Forecasts
WRF	The Weather Research and Forecasting
RMSE	Root Mean Square Error
MAE	Mean Absolute Error
MBE	Mean Bias Error
CRPS	Continuous Ranked Probability Score
PINAW	Prediction Interval Normalized Average Width
TDI	Temporal Distortion Index
TDM	Temporal Distortion Mixed

PLC	Programmable Logic Control
ASI	All-Sky Imager
CAPEX	Capital Expenditure
OPEX	Operating Expenditure

LIST OF FIGURES

LIST OF TABLES

Chapter 1

1. Introduction

1.1 The need of alternative energy to fossil energy

1.1.1 Actual state of world energy consumption

With the advancement of modernization and urbanization, energy needs increase rapidly: the world final energy consumption[1] has been multiplied by 8 in the past five decades [1], [2]. The Figure 1.1 shows this trend of energy need increasement. Besides, we can notice that fossil fuel energy still has a very high proportion in the final energy produced and plays an important role in the global Greenhouse Gas (GHG) emission. As shown in Figure 1.2, around three-quarters of the GHG emission comes from energy production sector.

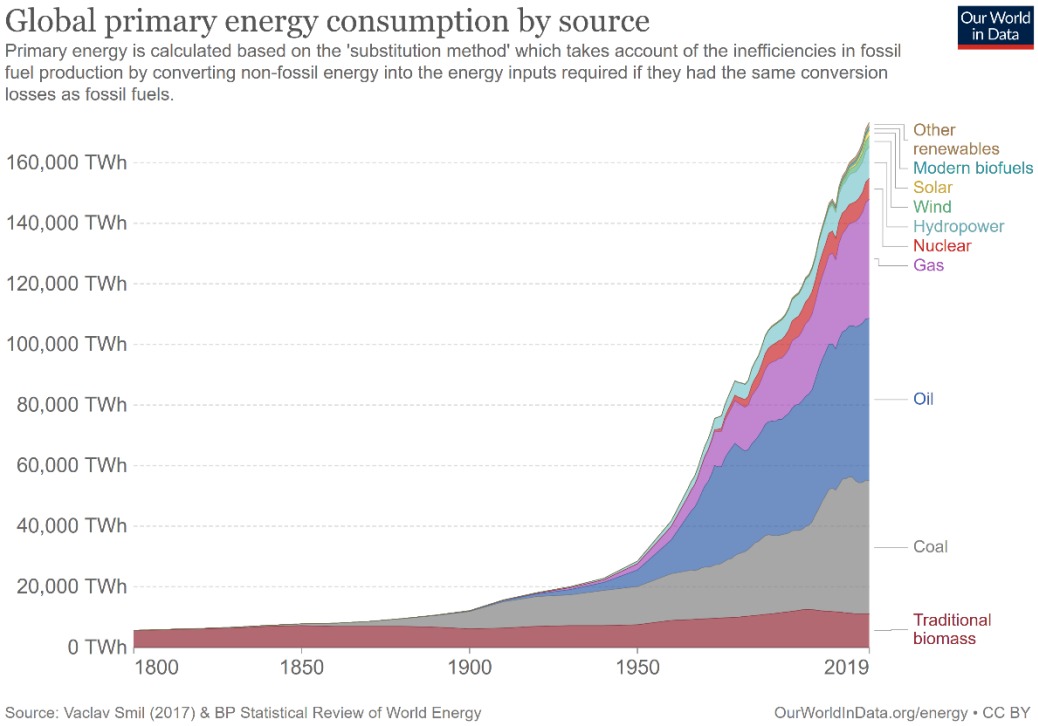


Figure 1.1 World energy consumption trend, [1]

[1] Final energy accounts for the consumption of all fuels, in final forms, consumed by an economy

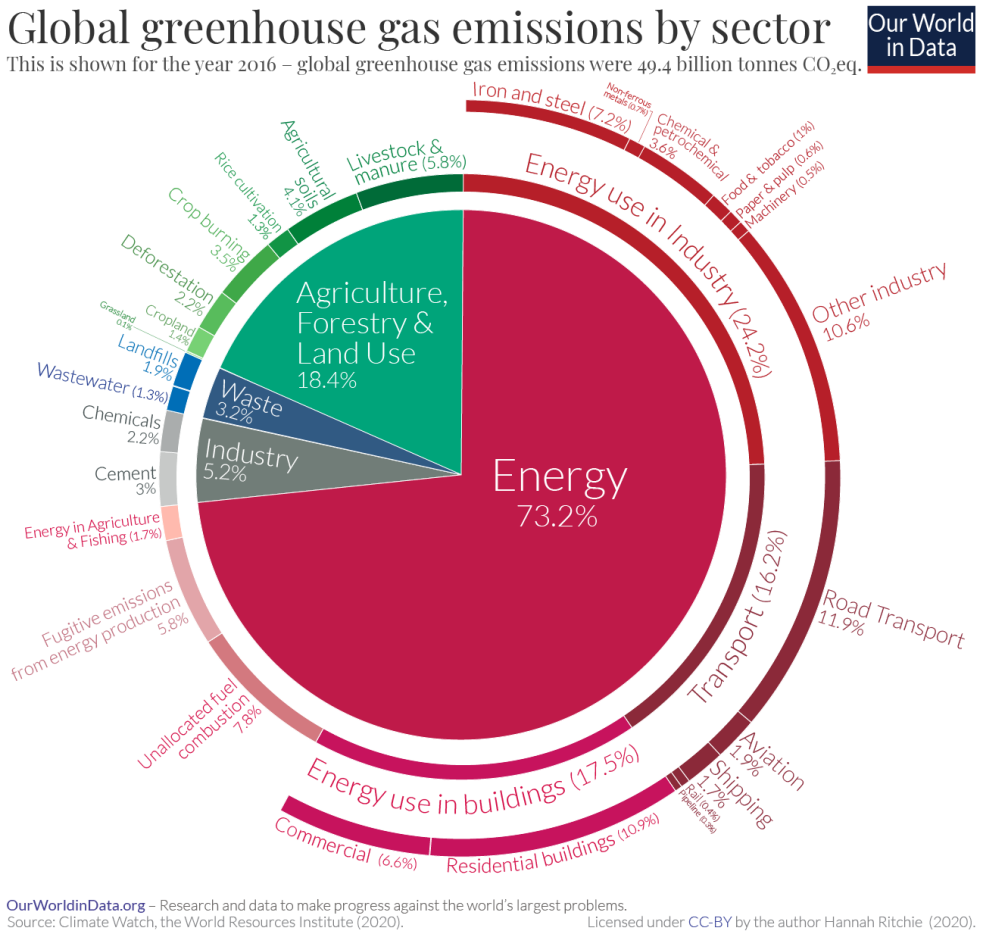


Figure 1.2 Global GHG Emission chart, (Climate Watch, 2020)

1.1.2 Why we need decarbonization

According to the report of [3], the Greenhouse Gas (GHG) emission of 2019 has increased more than 62% compared to 1990. And the use of fossil energy could eventually contribute to the greenhouse effect, which has several adverse environmental impacts, such as air pollution due to the particles in the air, the global warming due to the stuck of CO_2 and the following climate change, etc.

The Figure 1.3 shows the concentration of carbon dioxide (CO_2) in the atmosphere since 2003 that keeps increasing. Considering the increase of fossil energy consumption mentioned before, we can expect that if the fossil energy consumption is not replaced by a clean decarbonized energy, the CO_2 concentration will continue to raise.

Part of the emitted CO_2 is eventually absorbed by the ocean [4], which increases the ocean acidity and affect their biodiversity. The rest of the emitted CO_2 stays in the atmosphere for thousands of years, which increases slowly the temperature of our planet. In terms, with a high level of CO_2 stuck in the atmosphere, the temperature of Earth surface will increase and the biodiversity would probably be affected [5], [6]. In addition, it could probably melt the glacier, flood the city and create several social issues. In should be noticed that CO_2 is not the only GHG, methane, nitrous oxide and fluorinated gases also have global warming potential and their emission should thus be controlled.

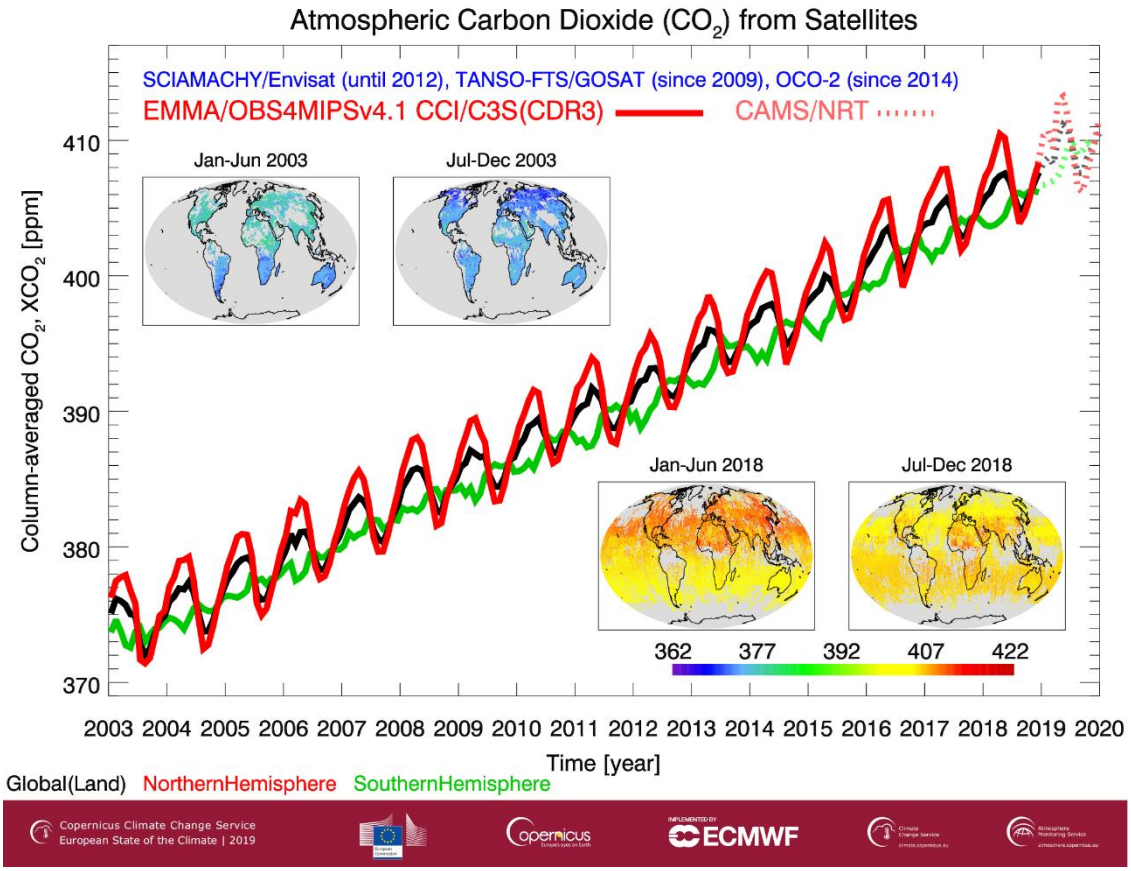


Figure 1.3 Atmosphere CO_2 concentration, Copernicus Atmosphere Monitoring Service (CAMS)

Facing these problems, different international agreements are signed, such as the famous "Kyoto Protocol" and "Paris Agreement", which aim to deal with the climate change by limiting the GHG emission and the temperature-rise change. More and more countries and regions have successively introduced some corresponding policies. For example, Japan and European Union have committed to achieve carbon neutrality in 2050 while China aims to achieve this goal in 2060.

1.1.3 How to achieve the decarbonization?

An ambitious objective could be achieved only by using efficient solution, such as electrifying the fossil fuel dependent sector in the energy demand and using renewable energy. Due to the lexical misuse between electricity and energy, some people might have the misunderstanding that we can totally reduce the GHG emission by switching to RES in electricity production aspect. In fact, the electricity production is only one part of the total energy production, heating and transport are other two main parts of the total energy production and they are harder to decarbonize since they are much more fossil fuel dependent, as shown in Figure 1.4. Electrify these sectors through clean renewable energy would be useful and necessary.

According to [7], which study the life cycle assessment of different types of energy production, the environmental impacts of renewable energy sources (RES) in terms of GHG emission is significantly lower than those of fossil energy, even when considering their whole life cycle. For example, for electric power production, the average emission value of CO_2 of oil is around 800 gCO_2eq/kWh, and for coal is even higher, around 1000 gCO_2eq/kWh for the average value. But for RES, the median value of

Photovoltaic energy is lower than 50 gCO_2eq/kWh, which is more than 15 times less [7]. Hence, RES is an efficient tool to reduce the environmental impacts.

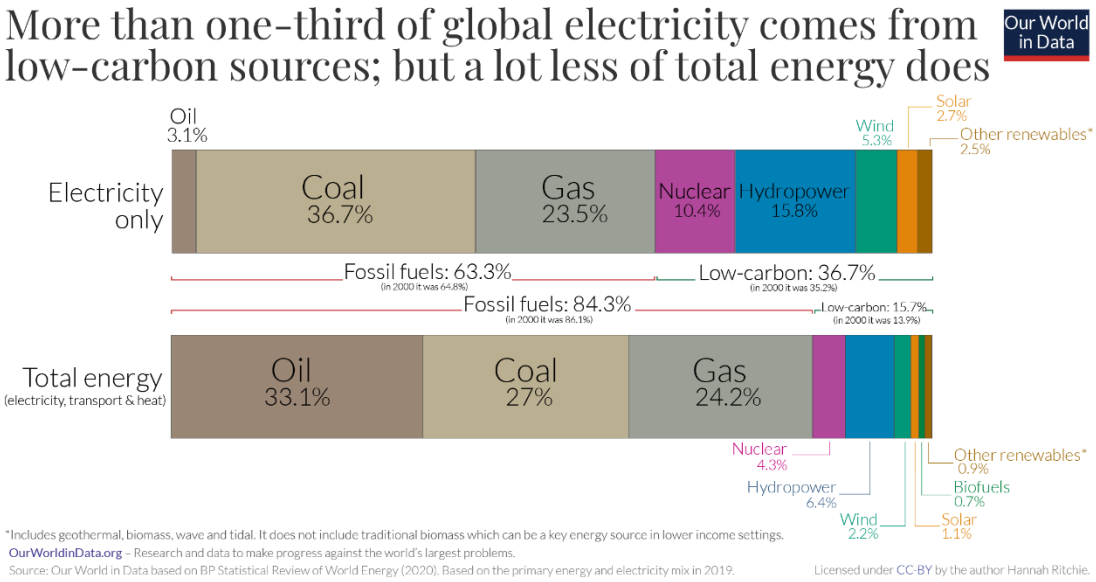


Figure 1.4 Electricity and world energy distribution, [2]

In the literature, there are many works study the RES energy used in specific sector, such as mining industry. Indeed, mining industry represents around one third of the final global industrial energy consumption (*e.g.* [8]). This energy consumption is huge, and it results in very important GHG emissions as well [9].

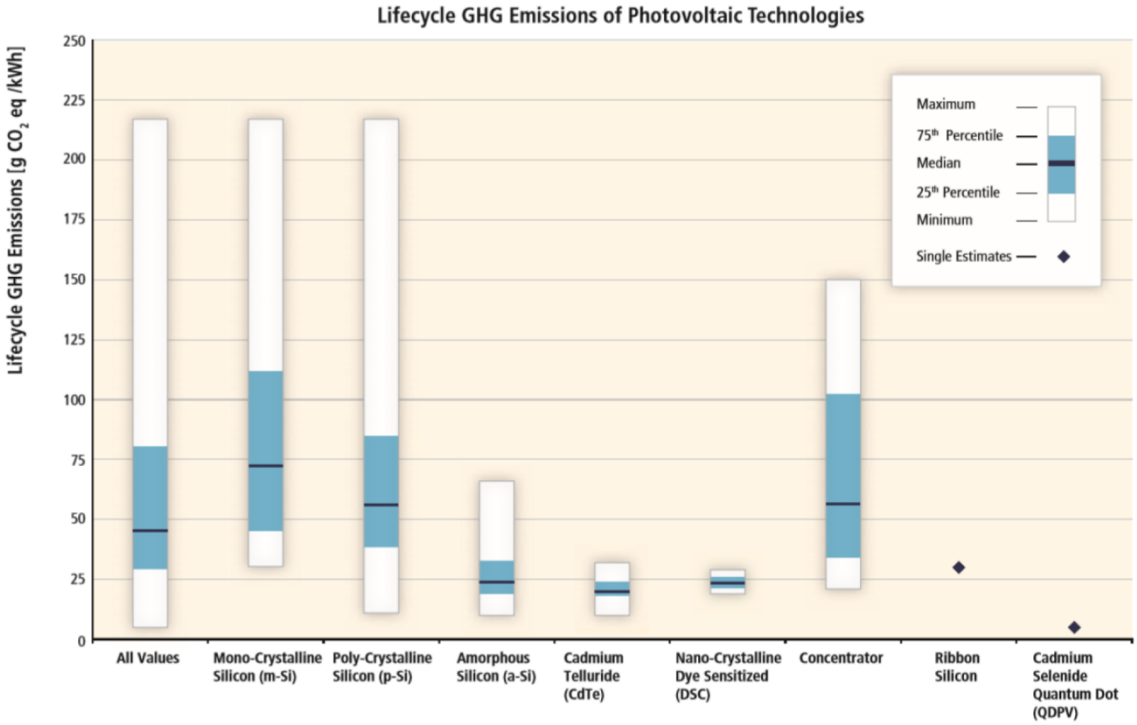


Figure 1.5 GHG emission of different Photovoltaic technology, [4]

Figure 1.6, taken from REN21 report, shows that the installed capacity of RES increases each year; RES is hence becoming more and more important for our future energy supply. Solar PV, wind power and hydropower are the most used RES in 2020 RES market and keep having high growing speed. Since

2016, solar PV technology surpasses wind power takes the first place in absolute installed capacity growth. Among all existing renewable energies, solar PV energy represents the largest absolute generated power growth according to IEA (International Energy Agency) solar PV report 2020. And the installation trend in the whole world is keep going fast according the REN21 report.

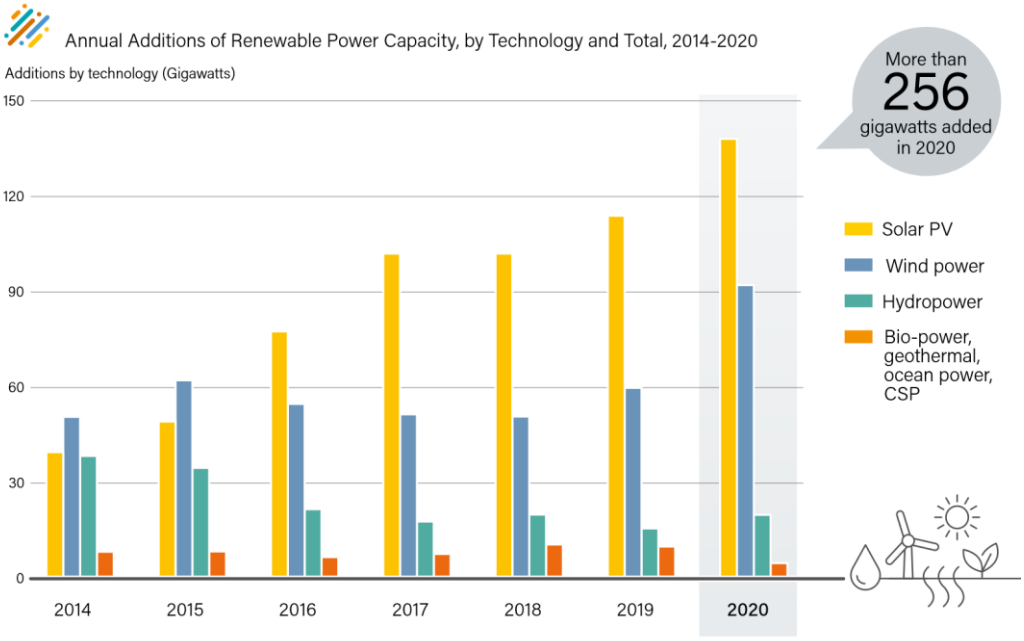


Figure 1.6 RES Global Capacity and Annual Additions (REN21, 2020), [10]

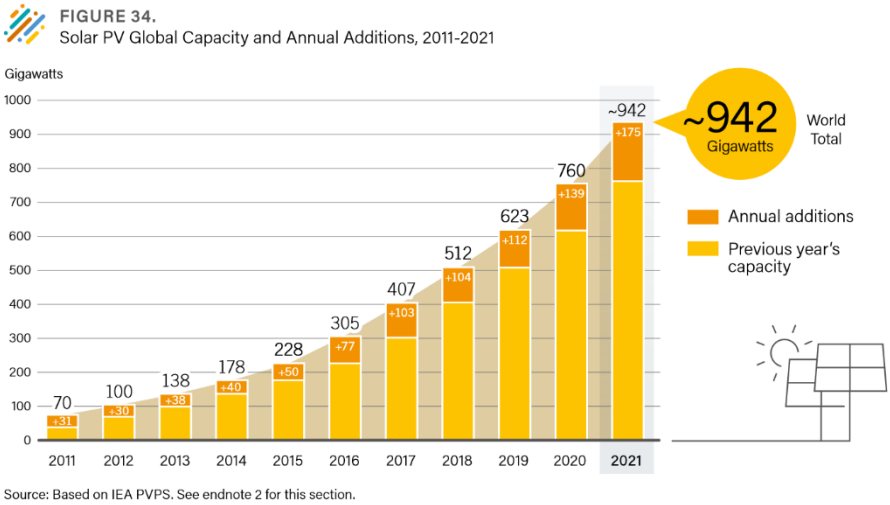


Figure 1.7 PV energy Global Capacity and Annual Additions (REN21, 2022)

1.2 The advantage of solar energy

[10] indicated that solar energy is the most adequate, clean and almost inexhaustible renewable energy so far. What we receive on Earth surface after atmosphere from the Sun in terms of power is around 82 x 10^6 GW, which is more than 5 000 times larger compared to what we consume at the same time. Photovoltaic (PV) technology is so far the most developed technology to make use of solar energy and has the most growth potential (IEA, 2019). Since the production cost of PV panel is getting lower and lower (Figure 1.8, IRENA), the condition of use is less restricted compared to concentrated solar power plant, not only for individual use but also for large-scale power plant, solar PV therefore becomes popular among the different RES.

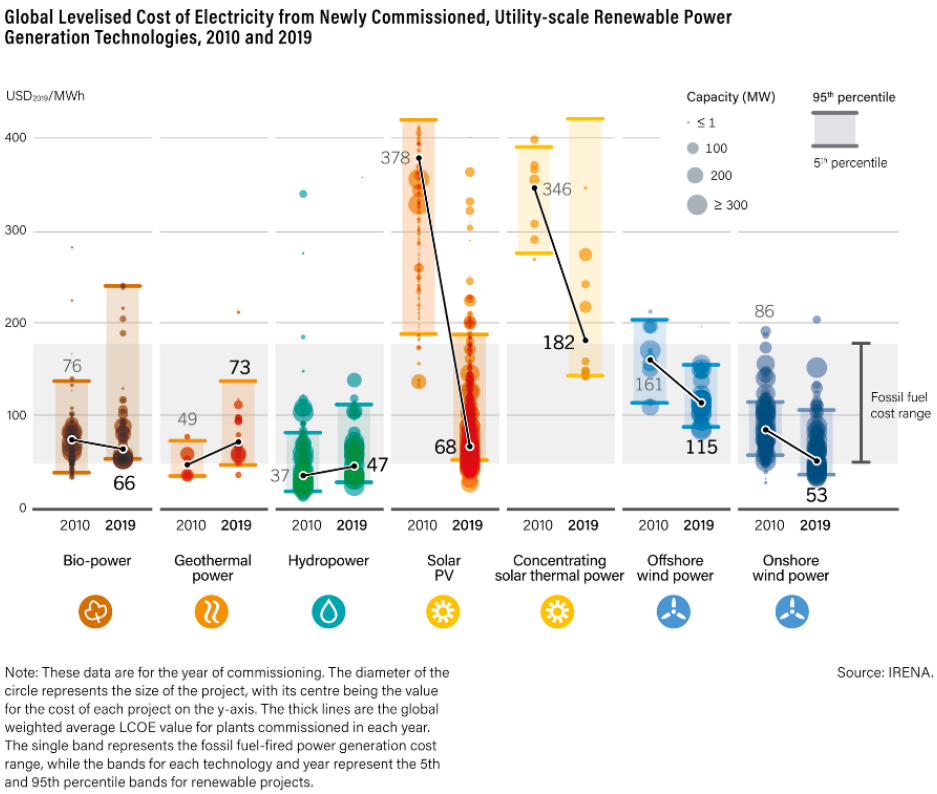


Figure 1.8 Global levelized cost of different RES (REN21, 2020)

Today, some mining companies try to make use of solar resource, especially in the zones with high solar potential such as South America, Australia [11], and, especially, Africa, where has the highest potentiality among all these areas [12].

The global horizontal irradiation map made by *Global Solar Atlas*, shown in Figure 1.9, makes it clear that the solar resource in Africa is very abundant. It means that the potentiality of solar energy development there is very high. Therefore, combining the advantage of lower cost in PV system and the abundant solar resource, in Africa, it is an ideal place to make use of the solar resource.

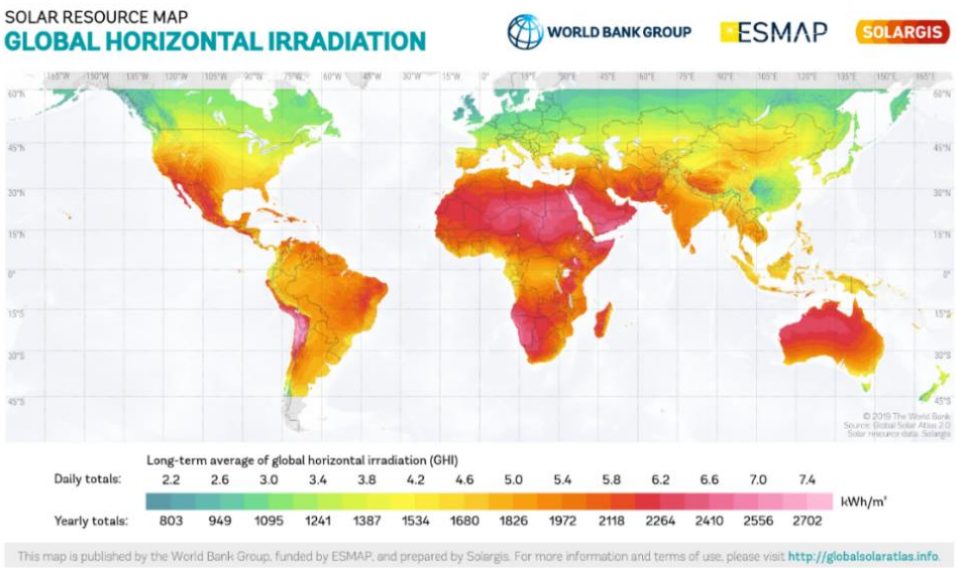


Figure 1.9 World solar resource map, source: Global Solar Atlas, World Bank Group

Moreover, [13] indicate that the Africa economic is growing fast, which implies that the electricity demand grows very fast as well. It is expected to be 3 times as today by 2030 according to the report of [14]. In order to meet this increase of energy demand and to respect the Paris agreement, PV technology is therefore chosen as fossil energy alternative. Through this method, these mining companies aim to reduce the environmental impacts, the fossil fuel consumption, the cost of electricity and reduce the fuel supply problem for some off-grid projects.

1.3 Hybrid energy system (HES)

The case study considered in this thesis is initially a thermal diesel power plant, which supplies the power demand of a mine in Mali (Africa). Presently the mine is (self-)supplied by an isolated thermal power plant of 50 MW, composed of 27 diesel generators and have been built and put in operation in 2016. Since the power need increased, the operator decided to add a solar PV power plant (24 MWp[2]) rather than using more diesel generators, in order to reduce the electricity cost in the whole installation. The new power system that mix diesel generators and solar PV power system is meant to build a Hybrid Energy System (HES). In this thesis, solar PV is discussed for the Hybrid Energy System (HES) study.

However, even though PV energy is abundant and clean, it has its own inconvenient – its variability, which could cause the instability of power system and should be mitigated.

1.3.1 Off-grid HES

When the hybrid (PV & Diesel) system is connected to the national grid, this variability issue might not be that problematic. Yet, most of mining zones are far away from residential and do not have access to the national or local grid; this is a so-called “off-grid” power system.

More precisely, unlike the grid-connected one, off-grid power system (also called isolated power system) cannot resort to the redundant power from national grid in response to the ramps of solar energy production variation, which means that the isolated network frequency and voltage must be maintained by the local power system itself. In a HES which combines diesel generators and PV power plant, the frequency mainly depends on the diesel generator because the frequency is related to the generator rotated speed.

1.3.2 Spinning reserve

In general, the power supply of most off-grid mines is based on diesel generator. As a result, if no other balancing mechanisms such as storage is implemented, the variability of renewable energy needs to be maintained by spinning reserve[3].

Traditionally, spinning reserve defines the reserve capacity of an energy system to mitigate the power supply variability due to some energy system equipment failure, such as the failure of diesel generator, inverter, etc. The spinning reserve is counted in Watt and represent the maximum reserve power that the system is able to supply in a very short delay in case of partially power supply loss. For example, with 1 MW spinning reserve, the energy system is able to mitigate the loss of generator with 1 MW nominal power. If the power supply loss is over the spinning reserve quantity, the energy system would be shutdown suddenly and all the generators would be turned offline. This energy outage situation is

[2] Wp: Watt-peak

[3] Spinning Reserve (SR) : A reserved capacity of diesel genset to solely manage the load, the sum of all SR should be superior to the nominal power of the biggest diesel generator

commonly called “blackout”. The spinning reserve is provided by running generators below their nominal power and the range to the maximum power supply capacity is the quantity of spinning reserve. Notice that the online spinning reserve is provided by online generators and cannot be supplied by offline generators because starting up generators takes much longer time than the energy outage limit. This is also the reason why this reserve capacity is called “spinning reserve”. In the HES of our case study, we expand and distinguish the spinning reserve in positive and negative one:

- the positive spinning reserve combines the reserved capacity in the online spinning machine and the PV curtailment capacity;
- the negative spinning reserve is the capacity of the online spinning machine to slow down and reduce the power output, as well as the load shedding possibility.

As shown in Figure 1.10, in case of the power shortage, the positive spinning reserve works similarly to the discharge process of the Battery Energy Storage System (BESS), the power comes out from the buffer to answer the power lack. In the other way around, the negative spinning reserve works like the charge process of BESS, except that it does not store the energy but reduce the energy production to mitigate the power excess. If the variability is under the security range of spinning reserve, the energy system will be stable. If the energy shortage could not be mitigated, the worse situation is having a “black out” case, which means energy outage and an important loss.

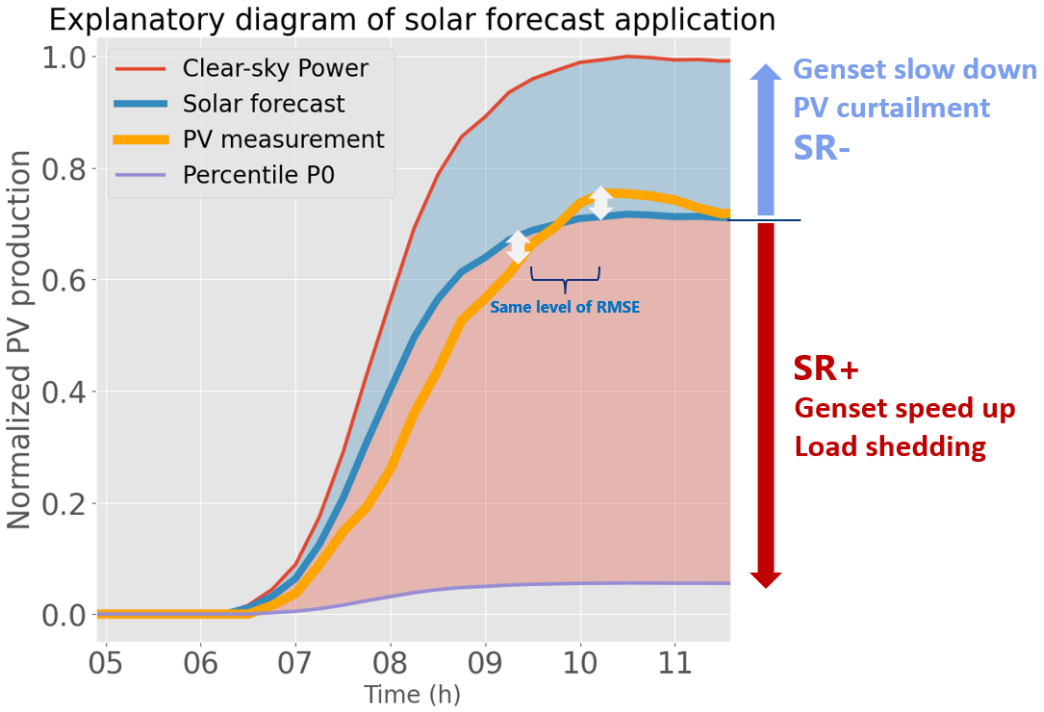


Figure 1.10 Graphical representation of short-term forecast and spinning reserve buffer

1.4 Why without Energy Storage System (ESS)?

In our case study, the PV penetration rate defined as the ratio between the PV installation capacity and the average total grid power demand – is around 48%. Theoretically, we could just keep an amount of negative / positive spinning reserve of diesel power plant to compensate the PV energy variability without other actions. Nevertheless, it could not be true anymore if PV penetration rate becomes too high since a high PV penetration rate ask a high Genset capacity for positive spinning reserve. According

to [15], if the diesel generator keeps its capacity as spinning reserve and works under the optimal function range to ensure the system stability, the economic aspect is worse than pure diesel generator case. This is because the specific fuel consumption (L/kWh) of generator is more important at low level of its nominal power, which means that the quantity of spinning reserve cannot be too high, and the benefits of adding PV into the power system becomes questionable in terms of cost reduction and environmental impacts. In addition, it should be noted that the quantity of spinning reserve increases when adding PV energy since additional balancing capacity is needed to cope with variability and forecast errors.

Presently, according to the electricity network infrastructure, the problem could be solved by using different kinds of storage system, such as Pumped-storage hydroelectricity (PSH) and Battery Energy Storage System (BESS) [16]. As mentioned in most of the reviews of hybrid energy system in the scientific literature (ex. [17], [18], [19], [20], [21]), the proposed solutions are all equipped with an energy storage system.

[15] says that energy storage system can compensate the lack of solar production, meet the peak load demand and smooth the power output. However, he also pointed out that some non-developed countries don't have any recycle policy when the battery arrives at the end of its life cycle, which actually could lead to a serious environmental impact problem. It is indeed functional and seems to be advantageous in terms of convenience and fuel consumption saving for a stand-alone power plant. However, using storage system has several constraints, such as extra investment cost and maintenance cost (e.g. replacement when battery arrives at its end of lifecycle). Moreover, the charge and discharge processes always come with energy losses and the storage system efficiency could fall down due to bad operating conditions. Therefore, if possible, an alternative of storage system could be interesting.

Unlike the study of HES with storage design, there is much less papers in the literature on hybrid energy system without energy storage system -or with a very limited one. However, some papers provide promising paths, such as [13], they have done an experimental study on this subject and provides a satisfying result. In the thesis of [9], the system costs section shows the optimal results of system costs for several generation mixes. Figure 1.11, taken from [9], shows the details of the LCOE[4] (Levelized Cost of energy) of different kinds of generations. We can see that the final production cost per kWh of HES (PV-Battery-Diesel) with battery is higher than those without battery (PV-Diesel). In 2020, the price of storage system installation is around 350 US dollar per kWh for an 4-hour system [22], which means that the power plant with battery system needs much more investment than simple PV & Diesel power plant. And this is an important constraint for some mining companies because they need to have enough cash flow to launch the project.

[4] LCOE (Levelized Cost of energy): the definition is the total life cycle cost divided by the total energy produced during the life cycle.

Table 5.1: Summary of key results across four mines and three mining regions

Atacama, Chile (Off-Grid)

Technological configurations	LCOE (US$/kWh)	% Renewable	TLCC Savings	CO_2 Savings
Wind & PV & CSP Tw (Gas backup) & Diesel	0.167	80%	41%	82%
Wind & PV & CSP Tw (Gas backup)	0.185	79%	33%	82%
Wind & PV & CSP Tw & Diesel	0.185	83%	34%	83%
CSP Tw & Diesel	0.192	83%	32%	83%
CSP Pt (Gas backup) & Diesel	0.210	62%	26%	66%
Wind & PV & CSP Pt & Diesel	0.220	73%	22%	73%
PV & CSP Pt & Diesel	0.223	71%	21%	71%
Wind & PV & Battery & Diesel	0.229	49%	19%	49%
Wind & PV & Diesel	0.232	41%	19%	41%
PV & Diesel	0.236	32%	17%	32%
CSP Pt & Diesel	0.236	71%	16%	71%
PV Tracking & Diesel	0.237	32%	16%	32%
PV & Battery & Diesel	0.238	52%	16%	52%
Wind & Diesel	0.240	32%	15%	32%
Diesel only (Baseline)	0.283	0%	0%	0%
Wind & PV & Battery	0.358	100%	-15%	100%
PV & Battery	0.382	100%	-26%	100%
Wind & Battery	0.404	100%	-30%	100%

Atacama, Chile (Grid-Connected)

Technological configurations	LCOE (US$/kWh)	% Renewable	TLCC Savings	CO_2 Savings
PV & Grid	0.130	27%	7%	27%
PV Tracking & Grid	0.131	27%	6%	27%
Wind & CSP Tw & Grid	0.134	38%	4%	38%
Wind & Grid	0.134	27%	4%	27%
CSP Tw & Grid	0.136	57%	3%	57%
Grid only (Baseline)	0.140	0%	0%	0%

Yukon, Canada (Off-Grid)

Technological configurations	LCOE (US$/kWh)	% Renewable	TLCC Savings	CO_2 Savings
CCGT & Diesel	0.150	0%	57%	47%
Wind & CCGT & Diesel (Near-Optimal 1)	0.151	8%	57%	51%
Wind & CCGT & Diesel (Near-Optimal 2)	0.155	25%	55%	60%
GT & Diesel	0.157	0%	53%	10%
Wind & PV & Diesel	0.256	48%	21%	48%
Wind & PV & CSP Tw & Diesel	0.256	50%	21%	50%
Wind & CSP Tw & Diesel	0.259	54%	20%	54%
Wind & Diesel	0.262	38%	20%	38%
PV & CSP Tw & Diesel	0.309	18%	5%	18%
PV Tracking & Diesel	0.309	18%	5%	18%
PV & Diesel	0.309	18%	5%	18%
CSP Tw & Diesel	0.316	32%	3%	32%
Diesel only (Baseline)	0.325	0%	0%	0%

North-Western Australia (Off-Grid)

Technological configurations	LCOE (US$/kWh)	% Renewable	TLCC Savings	CO_2 Savings
CSP Tw (Gas backup) & Diesel	0.185	75%	40%	68%
PV & CSP Tw (Gas backup)	0.206	68%	32%	59%
Wind & PV & CSP Tw & Diesel	0.208	84%	33%	84%
PV & CSP Tw & Diesel	0.209	85%	32%	85%
Wind & PV Tracking & CSP Tw & Diesel	0.209	84%	32%	84%
Wind & CSP Tw & Diesel	0.209	85%	32%	85%
CSP Tw & Diesel	0.210	86%	32%	86%
Wind & PV & CSP Pt & Diesel	0.244	62%	22%	62%
Wind & PV & Battery & Diesel	0.249	55%	20%	55%
Wind & PV & Diesel	0.251	52%	20%	52%
PV & CSP Pt & Diesel	0.255	66%	18%	66%
PV & Diesel	0.260	32%	16%	32%
PV Tracking & Diesel	0.262	32%	16%	32%
PV & Battery & Diesel	0.264	52%	15%	52%
Wind & Diesel	0.265	39%	15%	39%
CSP Pt & Diesel	0.265	75%	14%	75%
Diesel only (Baseline)	0.310	0%	0%	0%
Wind & PV & Battery	0.381	100%	-16%	100%
PV & Battery	0.490	100%	-53%	100%

Optimal mix including CSP Power

Optimal mix with no CSP Power

Baseline mix

CSP Tw = CSP Tower

CSP Pt = CSP Parabolic Trough

Figure 1.11 System cost of using different types of HES [9]

If the energy storage system were taken away, there is no doubt that we can reduce the initial investment cost and the maintenance cost, and the energy system would be more environmentally friendly. According to the Life Cycle Assessment (LCA) study of [7], the carbon footprint of a battery energy storage system is in the order of 100 kg CO_{2eq}/kW installed, which is around 10% of the PV energy system carbon footprint, with an order of 1t CO_{2eq}/kW installed. It means that a hybrid system without storage could at least reduce 10% carbon footprint less than those has storage system.

1.5 The alternative of ESS – solar forecasting with PMS

The primary function of storage system is to balance the dispatch error resulting from forecast errors as well as to mitigate issues resulting from the solar variability. When the PV energy share is low, the stability of electricity network is relatively easy to ensure, but when the PV penetration rate becomes more and more important, only using the spinning reserve seems to be insufficient [23].

[24] indicates that a reliable solar forecasting technique could help to improve the system performance of using fluctuating solar energy. [25] says that an accurate solar power forecasting can be very useful and important to build an optimum management plan for the power supply/demand balance control of solar PV power plant. Besides the advantage of ensuring the grid stability, [26] also points out that using solar forecasting can improve the so called “optimal availability of electricity” and help in reducing the uncertainty of power plant output. Therefore, it could be easily seen that the quantity of spinning reserve could be smartly and dynamically sized thanks to the solar forecasting information.

According to this principle, the need of storage system could be eventually reduced or even avoided by using a dedicated Power Managing System (PMS) with short-term probabilistic solar forecasting, which providing PV power forecasting along with uncertainty.

The function of this PMS is firstly to predict the PV production with the help of solar forecasting technique and then calculate the residual demand (total demand minus PV production), which needs to be smartly committed by diesel generators. The optimization algorithms for Unit commitment (UC) problem[5] can generate an optimal combination of diesel generators with different characteristics to meet the energy demand. Finally, maintaining the power supply/demand balance and ensuring the grid stability with a minimum energy production cost. Therefore, using solar forecasting technique with PMS to extenuate this effect is theatrically possible and solar forecasting technique is chosen for our study.

1.6 Research questions

The scientific literature already contains several studies about hybrid system (e.g. PV-Diesel), Energy dispatching (e.g. unit commitment) and solar forecasting. However, to the best of our knowledge, there is no study that combines all of them together to find out the feasibility of such hybrid system and the gain that could be obtained by applying these studies to an isolated fossil energy-based power system.

This work aims at filling this gap by addressing these three aspects to provide a methodology which allows people to follow the procedure from practical problem formulation, optimization problem simulation to HES performance analysis to find out the feasibility of certain scenario and intend to learn whether it is worth to do so.

In a nutshell, the main questions to be answered in this thesis could be listed as below:

What is the relationship between the quality of probabilistic solar forecasting and the HES performance?

1. **What would be the best practical metrics to assess HES performance?**
2. **How should a suitable platform to simulate the HES behavior be implemented?**
3. **What are the benefits of using a forecast-driven HES compared to a diesel-only setup case?**

All these underlying questions above are the essential element of this thesis, the research of these questions starts in this work.

[5] (UC) Unit commitment problem: a classical mathematical optimization problem, the detail is showed in the following section.

Chapter 2

2. Elements of building a storage-less HES

Contents

SUMMARY OF CHAPTER

This chapter introduces essential notions for the development of a Hybrid Energy System (HES) such as the Unit Commitment (UC) problem and solar forecasting.

Firstly, we introduce the role of solar forecasting in the operation of an isolated hybrid. Secondly, we demonstrate how an isolated storage-less HES works. Lastly, we show some the state of the art of solar forecasting and optimization methods for solving the UC problem.

In the solar forecasting section, we give an overview of different forecasting techniques with different spatial and temporal resolutions. The metrics to evaluate the forecasting performance are introduced and their advantages and limits are discussed.

In the unit commitment section, we start with an introduction to the basic concepts. We then present the UC problem and detail common optimization methods currently used.

The section on optimization methods can be divided into two main series: the classical mathematical optimization methods and the heuristic methods. In this chapter, the specificities of different optimization methods are briefly introduced.

At the end of this chapter, the major difficulties in building a storage-less HES are discussed.

2.1 Necessity of forecasting

As introduced introduction, a dedicated designed Power Management System (PMS) integrated with solar forecasting can in a certain extent replace the usage of ESS. This chapter introduces why solar forecasting is necessary and describes the essential blocks of a storage-less HES, such as the unit commitment problem and solar forecasting. The ideal HES functionality as well as the major difficulties of building this kind of HES are also introduced.

In a classic thermal energy system with a part of a renewable resource, there will be more unpiloted variations in terms of energy production. For a storage-less designed hybrid energy system, the system constraint becomes even tougher. The decrease of instantaneous PV power can be very fast, potentially much faster than the time needed to start a diesel generator since it needs a preheating process. Without an energy storage system as a buffer, the supply/demand balance is hard to maintain, and it could lead to stability issues and ultimately threaten the security of supply. Therefore, some additional actions must be taken and solar forecasting is one of the choices[25]. Integrating an accurate forecasting method can significantly reduce the system impact of solar energy variability by anticipating solar power generation. This anticipation can be made at two levels: 1) in the planning of the dispatchable generating units, and 2) in the sizing of the spinning reserve to mitigate fast variability and forecasts errors.

2.1.1 Forecasting for the dispatch of generating units

Solar forecasting provides information on future PV production and could be used to determine the residual demand. Residual demand is the difference between the total demand and PV production, also called net load, which should be committed by the diesel generators. In the example shown in Figure 2.1, the red line is the power provided by two 5 MW diesel generators, whose optimal operation range is from 70% to 100% of the nominal power. If at instant $(t + \delta t)$, the residual demand is lower than the power provided by these two generators, the situation is not optimal, even if the diesel generator power output is decreased to its minimum limit, or the PV power generation curtailed. Hence, instead of using big diesel generators with overproduction, we can use smaller diesel generators to meet the need without energy waste, which is less expensive and more environmentally friendly. However, starting up the diesel generator can't be done in one second, so we need to forecast to anticipate at instant (t) to make a good decision.

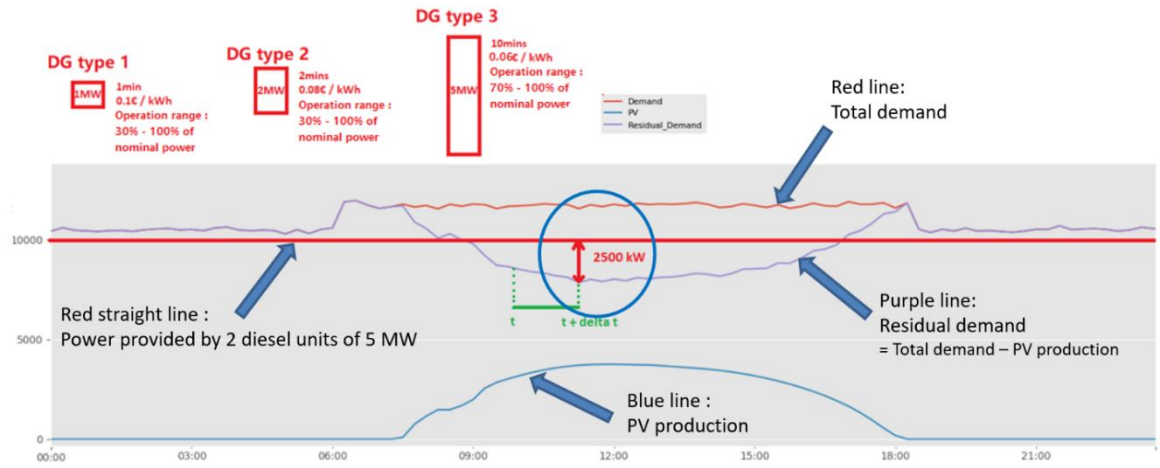


Figure 2.1 Simulation result of a simplified Hybrid Energy System

2.1.2 Forecasting for the sizing of spinning reserve

A reasonable spinning reserve sizing should not be too large, since it is costly, and nor too small, or it could not mitigate the energy variations. In a classic HES equipped with an energy storage system, the system variations can be absorbed by an energy storage system and extra spinning reserve. But our storage-less HES can only rely on the spinning reserve of diesel generators to ensure grid stability. A smart sizing of the spinning reserve is therefore paramount.

Without solar forecasting information, the quantity of spinning reserve would be oversized in most of the cases because it would be defined according to the possible worst situation. An excessive oversizing of the spinning reserve is uneconomical because keeping a spinning reserve has an extra cost. To smartly size the spinning reserve, the solar forecast is important, but knowing the possible forecast error is also necessary. To begin with, a deterministic forecast with uncertainty could already help, but the best option would be probabilistic solar forecasting, which provides different quantile levels of uncertainty for users. In an ideal case, the solar forecast could help to define a dynamic level of spinning reserve, which means that when the forecast uncertainty is small, the redundancy of spinning reserve could be smaller and when the uncertainty is high, the spinning reserve will be sized larger to ensure the power supply security.

In brief, solar forecasting with uncertainty is particularly important for storage-less HES since planning deviations are entirely balanced by the spinning reserve. The security and stability of the power supply could be ensured by a precisely designed power management system (PMS) with integrated solar forecasting. The ideal functionality of a storage-less HES with the help of PMS and solar forecasting is introduced in the following paragraph.

2.2 Ideal Storage-less HES functionality

Our storage-less HES is composed of three parts as shown in Figure 2.2: 1) the thermal power plant, made of several different diesel generators, 2) the PV power plant, 3) the power demand.

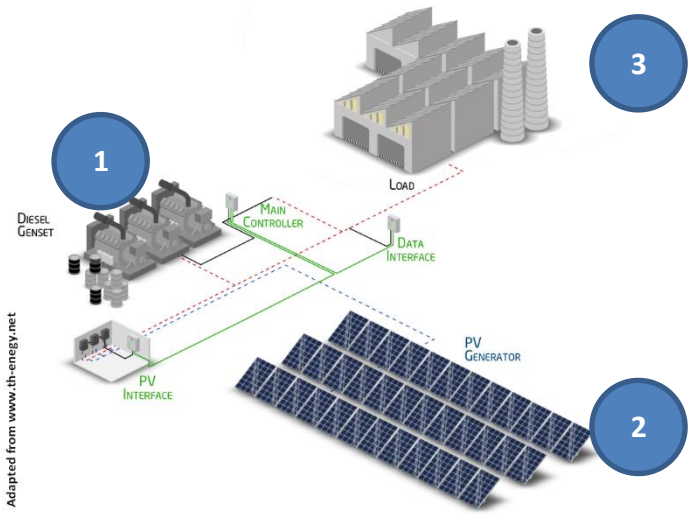


Figure 2.2 Illustration of the Hybrid Energy System (HES)

In this HES, the solar power is used in priority, where the residual demand is committed by the diesel thermal power plant. When the PV production is abundant, some diesel generators are turned off to save fossil fuel. If the time to turn off a generator is too short, the overproduction would be curtailed in this case.

Generally, in a diesel generator-based electrical grid, when there is an unexpected power supply/demand unbalance caused by a lack of energy production or a grid demand overload, the operator can use the spinning reserve to compensate this unbalance. If this action is not sufficient, the grid operator needs to take other solutions rapidly to mitigate the power imbalance. For example, immediately turn on some small and fast generators to balance the energy deficit. And if the power provided by these rapid generators is still not enough, the last solution would be voltage reduction[6] or load shedding[7]. In general, in an energy system without renewable energy, the main cause of imbalance is resulting from the energy demand forecast error or the generator failure, which usually causes a power generation outage.

In addition to the situations mentioned above, energy overproduction and low load situations should be considered, since the overproduction in an off-grid system will rise the network frequency. In these situations, a precise power management action based on diesel generators or PV power plants should be taken. The overproduction could be controlled by curtailing/decreasing the power production sources. For diesel generators, we can reduce the fuel input and generator speed to produce less power. In terms of PV curtailment, it could be achieved by an explicit command to the solar inverter. These two regulation methods can work together or separately, which is chosen in the smart dispatching action.

If we look closer into power management consideration, there are many more constraints to consider. For example, the power supply/demand balance should be ensured on a scale less than 100 milliseconds by industrial electronic devices like circuit breakers. Power electronic aspects, such as the transient balance, flickers, and harmonic issues also need to be considered. However, considering all the micro aspects of the power electronic is not the purpose of this thesis and it requires a huge amount of work that could be another thesis. Hence, in this thesis, I decided to study a steady-state (only the control law) in a macro aspect, which means that once the supply/demand balance is satisfied at a certain timescale, the inner diesel generator electronic reaction to maintain the balance is automatically achieved.

To conclude, the solar forecasting technique which predicts the future PV production is therefore very helpful to prepare the generating unit management. In the following section, a brief introduction of solar forecasting is shown.

2.3 Solar forecasting

There are several types of solar forecasting techniques for different time horizons and spatial horizons. Some work like [23] divide forecast methods into two types: physical models and statistical models. The former is based on the solar irradiation information and analytical equation to model PV production, and the latter is based on statistical or machine learning methods to directly predict the PV production. But mostly, the forecast methods are classified according to the variation of spatial and temporal resolution, such as explained in Figure 2.3.

[6] Voltage reduction: this method works only in a huge electricity network since the impedance is considered as the same, If the voltage drops, the current and the power decrease as well.

[7] Load shedding: One of the methods to maintain the electrical grid supply/demand balance, it means to shut down a part of the load to avoid the blackout of the entire electrical grid, it is always used as a last-resort method.

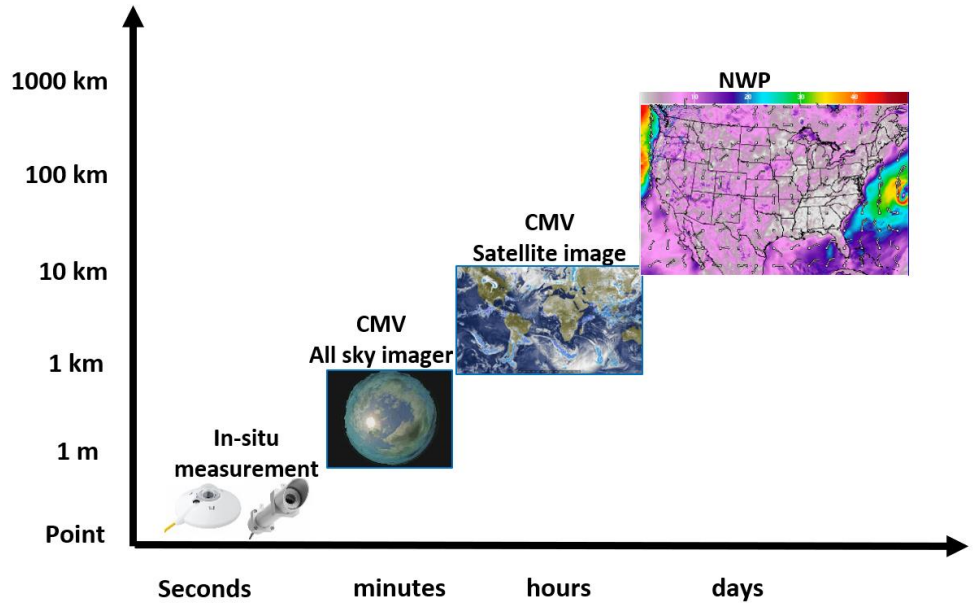


Figure 2.3 Solar forecast methods overview according to the spatial and temporal resolution, inspired from [27] and [28].

2.3.1 Different types of solar forecasting

As shown in the Figure 2.3, there are four major types of solar forecasting:

1) Nowcasting and very short-term forecasting based on irradiance time-series from in-situ measurement. These measurements are usually done by pyranometers, which can directly measure the Global Horizontal Irradiation (GHI). Their temporal horizon is normally under 1-minute. The most popular forecasting models in this range are probably the persistence and autoregressive models.
2) Very short-term forecasting based on All-Sky Imager (ASI). A typical ASI would be a hemispheric camera installed on the ground at the place of forecasting. The temporal horizon of these forecasts is normally from seconds to half hour. Forecasting models associated with ASI are based on image processing, such as Cloud Motion Vector (CMV)-based method, especially the optical flow approach.
3) Short-term forecasting based on satellite images. Satellite-derived estimations of irradiance have a larger spatial resolution than ASI. The temporal horizon of these forecasts usually ranges from 15-min to a few hours. In short-term forecasting, the satellite image is used because it can provide a larger spatial resolution, which could be derived as a longer forecast horizon.
4) Medium-term forecasting based on meteorological data. Largely relying on NWP (Numerical Weather Prediction) models, these forecasting models are developed by some laboratories (e.g. ECMWF, GFS, WRF, etc.) with lots of data and sophisticated mathematical models. It becomes more effective compared to other models with a temporal horizon after 3-4 hours.

Each forecasting technique has its own advantages and weaknesses, and they can be combined or used separately to accommodate the energy system variability. For example, one could use the all-sky imager with hemispheric camera for very-short term production forecasting, the satellite images for short term forecasting and historical data for medium-term production forecasting and consumption forecasting. In order to compare the performances of different forecasting techniques, several metrics are used. The following paragraph shows some classic metrics typically used in solar forecasting.

2.3.2 Evaluation of solar forecasting methods

2.3.2.1 Deterministic forecasts

In the literature, lots of metrics have been proposed to benchmark solar forecasts (e.g. [27]). Among them, the most used metrics for the evaluation of deterministic solar forecasting skills are the Root Mean Square Error (RMSE), the Mean Bias Error (MBE), and the Mean Absolute Error (MAE). They are defined as follows:

$$(1)\quad RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(I_i - \hat{I}_i)^2}$$

$$(2)\quad MBE = \frac{1}{N}\sum_{i=1}^{N}(I_i - \hat{I}_i)$$

$$(3)\quad MAE = \frac{1}{N}\sum_{i=1}^{N}\left|I_i - \hat{I}_i\right|$$

Where I is the predicted value of irradiance and Î is the measured reference value.

In the equation above, the irradiance I could refer to the GHI (Global Horizontal Irradiance), DNI (Direct Normal Irradiance), GTI (Global Tilted Irradiance), etc. RMSE is sensitive to the extreme value and can show the anomaly outlier. MBE shows the average bias error of forecast, but in a case the forecast value oscillates above and below the measurement value, the MBE would be small and cannot really reflect the system error. MAE is an adaptation of MBE for this situation by taking the absolute value of the error. But both RMSE and MAE result in a situation that the metric result could not distinguish the over-estimated and underestimated cases, which would lead to different results in an isolated storage-less HES. These metrics need to be normalized before comparison between different models.

The coefficient of determination R^2, defined in (4), and the Pearson correlation coefficient ρ, defined in (5) are also frequently used to measure the accuracy of the prediction model by normalizing the ratio between forecast and observation data. For the coefficient of determination, the closer this value to 1, the better the accuracy of the prediction model is.

$$(4)\quad R^2 = 1 - \frac{\sum_{i=1}^{N}(\hat{I}_i - I)^2}{\sum_{i=1}^{N}(\bar{I}_i - I)^2}$$

The coefficient of correlation of Pearson varies from -1 to 1, the closer this value to 1, the better linear correlated between the forecast and the measurement, which means better accuracy.

$$(5)\quad \rho = \frac{Cov(I,\hat{I})}{\sqrt{Var(\hat{I})Var(I)}}$$

Where $\hat{I}_i$ is the measured irradiance, $\bar{I}$ is the average value of Irradiance, and the function Cov is covariance.

However, even though the coefficient of Pearson is quite useful in describing the linear correlation, but for a case with an extreme outlier, ρ will be much smaller than the real correlation. And for a non-linear but high correlated case, ρ is not suitable to describe the correlation.

2.3.2.2 Probabilistic forecasts

Probabilistic forecast allows estimating different percentile levels of GHI or PV production, these values form an uncertainty range, which is also called the sharpness of probabilistic forecast. And the fraction of the observations is covered by the forecast uncertainty, is called forecast reliability. There are also

performance metrics specific to probabilistic forecasting. Rather than getting a score by comparing between two deterministic values, probabilistic forecast metrics evaluate the sharpness and the reliability of the probabilistic forecast.

The Prediction Integral Normalized Average Width (PINAW) is frequently used as it evaluates the sharpness of probabilistic forecast normalized by the mean value of the predicted variable. It is calculated as the average width of different inter-quantile levels for different confidence levels. It is defined as (6) :

$$(6)\quad PINAW(\gamma) = \frac{1}{N}\sum_{i=1}^{N} \frac{\hat{q}_{i_{1-\frac{\gamma}{2}}} - \hat{q}_{i_{\frac{\gamma}{2}}}}{\bar{y}}$$

Where γ is the confidence level, N is the number of forecasts, $\hat{q}_{i_{1-\frac{\gamma}{2}}}$ is the upper bound of a given inter-quantile around the P_{50} quantile, and $\hat{q}_{i_{\frac{\gamma}{2}}}$ is the lower bound, $\bar{y}$ is the average measurement of GHI or power generation according to the context.

Since PINAW offers a set of values for each forecast model, its average value across the various confidence levels was also calculated as it corresponds to the overall performance of each model and is easier to handle when plotting [29].

PINAW should not be used alone since it evaluates only the sharpness without considering the reliability.

The Prediction Interval Coverage Percentage (PICP) is hence introduced as a complement since it quantifies the reliability through calculating the fraction of observations inside a given forecast confidence level, as shown in (7) . For example, for a confidence level of 50%, the corresponding prediction interval is between P_{75} and P_{25} of a given probabilistic forecast (i.e., the upper and lower bounds centered in P_{50}). If the observation is inside this interval, q_i is 1, otherwise 0.

$$(7)\quad PICP = \frac{1}{N}\sum_{i=1}^{N} q_i$$

$$q_i = \begin{cases} 1, if\ y_i \in Q^{\alpha}(x_i) \\ 0\ if\ y_i \notin Q^{\alpha}(x_i) \end{cases}$$

Where N is the number of observations (and forecasts), y_i and Q^{α}correspond to a given observation and inter-quantile interval, respectively

The Continuous Ranked Probability Score (CRPS) is a popular metric to evaluate these two aspects [30], as it takes into account the reliability and resolution (i.e., the ability for a model to generate case-dependent forecasts) of the forecasting model, as well as the variability of the forecasted variable [31].

Although it can also be calculated through the integration of the Brier Score [31], the CRPS was here calculated based on the difference between the Cumulative Distribution Functions (CDF) of the probabilistic forecast and the observation (measurement), as defined in (6) and illustrated in Figure 2.4.

$$(8)\quad CRPS(F, x_0) = \int_{-\infty}^{\infty} (F(x) - \mathbb{1}(x - x_0))^2\, dx$$

Where F is the CDF of a probabilistic forecast x, such that $F(x) = P[x \geq x_0]$, x_0 is the observation, and $\mathbb{1}$ is the Heaviside function describing the observation CDF.

Since the observation is assumed to have no uncertainty, its CDF of observation jumps directly from 0 to 1 at the observed value (thus, described using a Heaviside function centered in that same value). And the CDF of the probabilistic forecast has a smoother and flatter curve. As shown in Figure 2.4, the left is the probability density function of forecast and observation, the right is the associated cumulative distribution function, the shadow area is the difference between the two CDFs, whereas the CRPS considers the square of such difference.

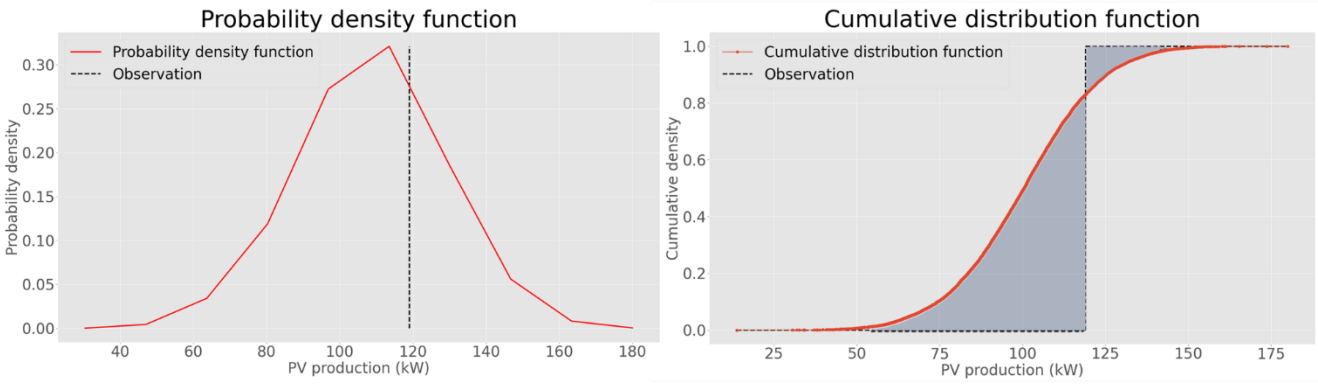


Figure 2.4 CRPS explanation diagram

2.3.2.3 Less common error metrics

The ratio values

Besides these classic metrics, some new metrics have been introduced to provide complementary information compared to these classic metrics by getting afterward statistical value. For example, [32] proposed a new metric, the ratio value s, that aims to evaluate the forecast performance considering the solar variability as well as the forecast uncertainty. The solar resource variability V and the forecasting uncertainty U are first calculated using the formula (9) and (10). Note that U is similar to the RMSE, but it has irradiation at clear sky conditions at the place of denominator as reference value.

$$(9)\quad V = \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left(\frac{I(t)}{I_{cls}} - \frac{I(t-1)}{I_{cls}(t-1)}\right)^2} = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(\Delta k_c(t)^2}$$

Δk_c here is the value change of clearness index during the cloud event, and I_{cls} is the irradiation under clear sky conditions.

$$(10)\qquad U = \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left(\frac{\hat{I}(t)-I(t)}{I_{cls,t}}\right)^2}$$

The ratio value s is then calculated as follows:

$$(11)\qquad s = \frac{V-U}{V} = 1 - \frac{U}{V}$$

S range from 0 to 1, with $s = 1$ corresponding to a best forecasting performance. Even though this metric has its advantage in considering the relationship between the solar variability and the forecast uncertainty, it is highly dependent on the quality of chosen clear sky model [26].

The Temporal Distortion Index

At the application level, the advance or delay of a forecast with respect to the actual value could bring significantly different results for a storage-less HES. [26] reminds that the knowledge of temporal distortion should be considered not only by the grid operator but also by the forecasting designer. Therefore, low temporal distortion should be one of the important criteria to select suitable solar forecasting in storage-less HES.

A metric that can tell the temporal distortion is proposed by [33]. This metric aims to decompose the forecasting error in temporal error and amplitude error, and an index called *Temporal Distortion Index* (TDI) is introduced to quantify the value of temporal distortion. The calculation of TDI is based on a classic algorithm named *Dynamic Time Warping* (DTW) [34], which is also a subbranch of dynamic programming[8].

[26] further decomposed TDI in 3 parts: advance, aligned, and delay. He distinguishes the percentage in advance and in delay of this temporal distortion index with $TDI_{advance}$ and TDI_{delay}.

$$(12) \qquad TDI = \frac{1}{N^2}\sum_{I=1}^{K-1}|(i_{I+1}-i_1)(i_{I+1}+i_1-j_{I+1}-j_1)|$$

$$(13) \qquad TDI_{advance} = \frac{1}{N^2}\sum_{\substack{I=1 \\ i_1 \geq j_1 \\ I_{I+1} \geq J_{I+1}}}^{K-1} |(i_{I+1}-i_1)(i_{I+1}+i_1-j_{I+1}-j_1)|$$

$$(14) \qquad TDI_{delay} = \frac{1}{N^2}\sum_{\substack{I=1 \\ i_1 \leq j_1 \\ I_{I+1} \leq J_{I+1}}}^{K-1} |(i_{I+1}-i_1)(i_{I+1}+i_1-j_{I+1}-j_1)|$$

[26] then proposed a new metric, the *Temporal Distortion Mixed* (TDM), which is calculated as follows:

$$(15) \qquad Advance = \frac{TDI_{advance}}{TDI_{total}} \qquad Delay = \frac{TDI_{delay}}{TDI_{total}}$$

$$(16) \qquad TDM = 2 \cdot Advance - 1 = 1 - 2 \cdot Delay$$

TDM close to -1, means that the forecasts are in advance, and inversely TDM close to 1 means that there are delayed. However, the case of forecast with same composition in advance and delay, which means the value of advance and delay are both 0.5, TDM equal to 0, it is still difficult to interpret the real situation is correctly aligned or half in advance and half in delay.

2.3.2.4 Metrics used in this thesis

For our case study, all these metrics have one or more inconvenient since in a storage-less HES, not only the magnitude of the error should be considered, but also the good anticipation of uncertainty and error is important. And these metrics could not provide the information for us to evaluate the practical performance of a forecast. In order to bypass the disadvantage of these metrics to evaluate the forecasting performance, I propose a method that uses the consumption of spinning reserve, load shedding quantity as well as the PV curtailment quantity to show the quantity of error, the detail of this method is developed in chapter 4.

[8] Dynamic programming: an algorithm decomposes a sophisticated question into several simple questions, which can normally significantly improve the system performance and reduce the calculation time.

To be able to evaluate the system performance, besides the metric used for performance evaluation, the notion Unit Commitment (UC) problem, is used to describe the energy system behavior and generate the generators dispatching order.

2.4 Unit Commitment (UC) problem

2.4.1 UC notion introduction

Unit Commitment (UC) problem is a classic mathematical optimization problem, which is fully described by an objective function, system constraints, and decision variables. The objective function can be mono-objective or multi-objective, such as minimizing the dispatching cost and the system emissions, or maximizing the energy generation revenue and the energy system reliability, etc. The optimal solution is the best combination of different generating units, which could be thermal units, hydro units, or renewable generation units.

The main idea of the UC problem is to find the optimal solution with the objective function when meeting a certain demand [35]. Physical and temporal constraints as well as the security of power supply also need to be considered. To clearly illustrate the UC problem, we consider a simple example of a thermal power system. As shown in Figure 2.5, we suppose that there are several diesel generators with different characteristics and that the demand is perfectly known for each moment in the future. Since there are several possibilities of combination of the use (switch on/off, requested power) of generators, the UC problem is how to select the combination of diesel units that minimize the objective function.

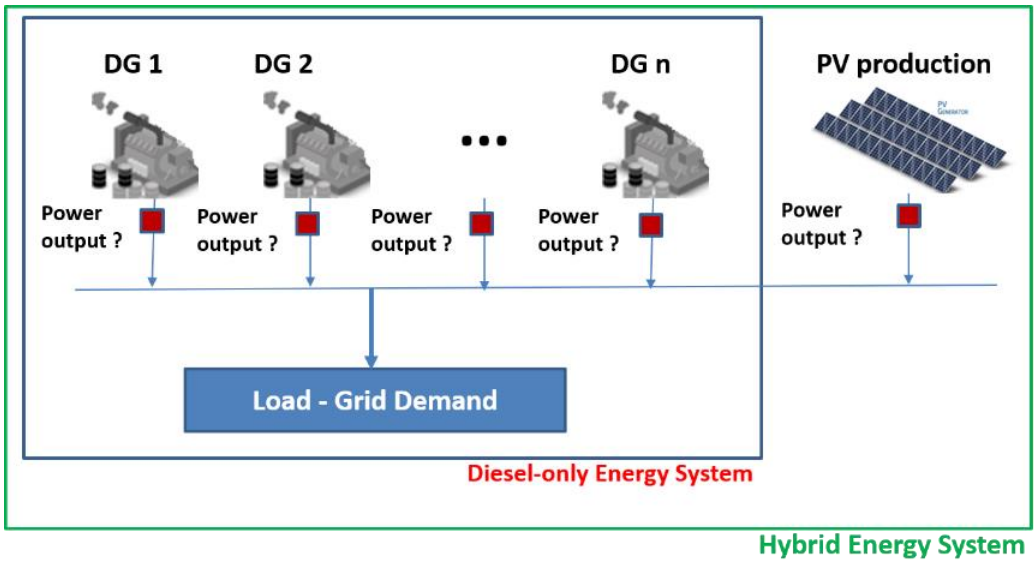


Figure 2.5 Illustration of UC problem

The development and application of algorithms to solve the UC problem is an important concern for the power industry, especially in power system management. The most common use in the electricity field is the day-ahead generating units start planning, which relies on medium-term forecasts to define the priority list of generating units.

The renewable energy penetration rate in world power generation has significantly increased in past years, the grid stability becomes more difficult to ensure [36], [37], especially for off-grid power systems. In hybrid (PV-Diesel) power system, the PV production forecasting error should be considered,

as well as the forecast uncertainty of power demand and the risk of generating unit failure, which are considered as system constraints for the UC problem.

2.4.2 Practical UC problem formulation in a mathematical way

The UC problem formulation in mathematical expression is outlined in the section, which aims to provide a clear overview of the whole energy system and the problem.

2.4.2.1 System overview

As mentioned in section 2.2, the energy system in our case study is composed of three parts: the diesel genset, the PV power plant, and the grid demand. Based on the hybrid power plant functionality, this UC problem could be considered as a mixed integer linear problem optimization with mono-objective function and several system constraints. The associated nomenclatures are detailed in the following section.

2.4.2.2 Nomenclature

Supposed that there are **n** diesel generators, each of them has different characteristics and a PV power plant integrated into the energy system. x_i represents the state of machines, p_i represents the power output of diesel generators. The concrete definitions of each part of HES are shown below:

Diesel generator commitment variables:

GS – Total number of diesel generators in the generator set

n – Subscript indices indicate generator with maximum values of GS

dt – index indicates time and with maximum values of MaxT

N – Number of timestep

$X_{n=1}^{GS}, t \in \{0,1\}$ – Status of diesel generator, a binary value (ON/OFF), which is the integer element of this UC problem

$S_{t,n}$ – Pulse Start-up of diesel generator

$SD_{t,n}$ – Pulse shut down of diesel generator

a_n, b_n, c_n, d_n, e_n – Price coefficient associate generator on run, fuel consumption, spinning reserve, start-up costs, and shut down costs

$SR_{t,n}Positive$ – Minimum total positive Spinning Reserve requirement for grid security

$SR_{t,n}Negative$ – Minimum Total negative Spinning Reserve requirement for grid security

PD_t – Active power demand

DR – Maximum power down rate (MW/h)

UR – Maximum power up rate (MW/h)

ΔT^{01} – Time-lapse to start generator at cold condition

$Uptime_{min}$ – Minimum uptime of generator before shut down

ΔT^{10} – Time-lapse to stop generator

$Downtime_{min}$ – Minimum downtime of generator before restart/commit

$P_{min,n}$ – Minimum active Power Output of generator

$P_{max,n}$ – Maximum active Power Output of generator

$P_{t,n}$ – Active Power Output of generator

C_{GS} – Final cost of power generation by generator set

$P_{max,nm}$– Nominal power of the biggest online generator

$startup, n$ – startup pulse of generator n

$N_{startup,n}$ – Limit number of startups of generator n

$shutdown, n$ – shutdown of generator n

$N_{shutdown,n}$ – Limit number of shutdowns of generator n

Photovoltaic power plant variables:

$P_{PV}max$ – Max active power after inverter

$P_{PV,t}$ – Available active power after inverter at instant t

$C_{PV}(P)$ – Hourly cost of active power generation

C_{PV} – Final cost of PV power generation

E_{PoA} – Irradiance level received on plan of array

T_c – PV cell temperature

Common variables:

C_{Buffer} – prohibitive cost of load shedding or PV production curtailing

C_{final} – Final cost of power generation

2.4.2.3 UC problem Objective

The first industrial and economical objective in the HES of our case study is to minimize the fuel energy production cost. This final cost is composed of three parts:

The cost of diesel generator set:

(17) $C_{GS,n}(t) = \sum_{i=t}^{N}\sum_{n=1}^{GS}(a_n * X_{t,n} + b_n * P_{t,n}) + \sum_{i=t}^{N}\sum_{n=1}^{GS}(c_n * SR) + \sum_{i=t}^{N}\sum_{n=1}^{GS}(d_n * S_{t,n}) + \sum_{i=t}^{N}\sum_{n=1}^{GS}(e_n * SD_{t,n}), \forall t$

The cost of PV energy production:

(18) $C_{PV}(t) = (C_{PV}(P) * P_{PV,t}), \forall t$

The cost of penalty action:

(19) $C_{Buffer}(t) = C_{PV\ curtailment}(t) + C_{diesel\ curtailment}(t) + C_{load\ shedding}(t) + C_{non-spinning\ reserve}(t)$

The final objective is to minimize the sum of these three costs:

$$\arg\min(C_{final}) = \arg\min\,(C_{GS,n} + C_{PV} + C_{Buffer}) \quad (20)$$

More precisely, is to minimize the final cost since instant t and upon N timestep, which can be written as:

$$C_{final} = dt \sum_{i=1}^{N-1} \sum_{n=1}^{GS} (C_{GS,n}(t+i)dt + C_{PV}(t+i)dt + C_{Buffer}(t+i)dt) \quad (21)$$

2.4.2.4 System Constraints

The system constraints are necessary to achieve a realistic simulation. In our case study, the system constraints are principally upon generating units and PV power plant.

1. Power balance between demand and production

The first constraint is to ensure the power supply/demand balance: the sum of diesel genset production (p) and PV production (P_{pv}) should equal to the grid demand ($demand$) at every instant:

$$\sum_{n=1}^{GS} P_{t,n} + P_{pv,t} = PD_t\,, \forall t \quad (22)$$

2. Power output limit

For diesel generator:

Each generator has its own optimal power output range (eg. From 40% to 100%), and $P_{t,n}$ should stay in the range :

$$P_{min,n} < P_{t,n} < P_{max,n}\,, \forall t, n \in Gs \quad (23)$$

For PV panel:

Where P_{pv} (PV production) depends on Irradiance E_{poa}, temperature T_c, PV panel characteristic, and power output percentage ε_k, the formulation could be simplified as:

$$P_{pv,t} = f\left(E_{PoA,t}, T_{c,t}\right) * \varepsilon_k\,, \forall t \quad (24)$$

3. Limit in Spinning reserve

In a storage-less HES like our case study, the spinning reserve is the most important regulation method. The limit definition in spinning reserve (SR) is therefore very important. And I classify them as positive one and negative one.

The **positive SR** is the reserved capacity of the rotating generator, which aims to compensate the energy deficits in avoiding the overload cases, such as negative solar forecasting errors (-), generator failures (-), and negative demand forecasting errors (-), etc. The positive spinning reserve constraint $SR_{t,n}Positive$ can be expressed as follows:

$$P_{max,nm} < SR_{t,n}Positive < \sum_{i=1}^{n} \left(P_{max,t} - p_{n,t}\right)\,, \forall t \quad (25)$$

The **negative reserve** is the capacity of PV production to be curtailed and generating units to decelerate. It aims to avoid overproduction cases, such as the positive solar forecasting errors (+) and positive demand forecasting errors (+), which both mean that the total production exceeds load demand,

resulting in a frequency rise. The negative spinning reserve constraint $SR_{t,n}Negative$ can be expressed as follows :

$$\sum_{i=1}^{n}\left(p_{n,t} - P_{max,t}\right) > SR_{t,n}Negative\ , \forall t \quad (26)$$

4. Limit in ramp rate

The power ramp rate is an important element of the electrical grid flexibility. The larger the ramp rate authorized, the more flexible the system is. The constraint of generating units ramp rate means that the power output change should not exceed the up/down rate limit between each two timesteps:

$$-DR * dt < \sum_{n=1}^{GS} P_{t,n} - P_{t-1,n} < \mathrm{UR} * \mathrm{dt}\ \ , \forall t \quad (27)$$

5. Limit in minimum up/downtime

All electricity generators need time for warm-up and excitation, but also have a constraint on the minimum up/downtime. The minimum uptime is the minimum time that a generator should be on before shutting down.

$$\Delta T^{01} > Uptime_{min} \quad (28)$$

And the minimum downtime is the minimum time that an off-state generator should wait before committing to generate power.

$$\Delta T^{10} > Downtime_{min} \quad (29)$$

6. Limit in startup/shut down number

For real-world applications, each startup will shorten the lifetime of the generating units, it is therefore necessary to limit the number of startup/shut down for each day. This limit is written as:

$$\sum_{i=1}^{t} startup, n < N_{startup,n}, \forall t, \forall n \quad (30)$$

$$\sum_{i=1}^{t} shutdown, n < N_{shutdown,n}, \forall t, \forall n \quad (31)$$

7. Initial conditions

In the case study of this thesis, all the generating units would be turned on at night to meet the almost constant grid demand, which means that the initial state of all generators would stay on at the beginning of the day. And for each intra-hour simulation after the initial simulation, the state of the machine will be kept for the next optimization.

8. Limit on generator dispatching

In our case study, there are several types of diesel generators, some could be started via PLC (Programmable Logic Control) which is called automatic, and some need to be operated manually. And they have different startup /shutdown times and costs. According to the feedback of the grid operator, each startup of the machine will shorten the machine's lifetime. In the simulation, a huge startup/shutdown cost is therefore given to the non-automatic generator, to avoid the frequent startup/shutdown operation.

$$d_n, e_{n_{non-automatic\ generator}} \gg activity\ cost(\ b_n) \quad (32)$$

The following expression is also used to minimize the state change of diesel generator to reduce the machine usage, and to minimize the number of machines at state ON to reduce the cost.

$$\min_{i=\{1\dots n\}}\left(\sum_{n=1}^{GS}|Xi(t)-Xi(t-1)|\right),\forall t \quad (33)$$

$$\min_{i=\{1\dots n\}}\left(\sum_{n=1}^{GS}Xi\right) \quad (34)$$

These two constraints are achieved through minimizing the final system cost.

Once the objective function and the system constraints are clearly defined, the next step is to select and apply a suitable optimization algorithm to solve the optimization problem. The following section aims at classifying the existing techniques and showing an overview of the state of the art of optimization methods for the UC problem.

2.4.3 Overview of optimization methods for solving UC problem

UC problem itself belongs to the category of Mixed Integer Problem (MLP) for combination optimization, it could be a mixed integer non-linear problem (MINLP) if we consider the startup part of generating unit, where the specific fuel consumption and the power output is not linear. For example, [38] said that a UC problem could be described as a combination of non-linear problems with economic constraints and pure integer non-linear problems by using Dantzig–Wolfe decomposition (Benders Decomposition) method. But in many applications, it is usually treated as a Mixed Integer linear problem (MILP) because it is much easier to solve and we can bypass the non-linear part by adding startup constraint. Or like [39] who treat the UC problem with a non-linear fuel cost by making an optimal linear approximation.

UC problem has been studied for several decades. In the literature, there exist several branches of methodology for solving the UC problem. In the review articles [40] and [41], the optimization methods are roughly divided into two parts: the classical/mathematical algorithms and the Heuristic algorithms.

The common point of classical/mathematical algorithms is that they are all based on mathematical foundations and are set to find out the global optimal solution. In general, for a small-scale problem, with few variables and few constraints, classic optimization methods such as MILP (Mixed Integer Linear programming), DP (Dynamic programming) and Branch and bound, are efficient enough to solve the problem and return a globally optimal result. However, when the number of variables and constraints increase, which makes the problem become a large-scale optimization problem, classic optimization problems sometimes may suffer from the computational time and computational space.

Heuristic algorithms are therefore introduced to solve large-scale NP problems, which means the problem is complex and could not be solved in a polynomial time. Heuristic algorithms are mostly based on human experiences or limited knowledge. For example, instead of spending huge computational and time resources to get slightly better results by comparing all the possibilities, heuristic algorithm is limited in its search zone according to the experiences to get an acceptable local optimal result. Their main idea is to start from a random feasible initial solution and then use an iterative improvement strategy to find a reasonable near-optimal solution in an acceptable computational time. The heuristic method is efficient and satisfying in solving large-scale problems; however, since heuristic algorithms are based on limited experience or incomplete information, the quality and performance of these methods are thus mainly depending on designers' experience and specific issues. Furthermore, the final solution provided is not a real globally optimal solution, and it needs a base of random feasible solutions to iterate, which takes time to prepare. Some heuristic methods like ANN (Artificial neural network) not only needs a minimum time but also a base training

dataset to train a model before it becomes useful, which means it might be hard to use for a real-time application. And the more the calculation time is short, the more likely the result is far from the optimal solution.

In this section, two types of commonly used methodologies are introduced.

2.4.3.1 Classic optimization algorithm

Exhaustive Enumeration and Priority Listing

Exhaustive Enumeration first enumerates all possible combinatorics of generator units, then calculates the cost of each possibility; the one with minimum cost is the optimal solution. Priority Listing is used to classify the generating units in priority order according to certain standards, such as the unit price of power generation. The unit with the lowest generation unit price would be used first and then the second, etc.

Both methods are simple and useful. However, for a large-scale problem or a problem with several decision variables, they are hard to implement and calculate. These methods are only suitable for small-scale simple optimization problems or working with other algorithms. For example, [42] proposed a method based on the improved priority listing and neighborhood search technique for ramp rate constrained UC problems.

Dynamic programming (DP)

DP is particularly strong at dividing complex problems into several sub-problems to simplify the process of solving the whole problem, which allows people to overcome or bypass the non-linear part and the non-convexity of the optimization problem. Thanks to this feature, DP has been widely used in the UC problem field [43].

DP could be considered as an improved version of recursion method, instead of calculating twice the overlapped sub-problem, DP can reuse the result already obtained for the same problem. The most famous and adaptive example would be the computation of Fibonacci numbers as shown in Figure 2.6: by keeping the previous result in the computer memory, the time complexity to compute the n first numbers is reduced from 2^n to n.

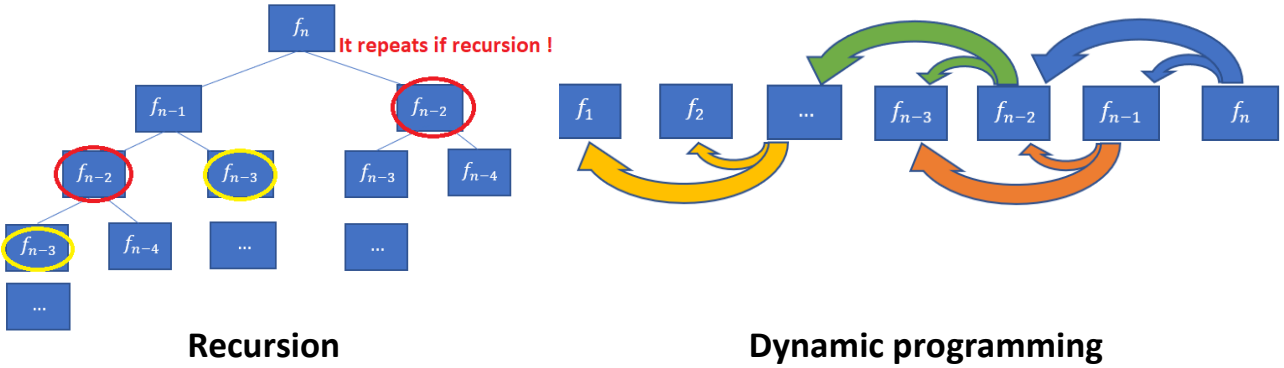


Figure 2.6 Example of computing Fibonacci number, recursion(left), Dynamic programming(right)

However, since the system constraints at different states might be different, this could be a limit when implementing DP algorithm because it is generally hard to realize. In addition, [44] claims that DP has downsides when dealing with high dimensionality problems.

Lagrangian relaxation (LR)

The Lagrangian relaxation method is widely applied in mathematical optimization constrained problems. The main idea is to put the constraint in the objective function to get an approximate Lagrangian subject. In this objective function, a penalty coefficient would be added to the new constraint variable and the final approximate function would be very close to the original problem since the penalty coefficient is hard to perfectly correspond to the real constraint.

This method is simple to use. However, since the final function obtained is still an approximative function compared to the original problem, the final result obtained is often suboptimal. [41] points out that LR method could give fault commitment order for some generating units in sensitivity problems.

Integer and Linear Programming

Linear programming is a classic method for optimization problems with both linear objective function and linear constraint. Below is a simple example of LP:

$$\min \{c^T x : Ax = b, x \geq 0\} \quad (35)$$

However, the solution of linear problem sometimes could be unrealistic since some cases in real world ask for an integer solution such as the generating unit state, which is a binary value and is discrete. Therefore, with an integer constraint, integer linear programming is introduced to find the integer solution of optimization problem. Below is a simple example of IP:

$$\min \{c^T x : Ax = b, x \geq 0, x \; is \; integer\} \quad (36)$$

Mixed Integer Linear Programming (MILP)

For optimization problems with several decision variables, which contain discrete and continuous variables in the same question, Mixed integer linear programming is one of the most suitable solutions. This approach aims to find the optimal solution respecting the discrete variable and continuous variable. Below is a simple example of MILP:

$$\min \{c^T x + d^T y : Ax + By = b, x, y \geq 0, x \; is \; integer\} \quad (37)$$

MILP is also widely used in UC problems for its deterministic characteristic, which could provide a globally optimal result. [45] proposed a MILP method to deal with price-based UC problem. In their work, they also introduce a comparison with Lagrangian relaxation method. [46] proposed a new way to formulate the MILP for UC problems: instead of adding numerous constraints to reduce the search space, his work provides a tight and compact MILP formulation method which could significantly reduce the computational time in a tight search space.

However, even though MILP is suitable for some real operational problems, it is hard to find the optimal solution if there are too many variables. The computation time complexity to obtain optimal solution is exponentially increasing with the number of parameters to handle.

Branch and Bound

The “Branch and Bound” algorithm is an Exhaustive Enumeration variant; the difference is that this method can cut off the useless branches according to the system constraint when it enumerates all the possibilities. The bound of the constraint function is refreshed frequently when obtaining a better solution. This method could be thus much faster than the classic Exhaustive enumeration method. [47]

introduced an approach that groups the generating units to overcome certain disadvantages to solve UC problems.

However, if there are lots of variables, this method requires a huge memory for stocking the branch results before calculation. Furthermore, the method or constraint to cut off useless branch need to be defined carefully to reduce the search space, or it would be similar to the exhaustive enumeration in an extreme case.

Simplex algorithm - Interior Point Optimization algorithm

The main idea of Simplex Optimization is firstly to find out the available zone of the mathematical optimization problem, then set a start point among all the possible convex points. After that, move this point towards the direction with the higher revenue increase or lower cost increase according to the objective function until meeting the system constraints. If the constraints are met, change the direction and repeat the action until getting the highest revenue or lowest cost in the available zone as optimal solution. In most cases, the simplex algorithm is an efficient algorithm, which can find the optimal solution in a limited time. However, its time complexity grows in an exponential scale with the increase of system constraints and decision variables.

Therefore, new polynomial-time algorithms were proposed. The Interior Point Optimization algorithm, also called barrier algorithm, was developed based on the simplex algorithm and Newton's method; the difference is that the interior point optimization algorithm starts from a point inside the available zone instead of a point on the board of available zone and that Newton's method is used to iterate the direction. Unlike simplex algorithm, this algorithm has pure polynomial time complexity even for a large-scale problem. Compared to classic simplex algorithms, this makes it very suitable for solving large linear problems and the convergence speed is much higher.

[48] proposed an approach based on interior point method to solve UC problem. In this work, it shows an encouraging result for large scale UC problem compared to Sub-gradient method and penalty-bundle method in performance in dual maximization.

2.4.3.2 Heuristic algorithm

Apart from the classic deterministic method, Heuristic algorithms are also popular to solve large-scale UC program problems. Heuristic algorithms are mainly the combination of random algorithms and local search algorithms. They are firstly designed to solve the (Non-Polynomial) NP-complete problem (intersection of NP and NP-hard problem), which is unsolvable by a traditional mathematical method in a polynomial time. Fortunately, heuristic algorithms provide the possibility to get an approximation solution instead of getting an expensive exact optimal solution.

The most used Heuristic algorithms are as follows:

1) Machine Learning (ML), including Artificial Neural Network (ANN), Reinforcement Learning (RL), etc.
2) Evolutionary algorithms, such as Genetic Algorithm (GA), Evolutionary programming, and Ant colony search Algorithm.
3) Fuzzy system.

All of these Heuristic algorithms are designed by imitating natural phenomena. Neural networks, for example, try to parallel the optimization process of a human brain; evolutionary algorithms try to

simulate real-world phenomena like how ants find the shortest way when moving food to find out the optimal solution for complicated problems and the fuzzy logic is to flexibly describe the variable characteristic like a human rather than in a binary way to make the system more realistic.

2.4.4 Choice for optimization solver

There is no one single optimal optimization method, the choice of the optimal method will depend on the scale and the nature of the optimization problem. In real-world applications, the behavior of diesel generators production is not 100% linear. However, we can add optimal operation range as a system constraint to bypass the non-linear startup part, which means that in the optimization, the generator would only be allowed to work in its optimal operation range, which has a linear behavior between the fuel consumption and the power output. Hence, the UC problem could be seen as a typical MILP. In our case study, there are more than 20 diesel generators, and this problem theoretically should be considered as a large-scale problem. But since some of the diesel generators have the same characteristic, and most of them are always on, so we can combine the same small generators as one bigger generator. The problem could be hence described in a case with same total genset capacity and energy demand, but less generators, which is easier to solve. For thesis simulation study, classic optimization algorithms that aim to find a global optimal solution are therefore chosen since all the calculation steps are trackable and supported by mathematics.

For this thesis, Python is chosen as the programming language since it is powerful, user-friendly, and especially rich in third-party library language. In terms of optimization solver, among the commercial solvers like “Gurobi" and "CPLEX" and open-source solvers like “GLPK” and “CBC”, “Gurobi” has been chosen because it is the most powerful solver in the current state [49], [50] and it is free for academic use. Inside “Gurobi”, there are two main algorithms available: barrier algorithm and simplex algorithm, to solve continuous problems and mixed integer problems. The barrier algorithm, also called interior point method as explained in section 2.4.2.3, is good at solving large-scale and difficult models. However, it sometimes suffers from numerical issues, which is a general name given for the situation that the results of an optimization problem are either erratic, inconsistent, unexpected, or plain poor performance. In these cases, simplex algorithm, one of the most used algorithms for linear problem [51], could be used as an alternative since it is much less sensitive to the numerical issue. But in the cases without numerical issue, it does not perform as well as the barrier algorithm. Under these two main algorithms, “Gurobi” provides 6 optimization methods for user to choose:

- “automatic”,
- “primal simplex”,
- “dual simplex”,
- “barrier”, “concurrent”,
- “deterministic concurrent”,
- “deterministic concurrent simplex”.

When the practical problem is modeled, I iterate all the optimization methods of “Gurobi” to find the fastest solution for a typical one-day simulation with 15-min resolution, 1-hour forecast lead time, 30-min update frequency, and 47 cycles in total, the detail of the simulation is developed in the methodology chapter. After comparing the converging speed, the speed between different methods is quite close.

One-day simulation with 15-min resolution, 1-hour forecast lead time, 30-min update time, and 47 cycles in total	
Primal simplex	**Dual simplex**
3.68 seconds	3.69 seconds
Parallel Barrier method	**Concurrent optimization**
3.99 seconds	3.81 seconds
Deterministic concurrent	**Deterministic concurrent simplex**
3.91 seconds	3.79 seconds

Table 2.1 Result of different optimization methods based on simplex algorithm

As explained in the beginning of this section, the suitable choice of the algorithm is totally depending on the problem to be solved; different types of optimization methods with different problems could have different results. For example, the dual simplex is good at solving mixed integer problems with root relaxation, the barrier method is for quadratic programming, and concurrent optimization is for linear problems without knowing the problem scale. Finally, the parameter of optimization method is set as "automatic", which first starts to solve the optimization problem with Simplex and Barrier algorithm, if the solving time is over certain level, other concurrent algorithms would be used to solve the problem, and the quickest algorithm is hence selected. In one word, the program could smartly choose the suitable algorithm for the corresponding problem.

2.5 The major difficulties of building a storage-less HES

Even though the previous sections describe a problem that seems easy to understand and to tackle, there are always some unavoidable difficulties, which are principally stochastic and sometimes difficult to properly forecast.

1) The variability of solar resources induced by clouds and/or water vapor, aerosols, etc. that raises difficulties in solar forecasting.
2) The variability due to different components of energy production system. For example, failures of diesel generators, soiling dust on the PV panel, dysfunctions of PV panels and inverters, etc.
3) The variability of the energy demand includes stochastic aspects.

Among these variabilities, the solar variability is the one that affects most the energy system stability and is the one that has the biggest interest to study since it could be less stochastic and has a certain rule to follow. Solar variability and solar predictability are discussed in the following chapter.

Chapter 3

3. The solar resource, its variability and how to deal with it

Contents

SUMMARY OF CHAPTER

This chapter introduces some essential points of solar resource and solar variability, which is the main challenge when using solar energy. First, the composition of solar resource that arrives on earth is explained. Then, the nature of solar variability, such as the different reasons and sources of solar variability are explained.

To better understand the solar variability, some methods of solar variability characterization are introduced and an analysis of solar variability is done as a demonstration of solar variability. A solar variability analysis work is also done, which uses the transition probability to describe the variability instead of classical statistical metrics like Root Mean Square Error (RMSE), and Mean Absolute Error (MAE). This work could not only be used to help to understand the variability, but also could be used as a reference level last-resort solar forecast approach.

Then, we discuss the notion of predictability. As a feature that determines the solar forecast performance, predictability is not only related to the solar variability, but also depends on the forecast method and the observable data that we have. In general, a geometry-induced variability and periodic variability are easy to predict, but a stochastic variability is much harder to predict. The possible consequences of unpredicted stochastic variability is also discussed.

After that, we introduce the principle of the solar forecast approaches used in this thesis work, including their specialty, the reason for selecting them, and how we compute these methods.

In the end of the chapter, we discuss the solution considered in this thesis to against the Hybrid Energy System uncertainty, to dynamically size the spinning reserve of diesel generators to ensure the power supply according to different sources of uncertainty. The way to quantify these uncertainties is also detailed.

3.1 Solar resource variability

3.1.1 Introduction to the solar resource

Not all the solar radiation that arrives at the top of the atmosphere reaches the surface of the Earth. As shown in Figure 3.1, part of it (around 29%) is reflected to the universe and some (around 23%) is absorbed in the atmosphere by water vapor, dust, and ozone. Only around 48% of the solar radiation the top of the atmosphere pass through the atmosphere and reaches the ground [52], [53].

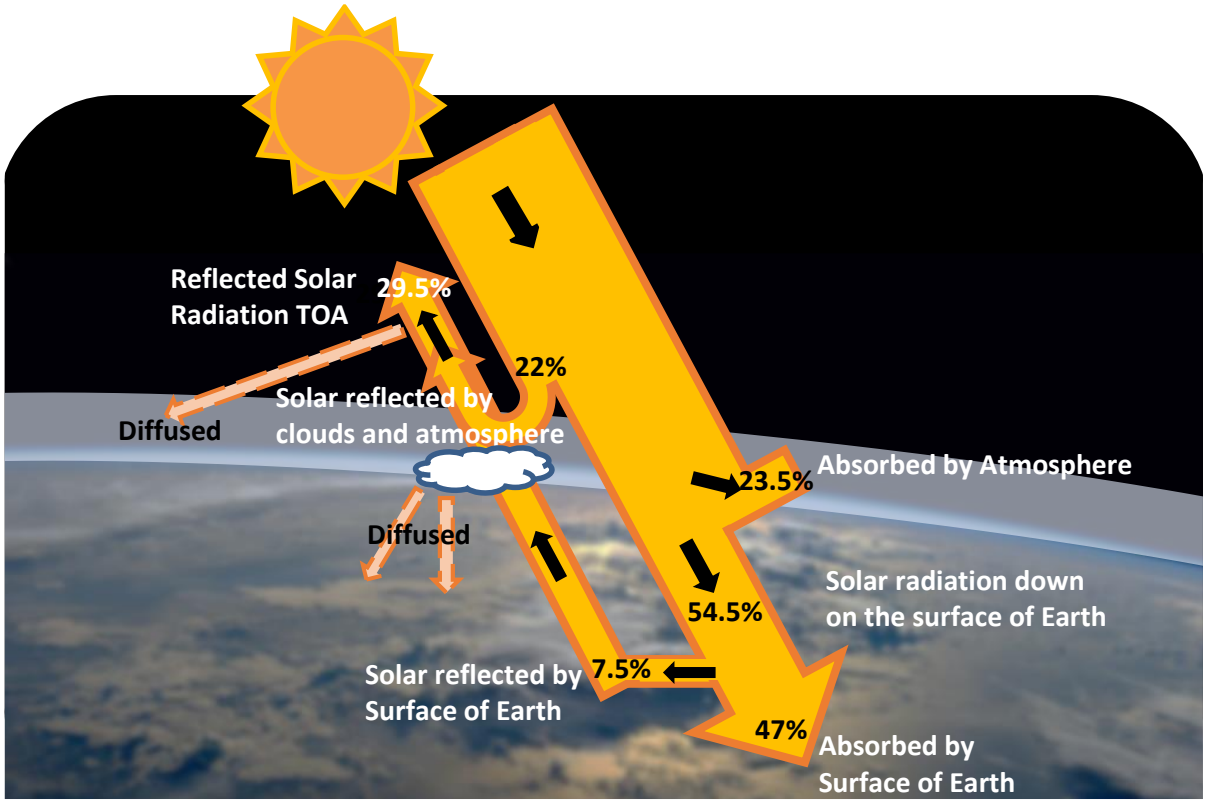


Figure 3.1 Solar radiation arrived on the ground, earth photo credit: NASA/Reid Wiseman

The energy of the solar radiation reaching the surface of the Earth can be harvested thanks to several technologies, such as photovoltaic technique. The photovoltaic process converts the photon energy to electric power. The efficiency of this conversion depends on several elements, such as the incident illuminance, the PV cell inherent efficiency, the PV cell temperature, etc.

In this thesis, only the global horizontal irradiance (GHI) is considered. According to [26], considering the GHI rather than GTI makes the study more independent of the geometry, which means it can better extrapolate to a different area. The GHI can be decomposed into the diffuse horizontal irradiance (DHI) and the direct normal irradiance (DNI) as shown in equation (38) and illustrated in Figure 3.2.

$$GHI = DHI + DNI \cos(\varphi) \tag{38}$$

Where φ is the zenith angle.

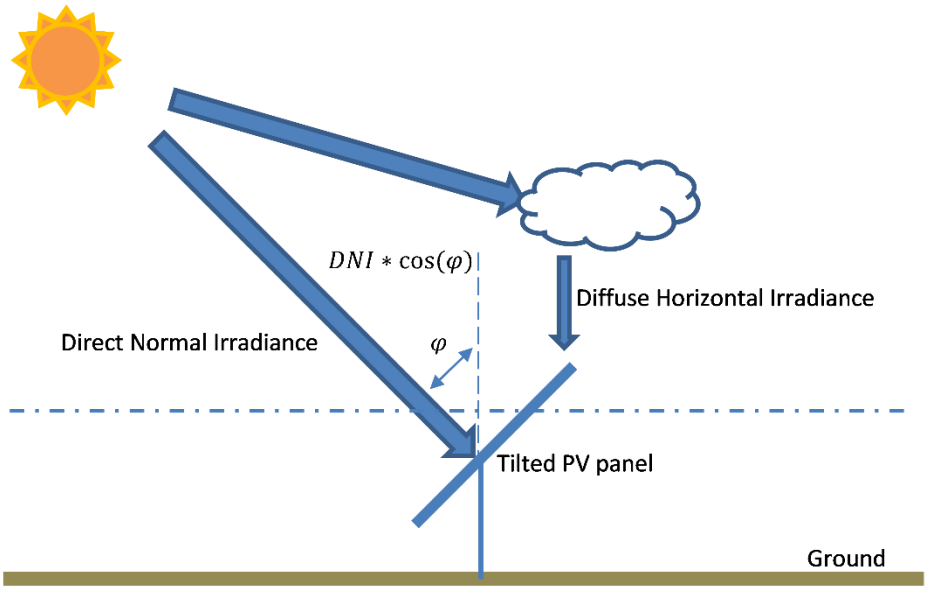


Figure 3.2 Explanation diagram of solar irradiance arrived on the ground

3.1.2 Sources of solar variability

Even though PV energy has lots of advantages compared to fossil fuel energy, and has been widely used, its intrinsic temporal variability characteristic must be considered when using PV energy. The causes of variation are diverse.

3.1.2.1 Geometrical causes

Geometry-induced variations are arguably the most prominent. The rotation of the Earth and its revolution around the sun indeed cause daily and yearly cycles in the solar surface irradiance. However, these variations are regular and thus manageable [54]. Figure 3.3 shows the regular change of irradiance level of our case study from 2005 till 2021.

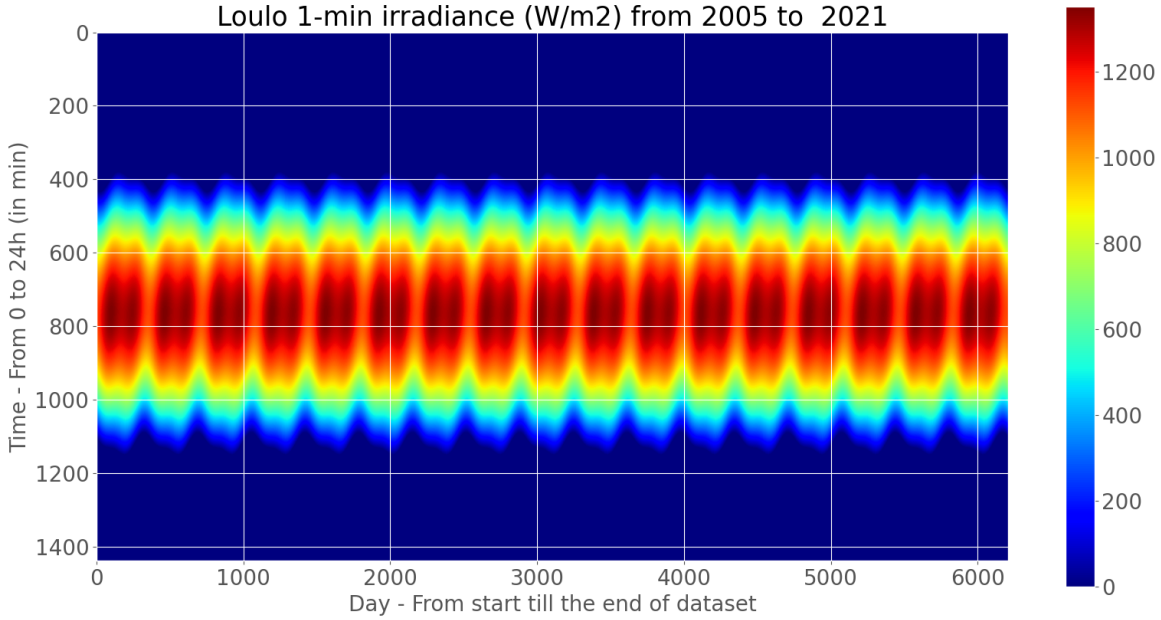


Figure 3.3 Solar resource received at the top of atmosphere

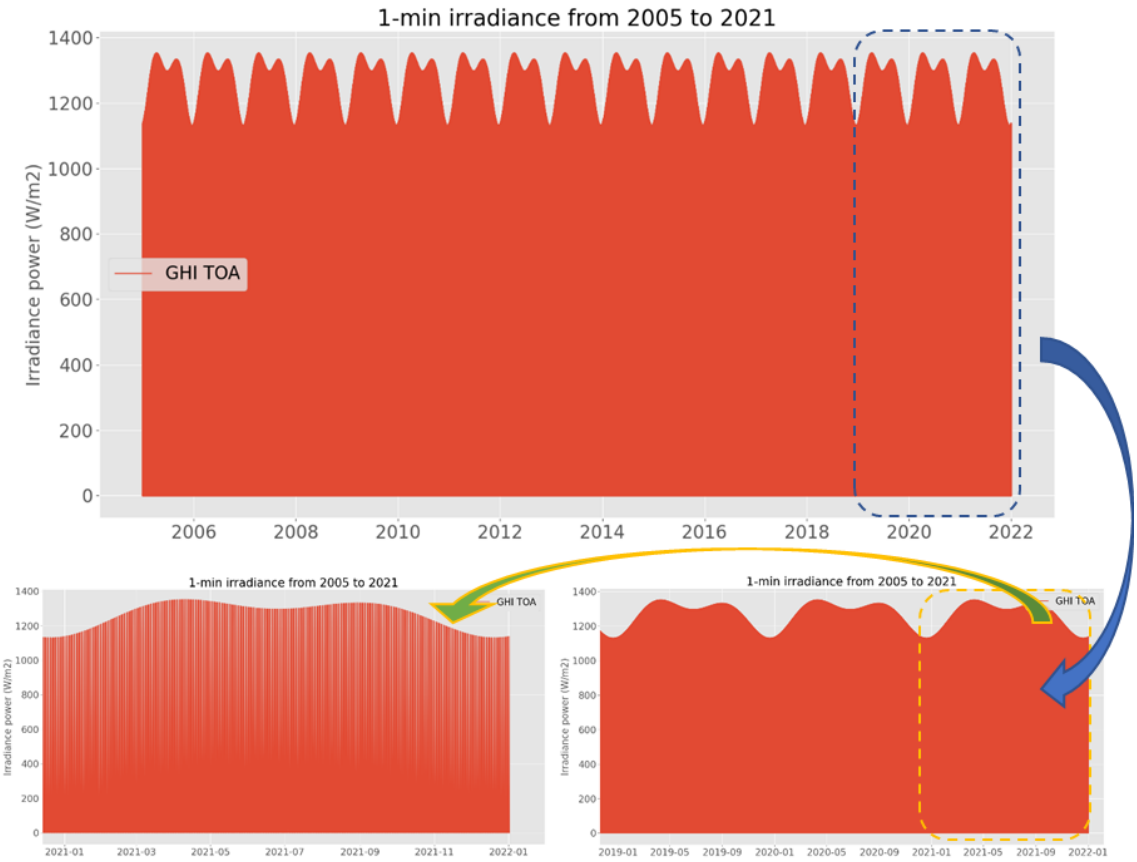


Figure 3.4 The variation of solar resource received at the top of atmosphere of case study located in Mali

Figure 3.4 clearly show that the solar irradiance of our case study located in Mali, arrives at the top of the atmosphere, noted GHI TOA, which has its variation but periodical and predictable, its detail is shown as Figure 3.5.

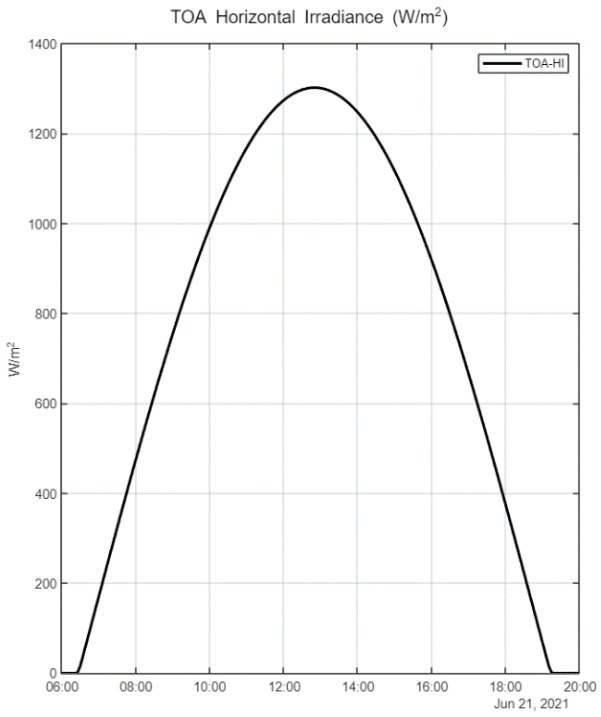


Figure 3.5 Detail of horizontal irradiance TOA of case study located in Mali

3.1.2.2 Clear sky irradiance variability

The GHI reaching the ground under clear sky conditions can be modelled and is often used as reference. In the literature, there exists several clear-sky models, such as [55], [56], [57], [58], [59]. Some of them like e.g. the European Solar Radiation Atlas (ESRA) model [60] do not consider the aerosol and water vapor, while models such as Mc-Clear model [59] account for ozone, water vapor and aerosol optical depth.

Figure 3.6 shows for an aerosol-free and dry clear sky situation; the GHI CLS is quite close to the GHI TOA, which is easy to predict. The reason of this gap is the atmosphere between two positions absorb a little bit of solar irradiance.

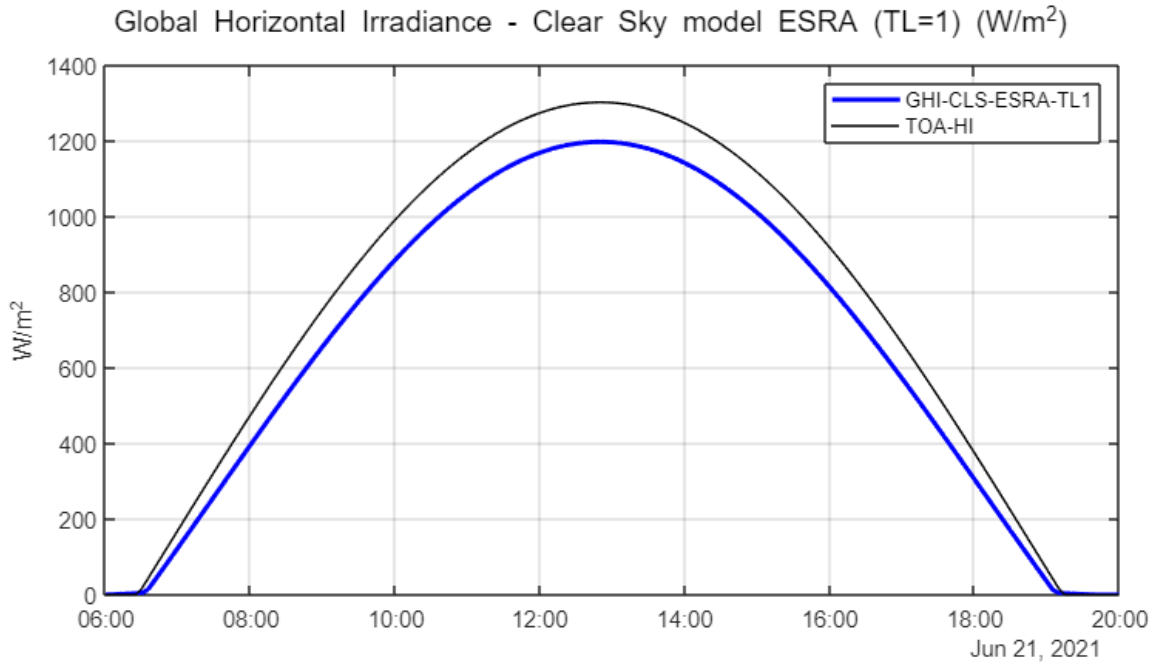


Figure 3.6 Comparison between GHI TOA and GHI CLS ESRA

Figure 3.7 shows the difference between GHI TOA and GHI clear-sky of Mc-Clear model, which, unlike ESRA clear-sky model, considers the aerosol and the water vapor in the atmosphere. We see that there is a difference between these two GHI estimations even in a cloud-free situation, which means that the cloud coverage is not the only driver of solar variability. The diffusion and absorption processes due to the water vapor, ozone and aerosol in the atmosphere are another unavoidable element that create the solar variability. Compared to cloud event, in a normally stable area without important environmental issue, water vapor and the aerosol have much less impact in terms of solar variability, it reduces the DNI by changing a little bit the transmittance of the atmosphere, but it could maybe increase the DHI in the same time. Hence, since the GHI is composed by DHI and DNI, the final GHI remains approximately the same in a clear sky situation and a considerably large area. However, even though the variability in GHI is not that huge, but in terms of BHI, this variability is quite important and has to be considered for the energy system management.

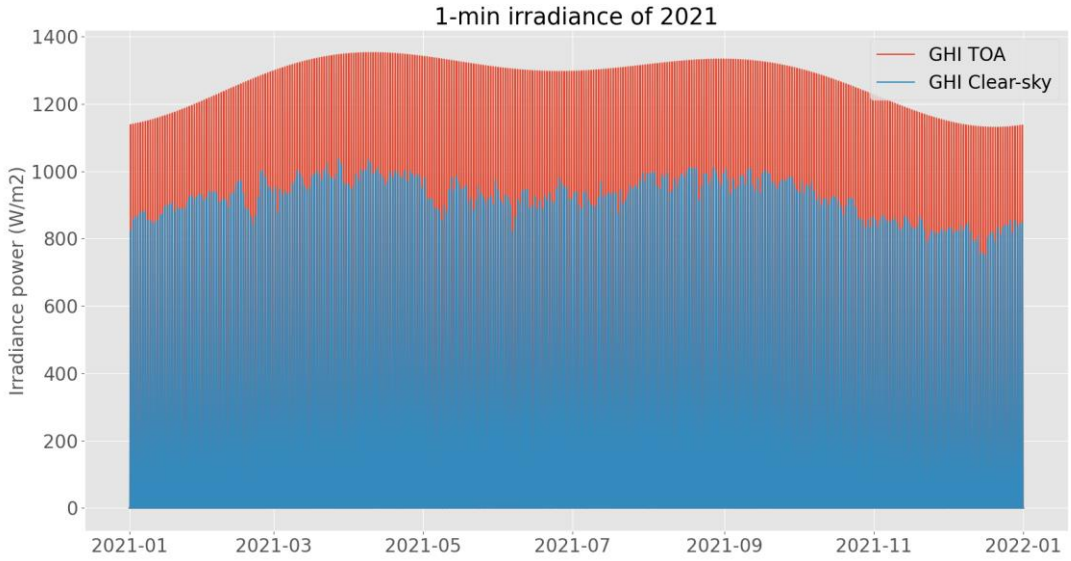


Figure 3.7 Solar resource received at the top of atmosphere and on earth in a no cloud situation

However, even though the essential elements are considered in these advanced clear sky models, due to the imperfect measurement and simulation model, there are still some differences between the simulated and measured GHI ground in clear sky situation, as shown in Figure 3.8. This kind of small errors is also considered as one of the sources of solar variability.

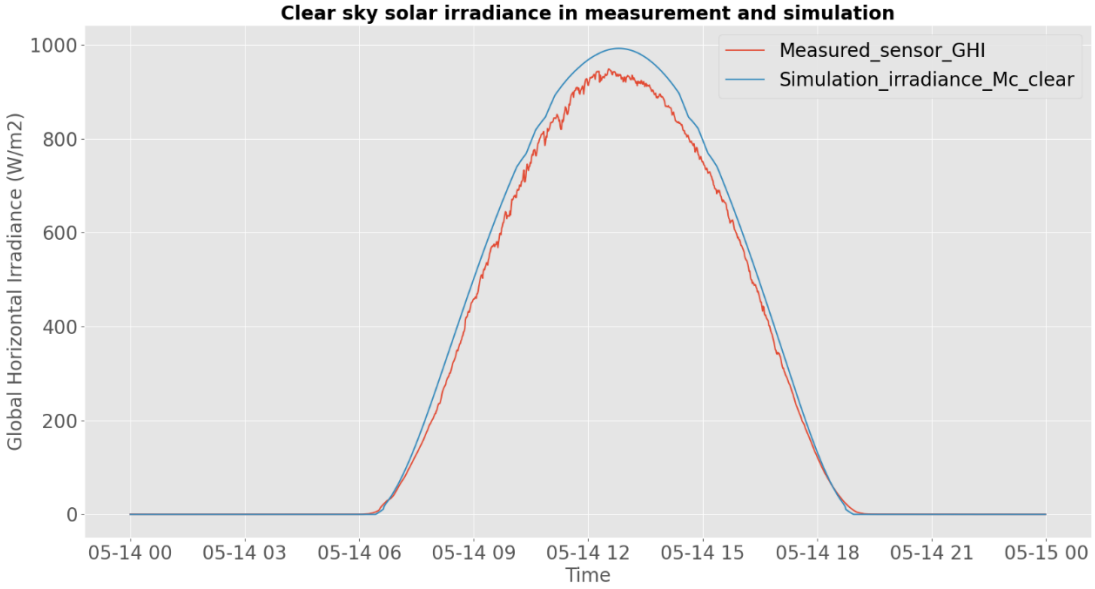


Figure 3.8 Difference between the simulated and ground measured GHI in clear sky situation

3.1.2.3 Stochastic variability

The variation of meteorological conditions like cloud coverage, sandstorm, dust aerosol, etc. also contribute to the solar variability. For a given point, these events can significantly change the solar resource level. Figure 3.9 shows the difference between GHI Clear-Sky and GHI ground measurements, which is mainly induced by the meteorological conditions. Among all these reasons, clouds have the highest impact on the final variability [61], [62]. However, the impact of the cloud on solar irradiance is hard to quantify since the thickness and the shape of the cloud are difficult to know, and it can not only absorb but also reflect the sunlight [63].

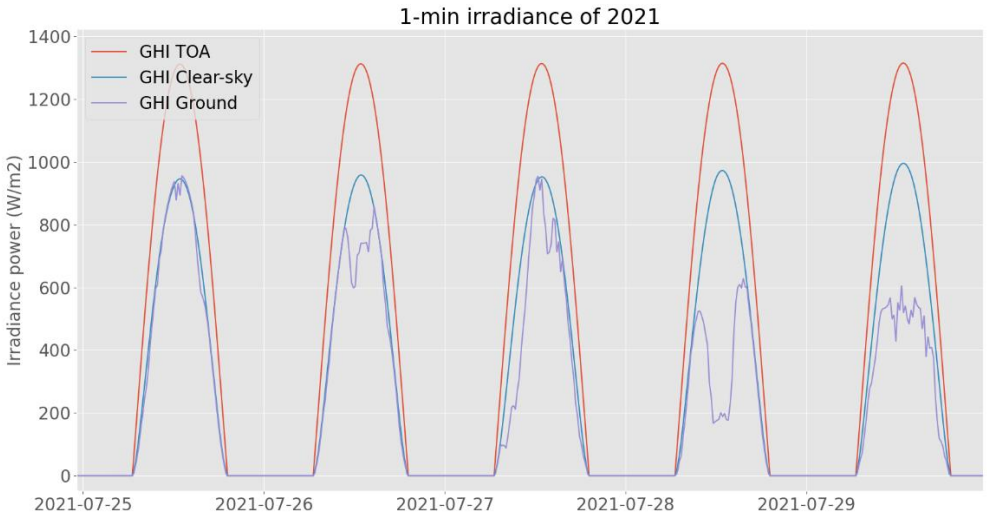


Figure 3.9 Solar resource received at different level of height and different sky situations

3.1.2.4 Impact of spatial and temporal averaging on variability

[54] indicates that the solar variability could largely be mitigated if the temporal and spatial scales could be enlarged to a much longer and larger level. For example, average value of 1-year solar irradiance data has much less variability than the 1-min data. Similarly, the average value of whole solar irradiance arriving on earth has much less variability than a single point. To study the variability of a point, like in our case study, a smaller temporal resolution should be used, since it provides more detail about the variability. As shown in Figure 3.10, GTI from Helioclim3 Version 5 (HC3v5) has a coarser temporal resolution compared to the measurement data. Even though the trend of the solar variation is mostly predicted by the HC3v5 data, due to this less precise temporal resolution, the actual variability is not shown.

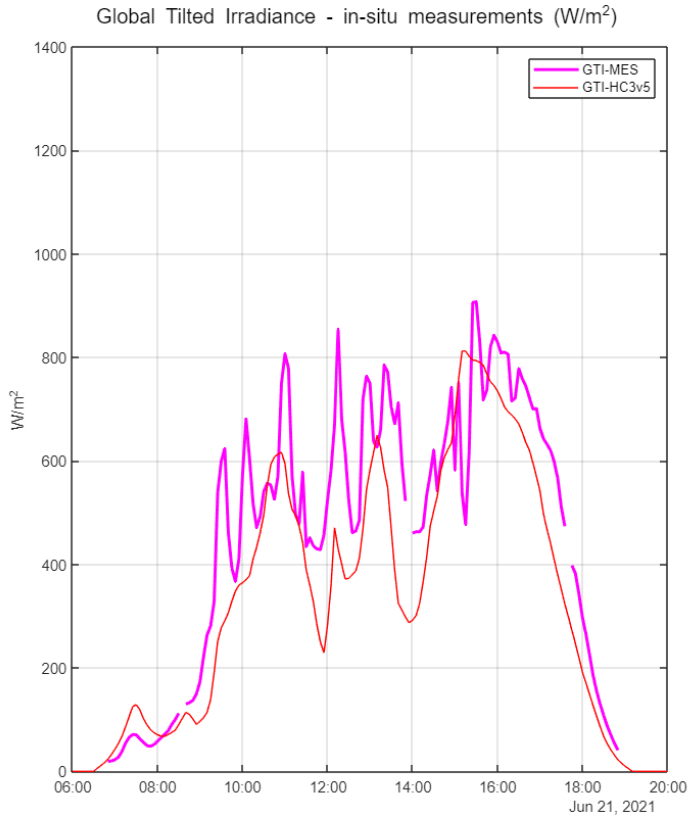


Figure 3.10 Different input data with different temporal resolution

In the literature, some works define the solar variability as the variation of the ratio between the GHI ground measurements and TOA GHI, or the ratio between GHI ground measurements and clear sky GHI [54], which is mostly chosen since it could bypass the solar geometry effect and quantify the solar variability in an index called clear sky index k_c, the ratio between the GHI ground measurements and the clear sky GHI, ranges between 0 and 1, defined as (39) .

$$k_c = \frac{GHI_{Ground}}{GHI_{Clear\ sky}} \quad (39)$$

where 1 means clear situation and 0 for totally cloudy situation.

3.1.3 Statistical characterization of solar variability

In the literature, there are several methods to characterize the solar variability. In this thesis, we follow the work of [64]. The authors of this work propose a method of sky state characterization, which classifies the sky state in 4 types (clear, slightly cloudy, cloudy and variable) by calculating the value k called the index of serenity and the value of a moving function, defined as (40) and (42) , respectively. The method is illustrated in Figure 3.11.

$$k(t) = \frac{|x_{observ}(t) - x_{CLS}(t)|}{x_{CLS}(t)} \quad (40)$$

$$MA = \frac{1}{N_m} \cdot \sum_{i=0}^{N_m - 1} k(t-i) \quad (41)$$

$$MF = \sum_{i=0}^{N_m - 1} |MA - k(t-i)| \quad (42)$$

Where t is the temporal resolution, x_{observ} is the observation, x_{CLS} is the clear-sky model [59]*, N_m is the total time steps used to calculate the Moving average.*

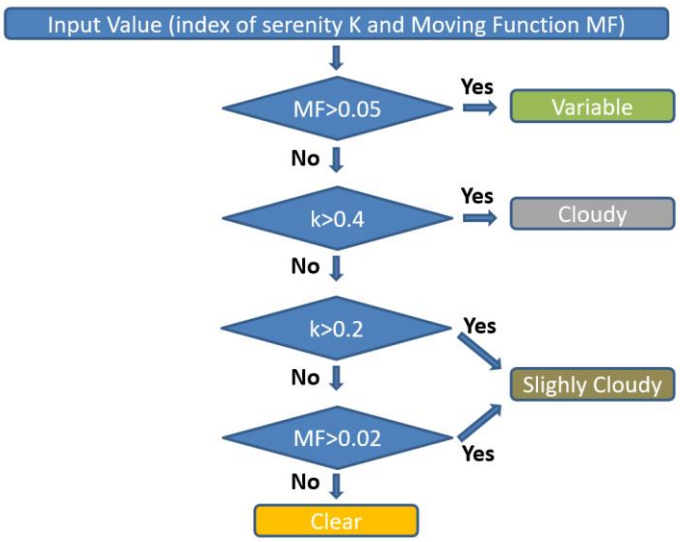


Figure 3.11 Reproduced diagram of sky state classification method proposed by [13]

To better understand the solar variability at our site of interest, we combine this characterization method with an analysis of the transition between the various sky states, using transition matrices. Firstly, a 2-year long 1-min resolution irradiation data is used, and this sky-state classification method

is applied. We thereby classify the whole timeseries into 4 types: Variable, Cloudy, Slightly Cloudy and Clear. The data is then grouped in 5-min blocks; we found that there are 31 possible (i.e. with non-zero probability) combinations of 4 classes within a 5-min block. They are illustrated in Figure 3.12.

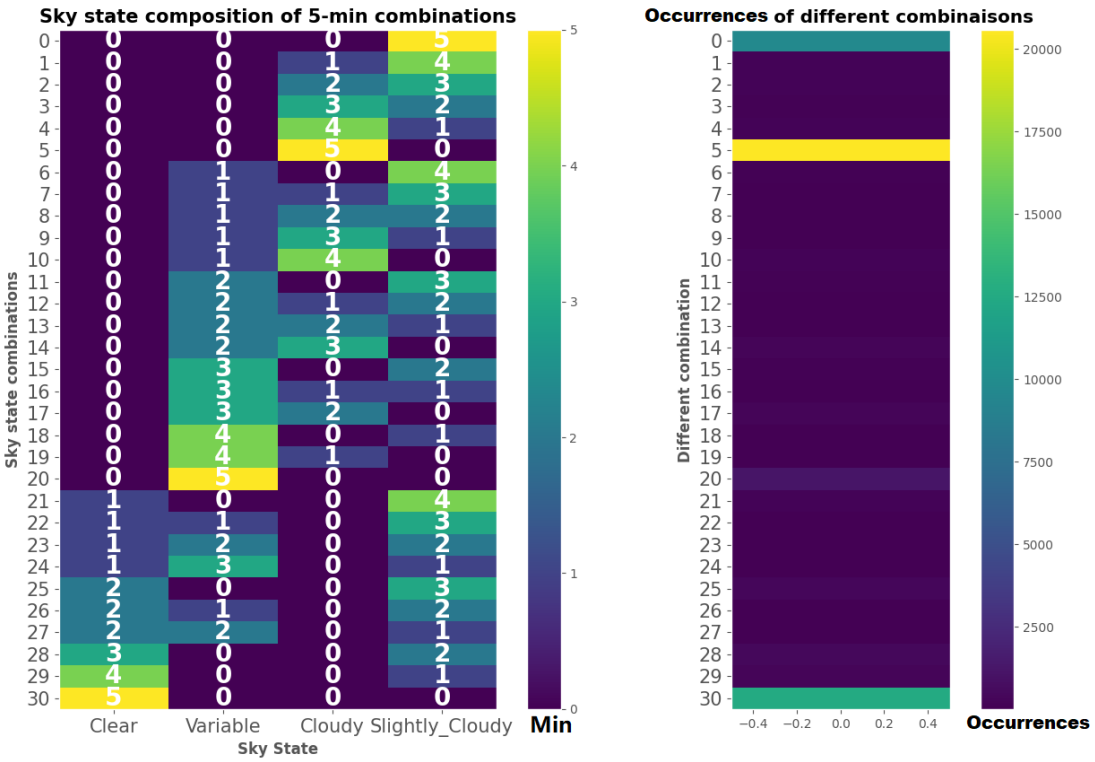


Figure 3.12 Combinations of sky states in 5-mins block of 2-year long 1-min resolution data

Many of these 31 combinations have low appearances; hence, after filtering the rare sky state combinations whose probability of occurrence is lower than 0.1%, there are 6 combinations left, as shown in Figure 3.13. From this figure, it is not difficult to see that the solar variability inside the 5-min block of the example area is not high.

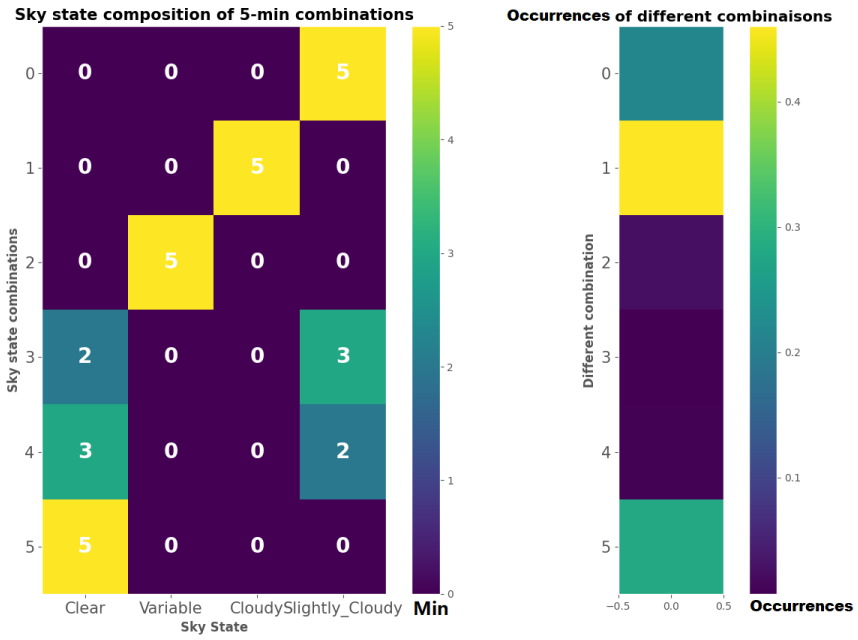


Figure 3.13 Combinations of sky states after filtering rare cases

To have a better image of the global variability, we compute the transition probability between sky states and represent it in a transition matrix, shown in Figure 3.14. We see that the combinations 0,1,2

and 5 have a quite high probability to stay in the same sky state, which means that the situation is stable, and the variability is relatively low. For sky state combinations 3 and 4, on the other hand, the variability is relatively higher. This transition matrix shows us a clear overview of the transition probability, which helps us to understand the variability of given data.

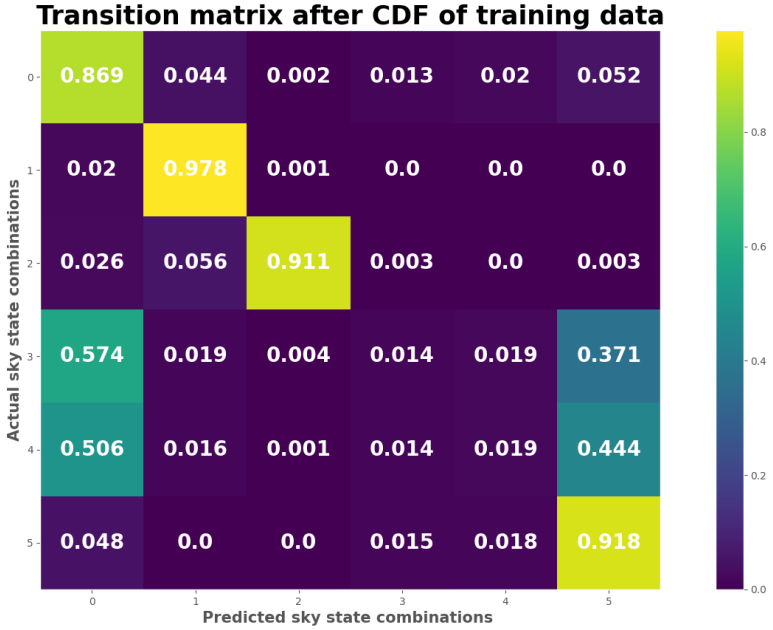


Figure 3.14 One step transition matrix of different sky state combination

This variability analysis method could also be used as a solar forecast method by firstly predicting the future sky-state, then processing the historical data to generate the distribution of PV production or irradiance in the same situation. And finally, select the quantile level of distribution according to the risk-gain management strategy. This method is particularly suitable as a complement for advanced forecast solutions or as a last-resort forecast solution for situations like loss of communication between the in-situ measurement instruments and the control system. In the latter analysis section, a variant of this Markov chain-based method is used, as shown in 3.3.3.

3.2 Variability, predictability, and forecasting

3.2.1 Solar variability and predictability

If solar forecasting is the tool to predict solar variability, predictability is the key feature that determines whether solar forecasting could be effective. In the literature, the notion of predictability and forecastability are frequently used and very close to each other in definition. Like [65] use the term of predictability to describe the forecast capacity, and forecastability to describe the potential deviation of the forecast. In this section, we describe the predictability only. As the name suggests, predictability means the capacity to predict. It is related but not equivalent to variability as we can have variable but predictable situations. The variations of GHI TOA, for example, are regular and periodic, thus totally predictable.

Precisely speaking, the predictability is based on the performance of forecast method and the observable that we have. Left part of Figure 3.15 shows Smart Persistence (SP) forecast based on measured GTI data and a scatter plot with measurements, the SP forecast here is achieved by keeping the clear sky index constant for consecutive instants and considering the change of solar geometry; Right part of Figure 3.15 shows the same plot but with SP forecast based on Helioclim3 v5 (HC3v5) data. In these figures, both the forecasts have a 30-min forecast horizon, which results in the characteristic delay compared to the measurement. The SP forecast based on HC3v5 has less variation

but a larger uncertainty range, which means that the predictability is not that satisfying. The SP based on measurement has a higher predictability and the statistical forecast performance is better that the other one, which prove the importance of good forecast method and data to predictability.

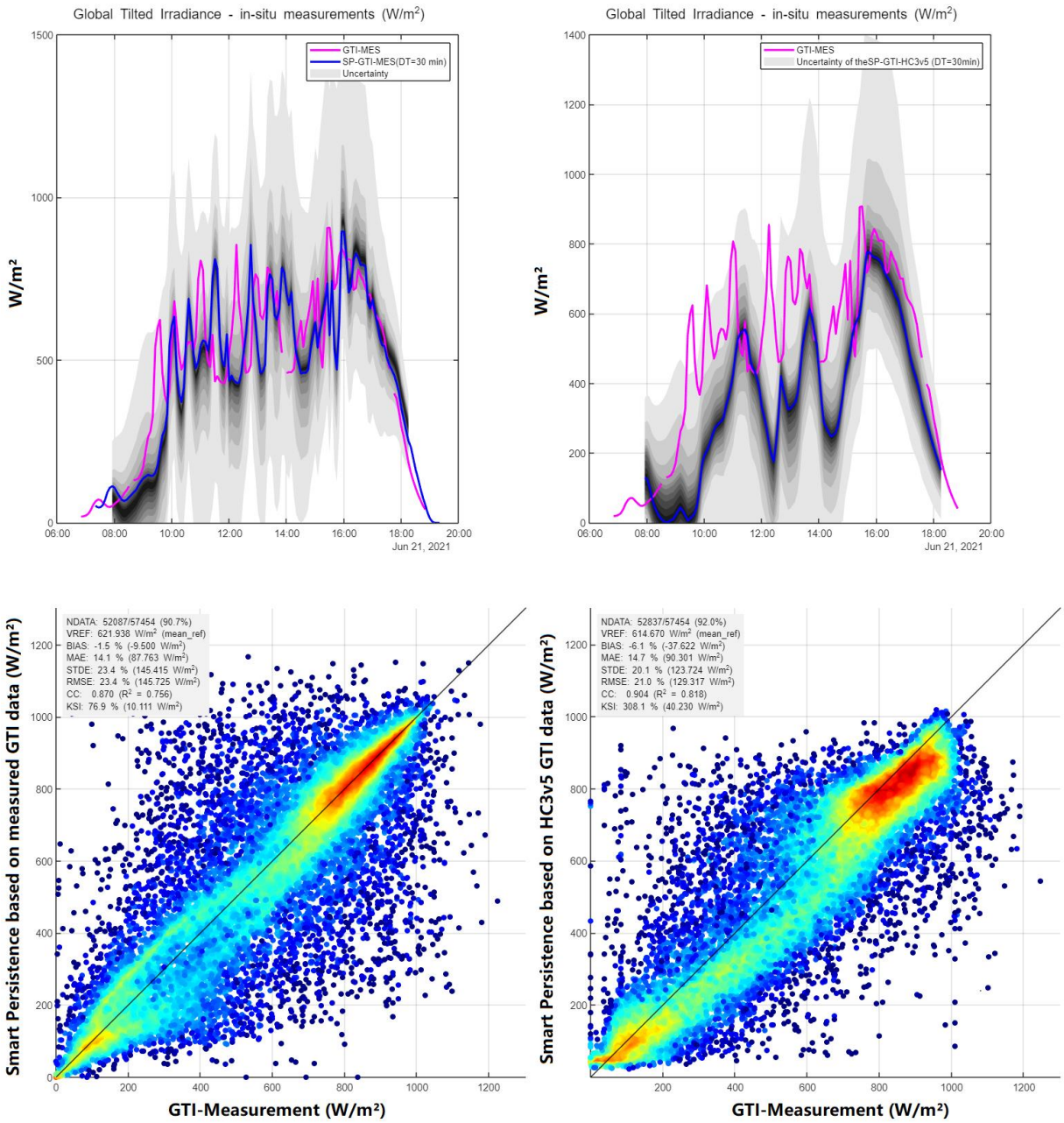


Figure 3.15 Probabilistic smart persistence based on measured GTI data (left), and Probabilistic smart persistence based on HC3v5 data (right)

The predictability mainly depends on unexpected events, which could strongly reduce the predictability since it is out of consideration and thus difficult to model and predict in a short time. Besides, it also depends on the information we have and the method we use. The predictability is somehow reflected in the uncertainty range, when the uncertainty range is large, the predictability is low, and the predictability would be high in other way. However, by adding in-situ hemispheric camera, the uncertainty range could probably be sharper and hence enhance the predictability. In general, the more information we have, the more likely we can analyze the situation to improve predictability.

3.2.2 The possible consequences of solar variability in isolated HES

In real-world applications, the solar variability is reflected in final PV production output. Besides the above reasons, the PV energy system components, such as different characteristics of PV module, different DC/AC inverters, and sun-tracking systems are also the reasons induce the final PV production.

In an isolated storage-less HES, without any energy buffer like battery energy storage system to mitigate the power fluctuation, solar variability is one of the main causes that lead to the unbalance between the power supply/demand when using solar energy.

In our case study, the solar forecast technique is used to predict the future PV production so as to define the genset planning and size of the spinning reserve. If the solar variability is not correctly predicted, redundant solutions would be taken to attenuate the fluctuation. For example, spinning reserve and even load shedding would be taken in case of energy lack. And for the situation of energy excess, the PV curtailment would be taken. All of these situations create extra costs, such as extra fuel consumption, extra running hour for diesel genset, waste of PV production, and maybe damage in case of load shedding.

If the power supply/demand balance is missed, the result would be much more serious, the energy outage situation, the so-called “Blackout” situation would occur and the economic loss could be light or very important according to the business type that needs power. If the solar variability is correctly predicted, these consequences could be mostly mitigated. Here, the predictability of the solar variability heavily affects the final result.

3.3 Solar forecast approaches used

This thesis uses three main forecast methods in the energy system simulation: the perfect prognosis forecast, the Complete History – Persistence Ensemble (CH-PeEn) forecast, and an advanced Markov Chain Mixture based model.

3.3.1 Perfect prognosis (PP) forecast

The perfect prognosis is considered as the best possible solar forecast, as it considers the forecast is the same as the real data. Hence, no matter the statistical metrics or the practical performance in simulation, the perfect prognosis provides the best performance compared to other solar forecasts. In our work, we consider a simulated PV production based on part of actual PV measurement, as a perfectly known timeseries, which is used as the PP forecast and the actual PV production. The detail of this simulated PV production is developed in chapter 4.

It should be noted that this method is not a realistic operational method but only serves as an upper bound (asymptotic) reference for simulation result analysis.

3.3.2 Complete History - Persistence Ensemble (CH-PeEn) forecast

Classic persistence forecast assumes that the future solar irradiance is the same as the present one. It is the simplest possible forecast and yet can be efficient for very short forecast horizons. However, for longer forecast horizons, its performance drops significantly, as illustrated in Figure 3.16. One important drawback is that the classical persistence does not account for the regular variations due to the sun position. Therefore, even on a clear sky day, which should be easy to predict, it performs not always satisfying.

To work around this limitation, the smart persistence approach has been introduced. Rather than persisting the irradiance, the clear-sky index is considered. The clear sky profile could be obtained through some mature products like McClear [59] or some simpler approach based on historical data like [66]. Yet, the performance still goes down when the forecast horizon gets longer due to the characteristic of persisting a same value.

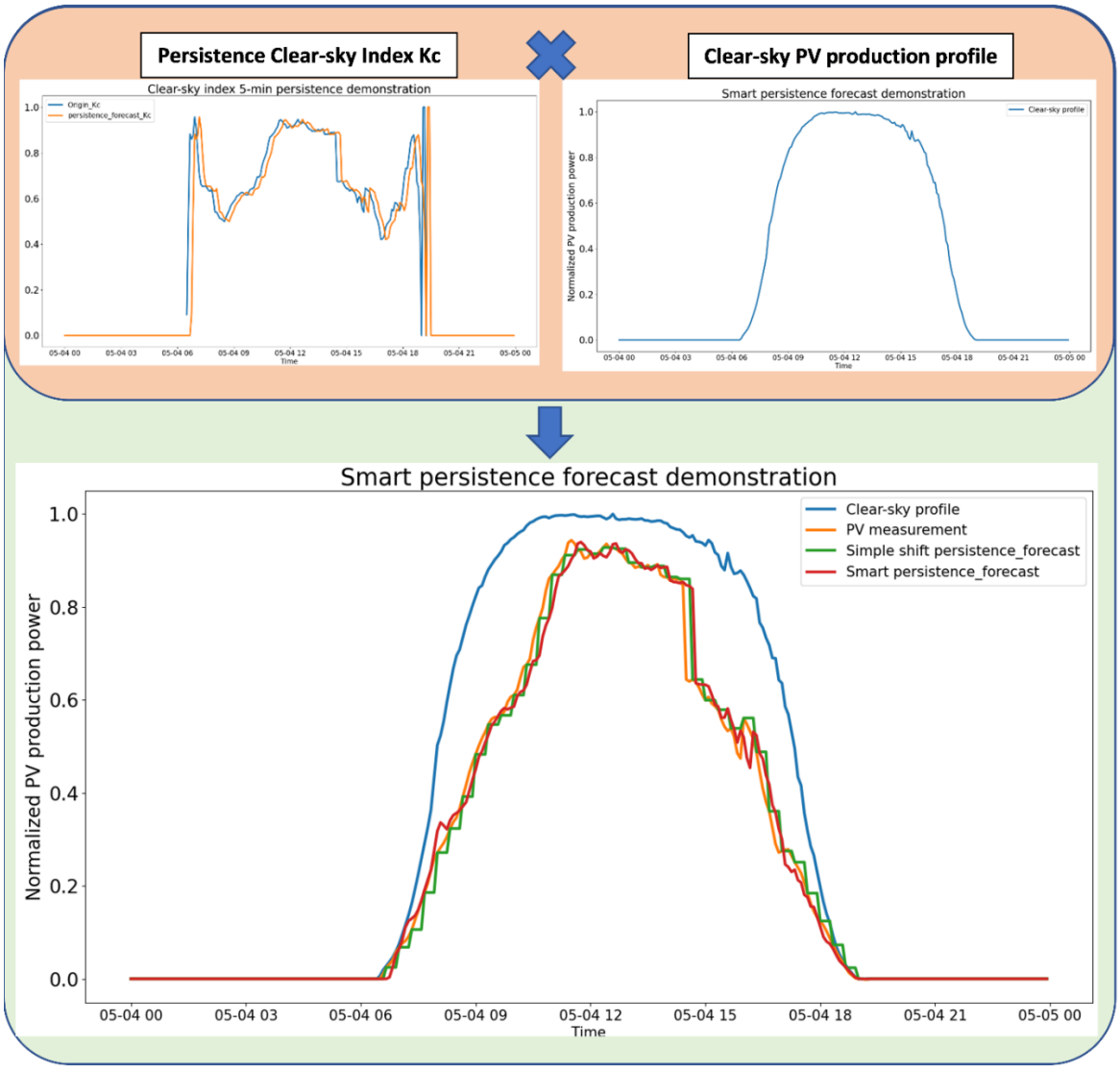


Figure 3.16 Scheme of smart persistence achievement

The previous approaches are designed for the deterministic forecast, when it comes to probabilistic forecast, Persistence Ensemble (PeEn) [67] is frequently used as probabilistic forecast reference. The main idea is to select the N closest clear-sky index observation to build a distribution for future clear-sky index forecast. Then apply the same process by multiplying the clear sky profile to obtain the forecast value. PeEn works fine as a reference case but when it is used in a real-world application for the spinning reserve sizing, its limited selection of N closest measurement forms a relatively small ensemble, which results in a small spinning reserve and hard to ensure the system stability.

Finally, an approach called Complete History - Persistence Ensemble (CH-PeEn) model proposed in [68] is chosen, as a benchmark baseline method for probabilistic forecast. The main idea behind CH-PeEn model is to use the whole data of the measurements, which has same instant of each day, to build an

ensemble of clear-sky index distribution. As an improved version of PeEn [69], CH-PeEn avoids the drawback of PeEn model of using the limited and subjective moving window size of recent data. Thanks to the design of using whole measurement data, CH-PeEn could provide a relatively stable performance in CRPS regardless the forecast horizon.

However, our case study has a particular climate, which is mainly composed of clear day and rainy variable day as shown in Figure 3.17. Hence, we made an adaption by processing the CH-PeEn for each month separately to avoid a bad sharpness. For high stability constraint cases, a large uncertainty range would ask high Genset capacity need, and this would be tough for higher PV penetration rate.

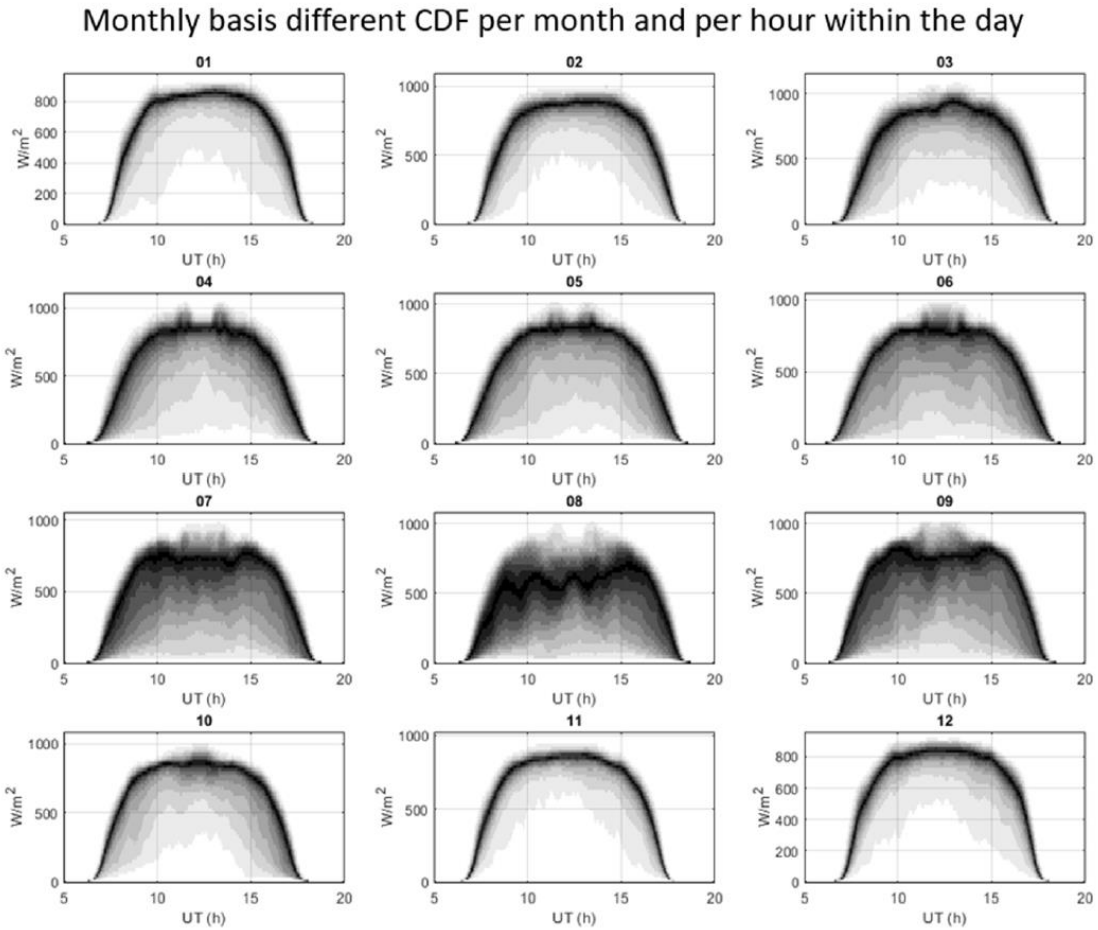


Figure 3.17 Monthly solar variability of case study

To build a typical profile of each month, we collect all the data of that month and sort the data of each instant by percentile level, and finally a typical daily profile with a complete distribution is obtained as shown in Figure 3.18. In this thesis work, this adapted CH-PeEn is chosen to provide the lower bound of HES performance.

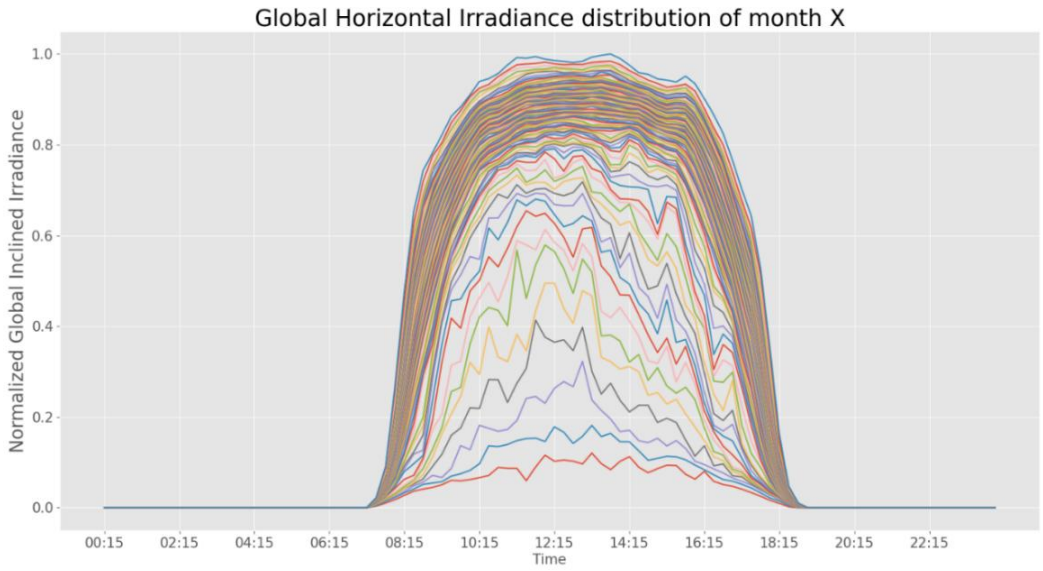


Figure 3.18 Example of a climatological forecast model for January of 10-year combined data

3.3.3 Markov chain mixture-based forecast

In addition to a "worse performer" (CH-PeEn) and a "perfect performer" (PP), we use a more realistic forecasting method, the Markov Chain Mixture (MCM) distribution model [70]. This model is simple yet robust and, similarly to the method used for the solar variability analysis of section 3.1.3, relies on a transition matrix.

Figure 3.19 shows the main idea of this approach, which is to divide the clear-sky index dataset in several classes (bins of magnitude) as shown in the left subfigure, and then find the transition probabilities of different classes between instant t and t+1, as shown in the middle subfigure, finally providing the distribution of different classes of clear-sky index, defined as (39) .

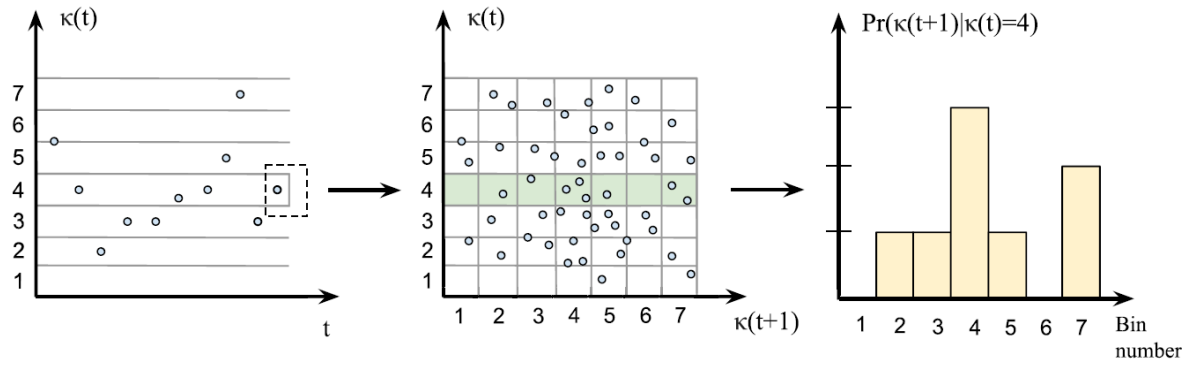


Figure 3.19 Explanation diagram of Markov Chain Mixture distribution method, [71]

The 1-step transition matrix shows the probability of next time step. For a further forecast instant, the forecast information could be obtained by multiplying this matrix itself corresponding times to have a new multi-step transition matrix, which provides the transition probability between actual instant and the forecast instant.

Once the distribution of the clear-sky index is obtained, we use a clear sky model to derive a the GHI:

$$GHI_{forecast} = K_{c_{forecast}} \times GHI_{CLS} \quad (43)$$

In this work, we get GHI_{CLS} from the Mc Clear model [59].

To be used in our simulation, the GHI forecast must be converted to global tilted irradiance (GTI). We use a model proposed by [72], which converts the DHI data to DTI data. The preparation work for this model is to first derive the DHI data from GHI data, then oversampling it to 1-min DHI data, and then process the DTI value. The GTI value could be hence derived from the 1-min DTI data.

Finally, we get the PV production forecast using the following irradiance to power transfer model [73], [74]:

$$P_{pv}(t) = S_{PV}\big(\rho GTI(t)\big)PR(1-\alpha(\frac{1}{\cos(\theta_i)}-1))(\eta_0+\beta\rho GTI(t))(1-\gamma(T_{cell}(t)-25)) \quad (44)$$

where S_{PV} is the PV plant surface, θ_i is the incident angle of solar irradiance over the PV panels, ρ is the fouling index which shows the difference between the measurement from professional instruments like pyranometer and the real irradiance received by the panel. PR is the global performance ratio between the actual delivered power and the power from solar irradiance, η_0 is the PV panel efficiency at standard test condition (STC).

The advantage of this method is that the process is relatively simple, it considers not only the historical data to give a better estimation of the uncertainty range but also the closest information for a better forecast.

Notice that the forecast result from this model is used as forecast input for the HES performance analysis. The final result of this forecast is generated each 5-min and is given in a format of 21 percentiles levels, 5-min temporal resolution, and 1-hour forecast lead time.

3.4 Solution for storages-less HES considering system uncertainty

Unlike an energy system with energy storage system, which has a fixed capacity of energy buffer to mitigate the solar variability, an isolated storage-less HES has to precisely define the energy buffer capacity – spinning reserve with the solar forecast and its uncertainty. With the help of short-term forecast, a suitable quantity of spinning reserve and smart generators dispatching could bypass the use of energy storage system. To achieve this purpose, the quantification of the uncertainty sources is necessary when sizing the spinning reserve and defining a solution to mitigate the system uncertainty and ensure the electrical network stability.

3.4.1 Quantification of solar forecast error

For our hybrid energy system, the first source of uncertainty is the solar forecasting error. Fortunately, the probabilistic solar forecasts provide an uncertainty range, which is particularly suitable for defining power management solutions that consider the uncertainty of solar forecast.

In this work, we mainly focus on the error overestimated case for our spinning reserve sizing, several intervals of different percentile levels under P_{50} of the probabilistic forecast are tested. The detail of choice of different scenarios is shown in the simulation analysis section of chapter 5.

3.4.2 Quantification of energy system failure

The second source of uncertainty are the energy system components failures, such as diesel generator failure, tracker system and inverter issues. This uncertainty mainly depends on the manufacturing quality and the durability of the materials that are used in the system; this value is normally provided by the manufacturer or the experience of the system operator. However, in our case study, this value is hard to quantify since the generators and other equipment have different health states. To deal with this kind of situation, in a classic diesel generator-based system, operators define spinning reserve equal to the nominal power of the largest online generator to be able to cover any failure of the online generators. In the industrial world, this rule is called the “n-1” principle.

In our simulation, however, the energy system failure probability is almost impossible to model. If we were to apply the “n-1” principle, the quantity of spinning reserve would be much higher than the solar forecast uncertainty. Most of the solar variability would thus be mitigated by this redundant spinning reserve and the simulation result could not show the real performance of different forecasts. Hence, in a first stage, the spinning reserve is sized according to the solar forecast uncertainty only. In a later stage, an extra energy buffer sized with a percentage of the PV installation capacity is considered, to secure the grid stability in case of the forecast is not well calibrated, where the spinning reserve sized with this forecast uncertainty is not sufficient. More details are shown in the simulation analysis in chapter 5 of this thesis.

3.4.3 Quantification of energy demand uncertainty

The third uncertainty source is the energy demand uncertainty. In order to find how to model energy demand uncertainty, an analysis of real power demand without PV system is done.

The industrial site of our study is operating all day and all year round, which means that the power demand never stops, except for maintenance and in case of system failure. From Figure 3.20, we see that the energy demand is roughly the same during a whole year, except for a few days that have much less demand. According to the feedback of site operators, these are the maintenance days or the energy outage situations. For maintenance days, around 26 MW power generation remains and for energy outage situations, the power generation drops to zero except in some cases that some small generators remain running.

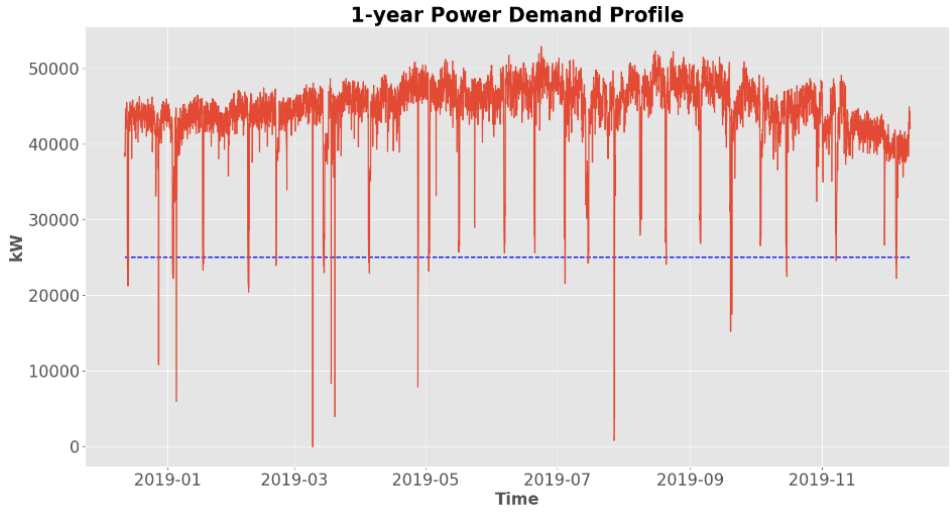


Figure 3.20 Real grid demand of an industrial site

To understand the maintenance frequency, a threshold of 26MW is applied to determine the day of maintenance. It is noticeable from Figure 3.21 that the probable maintenance day is mostly located on Thursday, which is the same as the operators' feedback. This information is very useful for the next day generators planning. A periodic maintenance means that these days could be excluded from the general strategy definition.

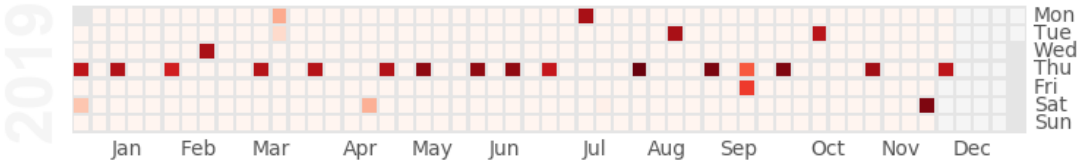


Figure 3.21 Grid demand of study case in west Africa with maintenance threshold

Excluding the maintenance days and abnormal energy outage situations in the dataset, the daily consumption profile is obtained like Figure 3.22, which shows the general situation of the energy demand.

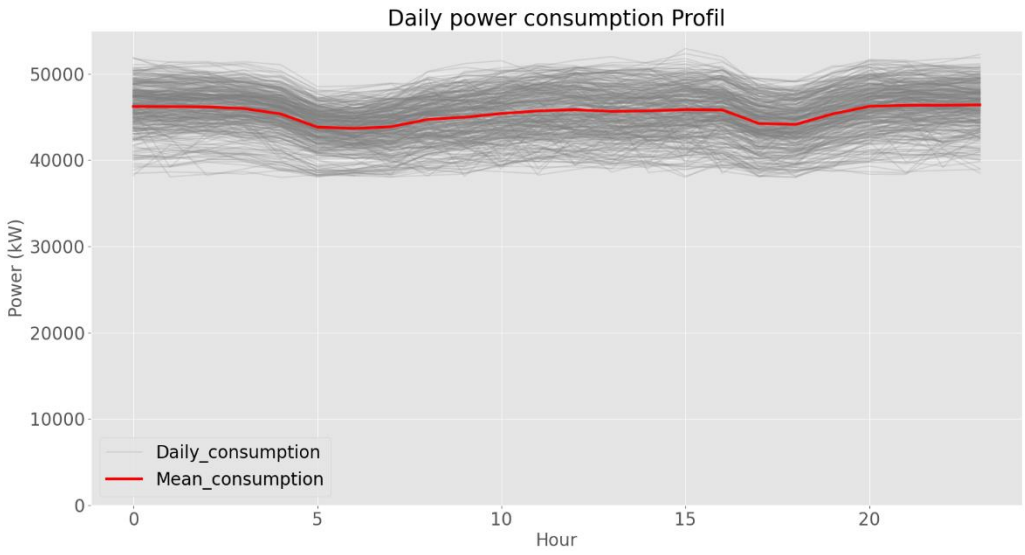


Figure 3.22 Daily consumption profile of 1-year data by excluding the maintenance cases

In brief, the grid demand of our case study mostly stays in a relatively small range (38 MW to 52 MW), except for the maintenance days and energy outage situations. Hence, the grid demand uncertainty is quite small. In order to focus on the solar variability in the simulation work, a light assumption which suppose the energy demand is all the time well predicted is made in this thesis. The grid demand uncertainty is not considered in our simulation.

Chapter 4

4. HES performance evaluation methodology

Contents

SUMMARY OF CHAPTER

In this chapter, we introduce the general methodology used to analyze the performances of a Hybrid Energy System (HES). This includes the key definitions to evaluate the Hybrid Energy System (HES) performance, and the simulation framework which explains how the simulation work is achieved.

Inside the key definitions, we firstly show the assumptions made to simplify the simulation work and focus on the influence of solar variability on system performance. Secondly, we introduce the criteria used for performance evaluation: in this work we consider classic criteria like stability and economic performance, but also additional criteria like solar forecast performance to adapt the HES. Lastly, we present the details of our calculation for each cost, which unifies all the performance indicators in a common unit through this cost-based simulation.

The simulation framework explains how to use solar forecast in a simulation of isolated storage-less HES is introduced in three parts. The first part is about the PV production and energy demand forecast. The second part is to generate the generator's order through an optimization solver according to the system constraints and the obtained forecast information. The third part is to evaluate the performance of our forecast compared to the actual measurements.

At the end of this chapter, we show a simplified typical case study to illustrate how our simulation works, and how our framework evaluates the HES performance.

4.1 Key definitions of HES performance evaluation

In the previous chapters, we have detailed the ideal functionality of HES, the constraints of using solar energy, and the difficulties of building a storage-less HES. In this section, we describe how we use the solar forecast in the HES simulation, and how to evaluate its practical performance without using classic statistical metrics. This simulation result is later used to analyze strategies for the integration of solar energy in the HES and to assess the performance of forecast techniques.

In the literature, there are many approaches to evaluate the performances of an energy system considering a single aspect, such as the HES stability performance, the HES economy performance, and the solar forecasting performance. For example, [75], [76] use a classic indicator like the voltage of HES to represent the system stability. [77] considers mostly the cost of different components of energy system and use its final cost as an economic indicator. For the forecast performance, [78], [79], [80], [81] use classic metrics like RMSE and MAE for deterministic forecast and CRPS (Continuous ranked probability score) and PINAW (Prediction Interval Normalized Average Width) for probabilistic forecast.

The metrics used in the works mentioned above focus on a specific aspect of forecast or HES but none of them embrace the different aspects of an energy system where system stability, economic performance, environmental impact and security of supply need to be considered equally.

Therefore, in this section, I would like to introduce a general HES performance evaluation methodology which allows us to evaluate the HES performance when using different solar forecast techniques. The first step of this general methodology is to set the simulation framework and define suitable criteria, which allow to answer the research question of this thesis.

4.1.1 Assumption of simulation

Several assumptions have been considered for the numerical simulation work in this thesis. In a real energy system, there are many small variations and power electronic behaviors occurring at temporal scale of milliseconds. Modelling the behavior of a power system at this scale requires an adapted model such as e.g. *PowerFactory, ETAP, etc*. Modelling the power system at such fine scale is out of the scope of this thesis. Instead, we focus on the short-term planning of the power system. We therefore assume that if the integrated power supply/demand balance is ensured for a 15-min interval, the supply/demand balance at finer scales within these 15-min is also respected.

Each day is considered as an independent event. Indeed, in our case study, there is no storage and the power demand remain at a high level even at night, so that all the generators are turned on at night to cover the power demand. As a result, the initial state of the HES is always the same at the beginning of the day. Assuming each day is an independent event allows us to simplify the simulation process.

4.1.2 Criteria and metric definition

In an interconnected energy system, the system performance is mostly based on criteria such as stability and economic efficiency.

4.1.2.1 Stability

In terms of stability evaluation, it is usually represented in frequency and voltage of the electrical grid. However, these quantities require a detailed modelling of the network which is out of the scope of this thesis (see previous section). Hence, in the simulation work of this thesis, rather than using frequency

and voltage, the number and quantity of energy excess/shortage situations are used to represent the HES stability performance.

- **System stability:**
  - Number and quantity of energy excess,
  - Number and quantity of energy shortages,
  - Number of "System imbalance – Load shedding" situations.

4.1.2.2 Economic perspective

From an economic perspective, the fuel consumption, the runtime of diesel generators, and the associated operation and maintenance costs are the classic indicators used in a classic thermal energy system [82], [83]. In our case study, besides these values, the fuel and cost saved by using PV energy are also considered.

- **Economic performance:**
  - PV energy used/fossil energy used,
  - Fossil fuel saved,
  - Cost saved by using PV energy,
  - Total marginal operation cost,
  - Operation & maintenance cost.

4.1.2.3 Forecast performance

To evaluate the solar forecast performance, rather than using classic statistical metrics, the spinning reserve usage and the corresponding costs are used instead. A good statistical value does not always mean a good practical economic performance [26]. For example, an overestimated forecast and an underestimated forecast may have the same RMSE, but their respective impact on the storage-less system could be significantly different. Similarly, a moderate forecast error can have a larger impact on the power system than a bad forecast depending on the configuration of the gensets (this aspect is further analyzed in chapter 6).

It should be noted that the quantity of spinning reserve used somehow reflects the accuracy of the forecast. Indeed, if the solar forecast is accurate, the reserved capacity is not used, but if it is inaccurate, the reserved capacity is used to maintain the power supply/demand balance. This metric could bypass the limits of classic metrics in distinguishing the overestimated and underestimated cases, and the solar forecast performance could be quantified in the numerical simulation. Indirectly, the influence of solar forecasting performance on the final HES performance could therefore be evaluated. In addition, using the spinning reserve cost allows a direct comparison with other cost components.

- **Solar forecasting performance:**
  - Cost of maintaining spinning reserve,
  - Cost of actual spinning reserve usage.

The cost calculation of using spinning reserve is based on the unit marginal cost (the CAPEX is not considered throughout this analysis) and the quantity of spinning reserve used, which is equal to the difference between residual demand and the actual diesel generator production. The detail of calculation is developed in section 4.1.3.

4.1.2.4 Summary

The criteria and the metrics used in HES performance evaluation in this work are summarized in Table 4.1. However, all the metrics used don't have a common indicator, which makes it hard to quantify the performance of the whole HES. To have a common metric, I represent the performance of all criteria in one cost-oriented criterion since all of them could be associated with a cost in the numerical simulations.

Criteria	Metric used in this work
Stability	Energy excess/shortage Blackout occurrence Spinning reserve availability
Economy performance	Fossil fuel savings Total marginal cost/OM cost
Solar forecasting performance	Use of spinning reserve & corresponding cost

Table 4.1 Metrics used for HES performance evaluation

For example, the energy excess in the simulation has an associated cost for wasting energy like PV curtailment, and the energy shortage cases have an associated cost for extra fuel consumption of using spinning reserve, or extra economic loss if there is an energy outage. The cost of energy outage could be configured in the simulation.

Thanks to the design of this unified criterion, the final HES performance is represented in cost, for each given day, when the solar forecast is perfect, the system cost would be the lowest. The detail of the calculation is developed in the following section.

4.1.3 HES Performance indicators calculation

In the composition of our case study, the cost of HES comes primarily from the diesel generator, PV energy system, and the loss of industrial revenue in case of energy outage. This section details the calculation principle of each of these costs.

4.1.3.1 Cost of diesel generators

The diesel genset cost is the most important cost in our case study. It is composed of two parts:

1) **CAPEX – Capital Expenditure**

   The first part is the CAPEX, which is the fixed investment cost of an installation. The unit fixed cost of each kWh produced is usually used to optimally size the installation. However, the installation sizing is not the purpose of this thesis, and the running state of the existing machines is hard to quantify. In this thesis, we aim to analyze the benefit of integrating a PV system into an existing installation, which means evaluating the difference between the same configuration with and without PV system. Since the genset installation is the same in both two cases, the CAPEX remains the same. As a result, the investment cost of diesel generators is hence not considered.

2) OPEX – operational expenditure

The second part is the operational expenditure, which could be divided into two parts in our simulation work:

- The operation and maintenance (OM) costs is the basic cost to maintain a HES installation. This cost is separated into two parts, the fixed OM cost, and the variable OM cost. The fixed OM cost is related to the installation capacity, it is also called fixed cost in a simulation. The unit could be [$/year] or [$/kW] if the annual cost could not be obtained easily. In our case, we use [$/kW] and the total nominal power to estimate the total OM cost of genset. For each day, the OM cost is the same for a given installation. The variable cost mainly represents the cost of starting and shutting down a generator.

$$Total\ OM\ cost\ [\$] = Fixed\ OM\ cost\ [\$] + Variable\ OM\ cost\ [\$] \quad (45)$$

$$Fixed\ OM\ cost\ [\$] = nominal\ power_{generators}\ [kW] * unit\ OM\ cost\ [\$/kW] \quad (46)$$

$$Variable\ OM\ cost\ [\$] = Number_{startup/shut\ down} * Cost_{startup/shut\ down}\ [\$] \quad (47)$$

Since the fleet of genset is constant in all simulations, the fixed OM costs remain constant. The relevant OPEX in our analysis is thus the variable OM costs that come from the additional startup/shutdown needed when integrating the PV energy.

- The **Marginal cost** is the cost of producing one more kWh (except for the parts of fixed and O&M cost of starting and shutting down a machine). It is the integral power output multiplied by the unit price of producing a kWh with corresponding fuel. As shown in Figure 4.1, the efficiency of the generator varies according to the state of charge. However, simulating precisely the generator's performance with efficiency is not the purpose of this thesis. On the contrary, we focus on the power output and the system cost. In this figure, the specific fuel consumption, which means hourly fuel consumption, its relation with the power output is linear except for the startup phase. Hence, in our simulation, we consider the linear function for our cost calculation, the startup phase is calculated separately by the optimization process with specific system constraints.

$$Total\ Marginal\ cost\ [\$] = hourly\ energy\ output\ [kWh] * unit\ marginal\ cost\ [\$/kWh] \quad (48)$$

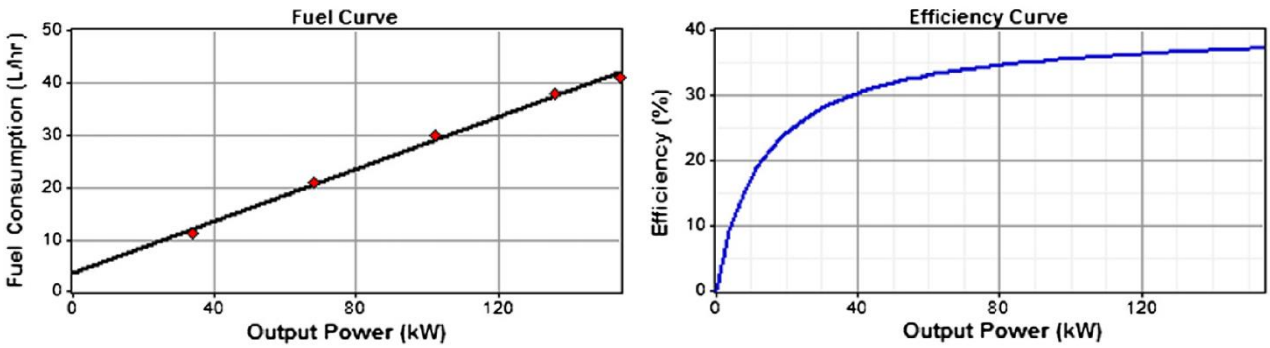


Figure 4.1 Example diesel generator characteristic curve, [84]

Remark: In general, each startup of the diesel genset is equivalent to 50 running hours lifetime, which should be counted in the case when considering the generators' CAPEX. In our simulation work, only the fuel consumption and the OM part of start-up cost (included in the general OM cost) is considered, the reduction of lifetime is out of the scope of this thesis.

4.1.3.2 Cost of PV energy system

Unlike the investment cost of diesel generators' installation, the investment cost of the PV power plant is considered to calculate the PV curtailment cost with the operation and maintenance cost. The simplified formula for investment cost with loan interest is defined as (49) :

$$Investement\ cost = loan * (1 + x)^n \quad (49)$$

Where loan is the rental amount, x is the annual interest rate and n is the loan period.

The total cost of PV power plant is calculated as (50) :

$$C = Investement\ cost + maintenance\ cost \quad (50)$$

In PV energy system, the unit cost of normal production and PV curtailment are the same. The main idea is to calculate the Levelized Cost of Energy (LCOE) of PV energy. The LCOE is the complete cost of producing one kilowatt-hour of energy, which considers the investment cost, operation, and maintenance cost. Thus, according to [85], knowing the discount rate of cost and energy production, the LCOE of each kWh of PV energy can be calculated as the formula (51) :

$$LCOE_{PV} = \frac{\sum_{t=0}^{T} \frac{C_t}{(1+r)^t}}{\sum_{t=0}^{T} \frac{E_t}{(1+r)^t}} \quad (51)$$

where, C_t is the total cost of a whole lifetime and the E_t is total energy produced through the whole lifetime, r is the discount rate of each year. The discount rate is the alternate investment return against which a project's return is evaluated (all financial conditions included). Hence, for a non-finished energy system, the discount rate could be variable.

Therefore, for the normal production of PV energy, the real cost is calculated with the LCOE and the PV energy effectively used, which is the difference between the total PV energy produced and the PV curtailment, also equal to the fossil energy saved. The PV curtailment also has a cost which is detailed in the section of system imbalance cost. The calculation of energy and cost saved are easily derived according to the fuel energy density, the fossil energy used, and the current fuel price.

4.1.3.3 Cost of system imbalance

When the solar forecast is underestimated, the final power production is higher than the demand. In this case, there is a cost for this energy excess situation if the overproduction could not be mitigated by the deceleration of diesel generators. This energy excess cost is come from the cost of wasting the PV energy by the PV curtailment action. The penalty cost of PV energy non-used is calculated according to the LCOE and the quantity of PV curtailment, as shown in (52) .

$$Excess\ cost = PV\ curtailment\ [kWh] * LCOE_{PV}\ [kWh/\$] \quad (52)$$

In the other way around, when a solar forecast is overestimated, energy shortage situation occurs. If the electrical network stability can be ensured by online spinning reserve, the cost is just the extra fuel consumption. Otherwise, there will be a load shedding action, which could potentially lead to a blackout situation. A load shedding or blackout situation causes a much more serious issue and higher loss. In our simulation, the load shedding cost is set very high, which force the optimizer to consider it as a last resort solution.

$$Actual\ SR\ used\ cost = Actual\ SR\ used\ [kWh] * \text{unit marginal cost [\$/kWh]} \quad (53)$$

$$Shortage\ Cost = Actual\ SR\ used\ cost + load\ shedding\ cost \quad (54)$$

$$Imbalance\ Cost = Shortage\ cost + Excess\ cost \quad (55)$$

4.2 Framework of simulation

In general, the process of solving an optimization problem in programming could be simplified in three steps [86], [87], as shown in Figure 4.2:

1) Preparing the input timeseries of the analysis period,
    a. PV forecast and PV measurement timeseries,
    b. Energy demand timeseries,
2) Modeling the energy system respecting the system constraints,
    a. Defining the system constraints to consider,
    b. Defining the system objective and decision variables,
    c. Programming the practical energy system as an optimization model in a mathematical way
3) Generating the solution of optimization problem with a solver, as explained in the Unit commitment section of chapter 2.
    a. Defining the optimization solver,
    b. Applying the optimization solver to the optimization model and generating the optimization results.

Figure 4.2 Illustration diagram of classic optimization process

However, this process is mainly for the pre-analysis because the input timeseries are normally complete for the whole analysis period. In an operational context, which integrates the use of solar forecast technique, it means that the solar production forecast timeseries would be refreshed at each forecast time rather than be given in one time for the whole analysis period. Hence, the actual information of each state that we have is always limited by the forecast lead time, and this process could not be used directly for the simulation of our case study, since it needs the residual demand timeseries of the whole period to generate the generators dispatching order and behavior.

In this thesis, I used two extra parameters and a corresponding system performance evaluation methodology to achieve the simulation of our case study. The two parameters are forecast lead time and genset update time respectively.

The first parameter, solar forecast lead time, needs to be considered for the HES performance evaluation because a long enough forecast is needed to generate a globally optimal solution. But at the same time, we prefer solar forecasting as short as possible to ensure a better accuracy performance. This parameter could therefore help us to find the best setting for a given energy system. Secondly, for an advanced forecast method, the forecast refresh time is always shorter than the forecast lead time, however, it is impossible to update the diesel genset status all the time, since the start/stop of diesel generators needs extra human resources and costs. Moreover, it would significantly reduce the diesel generator's lifetime which could be very costly. Hence, the second parameter is to set a compromise between each forecast refresh time and the forecast lead time, to update the genset state.

The scheme of the general methodology could be simplified as shown in Figure 4.3. Rather than using the data of the whole analysis period and solving the optimization problem in one time, we place the forecasting data and generate the optimal solution for each forecast period. Finally, the whole analysis period is done in several cycles. The data used in the simulation contains the solar forecasting and estimated energy demand timeseries with forecasting uncertainty.

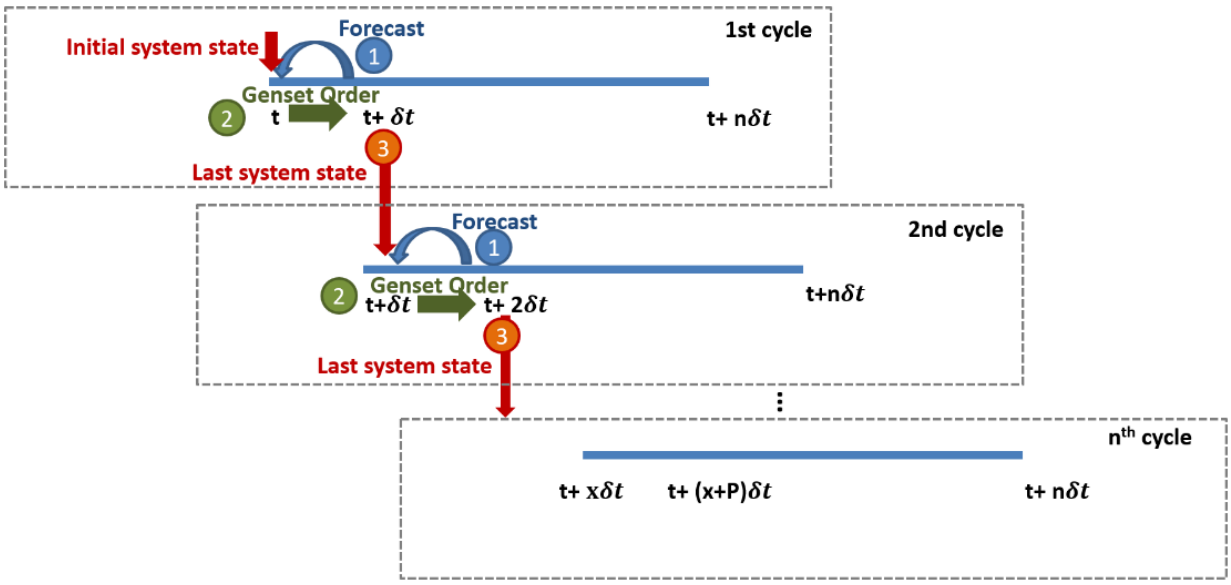


Figure 4.3 Illustration diagram of HES performance evaluation methodology with time notion

Nevertheless, due to the variability of solar resources and energy demand, it is almost impossible to have a perfect forecasting, which means that the actual state is not as the forecast said. Hence, the system cost obtained with forecasting information cannot reflect the real system performance. To obtain the real system performance of each period, a step called real-data evaluation is introduced in the general performance evaluation methodology, which could be used to evaluate all kinds of PV/diesel HES. As shown in Figure 4.4, the methodology is divided into three steps, the main idea of these steps is developed in the following section.

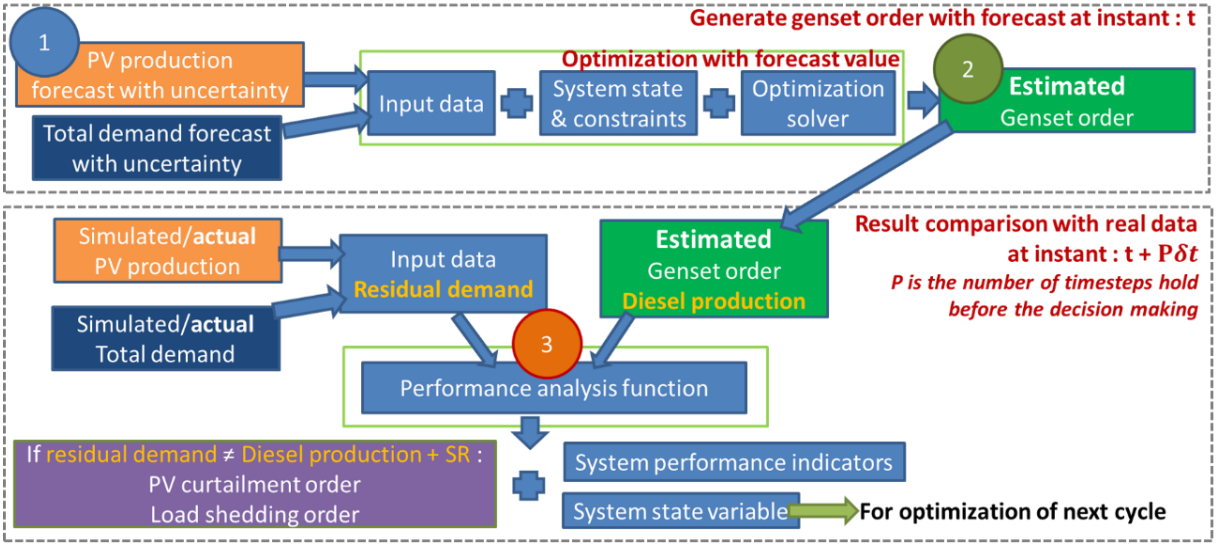


Figure 4.4 Illustration diagram of HES performance evaluation methodology

4.2.1 PV and residual load forecasting

In this methodology, the first step is to achieve a PV production and load forecast. In our case, the demand forecast is omitted by considering it is almost constant. In terms of PV production forecast, we use a probabilistic forecast which provides a deterministic value and an uncertainty range. The deterministic value is used for generating genset order dispatching and the uncertainty range is used to help define the spinning reserve quantity.

Figure 4.5 shows a concrete example of making a forecast; at instant t, we make a PV and grid demand forecast for instant $(t + \delta t)$, noted $\widehat{\boldsymbol{PV}}(\boldsymbol{t} + \boldsymbol{\delta t}|\boldsymbol{t})$ and $\widehat{\boldsymbol{C}}(\boldsymbol{t} + \boldsymbol{\delta t}|\boldsymbol{t})$ respectively. At instant $(t + \delta t)$, the residual demand, which is equal to the demand minus the PV production, should be committed by the diesel production $\widehat{\boldsymbol{DG}}(\boldsymbol{t} + \boldsymbol{\delta t}|\boldsymbol{t})$**.** Since the forecast lead time of our case study is not only a single point but a period ranging from $\boldsymbol{t}$ to $\boldsymbol{t} + \boldsymbol{n\delta t}$, the final PV forecast gives a timeseries of forecast results with different forecast horizons. This process is repeated every refresh time $\boldsymbol{\delta t}$**,** which means that there is a new forecast every refresh time. The forecast of different forecast horizons has different performances: the closer the horizon, the better the forecast accuracy.

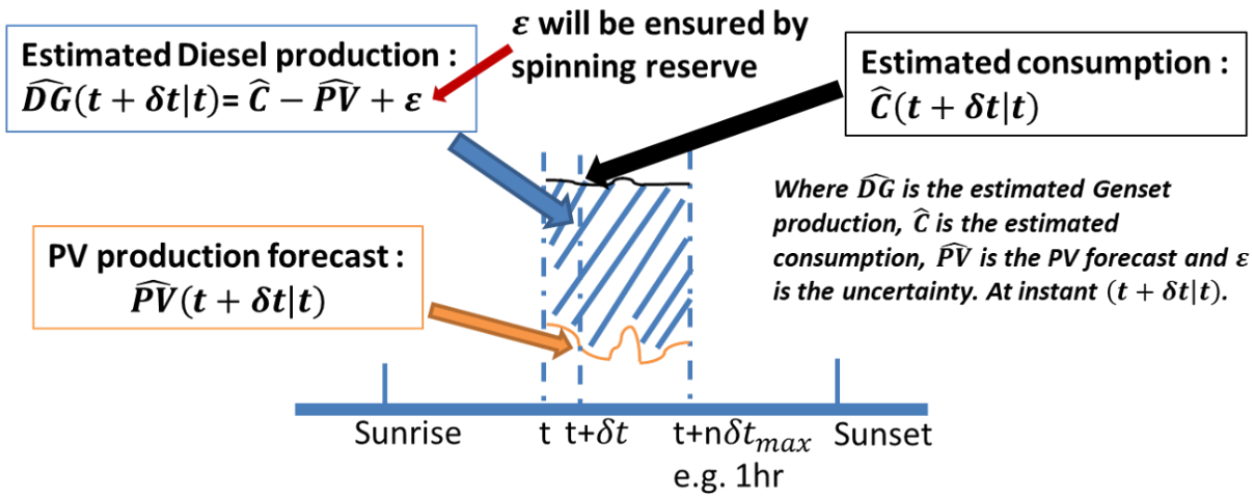


Figure 4.5 Illustration of PV forecasting application

4.2.2 Genset planning/dispatching

The second step of the methodology is to generate the prior genset planning with an optimization solver, according to the PV production forecast, load forecast information, and the **previous system state variable** X(t). For the next forecast refresh instant $t + \delta t$, the genset order is noted $O(t + \delta t|t)$, which is set for the whole forecast lead time $\boldsymbol{n\delta t}$, and there is hence n possible genset orders for each discrete timestep as shown in Figure 4.6. But as mentioned before, genset update time is introduced to avoid the frequent change of machine state, hence, among n genset orders, there is only 1 order would be updated at genset update time.

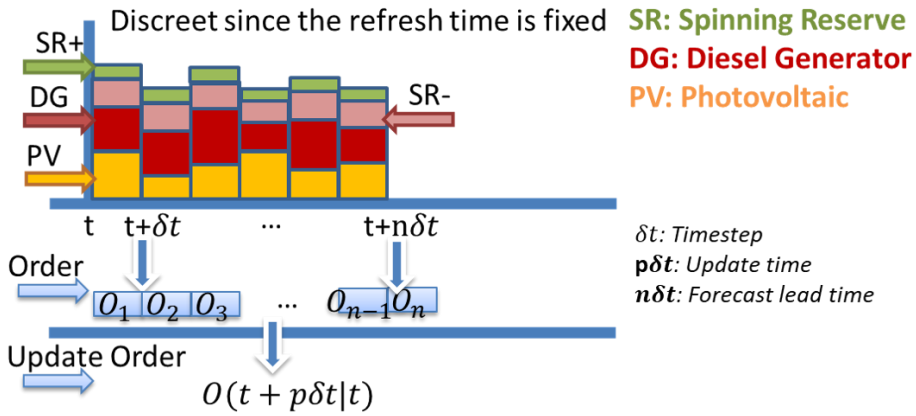


Figure 4.6 Illustration of genset order generation with PV forecasting and optimization solver

Supposing there are ***k*** generators, the system state variable and genset order could be written as:

State variable: $X(t) = Status_k \epsilon \{0, 1\}$

Genset Order: $O(t + \delta t|t) = \begin{cases} Status_k \epsilon \{0, 1\} \\ Power_k \epsilon [P_k^{Min}, P_k^{Max}] \end{cases}$

The system state is saved for the next optimization cycle and noted as **system state variable** $X(t + \delta t)$.

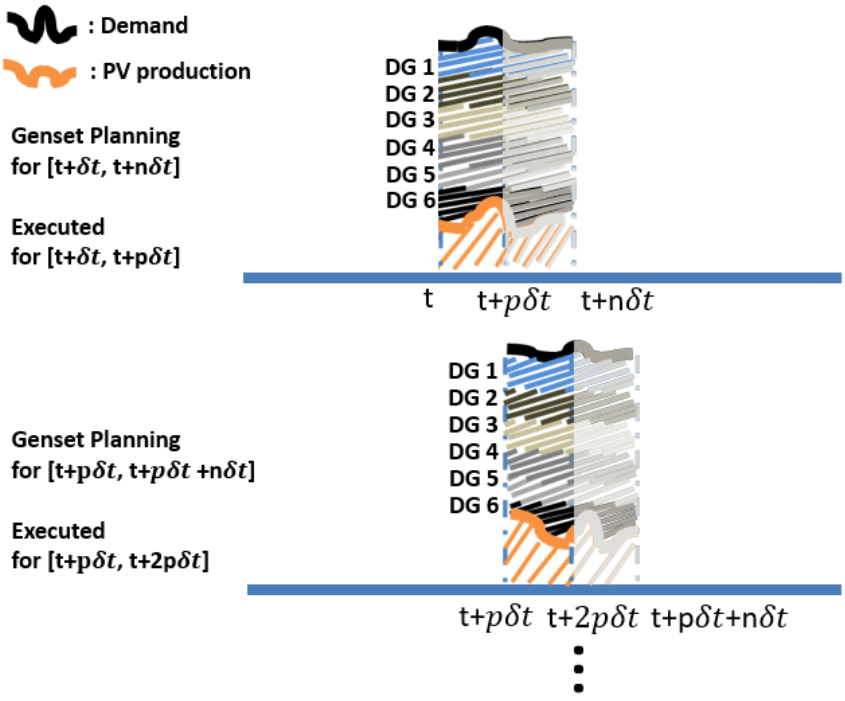


Figure 4.7 Illustration of simulation with forecast lead time and update time

As shown in Figure 4.7, each genset planning is generated for a period of $n\delta t$, but only the $[t,\ t+p\delta t]$ period of genset planning is effectively used, at instant $(t+p\delta t)$, a new planning is used for the next $n\delta t$ period. In a one-day simulation, this process is repeated many times.
Nevertheless, there is no perfect forecast, the difference between forecasting and reality is the key point for the HES performance evaluation, which is also the main content of the third and last step of the methodology.

4.2.3 Posterior evaluation

From the second step, we obtain the planning of genset order for the future instant $O(t+p\delta t|t)$, when we arrive in the future forecast instant $t+p\delta t$, the actual residual demand could be derived from the actual PV production $PV(t+p\delta t)$ and actual consumption $C(t+p\delta t)$.

As shown in Figure 4.8, for each update period, the performance evaluation process calculates the difference between the **planned production** $\widehat{\boldsymbol{P}}(\Delta \boldsymbol{t}), \Delta \boldsymbol{t} \in [\boldsymbol{t}, \boldsymbol{t}+\boldsymbol{p\delta t}]$ and the **actual residual demand** $\mathbf{RD}(\Delta \boldsymbol{t}), \Delta \boldsymbol{t} \in [\boldsymbol{t}, \boldsymbol{t}+\boldsymbol{p\delta t}]$**.** This difference is the power supply/demand imbalance, which could be energy excess or energy shortage situation.

Then, every system imbalance action is considered to calculate the different parts of the HES performances/costs. The final summarized system cost is therefore a performance indicator: the less the cost is, the better the performance. Moreover, this performance score could be potentially used as a feedback for the solar forecast of each optimization cycle.

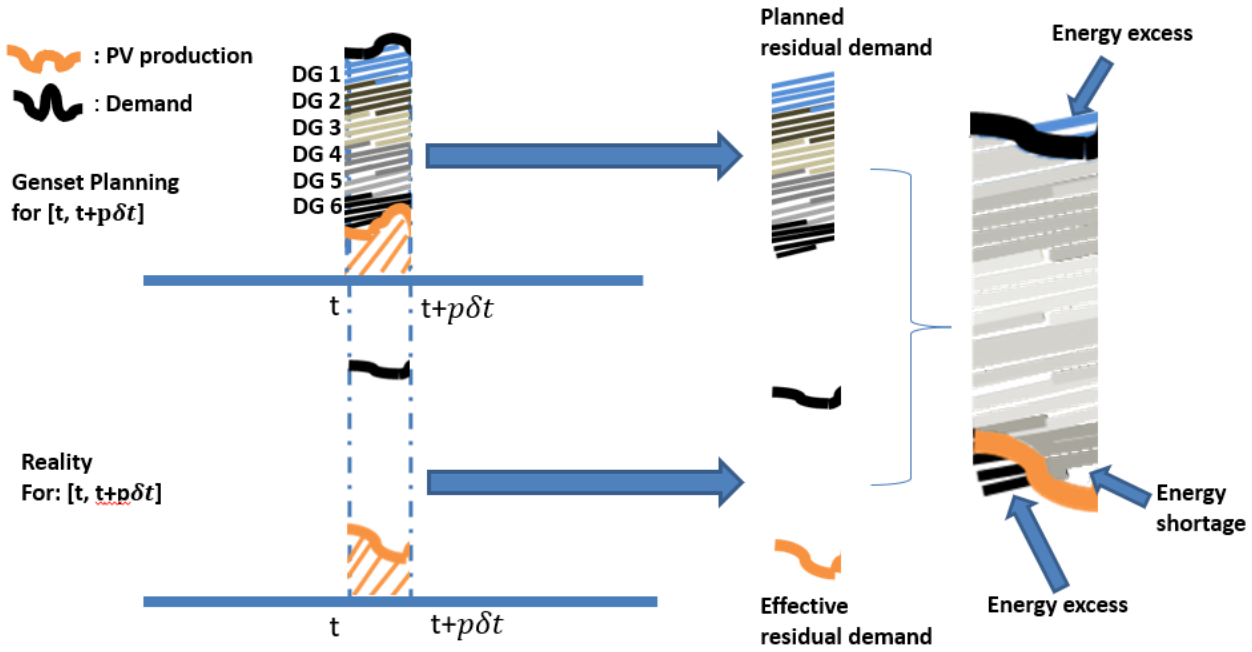


Figure 4.8 Illustration of real data evaluation process in simulation

In summary, the principle of this methodology is to achieve the classic optimization process for each forecast period and apply the genset planning every update time, as shown in Figure 4.9. At each posteriori evaluation, we evaluate the estimated optimal solution generated with forecasting information compared to the actual residual demand.

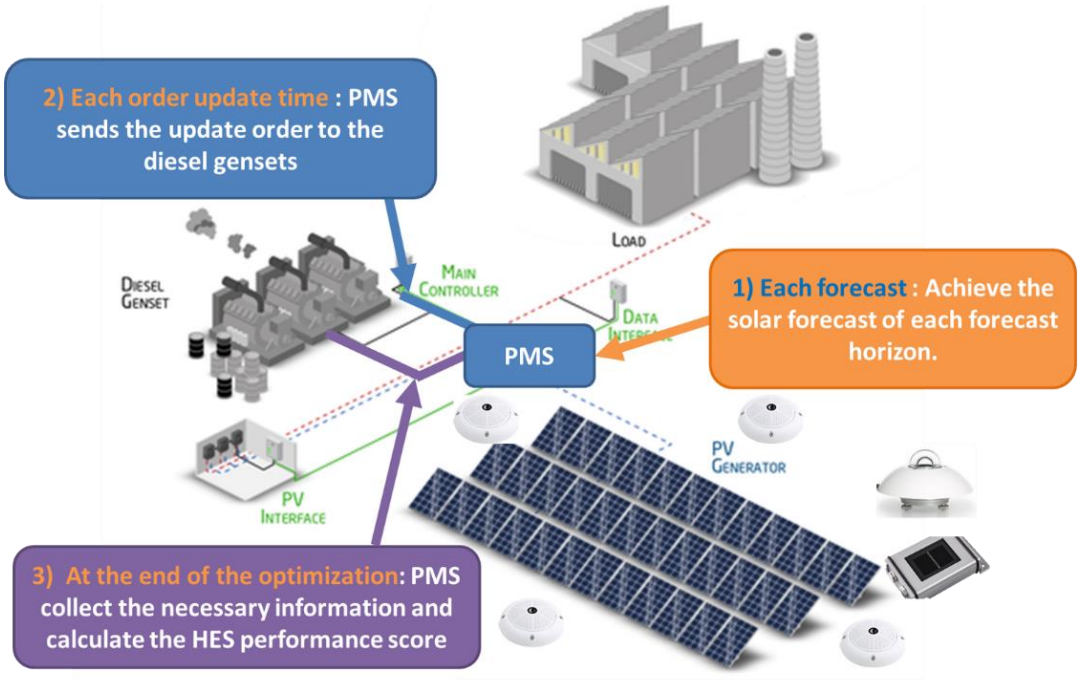


Figure 4.9 Illustration of PMS functionality with different genset order update time

To avoid integrating the influence of PV power production failure that is due to unexpected accident or human factor in the HES performance analysis, we decide to use PV power simulation based on a part of actual PV production measurement, satellite data and reanalysis temperature, achieved with a regression process for the posteriori evaluation of the HES performance.

In evaluating the statistical performance, the data with solar elevation angle smaller than 5 degrees are not considered. Here the solar elevation angle calculation is based on the Python library ***Solar Geometry 2*** [88]**.**

4.3 Example of application of the HES performance indicators calculation

4.3.1 Context

As Introduced before, there are more than 20 diesel generators in our case study. While simulating this system is not difficult, visualizing the results of the simulation is more challenging. Hence, to give a clearer explanation of this methodology, a simplified case is detailed here.

In our case study, the load-sharing rule is applied, which means that all the generators have a load equal to the same percentage of their nominal power. This rule makes the spinning reserve from each generator remain in the same percentage. Since the generators in our case study could be divided into 4 types, and the high energy demand makes the generators remain on most of the time, it becomes hence reasonable to use fewer generators in the simulation, but with at least 4 types of characteristics and with a total genset capacity equal to the actual case study. Hence, the number of generators is reduced to 6 but covers all the types of existing diesel generators with a certain scale. This simplified case study takes the same geographical position as introduced before, having a rich solar resource as shown in Figure 4.10.

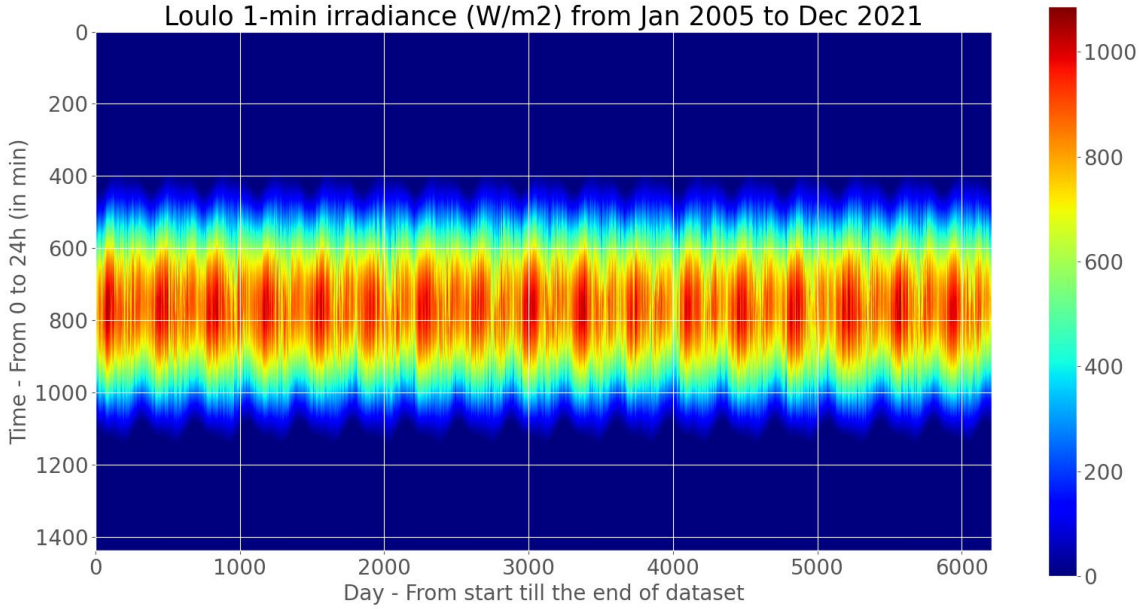


Figure 4.10 Solar resource overview of case study

The composition of the energy system remains the same, diesel generators, PV energy system, and the load. The PV production is a simulation based on a part of the reality as explained in chapter 3, to match the solar forecast. The load is also the same as introduced before, which is almost stable all the time. The only difference is the number of generators and their nominal power, which is scaled up to match the demand and the spinning reserve system constraint. The details of all the components in the simplified case study are shown in Table 4.2.

	Generator 1	Generator 2	Generator 3	Generator 4	Generator 5	Generator 6
Nominal power (kW)	3000	6000	8000	15000	18000	21000
Optimal operating range	40%-100%			60%-100%		
Unit cost (kwh/$)	0.188			0.146		
Fuel used	Diesel			Heavy fuel		
Fuel density (kg/L)	0.86*			0.96*		
Fuel energy density (MJ/kg)	45.6*			39.5*		
Fuel price ($/L)	0.73*			0.65*		
*Value provided by the grid operator of our case study						

Table 4.2 Details of the generators in the simplified case study

4.3.2 Simulation demonstration

The simulation work follows the principle of the methodology, a genset planning is generated according to the forecast information at each optimization cycle. As a system constraint in the optimization, the spinning reserve is sized according to the solar forecast uncertainty and a user-defined uncertainty range. An example of variable day is presented to show more detail about genset dispatching to maintain the power supply/demand balance. The structure of simulation work could be briefly considered as input and output part.

The Input:

The input of the simulation is the forecast information of PV production, energy demand, and the system constraint. Figure 4.11 shows an example of a probabilistic forecast. To answer the purpose of the research question of this thesis, the evaluation of the statistical performance of the forecast method is necessary.

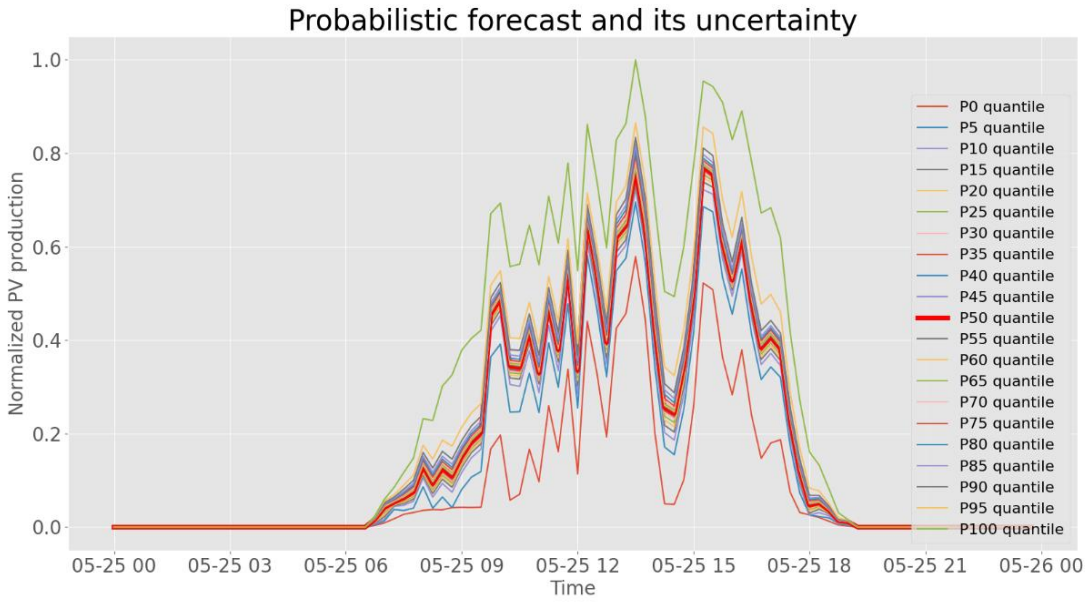


Figure 4.11 Display of probabilistic forecast and its uncertainty range

The Output:

Figure 4.12 shows the global dispatching planning (grey parts), the overlapped solar forecast information (green lines), the minimum quantity of spinning reserve (purple part), the demand (Blue line), and the forecasted PV production (yellow part).

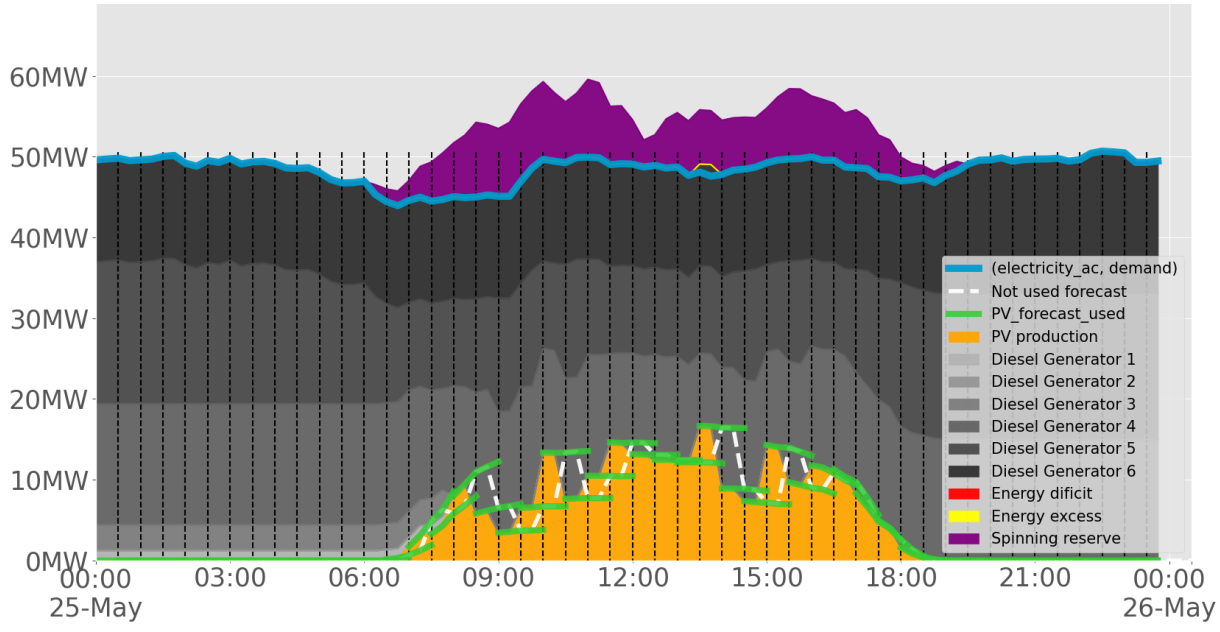


Figure 4.12 Phase of generating generators planning of typical case

The second step is to evaluate the HES performance by comparing the estimated value and the reality. As the result shown in Figure 4.13 and Figure 4.14, the real residual demand is not the same as the planned diesel power generation, which asks requires energy system to take some penalty actions like PV curtailment and using spinning reserve to maintain the power supply/demand balance.

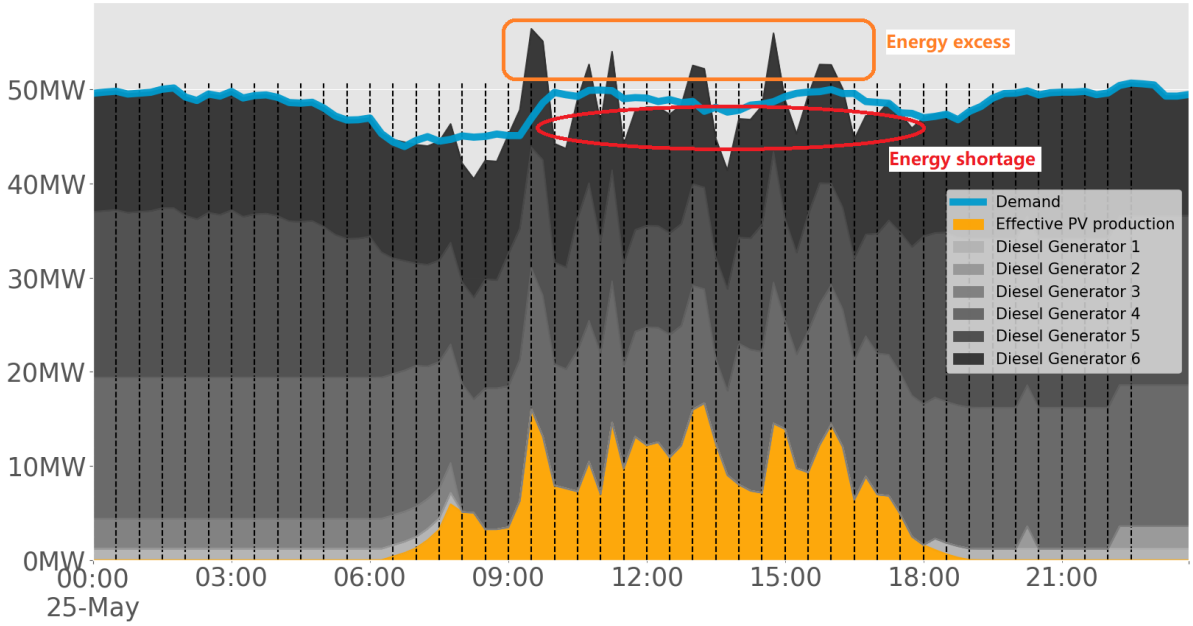


Figure 4.13 Phase of real data evaluation of typical case

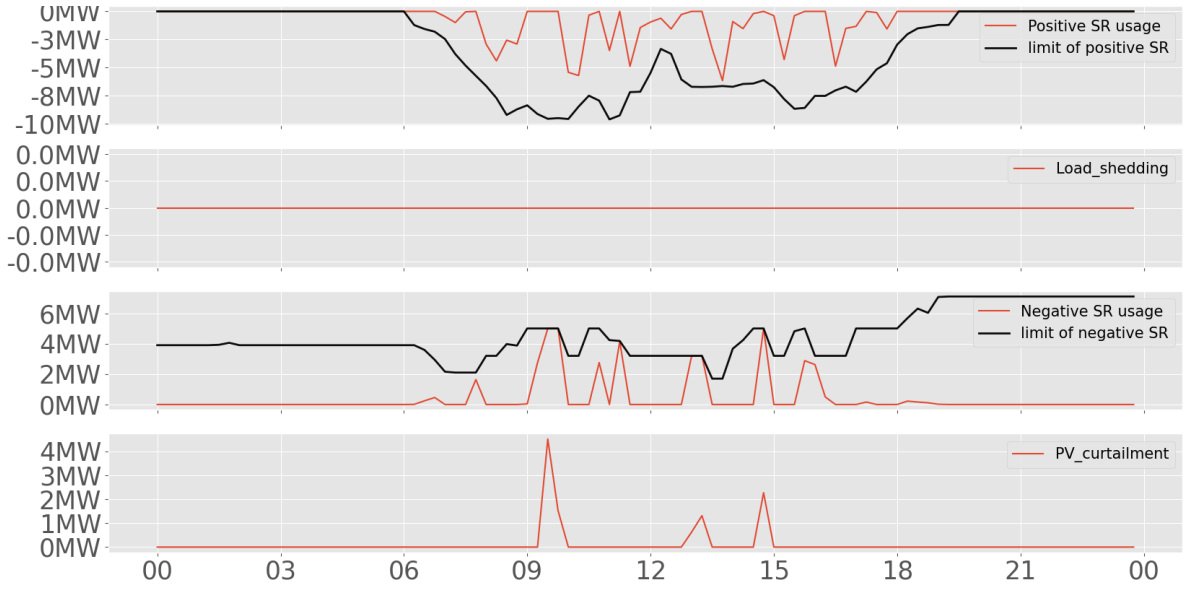


Figure 4.14 Detail of real data evaluation phase of typical case

Hence, as shown in Table 4.3, the final system cost is calculated according to the calculation detail of section 4.1.3. In the result of the simulation example, the extra fuel consumption by using positive spinning reserve costs 3,090 dollars, the PV curtailment costs 102 dollars. The final daily system cost is therefore the sum of genset cost, PV cost, and the cost of system imbalance, resulting in 181,990 dollars.

Energy system components		Example case			No PV case
		Energy produced (kwh)	Unit production cost ($/kwh)	Total cost ($)	Total cost ($)
Genset cost	Active SR cost	80791	0.0188	1519	0
	Diesel	N/A	0.188	156085	177456
	Heavy fuel	N/A	0.146		
	Startup cost	N/A	N/A	0	0
	Shutdown cost	N/A	N/A	280	0
	OM cost	N/A	0.01	17040	17040
PV Energy cost	Produced	99425	0.04	3977	0
	Effectively Used	96863	0.04	3875	0
	PV curtailment	2562	0.04	102	0
System imbalance cost	Use of negative spinning reserve	10029	-0.94	-943	0
	Use of positive spinning reserve	16434	0.188	3090	0
	Load shedding	0	9999	0	0
Total system		1161668	N/A	181990	194496
*Result based on a 24-hour simulation with 15-min temporal resolution					

Table 4.3 Example of HES cost composition

The reference case (same setup but no PV energy), shown in Figure 4.15 till Figure 4.17 in a format of overview, has a final system cost of 194,496 dollars, which is composed of genset fuel cost of 177,456 dollars and operating/maintenance cost of 17,040 dollars. Compared to this no PV case, the example case has an economic saving at around 6.4%, with a PV energy share around 8.5%.

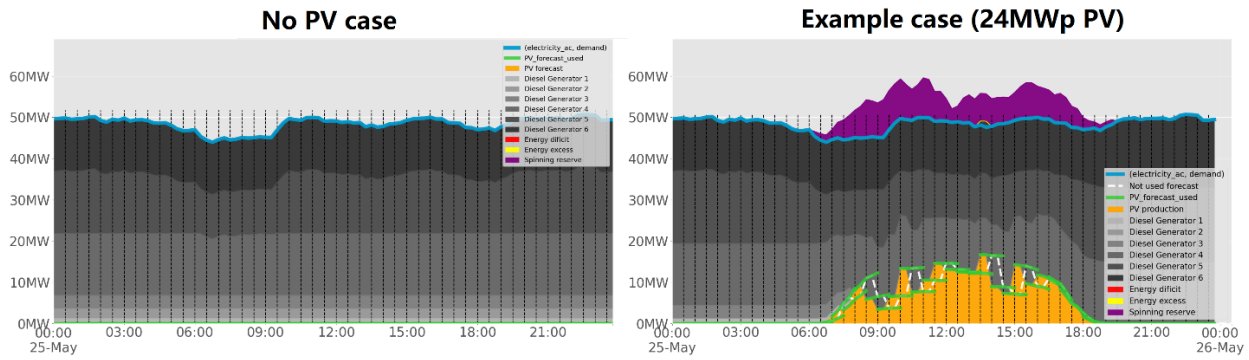


Figure 4.15 Phase of generating generators planning of no PV case

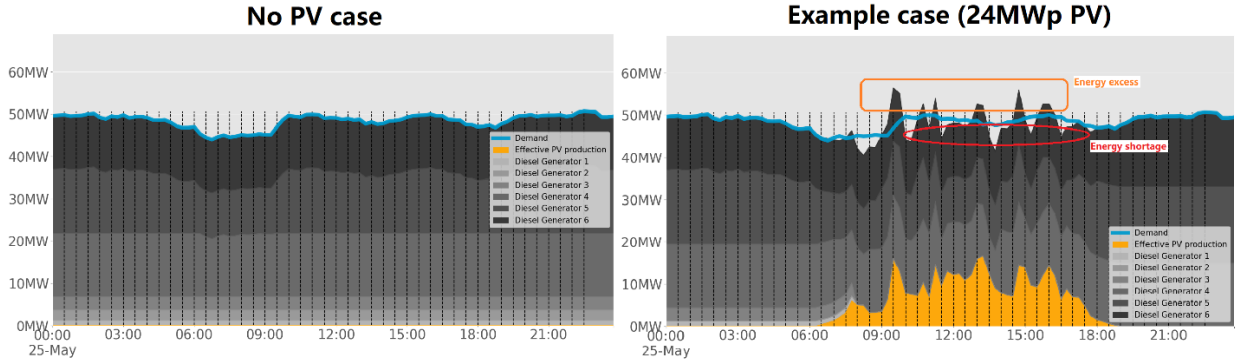


Figure 4.16 Phase of real data evaluation of no PV case

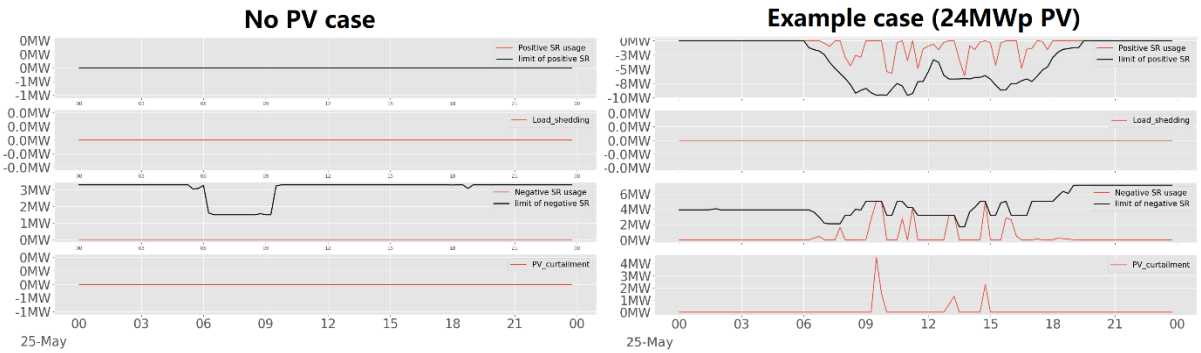


Figure 4.17 Detail of real data evaluation phase of no PV case

These results illustrate a typical case study and the calculation of total system cost. In chapter 5, more typical cases and scenarios for answering the research question of this thesis are tested and detailed.

Chapter 5

5. Analysis of the sensitivity of the HES performance to operational setup parameters

Contents

SUMMARY OF CHAPTER

This chapter presents the general approach of HES simulation, including the logic and general workflow, the forecast methods used, and finally the sensitivity to different parameters such as PV penetration rate, forecast uncertainty used for spinning reserve sizing, etc.

Following the general approach, we implement five scenario analysis designed for answering the research questions of this thesis using days representative of the local meteorological conditions. A test with annual data is achieved to give a more complete evaluation of different parameters in the HES, typical days are used to demonstrate the concrete influence of different parameters.

The first scenario aims to find the added value of using solar forecasts in a Hybrid Energy System (HES), three main forecast methods including the Complete History – Persistence Ensemble (CH-PeEn), Markov chain mixture (MCM) forecast, and Perfect forecast are used in the simulation to provide a full range of potential gain. The CH-PeEn and the Perfect forecast are chosen to be the lower and upper bound of the gain range since they are frequently used as a reference in statistical performance evaluation work. The case with the MCM approach method is supposed to be between the upper and lower bound of system performance range.

The second scenario is designed to explore the influence of PV penetration rate on the HES performance, with different forecast methods at different PV penetration rates, and aims to determine the optimal PV penetration rate.

In the third scenario, we study the influence of different reserved capacity sizing methods on the system performance, for different PV penetration rates. Here, the reserved capacity is sized with different uncertainty ranges of probabilistic forecast and extra buffer quantities. For the last two scenarios, we analyze the influence on system performance when using different forecast horizons and genset dispatch update times.

Some of the different scenarios listed above have only been tested on representative days to save computation time. Only for a selection of parameters, the analysis has been made on a year-round data to validate the results on a broader time range.

To conclude, this chapter shows the detail of the off-grid isolated HES performance when using different solar forecast methods in different setups. The results bring some guiding information for spinning reserve sizing and operating storage-less HES.

5.1 Approach

5.1.1 General approach and parameters

In this section, we set up a simulation framework to test and analyze different scenarios corresponding to different configurations of the HES to answer the different research questions of this thesis. In this framework, some parameters are fixed, and other parameters are configurable, allowing us to test different scenarios according to the different needs of evaluation.

Some parameters are related to the configuration of the HES itself, such as the level of power consumption, the characteristics of the different diesel generators, and the PV capacity. Only the impact of the PV capacity will be assessed in our scenarios. By changing this parameter, the optimal penetration rate and the economic gain using different forecast methods can be evaluated.

Other parameters are related to the operation of the HES: the solar forecast method and the parameters of the optimization process used by the PMS. The first parameter is the time horizon for the dispatching, equal to the forecast horizon. A longer forecast lead time could help to better anticipate the evolution of the PV power generation but forecasts having a lower accuracy at longer time horizon, it may be better not to use it for the dispatch. Yet, a forecast with shorter lead time would less anticipate the weather changes that could result in too frequent start/stop of gensets, reducing their lifetime. The determination of the optimal lead time for the dispatch is one of the goals of this chapter.

Hence, as the second parameter, we consider Genset order update time, which is a compromise between the forecast horizon and the Genset order update time for dispatch. In the real-world application of our case study, the forecast used has a 1-hour forecast horizon. Following this case, we firstly set the forecast lead time at 1-hour and the Genset order update time at its half of 30-min. We have tested different values for these two parameters to analyze the sensitivity of the performance of the PMS to the choice of these parameters.

For the sake of clarity, we have used three simple forecasting methods that are presented in chapter 3: the Perfect Prognosis (PP), the Complete History – Persistence Ensemble (CH-PeEn) and the Markov Chain Mixture (MCM) solar forecasting methods. The effect of using one or the other forecasting algorithm on the performance of the HES will be evaluated. The latter two are probabilistic forecasts: they provide several quantile levels of possible PV productions rather than one single “deterministic” timeseries. We decided to use as the main driver for the Genset dispatching, the P_{50} quantile level of the probabilistic forecast: if correctly calibrated, it means that the effective PV yield for the corresponding lead time has 50 % chance of being higher or lower than this forecasted quantile value.

Another forecasting parameter used by the HES is the forecast uncertainty range, also called interquartile level, which is the difference between two quantile levels of probabilistic forecast. This interquartile level is used to dynamically size the spinning reserve, which is used to balance the potential forecast errors. In general, the uncertainty range $P_{50} - P_{\varepsilon}$ below P_{50} is used to size the spinning reserve, where ε varies from 0 to 50 depending on the scenario needs. Different strategies for sizing the spinning reserve (different interquantiles and additional buffer) are evaluated.

Besides these parameters, the data used is also *configurable*. Some of the scenario listed above are tested using reference days in order to avoid computation time, while other are tested on a year-round data. In the end, the result obtained with reference days are tested in a year-round data

5.1.2 Complementary explanation of parameters

For an off-grid storage-less context, the choice of this uncertainty range is very important because the positive spinning reserve is the main source to mitigate the system variability with the curtailment or clipping capacity of the PV system. In our case study, an overestimated forecast has a different impact on our energy system compared to an underestimated forecast. The overestimation results in an energy shortage situation. In a storage-less off-grid energy system, the only source for short-term balance is spinning reserve (starting a diesel generator needs a much longer time). The energy shortage is hence firstly mitigated by the positive spinning reserve and, if still needed, by load shedding, which is highly undesirable. The forecast underestimation results in an energy excess compared to the expectation. This case can be easily regulated by the negative spinning reserve introduced in chapter 4, and if still needed, PV curtailment would be proceeded. Thus, there is a great contrast by the secondary balancing mechanism, between load shedding and blackout against the PV curtailment, which is considerably less costly. Hence, the focus in sizing spinning reserve is mostly aimed to address potential overestimation cases. For example, if a given quantile level (e.g., P_{50}) is considered as deterministic driver for the dispatch optimization, the uncertainty range used for SR sizing would be $P_{50} - P_0$, since we consider only the overestimated cases with the positive spinning reserve. If the forecast method is properly calibrated, the case of 50% forecast uncertainty counts between P_0 and P_{50} is a safe solution in terms of security, to cover all the overestimated cases, but it is also relatively costly in term of diesel consumption for the large spinning reserve.

As shown in Figure 5.1, the difference between $P_{50} - P_0$ and $P_{50} - P_5$ is significant –almost 15 % of the normalized PV production, due to heavy tailed distribution of the surface solar irradiance. But the cost of keeping extra spinning reserve to cover this difference is relatively high. [89] indicates that different probabilistic forecasts provide different information on uncertainty range, the use of forecast should be adapted to the forecast method. In some literature work like [90], it uses the P_{10} and P_{50} for the economics analysis by calculating the exceedance probabilities. Hence, we add several similar options that use a smaller uncertainty range for spinning reserve sizing like $P_{50} - P_5$, $P_{50} - P_{10}$, $P_{50} - P_{15}$, etc., to assess whether taking some moderate risk may reduce the system cost.

However, it doesn't mean that using higher uncertainty range is better: a too high uncertainty range for spinning reserve sizing requires a very important generator capacity, which is very costly and even unfeasible for high PV penetration rates. At the same time, the uncertainty range used could not be too small since it could probably not be enough to cover the system uncertainty.

For probabilistic forecasting, there is a focus on ensuring a good point forecast, but also providing a probabilistic confidence range with different percentiles. These forecast percentiles should correspond to the actual measurement percentile, to which is designated as calibration. An uncalibrated forecast can pose a challenge to spinning reserve sizing and hence compromise the energy system management.

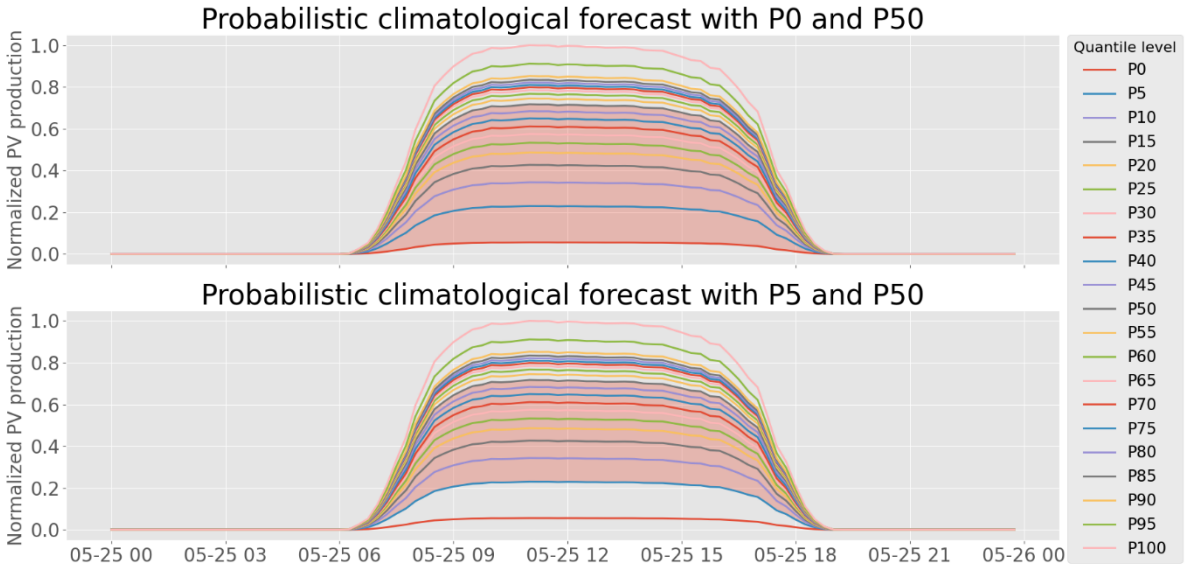


Figure 5.1 Example of inter-quantile level of probabilistic forecast

In fact, there is never a perfectly calibrated model; as [90] points out, even if the likelihood is small, a system imbalance situation could still occur even with a spinning reserve sized with uncertainty range $P_{50} - P_0$. Besides, if the quantile level used for optimization dispatch is not well calibrated, it results in breaking the minimization trade-off between PV curtailment and SR^+ quantity.

Therefore, for practical applications, not only the calibration of the probabilistic forecast methods used should be ensured, but an extra buffer of positive spinning reserve (SR+) should also be added. The aim of this additional buffer is to mitigate the uncertainty derived from the forecast method. In our simulation, an extra buffer equivalent to a fixed percentage of the PV installed capacity is added to the positive spinning reserve sizing. Here, different fixed percentages have been considered: 5% and 10%. Finally, there are 12 combinations in total for the strategy used to size the spinning reserve (6 different inter-quantiles and 2 values for the buffer). Hence, the final SR sized SR^+ is composed of two parts, a part of dynamic sizing SR_F^+ fed by the forecast method and a part of fixed extra buffer SR_0^+:

$$SR^+(t) = SR_F^+(t) + SR_0^+ \quad (56)$$

In brief, running the simulation with different parameter settings will yield different results. The detail of all parameters is summarized in Table 5.1. The consideration of the 19440 combinations resulting from the parameter values listed in Table 5.1 would be too computationally intensive (more than 1000 hours) and complex to analyze. For the sake of clarity, we decide to conduct a one-parameter-at-time analysis while considering multiple parameters when needed. Other external parameters like the fleet of gensets and the energy demand remain constant.

The first analysis (section 5.2) aims to demonstrate the general workflow of our analysis, and to find out the added value of using solar forecast method, through an annual data simulation.

The second analysis (section 5.3) is designed to explore the influence of PV penetration rate on the system cost with different sky states and forecast methods.

<table>
<tr><th>PV rate in capacity*</th><th>Forecast method</th><th colspan="2">Uncertainty range used for energy buffer sizing</th><th>Forecast Lead Time</th><th>Update Time</th><th>Type of day</th></tr>
<tr><td>0%</td><td rowspan="3">CH-Persistence Ensemble (CH-PeEn) forecast</td><td>P50-P25</td><td rowspan="9">With extra buffer = X% PV capacity</td><td>1-hour</td><td rowspan="2">30-min</td><td rowspan="3">Variable</td></tr>
<tr><td>25%</td><td>P50-P20</td><td rowspan="2">2-hour</td></tr>
<tr><td>50%</td><td>P50-P15</td><td rowspan="4">1-hour</td></tr>
<tr><td>75%</td><td rowspan="3">Markov chain based forecast</td><td>P50-P10</td><td>3-hour</td><td rowspan="3">Clear</td></tr>
<tr><td>100%</td><td>P50-P5</td><td rowspan="2">6-hour</td><td rowspan="3">2-hour</td></tr>
<tr><td>125%</td><td>P50-P0</td></tr>
<tr><td>150%</td><td rowspan="3">Perfect forecast</td><td rowspan="3">12-hour</td><td rowspan="3">3-hour</td><td rowspan="3">Year round data</td></tr>
<tr><td>175%</td></tr>
<tr><td>200%</td></tr>
</table>

Table 5.1 Parameters of the PMS simulation considered in this analysis

The third one (section 5.4) aims to find out the influence of positive spinning reserve sizing strategies on the optimal PV penetration capacity and resulting costs of the HES.

The last two analysis (section 5.5 and section 5.6) are set to find out the influence of different forecast lead times and Genset update times on the system cost.

In section 5.7, the best combination of parameters of the setup are summarized and discussed.

5.2 Quantifying the added value of using solar forecasts

This analysis aims to quantify the added value of using different forecast methods in an isolated storage-less HES. However, it is difficult to define the baseline of a HES without forecast since the system performance would be quite different according to the different management strategies. Hence, this was addressed by comparing three different models with a varying degree of accuracy.

The best case means that all the PV energy produced could be totally used, and hence less diesel used and less PV curtailment. The worst but realistic case is to use an easy-implemented forecast method, which results in higher spinning reserve need and diesel consumption, also more PV curtailment.

5.2.1 Simulation setups

The simulation setup of this scenario is shown in Table 5.2. The no PV case is a reference to show the gain of using PV energy. A PV penetration rate of 50% is here considered since it is close to the actual PV installation of the mine in Mali of our case study. The forecast methods are chosen to provide a full range of economic gains when using different forecast methods, including the CH-PeEn, the MCM model, and the Perfect Prognosis forecast, which were introduced in chapter 3. The uncertainty range $P_{50} - P_0$ is chosen as a reference for SR sizing, with an additional buffer of 10% PV installation capacity, aiming to cover potentially underestimated uncertainty (e.g., an over-estimated P_0). The annual data is used to get a broader view of the system performance, two types of sky states are tested to demonstrate the concrete daily situation. In the real-world application of case study, the forecasts have a 1-hour horizon. Following this case, we firstly set the forecast lead time at 1-hour and the Genset order update time at 30-min.

PV rate in capacity*	Forecast method	Uncertainty range used for energy buffer sizing		Forecast Lead Time	Update Time	Type of day
0%	CH-Persistence Ensemble (CH-PeEn) forecast	P50-P25	With extra buffer = X% PV capacity	1-hour	30-min	Variable
25%		P50-P20		2-hour	1-hour	
50%	Markov chain based forecast	P50-P15		3-hour		Clear
75%		P50-P10			2-hour	
100%		P50-P5		6-hour		
125%	Perfect forecast	P50-P0		12-hour	3-hour	Year round data
150%						
175%						
200%						

Table 5.2 Simulation setup for quantifying the added value of using different forecasts

5.2.2 Result

The simulation result with above setting is shown in Figure 5.2. The no PV case is composed of the Genset fuel cost and the Genset maintenance cost, which represent around 9.5% of the total cost. Notice that this cost is fixed for all the cases since it is the cost of maintaining the Genset installation and it is the same for all the cases.

Compared to the no PV case, with 50% PV penetration rate (which corresponds to around 12% PV energy share), the best case is obtained - as expected - with the Perfect forecast, which reduces the total system cost by around 8.0%. By using MCM model forecast, the system cost could be saved by around 5.2%, which represents 65% of the maximal cost saving obtained with the perfect prognosis. And the case using CH-PeEn forecast reduce 2.5% of annual system cost, which represents 31.25% of the optimal case. The CH-PeEn method could be used as a baseline of system performance reference since this method is easy to compute and only based on the past and current data.

In terms of the range of the gain of using solar forecast, the system cost is reduced between 2.5% and 8.0% depending on the considered forecast model and compared to a no PV case. By looking into the detail of this figure, it's possible to observe that the basic fuel consumption is quite similar for the three cases; the difference in system cost comes from the spinning reserve sizing and some PV curtailment. Since the investment cost of the PV system is considered, wasting the PV energy through curtailment has a relatively important penalty cost.

The perfect forecast can accommodate a higher PV energy share since there is neither PV curtailment nor SR sizing, the final system cost is hence the lowest. The MCM forecast has a slightly higher fuel cost and spinning reserve cost, which results in a higher final system cost. The CH-PeEn model is a more conservative mode in terms of forecast sharpness (similar to a climatological forecast) and, thus, has a greater uncertainty, resulting in a higher spinning reserve sizing cost and, consequently, the highest final system cost among three cases.

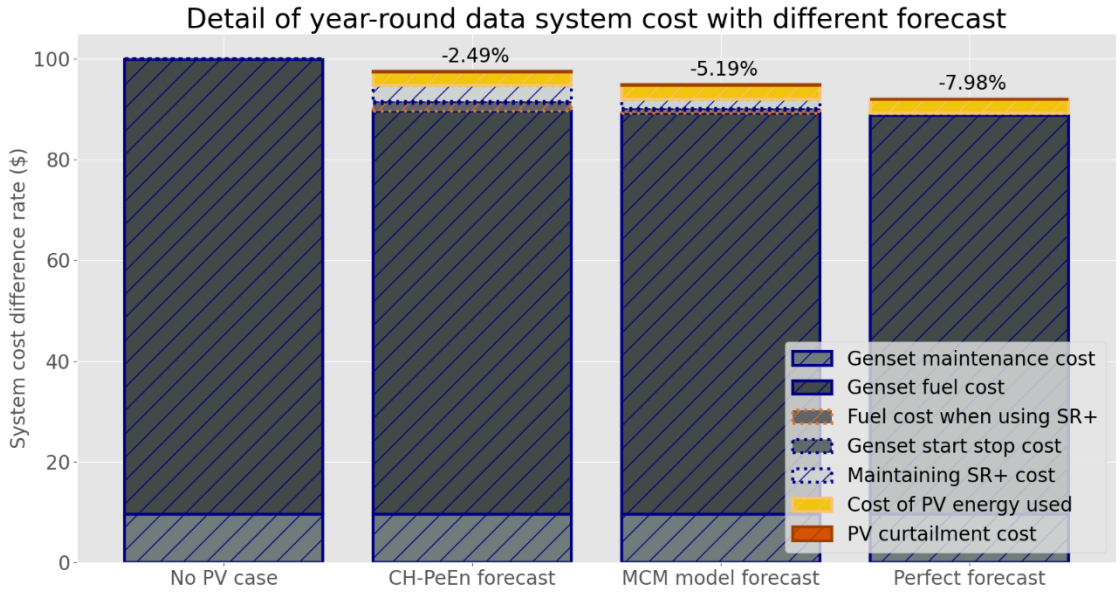


Figure 5.2 Cost saving of annual situation compared to no PV case by using different forecasts

Then, as shown in Figure 5.3, the top of the figure above zero is the sum of the additional cost, the bottom is the sum of the cost saving compared to no PV case. By looking into the details, the cost of spinning reserve sizing when using CH-PeEn is much more important than the cost of actual usage of spinning reserve, since the uncertainty provided by CH-PeEn is large and the actual variability at clear day is small. To mitigate this, either using other more accurate forecast methods, which could provide a lower accuracy (sharper uncertainty range), or taking a more risk-favorable dispatch (with possible imbalance situations), using a spinning reserve defined with a smaller inter-quantile range. Then, we can also observe that a more accurate forecast method could allow a higher PV penetration rate.

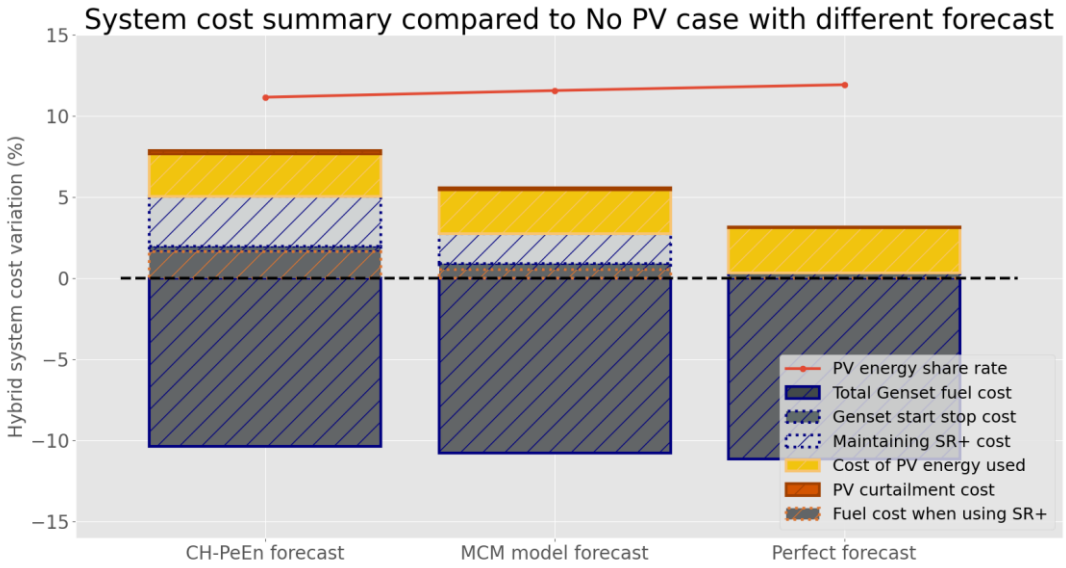


Figure 5.3 Cost difference of annual situation compared to no PV case by using different forecasts

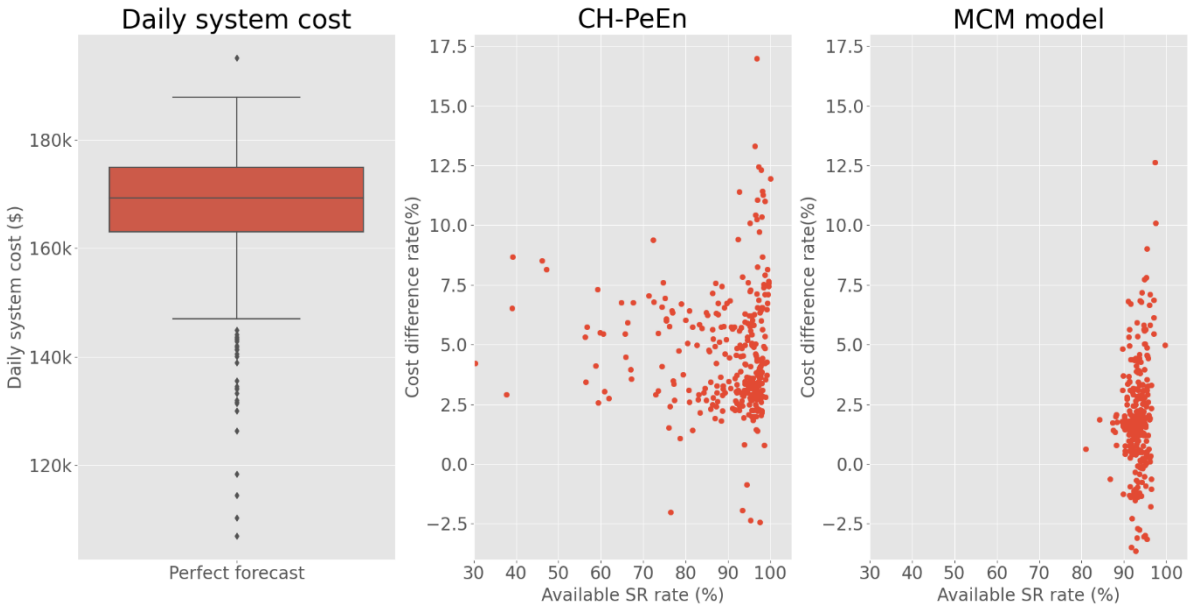


Figure 5.4 Link between risk and cost difference of differents methods compared to Perfect case

In terms of the relationship between the risk and the economic performance, the available spinning reserve (difference between the maximal and actual production of the gensets) is used to represent the risk: the lower this value and higher the risk of a shortage. The results of the annual analysis are displayed in Figure 5.4, where the boxplot in the left is the distribution of daily system cost using Perfect forecast, the scatter plots in the middle and right are the relationship between available SR and cost difference compared to Perfect case. For perfect forecast, there is no SR sized since the perfect prognosis has by definition no uncertainty. Between MCM model forecast and CH-PeEn forecast, the MCM model forecast has higher risk compared to the CH-PeEn case, as the histogram shown in Figure 5.5, where the available SR of CH-PeEn cases has a larger spread but MCM model cases are concentrated around 93%. Yet, the annual system cost of MCM model case is also slightly lower, which signifies an underlying relationship between the risk and economic gain.

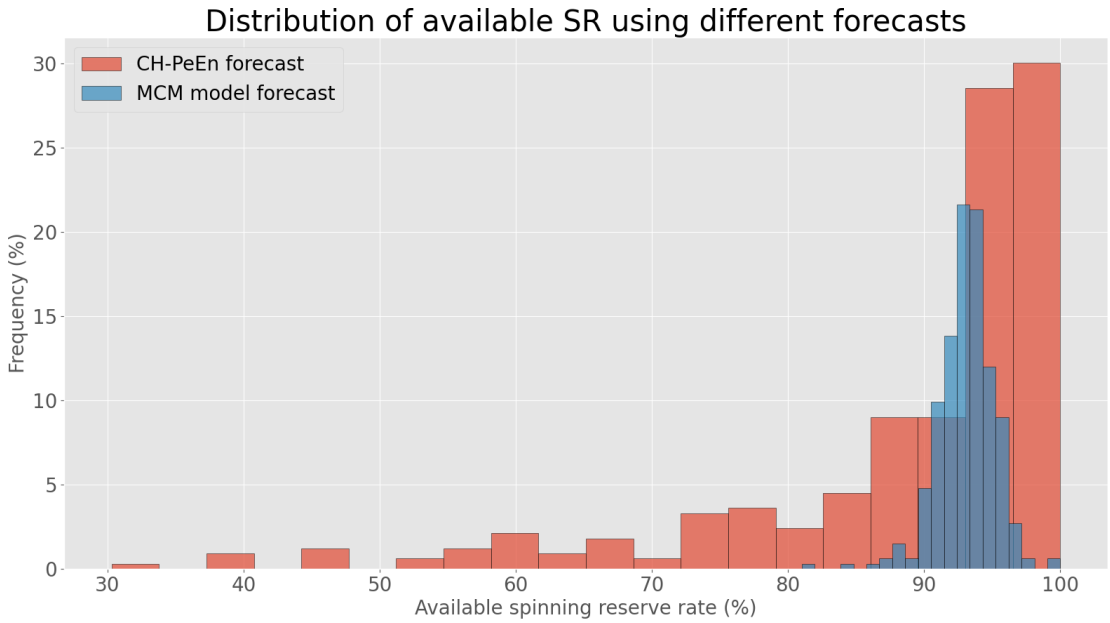


Figure 5.5 Distribution of available spinning reserve rate using different forecasts

This general result is based on the chosen parameters, with a redundantly sized energy buffer, and with a 1-hour forecast horizon. If the SR is sized smaller, or if there is another better combination of forecast lead time and Genset Update, the general system performance could possibly be improved. For this purpose, in the following section, we explore the concrete influence on system cost with different parameters.

5.3 Influence of different PV penetration rates

5.3.1 Simulation setups

The scenario is set to find out the influence of PV penetration rate on the HES performance, while considering different sky states and the different forecast methods detailed before in chapter 3. To avoid any system imbalance situations, the spinning reserve is sized with $P_{50} - P_0$ uncertainty range with an extra buffer as explained in section 5.1.2. The update time remains the same as before, set at 30-min. The forecast lead time is mainly set for 1-hour, a 12-hour test is also achieved to evaluate the impact of forecast lead time. Then, both variable and clear day are tested in this scenario to give a complete evaluation, the final setup is shown in Table 5.3.

PV rate in capacity*	Forecast method	Uncertainty range used for energy buffer sizing		Forecast Lead Time	Update Time	Type of day
0%	CH-Persistence Ensemble (CH-PeEn) forecast	P50-P25	With extra buffer = X% PV capacity	1-hour	30-min	Variable
25%		P50-P20		2-hour	1-hour	
50%		P50-P15		3-hour		Clear
75%	Markov chain based forecast	P50-P10		6-hour	2-hour	
100%		P50-P5				
125%						
150%	Perfect forecast	P50-P0		12-hour	3-hour	Year round data
175%						
200%						

Table 5.3 Simulation setup selected to assess the influence of different PV penetration rates

5.3.2 Results

Clear-day

The simulation results of a typical clear-sky day are shown for different forecast methods in Figure 5.6. It can be seen that for this given data and setup, the system cost decreases for PV penetration rate ranging from 0 to 50% both for CH-PeEn and MCM forecast. Above 50 % penetration rate, system cost starts to increase: the storage-less HES cannot make use of surplus PV generation and the HES has to curtail PV power production to ensure the power supply/demand balance. As a result, as PV penetration increases, the growing PV costs are each time met with more wasteful PV generation and lower savings in fuel needs. For CH-PeEn and MCM, the dispatch could ensure a system balance until respectively 75 and 100% PV penetration. After these values, the dispatch is unsolvable since the forecast uncertainty becomes too important. As introduced before, the uncertainty range is used to size spinning reserve. A high uncertainty reflects in high power need, which makes the situation becomes unfeasible since the Genset installation has a constraint in operation range and limited nominal power. Hence, the power management strategy for an isolated storage-less HES should be precisely designed according to the characteristic of the forecast method, PV penetration rate, and the

Genset installation of energy system. When using Perfect forecast, the system cost decrease from 0 to 175% PV penetration rate. Then the system cost goes up due to the PV curtailment and the reason of storage-less design, which has less capacity in making use of PV energy. Its optimal penetration rate is hence 175% for this given condition.

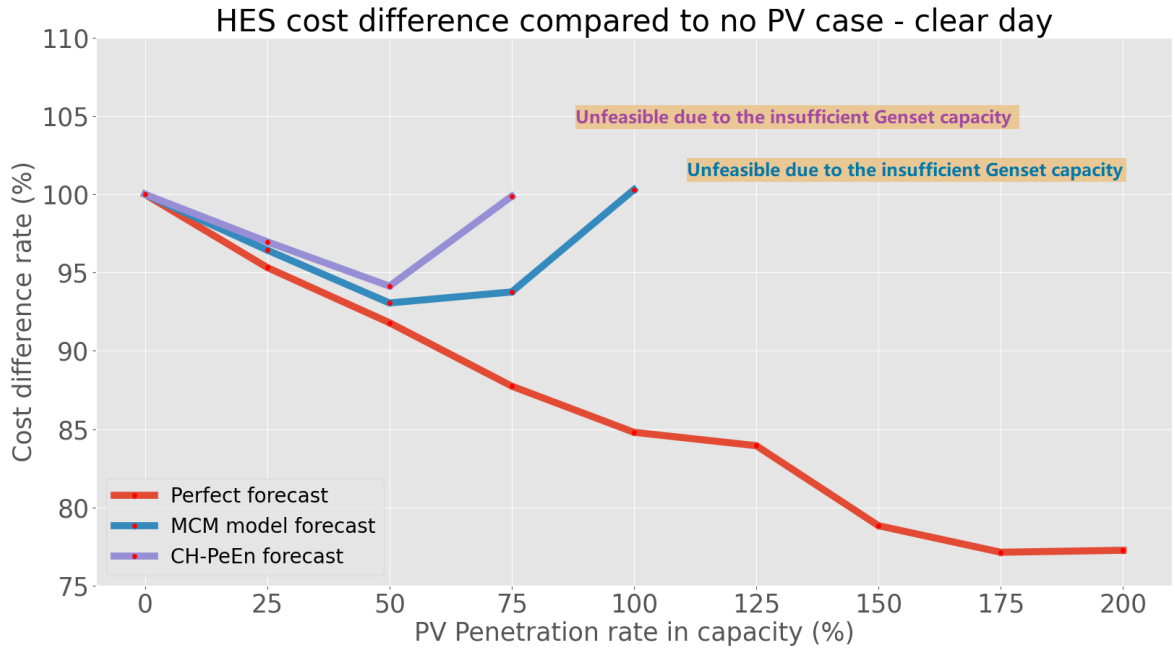


Figure 5.6 Simulation result of clear day situation with P0-P50 and 10% buffer

As the detail of the system cost composition shown in Figure 5.7, where the CH-PeEn forecast shows the highest system cost due to its conservative nature as mentioned in section 5.2.2, it brings maximum 5.6% economic saving with 50% PV penetration. When using the MCM model forecast, the energy system is able to accommodate a little bit more PV than the previous case since its forecasts are sharper (i.e. smaller uncertainty range), hence smaller need for Genset capacity, it allows to have around 6.2% economic saving at 50% PV penetration. And the perfect forecast case reduces around 22.9% system cost at 175% PV penetration.

When the PV penetration increase and the PV production is too large to be absorbed by the HES, as explained in chapter 4, the negative spinning reserve would be used first, when it becomes too important, it comes with PV curtailment, up to a point where the system cost is even higher than the no PV case. However, we can also notice that even though there is some PV curtailment, the Genset fuel consumption has a trend to decrease.

However, when the system arrives at its saturation, the PV energy would be curtailed as explained in section 5.1.2. Also, as mentioned before, storage-less design limits the capacity to accommodate the PV energy, which means that a higher PV penetration rate does not mean a better economic gain.

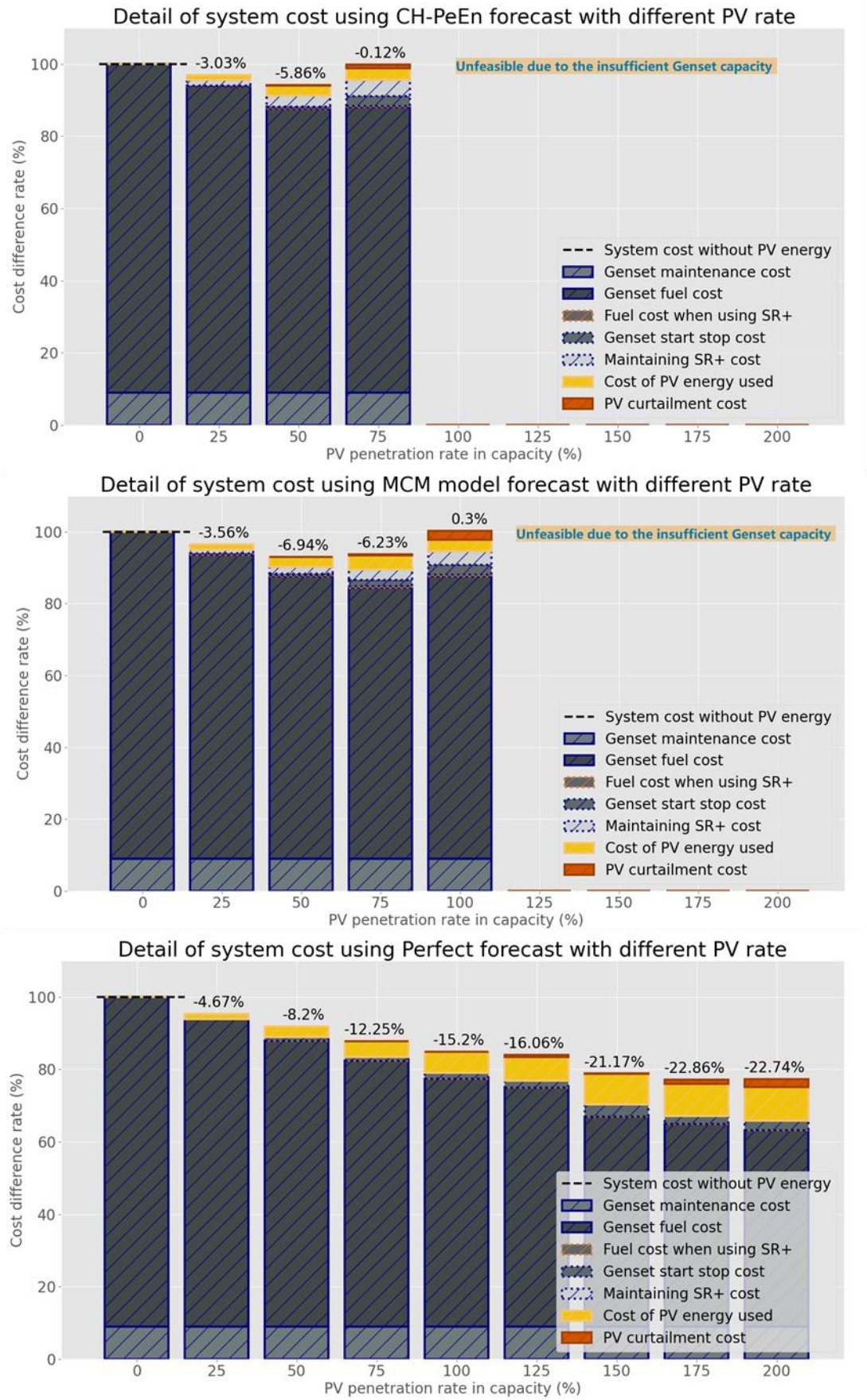


Figure 5.7 Simulation result detail of clear day situation

By looking into the detail of Figure 5.6, the decreasing slope of perfect forecast become strangely smaller at 50% PV penetration, but the MCM case does not. This situation is caused by the limited forecast horizon, hence, I run the simulation with 1 and 12-hour horizon for perfect forecast and 1-hour for MCM model to clarify the reason.

As shown in Figure 5.8, for Perfect forecast dispatch order, some generators are turned off in the middle of the day and returned on later to avoid energy excess. This action increases the machine ageing and extra fuel consumption as explained in chapter 4. But in the dispatch order when using MCM model forecast shown in Figure 5.9, the main generators are switched on almost all the time, which avoids the change of the machine state. This fact makes the system cost of using perfect forecast a little bit higher than it should be since the startup and shutdown cost is considered and the CAPEX of Genset installation is not considered in our simulation.

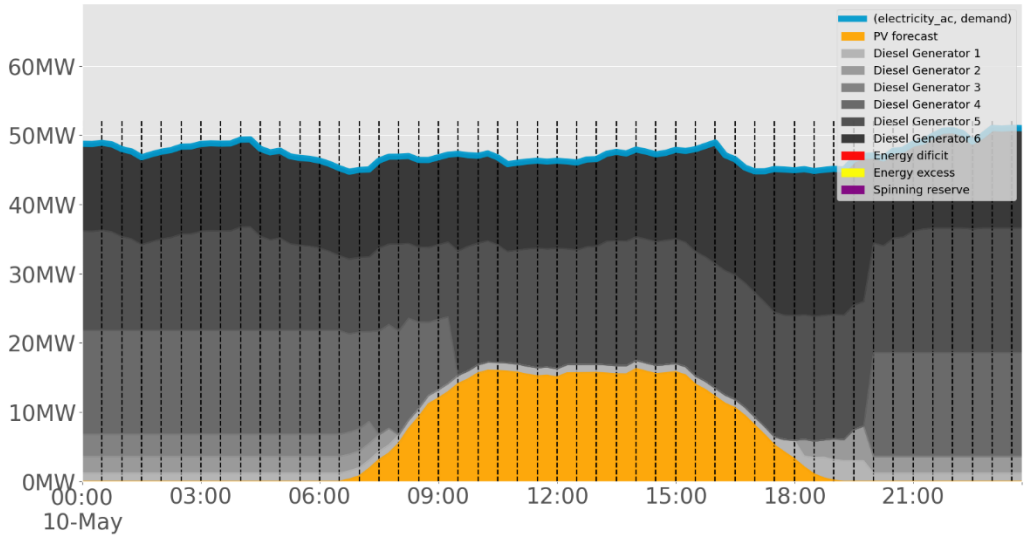


Figure 5.8 Simulation dispatch order with perfect forecast and 1-hour forecast horizon

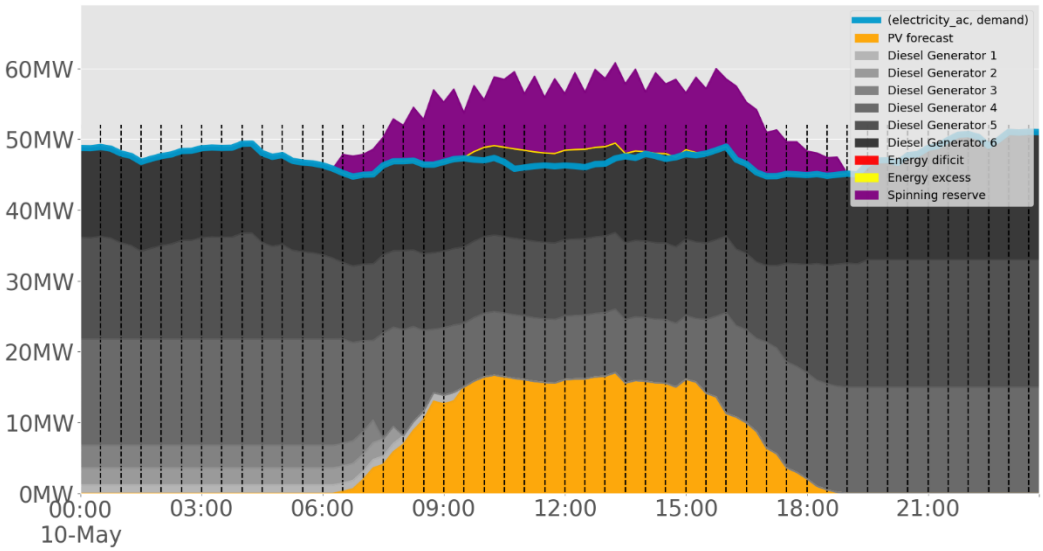


Figure 5.9 Simulation dispatch order with MCM model forecast and 1-hour forecast horizon

But when the forecast horizon is extended to 12-hour as shown in Figure 5.10, the dispatching of generators is much simpler and can avoid the unnecessary start/stop process and the system cost also drops. For more detail on the forecast horizon influence on system cost, an analysis in section 5.5.1 is shown in the latter part of this chapter.

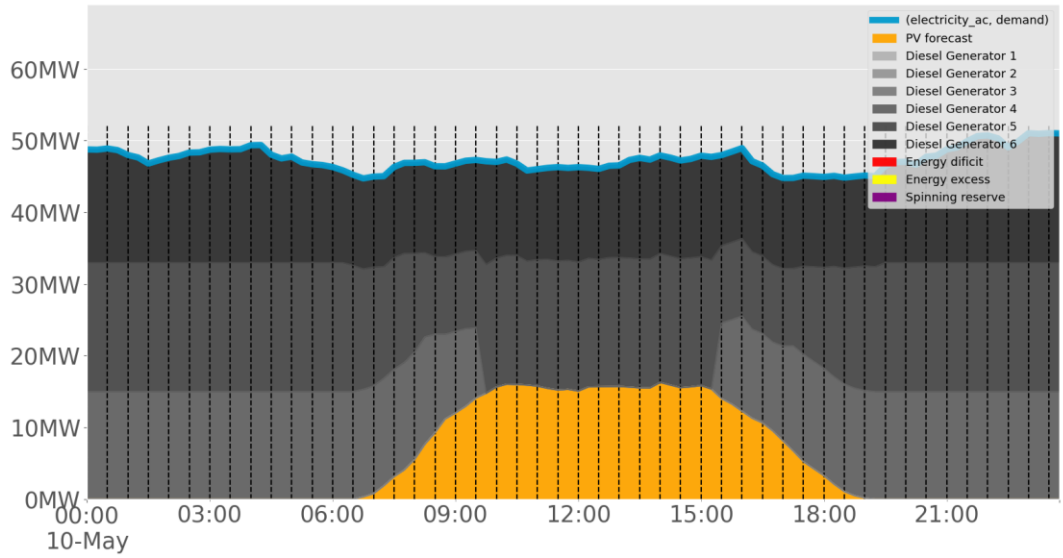


Figure 5.10 Simulation dispatch order with perfect forecast and 12-hour horizon

Variable-day

In a variable situation, the simulation result is shown in Figure 5.11, where the general pattern of using Ch-PeEn and MCM model is similar to clear day, the system cost decrease from 0 to 50% penetration rate and then goes up due to the PV curtailment. Continuing to increase the PV penetration rate, the system becomes unfeasible from 100% penetration rate for CH-PeEn forecast and 125% for MCM model since the need of Genset capacity becomes too high as explained in clear day situation.

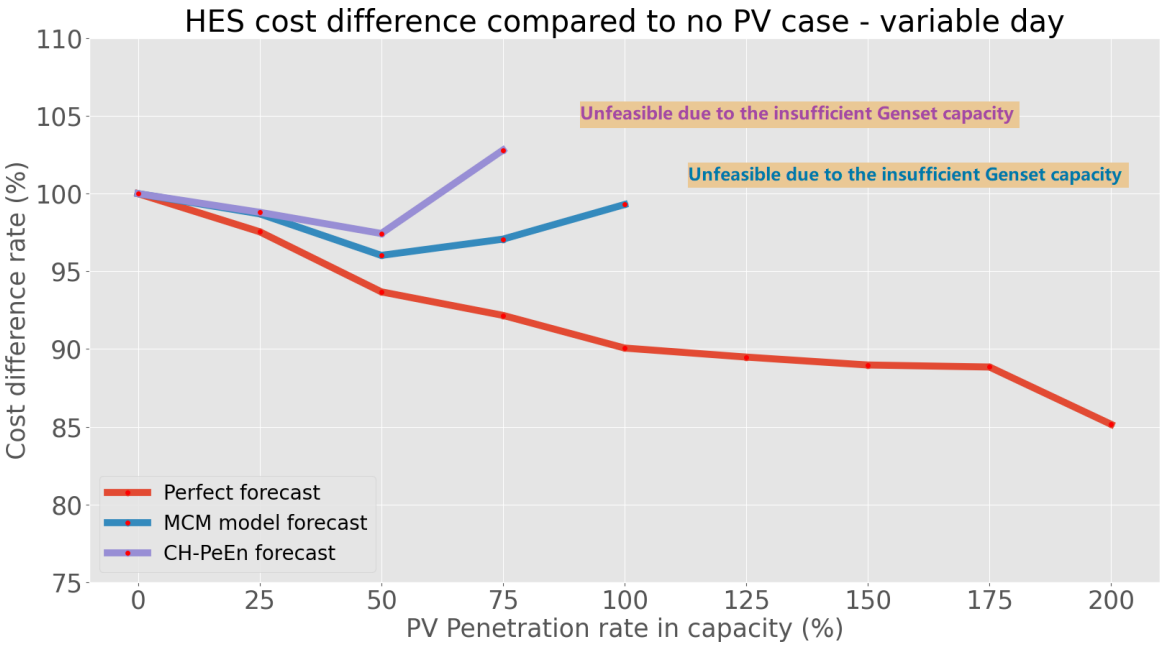


Figure 5.11 Simulation result of variable day situation with P0-P50 and 10% buffer

But for perfect forecast, the maximum PV penetration could be increased to 200 %, which is the highest setup. This penetration rate is higher than clear day situations since the solar resource in variable day is lower. Due to this reason, the general economic gain of each model is less important, a higher PV penetration rate could be hence achieved and reduce the system cost by reducing the fuel consumption. In this weather situation, as shown by the Perfect forecast, using an accurate forecast model would bring a much higher economic gain than other models.

Besides, we can also notice that at 25% PV penetration rate, the economic gain by using MCM or CH-PeEn model is small since the PV energy share rate is similar and low. The detail plot of variable day is not shown here since the detail is similar to the clear day situation.

5.4 Influence of probabilistic forecast uncertainty

5.4.1 Simulation setups

This setup aims to find out the influence of the SR sizing strategy on the optimal PV penetration capacity and resulting costs of the HES. To explore this, we vary the size of spinning reserve by quantifying the forecast uncertainty with different minimum percentiles (from P_0 to P_{25}) while keeping the maximum percentile at P_{50} for Genset dispatching optimization. Additionally, increasing the spinning reserve with a buffer is considered. The simulation setups are shown in Table 5.4, where we only focus on the variable day since clear day has less uncertainty and doesn't need a very high spinning reserve. Since clear day works even with a small SR, changing the SR quantity just means changing the cost of keeping different SR. Then, both CH-PeEn and MCM forecast are chosen for the simulation since both of them come with forecast uncertainty, which make it valuable to explore the influence of the forecast uncertainty on system cost. The Genset update time and the forecast lead time are set at 30-min and 1-hour, respectively.

PV rate in capacity*	Forecast method	Uncertainty range used for energy buffer sizing		Forecast Lead Time	Update Time	Type of day
0%	CH-Persistence Ensemble (CH-PeEn) forecast	P50-P25	With extra buffer = X% PV capacity	1-hour	30-min	Variable
25%		P50-P20		2-hour	1-hour	Clear
50%		P50-P15		3-hour	2-hour	
75%	Markov chain based forecast	P50-P10		6-hour	3-hour	Year round data
100%		P50-P5		12-hour		
125%		P50-P0				
150%	Perfect forecast					
175%						
200%						

Table 5.4 Simulation setup selected to assess the influence of different energy buffer quantities

5.4.2 Results

The cost saving achieved using CH-PeEn and MCM forecast for the feasible PV penetration rates is shown in Figure 5.13 and Figure 5.14, respectively. When no extra spinning reserve buffer is considered, we observe system imbalance in all cases. This highlights a lack of calibration on the forecasting models, as the system imbalance results from actual PV production below the minimal quantile predicted (in other words, P_0 does not actually reflect the minimum percentile). The same occurs when a very risky sizing is assumed by considering P_{20} and P_{25} as minimum percentiles for the uncertainty quantification.

In Figure 5.13, we consider the result with $P_{50} - P_0$ and 10% buffer (the context of example scenario) as a reference. We aim to explore the influence of reducing the quantity of extra spinning reserve buffer, hence another option, which is 5% of the installed PV capacity (half of the reference case), is used for both forecast methods. We can see from the figure that with 5% extra buffer, the system cost could be lower till 2%, and the PV penetration could be higher compared to 10% extra buffer case.

In Figure 5.14, we explore the impact of using different minimum percentiles level, when there are some missing points, it means that there is a load shedding or unfeasible situation, as shown, by increasing the minimum percentile from P_0 to P_{15}, the system cost get lower with maximum 4% reduction for MCM model at 100% penetration rate, the PV penetration rate is extended for a higher level till 175% with a reduction of cost when the SR is sized with $P_{50} - P_{15}$ with 5% buffer. However, when using a higher percentile of lower bound, load shedding appears for the low PV penetration rate, since the uncertainty outliers are mostly appeared in the beginning and the end of the day as shown in Figure 5.12, which are mitigated by the extra buffer. But this buffer is sized proportionally to the PV capacity, when the penetration is small, this buffer becomes insufficient to cover all the uncertainty. The safety range is hence too small and the system imbalance cases would occur.

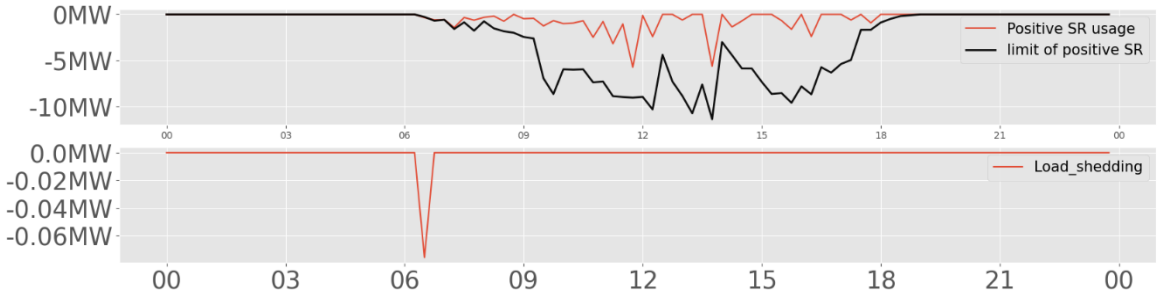


Figure 5.12 Detail in system imbalance at the beginning of the day

When looking into the result in Figure 5.14, the system cost goes up and down, the reason of this situation is based on several facts. Firstly, due to the PV curtailment, system cost goes higher. Then, by increasing the PV penetration rate, some generators all turned off during daytime and their cost is hence avoid and the system cost goes down. But if the PV penetration rate keep increasing, the PV curtailment will increase the system cost again.

For the cases of $P_{50} - P_5$ and $P_{50} - P_{10}$with 10% buffer, the system cost with a smaller SR is even higher, that is due to the forecast horizon in this analysis is limited at 1-hour, some small generators are turned on and a big generator is turned off to meet the energy demand and avoid the energy excess, but the case with greater SR need just keep the larger generator running and hence avoid the startup cost of small generators. For the detail of forecast horizon influence, section 5.5.1 explore its impact.

When the spinning reserve energy buffer is sized with $P_{50} - P_{15}$ with 10% buffer, the PV penetration rate could be extended till 175% for both forecast methods, which is much higher than the initial situation. However, to notice that at 175% PV penetration rate, the system cost is close or even higher than the no PV case due to the PV curtailment as explained before. This fact shows that a higher PV penetration doesn't mean a higher economic gain.

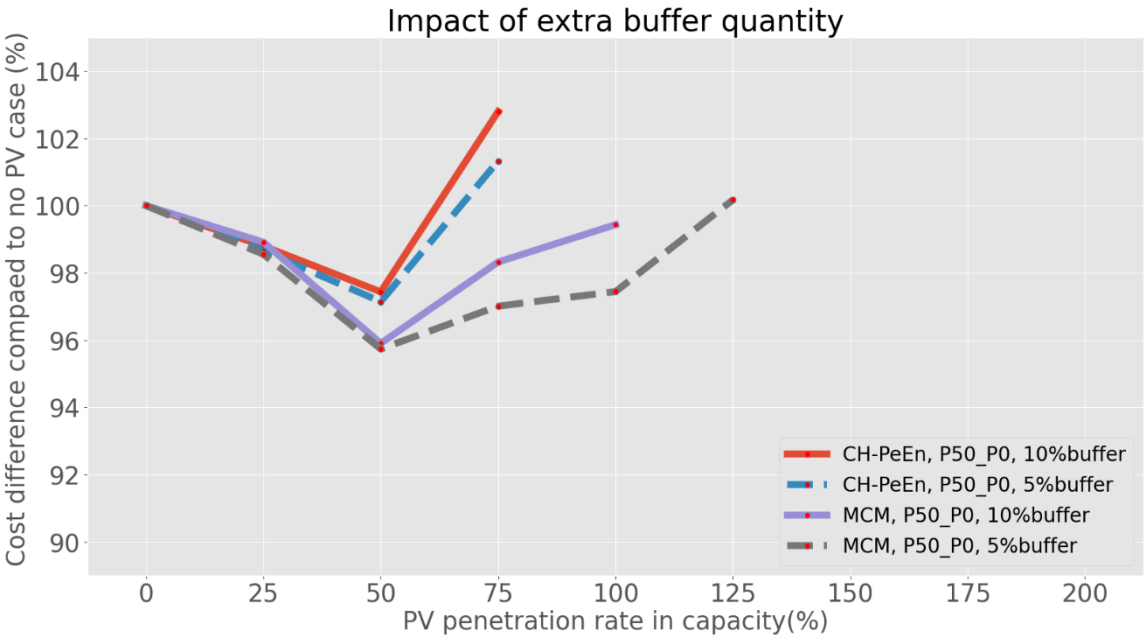


Figure 5.13 Cost difference compared to no PV case with different extra buffer quantities at variable situation

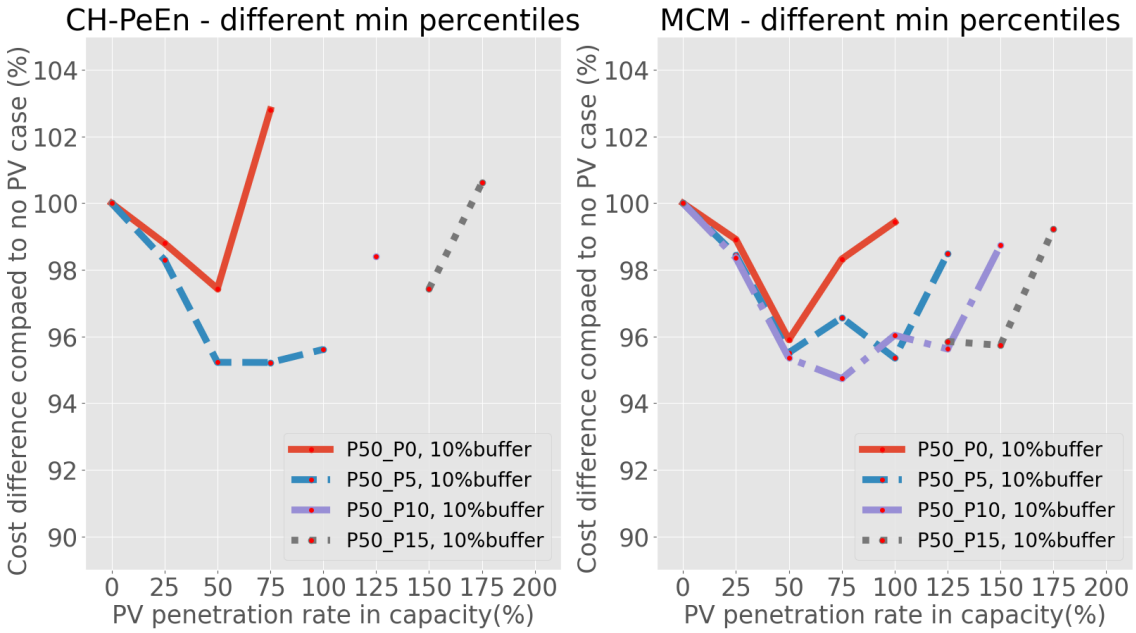


Figure 5.14 Cost difference compared to no PV case with different minimum percentiles at variable situation

From above simulation cases, the detail of available spinning reserve rate of each forecast method with different buffer quantities and minimum percentiles at different PV penetration rate is shown in Figure 5.15 and Figure 5.16. From these figures, several conclusions could be drawn:

Firstly, as shown in Figure 5.15, when the energy buffer quantity decreases, the available spinning reserve rate decreases as well, which signifies the energy system is taking more risk on system stability.

Secondly, as shown in Figure 5.16, the available SR rate gets higher when the PV penetration rate increase since the extra buffer is proportionally sized according to the PV installed capacity, which means that the way to size extra buffer is a redundant design.

Thirdly, for a SR sized with $P_{50} - P_0$ uncertainty range, the difference in available SR rate with 5% and 10% extra buffer is small, if some system imbalance cases are acceptable, sizing a smaller SR could allow a higher economic saving.

Finally, the available SR rate of MCM shown in Figure 5.16 seems too redundant, since the daily average values are mostly higher than 80%. Similar to the CH-PeEn and MCM cases shown in Figure 5.7, where the cost of maintaining SR is more important than the actual SR usage cost. Notice that this kind of redundant spinning reserve design is very costly. If the actual need of SR is always less than 20% of the sized SR, the SR quantity could be sized smaller to reduce the fossil fuel consumption and save the system cost. Therefore, a smart compromise between risk and gain is needed for energy system management, we process hence a deeper look in the link between SR quantity and available SR rate.

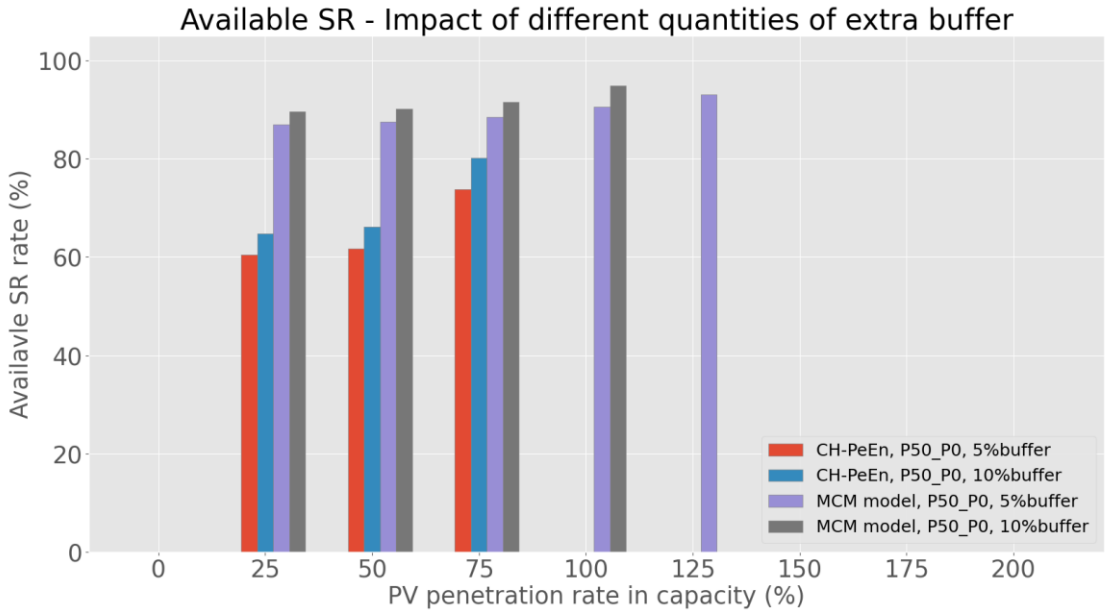


Figure 5.15 HES available SR rate in different energy buffers and PV penetration rates

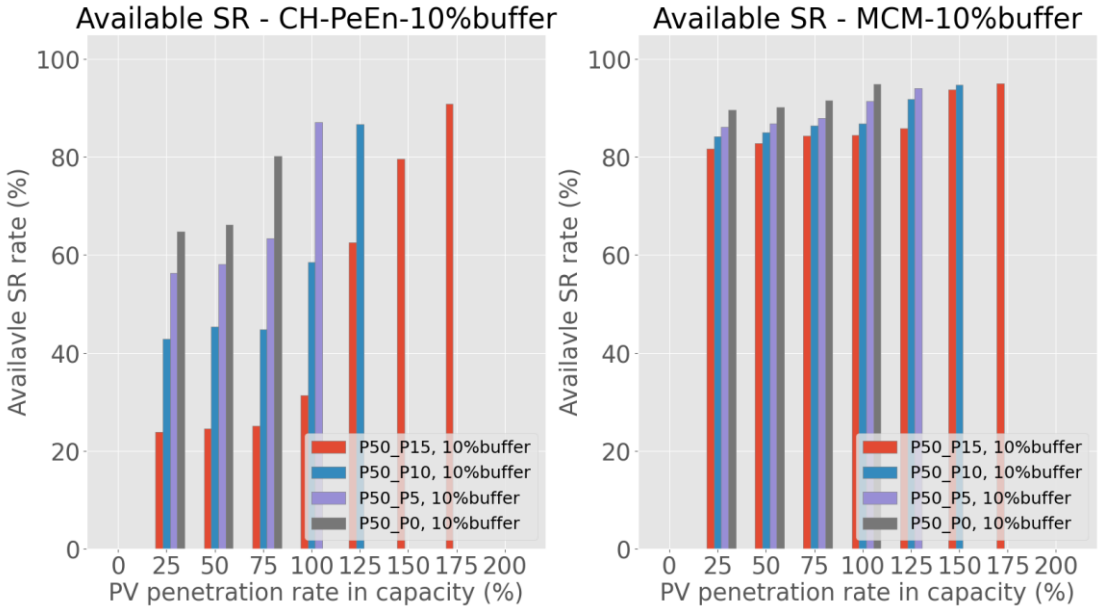


Figure 5.16 HES available SR rate in different minimum percentiles and PV penetration rates

Figure 5.17 and Figure 5.18 show the result obtained with CH-PeEn and MCM model method at 50% PV penetration rate, their optimal spinning reserve quantity: for CH-PeEn and MCM forecast the optimal spinning reserve at this PV penetration rate is obtained using the interquantile $P_{50} - P_5$ with an extra buffer of 5% PV capacity. To notice that the daily stable rate is the fraction of stable time step over total time step. When daily stable rate is 100%, it means the HES has no load shedding case. From these figures, it shows that not only the different energy buffer quantities influence the system stability, but also the spinning reserve management. As shown in Figure 5.18, for MCM model, the $P_{50} - P_5$* and $P_{50} - P_{15}$** cases have same amount of spinning reserve in one day, but their stability result is different due to the wrong spinning reserve organization.

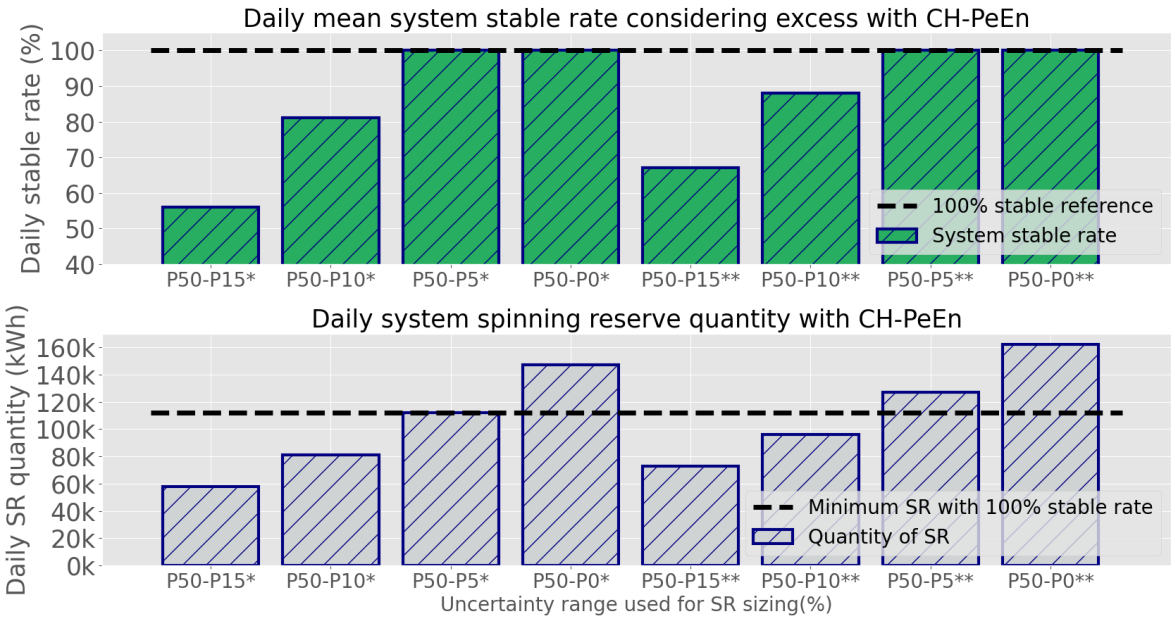


Figure 5.17 System stability and its associated spinning reserve quantity using CH-PeEn forecast with different uncertainties at variable day situation (* means with extra buffer of 5% PV capacity for SR sizing, ** for 10%)

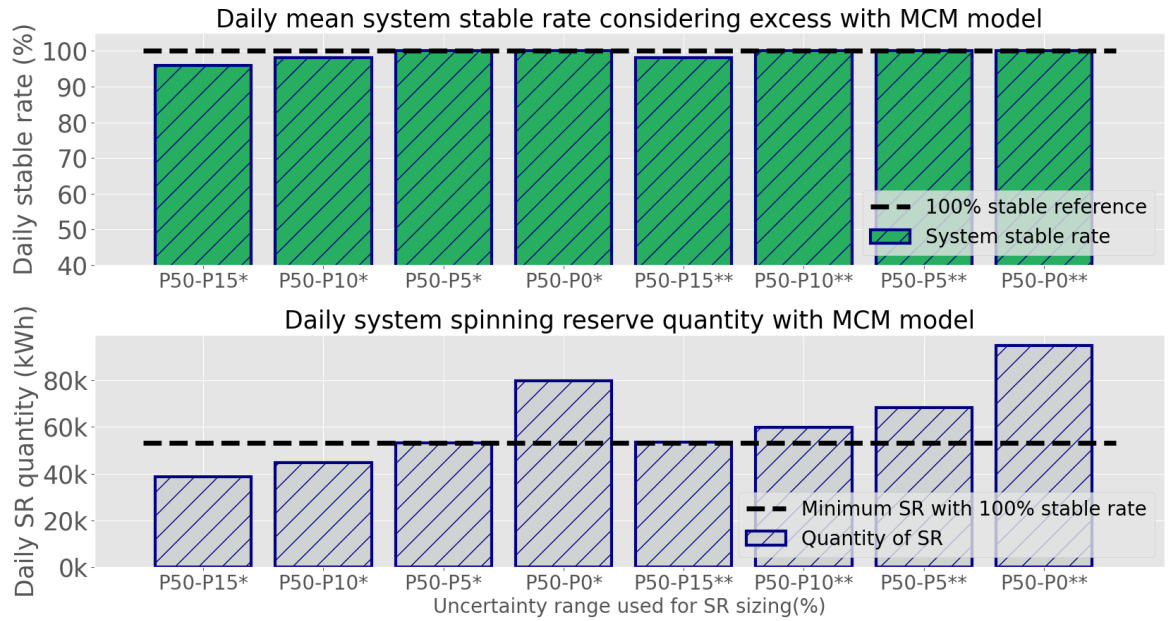


Figure 5.18 System stability and its associated spinning reserve quantity using MCM forecast with different uncertainties at variable day situation (* means with extra buffer of 5% PV capacity for SR sizing, ** for 10%)

To conclude, the system costs get higher with a higher spinning reserve, but the system is also more stable (with lower probability of a system imbalance due to badly sized spinning reserve). Low uncertainty range for SR sizing could reduce the system cost, but if it is too low and insufficient to cover all the variabilities, system cost will be more important due to the load shedding. But when the stable rate arrives at 100%, continuing to add SR just leads to a higher system cost for nothing. A suitable SR sizing could be achieved with an energy demand forecast and a relatively accurate forecast with a longer forecast horizon, to ensure the grid stability.

5.5 Influence of forecast horizon

5.5.1 Simulation setups for forecast horizon

In this section we assess the influence of forecast horizon in the planning and operation of an off-grid storage-less HES. Three forecast horizons of 1-hour, 2-hour, and 3-hour are tested for all the forecast methods, and two longer forecast horizons of 6 and 12-hour are tested additionally for Perfect forecast, which are highlighted in deeper orange color.

The setup of this test is shown in Table 5.5, the PV penetration rate is set at 50%, the spinning reserve sized as $P_{50} - P_0$ (the more conservative size considered in this work) and extra buffer of 10% PV capacity. The update time is set at 30-min to keep the same condition as the previous scenarios, and finally, both clear and variable sky situation are tested.

PV rate in capacity*	Forecast method	Uncertainty range used for energy buffer sizing		Forecast Lead Time	Update Time	Type of day
0%	CH-Persistence Ensemble (CH-PeEn) forecast	P50-P25	With extra buffer = X% PV capacity	1-hour	30-min	Variable
25%		P50-P20		2-hour		
50%		P50-P15		3-hour	1-hour	Clear
75%	Markov chain based forecast	P50-P10		6-hour	2-hour	
100%		P50-P5		12-hour	3-hour	Year round data
125%	Perfect forecast	P50-P0				
150%						
175%						
200%						

Table 5.5 Simulation setup selected to assess the influence of different forecast horizons

5.5.2 Results

As shown in Figure 5.19, in general, longer forecast horizons lead to lower system cost. While a 6-hour horizon leads to an additional improvement of 10.5%, further increasing the horizon to 12-hour did not further improve the HES system operation. The case of MCM model at clear day is due to the design of methodology as explained in section 5.4.2, which generates an optimal result for the forecast period, but not the whole period. Hence, optimizer turn off one generator to avoid energy excess, and then turn it on again to meet the energy demand, which results in a higher system cost.

As shown in Figure 5.19, with a Perfect forecast, a long forecast horizon is beneficial, since there is no uncertainty, using longer temporal information in the dispatch decrease the fuel consumption. The reason is that a longer forecast horizon could allow the optimizer to better organize the Genset dispatching, for example, a sufficiently large forecast horizon, the optimizer keeps the Genset always on to avoid the startup and shutdown cost. However, as [91] mentions, increasing forecast horizon is

not always positive, the benefit from longer forecast horizon could be decreased due to the reduce of forecast skill, and hence keeping a Genset on could be unnecessary, and results in a higher system cost.

Using Perfect forecast leads to a system cost reduction from 5.8% to 10.5%. For other forecast methods, the gain of using a longer forecast horizon is smaller, since their uncertainty is larger and their cost from SR, fuel consumption and PV curtailment are also higher. The difference between the perfect forecast and other two forecasts shows the effect of the forecast uncertainty which increases with lead time for MCM and CH-PeEN.

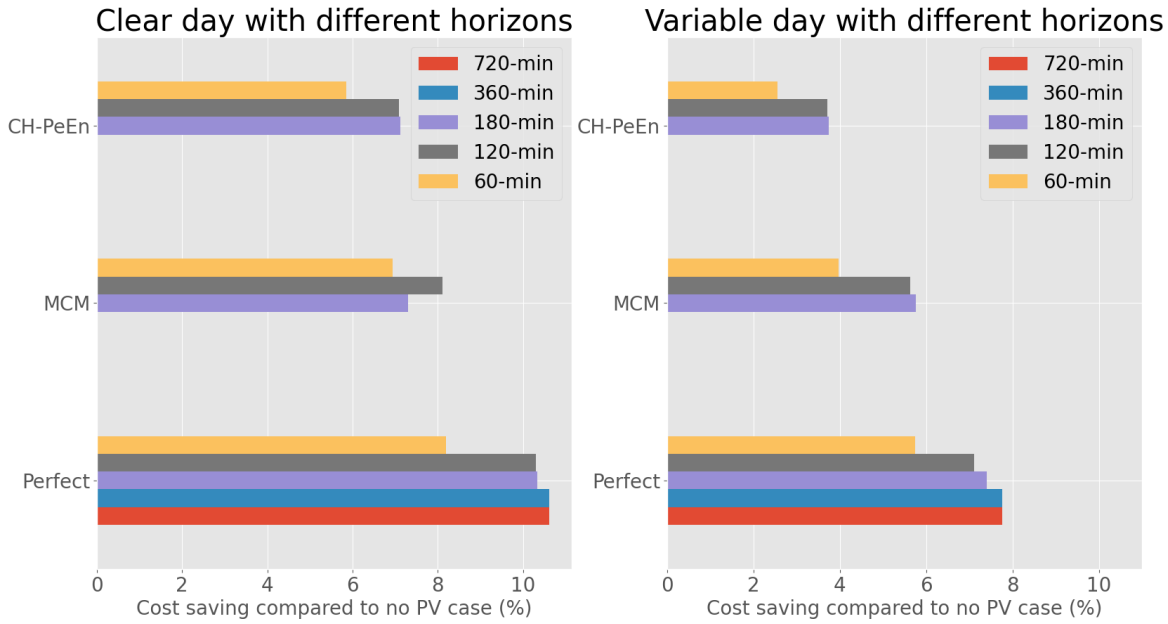


Figure 5.19 Economic cost saving compared to no PV case by using different forecasts with different forecast horizons

5.6 Influence of genset update time

5.6.1 Simulation setups for genset update time

For evaluating the influence of update time on system performances, the main configuration remains the same as section 5.5.1. As shown in Table 5.6, the forecast horizon is set at 180-minute so that there is a broader range of update times to test. Thus, Genset update time is set from 30 to 180-minute. In this test, all the different sky states are tested.

PV rate in capacity*	Forecast method	Uncertainty range used for energy buffer sizing		Forecast Lead Time	Update Time	Type of day
0%	CH-Persistence Ensemble (CH-PeEn) forecast	P50-P25	With extra buffer = X% PV capacity	1-hour	30-min	Variable
25%		P50-P20		2-hour		
50%		P50-P15		3-hour	1-hour	Clear
75%	Markov chain based forecast	P50-P10		6-hour		
100%		P50-P5			2-hour	
125%						
150%	Perfect forecast	P50-P0		12-hour	3-hour	Year round data
175%						
200%						

Table 5.6 Simulation setup selected to assess the influence of different Genset Update times

5.6.2 Results

For a given forecast horizon, the daily HES cost saving compared to no PV case with different genset update times is shown in Figure 5.20.

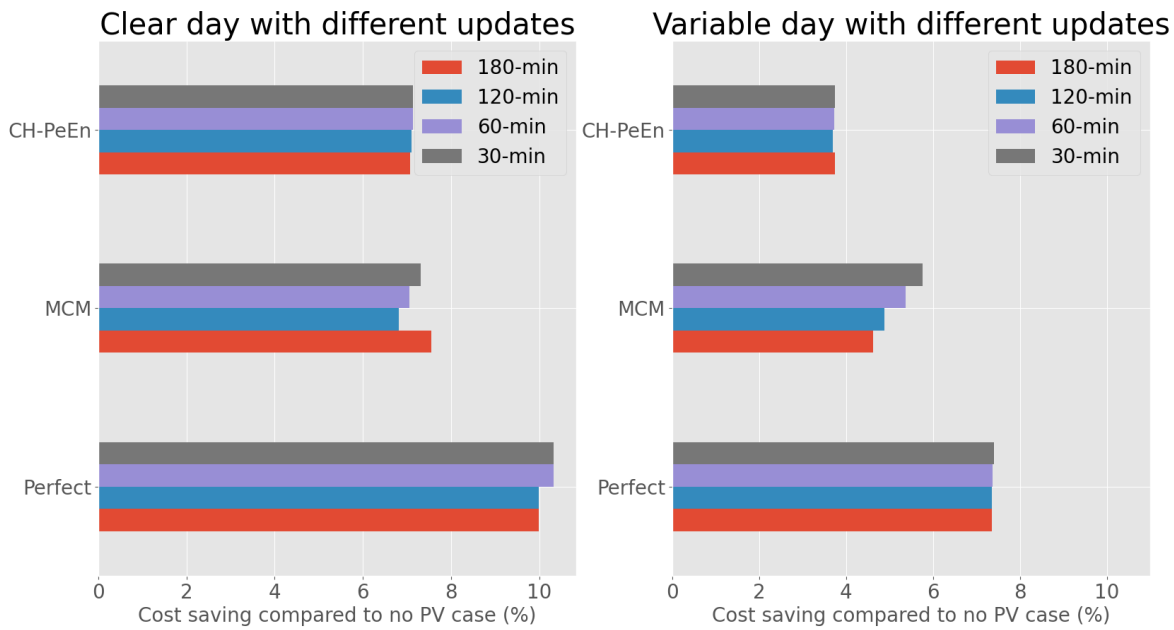


Figure 5.20 Economic cost saving compared to no PV case by using different forecasts with different update times

As shown, when using CH-PeEn forecast, the system cost does not change that much with different Genset update times. Since CH-PeEN model is similar to a climatological forecast, its forecast information including uncertainty range does not change much for a same forecast horizon. Therefore, for a same forecast horizon, the economic gain is similar no matter which genset update time is chosen.

When using MCM model forecast, a relatively short update time reduces the final economic gain, which means increase the final system cost. In fact, a shorter update time means a higher probability in changing the state of the diesel generators. In our simulation work, each startup and stop of generators has a cost. Hence, a more often genset update may increase the system cost. Notice that the reduction of the genset lifetime by using short update time is not considered, since the generator installation CAPEX is not considered in our simulation. If the genset runtime is translated in cost and considered in the simulation, the system cost with longer genset update time would be higher since the Genset stay on longer.

In terms of Perfect forecast, no matter in which update time, since the timeseries is already known and the information is always the same, the key factor of system cost is the forecast horizon. Hence, for a given forecast horizon, the system cost remains as the same with different genset update frequencies.

In conclusion, a relatively shorter genset update time frequency provides generally a higher economic saving whatever the forecast method, since the longer update time could avoid too much start and stop process by keeping the genset on with acceptable energy excess cost. The exceptional case of MCM model at clear day with 180-min genset update time, it occurs due to the similar reason of methodology design as explained in section 5.5.2, the optimizer generates a local optimal result but not a global one. When the Genset update time becomes too short, the system cost increases due to

the already mentioned shutdown and startup of diesel generators throughout the day. This update time has a more effective influence in system cost for better performing forecast methods than the forecast methods that provide same information all the time.

5.7 Validation of simulation setup parameters

5.7.1 Simulation setups with annual data

After exploring the influence of different parameters, we select the most advantageous parameters from our previous simulation analysis to re-explore the added value of using solar forecast. The PV penetration rate, the forecast methods and the annual data used would be the same as section 5.2. The uncertainty range $P_{50} - P_0$, with an additional buffer of 10% installed PV capacity, is still chosen since it is the most secure option to avoid load shedding situation. However, according to section 5.5.1, we know that a relatively long forecast could significantly improve the system performance, especially for Perfect forecast. Hence, the Perfect forecast in this case would be tested both in 3-hour and 12-hour forecast horizon, and the other two methods would use their maximum forecast horizon at 3-hour. Then, I choose the Genset update time at 30-min since section 5.6 tells that a shorter Genset update helps in improving the system performance in our simulation setup. The final simulation setup of this scenario is chosen as shown in Table 5.7.

PV rate in capacity*	Forecast method	Uncertainty range used for energy buffer sizing		Forecast Lead Time	Update Time	Type of day
0%	CH-Persistence Ensemble (CH-PeEn) forecast	P50-P25	With extra buffer = X% PV capacity	1-hour	30-min	Variable
25%		P50-P20		2-hour	1-hour	
50%	Markov chain based forecast	P50-P15		3-hour	2-hour	Clear
75%		P50-P10		6-hour	3-hour	
100%	Perfect forecast	P50-P5		12-hour		Year round data
125%		P50-P0				
150%						
175%						
200%						

Table 5.7 Simulation setup with selected advantageous parameters for quantifying the added value of using forecast

5.7.2 Results

The simulation result with above setting is shown in Figure 5.21, the annual no PV case is used again as a reference, which is composed of Genset fuel cost and the Genset maintenance cost.

Compared to the no PV case, the performance of using CH-PeEn is slightly better than section 5.2, till around 2.8% cost reduction. Even though the advantageous parameters like forecast horizon is improved, the economic gain from it is limited at around 1% like the result at section 5.5.1 shown, the advantage of longer forecast horizon is smaller for the basic forecast method.

When using MCM model, forecast, compared to the result of section 5.2, the result is even slightly lower (around 0.2%), at around 5.2% cost reduction. The reason for this result is due to the methodology design explained in section 5.5.2, the result generated by the optimizer is locally optimal (only for that optimization period) but not globally.

When using Perfect forecast, with 3-hour and 12-hour forecast horizon, all of their performance is better than the case with 1-hour horizon at section 5.2, for around 8.5% and 11.1% cost reduction, respectively. This result is also coherent to the result analysis from section 5.5.2, where a longer forecast horizon is useful and positive for high quality accurate forecast.

In terms of the range of the gain of using solar forecast, for 50% PV penetration rate (which corresponds to around 11.9% PV energy share), the system cost is reduced between 2.8% and 11.1% depending on the considered forecast model and compared to a no PV case. From the figure, we can also notice that a better performing forecast method could indeed reduce the fossil fuel consumption, directly and indirectly. The former one is by smartly dispatch the generating units, and the latter one is to reduce the quantity of SR needed through a sharper uncertainty range.

Briefly conclude, as explained in section 5.2, the perfect forecast has a highest PV energy share and economic saving since there is neither PV curtailment nor SR sizing. A longer forecast horizon can even improve its performance. The MCM forecast has a lower economic saving due to the fuel cost to covering its uncertainty and the PV curtailment cost. The CH-PeEn model has a lowest economic saving due to its conservative design, which has a bigger uncertainty, and results in a higher system cost.

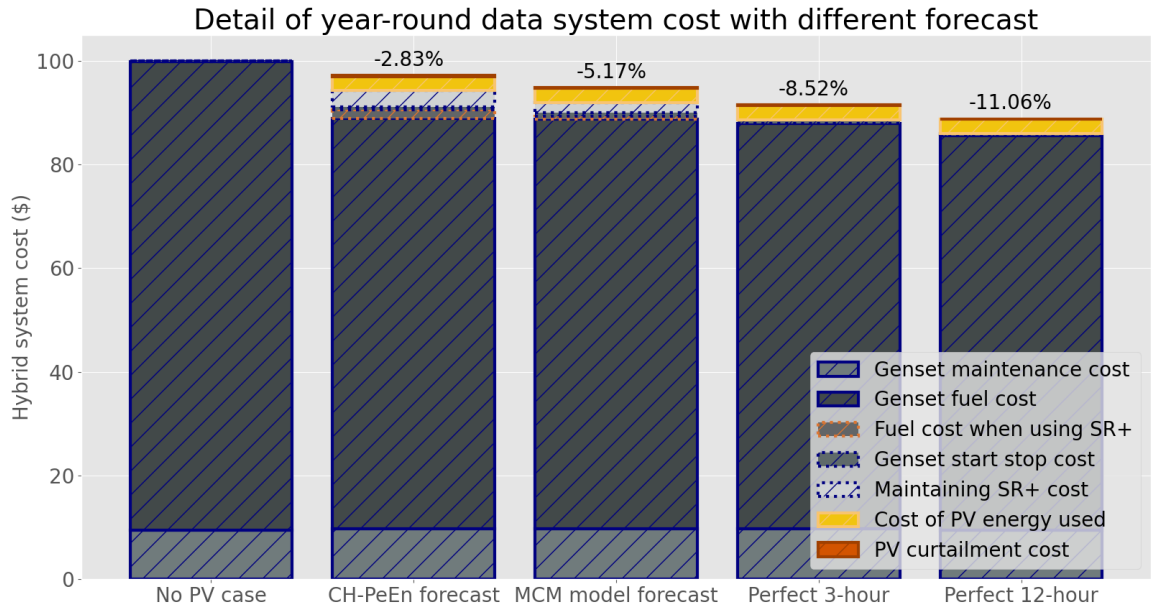


Figure 5.21 Economic saving of annual situation compared to no PV case by using selected advantageous parameters

Figure 5.22 shows the detail of the cost variation compared to the no PV case. As explained before, the top above zero is the sum of additional costs, the bottom is the sum of cost savings. In this figure, we can see that the SR of CH-PeEn anc MCM model is redundantly sized, since their cost of maintaining SR is much higher than the actual use cost. We can also notice that a more accurate forecast can not only help in reducing the fuel consumption, but also improving the PV energy share rate, which is coherent to the conclusion of section 5.2.2.

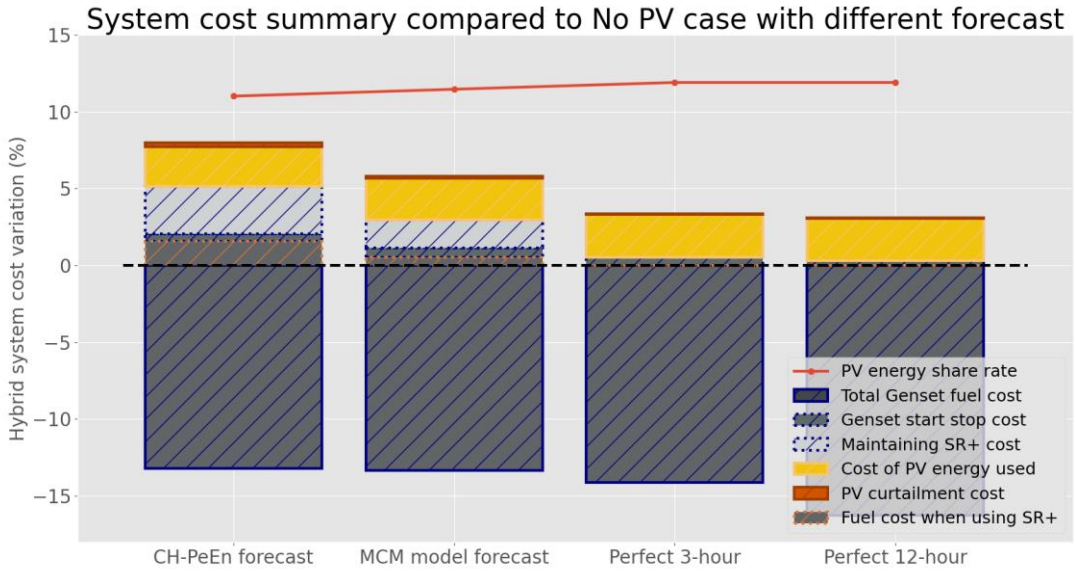


Figure 5.22 Cost difference of annual data situation compared to no PV case by using selected advantageous parameters

Another interesting point is the amortization time for this kind of energy system. For a HES like our case study, with 50 MW energy demand, 24MW_p PV installation. For a raw calculation, 24 MW_p system needs around 24 million dollars, if the economic gain of each day is 11%, with an around 170k dollar daily system cost, the amortization of the PV system could be achieved in around 3.7 years. For a real-world application, even though the economic gain of each day would be less due to other unexpected events and human factors, the amortization time will still be in a similar magnitude.

5.8 Conclusion

As shown in the section 5.2 and 5.7, the gain of using forecast method depends on several elements. The most important one is the forecast method performance. However, the sky state situation also plays an important role in system economic cost, for example, the CH-PeEn forecast performs obviously relatively well compared to other advanced forecasts in clear day, but in a variable day situation, the economic gain is less than the half of clear day case.

To notice that the above results are based on a redundantly sized spinning reserve, which is used to mitigate the system unbalance. Yet, high security also signifies a higher system cost. As shown in the analysis of section 5.4, different quantities of energy buffer for an isolated storage-less HESS could lead to quite different results. An insufficient energy buffer could not cover the energy shortage imbalance, but a too-redundant energy buffer asks a high genset capacity and has a higher system cost. Hence, the compromise between uncertainty range used for spinning reserve sizing and the system cost should be studied carefully in a real-world application, according to the characteristic of forecast method used, the genset capacity, and the tolerance of system imbalance. This should be particularly considered for high PV penetration rate cases, besides the forecast method performance.

The PV penetration rate is another important element plays an important role, for a low PV share part, even though the forecast is perfect, the gain of using forecast is limited. Relatively, if the forecast has a lower performance, the influence on the whole system gain would be low as well. Indirectly speaking, rather than spending costs in finding a high-quality solar forecast method for low PV penetration rate, adding more PV installation before the saturation rate is a better option.

The final but the most important element is a suitable power management system. As shown in section 5.4, a case with high PV penetration rate could have a system cost higher than no PV case, which means that investing a lot in building a huge PV system but without good management strategy, unless we don't consider the cost in building PV system, otherwise we'd better not to install the PV system, since it is not worthy both in economic and in environmental aspect.

Hence, the general gain of using forecast method should refer to a reference level. In a nutshell, the upper reference should be the case using Perfect forecast in a same situation, the lower bound could be the case of using basic model like CH-PeEN forecast or even basic persistence forecast. So far, many real-world applications of solar forecast are still directed by the classic statistical metric like RMSE, MAE, etc. Yet, as mentioned before, a same amount of over-estimated and under-estimated forecasts provides same value of RMSE or MAE, but the final system result is totally different. To have a further understanding, the following chapter studies the link between the classic metric and the system performance.

Chapter 6

6. Statistical and economic performance

Contents

SUMMARY OF CHAPTER

This chapter explores the link between the statistical performance of solar forecast method and the practical economic performance of Hybrid Energy System (HES).

At the beginning of this chapter, the general statistical performance of the chosen forecast method is shown. In this work, both deterministic and probabilistic metrics are considered since both of them are used in our simulation work. The deterministic forecast is used to generate the genset order.

The probabilistic uncertainty range is used to size the spinning reserve as an optimization constraint in our simulation. Hence, the relevancy of classic statistical metrics like RMSE and MAE for deterministic metrics, PINAW and CRPS for probabilistic forecast metrics can be explored and discussed.

Then a year-around HES simulation has been achieved, and the associated indicators like daily overall system costs are obtained. The statistical-based metrics of the solar forecasting part and HES economic performance results are put together to explore and discuss the relationships between them.

According to these statistical and economic results, the classic metrics could tell the general guideline of practical economic performance. Nevertheless, the proportional difference in statistical results could not be reflected in the economic performance.

6.1 Simulation setups with year-round data

As seen in Chapter 5, even though the classic statistical evaluation metrics can reflect the solar and system variability, it does not correctly describe the practical HES economic performance. Thus, we introduced a cost-based approach to correctly evaluate the HES economic performance in chapter 5. This chapter aims to question the equivalence between solar variability, solar forecast uncertainty, and the final practical economic performance using our cost-based approach.

To be able to explore the link between different results, we used a yearly time series to provide a more complete evaluation. We set the PV penetration rate at 50% to correspond to the actual PV penetration of our case study. For the sake of simplicity, we have chosen the MCM model forecast as the main forecast method, framed by the CH-PeEn and the perfect prognostic forecast methods. We performed RMSE, MAE deterministic analysis and PINAW, CRPS probabilistic analysis over the whole uncertainty range of the forecast, but for HES economic simulation we fixed the spinning reserve with the $P_0 - P_{50}$ plus an extra buffer of 10% PV installation capacity, which is sized to avoid any load-shedding action. For this work, we also ran the dispatch using perfect forecast with a 12-hour forecast horizon to get the best optimization, *i.e.* the genset scheduling with the lowest cost. As seen in chapter 5, the forecast horizon has some impacts on the final performance. Finally, we fixed the update time to 30 min and the forecast lead time for MCM forecast to 3-hour. Table 6.1 summarizes the parameters used in this simulation, and are highlighted in orange. The parameters highlighted in deep orange are only used for perfect prognosis forecast.

PV rate in capacity*	Forecast method	Uncertainty range used for energy buffer sizing		Forecast Lead Time	Update Time	Type of day
0%	CH-Persistence Ensemble (CH-PeEn) forecast	P50-P25	With extra buffer = X% PV capacity	1-hour	30-min	Variable
25%		P50-P20		2-hour		
50%		P50-P15		3-hour	1-hour	Clear
75%	Markov chain based forecast	P50-P10		6-hour	2-hour	
100%		P50-P5				
125%						
150%	Perfect forecast				3-hour	Year round data
175%		P50-P0		12-hour		
200%						

Table 6.1 Simulation setup selected for 1-year long dataset complete evaluation

6.2 Statistical performances

The solar forecasts used in our simulation are probabilistic forecasts. For this reason, we evaluate them with probabilistic metrics such as PINAW and CRPS [92]. We used the P50 of the forecast for the optimization process of genset order generation. The other quantile levels of the probabilistic forecast are used for the sizing of the spinning reserve as a constraint for the optimization. Hence, we used both deterministic and probabilistic forecast metrics to completely evaluate the forecast performance. In this work, the classic deterministic metrics like RMSE and MAE [30] are used to evaluate the deterministic forecast performance, the classic probabilistic forecast metrics like CRPS and PINAW are used to evaluate the probabilistic part of the forecast method.

The forecast used in our work has a 1-hour forecast horizon and 15-min temporal resolution, in a classic evaluation way, the performance of different forecast horizons could be evaluated separately, such as the performance of 15-min horizon, 30-min horizon, etc. But since the forecast horizon used is 1-hour, the information of 15, 30, 45, and 60-min horizon are used together for each optimization process. The statistical evaluation in this chapter hence combines all the forecast horizons from 15-min to 60-min head.

6.2.1 Deterministic metrics

To give comparable results, all the daily metrics are provided in normalized value, where the normalization reference is the average of daily PV production measurement from the whole available year. On purpose, this normalization makes the metrics relative in a same order, thus lower nRMSE and lower nMAE is better. The Table 6.2 shows the detail of normalized RMSE and normalized MAE that span all the data points. As shown, the nRMSE and nMAR of CH-PeEn forecast are about twice higher as MCM forecast case.

Normalized Root Mean Square Root (nRMSE) and Normalized Root Mean Square Root (nMAE)

	CH-PeEn[9] forecast (%)	MCM[10] model forecast (%)	Perfect forecast (%)
Mean of nRMSE (%)	**47.7**	**28.9**	**0**
Best of nRMSE (%)	**4.9**	**6.8**	**0**
Mean of nMAE (%)	**38.3**	**20.7**	**0**
Best of nMAE (%)	**3.7**	**4.8**	**0**

Table 6.2 Summary of statistical performance of deterministic metrics with different forecast, all forecast horizons combined from 15-min to 60-min head.

Figure 6.1 present the general distribution of these daily error metrics. The general situation corresponds to the annual average. We can observe that the CH-PeEn forecast has much higher values of daily error metrics, compared to MCM forecast, which means that – as expected - the P_{50} of the MCM model forecast outperform the one of CH-PeEn forecast.

The quantile level of CH-PeEn is defined according to the whole historical data to cover all the cases, which becomes conservative for a daily short-term forecast. That makes the difference particularly large between forecast and measurement. In terms of nRMSE comparison, nRMSE of MCM model is lower than the one of CH-PeEn in more than 81% cases. The situation in nMAE is similar, which is more than 82%.

[9] CH-PeEn: Complete History – Persistence Ensemble forecast
[10] MCM model: Markov Chain Mixture Model forecast

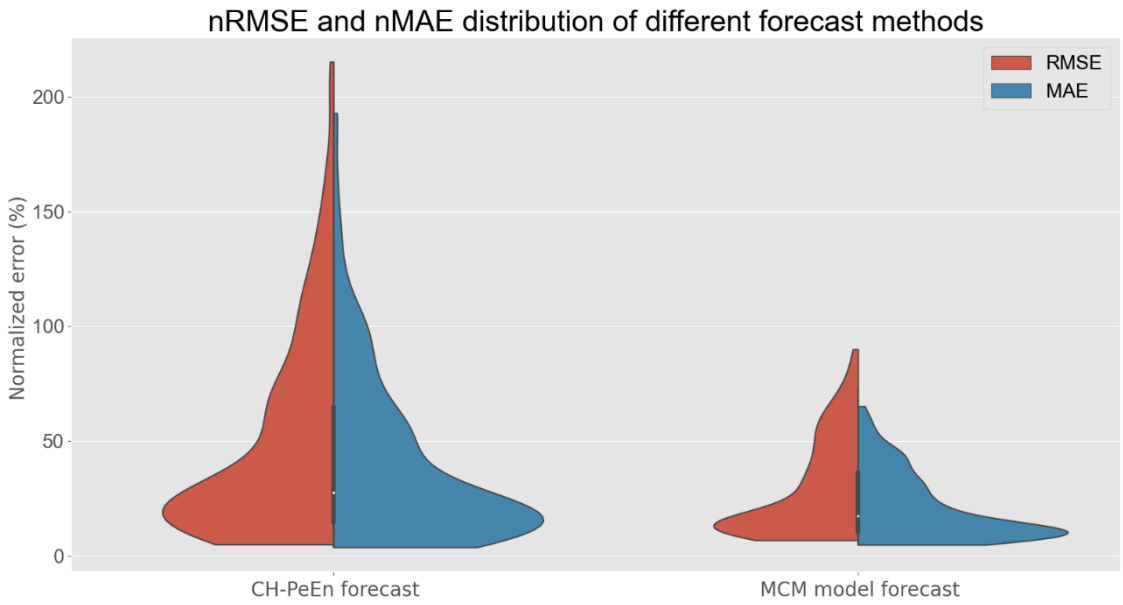


Figure 6.1 Histogram of nRMSE and nMAE of different forecast methods

6.2.2 Probabilistic metrics

As mentioned before, we have chosen PINAW with PICP and CRPS to show the statistical performance of probabilistic forecast methods. These metrics are calculated in their classic way [92], [80], which means the confidence levels for PICP and PINAW are calculated with the uncertainty around P50 level of probabilistic forecast. for example, 50% confidence level means the uncertainty range of $P_{25} - P_{75}$.

Standard Prediction Interval Coverage percentage (PICP) and Prediction Interval Normalize Average Width (PINAW)

Confidence level (%)		10	20	30	40	50	60	70	80	90	100
Mean PICP (%)	CH-PeEn	11.1	21.9	32.6	43.5	53.9	63.9	73.9	83.0	91.7	99.3
	MCM	10.6	20.9	31.7	42.1	52.2	62.4	72.5	83.1	92.1	99.3
Mean PINAW (%)	CH-PeEn	13.7	27.7	41.9	57.2	73.7	91.6	112.8	141.9	190.3	271.6
	MCM	6.6	13.3	20.3	27.9	36.6	46.9	59.3	75.8	102.9	184.9

Table 6.3 Summary of PICP and PINAW for the CH-PeEn and MCM probabilistic forecast method.

The average value of PICP and PINAW at different confidence levels is shown in Figure 6.2, where the CH-PeEn forecast has a quite similar PICP compared to MCM forecast. Yet, both of them have a distance to the perfect case, especially towards a high confidence level. The general PINAW is shown in Figure 6.2, where the PINAW of MCM forecast is significantly smaller than CH-PeEn forecast. CH-PeEn has a large PINAW due to its conservative characteristic, with an intrinsic lack of resolution. Notice the fact that a large PINAW means a large uncertainty range, which may imply a large spinning reserve in HES simulations: hence, this could mitigate more system variability, but also could result in higher system costs.

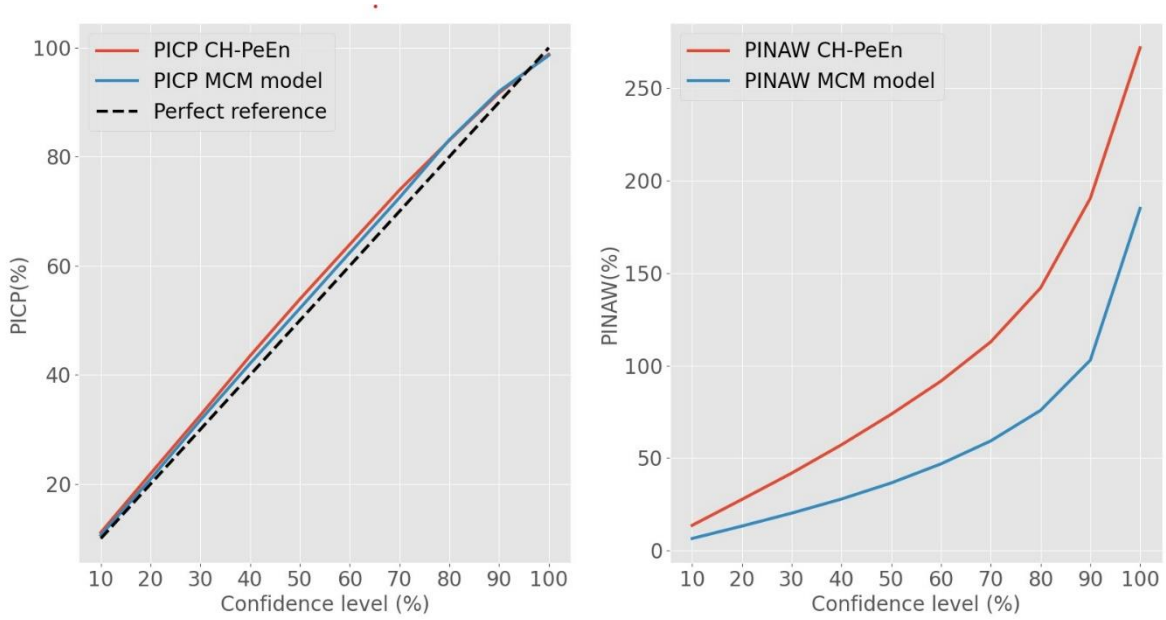


Figure 6.2 Average PICP and PINAW of annual simulation for CH-PeEn and MCM forecast methods (30% means that the spinning reserve is sized with P35-P65 and 10% PV capacity extra buffer)

Continuous Ranked Probability Score (CRPS)

	Statistical performance
Annual Mean nCRPS of CH-PeEn (%)	**31.8**
Annual Mean nCRPS of MCM (%)	**17.8**

Table 6.4 Summary of nCRPS with different forecast methods

In terms of CRPS evaluation, the average nCPRS of MCM model is around twice lower as the CH-PeEn case. The daily average nCRPS of annual data is shown in Figure 6.3, where the error spread of MCM forecast is mainly centered at under 20%, but the error spread of CH-PeEn is much larger, with a maximum value higher than 110%. It means that the MCM model forecast outperforms the CH-PeEn forecast in general, with a sharper uncertainty and higher reliability.

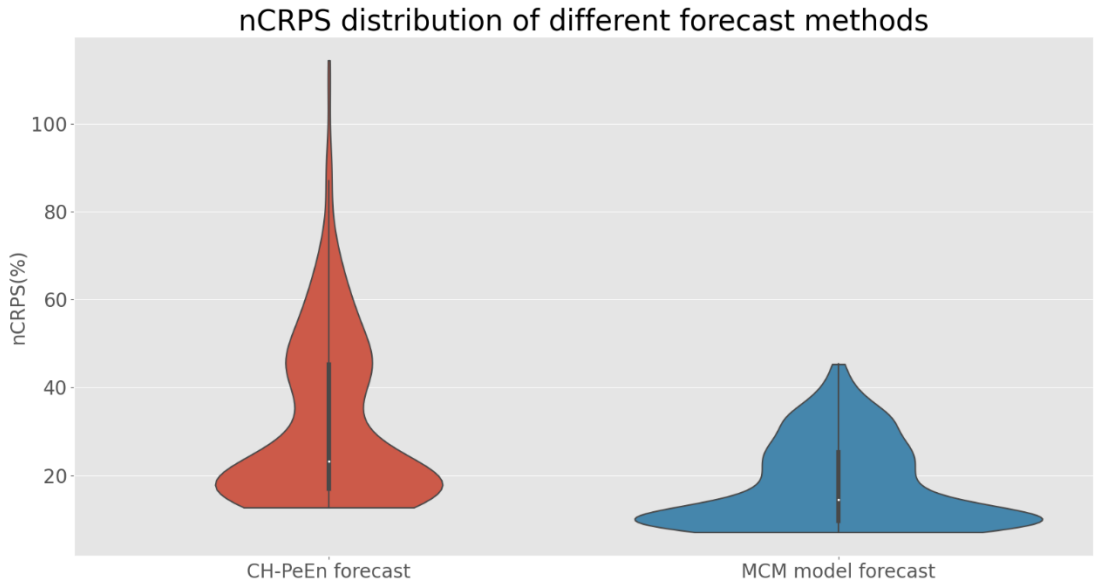


Figure 6.3 Average nCRPS of annual simulation with different forecast methods

6.2.3 Conclusion

After comparing the different statistical metrics of two forecast methods used in our simulation, we can see that, except for the PICP, the general performance pattern of these metrics is similar: as expected, the MCM forecast provides much better performance than the CH-PeEn, both on deterministic and probabilistic metrics. In terms of PICP, these two methods have a similar level of performance, but they could not be used as a single indicator to guide the application since their PINAW are significantly different.

On average, owing to these standard statistical metrics, the ratio of outperformance of the MCM method compared to CH-PeEN is approximately two.

This clear outperformance may lead users to consider the potential practical economic performance difference to be more or less twice as well. In the following section, we explore the simulation results based on the dataset with two forecast methods to see the practical economic performance.

6.3 Practical economic performance

Figure 6.4 shows the two global histograms of the daily system costs from the year-around HES simulation, using of the three forecast methods. We can observe that, as expected, the perfect prognosis (12-hour ahead) cases have the lowest system cost, the MCM model performs a little bit better than the CH-PeEn case, which has the highest system cost distribution. To notice that the MCM model and CH-PeEn forecast have 1-hour forecast horizon, which is less advantageous compared to 12-hour. The result of perfect prognosis is not that far ahead than the other two methods since the PV energy share rate is only around 9% of the total demand.

	CH-PeEn forecast ($)	MCM model forecast ($)	Perfect forecast ($)
Average daily system cost	**175k**	**170k**	**161k**

Table 6.5 Average daily system cost with different forecast methods

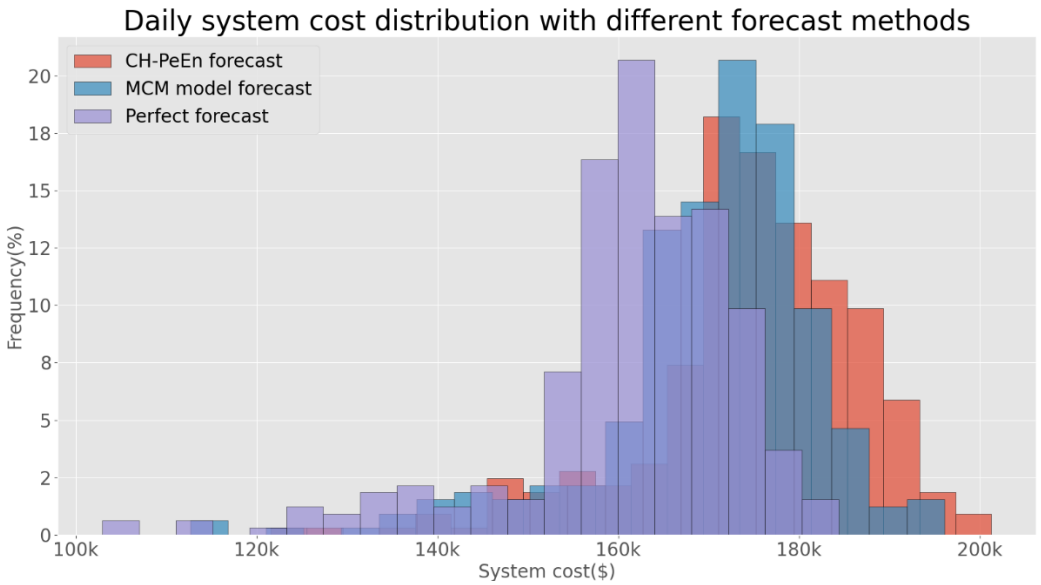


Figure 6.4 Distribution of daily system cost with perfect forecast

With the setup explained in section 6.1, the spinning reserve is over-sized with uncertainty range $P_0 - P_{50}$ and extra buffer of 10% PV installation capacity. In that case, there is no system imbalance

with a significant increase the general system cost. With a positive spinning reserve sized smaller, the result would have had much more system imbalance and load shedding situations.

Considering the case without PV (only genset) as a reference, we computed a general daily cost ratio, expressed in percentage. The average value is shown in Table 6.6. Where the perfect case could provide maximum 11.7% economic gain, notice that this perfect case is obtained with 12-hour horizons, which could significantly improve the performance and provide a real upper bound of economic gain.

	CH-PeEn forecast	MCM model forecast	Perfect forecast
Average economic saving	**2.8%**	**5.6%**	**11.7%**

Table 6.6 Average daily system cost saving ratios with different forecast methods

Figure 6.5 presents the comparison of the observed distributions of the daily system cost saving ratios for the three forecast methods.

The perfect case is always positive, which means always better than the case without PV. This ratio can go up to 17.7 % with an interquartile range (IQR) between P_{25} and P_{75} of 9.2% to 14.5%.

We can also observe that with an IQR between 3.2% and 8.1% the MCM forecast provides an overall intermediate performance between the CH-PeEN (IQR: 0% to 5.9%) and the perfect forecast.

Furthermore, using MCM or CH-PeEn forecasts implies to have some cases that use PV energy and these forecast methods are even worse than not using PV, due to the oversized positive spinning reserve. For the CH-PeEn methods, it corresponds to 24.7 % of the days with the year-around HES simulation and only 5.9% for the MCM methods.

Indeed, as explained in chapter 5, for an isolated storage-less HES, the system stability is mainly secured by the spinning reserve, using redundant SR to cover all the over-estimated forecasts would be costly, and the occurrence chance of extreme cases is not always high. For a highly stable constraint storage-less application, having some days with low or even negative cost saving ratios to ensure global security, while having positive ratios for the majority of days may be acceptable. The choice of the probabilistic forecast is of crucial importance when considering this ratio.

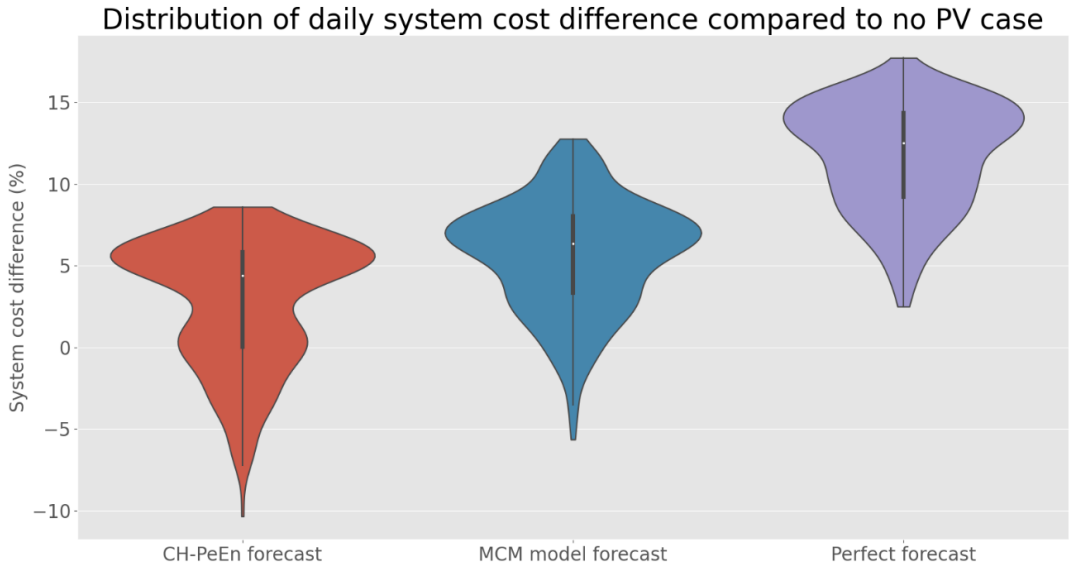


Figure 6.5 Distribution of daily system cost with different forecast methods

By simply comparing the cost saving by using MCM model and CH-PeEn, the difference in economic performance does not simply match the statistical performance. The purpose of the next section is to further explore potential relationships between statistical and practical economic performances.

6.4 Relationships between statistical and practical economic performance

In this section, for the sake of clarity, the reference year-around simulation case is no longer the one without PV. Indeed, to better contrast the difference of daily costs from the different forecasts, we have chosen in this section to use the case using perfect forecast with *12-hour horizons* and the case without PV energy as the reference and to define a skill cost as (57) :

$$Skill\ score_{Cost} = \frac{Cost_{no\ PV\ case} - Cost_{forecast}}{Cost_{no\ PV\ case} - Cost_{perfect\ prognosis,ref}} \quad (57)$$

The denominator corresponds to the biggest gain that we can have for a given PV penetration rate, and the numerator is the actual gain of each forecast method compared to no PV case. Hence, when this s*kill score* is equal to 1, it means that the forecast method used has same system performance as that obtained with the perfect prognosis. When this *skill score* has a negative value*, it means* the forecast method could have a negative effect and its cost is even higher than the case without PV.

6.4.1 Cost skill score exploration

To justify the usefulness of this skill score cost, we process a comparison between the sky state and the skill score of each forecast method, as shown in Figure 6.6.

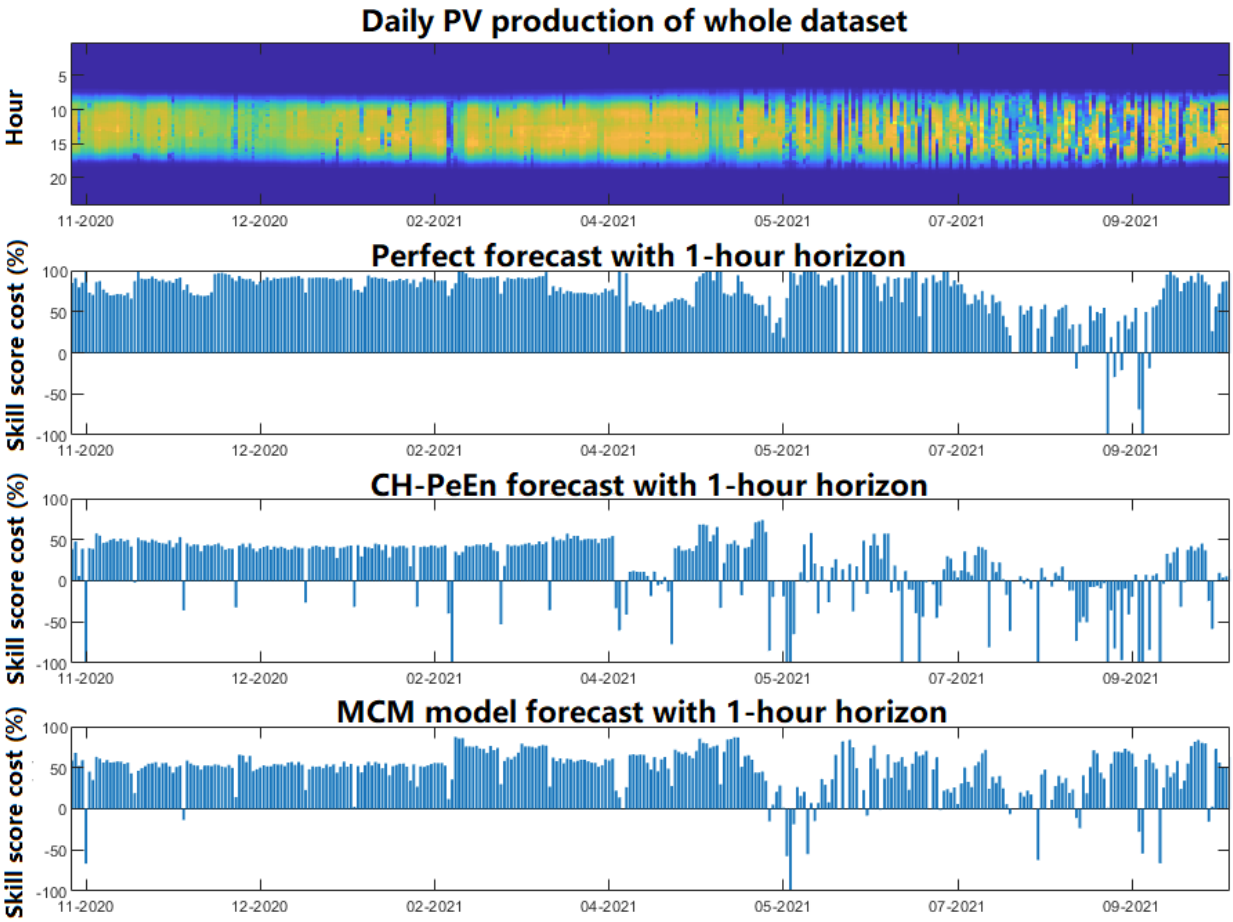


Figure 6.6 Link between sky state and daily cost skill score using different forecast methods

These results are obtained with 1-hour forecast horizon and 30-min genset update time with 50% PV penetration rate. In this figure, the figure on the top is the PV production of each day, the horizontal axis is the date and the vertical axis is the time of a day (from 0 to 24h). The three figures below have the same horizontal axis, and their vertical axis is the skill score cost over the daily energy demand to normalize the result.

As shown in this figure, the beginning of the dataset is mainly composed of clear sky states, and the cost skill score of each forecast method is mostly higher than 50%, which means having a good performance, and the cost skill score of each method also correspond to the statistical performance that we have shown in section 6.2. When we arrive at around May 2021 of the dataset, we observe that the PV production starts to become variable due to the rainy season of our case study. As a consequence, the cost skill score of each model decreases as well, and their performance is also coherent to the statistical performance analysis.

Besides, we also did a harrow diagram, with the horizontal axis which is the average PV production over the average PV production in clear sky conditions, and the vertical axis is the length of the day over the length of the day in clear sky condition as well. This figure provides us a way to separate the clear sky day and overcast day. By using the cost skill score of each day for the hue, we can obtain a result as shown in Figure 6.7 and Figure 6.8.

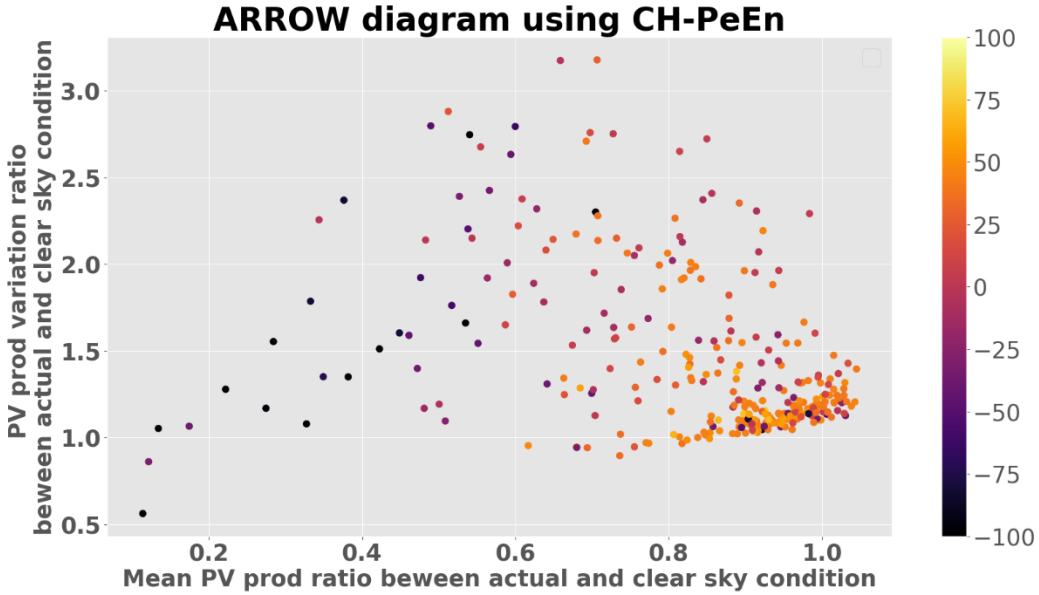


Figure 6.7 Harrow diagram with cost skill score using CH-PeEn forecast

In these figures, the clear sky day is at the right of the figure, since the ratio between PV production in actual and clear sky conditions is towards to 1. The left part of the figure corresponds to overcast situations. From this figure, we can obtain similar information as Figure 6.6, where the clear day has a higher cost skill score and the overcast day doesn't. And we can also see that the MCM model outperforms the CH-PeEn forecast, since MCM have relatively good performance both in clear and overcast days, but the CH-PeEn works well in clear days but not in overcast situations. From the results obtained, we can say that this cost skill score could relatively associate the system performance with forecast performance.

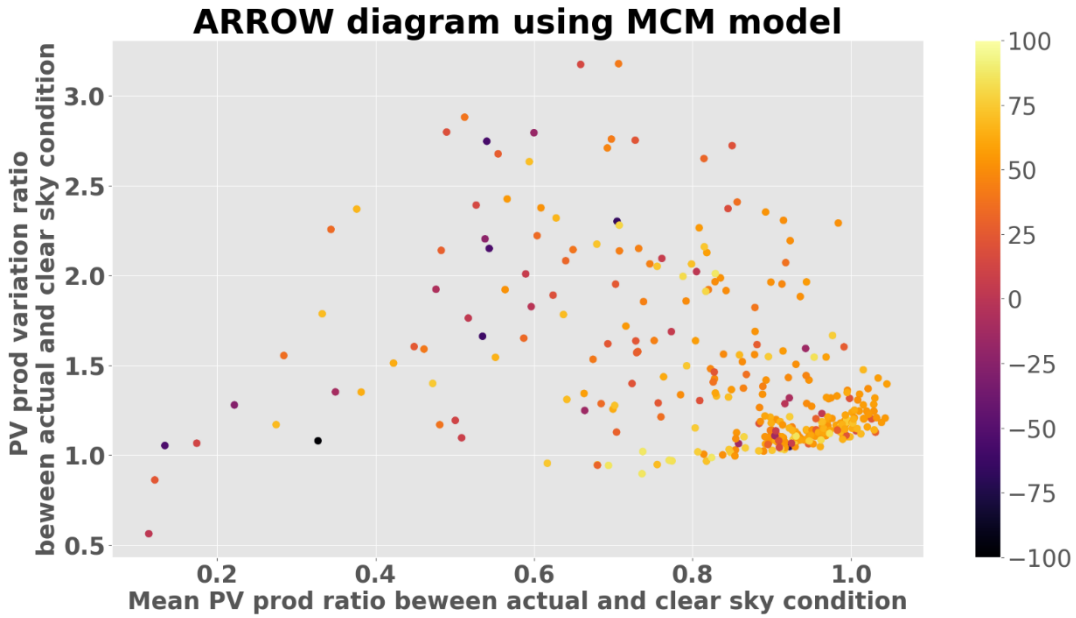


Figure 6.8 Harrow diagram with cost skill score using MCM model forecast

6.4.2 Deterministic metrics evaluation vs economic differences

Besides this skill score about the cost, four statistical metrics are used in this work to provide a complete overview.

In the plots from Figure 6.9 to Figure 6.14, the horizontal axis is the normalized error of different metrics, and the vertical axis is the cost skill score. Even though these metrics are evaluated allow a comparison with a same indicator, due to the characteristic of different indicators, this section is divided into two parts, including the deterministic metrics and probabilistic metrics, to give a relatively concrete conclusion.

nRMSE vs cost skill score

The link between nRMSE and system cost skill is shown in Figure 6.9Figure 6.9, as shown, three conclusions can be drawn:

- In Figure 6.9, we observe that the error of MCM model are mostly under 100%, but CH-PeEn has the error till 200%. Hence, the first one is the general error spread of CH-PeEn is larger and higher than the MCM model forecast, which means that the statistical performance of CH-PeEn model is worse than the MCM model,
- In the same figure, we also see that the cost skill score of MCM model are mostly closer to 1 compared to CH-PeEn, the following conclusion is hence the economic gain by using CH-PeEn forecast is generally lower than MCM forecast, but there are also some exceptional cases,
- The last and most important conclusion is obtained from Figure 6.10, that is for a same level of statistical error, the economic performance could be very different, and their economic difference is not proportional to the statistical performance.

For example, as shown in Figure 6.10, at 20 % nRMSE level, with the same nRMSE, their skill score cost is very different, most of the MCM cases perform better than the CH-PeEn, but sometimes the CH-PeEn cases have better performance, these situations signify a fact that statistical metrics could not always reflect the practical performance.

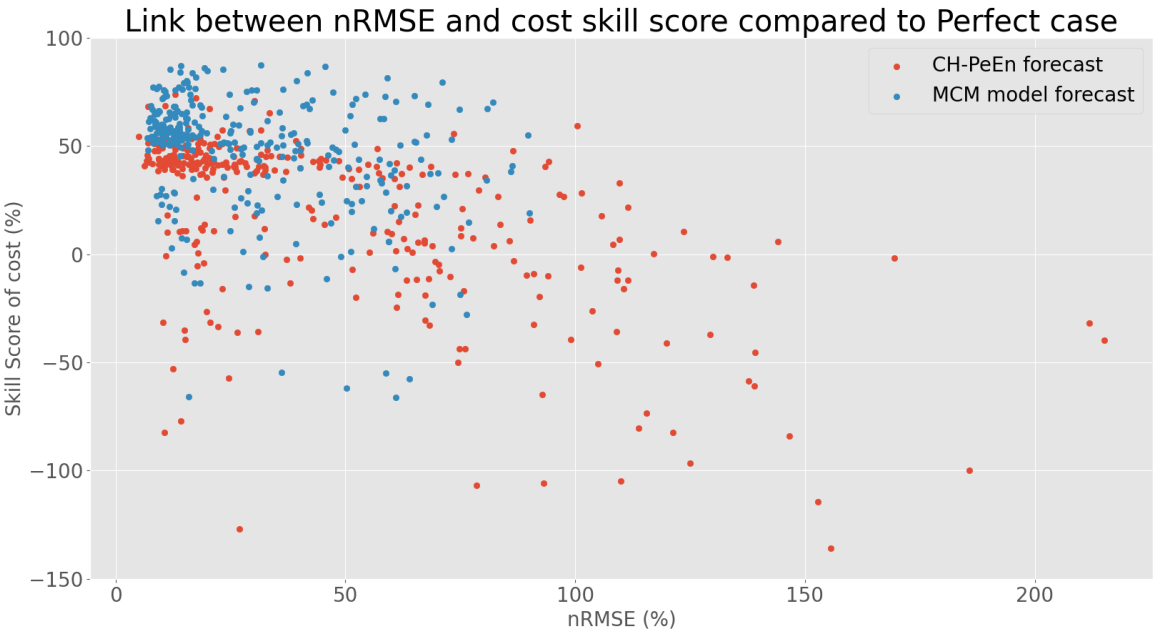


Figure 6.9 Link between nRMSE and daily cost skill score using different forecasts

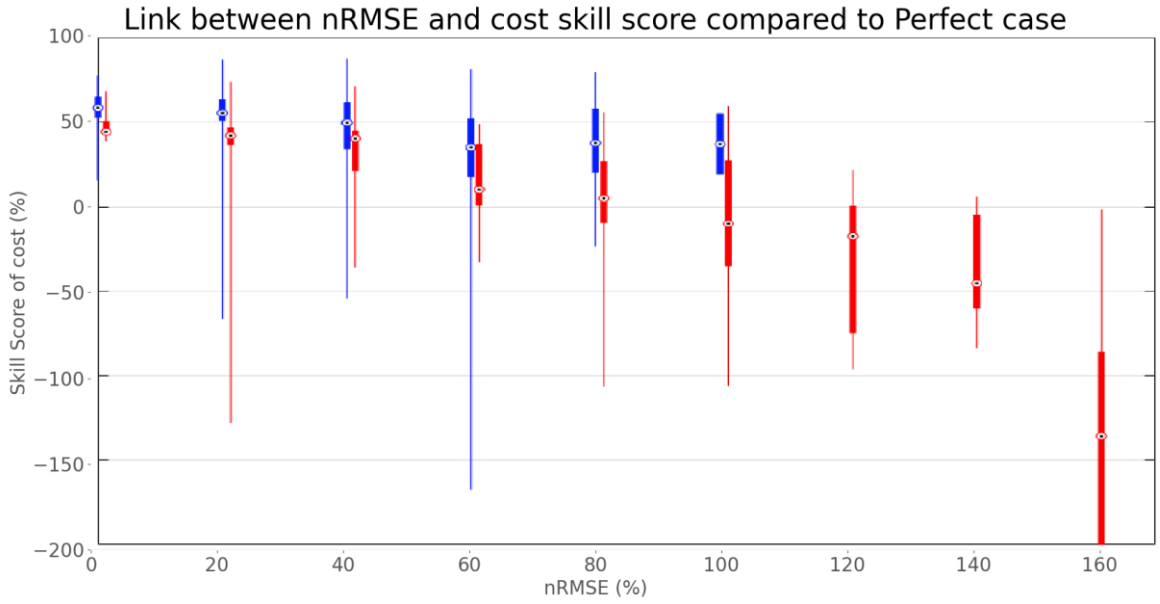


Figure 6.10 Link between nRMSE and daily cost skill score using different forecasts in boxplots

nMAE vs cost skill score

After looking into the nRMSE evaluation, we process the same evaluation of nMAE, whose general pattern is similar to the nRMSE one. The small difference between RMSE and MAE is hard to distinguish, hence the detail is not shown here.

6.4.3 Probabilistic metrics evaluation vs economic differences

PICP vs cost skill score

In terms of PICP evaluation, for 50% confidence level, unlike the annual PICP shown in the previous statistical performance evaluation, the performance of MCM model and CH-PeEn forecast are no longer so similar to each other as shown in section 6.2.

In fact, the result in this section is daily PICP, but the result shown before is annual PICP curve, which is calculated through a whole year of data and the variation could be largely mitigated.

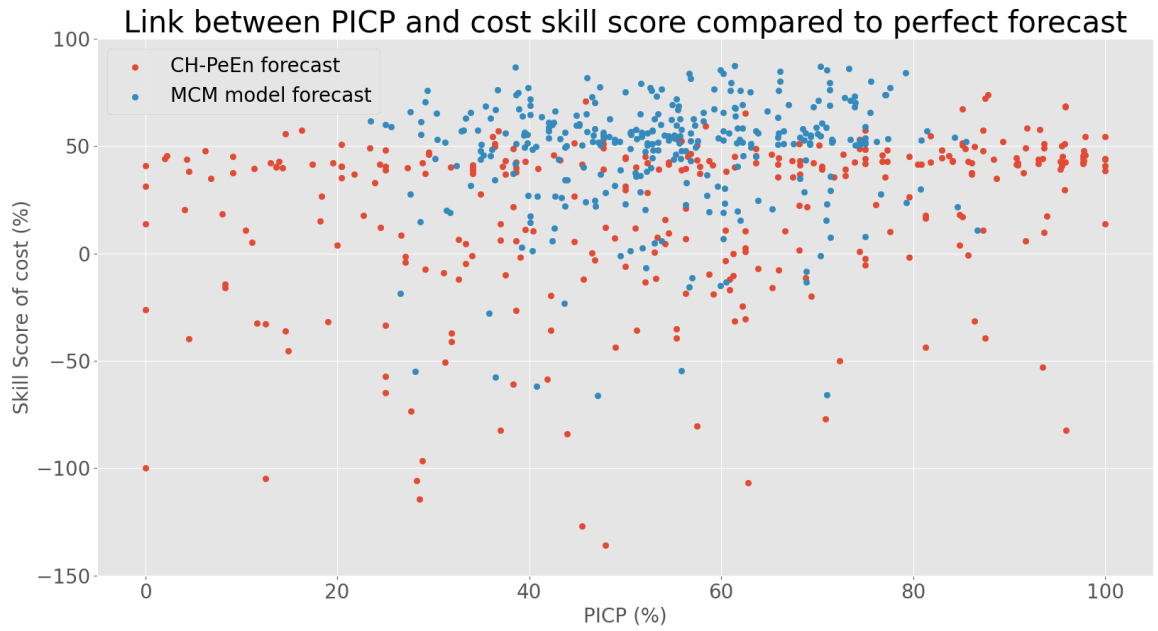


Figure 6.11 Link between PICP and daily cost skill score

From this figure, we can draw the following conclusions:

- As shown in Figure 6.11, the MCM model forecast cases are mainly centered around 50% PICP, but the CH-PeEn cases have a much larger spread, which means that the performance of MCM model is more coherent and better than CH-PeEn.
- For the same confidence level, MCM model provides a better economic result as well since the cost skill score of MCM model shown in Figure 6.11 are closer to 1.
- Finally, for a same level of PICP, the economic performance could be very different as the case of RMSE, which means that PICP is not able to correctly reflect the economic performance.

PINAW vs cost skill score

The PINAW evaluation with the cost skill score is shown in Figure 6.13, where the PINAW spread of CH-PeEn is a bit different compared to the evaluation of previous metrics. Unlike the MCM model which has its PINAW all under 100%, CH-PeEn has some cases with a much higher PINAW. This is also due to the characteristic of the CH-PeEn forecast used, as mentioned, the distribution is generated separately for each month. And the weather characteristic of our case study location has a rainy season and its variability is much more important than other seasons, the details of the annual daily production profile and the associated PINAW curve are shown in Figure 6.12. As the figure shows, since July of the year, the PV production has had much more variability and its associated PINAW grow both in small and big uncertainty ranges.

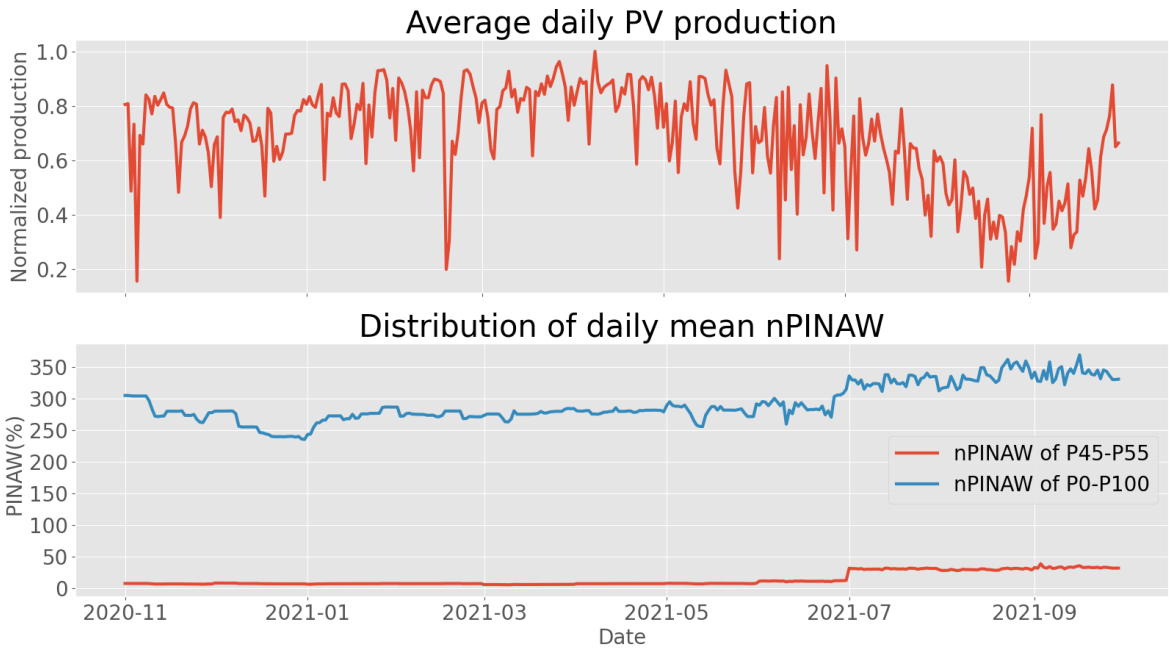


Figure 6.12 Explanation of CH-PeEn PINAW sudden growth

Besides this rainy season effect, the MCM model still performs better than CH-PeEn forecast, with a better economic gain and smaller general PINAW, which needs less cost for spinning reserve.

By combining the result from PICP and PINAW, as classic statistical metrics, they are able to distinguish the general performance of two forecast methods, but in terms of telling the concrete performance difference, they are far from enough to describe the practical economic performance.

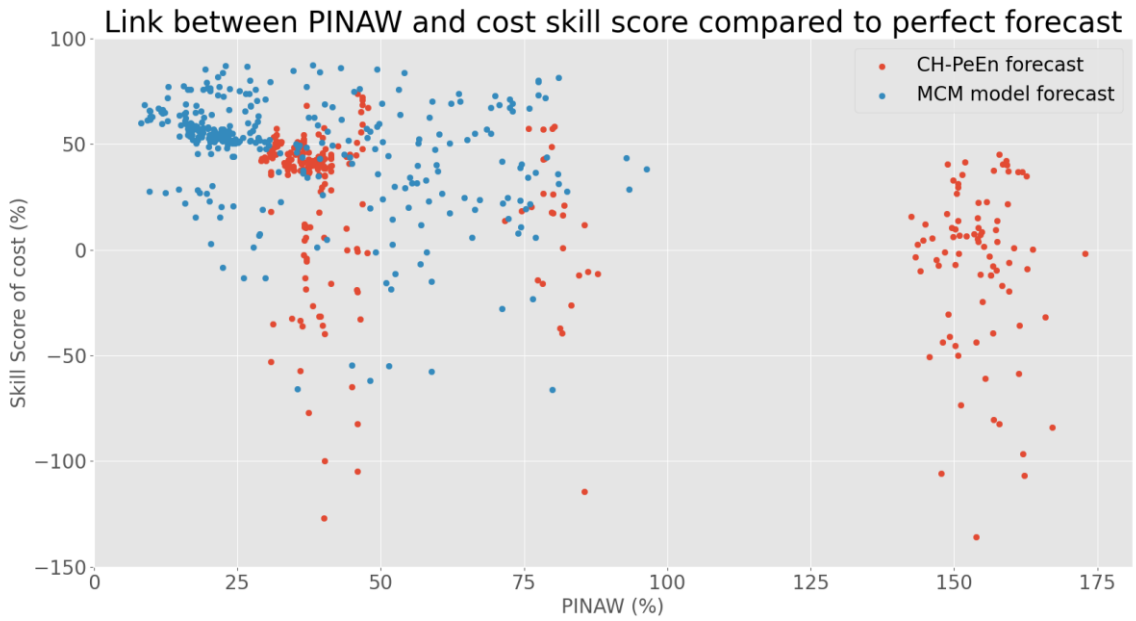


Figure 6.13 Link between PINAW and daily cost skill score

nCRPS vs cost skill score

In the end, the nCRPS evaluation and its link with system economic performance are shown in Figure 6.14, unlike PICP and PINAW which have their performance in different confidence level, CRPS evaluate all the quantile level of forecast uncertainty which could tell the general performance of a probabilistic forecast by combining the accuracy and the sharpness in a same time.

From this figure, we can conclude the following points:

- As shown in Figure 6.14, the MCM model has a better statistical performance than the CH-PeEn forecast, since the CRPS of MCM model is generally lower and closer to 0.
- Similarly, MCM model, a better statistical performance model has better economic performance as well, as shown in the figure, its cost skill score is closer to 1.
- However, the CRPS can only give general distinguishing information but not able to tell the concrete difference of economic performance. A same issue for all the classic statistical metrics, at a same level of error, the practical economic performance could be very different.

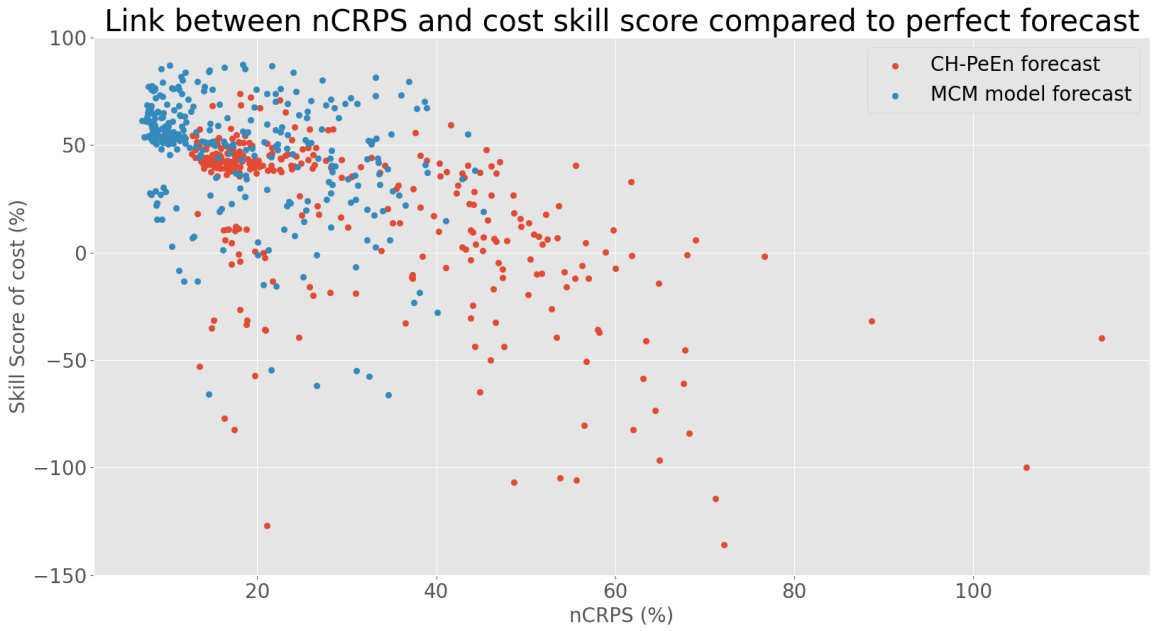


Figure 6.14 Link between nCRPS and daily system cost skill

6.4.4 Summary of results

From the results obtained in section 6.4, we observe that these statistical metrics have similar behavior and the conclusions drawn are similar:

- The classic statistical metrics are able to distinguish the performance of different forecast methods. For example, a higher quality MCM model has a more consistent performance in all statistical metrics compared to CH-PeEn forecast.
- The statistical metrics are not able to tell the practical HES economic performance. For the same level of error, the HES economic performance could be very different. However, a higher statistical quality model could generally have a better economic performance, which means there is perhaps a correlation.

Chapter 7

7. Conclusion and perspectives

7.1 Conclusion

This thesis is motivated by an industrial project of HES which is composed of 24 MW_p PV installation and several diesel generators for a total capacity of 65 MW in Mali. This HES supplies electricity to an off-grid gold mine, which has a quite high and consistent electric power demand 24/7, with an average of around 48 MW.

Unlike the grid connected HES with a normal domestique demand profile, which only needs a small energy storage system (ESS) to ensure the system stability [93], [94]. In an isolated HES system equipped with large share of RES and ESS, balancing the mismatch between variable energy supply and this high but almost constant energy demand profile would require a huge storage capacity both in power and energy, which raises considerable constraints of high investment costs, additional environmental impacts, potential maintenance issues, etc.

This thesis studies the possibility to avoid such costs by evaluating the performances of a storage-less HES for a large-scale off-grid industry. However, it is important to note that beyond this concrete industrial application in Mali, there is a considerable potential for storage-less HES designs to unlock the penetration of clean and affordable solar energy for such high energy-consuming industry in off-grid contexts (which is in line with the call of *GOAL 7* defined by the United Nations Environment Program (UNEP).

Not having an energy storage system in HES raises the challenge of compensating the solar energy variability through positive and negative spinning reserves and PV system curtailment. These compensation mechanisms are driven by intra-day solar forecasts complemented with an associated uncertainty, which allows to dynamically dispatch – start/stop/regulate – the genset, size the spinning reserves, or give curtailment orders. To efficiently leverage most of the PV production, and thus reduce the fuel consumption, the dynamic dispatching of the HES should be fed by high quality probabilistic forecasts. In order to manage the HES uncertainty induced by solar variability, we used a cost-based dynamic Unit Commitment approach using probabilistic solar forecasts as presented in chapter 4, which is driven by an objective function of minimizing the system cost.

For a proper operational evaluation of this dispatch, the deviations between the forecasted PV yield and the effective PV production were considered, inducing some additional costs compared to the

optimized system cost, coming from the use of positive spinning reserve, up to load shedding or even potential blackout, and PV curtailment in case of the negative spinning reserve limitation. The overall effective system cost is hence dependent on the quality of the forecasting method and the underlying solar variability, and is considered as the basis of our metrics to assess the HES performance.

Solar forecast is a cornerstone of off-grid energy system, and its accuracy is crucial for an efficient and cost-effective operation. The performances of a forecast algorithm are generally quantified by a set of error metrics, measuring the distance between actual and forecasted values. Considering the operation of a HES, the economic impact of a given forecast error can be different depending on the level of the energy demand and the states of the gensets. As a result, we may question **what is the relationship between the quality of probabilistic solar forecasting and the HES performance?**

To answer this main research question, we have developed a simulation framework to simulate the HES behavior using different probabilistic forecast methods, realistic PV production, real energy demand, and functional optimization algorithm for Unit Commitment problem. This simulation platform allows us to configure different parameters to be able to run different scenarios as explained in chapter 5, exploring notably the PV penetration rate, the update frequency of the genset dispatch order, the forecast horizon of the different probabilistic forecast methods, and different dynamic sizing approaches for the positive spinning reserve.

The analysis of the different scenarios enabled us to answer the following three research sub-questions to contribute to the main one:

- **What would be the best practical metrics to assess HES performance?**
- **How should a suitable platform to simulate the HES behavior be implemented?**
- **What are the benefits of using a forecast-driven HES compared to a diesel-only setup case?**

The sub-question 1: What would be the best practical metrics to assess HES performance?

As introduced in chapter 4, the classic way to assess the energy system performance is based on its frequency and voltage for stability and fuel consumption for economic performance. However, this thesis focuses on the evaluation of a storage-less design from an energy balancing perspective, raising the need for alternative metrics. For stability assessment, we consider the occurrence and magnitude of energy shortage/excess situations; whereas for economic performance evaluation, we use the effective final system cost, which is composed by the genset operating cost, PV system cost and system imbalance cost. Besides the classic criteria, the amount of spinning reserve used is also considered as a proxy to the practical forecast performance.

All of the associated actions inside the defined criteria could be simulated and given a corresponding cost as presented in chapter 4. For example, the PV curtailment cost is assigned to the energy excess cost, the load shedding cost is assigned to the energy shortage cost and the use of positive spinning reserve is assigned to the stability cost.

Finally, the effective total system cost derived from the associated cost of system stability, economic performance, and solar forecast performance, is used to represent the system performance. Compared to a reference case, the lower the system cost, the better the performance.

The sub-question 2: How should a suitable platform to simulate the HES behavior be implemented?

In terms of achieving the HES behavior simulation, a Unit Commitment framework is most frequently used. To the best of our knowledge, no study in the literature analyses the operational application of solar forecast in a HES, considering the deviations between the forecasts driving the HES planning and the actual observations. Some commercial software like HOMER [95] and PVsyst [96] provide a possibility of energy system simulation, however, they could not show the complete calculation detail to users and are less flexible to simulate the HES behavior using solar forecasting. In this thesis, we introduce an innovative way to simulate the HES using short-term solar forecast information to dynamically dispatch the diesel generators. Rather than restraining the simulation and analysis only to the planning of the energy system (unit commitment using the forecast and the balancing mechanisms of the HES), our method goes beyond and analyses the actual operation dispatch and performs several intra-day optimization cycles in order to ingest the latest forecast information available at each genset update time. By doing so, we aim to reflect the actual use of the HES in an industrial context.

The forecast information used is updated every 15-min and the genset order could be generated at each forecast update. Using frequently updated forecasts could provide more recent and, thus, accurate information, but a too frequent genset update could significantly reduce its lifetime. Hence, two parameters of forecast horizon and genset update time are introduced in our simulation to balance the forecast accuracy and the genset lifetime.

Our simulation platform is a Python-based simulation tool, using a cost-based approach, where the objective function minimizes the system cost. It allows users to simulate the actual HES behavior when using solar forecasts, and provides a comparable, easily understandable system cost as performance indicator. It is particularly suitable for evaluating the performances of off-grid HES that use short-term solar forecasting as input data. This is very important since classic forecast error metrics cannot correctly describe its practical economic performance, making this simulation tool a more direct and intuitive way to do so. Notice that the simulation implementation should be precisely tuned according to the considered case study, then the associated parameters for each simulation should be precisely configured according to the actual need of practical case.

The sub-question 3: What are the benefits of using a forecast-driven HES compared to a diesel-only setup case?

As shown in chapter 5 and chapter 6, we can verify that the gain of using solar forecast depends on several factors, such as the presence of the weather situation, the performances of the forecast, the characteristics of the energy system, the PV penetration rate, and so on.

Furthermore, this gain should be evaluated relative to a reference case. In this thesis, a scenario where no PV is installed is chosen as the reference case. For a 50% PV penetration rate, around 12% PV energy share, the economic gain ranges between 2.7% for a simple baseline CH-PeEn forecast and 11.7% for a perfect prognosis.

In general, more PV capacity means a higher probability to get higher economic gain. However, results show a saturation level of PV penetration beyond which does not happen: the storage-less HES has a limited ability to accommodate PV energy, leading to considerable PV curtailment. This degrades the

economic and environmental impact of PV, up to the point of the HES being unprofitable compared to the no PV case.

Furthermore, installing a huge PV system but without a good management strategy in PV power using and spinning reserve sizing would increase the unnecessary cost, which is not worthy both in economic and environmental aspect.

The forecast method should also be chosen carefully. In general, a high-quality forecast method could provide a relatively more consistent performance both in clear and variable day. But as shown in chapter 5, a conservative method has a correct performance in clear day, but in variable day, although it provides similar reliability, it comes with a lot more PV curtailment, and increases the system cost. Nevertheless, for a low PV penetration rate, the difference in economic gain between forecast methods with different qualities is very small. This tells us that for PV penetration levels below the saturation point, rather than investing in improving the quality of forecast method, increasing the PV penetration should be prioritized to reduce the system cost.

Besides the gain in direct economic savings, solar forecasting could also help in extending the service life of genset, by better organizing the genset operation to stay at its optimal operating range, which brings economic savings indirectly.

The main question: What is the relationship between the quality of probabilistic solar forecasting and the HES performance?

The classic way of evaluating solar forecast performance is to use a statistical metric like RMSE, MAE for deterministic forecasts or PINAW, PICP, CRPS for probabilistic ones. Yet, as discussed in chapter 4, these are insufficient to consider the practical performance of the isolated storage-less HES itself. For example, an overestimated and underestimated forecast could lead to quite different results. As shown in chapter 5 and chapter 6, the classic statistical metrics could tell the general performance difference between two forecast methods, but they cannot correctly reflect the practical economic performance. For example, the same level of daily RMSE could have a huge difference in practical economic performance.

Therefore, we use a newly defined system cost skill score as shown in chapter 6, to correctly reflect the practical economic gain compared to the perfect forecast and no PV case when using different forecast methods. The difference between these two cases provides the biggest range in terms of economic saving, which is used as the denominator of this cost skill score, whereas the difference between the no PV case and forecast model under consideration is used as the numerator, to reflect the actual economic performance.

Thus, in chapter 6 we have searched for possible correlations between the cost skill score and the statistical forecasting performance, according to different metrics. The results shown prove again our conclusions:

- The high-quality forecast can provide more consistent performance in different weather situations, and provide a better cost skill score compared to a more basic method,
- The quality of the forecast method is not the only element that influences the HES economic performance; below the PV optimal penetration level, increasing PV capacity can be more impactful,

- The classic statistical error metrics can give distinguishing information between two forecast methods, but cannot predict their practical economic performance.

Besides, we have also done several attempts in searching a dedicated metrics modified from the classic metrics to properly evaluate the HES economic performance. For example, as shown in Figure 7.1, for a HES management strategy where P50 is used for genset dispatching and the uncertainty range of $P_{50} - P_{\varepsilon}$ is used for spinning reserve sizing, rather than focusing on the uncertainty range around P50, we separate the whole uncertainty range into 3 parts according to the way we use solar forecast information in our HES simulation:

- The upper zone, called MOR (Measurement Over Rate) zone - for the underestimated forecast cases,
- The middle zone for the main target coverage zone,
- The lower zone under P_{ε} called MUR (Measurement Under Rate) zone - for load shedding cases.

Then, we process the RMSE, MAE for these three zones and try to link them with the corresponding cost skill score, but there is no clear correlation has been found.

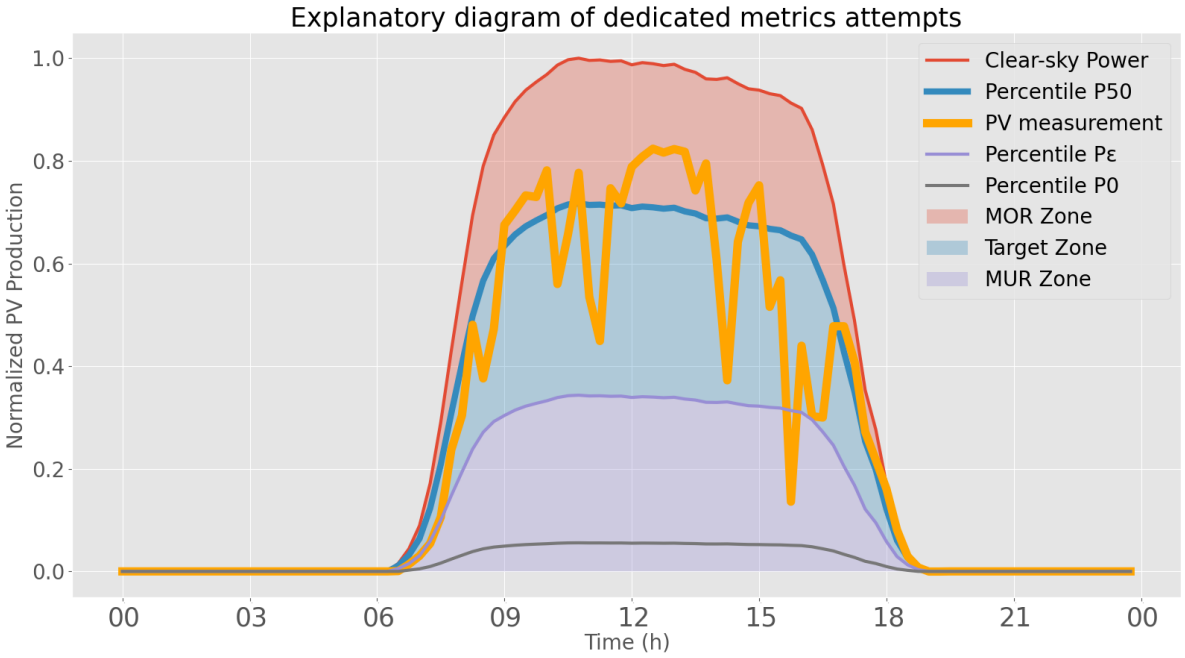


Figure 7.1 Explanatory diagram of dedicated metrics attempts

To conclude, there is still a long way to go for determining a suitable metric of practical performance evaluation. The linear or quasi-linear characteristic of statistical metrics has its limit in describing the practical performance, and hence should not be used as only performance indicator.

Finally, rather than using standard error metrics, using a simulation tool like ours is perhaps a more direct and intuitive way to assess the HES economic performance, since it simulates the actual behavior of solar forecast application in a HES. Accompanied by the newly introduced cost skill score, the simulation result based on our cost-based approach could correctly reflect the system performance and the influence of solar forecast used. However, these results and conclusions should be considered with caution, being aware of the limitations of the proposed methodology. That is the aim of the following section.

7.2 Limitations of simulation of this thesis

The simulation framework proposed in this thesis has some limitations, mainly driven by the need of modelling assumptions and simplifications as well as limitations in the data. The first one is the aspect of electronic dynamic behavior which is not considered. In real-world application, the system imbalance situation would occur in a scale of 100 milliseconds, which is much shorter than the 15-min temporal resolution that we choose in our simulation work. However, as explained in chapter 4, modeling a HES in such a fine scale requires specific modeling tool and is out of the scope of this thesis. Hence, to simplify the work, we assumed that if the power supply/demand balance is respected in a 15-min scale, the dynamic aspect of power electronics is respected as well. One should note that this is a sufficient but unnecessary condition, which means that when the higher scale balance is ensured, the smaller scale one is automatically achieved, but when a smaller scale balance is ensured, it does not mean the balance is ensured for a longer period.

The second limitation is that the potential failure of energy system component and the uncertainty of energy demand are not considered. As detailed in chapter 3, the “n-1” principle, which means that the spinning reserve sizing should be equal or higher than the largest online genset. This principle is not considered in the simulation, or it would be much higher and cover all the solar forecast error, which makes it difficult to evaluate the forecast performance. In case of considering the potential failure of diesel generating unit, the “n-1” principle has to be considered.

The third limitation is that the genset installation cost (also designated as capital expenditure or CAPEX) is not considered, as explained in chapter 4. Thus, the optimization disregards the impact of changes in genset lifetime. However, this should have a minimal impact since the optimization results show a mostly stable genset operating time, since the demand under consideration is quite stable.

The fourth limitation comes from the data used in the simulation. As explained in chapter 4, the PV production time-series used are reproduced from a part of real measurements, to avoid the PV production variation which is influenced by external factors such as site maintenance or the actual operation that the PV plant followed (e.g., curtailment, shutdown, maintenance, etc.). This option allows to evaluate the forecast method performance against natural solar resource variability but not other factors. Besides, the energy demand in our case is quite stable and assumed to be perfectly known, which greatly simplifies the power management strategy and simulation process. For a case that needs to consider the energy demand uncertainty, the need of reserve capacity and the resulting economic gain could be different.

The last limitation is that we consider each day of our simulation as an independent day. This is driven by the fact that in a storage-less design, there is no storage state of charge which needs to be modelled in a continuous way. The stability of the energy demand also makes that all the diesel generators have to be turned on at night to commit the demand and results in a quite similar state of operation at the beginning of each day. Otherwise, the simulation work would be more complex and more computationally demanding.

7.3 Perspectives

A HES with battery design

This thesis provides a simulation study of isolated storage-less HES, proposing a modelling and evaluation frameworks. Since the storage-less design limits the capacity of accommodating PV energy, but also of the valorization of solar forecast, future work should consider the implementation and evaluation of the same scenarios but with a HES equipped with an ESS. It could explore again the optimal PV penetration rate, and see if the ESS could extend it, how the size of ESS changes the gain of using PV energy, and the impact of the solar forecast used in economic performance. We could also compare the difference in economic gain by using solar forecast in a storage-less HES and a HES equipped with an ESS.

At which point the PV overbuilding approach could be achieved for our case study?

According to the firm power approach proposed in [97], overbuilding the PV installation could minimize the need for ESS and even avoid the need for solar forecast (and their associated costs). However, for the isolated HES like our case study, with a considerably and stable demand profile (almost stable at 50MW) even at nighttime context, it would be different and difficult to apply the 100% PV with ESS design since the investment cost is too important.

However, it would be interesting to see at which point this approach could be applied in our case study, coupled with an ESS, to find out the optimal compromise of PV penetration and ESS size. Besides, one could also consider also the environmental efficiency like the reduction of Greenhouse Gas (GHG) emissions, to see how economic cost and environmental impacts could be jointly minimized.

What about using a Numerical Weather Prediction (NWP) forecast model?

To notice that non-baseline and non-perfect forecast method used in the thesis - the MCM approach - is also limited in terms of forecast horizon, since it is mostly suited for shorter term horizons. To address this, testing the integration of Numerical Weather Prediction (NWP) forecasts would allow to better understand the value of an operational solution suitable for longer term horizons to the management of the genset dispatching.

Decision-making on the energy transition aspect

Due to the considerable investment cost associated with ESS, the possibility of replacing the diesel generators by an ESS with additional PV capacity has never been raised by the industrial partner of this thesis. Thus, an economic and environmental evaluation of a diesel-based setup against the storage-less and the 100% renewable ESS-based configurations would support the investors' decision-making process. The environmental component could be achieved through a Life Cycle Assessment (LCA) [7], which takes into account not only the fuel-related emissions but also the generators (both diesel and PV) embedded impact. It is possible that such analysis would position the storage-less concept as a reference short-term transitory solution, where an isolated diesel-only system is first coupled with PV capacity and diesel generators are replaced by additional PV and ESS as they reach their end of life.

Since the energy demand of case study at night is similar to daytime, having 100% PV with ESS is almost impossible only if we accept to have a huge ESS with a very high investment and maintenance cost. Moreover, the state, the lifetime, and the total environmental impact of diesel generators would

change a lot in the final result. In brief, besides the performance evaluation, the LCA environmental evaluation would be very interesting and necessary to do for decision-making in the future.

Potential application

By design, this thesis is more applicative compared to other fundamental researches. The simulation platform implemented in this thesis allows the industrial user to simulate and estimate the HES performance elsewhere of our case study. And the application of this simulation with dynamic genset dispatching could potentially allow a HES to be run in a 100% automatic way, with a pre-condition that all the generating units used should be automatically dispatchable. Besides, the forecast method used should a relatively high-quality one, and be correctly calibrated since the spinning reserve is sized with the forecast uncertainty range.

References

[1] V. Smil, *Energy Transitions: History, Requirements, Prospects*. ABC-CLIO, 2010.

[2] BP Statistical Review of World Energy, "bp-stats-review-2021-full-report.pdf," 2021. https://www.bp.com/content/dam/bp/business-sites/en/global/corporate/pdfs/energy-economics/statistical-review/bp-stats-review-2021-full-report.pdf (accessed Oct. 20, 2021).

[3] IEA, "Global Energy Review: CO2 Emissions in 2020 – Analysis - IEA," p. 14, 2020.

[4] IPCC, "AR5 Climate Change 2013: The Physical Science Basis — IPCC," 2013. https://www.ipcc.ch/report/ar5/wg1/ (accessed Jan. 12, 2022).

[5] N. Eisenhauer, S. Cesarz, R. Koller, K. Worm, and P. B. Reich, "Global change belowground: impacts of elevated CO2, nitrogen, and summer drought on soil food webs and biodiversity," *Glob. Change Biol.*, vol. 18, no. 2, pp. 435–447, 2012, doi: 10.1111/j.1365-2486.2011.02555.x.

[6] N. Teixidó *et al.*, "Functional biodiversity loss along natural CO2 gradients," *Nat. Commun.*, vol. 9, no. 1, p. 5149, Dec. 2018, doi: 10.1038/s41467-018-07592-1.

[7] R. Besseau, "Analyse de cycle de vie de scénarios énergétiques intégrant la contrainte d'adéquation temporelle production-consommation," 2019.

[8] B. C. McLellan, G. D. Corder, D. P. Giurco, and K. N. Ishihara, "Renewable energy in the minerals industry: a review of global potential," *J. Clean. Prod.*, vol. 32, pp. 32–44, Sep. 2012, doi: 10.1016/j.jclepro.2012.03.016.

[9] J. J. S. Guilbaud, "Hybrid Renewable Power Systems for the Mining Industry: System Costs, Reliability Costs, and Portfolio Cost Risks," Doctoral, UCL (University College London), 2016. Accessed: Jul. 01, 2019. [Online]. Available: http://discovery.ucl.ac.uk/1528681/

[10] B. Parida, S. Iniyan, and R. Goic, "A review of solar photovoltaic technologies," *Renew. Sustain. Energy Rev.*, vol. 15, no. 3, pp. 1625–1636, Apr. 2011, doi: 10.1016/j.rser.2010.11.032.

[11] K. Kawajiri, T. Oozeki, and Y. Genchi, "Effect of Temperature on PV Potential in the World," *Environ. Sci. Technol.*, vol. 45, no. 20, pp. 9030–9035, Oct. 2011, doi: 10.1021/es200635x.

[12] S. Baurzhan and G. P. Jenkins, "Off-grid solar PV: Is it an affordable or appropriate solution for rural electrification in Sub-Saharan African countries?," *Renew. Sustain. Energy Rev.*, vol. 60, pp. 1405–1418, Jul. 2016, doi: 10.1016/j.rser.2016.03.016.

[13] A. Bichet, B. Hingray, G. Evin, A. Diedhiou, C. M. F. Kebe, and S. Anquetin, "Potential impact of climate change on solar resource in Africa for photovoltaic energy: analyses from CORDEX-AFRICA climate experiments," *Environ. Res. Lett.*, vol. 14, no. 12, p. 124039, Dec. 2019, doi: 10.1088/1748-9326/ab500a.

[14] INREA, "Africa 2030: Roadmap for a Renewable Energy Future," p. 72, 2015.

[15] D. Yamegueu, Y. Azoumah, X. Py, and N. Zongo, "Experimental study of electricity generation by Solar PV/diesel hybrid systems without battery storage for off-grid areas," *Renew. Energy*, vol. 36, no. 6, pp. 1780–1787, Jun. 2011, doi: 10.1016/j.renene.2010.11.011.

[16] D. Lew *et al.*, "The Western Wind and Solar Integration Study Phase 2," NREL/TP--5500-55588, 1220243, Sep. 2013. doi: 10.2172/1220243.

[17] M. K. Deshmukh and S. S. Deshmukh, "Modeling of hybrid renewable energy systems," *Renew. Sustain. Energy Rev.*, vol. 12, no. 1, pp. 235–249, Jan. 2008, doi: 10.1016/j.rser.2006.07.011.

[18] S. M. Shaahid and I. El-Amin, "Techno-economic evaluation of off-grid hybrid photovoltaic–diesel–battery power systems for rural electrification in Saudi Arabia—A way forward for sustainable development," *Renew. Sustain. Energy Rev.*, vol. 13, no. 3, pp. 625–633, Apr. 2009, doi: 10.1016/j.rser.2007.11.017.

[19] S. Rehman and L. M. Al-Hadhrami, "Study of a solar PV–diesel–battery hybrid power system for a remotely located population near Rafha, Saudi Arabia," *Energy*, vol. 35, no. 12, pp. 4986–4995, Dec. 2010, doi: 10.1016/j.energy.2010.08.025.

[20] A. Urtasun, P. Sanchis, D. Barricarte, and L. Marroyo, "Energy management strategy for a battery-diesel stand-alone system with distributed PV generation based on grid frequency

modulation," *Renew. Energy*, vol. 66, pp. 325–336, Jun. 2014, doi: 10.1016/j.renene.2013.12.020.
[21] N. Agarwal, A. Kumar, and Varun, "Optimization of grid independent hybrid PV–diesel–battery system for power generation in remote villages of Uttar Pradesh, India," *Energy Sustain. Dev.*, vol. 17, no. 3, pp. 210–219, Jun. 2013, doi: 10.1016/j.esd.2013.02.002.
[22] W. Cole, A. W. Frazier, and C. Augustine, "Cost Projections for Utility-Scale Battery Storage: 2021 Update," *Renew. Energy*, p. 21, 2021.
[23] J. Antonanzas, N. Osorio, R. Escobar, R. Urraca, F. J. Martinez-de-Pison, and F. Antonanzas-Torres, "Review of photovoltaic power forecasting," *Sol. Energy*, vol. 136, pp. 78–111, Oct. 2016, doi: 10.1016/j.solener.2016.06.069.
[24] M. Diagne, M. David, P. Lauret, J. Boland, and N. Schmutz, "Review of solar irradiance forecasting methods and a proposition for small-scale insular grids," *Renew. Sustain. Energy Rev.*, vol. 27, pp. 65–76, Nov. 2013, doi: 10.1016/j.rser.2013.06.042.
[25] S. Sobri, S. Koohi-Kamali, and N. Abd. Rahim, "Solar photovoltaic generation forecasting methods: A review," *Energy Convers. Manag.*, vol. 156, pp. 459–497, Jan. 2018, doi: 10.1016/j.enconman.2017.11.019.
[26] L. Vallance, "Synergie des mesures pyranométriques et des images hémisphériques in situ avec des images satellites météorologiques pour la prévision photovoltaïque," 2018. Accessed: Jun. 28, 2019. [Online]. Available: https://tel.archives-ouvertes.fr/tel-02078908
[27] J. Zhang, B.-M. Hodge, and A. Florita, "Metrics for Evaluating the Accuracy of Solar Power Forecasting: Preprint," p. 10, 2013.
[28] D. Yang, J. Kleissl, C. A. Gueymard, H. T. C. Pedro, and C. F. M. Coimbra, "History and trends in solar irradiance and PV power forecasting: A preliminary assessment and review using text mining," *Sol. Energy*, vol. 168, pp. 60–101, Jul. 2018, doi: 10.1016/j.solener.2017.11.023.
[29] T. Carrière, R. Amaro e Silva, F. Zhuang, Y.-M. Saint-Drenan, and P. Blanc, "A New Approach for Satellite-Based Probabilistic Solar Forecasting with Cloud Motion Vectors," *Energies*, vol. 14, no. 16, p. 4951, Aug. 2021, doi: 10.3390/en14164951.
[30] D. Yang, J. Kleissl, C. A. Gueymard, H. T. C. Pedro, and C. F. M. Coimbra, "History and trends in solar irradiance and PV power forecasting: A preliminary assessment and review using text mining," *Sol. Energy*, vol. 168, pp. 60–101, Jul. 2018, doi: 10.1016/j.solener.2017.11.023.
[31] P. Lauret, M. David, and P. Pinson, "Verification of solar irradiance probabilistic forecasts," *Sol. Energy*, vol. 194, pp. 254–271, Dec. 2019, doi: 10.1016/j.solener.2019.10.041.
[32] R. Marquez and C. F. M. Coimbra, "Proposed Metric for Evaluation of Solar Forecasting Models," *J. Sol. Energy Eng.*, vol. 135, no. 1, Feb. 2013, doi: 10.1115/1.4007496.
[33] L. Frías-Paredes, F. Mallor, M. Gastón-Romeo, and T. León, "Assessing energy forecasting inaccuracy by simultaneously considering temporal and absolute errors," *Energy Convers. Manag.*, vol. 142, pp. 533–546, Jun. 2017, doi: 10.1016/j.enconman.2017.03.056.
[34] H. Sakoe and S. Chiba, "Dynamic programming algorithm optimization for spoken word recognition," *IEEE Trans. Acoust. Speech Signal Process.*, vol. 26, no. 1, pp. 43–49, Feb. 1978, doi: 10.1109/TASSP.1978.1163055.
[35] N. P. Padhy, "Unit Commitment—A Bibliographical Survey," *IEEE Trans. Power Syst.*, vol. 19, no. 2, pp. 1196–1205, May 2004, doi: 10.1109/TPWRS.2003.821611.
[36] M. Lave, R. J. Broderick, and M. J. Reno, "Solar variability zones: Satellite-derived zones that represent high-frequency ground variability," *Sol. Energy*, vol. 151, pp. 119–128, Jul. 2017, doi: 10.1016/j.solener.2017.05.005.
[37] B. Palminter, B. Mather, M. Coddington, K. Baker, and F. Ding, "On the Path to SunShot: Emerging Issues and Challenges in Integrating Solar with the Distribution System," p. 99.
[38] G. B. Sheble and G. N. Fahd, "Unit commitment literature synopsis," *IEEE Trans. Power Syst.*, vol. 9, no. 1, pp. 128–135, Feb. 1994, doi: 10.1109/59.317549.
[39] Q. Zhai, X. Guan, and J. Yang, "Fast unit commitment based on optimal linear approximation to nonlinear fuel cost: Error analysis and applications," *Electr. Power Syst. Res.*, vol. 79, no. 11, pp. 1604–1613, Nov. 2009, doi: 10.1016/j.epsr.2009.06.005.

[40] N. P. Padhy, “Unit commitment-a bibliographical survey,” *IEEE Trans. Power Syst.*, vol. 19, no. 2, pp. 1196–1205, May 2004, doi: 10.1109/TPWRS.2003.821611.
[41] S. Y. Abujarad, M. W. Mustafa, and J. J. Jamian, “Recent approaches of unit commitment in the presence of intermittent renewable energy resources: A review,” *Renew. Sustain. Energy Rev.*, vol. 70, pp. 215–223, Apr. 2017, doi: 10.1016/j.rser.2016.11.246.
[42] R. Quan, J. Jian, and L. Yang, “An improved priority list and neighborhood search method for unit commitment,” *Int. J. Electr. Power Energy Syst.*, vol. 67, pp. 278–285, May 2015, doi: 10.1016/j.ijepes.2014.11.025.
[43] P. G. Lowery, “Generating Unit Commitment by Dynamic Programming,” *IEEE Trans. Power Appar. Syst.*, vol. PAS-85, no. 5, pp. 422–426, May 1966, doi: 10.1109/TPAS.1966.291679.
[44] I. A. Farhat and M. E. El-Hawary, “Optimization methods applied for solving the short-term hydrothermal coordination problem,” *Electr. Power Syst. Res.*, vol. 79, no. 9, pp. 1308–1320, Sep. 2009, doi: 10.1016/j.epsr.2009.04.001.
[45] T. Li and M. Shahidehpour, “Price-Based Unit Commitment: A Case of Lagrangian Relaxation Versus Mixed Integer Programming,” *IEEE Trans. Power Syst.*, vol. 20, no. 4, pp. 2015–2025, Nov. 2005, doi: 10.1109/TPWRS.2005.857391.
[46] Morales-Espana, J. M. Latorre, and A. Ramos, “Tight and Compact MILP Formulation for the Thermal Unit Commitment Problem,” *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 4897–4908, Nov. 2013, doi: 10.1109/TPWRS.2013.2251373.
[47] D. Palis and S. Palis, “Efficient Unit Commitment - A modified branch-and-bound approach,” in *2016 IEEE Region 10 Conference (TENCON)*, Nov. 2016, pp. 267–271. doi: 10.1109/TENCON.2016.7848004.
[48] M. Madrigal and V. H. Quintana, “An interior-point/cutting-plane method to solve unit commitment problems,” in *Proceedings of the 21st International Conference on Power Industry Computer Applications. Connecting Utilities. PICA 99. To the Millennium and Beyond (Cat. No.99CH36351)*, Santa Clara, CA, USA, 1999, pp. 203–209. doi: 10.1109/PICA.1999.779404.
[49] J. Jablonský, “Benchmarks for Current Linear and Mixed Integer Optimization Solvers,” *Acta Univ. Agric. Silvic. Mendel. Brun.*, vol. 63, no. 6, pp. 1923–1928, Jan. 2016, doi: 10.11118/actaun201563061923.
[50] W. E. Helm and J.-E. Justkowiak, “Extension of Mittelmann’s Benchmarks: Comparing the Solvers of SAS and Gurobi,” in *Operations Research Proceedings 2016*, A. Fink, A. Fügenschuh, and M. J. Geiger, Eds. Cham: Springer International Publishing, 2018, pp. 607–613. doi: 10.1007/978-3-319-55702-1_80.
[51] H. Nabli, “An overview on the simplex algorithm,” *Appl. Math. Comput.*, vol. 210, no. 2, pp. 479–489, Apr. 2009, doi: 10.1016/j.amc.2009.01.013.
[52] K. E. Trenberth, J. T. Fasullo, and J. Kiehl, “Earth’s Global Energy Budget,” *Bull. Am. Meteorol. Soc.*, vol. 90, no. 3, pp. 311–324, Mar. 2009, doi: 10.1175/2008BAMS2634.1.
[53] M. Wild *et al.*, “The energy balance over land and oceans: an assessment based on direct observations and CMIP5 climate models,” *Clim. Dyn.*, vol. 44, no. 11, pp. 3393–3429, Jun. 2015, doi: 10.1007/s00382-014-2430-z.
[54] R. Perez, P. Lauret, M. Perez, M. David, T. E. Hoff, and S. Kivalov, “Solar Resource Variability,” in *Wind Field and Solar Radiation Characterization and Forecasting*, R. Perez, Ed. Cham: Springer International Publishing, 2018, pp. 149–170. doi: 10.1007/978-3-319-76876-2_7.
[55] R. H. Inman, H. T. C. Pedro, and C. F. M. Coimbra, “Solar forecasting methods for renewable energy integration,” *Prog. Energy Combust. Sci.*, vol. 39, no. 6, pp. 535–576, Dec. 2013, doi: 10.1016/j.pecs.2013.06.002.
[56] V. Badescu *et al.*, “Accuracy analysis for fifty-four clear-sky solar radiation models using routine hourly global irradiance measurements in Romania,” *Renew. Energy*, vol. 55, pp. 85–103, Jul. 2013, doi: 10.1016/j.renene.2012.11.037.
[57] N. A. Engerer and F. P. Mills, “Validating nine clear sky radiation models in Australia,” *Sol. Energy*, vol. 120, pp. 9–24, Oct. 2015, doi: 10.1016/j.solener.2015.06.044.

[58] P. Ineichen, “Validation of models that estimate the clear sky global and beam solar irradiance,” *Sol. Energy*, vol. 132, pp. 332–344, Jul. 2016, doi: 10.1016/j.solener.2016.03.017.
[59] M. Lefèvre *et al.*, “McClear: a new model estimating downwelling solar radiation at ground level in clear-sky conditions,” *Atmospheric Meas. Tech.*, vol. 6, no. 9, pp. 2403–2418, Sep. 2013, doi: 10.5194/amt-6-2403-2013.
[60] C. Rigollier, O. Bauer, and L. Wald, “On the clear sky model of the ESRA — European Solar Radiation Atlas — with respect to the heliosat method,” *Sol. Energy*, vol. 68, no. 1, pp. 33–48, Jan. 2000, doi: 10.1016/S0038-092X(99)00055-9.
[61] A. Mills and R. Wiser, “Implications of Wide-Area Geographic Diversity for Short- Term Variability of Solar Power,” LBNL-3884E, 986925, Aug. 2010. doi: 10.2172/986925.
[62] T. E. Hoff and R. Perez, “Modeling PV fleet output variability,” *Sol. Energy*, vol. 86, no. 8, pp. 2177–2189, Aug. 2012, doi: 10.1016/j.solener.2011.11.005.
[63] R. Dambreville, “Prévision du rayonnement solaire global par télédétection pour la gestion de la production d’énergie photovoltaïque,” phdthesis, Université de Grenoble, 2014. Accessed: Mar. 03, 2022. [Online]. Available: https://tel.archives-ouvertes.fr/tel-01130251
[64] G. M. Tina, S. De Fiore, and C. Ventura, “Analysis of forecast errors for irradiance on the horizontal plane,” *Energy Convers. Manag.*, vol. 64, pp. 533–540, Dec. 2012, doi: 10.1016/j.enconman.2012.05.031.
[65] C. Voyant *et al.*, “A Monte Carlo based solar radiation forecastability estimation,” *J. Renew. Sustain. Energy*, vol. 13, no. 2, p. 026501, Mar. 2021, doi: 10.1063/5.0042710.
[66] V. P. Lonij, A. E. Brooks, K. Koch, and A. D. Cronin, “Analysis of 80 rooftop PV systems in the Tucson, AZ area,” in *2012 38th IEEE Photovoltaic Specialists Conference*, Jun. 2012, pp. 000549–000553. doi: 10.1109/PVSC.2012.6317674.
[67] S. Alessandrini, L. Delle Monache, S. Sperati, and G. Cervone, “An analog ensemble for short-term probabilistic solar power forecast,” *Appl. Energy*, vol. 157, pp. 95–110, Nov. 2015, doi: 10.1016/j.apenergy.2015.08.011.
[68] D. Yang, “A universal benchmarking method for probabilistic solar irradiance forecasting,” *Sol. Energy*, vol. 184, pp. 410–416, May 2019, doi: 10.1016/j.solener.2019.04.018.
[69] Y. Chu and C. F. M. Coimbra, “Short-term probabilistic forecasts for Direct Normal Irradiance,” *Renew. Energy*, vol. 101, pp. 526–536, Feb. 2017, doi: 10.1016/j.renene.2016.09.012.
[70] J. Munkhammar and J. Widén, “A Markov-chain probability distribution mixture approach to the clear-sky index,” *Sol. Energy*, vol. 170, pp. 174–183, Aug. 2018, doi: 10.1016/j.solener.2018.05.055.
[71] J. Munkhammar, D. van der Meer, and J. Widén, “Probabilistic forecasting of high-resolution clear-sky index time-series using a Markov-chain mixture distribution model,” *Sol. Energy*, vol. 184, pp. 688–695, May 2019, doi: 10.1016/j.solener.2019.04.014.
[72] T. Muneer, “Solar radiation model for Europe,” *Build. Serv. Eng. Res. Technol.*, vol. 11, no. 4, pp. 153–163, Nov. 1990, doi: 10.1177/014362449001100405.
[73] S. R. Williams, T. R. Betts, T. Helf, R. Gottschalg, H. G. Beyer, and D. G. Infield, “Modelling long-term module performance based on realistic reporting conditions with consideration to spectral effects,” p. 4.
[74] I. de la Parra, M. Muñoz, E. Lorenzo, M. García, J. Marcos, and F. Martínez-Moreno, “PV performance modelling: A review in the light of quality assurance for large PV plants,” *Renew. Sustain. Energy Rev.*, vol. 78, pp. 780–797, Oct. 2017, doi: 10.1016/j.rser.2017.04.080.
[75] K. Shivarama Krishna and K. Sathish Kumar, “A review on hybrid renewable energy systems,” *Renew. Sustain. Energy Rev.*, vol. 52, pp. 907–916, Dec. 2015, doi: 10.1016/j.rser.2015.07.187.
[76] C. Ren, Y. Xu, Y. Zhang, and R. Zhang, “A Hybrid Randomized Learning System for Temporal-Adaptive Voltage Stability Assessment of Power Systems,” *IEEE Trans. Ind. Inform.*, vol. 16, no. 6, p. 13, 2020.
[77] W. Margaret Amutha and V. Rajini, “Techno-economic evaluation of various hybrid power systems for rural telecom,” *Renew. Sustain. Energy Rev.*, vol. 43, pp. 553–561, Mar. 2015, doi: 10.1016/j.rser.2014.10.103.

[78] J. Zhang, B.-M. Hodge, and A. Florita, “Metrics for Evaluating the Accuracy of Solar Power Forecasting: Preprint,” p. 10.
[79] D. Yang *et al.*, “Verification of deterministic solar forecasts,” *Sol. Energy*, vol. 210, pp. 20–37, Nov. 2020, doi: 10.1016/j.solener.2020.04.019.
[80] P. Lauret, M. David, and P. Pinson, “Verification of solar irradiance probabilistic forecasts,” *Sol. Energy*, vol. 194, pp. 254–271, Dec. 2019, doi: 10.1016/j.solener.2019.10.041.
[81] T. Gneiting and A. E. Raftery, “Strictly Proper Scoring Rules, Prediction, and Estimation,” *J. Am. Stat. Assoc.*, vol. 102, no. 477, pp. 359–378, Mar. 2007, doi: 10.1198/016214506000001437.
[82] M. S. Adaramola, M. Agelin-Chaab, and S. S. Paul, “Analysis of hybrid energy systems for application in southern Ghana,” *Energy Convers. Manag.*, vol. 88, pp. 284–295, Dec. 2014, doi: 10.1016/j.enconman.2014.08.029.
[83] B. E. Türkay and A. Y. Telli, “Economic analysis of standalone and grid connected hybrid energy systems,” *Renew. Energy*, vol. 36, no. 7, pp. 1931–1943, Jul. 2011, doi: 10.1016/j.renene.2010.12.007.
[84] M. S. Adaramola, S. S. Paul, and O. M. Oyewola, “Assessment of decentralized hybrid PV solar-diesel power system for applications in Northern part of Nigeria,” *Energy Sustain. Dev.*, vol. 19, pp. 72–82, Apr. 2014, doi: 10.1016/j.esd.2013.12.007.
[85] K. Branker, M. J. M. Pathak, and J. M. Pearce, “A review of solar photovoltaic levelized cost of electricity,” *Renew. Sustain. Energy Rev.*, vol. 15, no. 9, pp. 4470–4482, Dec. 2011, doi: 10.1016/j.rser.2011.07.104.
[86] R. Kerr, J. Scheidt, A. Fontanna, and J. Wiley, “Unit Commitment,” *IEEE Trans. Power Appar. Syst.*, vol. PAS-85, no. 5, pp. 417–421, May 1966, doi: 10.1109/TPAS.1966.291678.
[87] W. Ongsakul and N. Petcharaks, “Unit Commitment by Enhanced Adaptive Lagrangian Relaxation,” *IEEE Trans. Power Syst.*, vol. 19, no. 1, pp. 620–628, Feb. 2004, doi: 10.1109/TPWRS.2003.820707.
[88] Ph. Blanc and L. Wald, “The SG2 algorithm for a fast and accurate computation of the position of the Sun for multi-decadal time period,” *Sol. Energy*, vol. 86, no. 10, pp. 3072–3083, Oct. 2012, doi: 10.1016/j.solener.2012.07.018.
[89] S. E. Haupt *et al.*, “The Use of Probabilistic Forecasts: Applying Them in Theory and Practice,” *IEEE Power Energy Mag.*, vol. 17, no. 6, pp. 46–57, Nov. 2019, doi: 10.1109/MPE.2019.2932639.
[90] A. P. Dobos, P. Gilman, and M. Kasberg, “P50/P90 Analysis for Solar Energy Systems Using the System Advisor Model: Preprint,” National Renewable Energy Lab. (NREL), Golden, CO (United States), NREL/CP-6A20-54488, Jun. 2012. Accessed: Aug. 08, 2022. [Online]. Available: https://www.osti.gov/biblio/1044455
[91] A. Kaur, L. Nonnenmacher, H. T. C. Pedro, and C. F. M. Coimbra, “Benefits of solar forecasting for energy imbalance markets,” *Renew. Energy*, vol. 86, pp. 819–830, Feb. 2016, doi: 10.1016/j.renene.2015.09.011.
[92] D. Yang, “A universal benchmarking method for probabilistic solar irradiance forecasting,” *Sol. Energy*, vol. 184, pp. 410–416, May 2019, doi: 10.1016/j.solener.2019.04.018.
[93] M. Pierro, R. Perez, M. Perez, D. Moser, and C. Cornaro, “Imbalance mitigation strategy via flexible PV ancillary services: The Italian case study,” *Renew. Energy*, vol. 179, pp. 1694–1705, Dec. 2021, doi: 10.1016/j.renene.2021.07.074.
[94] R. Perez *et al.*, “From Firm Solar Power Forecasts to Firm Solar Power Generation an Effective Path to Ultra-High Renewable Penetration a New York Case Study,” *Energies*, vol. 13, no. 17, p. 4489, Aug. 2020, doi: 10.3390/en13174489.
[95] T. Lambert, “HOMER® Energy Modeling Software 2003,” National Renewable Energy Lab. (NREL), Golden, CO (United States), HOMER® 2003; 002642IBMPC03, Dec. 2003. Accessed: Oct. 19, 2022. [Online]. Available: https://www.osti.gov/biblio/1231444
[96] S. Kumar, P. Upadhyaya, and A. Kumar, “Performance Analysis of Solar Energy Harnessing System Using Homer Energy Software and PV Syst Software,” in *2019 2nd International Conference on Power Energy, Environment and Intelligent Control (PEEIC)*, Oct. 2019, pp. 156–159. doi: 10.1109/PEEIC47157.2019.8976665.

[97] M. Perez, R. Perez, K. R. Rábago, and M. Putnam, “Overbuilding & curtailment: The cost-effective enablers of firm PV generation,” *Sol. Energy*, vol. 180, pp. 412–422, Mar. 2019, doi: 10.1016/j.solener.2018.12.074.

RÉSUMÉ

Ces travaux de thèse portent sur l'intégration de l'énergie solaire dans les systèmes d'énergie hybride (HES) isolés sans stockage. Dans ce type de système qui est composé par des groupes électrogènes diesel et système Photovoltaïque (PV), la variabilité solaire doit être considérée car il pose des problèmes de la stabilité du réseau électrique. L'utilisation de systèmes de stockage d'énergie (ESS) est couramment recommandée pour atténuer cette variabilité. Cependant, l'utilisation des stockages venant toujours avec des contraintes telles que des coûts d'investissement élevés, des impacts environnementaux supplémentaires et des problèmes de maintenance potentiels. Le concept d'un dédié Power Management System (PMS) avec une prévision solaire court-terme est proposé pour réduire ou même remplacer l'utilisation des ESS. Cette thèse vise à explorer la faisabilité de ce concept et les gains potentiels qu'il pourrait apporter.

Pour évaluer la performance économique du HES, nous avons mis en place un simulateur numérique basée sur une approche de coût, qui évalue la performance du HES en coût, où le coût final du système est utilisé comme indicateur de performance. Ce simulateur dispose de plusieurs paramètres configurables pour différents scénarios, tels que le taux de pénétration du PV, le temps de mise à jour des dispatchs des génératrices, l'horizon de prévision, différentes approches de dimensionnement pour la réserve tournante et différentes méthodes de prévision. Étant donné que le coût du système est pertinent pour la qualité de la prévision et la variabilité solaire, nous explorons également la relation entre la performance statistique de la prévision et la performance économique du système. Enfin, les résultats obtenus montrent que les métriques classiques ne sont pas suffisantes pour représenter correctement la performance économique. Par conséquent, l'utilisation d'un outil de simulation comme le nôtre est peut-être une option plus directe et intuitive pour évaluer l'impact de la qualité de la prévision sur la performance économique du système.

MOTS CLÉS

Intégration PV, Unit Commitment, Système d'Energie Hybride, Gestion de puissance

ABSTRACT

When using solar energy in a hybrid energy system (HES), its temporal variability must be considered to ensure grid stability. In an off-grid HES which is composed of diesel generators and PV systems, energy storage systems (ESS) are commonly used to mitigate the variability. However, in large-scale off-gird HESs, the benefits stemming from ESS are accompanied by some drawbacks such as additional investment costs, extra environmental impacts, maintenance issues, etc. The use of a Power Management System (PMS) with short-term probabilistic forecasts may reduce or eliminate the need for an ESS. However, while numerous studies exist on solar forecasting and algorithms for Unit commit (UC) problems, there are few work studies on a concrete problem for this kind of insular storage-less HES. Hence, this thesis aims to find out the feasibility of this concept and the potential benefits of using a PMS with solar forecasting.

To evaluate the HES economic performance, we developed a simulation platform based on a cost-based approach, which assesses the HES performance in cost, and the final system cost is used as the performance indicator. This simulation platform has several configurable parameters for different scenarios, including the PV penetration rate, the generator dispatch update time, the forecast horizon, different sizing approaches for spinning reserve and different forecast methods. As the effective system cost is linked to both the forecast method quality and the solar variability, we examine the relationship between the statistical performance of forecasts and the economic performance of the system. Finally, the obtained results show that the classic statistical metrics are inadequate to correctly reflect the practical economic performance. Therefore, using a cost-based simulation tool maybe is a more direct and intuitive option to evaluate the influence of forecast method quality on final system economic performance.

KEYWORDS

PV integration, Unit Commitment, Hybrid Energy System, Power Management System