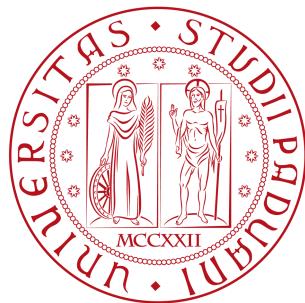


Università degli Studi di Padova

Dipartimento di Scienze Statistiche

Corso di Laurea Magistrale in
Scienze statistiche



La *Duck Curve* e la scomposizione della stagionalità dei consumi elettrici. Aspetti teorici e modelli statistici

Relatore: prof.ssa Mariangela Guidolin

Dipartimento di Scienze Statistiche, UNIPD

Correlatore: prof. Luigi Grossi

Dipartimento di Ingegneria e Architettura, UNIPR

Laureanda: **Bisetto Francesca**

Matricola: **2056766**

Anno Accademico 2022/2023

Indice

Introduzione	1
1 I consumi di energia elettrica in Italia	5
1.1 Il mercato energetico italiano	5
1.1.1 Le zone del mercato elettrico italiano	7
1.1.2 Gli attori del mercato energetico	8
1.2 Andamento del consumo di energia in Italia	9
1.2.1 I dati di Terna	9
1.2.2 Valutazioni generali	12
1.3 Le fonti rinnovabili: Energia solare	15
1.3.1 La diffusione dell'energia solare in Italia	17
1.4 La <i>Duck Curve</i>	18
1.4.1 La <i>Duck Curve</i> in California	21
1.4.2 La <i>Duck Curve</i> in Italia	22
1.4.3 Soluzioni al problema della <i>Duck Curve</i>	25
2 Aspetti metodologici	27
2.1 L'indice di Gini per valutare la stagionalità	28
2.2 Scomposizione della variazione di concentrazione stagionale	29
2.3 Test sul cambiamento della variazione stagionale	33

2.3.1	Test di significatività nelle due componenti	35
3	Analisi empirica	39
3.1	Applicazione ai dati giornalieri	39
3.1.1	Confronto con giorno medio in ogni anno	43
3.1.2	Confronto con giorno medio in ogni mese	45
3.1.3	Confronto con giorno medio feriale e nel fine settimana . . .	46
3.1.4	Confronto con giorno della settimana	48
3.2	Applicazione ai dati mensili	50
3.3	Modelli previsivi per ΔG	52
	Conclusioni	57

Introduzione

In un mondo fortemente dipendente dalla produzione di energia elettrica, le società di servizi pubblici sono migliorate nell'utilizzo dei dati per anticipare la domanda, cercando di operare nel modo più efficiente possibile. Nel panorama energetico contemporaneo, la transizione verso fonti di energia più sostenibili è diventata una priorità globale. Tra le fonti alternative, l'energia solare è emersa come una delle risorse più promettenti per ridurre l'impatto ambientale e garantire una fornitura energetica affidabile; si tratta infatti di una fonte sempre disponibile e facilmente prevedibile in quanto dipende dal ciclo dell'attività solare. Tuttavia, mentre l'energia solare contribuisce in modo significativo ad una produzione energetica pulita, essa presenta sfide uniche per la gestione delle reti elettriche. Uno dei fenomeni più intriganti e complessi in questo contesto è quello della *Duck Curve*.

La *Duck Curve* è una rappresentazione grafica della variazione del consumo di energia elettrica durante la giornata, che emerge quando si integra l'energia solare nelle reti elettriche tradizionali. Il suo nome deriva dalla somiglianza del grafico della domanda di energia elettrica residua rispetto a una curva che ricorda il corpo di una papera. Questo grafico rivela come la produzione solare raggiunga il suo apice durante le ore diurne, quando il sole è al massimo della sua potenza, mentre la domanda energetica presenta picchi tipicamente al tramonto, quando l'illuminazione artificiale diventa predominante. La *Duck Curve* rappresenta una sfida unica per gli operatori di rete elettrica e gli esperti di pianificazione energetica. Gestire

questa variazione giornaliera nella domanda e nella produzione energetica richiede una pianificazione avanzata e un'accurata previsione del consumo energetico netto, ovvero il consumo al netto dell'energia solare prodotta. La ricerca in questo campo è fondamentale per garantire l'efficienza delle reti elettriche, ridurre gli sprechi e ottimizzare l'integrazione dell'energia solare nella rete elettrica.

Le analisi e gli studi che vengono svolti nell'ambito della presente tesi si riferiscono ai consumi netti orari nel territorio italiano, nel periodo che va da gennaio 2019 a luglio 2023. La differenza tra l'energia consumata e quella prodotta, in una determinata ora, prende il nome di *Duck Curve* ed è caratterizzata da un andamento non regolare che rappresenta un problema per i gestori della rete. Questa tesi si propone di esplorare il fenomeno della *Duck Curve* utilizzando una scomposizione della curva che si basa sull'indice di Gini, per poi prevedere il consumo energetico netto.

L'elaborato è suddiviso in capitoli.

Nel primo capitolo si presenta una panoramica generale sul sistema elettrico italiano e la generazione di energia da fonti rinnovabili. Si introduce l'insieme di dati fornito da Terna S.p.A. utilizzato per le analisi, ponendo l'attenzione su alcune prime valutazioni generali.

Nel secondo capitolo vengono introdotti gli aspetti metodologici utili per la scomposizione dell'indice di Gini, proposti da Grossi & Mussini (2021), utili per valutare la stagionalità della *Duck Curve* ed i relativi test statistici.

Nel terzo capitolo vengono applicate queste metodologie ai dati, utilizzando dati a frequenza giornaliera e mensile. Il metodo utilizzato viene testato sui dati scegliendo di volta in volta una curva media di riferimento differente, che serve come misura riassuntiva del fenomeno. È stato poi adattato un modello che considera la stagionalità intrinseca nei dati, con l'obiettivo di prevedere nel breve termine il valore relativo all'indice di variazione della componente stagionale.

In questo contesto, la tesi rappresenta un contributo significativo per comprendere e affrontare le sfide connesse all'integrazione dell'energia solare nelle reti elettriche, sottolineando l'importanza di previsioni accurate del consumo energetico netto. La *Duck Curve* potrebbe essere un ostacolo, ma offre anche opportunità uniche per una migliore gestione dell'energia e una maggiore sostenibilità nel settore energetico.

Capitolo 1

I consumi di energia elettrica in Italia

L'energia elettrica rappresenta la linfa vitale della società moderna, alimentando una vasta gamma di attività, dall'illuminazione delle case all'industria, dai servizi pubblici alla crescente domanda di carica per i veicoli elettrici. L'analisi dei consumi rivela un quadro complesso ed in costante evoluzione, con implicazioni significative per l'ambiente, l'economia e la politica energetica; per questo motivo è importante analizzare il comportamento dei consumatori e cercare di prevederne l'evoluzione.

1.1 Il mercato energetico italiano

La filiera elettrica è un processo composto da una serie di fasi che intercorrono tra la produzione ed il consumo finale. Nel momento in cui si accende una lampadina, ad esempio, ci si trova nell'ultima fase. I passaggi intermedi variano a seconda che il periodo di riferimento sia prima o dopo la *liberalizzazione del mercato dell'energia*. A partire dal 1999 infatti il Decreto Bersani ha consentito il passaggio alla privatizzazione energetica italiana, prima di allora era sotto il monopolio statale; per cui

ogni cittadino è libero di scegliere il proprio fornitore di energia elettrica. La filiera del mercato energetico italiano è un sistema articolato e complesso, le principali fasi che la compongono sono: produzione, trasmissione e dispacciamento, distribuzione ed infine l'utilizzo. L'energia elettrica viene generata da migliaia di produttori, ovvero dalle centrali elettriche sparse in tutto il territorio nazionale. Questi si occupano della fase di generazione trasformando le fonti rinnovabili o fossili presenti in natura in energia elettrica; esistono varie tipologie di centrali in base alle fonti di produzione. Questa viene quindi trasportata sulla rete nazionale ad alta tensione gestita da Terna S.p.A. che opera in monopolio per garantire la sicurezza e l'efficienza del sistema. Un aspetto critico di questa fase riguarda il bilanciamento dei flussi di energia elettrica attraverso la rete che prende il nome di *dispacciamento* e consente di garantire il corretto equilibrio tra domanda ed offerta. Il dispacciamento svolge una funzione fondamentale in quanto l'energia elettrica non può essere immagazzinata, quindi è necessario che venga prodotta e consumata continuamente. Terna si occupa di questa funzione garantendo il costante equilibrio tra l'energia prodotta nei vari impianti e quella consumata da parte degli utilizzatori finali. La distribuzione dell'energia generata è affidata ad altre società che in primo luogo la trasformano da alta a medio/bassa tensione e successivamente fanno in modo che possa essere usufruita dal cliente finale.

In Italia soltanto il 13% del fabbisogno annuale di energia elettrica è importato dall'estero, il restante 87% è soddisfatto dalla produzione nazionale. È importante dunque analizzare e prevedere il comportamento dei consumatori, in modo tale che il dispacciamento garantisca un approvvigionamento energetico affidabile, sostenibile ed efficiente. L'analisi dei dati relativi al consumo energetico è di fondamentale importanza per svariati motivi. Ad esempio, la previsione dei consumi consente alle autorità di pianificare l'offerta dell'energia in modo efficiente. D'altro canto, la previsione dei picchi di consumo consente una migliore gestione della rete. In

questo elaborato si analizzano i consumi elettrici con lo scopo di monitorare l'uso delle fonti rinnovabili e valutare l'impatto ambientale della generazione di energia elettrica.

1.1.1 Le zone del mercato elettrico italiano

Il sistema energetico italiano viene suddiviso in aree dove consumatori e produttori possono vendere ed acquistare energia elettrica liberamente, mentre tra le diverse zone esistono delle limitazioni sulla compravendita. L'Italia è uno dei pochi paesi europei in cui esiste questo tipo di suddivisione. L'obiettivo è di cogliere correttamente l'andamento dei flussi a seconda delle condizioni di domanda ed offerta. Il sistema attualmente in uso è stato ideato per differenziare i prezzi di acquisto a seconda del bilancio tra capacità di generazione di energia e domanda che varia da zona a zona. Questo aspetto riguarda principalmente i produttori, infatti per i consumatori il costo è fissato attraverso il prezzo unico nazionale (PNU).



Figura 1.1: A sinistra le zone di mercato in vigore fino al 31 dicembre 2020, a destra la nuova configurazione zonale (Fonte: Terna S.p.A.).

A partire dal 1 gennaio 2021 è stata cambiata la configurazione delle zone di mercato italiane, come si può vedere in Figura 1.1, con lo scopo di massimizzare

l'efficienza del mercato, permettendo agli operatori di ottimizzare le contrattazioni evitando problemi di sicurezza nella rete. Le nuove zone sono quindi: Nord, Centro-Nord, Centro-Sud, Calabria, Sicilia, Sardegna. Ognuna è pensata per essere interconnessa con quelle dei paesi europei vicini.

1.1.2 Gli attori del mercato energetico

Il processo di liberalizzazione dei mercati energetici costituisce uno strumento vantaggioso per i consumatori finali. Essi infatti possono scegliere liberamente il fornitore di energia elettrica sia nel mercato tutelato, che in quello libero. Nonostante il mercato sia liberalizzato, i meccanismi che ne regolano il funzionamento sono stabiliti da enti, ministeri ed autorità che collaborano tra loro al fine di rendere il meccanismo più efficiente. I soggetti istituzionali coinvolti sono:

- **Ministero dell'Economia e delle Finanze (MEF).** Ha il compito di promuovere lo sviluppo dei mercati all'ingrosso e al dettaglio dell'energia. Ha sotto il suo controllo anche tutti gli altri enti che verranno descritti in seguito.
- **Acquirente Unico (AU).** Ha il compito di assicurare la fornitura di energia elettrica ai clienti vincolati, a prezzi competitivi e in condizioni di continuità. Svolge un'attività di approvvigionamento di energia per quei clienti che non hanno ancora scelto un fornitore sul mercato libero.
- **Gestore del mercato energetico (GME).** Gestisce ed organizza i mercati di energia elettrica e gas. Ha il ruolo di gestione ed organizzazione del mercato elettrico rispetto ai principi di neutralità, trasparenza e concorrenzialità.
- **Rete elettrica nazionale (TERNA).** Società responsabile della trasmissione e del dispacciamento dell'energia elettrica sulla rete ad alta tensione di tutto il territorio nazionale.

1.2 Andamento del consumo di energia in Italia

L'acquisizione dei dati ed il successivo processamento sono delle tecniche sempre più utilizzate negli ultimi anni. Anche nel settore energetico vengono continuamente raccolte informazioni relativamente al consumo ed alla produzione di energia. L'analisi dei dati in questo contesto risulta essere molto utile, la produzione oramai non è più totalmente centralizzata e basata sulle risorse di origine fossile ma sta diventando quasi completamente distribuita; il numero degli impianti rinnovabili presenti nel territorio nazionale crescono di anno in anno. L'avanzamento della transizione energetica mette gli addetti ai lavori davanti ad un duplice compito, da una parte continuare ad assicurare il corretto funzionamento della rete e dall'altra favorire l'integrazione con le nuove fonti. L'analisi dei *Big Data* viene in aiuto per poter comprendere il comportamento di queste fonti d'energia, legate sia alla disponibilità delle materie prime ma anche all'elevata volatilità dei prezzi delle fonti fossili.

1.2.1 I dati di Terna

Terna S.p.A. è l'azienda italiana che si occupa di attività di pianificazione, sviluppo e manutenzione della rete per gestire al meglio la trasmissione di energia elettrica in alta tensione. Un altro compito è quello di assicurare il dispacciamento, ovvero tutte quelle attività necessarie al fine di mantenere l'equilibrio tra domanda ed offerta di elettricità. Uno degli obiettivi che l'azienda si pone è l'integrazione tra le principali reti elettriche d'Europa, così da creare un sistema continentale sicuro ed efficiente.

Il ruolo che Terna ricopre in Italia è quello di registro naturale, ovvero tiene traccia di tutta la transizione energetica integrando anche le fonti rinnovabili. Essa raccoglie e consente di analizzare i dati relativi alla produzione, al consumo, trasmissione e distribuzione di energia elettrica. L'azienda raccoglie i dati a partire

#	Date	Generation	Load	Delta	Hour	Season	Mese	Anno	Giorno	Perc
1	2019-01-01	0.00	24370.00	24370.00	00:00:00	Winter	gen	2019.00	martedì	0.00
2	2019-01-01	0.00	23599.00	23599.00	01:00:00	Winter	gen	2019.00	martedì	0.00
3	2019-01-01	0.00	22501.00	22501.00	02:00:00	Winter	gen	2019.00	martedì	0.00
4	2019-01-01	0.00	21038.00	21038.00	03:00:00	Winter	gen	2019.00	martedì	0.00
5	2019-01-01	0.00	19871.00	19871.00	04:00:00	Winter	gen	2019.00	martedì	0.00
6	2019-01-01	0.00	19791.00	19791.00	05:00:00	Winter	gen	2019.00	martedì	0.00
7	2019-01-01	0.00	20348.00	20348.00	06:00:00	Winter	gen	2019.00	martedì	0.00
8	2019-01-01	40.00	21357.00	21317.00	07:00:00	Winter	gen	2019.00	martedì	0.19
9	2019-01-01	950.00	21027.00	20077.00	08:00:00	Winter	gen	2019.00	martedì	4.52
10	2019-01-01	2510.00	22547.00	20037.00	09:00:00	Winter	gen	2019.00	martedì	11.13
11	2019-01-01	4140.00	24092.00	19952.00	10:00:00	Winter	gen	2019.00	martedì	17.18
12	2019-01-01	4720.00	25050.00	20330.00	11:00:00	Winter	gen	2019.00	martedì	18.84
13	2019-01-01	4920.00	26536.00	21616.00	12:00:00	Winter	gen	2019.00	martedì	18.54
14	2019-01-01	4460.00	25709.00	21249.00	13:00:00	Winter	gen	2019.00	martedì	17.35
15	2019-01-01	3090.00	24190.00	21100.00	14:00:00	Winter	gen	2019.00	martedì	12.77
16	2019-01-01	1590.00	24264.00	22674.00	15:00:00	Winter	gen	2019.00	martedì	6.55

Figura 1.2: Prime osservazioni del dataset in esame.

da una grande varietà di fonti, tra cui i dispositivi di misurazione installati nella rete elettrica, dati provenienti dalle centrali elettriche, le informazioni sul mercato dell'energia e i dati meteo. Una volta raccolti, questi vengono elaborati, archiviati e resi disponibili al pubblico.

Nelle analisi svolte in questo elaborato si utilizza un *dataset* prodotto a partire dai dati aventi la struttura mostrata in Figura 1.2. Il periodo analizzato va da gennaio 2019 a giugno 2023; le osservazioni che lo compongono si riferiscono al fabbisogno orario di energia elettrica nel territorio nazionale e alla produzione di energia fotovoltaica.

Nella tabella in Figura 1.2 vengono mostrate le prime osservazioni del *dataset*, in particolare le variabili utilizzate sono:

- *Date*: data in cui è stata effettuata la rilevazione.
- *Generation*: produzione di energia da fonte fotovoltaica nel territorio italiano. È stata trasformata in Megawatt da Gigawatt (*1000).
- *Load*: quantità di energia elettrica, misurata in Megawatt, necessaria al fine di soddisfare il fabbisogno della rete. I dati forniti da Terna presentano una rilevazione ogni quarto d'ora, al fine di renderli confrontabili con quelli della produzione fotovoltaica, sono state selezionate solo le osservazioni relative al minuto 00.
- *Delta*: *Load-Generation*. Carico elettrico netto.
- *Hour*: orario in cui è stata effettuata la misurazione.
- *Season*: stagione della misurazione.
- *Mese*: mese della misurazione.
- *Anno*: anno della misurazione.

- *Giorno*: giorno della settimana della misurazione.
- *Perc*: $\text{Generation} / \text{Load} * 100$. Prende il nome di *Frazione di copertura* e rappresenta la percentuale di energia richiesta dal carico totale che viene soddisfatta dall'energia generata dal sistema fotovoltaico. Indicatore importante per valutare quanto l'energia solare stia contribuendo a coprire il fabbisogno energetico.

1.2.2 Valutazioni generali

A partire da questi dati sono state effettuate alcune prime analisi per monitorare come sono cambiati i consumi e la produzione di elettricità durante i vari anni presi in esame. Questo tipo di dati è caratterizzato da una forte stagionalità, sia all'interno della singola giornata, sia nei vari anni.

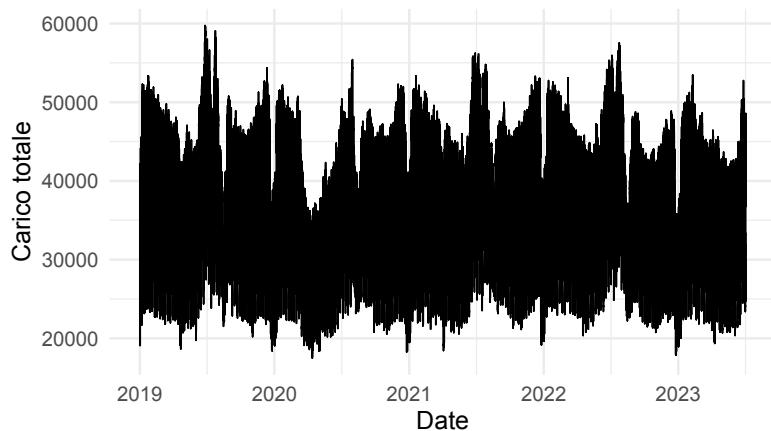


Figura 1.3: Serie storica giornaliera del carico totale di energia dal 2019 al 2023 in Italia.

Inizialmente si vuole valutare l'andamento generale a livello grafico della serie data del *carico totale*, ovvero la quantità complessiva di energia richiesta o consumata (Figura 1.3). Il *carico totale* tiene in considerazione tutte le richieste di

energia elettrica, sia da un punto di vista domestico, sia industriale o commerciale. Dal grafico in Figura 1.3 si può facilmente notare come, durante la pandemia di COVID-19 (da marzo a maggio 2020 circa), la richiesta di energia elettrica ha subito una forte variazione a causa degli effetti della pandemia stessa e alla misure adottate per contenerla. Tradizionalmente i comportamenti sociali e produttivi si riflettono sul sistema elettrico nazionale e questo vale anche per le misure messe in atto con lo scopo di contenere l'epidemia. Il 13 aprile 2020 è stato registrato il minor consumo di energia: è stato riscontrata una diminuzione del 45% rispetto allo stesso giorno dell'anno precedente. Queste conclusioni sono state tratte senza considerare altri fenomeni che avrebbero potuto influenzare la diminuzione della domanda, quali, ad esempio, le condizioni meteorologiche di quel periodo. Tuttavia, l'emergenza sanitaria e le procedure implementate per la gestione hanno avuto un impatto importante sulle abitudini e sulle attività sociali e produttive di tutto il territorio nazionale.

Analizzando più nello specifico la serie, si vuole studiare il comportamento dei consumatori. Nei grafici in Figura 1.4, studiando in particolare l'andamento orario durante ogni giornata (Figura 1.4(a)) si nota una forte componente stagionale. La domanda di energia elettrica inizia ad aumentare a partire dalle 6.00, ovvero quando le persone iniziano ad alzarsi la mattina, raggiunge poi il suo picco massimo durante le ore diurne fino alle 20.00, per poi decrescere durante la notte, quando la maggior parte delle persone riposa. Questo picco è causato dalla combinazione di consumi residenziali, commerciali ed industriali, tra i quali il riscaldamento, l'illuminazione ed il funzionamento dei macchinari nelle fabbriche. Ponendo l'attenzione ora sull'andamento giornaliero (Figura 1.4(b)), si nota come la richiesta vari notevolmente a seconda che si tratti di una giornata lavorativa rispetto al fine settimana. Nei giorni lavorativi infatti la domanda tende ad essere più alta a causa delle attività industriali e commerciali che richiedono energia per alimentare

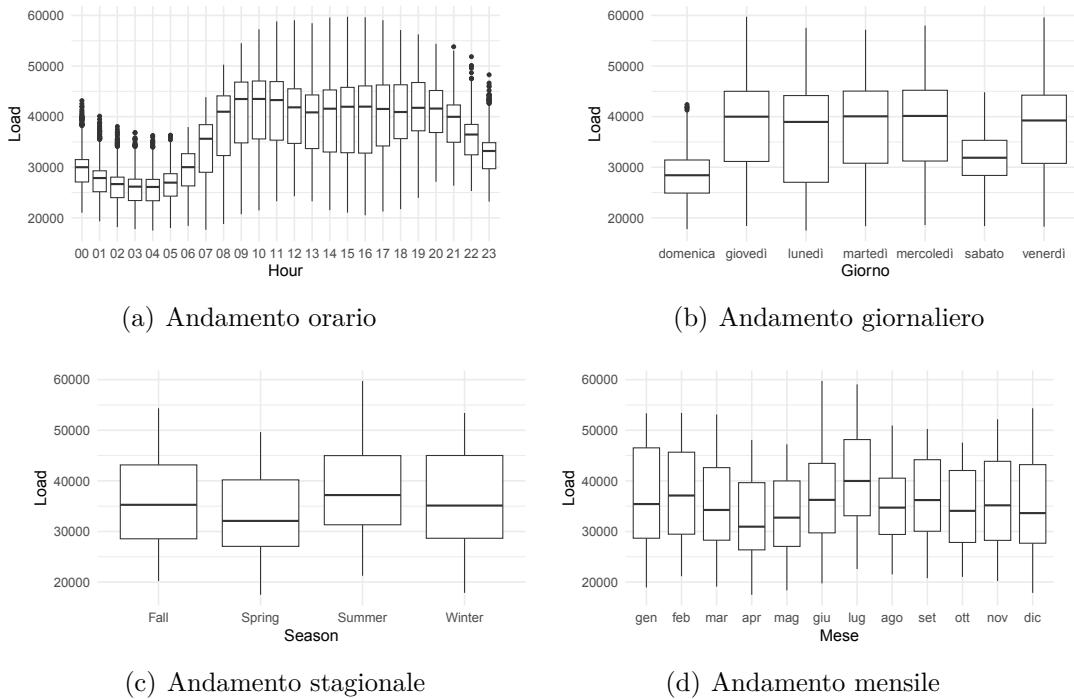


Figura 1.4: Andamento nel dettaglio del *carico totale*.

le fabbriche, gli uffici e le attrezzature. Nel fine settimana invece si osserva una tendenza differente in quanto molte aziende sono chiuse oppure operano a carico ridotto. Relativamente alle variazioni stagionali (Figura 1.4(c)), si nota come in primavera la domanda di energia sia inferiore rispetto agli altri mesi, questo può essere spiegato dal fatto che durante quei mesi la produzione di energia proveniente da fonti rinnovabili, in particolare con i sistemi fotovoltaici, raggiunge il suo massimo. In estate ed inverno invece l'uso dei sistemi di raffreddamento e riscaldamento possono aumentare la richiesta. Valutando ora il comportamento dei consumatori nei vari mesi dell'anno (Figura 1.4(d)) si osserva anche in questo caso una differenza sostanziale tra i periodi. La situazione che si osserva è simile al grafico precedente, infatti nei mesi primaverili ed autunnali la domanda di energia diminuisce in quanto la temperatura esterna è più mite rispetto agli altri mesi.

1.3 Le fonti rinnovabili: Energia solare

L'energia solare è la principale fonte di energia rinnovabile presente in natura. Essa è inesauribile in quanto il nostro pianeta è costantemente irradiato dai raggi solari ed è pulita al 100%, non comporta emissioni di anidride carbonica o polveri sottili. Essendo una fonte molto importante può essere impiegata in vari modi quali la produzione di energia elettrica, tramite pannelli fotovoltaici e termica, con pannelli solari. Negli ultimi anni lo sviluppo tecnologico degli impianti solari ha raggiunto livelli importanti, rispetto alle altre fonti rinnovabili, questa in particolare cresce in maniera proporzionalmente maggiore consentendo di creare centrali sempre più efficienti.

Il solare fotovoltaico. L'energia fotovoltaica sfrutta i vantaggi della fonte solare attraverso dei pannelli fotovoltaici che assorbono la luce del sole e la trasformano in energia. Il processo avviene all'interno della cella fotovoltaica, costituita da un materiale semiconduttore, il silicio. Si sfrutta quindi l'irraggiamento solare con l'obiettivo di produrre energia elettrica.

L'*effetto fotovoltaico* si manifesta nel momento in cui i raggi del sole colpiscono la cella fotovoltaica e parte dell'energia contenuta nella luce catturata viene trasformata in elettricità. Ogni cella fotovoltaica converte l'energia della radiazione solare in energia elettrica in corrente continua. In un impianto sono presenti diversi pannelli collegati tra loro che generano una potenza complessiva pari alla somma dei singoli pannelli. La potenza di un pannello è importante in quanto indica quanta energia è in grado di produrre, ad oggi si aggira attorno ai 300-400 Wp (potenza di picco).

L'utilizzo di questa fonte rinnovabile è un'alternativa ecosostenibile ad altre forme di energia, ne permette infatti la produzione a basso impatto ambientale, senza l'impiego di combustibili fossili. Tuttavia, al momento, i rendimenti degli impianti

ti sono bassi, si riesce a produrre circa un terzo dell'energia che teoricamente si potrebbe produrre. Al momento, il più grande problema riscontrato legato all'utilizzo dell'energia solare riguarda la discontinuità di produzione, non sempre ci sono giornate soleggiate e durante le ore notturne non vi è alcuna produzione. Sono disponibili delle batterie di accumulo, ma non sono ancora in grado di garantire una copertura energetica abbastanza grande da poter risolvere questo problema. In una giornata di sole la produzione maggiore si otterrà nelle ore centrali, in cui l'irraggiamento solare è massimo. L'efficienza dei pannelli è influenzata da diversi fattori tra i quali l'orientamento e l'inclinazione con cui sono stati installati; così facendo è possibile sfruttare al meglio questa tecnologia. La maggior parte di essi offre un maggior rendimento ad una temperatura ambientale di 25°C, si ottiene una maggior efficienza in Italia durante la primavera e l'autunno.

Il solare termico. In questo caso l'energia solare viene catturata e trasformata in energia termica. Il calore ricavato può essere usato per il riscaldamento o per il raffreddamento dell'acqua. Si tratta di una tecnologia più semplice e meno costosa rispetto al fotovoltaico. Il suo funzionamento è suddiviso in tre fasi: il collettore termico assorbe la luce del sole tramite l'assorbitore, qui uno speciale fluido termovettore viene riscaldato, successivamente una pompa trasporta il fluido verso lo scambiatore di calore del bollitore solare, qui l'energia termica viene trasmessa ad un serbatoio di stoccaggio. Nel caso in cui la radiazione solare dovesse risultare insufficiente per riscaldare l'acqua, un sistema di riscaldamento convenzionale scalderà il serbatoio di stoccaggio alla temperatura desiderata. Oltre alla produzione di acqua calda sanitaria, il fluido riscaldato può essere utilizzato per integrare il sistema di riscaldamento. Questo sistema consente di risparmiare energia elettrica, ma non ne genera.

Al fine di massimizzare la produzione dei singoli impianti, è di interesse valutare come la produzione totale di energia elettrica generata dal sole in Italia dipenda dalle

variabili meteorologiche. Alla luce di quanto detto precedentemente si può intuire che tale produzione presenta una doppia stagionalità, legata alla disponibilità di radiazioni solari sia rispetto durante l'anno e all'interno della singola giornata.

1.3.1 La diffusione dell'energia solare in Italia

L'energia solare in Italia ha avuto un incremento significativo a partire dal XXI secolo, ottenendo la quinta posizione a livello mondiale. Tra il 2009 ed il 2013 si è assistito ad un aumento della capacità nominale fotovoltaica installata, raggiungendo alla fine del 2013 la i 17,928 MW. Grazie ai numerosi incentivi per la produzione di energia solare fotovoltaica, i cosiddetti Conto Energia, è stato possibile ottenere questi risultati nel paese. Tuttavia, dopo il 2013 la crescita della capacità si è arrestata a causa della cessazione dei programmi di sussidi governativi.

La distribuzione geografica del fotovoltaico italiano è piuttosto complessa. Non esiste infatti una netta distinzione tra le zone come può verificarsi per altri tipi di energie, quale ad esempio l'eolico. La diffusione può essere valutata relativamente a tre parametri differenti ovvero: il numero di impianti, la potenza nominale installata e la potenza effettiva raggiunta. Secondo i dati forniti nel 2020 dal Gestore di servizi energetici, GSE, Lombardia, Veneto ed Emilia Romagna sono le regioni con il maggior numero di impianti installati.

In Italia il traguardo dei 20mila megawatt di potenza fotovoltaica installata è stato raggiunto nel 2018. Analizzando gli ultimi rapporti statistici forniti da GSE, viene evidenziato un +16.3% di impianti fotovoltaici in funzione, rispetto alla fine del 2022, per una potenza complessiva di 27GW. Valutando la produzione lorda di energia nell'ultimo semestre del 2023 risulta essere stata pari a 15214GWh, aumentata del 3.4% rispetto allo stesso periodo dell'anno precedente. È possibile dunque affermare che la transizione verso le energie rinnovabili, in particolare a quella solare, sia un fenomeno in costante evoluzione. In questi ultimi anni sono

molteplici le imprese che stanno studiando tecnologie e materiali che consentano un incremento della producibilità.

Nella Figura 1.5 si può osservare l'incidenza del fotovoltaico sul consumo totale di energia, suddiviso nei cinque anni di osservazione. In particolare, per ogni mese, viene mostrato il carico totale richiesto dalla rete e quanto invece viene prodotto attraverso i pannelli fotovoltaici. Si può facilmente osservare come la quantità di energia richiesta mensilmente dalla rete sia notevolmente maggiore rispetto alla produzione attraverso i raggi solari. È necessario anche considerare che l'energia solare è una delle fonti rinnovabili disponibili in natura, ad essa si aggiungono anche l'idroelettrica, l'eolica e quella geotermica. A partire dal 2019 si nota come la generazione dell'energia in Italia sia in costante crescita, ma allo stesso tempo possa essere influenzata da vari fattori tra cui ad esempio le condizioni meteorologiche. Come si sottolineava già precedentemente durante i mesi primaverili ed estivi la produzione raggiunge il suo picco, per poi descrescere durante l'autunno e l'inverno. Si può affermare quindi che essa ha un impatto positivo sulla transizione verso un sistema energetico più sostenibile e pulito. Tuttavia, l'effettiva percentuale di energia solare rispetto al carico totale può variare notevolmente durante i mesi dell'anno per questo motivo è importante potenziare la rete con sistemi di accumulo in modo da massimizzare il contributo dato dall'energia solare.

1.4 La *Duck Curve*

Negli ultimi anni le energie rinnovabili si stanno diffondendo sempre di più sia presso le imprese sia presso i singoli cittadini. Sfruttando l'energia prodotta dal sole si cerca di contribuire alla creazione di un futuro più sostenibile. Tuttavia, questo elevato assorbimento di energia solare non è esente da sfide. Queste tecnologie infatti hanno ridotto la domanda da fonti non rinnovabili durante le ore diurne, perché viene

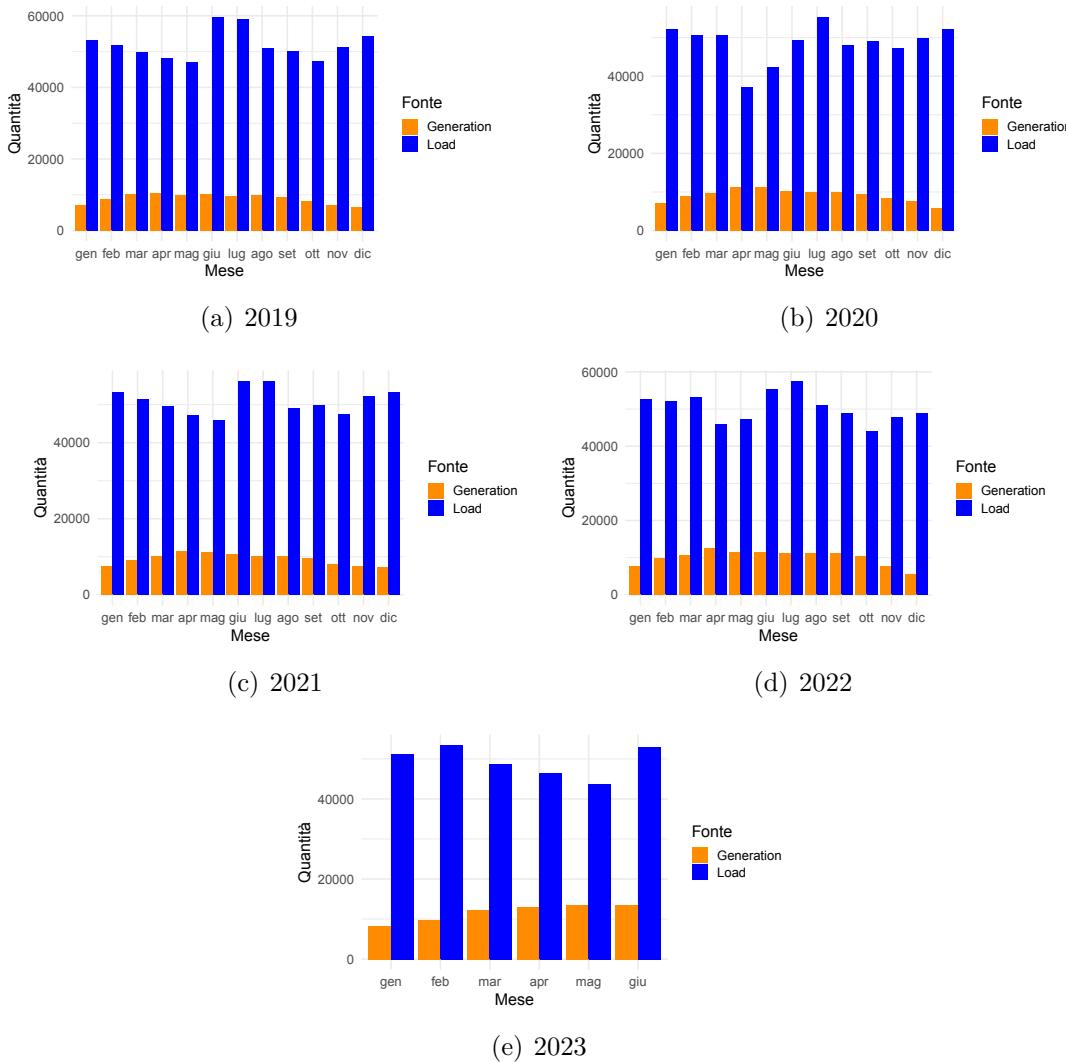


Figura 1.5: Incidenza del fotovoltaico sul consumo totale di energia nei 5 anni.

sfruttata l'energia prodotta dai sistemi solari fotovoltaici. Vi è quindi una maggiore tensione sulla rete poichè lavora per tenere il passo con le fluttuazioni della domanda e dell'offerta, che possono creare instabilità. L'introduzione dell'energia solare nelle nostre vite ha creato problemi nella stima dei modelli di curva di domanda. Le società di servizio pubblico infatti utilizzano modelli per prevedere la domanda così

da operare in modo efficiente garantendo energia nei periodi in cui la domanda è maggiore. Durante la giornata la richiesta di energia è bassa nelle ore notturne, aumenta di intensità nelle prime ore del giorno, quando le persone si svegliano e le aziende iniziano ad essere operative, e raggiunge un picco al tramonto. L'adozione su larga scala dell'energia solare ha portato a un notevole cambiamento nei modelli di consumo energetico. In particolare, la produzione di energia solare raggiunge il suo picco intorno a mezzogiorno, un momento in cui la domanda di elettricità è relativamente bassa. Di conseguenza, si verifica un eccesso di energia rispetto a quella richiesta, con una diminuzione della domanda netta. Tuttavia, alla sera, con la riduzione della produzione solare, la domanda netta tende ad aumentare.

La *Duck Curve*, letteralmente "curva a papera", è un termine utilizzato per descrivere questo modello di consumo e produzione di energia elettrica che si verifica in alcuni sistemi con una crescente adozione di energie rinnovabili. È una rappresentazione grafica della domanda di elettricità della rete nei giorni in cui la produzione di energia solare è elevata e la domanda è bassa. È un potenziale di instabilità del sistema energetico, in quanto, durante una singola giornata avvengono più cambiamenti nella domanda. Il fenomeno qui descritto si verifica principalmente per due ragioni principali. In primo luogo, riguarda la produzione di energia rinnovabile, specialmente durante le ore diurne, quando è possibile generare quantità significative di energia solare, a condizione che le condizioni meteorologiche siano favorevoli. In secondo luogo, questo fenomeno è influenzato dalla crescente domanda di energia nelle ore serali, quando il sole tramonta e la produzione di energia rinnovabile diminuisce. In altre parole, si tratta di un equilibrio complesso tra l'offerta di energia rinnovabile e la domanda energetica che varia nel corso della giornata.

1.4.1 La Duck Curve in California

Nel 2013, il California Independent System Operator (CAISO) ha pubblicato un grafico che metteva in evidenza l'impatto della generazione di energia tramite i pannelli fotovoltaici. Questa curva ha preso il nome di *Duck Curve* per la sua forma particolare: essa mostra la differenza tra la domanda di elettricità e la quantità di energia solare utilizzabile durante il giorno, come si può osservare nel grafico in Figura 1.6.

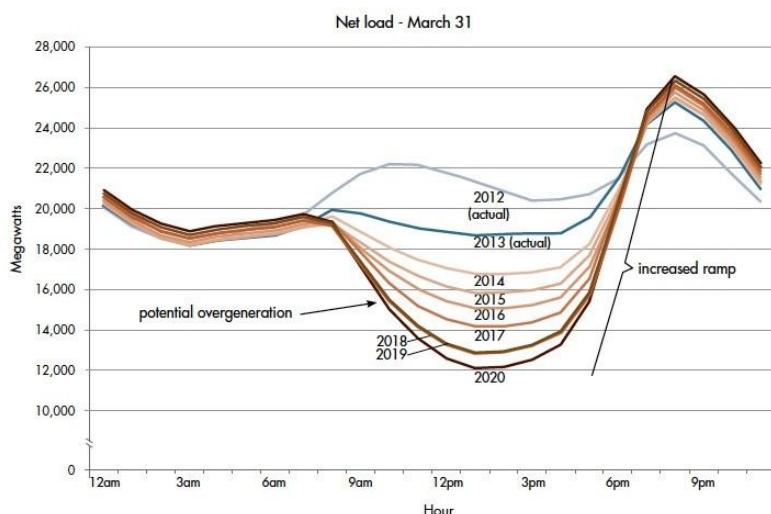


Figura 1.6: Grafico della *Duck Curve* in California dall'articolo (Jones-Albertus, 2017)

Questo fenomeno è stato studiato per la prima volta negli Stati Uniti, in particolare in California, paese in cui l'energia solare si è diffusa rapidamente a causa dell'elevata isolazione e della diminuzione dei costi per l'installazione. Il 2019 è considerato l'anno della svolta per gli Stati Uniti poichè il consumo di energia proveniente da fonti rinnovabili ha superato quello del carbone. Nel 2020 la domanda di elettricità è diminuita del 5% mentre la produzione da fonti rinnovabili è aumentata del 7% rispetto all'anno precedente. Alla fine del 2022 lo Stato vantava una capacità di energia solare installata che garantiva un'alimentazione per 10,5 milioni

di case. La curva rappresenta un punto di transizione per l'energia solare. È necessario ora pensare ad una pianificazione dell'energia considerando anche questa sempre più diffusa alternativa ecosostenibile. Man mano che l'adozione del solare cresce, la curva presenta una depressione nelle ore centrali della giornata sempre più accentuata. Questo è un grande problema per l'ente che gestisce il servizio in quanto rende più difficile bilanciare la rete. Il fenomeno risulta essere più intenso soprattutto in primavera ed estate: è stato necessario ricorrere ad una soluzione per evitare di spegnere una parte dei pannelli così da non sovraccaricare e danneggiare la rete elettrica. La *Duck Curve* in California sottolinea quanto sia cruciale gestire attentamente la produzione di energia solare e trovare soluzioni innovative per affrontare il problema del sovraccarico della rete elettrica durante le ore di punta serali. Questa rappresentazione visiva della sfida energetica è un richiamo per concentrarsi sull'implementazione di tecnologie di stoccaggio energetico avanzate, strategie intelligenti di gestione dell'energia e una pianificazione energetica mirata. Questi elementi sono essenziali per garantire una transizione verso un sistema energetico più sostenibile e sicuro.

1.4.2 La *Duck Curve* in Italia

Negli ultimi quattro anni in Italia si sta assistendo ad una progressiva riduzione del carico elettrico residuo, ovvero la quantità di energia elettrica richiesta dalla rete in un determinato momento, al netto delle produzioni provenienti da fonti rinnovabili, come l'energia solare.

Nel grafico in Figura 1.7 si può osservare come il profilo del carico residuo medio sembra notevolmente cambiato dal 2019 ad oggi, si può infatti notare un calo specialmente nelle ore diurne, coerente con il contributo dato dall'energia solare prodotta attraverso i pannelli fotovoltaici. La curva disegnata è data dalla differenza tra il carico totale di energia elettrica richiesta e la quantità generata attraverso

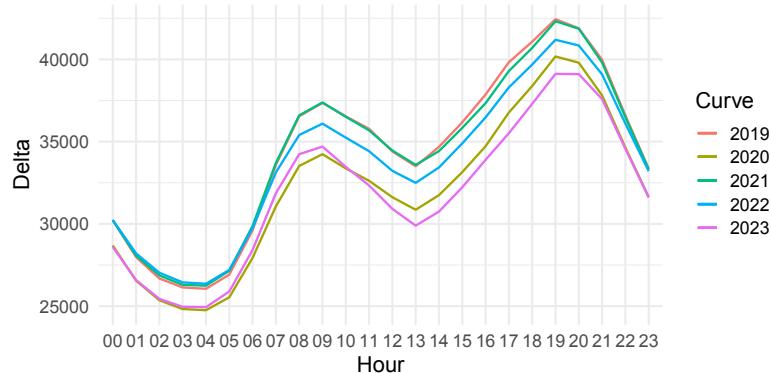


Figura 1.7: Profilo orario del carico residuo medio in Italia (*Duck Curve*)

i pannelli fotovoltaici. È possibile osservare come la forma assunta dalla curva assomigli a quella di un'anatra, da cui il nome di *Duck Curve*. Osservando il grafico in Figura 1.7 si può notare che la prepotente discesa e risalita del carico cominciano e terminano, rispettivamente, in corrispondenza dell'alba e del tramonto. Risulta lampante il legame tra il carico netto e l'energia solare: tra le ore 8 e le ore 14 circa la rete si sta rifornendo sempre di più dagli impianti fotovoltaici, abbassando quindi la richiesta di energia da combustibili fossili. La curva relativa all'anno 2020 può essere risultata leggermente diversa dalle altre in quanto i consumi risultano falsati a causa delle restrizioni imposte del governo durante la pandemia, gran parte delle attività produttive sono state chiuse per diversi mesi abbassando così la richiesta di energia.

Nei grafici in Figura 1.8 si è voluto mostrare la differenza tra i vari anni prendendo come giorno di riferimento il 6 maggio. È stato scelto a causa della produzione elevata di energia solare, rappresentata dalla curva blu. Come si può notare infatti negli anni precedenti la quantità di energia prodotta dai pannelli solari è risultata notevolmente inferiore. Si può anche osservare come durante il corso degli anni la curva verde, rappresentante il carico netto, assuma sempre più la forma della *duck curve*. Questo si osserva perché negli ultimi anni il governo offre sempre più incen-

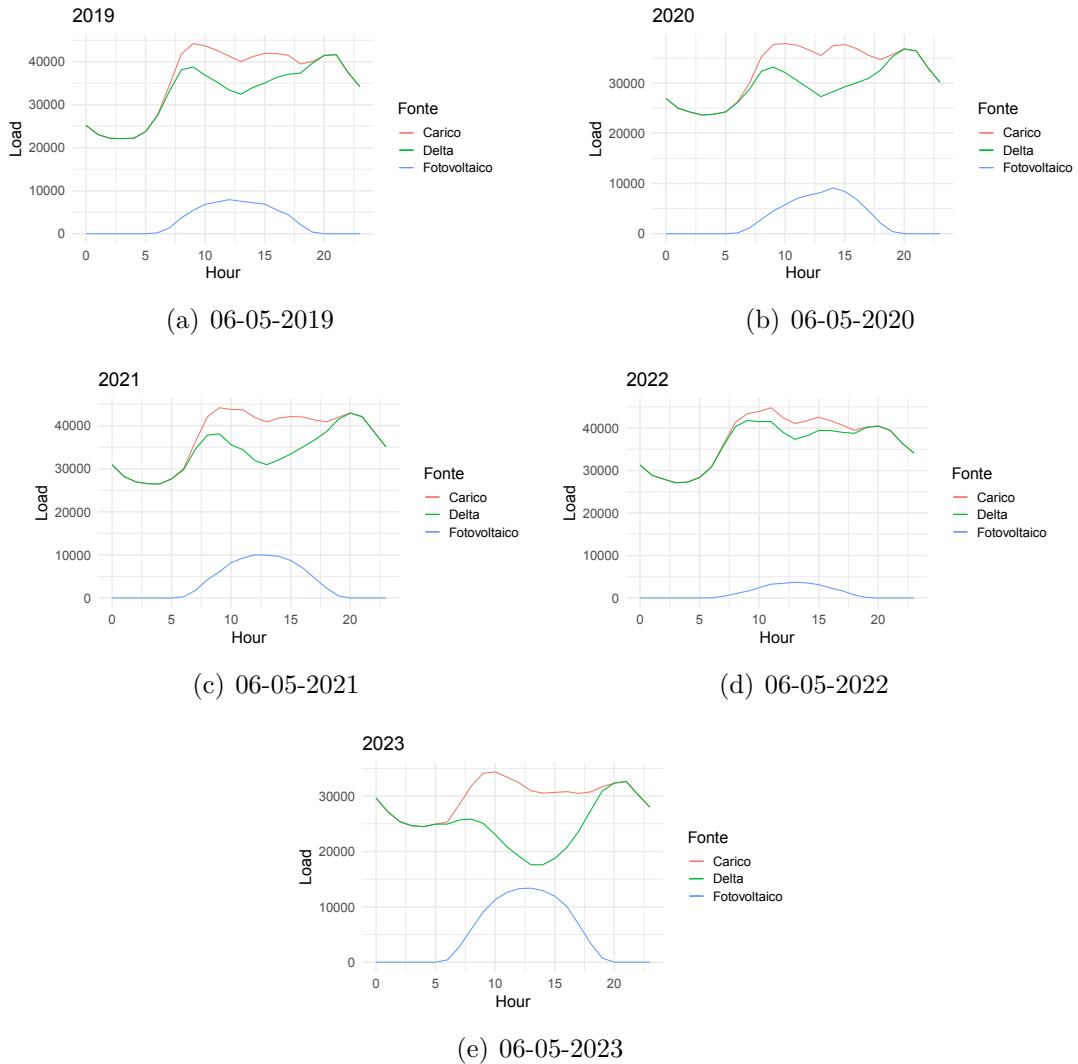


Figura 1.8: Curve di carico totale, carico netto e produzione fotovaltaica il 6 maggio dei cinque anni in esame.

tivi per l'installazione di impianti fotovoltaici, così da sfruttare l'energia solare che è una delle fonti più pulite di energia disponibile.

1.4.3 Soluzioni al problema della *Duck Curve*

Maggiore è la diffusione delle energie rinnovabili, più le questioni legate all'affidabilità e all'integrazione con le reti diventano importanti. L'energia solare è legata alla quantità di irraggiamento che è intermittente per sua natura. Non è quindi semplice conoscere quanta energia sarà disponibile in un dato momento. Quindi, la questione della *Duck Curve* è ancora studiata per trovarne una soluzione.

Uno degli sviluppi più recenti riguarda le batterie di stoccaggio energetico; in questo modo la sovrapproduzione durante le ore diurne può essere immagazzinata e riutilizzata nel momento in cui il sole tramonta. Da solo, lo stoccaggio dell'energia non può superare il problema. Al fine di migliorare la transizione energetica bisogna considerare anche alcune soluzioni dal lato della domanda come, ad esempio, la flessibilità del carico, ovvero la gestione dei flussi di energia agendo su produzione e consumo da parte delle aziende. Per ridurre i consumi di elettricità i miglioramenti vanno effettuati sull'isolamento degli edifici e migliorando gli impianti di illuminazione.

Capitolo 2

Aspetti metodologici

Nel seguente capitolo vengono presentati i metodi utilizzati per analizzare e prevedere la *Duck Curve*. È importante studiare il comportamento della curva al fine di ottimizzare i sistemi energetici in cui sono ampiamente presenti fonti di energia rinnovabile. L'approccio che viene sviluppato in questo elaborato si basa sull'indice di Gini ed in particolare su una sua scomposizione, in modo da valutare la variazione della stagionalità (Grossi et al., 2022). Le due componenti che si ottengono permettono di valutare il cambiamento nelle fluttuazioni stagionali e il cambiamento della loro ampiezza. Successivamente, viene implementata una tecnica che consente di testare la significatività delle componenti. La metodologia che verrà studiata ed implementata segue la trattazione teorica sviluppata da Grossi & Mussini (2021) nell'articolo *"Seasonality in tourism flow: Decomposition and testing changes in seasonal concentration"*, adattandola al caso della domanda oraria di energia al netto della produzione fotovoltaica. Utilizzando tali metodi si vuole analizzare la stagionalità presente nei dati. Verranno poi stimati alcuni modelli per serie storiche con l'obiettivo di prevedere il comportamento della serie nel breve termine.

2.1 L'indice di Gini per valutare la stagionalità

L'indice di Gini prende il nome da colui che per primo lo introdusse in letteratura (Corrado Gini) ed è una misura statistica ampiamente utilizzata per valutare la disuguaglianza economica o sociale all'interno di una popolazione. È comunemente utilizzato come indicatore di concentrazione per quantificare la disuguaglianza nella distribuzione del reddito e della ricchezza. Esso assume valori tra 0 e 1, dove 0 indica una situazione di equidistribuzione mentre 1 corrisponde alla massima concentrazione, ovvero massima disuguaglianza tra le distribuzioni. L'indice di Gini nasce per essere applicato in ambito sociale al fine di quantificare la disuguaglianza delle ricchezze. È una misura relativa di variabilità pari al doppio della distanza tra la curva della distribuzione effettiva della ricchezza (nota come curva di Lorenz) e la curva di equidistribuzione. A partire dal grafico in Figura 2.1, l'indice di Gini può essere calcolato come rapporto tra l'area delimitata dalla linea di equidistribuzione e la curva di Lorenz (A) divisa per l'area totale sottesa dalla linea di equidistribuzione (A+B). Esso consente di riassumere in un solo valore il grado di concentrazione della distribuzione del reddito.

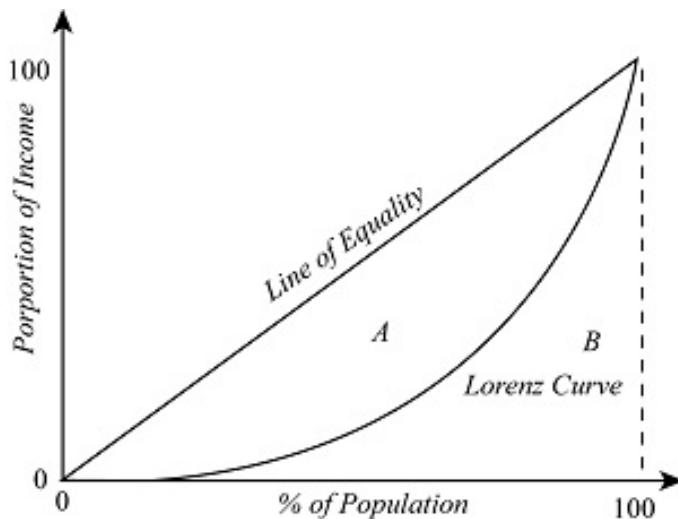


Figura 2.1: Curva di Lorenz

La concentrazione stagionale dei consumi energetici è un fenomeno che caratterizza i dati che vengono analizzati in questo elabotato. Si riferisce alle variazioni regolari che si verificano durante le stagioni dell'anno e possono essere influenzate da diversi fattori, tra cui le condizioni meteorologiche o le festività. Questo fenomeno porta ad uno squilibrio nella rete in quanto la domanda di energia non risulta costante nelle varie ore del giorno. La stagionalità dovrebbe essere misurata attraverso diversi indicatori di squilibrio temporale (Burtler, 1994). Un indicatore molto utilizzato per analizzare questo fenomeno è l'indice di Gini. In letteratura sono stati sviluppati studi che utilizzano in maniera alternativa il coefficiente di variazione (CV)¹ e la famiglia di indici di Theil (Fernández-Morales, 2003; Fernández-Morales et al., 2016; Duro, 2016; Lau et al., 2017; Rosselló and Sansó, 2017). I risultati che si ottengono confrontando le analisi svolte con i tre diversi indicatori possono essere molto diversi in quanto ognuno di essi pondera in modo diverso le distanze tra le osservazioni.

2.2 Scomposizione della variazione di concentrazione stagionale

Oltre alla misurazione globale della stagionalità, è utile ricorrere alla scomposizione degli indici. In questo elaborato si usa la scomposizione della variazione dell'indice di Gini proposto da Grossi e Mussini nel 2021, scomponendo la variazione dell'indice di Gini. Il valore della variazione dell'indice, calcolato in due istanti temporali differenti, può essere riscritto come somma di due diverse componenti, la prima che misura la stabilità stagionale (R) mentre la seconda (M) la grandezza delle fluttuazioni stagionali. Si assume che l'andamento delle fluttuazioni sia stabile

¹CV : $\sigma^* = \frac{\sigma}{|\mu|}$, σ è la deviazione standard e μ la media della popolazione.

negli anni.

Si indica con Y la variabile che rappresenta il consumo orario di energia elettrica al netto della produzione fotovoltaica, per ogni giornata dal 2019 al 2023. Sia $y_{1,t}, \dots, y_{n,t}$ la serie storica dei consumi nel periodo n (che va da 0 a 23) del giorno t ; dunque $y_{i,t}$ rappresenta i MW consumati nell'ora i del giorno t . $r(y_{i,t})$ indica il rango del periodo i secondo l'ordine crescente per MW consumati in uno specifico giorno dell'anno t . A partire da questi è possibile calcolare l'indice di Gini come funzione della covarianza tra $y_{i,t}$ e $r(y_{i,t})$ nel giorno t :

$$G_t = \frac{2 \operatorname{cov} [y_{i,t}, r(y_{i,t})]}{n \bar{y}_t} \quad (2.1)$$

con \bar{y}_t consumo giornaliero medio di energia elettrica.

La variazione delle concentrazione del consumo di energia dal giorno t al giorno $t+k$ è misurata come la differenza tra l'indice di Gini in $t+k$ e l'indice di Gini in t :

$$\Delta G = G_{t+k} - G_t = \frac{2 \operatorname{cov} [y_{i,t+k}, r(y_{i,t+k})]}{n \bar{y}_{t+k}} - \frac{2 \operatorname{cov} [y_{i,t}, r(y_{i,t})]}{n \bar{y}_t} \quad (2.2)$$

Sia $C_{t+k|t}$ il coefficiente di concentrazione calcolato nel giorno $t+k$ ordinando i periodi in base alla loro posizione in t

$$C_{t+k|t} = \frac{2 \operatorname{cov} [y_{i,t+k}, r(y_{i,t})]}{n \bar{y}_{t+k}} \quad (2.3)$$

A questo punto si può ottenere una scomposizione della differenza nell'indice di Gini aggiungendo e sottraendo $C_{t+k|t}$ nell'equazione (2.2).

$$\Delta G = \left(\frac{2 \operatorname{cov} [y_{i,t+k}, r(y_{i,t+k})]}{n\bar{y}_{t+k}} - C_{t+k|t} \right) + \left(C_{t+k|t} - \frac{2 \operatorname{cov} [y_{i,t}, r(y_{i,t})]}{n\bar{y}_t} \right) \quad (2.4)$$

$$\Delta G = \left\{ \frac{2 \operatorname{cov} [y_{i,t+k}, r(y_{i,t+k})]}{n\bar{y}_{t+k}} - \frac{2 \operatorname{cov} [y_{i,t+k}, r(y_{i,t})]}{n\bar{y}_{t+k}} \right\} + \quad (2.5)$$

$$\left\{ \frac{2 \operatorname{cov} [y_{i,t+k}, r(y_{i,t})]}{n\bar{y}_{t+k}} - \frac{2 \operatorname{cov} [y_{i,t}, r(y_{i,t})]}{n\bar{y}_t} \right\} \quad (2.6)$$

$$\Delta G = R + M \quad (2.7)$$

Così facendo si è ottenuta una scomposizione della variazione della concentrazione in due componenti denominate R e M . Nell'equazione (2.7), R è la componente di ri-classificazione che misura la variazione di concentrazione dovuta allo scambio di posizioni nella classifica dei periodi da t a $t + k$. R è pari a zero se la classifica dei periodi in $t + k$ è esattamente la stessa di quella in t . Raggiunge invece valore massimo, cioè $2G_{t+k}$, quando la classifica dei periodi in $t + k$ è completamente invertita rispetto alla classifica in t . La componente R consente di catturare qualsiasi deviazione dal modello stagionale, per questo motivo è vista come una misura della stabilità stagionale; più essa è grande, più il modello stagionale è instabile. Sempre considerando l'equazione (2.7) il termine M indica la componente di magnitudo e consente di misurare quanto il consumo di energia nel giorno $t + k$ è più o meno concentrato sui periodi con il rango più alto del giorno t . La componente di magnitudo viene calcolata tenendo i periodi ordinati in base al loro rango in t , per questo motivo si suppone che l'andamento stagionale sia lo stesso in t e in $t + k$. Assume

valore pari a zero se non ci sono stati cambiamenti nella distribuzione dei consumi o se i cambiamenti sono stati proporzionali in tutte le ore del giorno. Essa assume valori positivi nel momento in cui vi è un aumento della differenza (magnitudine) tra il picco e il punto di minimo dei consumi, assume invece valori negativi se il divario è diminuito.

Considerare solamente la variazione dell'indice di Gini può portare a conclusioni limitate e fuorvianti. Le due componenti infatti potrebbero annullarsi a vicenda, ma la differenza tra le due non va considerata come assenza di cambiamenti stagionali. Si esaminano gli 8 casi possibili, ottenuti come combinazione dei valori che R e M possono assumere, tenendo presente che il coefficiente associato alla componente di ri-classificazione è sempre non negativa.

1. $\Delta G < 0$ con $M < 0 \wedge R = 0$, l'intera variazione dell'indice di Gini è dovuta solamente alla componente di magnitudine.
2. $\Delta G = 0$ con $M = 0 \wedge R = 0$.
3. $\Delta G > 0$ con $M > 0 \wedge R = 0$, l'intera variazione dell'indice è dovuta solamente alla componente di magnitudine.
4. $\Delta G < 0$ con $M < 0 \wedge R > 0$, contributi opposti di R e M .
5. $\Delta G = 0$ con $M < 0 \wedge R > 0$, i contributi delle due componenti sono esattamente opposti e portano all'annullamento di ΔG .
6. $\Delta G > 0$ con $M < 0 \wedge R > 0$, contributi opposti di R e M .
7. $\Delta G > 0$ con $M = 0 \wedge R > 0$, aumento dell'indice è dovuto alla sola variazione dell'andamento stagionale, la magnitudine invece è stabile.

8. $\Delta G > 0$ con $M > 0 \wedge R > 0$, aumento della concentrazione stagionale dato dalla somma degli effetti di aumento della magnitudo stagionale e variazione dell'andamento stagionale.

2.3 Test sul cambiamento della variazione stagionale

L'indice di Gini, così come le sue componenti, sono strumenti statistici descrittivi e quindi non è possibile utilizzare le tecniche convenzionali per testare le ipotesi di significatività sul cambiamento rilevato. Per poter applicare gli usuali test sulle ipotesi relative alle misure di disuguaglianza è necessario che vengano verificate alcune assunzioni, tra cui la disponibilità di campioni ampi e l'indipendenza delle osservazioni. Tuttavia, i dati oggetto di studio non rispettano entrambi i requisiti in quanto sono serie storiche a cadenza oraria. Per valutare la significatività delle sue componenti separatamente, Grossi e Mussini (2021) hanno sviluppato una procedura basata sull'approccio suggerito da Rey e Sasté-Guitiérrez (2010).

Il metodo non parametrico proposto si basa su permutazioni casuali, al fine di verificare se la concentrazione è cambiata significativamente nel tempo. Inizialmente, l'approccio proposto da Rey e Sasté-Guitiérrez utilizzava l'indice di Theil, tuttavia può essere generalizzato a qualsiasi indice di concentrazione, quindi anche all'indice di Gini.

Siano I_{t+k} e I_t i valori di un indice di concentrazione al tempo $t+k$ e t rispettivamente. La statistica test è $\Delta I = I_{t+k} - I_t$, l'ipotesi su cui viene formulato il test è

$$\begin{cases} H_0 : & I_{t+k} - I_t = 0 \\ H_1 : & I_{t+k} - I_t \neq 0 \end{cases}$$

t	$t + k$	t^*	$t + k^*$
$y_{1,t}$	$y_{1,t+k}$	$y_{1,t}$	$y_{1,t+k}$
$y_{2,t}$	$y_{2,t+k}$	$y_{2,t+k}$	$y_{2,t}$
\vdots	\vdots	\vdots	\vdots
$y_{i,t}$	$y_{i,t+k}$	$y_{i,t}$	$y_{i,t+k}$
\vdots	\vdots	\vdots	\vdots
$y_{n,t}$	$y_{n,t+k}$	$y_{n,t+k}$	$y_{n,t}$

Tabella 2.1: Permutazioni casuali nelle coppie di osservazioni dalle distribuzioni $t + k$ e t .

La distribuzione campionaria di ΔI sotto l’ipotesi nulla viene costruita riassegnando casualmente gli indici temporali delle osservazioni, come si può vedere in Tabella 2.1. Le prime due colonne indicano i valori osservati in t e $t + k$ mentre la terza e la quarta mostrano due distribuzioni miste, ottenute assegnando casualmente gli elementi di ciascuna coppia di osservazioni riferite al periodo $i = 1, \dots, n$. Permutando le osservazioni all’interno delle coppie, le distribuzioni miste in t^* e $t + k^*$ si ottengono calcolando la correlazione tra i dati relativi allo stesso periodo. Rey e Sasté-Gutiérrez sostengono che t^* e $t+k^*$ sono i valori osservati in un processo in cui la concentrazione del fenomeno dovrebbe rimanere invariata nel tempo. Dopo aver effettuato la permutazione casuale degli indici temporali, si calcolano gli indici di concentrazione per le due nuove distribuzioni e la loro differenza. Eseguendo molte volte (J) questi passaggi di riassegnazione casuale si ottiene una distribuzione per la statistica test ΔI sotto l’ipotesi nulla. Dato $\Delta I_j = I_{j,t+k^*} - I_{j,t^*}$ la differenza calcolata per la j -esima riassegnazione casuale, con $j = 1, \dots, J$, sia d_j una variabile indicatrice che rappresenta il risultato del confronto tra ΔI_j e ΔI :

$$d_j = \begin{cases} 1 & \text{se } |\Delta I_j| \geq |\Delta I| \\ 0 & \text{se } |\Delta I_j| < |\Delta I| \end{cases} \quad (2.8)$$

Sia $D = \sum_{j=1}^J d_j$ il numero di volte in cui le differenze ottenute dalle J permutazioni casuali sono maggiori o uguali al ΔI osservato. A partire da questo si può ottenere un pseudo $p - value$ per il sistema di ipotesi specificato precedentemente nel seguente modo:

$$p(\Delta I | H_0) = \frac{1+D}{1+J} \quad (2.9)$$

Si rifiuta l'ipotesi nulla nel caso in cui lo pseudo $p - value$ è inferiore rispetto al livello di significatività fissato (0.05). Nel caso in cui H_0 viene rigettata si può affermare che c'è stato un cambiamento significativo nella concentrazione nell'arco temporale da t a $t+k$.

2.3.1 Test di significatività nelle due componenti

Si è visto precedentemente che possono esserci cambiamenti nell'andamento e nell'entità stagionale anche quando $\Delta G = 0$. La valutazione della significatività delle singole componenti R e M viene sviluppata da Grossi e Mussini (2021) attraverso una modifica del metodo si Rey and Sastré-Gutiérrez che consente di testare le ipotesi:

$$\begin{cases} H_0 : R = 0 & \left\{ \begin{array}{l} H_0 : M = 0 \\ H_1 : M \neq 0 \end{array} \right. \\ H_1 : R \neq 0 & \end{cases}$$

R e M sono pari a zero se i consumi di energia sono rimasti invariati o sono cambiati in modo proporzionale tra i tempi t e $t+k$. La distribuzione sotto l'ipotesi nulla per ogni statistica test viene costruita attraverso delle permutazioni casuali tra n coppie di dati osservati e ipotizzati in $t+k$. Sia $\lambda = \bar{y}_{t+k}/\bar{y}_t$ il rapporto tra il

t	$t + k$	$t + k^\lambda$	$t + k^*$
$y_{1,t}$	$y_{1,t+k}$	$y_{1,t+k^\lambda}$	$y_{1,t+k}$
$y_{2,t}$	$y_{2,t+k}$	$y_{2,t+k^\lambda}$	$y_{2,t+k^\lambda}$
\vdots	\vdots	\vdots	\vdots
$y_{i,t}$	$y_{i,t+k}$	$y_{i,t+k^\lambda}$	$y_{i,t+k}$
\vdots	\vdots	\vdots	\vdots
$y_{n,t}$	$y_{n,t+k}$	$y_{n,t+k^\lambda}$	$y_{n,t+k^\lambda}$

Tabella 2.2: Permutazioni casuali all'interno delle coppie di osservazioni delle serie $t + k$ e $t + k^\lambda$

consumo medio al tempo $t + k$ e quello al tempo t , moltiplicando i vari $y_{i,t}$ per λ si ottiene la distribuzione $t + k^\lambda$ che si può osservare nella Tabella 2.2.

La distribuzione $t + k^\lambda$ rappresenta la distribuzione che si sarebbe osservata in $t + k$ se la componente stagionale fosse rimasta stabile e con magnitudine pari a quella di t , il livello di concentrazione in $t + k^\lambda$ è per costruzione uguale a quello in t . Confrontando la distribuzione di $t + k^\lambda$ con quella in t si ha che l'andamento stagionale e la magnitudine sono invarianti nel tempo ed i consumi medi in $t + k^\lambda$ e $t + k$ sono per costruzione uguali. La distribuzione che si utilizza è quella rappresentata da $t + k^*$ (quarta colonna della Tabella 2.2), ottenuta attraverso una permutazione casuale all'interno di ogni coppia di osservazioni per lo stesso periodo i delle distribuzioni $t + k^\lambda$ e $t + k$ ($y_{i,t+k^\lambda}, y_{i,t+k}$). Si ottiene una distribuzione permutata solo per il tempo $t + k$ mentre la distribuzione al tempo t è quella osservata. I cambiamenti che si osservano nell'andamento stagionale e nella magnitudo possono essere valutati solo in riferimento alla distribuzione effettiva in t . Per una permutazione casuale, le due componenti sono ricalcolate sostituendo la distribuzione al tempo $t + k$ con quella al tempo $t + k^*$. Coincide con la distribuzione $t + k^\lambda$ se i consumi sono cambiati in modo proporzionale nel tempo, qui le componenti R e M risultano essere pari a zero poiché l'andamento stagionale e la magnitudine sono conservati. Ripetendo la permutazione casuale molte volte (J) si costruisce una distribuzione

per la statistica test sotto l'ipotesi nulla sia per la componente del cambiamento di *ranking* e della magnitudine.

R_j è la componente di *re-ranking* calcolata per la j -esima permutazione casuale ($j = 1, \dots, J$) e si crea una variaile indicatrice e_j che rappresenta il risultato del confronto tra R_j e il valore di R osservato:

$$e_j = \begin{cases} 1 & \text{se } R_j \geq R \\ 0 & \text{se } R_j < R \end{cases}$$

Sia $E = \sum_{j=1}^J e_j$ il numero di volte in cui il *re-ranking* è almeno uguale a R osservato. Lo pseudo *p-value* per la statistica può essere calcolato come

$$p(R | H_0) = \frac{1+E}{1+J} \quad (2.10)$$

Per un valore piccolo del *p-value* (minore di 0.05) si rifiuta l'ipotesi nulla per R , i dati quindi indicano che la stagionalità ha un andamento instabile nel corso del tempo (tra t e $t+k$).

M_j è la componente di grandezza calcolata per la j -esima permutazione casuale ($j = 1, \dots, J$), si crea allo stesso modo una variabile indicatrice f_j che rappresenta il risultato del confronto tra M_j e il valore di M osservato:

$$f_j = \begin{cases} 1 & \text{se } |M_j| \geq |M| \\ 0 & \text{se } |M_j| < |M| \end{cases}$$

Sia $F = \sum_{j=1}^J f_j$ il numero di volte in cui il modulo della magnitudine è almeno uguale a M osservato, lo pseudo *p-value* per la statistica test è

$$p(M | H_0) = \frac{1+F}{1+J} \quad (2.11)$$

Un valore piccolo del *p-value* (minore di 0.05) consente di rifiutare l'ipotesi nulla per M , i dati evidenziano quindi che la magnitudine è cambiata significativamente nel tempo (tra t e $t + k$).

Utilizzando lo stesso approccio è possibile testare anche la significatività di ΔG . Come già detto in precedenza, testare la significatività di ΔG significa verificare se il livello di concentrazione è cambiato ma non fornisce informazioni sulla significatività dei cambiamenti nell'andamento e nella magnitudine, misurati rispettivamente da R e M . Un ΔG piccolo o nullo può essere il risultato dei contributi fuori scala di R e M , che implicano cambiamenti notevoli nell'andamento stagionale e della magnitudo.

Sia $\Delta G_j = G_{t+k^*} - G_t$, ($j = 1, \dots, J$) la variazione dell'indice di Gini e h_j la variabile indicatrice che rappresenta il risultato del confronto tra ΔG_j e il ΔG osservato

$$h_j = \begin{cases} 1 & \text{se } |\Delta G_j| \geq |\Delta G| \\ 0 & \text{se } |\Delta G_j| < |\Delta G| \end{cases}$$

Sia $H = \sum_{j=1}^J h_j$ il numero di volte in cui il modulo di ΔG_j nelle J permutazioni casuali è maggiore o uguale al ΔG osservato, lo psuedo *p-value* viene calcolato come:

$$p(\Delta G | H_0) = \frac{1 + H}{1 + J} \tag{2.12}$$

Capitolo 3

Analisi empirica

Nel presente capitolo verrà effettuata un'applicazione pratica dei concetti relativi alla scomposizione dell'indice di Gini, includendo anche i test basati su permutazioni delle diverse componenti ΔG , R e M . I dati oggetto di studio sono stati presentati nella sezione 1.2.1 e fanno riferimento al consumo orario di energia elettrica al netto della produzione fotovoltaica. Si effettua un'analisi sulle variazioni delle componenti, sia in termini giornalieri, sia mensili, negli anni dal 2019 al 2023. L'obiettivo è di valutare la significatività delle componenti dell'indice di Gini per capire la dinamica temporale della *Duck Curve* sul mercato elettrico italiano ipotizzando poi quali politiche possono essere messe in atto al fine di migliorare la stabilità della rete.

3.1 Applicazione ai dati giornalieri

Partendo dai dati orari scaricati dalla piattaforma messa a disposizione da parte di Terna² dal 2019 al 2023 si vuole sintetizzare ogni curva giornaliera attraverso l'indice di Gini, esso verrà poi scomposto nelle sue componenti di ri-classificazione e magnitudine. Inizialmente la significatività delle componenti viene valutata con-

²[Download center](#) di Terna

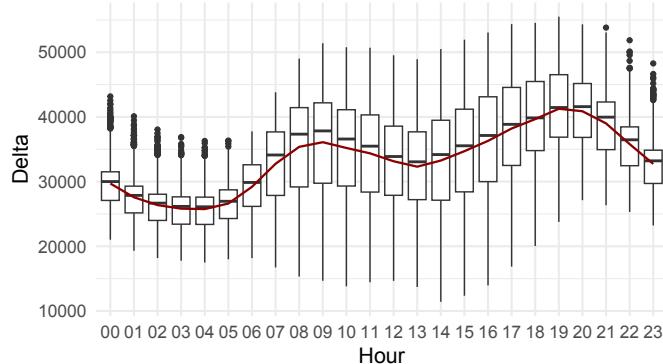


Figura 3.1: Profilo medio giornaliero della *Duck Curve*

frontando il valore dell’indice calcolato per ogni giorno con la curva oraria media ottenuta sintetizzando i 5 anni oggetto di studio, come si vede in Figura 3.1. Si può facilmente osservare come nelle ore centrali della giornata sia presente una variabilità maggiore, si tratta infatti di quei momenti in cui le persone sono più attive ed allo stesso tempo la produzione di energia è più elevata.

A partire dal profilo medio, considerato come tempo t_0 , sono stati calcolati l’indice di variazione ΔG , la componente di *ranking* R e quella di magnitudine M per ogni giorno. Successivamente attraverso le procedure di permutazione presentate nel capitolo precedente sono stati calcolati i valori del *p-value* relativi ai test. Nei grafici in Figura 3.2 si mostra la frequenza dei giorni significativi ($p\text{-value} < 0.05$) per ogni mese, relativamente alle tre componenti. È evidente come la componente di ri-classificazione, sia molto diversa da un giorno all’altro, in più dell’85% dei giorni in un mese si osserva un valore significativamente diverso da 0. Per quanto riguarda invece la magnitudine (M) si nota un risultato differente, in alcuni mesi è presente una frequenza bassa di giorni con un valore di M significativo; questi si concentrano principalmente nei mesi di marzo-aprile e novembre-dicembre dove si può affermare che l’ampiezza della curva giornaliera è cambiata in modo proporzionale a quella dei consumi medi. La frequenza di *p-value* significativi per ΔG sembra rispettare

uno schema che si ripete durante i cinque anni. Nei mesi più miti, ovvero quelli primaverili ed autunnali la frequenza dei *p-value* significativi si abbassa rispetto ai mesi estivi ed invernali, questo perchè diminuiscono i consumi e allo stesso tempo si ha una cospicua produzione solare.

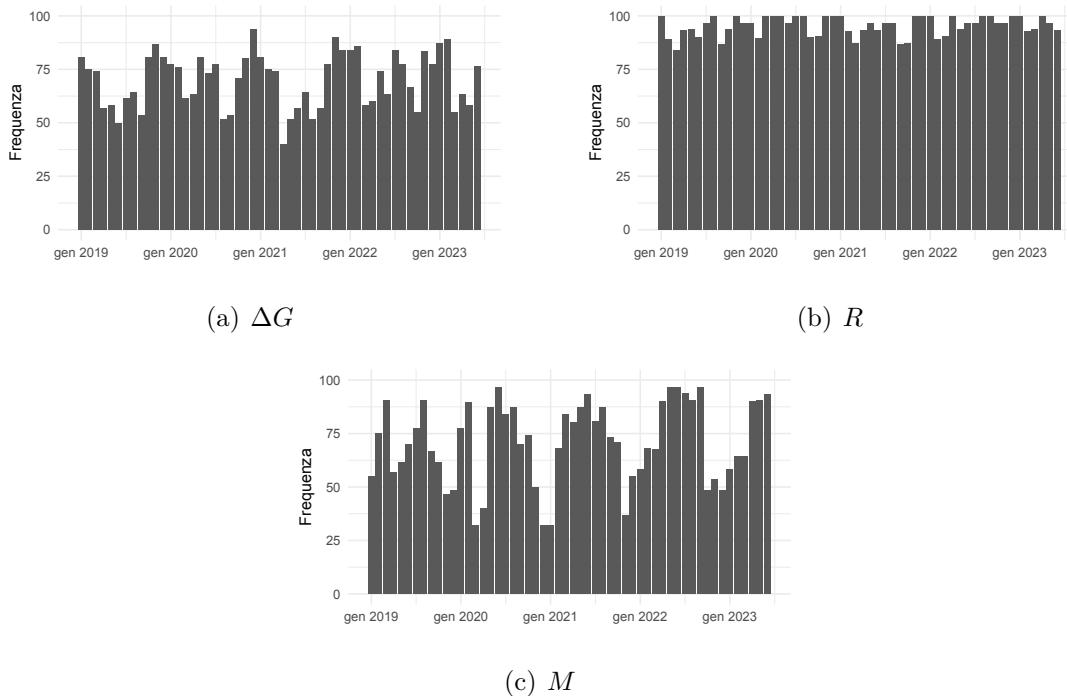


Figura 3.2: Frequenza dei giorni significativi ($p\text{-value} < 0.05$) per le tre componenti.

Vediamo ora più nel dettaglio il comportamento delle curve il 5 maggio, giorno in cui si è ottenuta la produzione massima di energia (Sezione 1.4.2), confrontandole con la curva media oraria. Nella Tabella 3.1 si possono vedere i risultati ottenuti nei cinque giorni; il valore di ΔG non risulta significativo nel 2021 e 2022, in particolare nel 2021 anche il valore associato alla componente di magnitudo è pari a 0. Nei grafici in Figura 3.3 si può vedere come variano le curve relative ai giorni analizzati. La curva in rosso raffigura l'andamento medio, ovvero il riferimento con cui viene fatto il confronto. La curva verde, relativa all'anno 2021, ha una forma

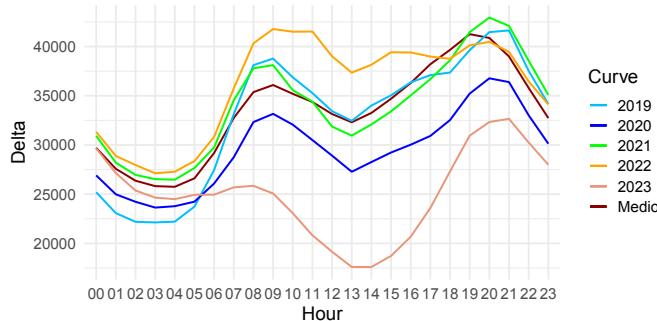


Figura 3.3: Confronto tra la *Duck curve* media e la *Duck curve* ottenuta in ogni anno il 5 maggio.

molto simile a quella del consumo orario medio, infatti la componente M risulta non significativa, i cambiamenti del 5 maggio 2021 sono stati proporzionali in tutte le ore del giorno rispetto ai consumi medi orari nei 5 anni. La curva rosa, che raffigura i consumi nel 5 maggio 2023, ha una forma nettamente differente da tutte le altre. In quel giorno infatti si è osservata la produzione massima di energia solare nelle ore centrali della giornata. Le componenti ΔG , R e M risultano essere tutte significativamente diverse da zero. Confrontando questo giorno con il consumo medio, si osserva che le curve presentano un andamento differente: qui il valore dato dalla differenza dell'indice di Gini nei due istanti temporali è positivo, ma i contributi delle componenti di ri-classificazione e magnitudine sono opposti ($R > 0$ e $M < 0$).

	Data	deltaG	R	M	pvalue.deltaG	pvalue.R	pvalue.M
1	2019-05-06	0.028	0.003	0.025	0.003	0.001	0.009
2	2020-05-06	-0.005	0.004	-0.009	0.039	0.001	0.001
3	2021-05-06	0.003	0.004	-0.001	0.188	0.001	0.585
4	2022-05-06	-0.003	0.008	-0.011	0.164	0.001	0.001
5	2023-05-06	0.018	0.069	-0.051	0.002	0.001	0.001

Tabella 3.1: Valori ottenuti dai test di permutazione per le varie componenti, confrontando ogni giorno con il giorno medio.

3.1.1 Confronto con giorno medio in ogni anno

A partire dall'analisi svolta precedentemente, sono state sviluppate altre varianti per valutare la robustezza del metodo, cercando di capire se fosse possibile prevedere il comportamento del valore dell'indice di concentrazione nel tempo. In questo caso si è scelto di utilizzare un giorno medio differente per ogni anno. I confronti che sono stati effettuati variano quindi a seconda del *benchmark* utilizzato.

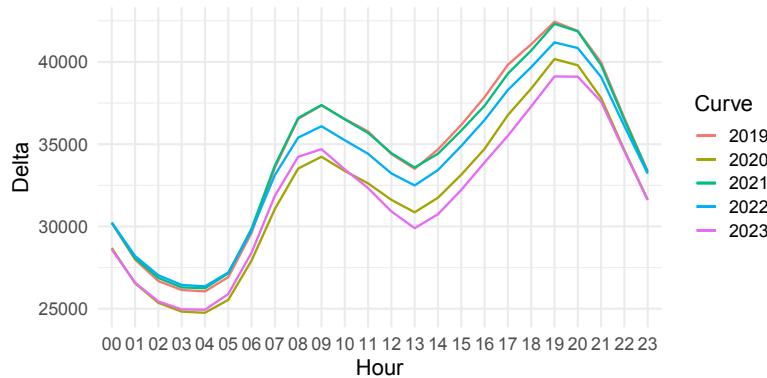


Figura 3.4: Profilo orario del carico residuo medio in Italia (*Duck Curve*)

Si può osservare il profilo medio orario nei cinque anni in Figura 3.4, la forma che assume la curva nel quinquennio non sembra aver subito variazioni; tuttavia nel 2020 e nel 2023 assume valori inferiori rispetto agli altri anni, questo è dovuto ad una maggiore penetrazione dell'energia solare. I grafici in Figura 3.5 sono analoghi a quelli analizzati precedentemente. In ogni mese si misura la frequenza di giorni in cui le componenti ΔG , R e M sono significativamente diverse da zero. Si vede chiaramente come per la componente di ri-classificazione la frequenza di giorni significativi in ogni mese è superiore a 75%, questo implica che i consumi cambiano tra i vari giorni e la curva media annuale non è in grado di riassumere in modo corretto il comportamento. Anche per la differenza tra l'indice di Gini nei due istanti e la componente di magnitudo si osserva un risultato analogo a quello precedente, in particolare le frequenze dei ΔG significativi diminuiscono leggermente per alcuni

mesi, se confrontate con quanto osservato in Figura 3.2. In alcuni mesi si osserva invece un risultato contrario per M , qui la frequenza dei giorni significativi aumenta, come ad esempio ad aprile 2019 che passa dal 56% al 67% di giornate in cui la componente di magnitudo è diversa da zero. La variazione dell'ampiezza delle curve non risulta essere correttamente colta da questo specifico riferimento; andando sempre più nello specifico, quello che ci si aspetta è di avere risultati sempre più vicini alla curva scelta come riferimento, in modo tale che la frequenza di valori significativi diminuisca sempre di più.

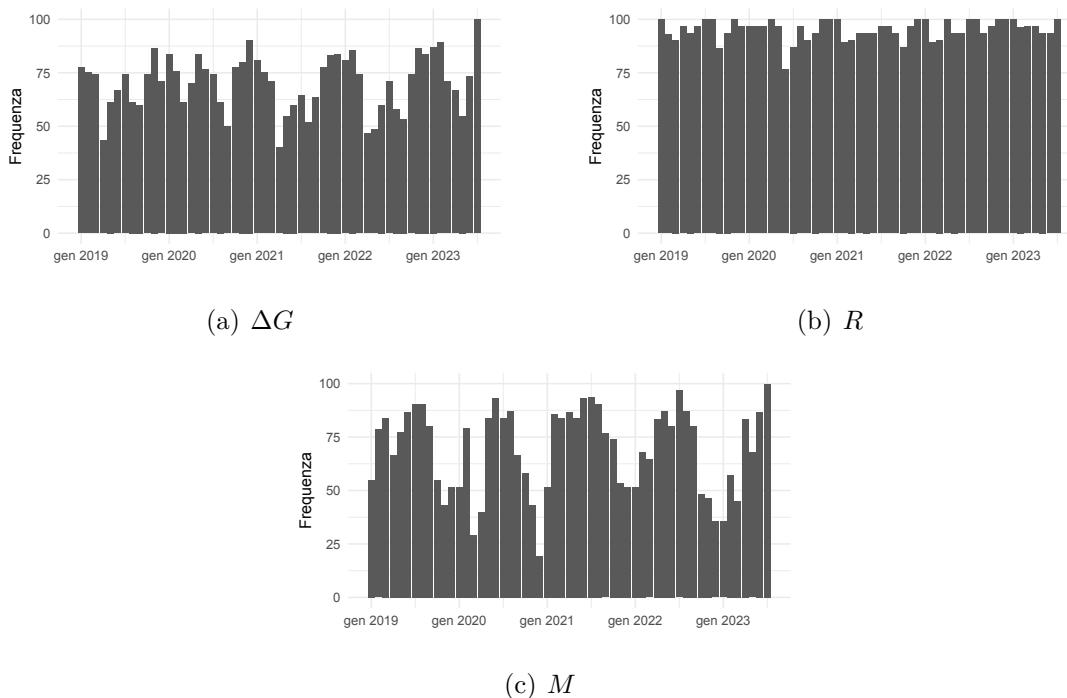


Figura 3.5: Frequenza dei giorni significativi ($p\text{-value} < 0.05$) per le tre componenti confrontando ogni giorno con il giorno medio per ogni anno.

3.1.2 Confronto con giorno medio in ogni mese

Cercando di analizzare sempre più nel dettaglio il comportamento dei consumatori si è scelto come riferimento il giorno medio per ogni mese durante i cinque anni analizzati. Ogni giornata quindi è stata confrontata con il profilo medio dello stesso mese a cui il giorno appartiene, ottenendo così 55 *benchmark*. In Figura 3.6 si può osservare la frequenza con cui vengono evidenziati dei giorni significativi per le tre componenti studiate. R risulta essere sempre fortemente significativa mentre ΔG e M molto più variabili. Utilizzando per le analisi diversi riferimenti si abbassa notevolmente il rumore proveniente dai dati, tuttavia questo non sembra cambiare in modo rilevante i risultati ottenuti.

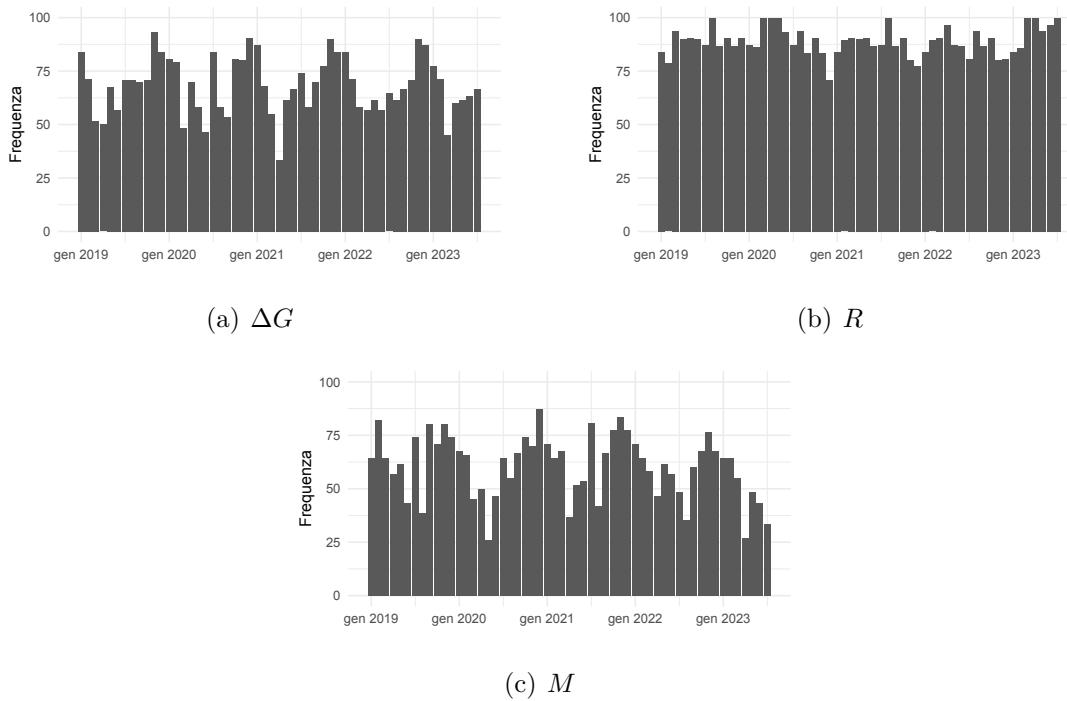


Figura 3.6: Frequenza dei giorni significativi ($p\text{-value} < 0.05$) per le tre componenti confrontando ogni giorno con il giorno medio per mese in ogni anno.

3.1.3 Confronto con giorno medio feriale e nel fine settimana

Essendo i dati a disposizione relativi al consumo netto di energia elettrica sembra ragionevole ipotizzare che il comportamento dei consumatori sia differente tra giorni feriali e fine settimana, dove per fine settimana si intendono il sabato e la domenica. In Figura 3.7 si può vedere il profilo medio delle due curve ottenute considerando da una parte solamente i giorni feriali (da lunedì al venerdì) e dall'altra il fine settimana (sabato e domenica). Si nota chiaramente come la forma della curva sia la stessa ma durante il fine settimana il consumo netto sia inferiore rispetto agli altri giorni. Nel fine settimana la maggior parte delle attività produttive restano chiuse quindi diminuisce in modo considerevole il consumo di energia, allo stesso modo però la quantità di energia solare media prodotta non cambia rispetto agli altri giorni, per cui il consumo netto diminuisce.

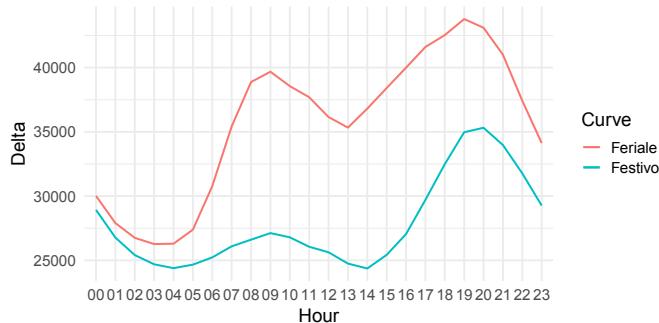


Figura 3.7: Profilo della *Duck Curve* per giorni feriali e fine settimana.

Partendo da questa evidenza si è deciso di studiare il comportamento delle tre componenti ΔG , R e M discriminando tra giorni feriali e il fine settimana; ogni giornata è stata confrontata con il corrispondente profilo medio a seconda che si trattasse di un giorno durante la settimana o nel *weekend*. I risultati che emergono sono rappresentati in Figura 3.8 dove si può osservare la frequenza mensile dei

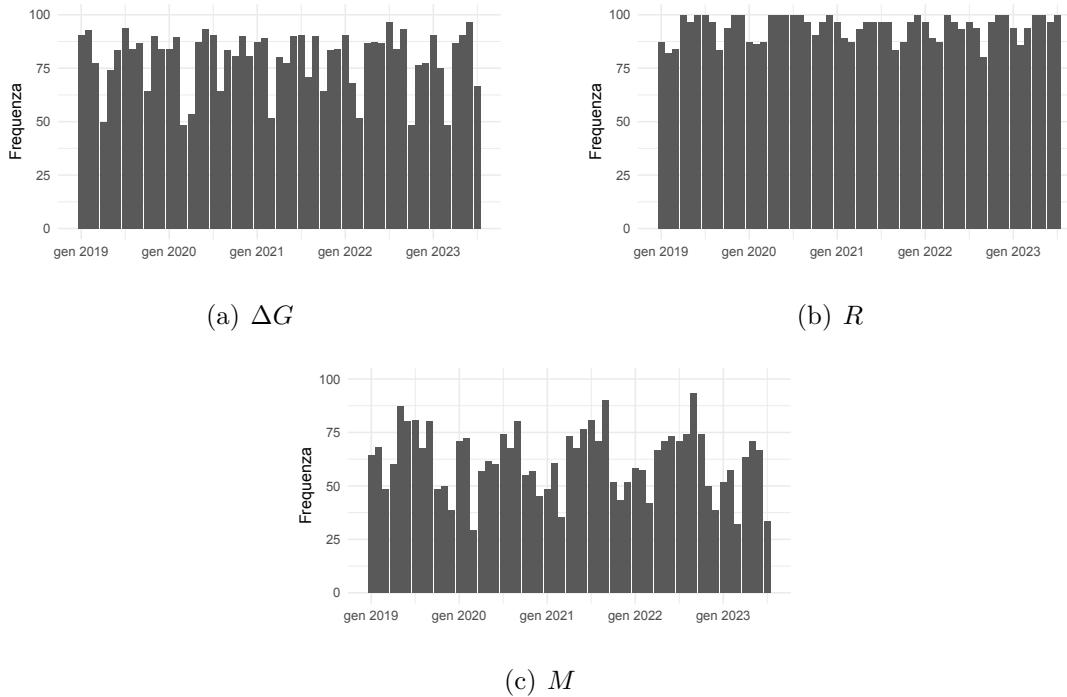


Figura 3.8: Frequenza dei giorni significativi ($p\text{-value} < 0.05$) per le tre componenti confrontando tra profili medi feriali e fine settimana.

giorni in cui la componente è significativamente diversa da zero. La componente di *ranking* si mantiene sempre fortemente significativa, in alcuni mesi addirittura tutti i giorni risultano differenti rispetto al profilo di riferimento. Sia la magnitudine che la differenza tra gli indici di variazione della concentrazione si mantengono variabili durante gli anni, in modo coerente con quanto evidenziato nei risultati delle analisi svolte precedentemente. Un aspetto interessante da evidenziare è come, per alcuni mesi, si osservi una frequenza maggiore di giorni significativi per la componente di magnitudo, rispetto agli stessi risultati ottenuti confrontando ogni giornata con il giorno medio che varia di anno in anno. Questo tuttavia resta coerente con quanto emerso nel momento in cui il confronto veniva effettuato con il giorno medio differente per ogni mese. Il motivo di questa evidenza può essere

spiegata analizzando i mesi in cui ciò si verifica, ovvero quando sono presenti dei periodi di vacanza come ad esempio Natale, che però non ricadono in un giorno classificato come un giorno nel fine settimana. Il comportamento dei consumatori in quella specifica giornata sarà differente rispetto al profilo medio generato in un primo momento.

3.1.4 Confronto con giorno della settimana

Rispetto alle analisi svolte finora, qui si vuole confrontare i consumi di ogni giornata con il profilo medio di uno specifico giorno della settimana. Nel grafico in Figura 3.9 si possono vedere i sette profili medi orari utilizzati per effettuare i test sulle componenti ΔG , R e M .

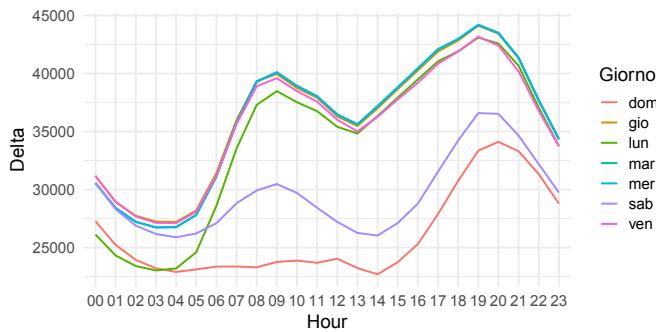


Figura 3.9: Profilo della *Duck Curve* per i 7 giorni della settimana.

La domenica risulta essere il giorno in cui il consumo netto di energia elettrica è inferiore, la maggior parte delle aziende sono chiuse e i consumatori tendono a passare più tempo fuori casa riducendo quindi l'uso dei sistemi di riscaldamento o raffreddamento. Per lo stesso motivo anche nella giornata di sabato i consumi risultano inferiori rispetto agli altri giorni della settimana seppur leggermente superiore se confrontato con la domenica. I giorni che vanno dal lunedì al venerdì invece presentano un comportamento pressoché simile durante tutte le ore, un caso particolare è il lunedì tra mezzanotte e le 8, orario in cui le attività produttive non

sono ancora pienamente operative dopo il fine settimana e si osserva quindi un *Delta* inferiore rispetto agli altri giorni.

A partire da questi dati sono stati effettuati i test di permutazione per ottenere una misura della significatività delle tre componenti, ogni giorno da gennaio 2019 a luglio 2023 è stato confrontato con il profilo medio del giorno della settimana corrispondente. In Figura 3.10 si può osservare la frequenza mensile con cui viene osservato un giorno in cui la specifica componente è significativamente diversa da zero. La componente R presenta un'elevata frequenza di giorni significativi, ancora maggiore rispetto alle altre analisi svolte. Sia ΔG che M risultano molto variabili, sia a marzo 2019 che a giugno 2020 ΔG risulta significativo in tutti i giorni questo vuol dire che la curva utilizzata come riferimento non consente di catturare in modo corretto le variazioni giornaliere che invece cambiano molto. Per quanto riguarda la componente di magnitudo invece si nota che per alcuni mesi diminuisce la significatività, quindi si coglie una certa regolarità nell'ampiezza della curva nei vari giorni della settimana.

Le analisi che sono state qui riportate, scegliendo di volta in volta un riferimento con cui fare il confronto sempre più specifico, al fine di eliminare il rumore presente nei dati giornalieri, hanno portato alla conclusione che il metodo utilizzato risulta essere robusto. Le conclusioni tratte in ogni sezione ci portano a dire che questa caratterizzazione per la *Duck Curve* in Italia non risulta essere così nitida, si veda l'elevata frequenza di giorni significativi in tutti i mesi. La penetrazione del solare non è ancora tale da riuscire a cambiare completamente i profili della domanda residua.

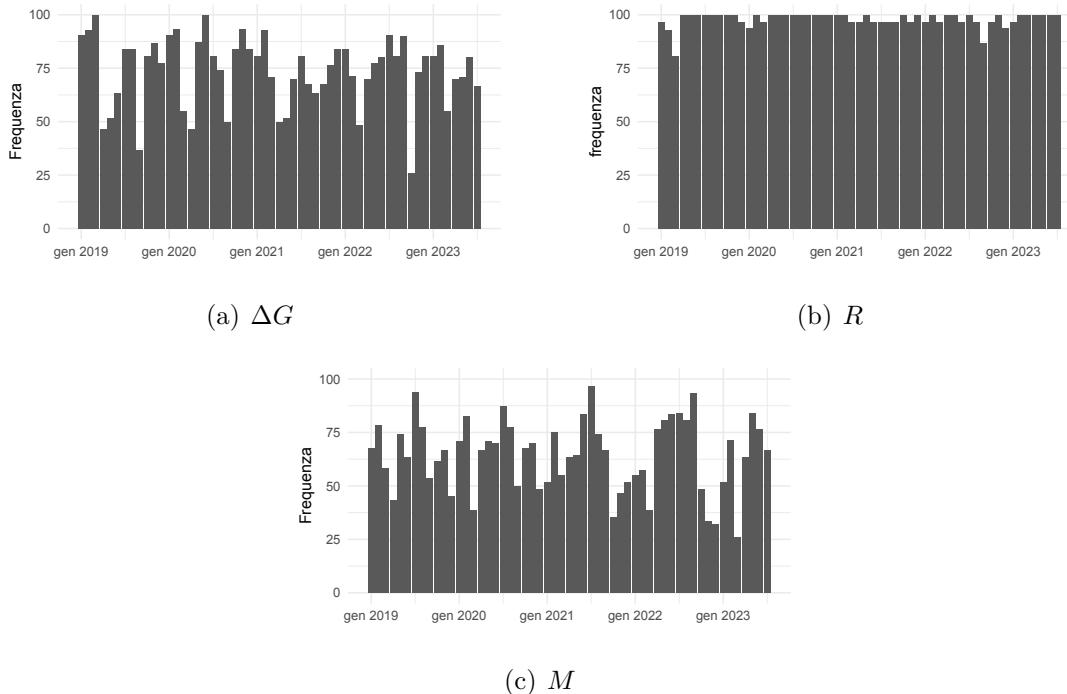


Figura 3.10: Frequenza dei giorni significativi ($p\text{-value} < 0.05$) per le tre componenti confrontando ogni giorno con il profilo medio di quel giorno della settimana.

3.2 Applicazione ai dati mensili

Un’analisi analoga è stata svolta sui dati aggregati, invece di considerare i singoli giorni si è presa come unità di riferimento i vari mesi ottenuti come media dei consumi orari nell’arco del periodo. Nel grafico in Figura 3.11 si può osservare il profilo medio utilizzato per svolgere le analisi, in particolare inizialmente ogni mese è stato confrontato con un mese medio calcolato su tutti e 5 gli anni (curva rossa). Successivamente è stato individuato un profilo medio per ogni anno e la significatività delle tre componenti in ogni mese è stata ottenuta confrontando ogni mese con il rispettivo profilo medio annuale. Si osserva che tutte le curve sono molto simili tra loro sia nella forma che nei valori assunti, il 2020 e 2023 sono gli anni in

cui mediamente il consumo netto di energia è leggermente inferiore.

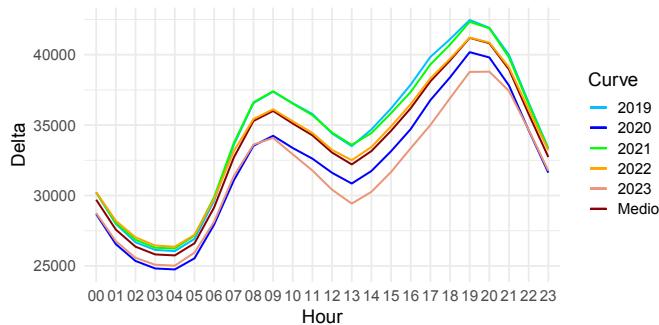


Figura 3.11: Profilo della *Duck Curve* calcolata a partire dai dati mensili.

A partire dai dati con cadenza mensile sono stati svolti i test di permutazione per valutare la significatività delle componenti ΔG , R e M . I risultati ottenuti sono rappresentati in Figura 3.12. In particolare si può osservare il valore del *p-value* ottenuto in ogni mese sia utilizzando un mese medio calcolato su tutti e cinque gli anni (a sinistra), sia un riferimento differente per ogni anno (a destra). In entrambi le analisi si ottiene una componente di ri-classificazione fortemente significativa in tutti i mesi, ad esclusione di marzo 2020 periodo nel quale è scoppiata la pandemia di *COVID-19* e i consumi di energia sono cambiati. Anche per la componente di magnitudo e per la differenza di concentrazione dell'indice di Gini si osservano comportamenti pressoché simili nelle due analisi; i mesi più miti ovvero in primavera ed autunno risultano essere più stabili e prevedibili.

Anche l'aggregazione dei dati porta a risultati coerenti con quanto detto precedentemente con le osservazioni giornaliere. Il metodo utilizzato è robusto, la *Duck Curve* cambia nel tempo e la componente di *ranking* cambia sempre, per cui non è mai possibile catturarla in modo corretto.

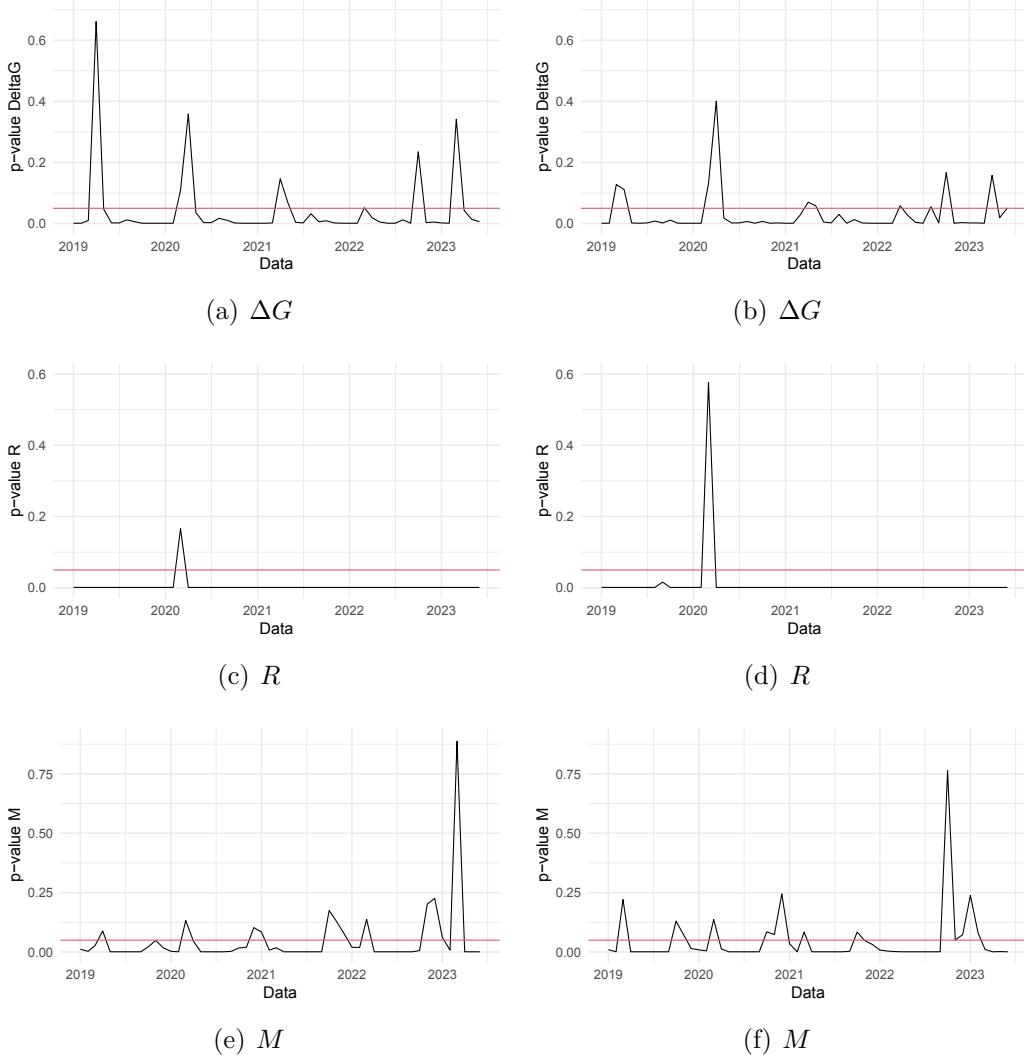


Figura 3.12: Nella colonna di sinistra sono rappresentati i *p-value* ottenuti confrontando ogni mese con il mese medio nei 5 anni mentre in quella a destra il confronto è stato fatto con un mese medio differente per ogni anno.

3.3 Modelli previsivi per ΔG

Utilizzando i dati a cadenza giornaliera ed in particolare la serie storica degli indici di variazione stagionale ΔG è stato stimato un modello con lo scopo di prevedere

l'andamento dei consumi nel corso degli anni.

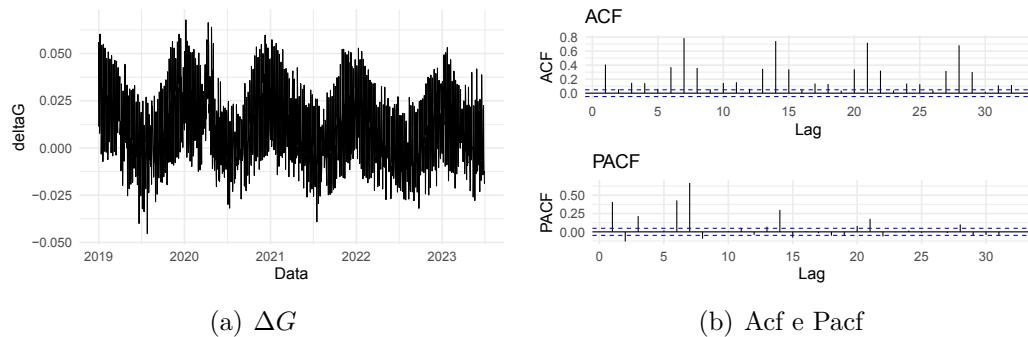


Figura 3.13: Analisi della serie storica di ΔG con autocorrelazione globale e parziale.

Ogni giorno è stato calcolato il valore dell'indice di variazione, si può osservare in Figura 3.13 l'andamento nei cinque anni e i grafici relativi all'analisi dell'autocorrelazione sia globale che parziale. Per l'analisi dei dati e la definizione di un modello è stato considerato il periodo che va da gennaio 2019 fino a dicembre 2022, mentre i dati relativi al 2023 sono stati utilizzati per effettuare le previsioni e valutare il modello. In questi dati è evidente la forte presenza di stagionalità, data anche dalla struttura stessa dei dati, in particolare si tratta di serie temporali multistagionali. Poichè si tratta di dati giornalieri si nota uno schema ricorrente sia settimanale che annuale che verrà inserito nel momento nella specificazione del modello.

Il metodo utilizzato per la scomposizione della serie prende il nome di *Seasonal and Trend decompositon using Loess*, STL, sviluppato da R.B. Cleveland, Cleveland, McRae & Terpenning. Si tratta di un metodo versatile e robusto che permette di stimare anche relazioni non lineari e consente di gestire qualsiasi tipo di stagionalità. Dopo aver applicato la scomposizione STL si modellano i dati de-stagionalizzati adattando un modello Arima che permette di stimare l'andamento della componente in esame. A partire dal modello stimato sono state effettuate delle previsioni in *rolling window*, un approccio ampiamente utilizzato nell'analisi

delle serie temporali. Secondo tale tecnica le previsioni vengono ottenute in modo sequenziale, spostando di volta in volta la finestra temporale su cui le previsioni si basano. Si usa questa tecnica per effettuare le previsioni in quanto consente di valutare l'efficacia dei modelli nel tempo. La tecnica è particolarmente efficace quando i dati hanno stagionalità e le tendenze che possono cambiare nel periodo. In questo studio si è scelto di utilizzare una finestra di ampiezza 1 in modo che si possano prevedere i consumi solamente per il giorno successivo, ciò prende il nome di *one-step ahead forecasting*. Questa metodologia è utile per valutare quanto il modello riesce a catturare le fluttuazioni e le dinamiche a breve termine nei dati.

Nel grafico in figura 3.14 si può osservare l'adattamento delle previsioni effettuate rispetto ai valori osservati, in particolare è stato effettuato uno zoom (Figura 3.14(b)) in cui si vuole porre l'attenzione sulle previsioni ottenute con il metodo di *rolling window*. Si può affermare che il modello coglie correttamente le fluttuazioni dei dati in alcuni giorni mentre per altri, in cui si osserva un comportamento differente, non riesce a catturare tali differenze. Il modello che utilizza il metodo STL per la decomposizione della serie risulta essere quello che si adatta meglio ai dati e presenta un errore percentuale medio inferiore, ma comunque molto alto (pari a 68.32). L'errore percentuale medio, MPE, è una metrica utilizzata per valutare l'accuratezza del modello di previsione; il valore ottenuto indica che l'errore medio delle previsioni del modello è pari al 68% rispetto ai dati osservati. Un errore così alto ci porta a dire che le previsioni calcolate non sono affidabili, in particolare tende a sovrastimare i valori effettivi.

A partire da questo modello è possibile apportare delle modifiche, aggiungendo anche altre caratteristiche proprie dei dati sulla produzione di energia solare, quali ad esempio le previsioni meteorologiche per il giorno successivo, al fine di migliorare sempre di più le previsioni. L'utilizzo dei dati giornalieri introduce una forte variabilità per cui sarà necessario inserire delle variabili che consentono di control-

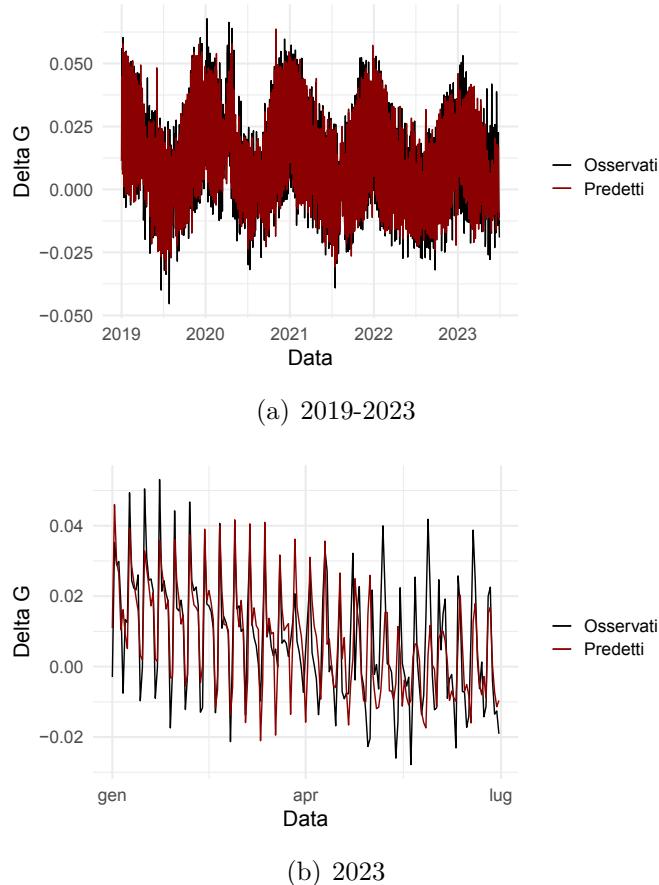


Figura 3.14: Confronto tra valori osservati e predetti nei 5 anni 3.14(a) e più nello specifico durante il 2023 3.14(b).

larla nel modo migliore. Con queste analisi si è cercato di studiare la diffusione del fenomeno dell’*Duck Curve* in Italia. La curva prodotta giornalmente varia nel tempo ed in particolare non si riesce mai a cogliere in modo esatto la componente di ri-classificazione. È difficile decidere quindi che politiche adottare, se la scelta si basa esclusivamente su questo schema, nel nostro paese infatti la penetrazione del solare non è così incisiva da creare problemi di stabilità nella rete, come invece è accaduto in California. Per questo motivo è necessario investire maggiormente su questa fonte rinnovabile e pulita.

Conclusioni

La ricerca condotta in questa tesi ha gettato luce sul fenomeno complesso ed in continua evoluzione noto come *Duck Curve* nel contesto dell'energia solare e delle reti elettriche. L'obiettivo principale di questo studio è quello di analizzare per poi prevedere l'andamento della curva attraverso una scomposizione dell'indice di Gini, proposta da Grossi & Mussini nel 2021, nelle due sue componenti di *ranking* e magnitudine al fine di prevedere meglio il suo comportamento e le sue implicazioni.

I risultati di questa ricerca hanno rivelato una caratteristica distintiva della *Duck Curve*: la mancanza di uno schema regolare nel tempo e prevedibile. In altre parole, la curva si presenta come un fenomeno altamente variabile nel tempo, che rappresenta una notevole sfida per gli operatori delle reti elettriche e gli esperti di pianificazione energetica. Questa variabilità è in gran parte dovuta alla natura intermittente dell'energia solare e alla fluttuazione della domanda energetica nei diversi giorni. Nel Capitolo 3 sono state effettuate varie analisi con l'obiettivo di testare la metodologia basata sulla scomposizione dell'indice di Gini per poter prevedere l'andamento dei consumi in Italia dal 2019 al 2023. Quello che è emerso, sia utilizzando i dati giornalieri che quelli mensili, è la robustezza del metodo utilizzato: in tutte le analisi le conclusioni tratte sono molto simili. È possibile affermare che la *Duck Curve* cambia spesso nel tempo ed in particolare la componente di ri-classificazione cambia sempre per cui non è mai possibile catturarla in modo affidabile.

Un aspetto interessante che emerge dall'analisi qui svolta è il fatto che in Italia il problema della *Duck Curve* non è ancora così acuto come in altri contesti, sebbene l'energia solare stia guadagnando terreno come fonte energetica. Questo problema potrebbe essere visto come un'opportunità unica per il paese. L'investimento nelle infrastrutture solari e la pianificazione anticipata potrebbero aiutare a gestire meglio la transizione verso un futuro energetico più sostenibile, evitando alcune delle sfide che gli altri paesi stanno affrontando, come il caso della California. Questa tesi vuole sottolineare l'importanza di promuovere l'adozione dell'energia solare e di investire in infrastrutture fotovoltaiche. Anche se la *Duck Curve* non è ancora una preoccupazione immediata in Italia, è essenziale prepararsi per le sfide future. L'energia solare offre un potenziale significativo per ridurre l'impatto ambientale, ridurre i costi energetici e garantire una fornitura energetica affidabile.

In conclusione, questa ricerca ha contribuito a comprendere meglio la *Duck Curve* e il suo impatto sulle reti elettriche, evidenziando la necessità di una pianificazione strategica e di investimenti mirati. L'Italia ha l'opportunità di diventare un leader nell'adozione dell'energia solare e di modellare un futuro energetico più sostenibile. La sfida ora è quella di tradurre questa consapevolezza in azioni concrete per un sistema energetico migliore e più verde.

Bibliografia

BROŻYNA, J., Mentel, G., Szetela, B. & Strielkowski, W. (2018), ‘MULTI-SEASONALITY IN THE TBATS MODEL USING DEMAND FOR ELECTRIC ENERGY AS A CASE STUDY’, *Economic Computation & Economic Cybernetics Studies & Research* **52**(1).

Calero, I., Canizares, C. A., Bhattacharya, K. & Baldick, R. (2022), ‘Duck-Curve Mitigation in Power Grids With High Penetration of PV Generation’, *IEEE Transactions on Smart Grid* **13**(1), 314–329.

Confronting the Duck Curve: How to Address Over-Generation of Solar Energy (n.d.).

Duro, J. A. (2016), ‘Seasonality of hotel demand in the main Spanish provinces: Measurements and decomposition exercises’, *Tourism Management* **52**, 52–63.

Duro, J. A. (2018), ‘Seasonality of tourism: A new decomposition’, *Tourism Economics* **24**(5), 615–621. Publisher: SAGE Publications Ltd.

Fernández-Morales, A. (2003), ‘Decomposing seasonal concentration’, *Annals of Tourism Research* **30**(4), 942–956.

- Fernández-Morales, A., Cisneros-Martínez, J. D. & McCabe, S. (2016), ‘Seasonal concentration of tourism demand: Decomposition analysis and marketing implications’, *Tourism Management* **56**, 172–190.
- Furman, E., Kye, Y. & Su, J. (2019), ‘Computing the Gini index: A note’, *Economics Letters* **185**, 108753.
- Gastwirth, J. L. (1972), ‘The Estimation of the Lorenz Curve and Gini Index’, *The Review of Economics and Statistics* **54**(3), 306–316. Publisher: The MIT Press.
- Grossi L & Mussini M (2021), ‘Seasonality in tourist flows: Decomposing and testing changes in seasonal concentration’. Publisher: Elsevier Ltd.
- Haider, R., Ferro, G., Robba, M. & Annaswamy, A. M. (2022), Flattening the Duck Curve: A Case for Distributed Decision Making, in ‘2022 IEEE Power & Energy Society General Meeting (PESGM)’, IEEE, Denver, CO, USA, pp. 01–05.
- Hou, Q., Zhang, N., Du, E., Miao, M., Peng, F. & Kang, C. (2019), ‘Probabilistic duck curve in high PV penetration power system: Concept, modeling, and empirical analysis in China’, *Applied Energy* **242**, 205–215.
- Kalair, A. R., Abas, N., Seyedmahmoudian, M., Rauf, S., Stojcevski, A. & Khan, N. (2021), ‘Duck curve leveling in renewable energy integrated grids using internet of relays’, *Journal of Cleaner Production* **294**, 126294.
- Metrics to measure the geographic characteristics of tourism markets: An integrated approach based on Gini index decomposition - ScienceDirect* (n.d.).
- Olczak, P., Jaśko, P., Kryzia, D., Matuszewska, D., Fyk, M. I. & Dyczko, A. (2021), ‘Analyses of duck curve phenomena potential in polish PV prosumer households’ installations’, *Energy Reports* **7**, 4609–4622.

- Pitra, G. M. & Musti, K. S. (2021), Duck Curve with Renewable Energies and Storage Technologies, in ‘2021 13th International Conference on Computational Intelligence and Communication Networks (CICN)’, IEEE, Lima, Peru, pp. 66–71.
- Rosselló, J. & Sansó, A. (2017), ‘Yearly, monthly and weekly seasonality of tourism demand: A decomposition analysis’, *Tourism Management* **60**, 379–389.
- Sarkar, S. (2023), ‘Gini decomposition: An inequality of opportunity perspective’, *Economics letters* **223**, 110975–. Publisher: Elsevier B.V.
- Schmalensee, R. (2022), ‘Competitive Energy Storage and the Duck Curve’, *The Energy Journal* **43**(2).
- Wang, Chang, Bai, Liu, Dai & Tang (2019), ‘Mitigation Strategy for Duck Curve in High Photovoltaic Penetration Power System Using Concentrating Solar Power Station’, *Energies* **12**(18), 3521.
- Weron, R. (2006), *Modeling and Forecasting Electricity Loads and Prices: A Statistical Approach*, 1 edn, Wiley.