

UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Fisica e Astronomia "Galileo Galilei"

Dipartimento di Matematica "Tullio Levi-Civita"

Corso di Laurea in Fisica

Tesi di Laurea

Confronto tra modelli stocastici: dall'irraggiamento al mercato dell'energia solare

Relatore

Prof./Dr. Fulvio Baldovin

Correlatore

Prof./Dr. Giorgia Callegaro

Laureando

Federico Passuello

Anno Accademico 2022/2023

Abstract

In questo lavoro di tesi, confrontiamo 2 possibili modelli per il calcolo del Production-at-Risk di un parco solare situato ad Amburgo, in Germania.

I dati raccolti, tramite CAMS, sezione dell'Esa (European Space Agency), sono relativi all'anno 2019 e l'analisi viene effettuata a tempo discreto.

Nella tesi introduciamo un modello stocastico che permette, dopo una calibrazione dei suoi parametri ai dati reali, una previsione dell'irraggiamento giornaliero per ogni coordinata spaziale desiderata. La funzione di irraggiamento è descritta attraverso una componente additiva deterministica stagionale ed una stocastica, basata sulle serie autoregressive.

Attraverso l'utilizzo di processi stocastici che descrivono il sistema complesso dell'irraggiamento, siamo in grado di calcolare il P@R (Production-at-Risk) di un parco solare, mediante procedura di simulazione. Nel lavoro svolto, sfruttiamo i processi ARCH nella descrizione della funzione d'irraggiamento puntuale. Confrontiamo questa proposta con un modello da noi sviluppato, su base statistica, che parte da una verifica del modello precedente, applicato non più ad un punto, ma ad una superficie. Otteniamo il risultato di aver dimostrato l'esistenza di un tetto massimo al costo dell'energia solare. Il nostro modello si basa sulla proposta di essere nella posizione economica più conservativa possibile. Partendo dalla richiesta di accettare solo i valori superiori alla soglia di minimo, otteniamo una descrizione nella quale l'irraggiamento registrato è sempre maggiore o uguale alla richiesta d'irraggiamento minimo prevista.

Riportando la situazione descritta al mercato dell'energia, non vi sarà, se si limitano le considerazioni al mercato del solare, uno squilibrio tra domanda in difetto rispetto all'offerta.

Nell'appendice vengono introdotti: l'angolo solare orario, la nozione di autocorrelazione, i processi gaussiani e le catene di Markov omogenee a tempo discreto.

Indice

1	Introduzione	1
2	Il mercato energetico	3
3	Il modello matematico	6
4	Caratterizzazione dell'irraggiamento	9
4.1	Componente stagionale	10
4.2	Componente stocastica	10
4.2.1	Volatilità condizionata	11
4.2.2	Driving noise process	11
4.2.3	Irraggiamento previsto	13
5	Applicazione al mercato dell'energia solare del modello statistico	15
5.0.1	Production-at-risk P@R	16
6	Proposta del modello statistico	18
7	Conclusioni	22
A	Descrizione dell'angolo solare orario	23
B	Teorema di Kolmogorov	24
C	Processi di autocorrelazione	25
D	Il processo gaussiano	26
D.1	Le catene di Markov omogenee a tempo discreto	26

Capitolo 1

Introduzione

Una delle grandi sfide che siamo chiamati ad affrontare è l'approvvigionamento energetico.

Come si può vedere, l'ECpc (*Energy Consumption pro capite*) calcolato su tutto il Mondo ha un valore costante da oltre 30 anni. La necessità di rispondere alla domanda energetica, richiede una riflessione approfondita sulle fonti.

L'indice che più però dovrebbe essere al centro degli studi è il TPEF (*total primary energy footprint*): riflette il consumo delle risorse naturali in tutto il mondo per soddisfare la domanda finale. Attraverso questi indici siamo in grado di descrivere con maggiore chiarezza l'impronta del consumo energetico di ciascuna persona, parametrizzando quanto incide sulle fonti energetiche.

La necessità di abbandonare le fonti non rinnovabili si inserisce nella realtà economica che viviamo. Nell'ottica di una transizione energetica, a dicembre del 2018 il Consiglio dell'Unione Europea ha approvato il regolamento sulla governance, che definisce il quadro dell'energia e dell'azione per il clima. Entro il 2030 l'UE dovrà ricavare il 32% della sua energia da fonti rinnovabili e raggiungere l'obiettivo principale di efficienza energetica del 32,5%. Questa volontà di lavorare per un cambiamento delle fonti energetiche è stata ribadita alla Cop27 di Sharm el-Sheikh del novembre 2022.

In questo quadro è interessante vedere come il 2019 sia stato il primo anno della storia in cui si è riusciti a produrre oltre della metà dell'energia utilizzata in Germania non dalle fonti fossili.

Indipendentemente dallo scenario che ci sarà, la Germania ha programmato di azzerare l'energia proveniente da combustibili fossili, incrementando l'utilizzo di forme alternative, quali nucleare e rinnovabili entro il 2045.

Diventa quindi necessario investire nelle fonti rinnovabili, sia in ambito tecnologico che in ambito educativo per il futuro.

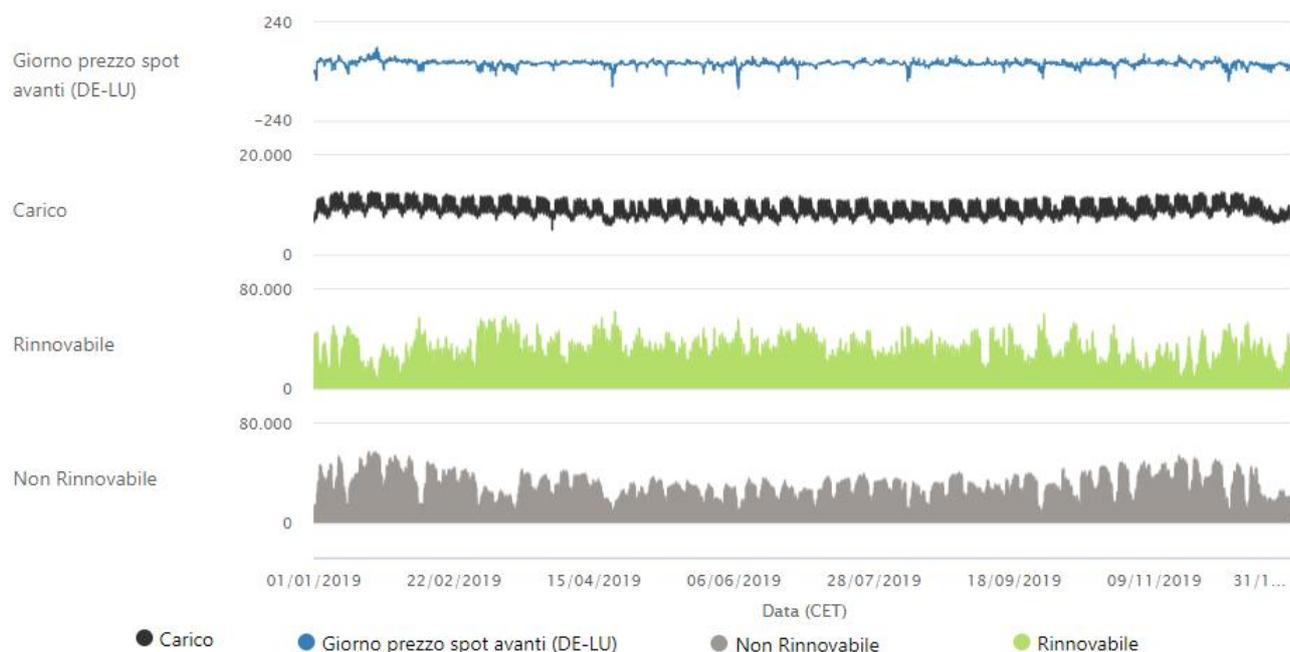


Figura 1.1: Confronto tra i volumi energetici prodotti in relazione al carico e al costo giornaliero dell'elettricità

Come si vede dal grafico, il carico di energia fornita è una funzione costante durante la settimana, che decresce il sabato ed ha il minimo la domenica, ciò è dovuto alla chiusura degli impianti produttivi. Durante l'anno il carico segue questa regolarità settimanale, con valori pressoché stabili. Questa osservazione trova conferma nel clima continentale e nella distribuzione sul lungo periodo delle chiusure per le festività. Guardando i prezzi del mercato, si osserva un andamento coerente che presenta dei minimi, a volte negativi, le domeniche. Questa è la situazione in cui l'offerta energetica è di molto maggiore della domanda.

Poiché il volume, cioè la produzione energetica delle fonti rinnovabili e di quelle non rinnovabili, deve essere costante, si osserva che l'energia derivata dalle fonti modulabili compensa le fonti rinnovabili variabili.

Le fluttuazioni dell'energia solare prodotta sono essenzialmente dovute alla radiazione solare e alla copertura del cielo. Viste le grandi fluttuazioni che si possono registrare su una superficie ampia, tanto più sarà specifica la zona in esame, tanto più sarà precisa l'analisi.

Capitolo 2

Il mercato energetico

Il mercato dell'energia elettrica è composto da vari sottomercati, con i propri prezzi caratteristici, sui quali produttori e consumatori basano la propria pianificazione. I gestori dei sistemi di trasmissione, per garantire una fornitura costante di energia elettrica, utilizzano la capacità di bilanciamento, cioè l'ottimizzazione di utilizzo delle fonti energetiche, per correggere le deviazioni dovute ai processi stocastici. Essi sono il principale problema nella gestione delle fonti rinnovabili, in particolare l'eolico e il solare.

Il gruppo di bilanciamento e il sistema energetico di bilanciamento assicurano che l'offerta e la domanda siano impiegate nel modo più conveniente possibile. Il mercato elettrico premia così la produzione e la capacità.

Il Capacity Market è un meccanismo di regolazione della capacità del mercato, che prevede una remunerazione della potenza flessibile. I grandi impianti di produzione verranno pagati per la loro disponibilità a produrre energia in caso di necessità o, in alternativa, gli operatori della gestione della domanda saranno remunerati per la disponibilità a ridurre i consumi.

La generazione da fonti rinnovabili può dare un apporto importante, ma la non programmabilità e soprattutto la non disponibilità con continuità della risorsa naturale, come ad esempio l'irradiazione solare nelle ore serali, è tuttora un limite.

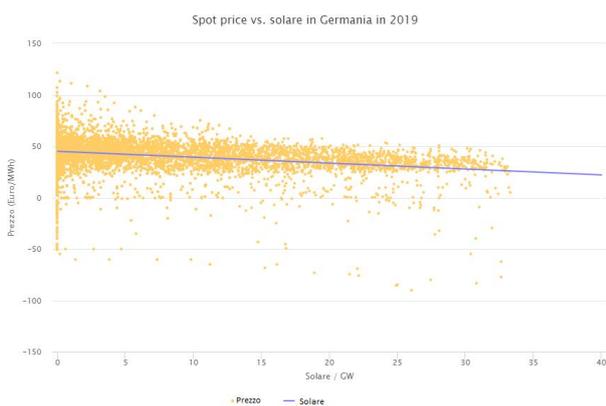


Figura 2.1: Costo dell'energia in Germania in relazione alla produzione solare nel 2019



Figura 2.2: Costo dell'energia in Germania in relazione alla produzione solare nel 2020

Abbiamo utilizzato i dati forniti da Fraunhofer-Institut für Solare Energiesysteme ISE [1] per individuare la relazione tra irraggiamento e costo dell'energia elettrica. Interpolando i punti raccolti, possiamo individuare una relazione di dipendenza lineare tra irraggiamento e costo:

- Fig.2.1: $\text{Prezzo} = 45.12\text{Euro/MWh} - 0.58\text{Euro/MWh} \bullet \text{Solare/GW}$
- Fig.2.2: $\text{Prezzo} = 39.04\text{Euro/MWh} - 0.77\text{Euro/MWh} \bullet \text{Solare/GW}$

In questa dipendenza lineare, tolti i valori più prossimi all'ordinata, dove entrano in gioco altre considerazioni, osserviamo che per un'intensità che tende a valori molto grandi, il costo ha come tendenza lo 0. Esso non potrà però mai essere raggiunto perchè altrimenti si arriverebbe ad una situazione che comporterebbe l'arbitraggio da parte di chi fa investimenti sulla fornitura energetica.

A differenza di altri beni, l'elettricità è molto difficile da immagazzinare, poichè le tecnologie di stoccaggio dell'elettricità su larga scala sono ancora limitate. Ciò significa che l'elettricità deve essere consumata nello stesso momento in cui viene prodotta.

Uno dei risultati è che il prezzo dell'elettricità oscilla durante il giorno. Il rapporto tra domanda e offerta, e quindi i prezzi, possono cambiare in tempi relativamente brevi sul mercato dell'elettricità.



Figura 2.3: Costo orario dell'energia in Germania (verde) e nel Nord Italia (giallo)

Anche la posizione influenza il costo dell'energia: in una regione più vicina all'Equatore, la produzione di energia dal solare sarà più costante durante l'anno rispetto ad una più a Nord. Oltre a ciò, ad influenzare il costo, è la domanda. Similmente alla produzione, più ci si sposta dall'Equatore, maggiore sarà il consumo di soglia. In questo modo abbiamo descritto come la domanda energetica sia determinata da molteplici fattori.

Nel grafico precedente si vede, quindi, il costo medio orario dell'energia come il risultato di domanda ed offerta nel periodo che intercorre tra il 01/01/2019 ed l'11/01/2019.

L'elettricità può essere scambiata con diversi anni di anticipo sul mercato dei futures: gli acquirenti utilizzano questi contratti a lungo termine per proteggersi dal rischio di aumento dei prezzi. Per questa certezza di pianificazione, pagano un premio, che i venditori registrano poi come ricavo aggiuntivo. I contratti a lungo termine garantiscono entrate ai produttori.

I volumi effettivi di consumo e generazione possono essere previsti in modo accurato, quindi il mercato spot a breve termine è costituito da due mercati con tempi di consegna diversi: il mercato del giorno prima e quello intraday. Gli operatori del mercato del giorno prima scambiano energia elettrica per il giorno successivo. Il mercato del giorno prima fornisce la maggior parte della fornitura elettrica necessaria, definendo così un prezzo di riferimento. L'intraday serve a gestire le fluttuazioni della domanda e del prezzo nel mercato, sfruttando un lead time, cioè il tempo che intercorre tra la richiesta, da parte del cliente, e la fornitura, da parte del fornitore, di energia elettrica, di 5 minuti.

Le energie rinnovabili si suddividono in stabili (biomassa) e variabili (eolico, solare, etc.).

Le fluttuazioni della produzione solare ed eolica dovute alle condizioni meteorologiche complicano la suddivisione dell'approvvigionamento energetico e hanno impatto significativo sulla volatilità dei prezzi.

Lavoreremo sulla produzione e sulla previsione dell'irraggiamento solare a partire dai dati raccolti da CAMS (Copernicus Atmosphere Monitoring Service), il quale fa parte del programma di Previsione di Medio-Lungo periodo dell'Unione Europea (ECMWF), e sull'analisi di K. Larsson, R. Green e F.E.

Benth [2]. Gli autori, partendo dai dati raccolti in 10 anni da CAMS in diverse località della Germania, sono riusciti a definire una funzione in grado di descrivere l'irraggiamento solare, considerando la componente stocastica.

L'analisi che effettueremo si concentrerà sulla città di Amburgo, in continuità con [2].

Capitolo 3

Il modello matematico

Lo spazio matematico sul quale lavoriamo è chiamato spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$.

Definizione 1. *Uno spazio campionario Ω è definito come l'insieme di tutti gli esiti possibili di un esperimento. Esso è composto da un numero finito di elementi ω e chiamiamo evento ogni sottoinsieme di Ω .*

Definizione 2. *\mathcal{F} è una σ -algebra di Ω degli eventi che vogliamo studiare. Le principali caratteristiche di una σ -algebra sono:*

- $\emptyset \in \mathcal{F}$
- Se $B \in \mathcal{F}$, allora anche il complementare $B^c \in \mathcal{F}$
- Se B_1, B_2, \dots è una collezione numerabile di set in \mathcal{F} , allora $\bigcup_{i=1}^{\infty} B_i \in \mathcal{F}$

Definizione 3. *\mathbb{P} è una funzione, detta misura di probabilità. Questa funzione è tale da assegnare ad ogni elemento della σ -algebra un valore compreso in $[0, 1]$: $\mathbb{P} : \mathcal{F} \rightarrow [0, 1]$.*

Definiamo che \mathbb{P} sia la probabilità a priori di essere nello stato aleatorio di nuvole. Il modello che andremo ad utilizzare è definito tramite gli stati aleatori nuvoloso ($m = 1$) e soleggiato ($m = 2$).

La probabilità di essere nello stato aleatorio nuvoloso è definita in funzione della stagione considerata ed è costante nell'intervallo stagionale. In questo modo definiamo la probabilità $\mathbb{P}_S = 0,3951$ di essere nello stato nuvoloso ($m = 1$) per l'estate ($t = 60, \dots, 303$) e similmente $\mathbb{P}_W = 0,5642$ per l'inverno ($t = 1, \dots, 59 \vee 304, \dots, 365$).

Fissiamo i tempi discreti crescenti $t_0, t_1, \dots, t_N \in \mathbb{R}$.

La σ -algebra \mathcal{F}_t rappresenta il set di informazioni disponibili al tempo t , cioè l'insieme degli eventi che sono avvenuti fino al tempo t .

Definizione 4. *La filtrazione $(\mathcal{F}_t)_{t \in \{t_0, t_1, \dots\}}$, cioè la famiglia di σ -algrebre tale che $\mathcal{F}_{t_0} \subseteq \mathcal{F}_{t_1} \subseteq \dots \subseteq \mathcal{F}_{t_N}, \mathcal{F}_{t_i} \subseteq \mathcal{F}$, rappresenta l'evoluzione dell'informazione o della conoscenza del sistema nel tempo.*

Le funzioni a valori reali definite in uno spazio campionario Ω e misurabili rispetto \mathcal{F} , sono dette variabili aleatorie.

Un processo stocastico Z a tempo discreto, è una famiglia di variabili aleatorie indicizzata del tempo, $(Z(t))_{t \in \tau}$, con $\tau \in \{t_0, t_1, \dots, t_N\}$, definita sullo spazio Ω .

Un processo stocastico Z è una funzione di 2 variabili.

- Per un istante fissato t , è una variabile casuale: $Z_t = Z_t(\omega), \omega \in \Omega$.
- Per un risultato casuale fissato $\omega \in \Omega$, è una funzione del tempo: $Z_t = Z_t(\omega), t \in \tau$.

Questa funzione è detta traiettoria del processo Z .

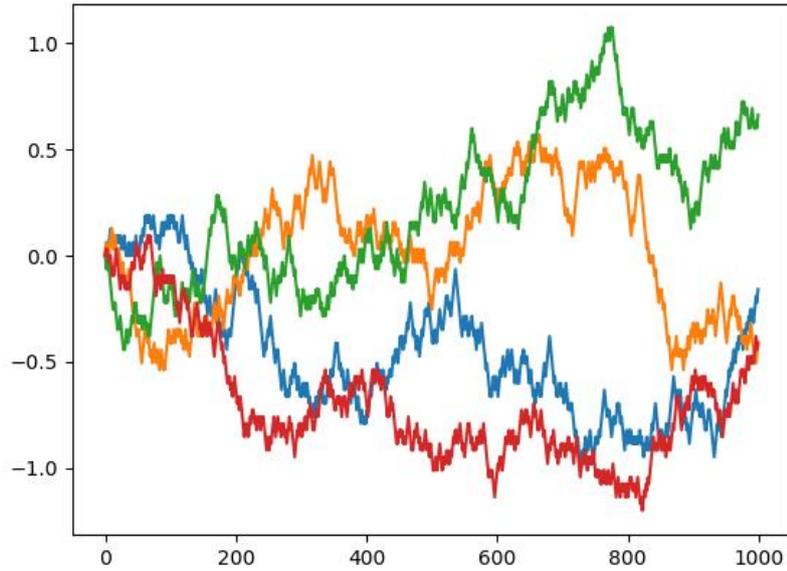


Figura 3.1: Esempio di processo stocastico a tempo

Il grafico in Fig.3.1, rappresenta un processo di Wiener, detto anche moto browniano, nel quale abbiamo messo in evidenza l'evoluzione a tempo discreto, differenziando le traiettorie. Il plot che abbiamo fatto con Python, evidenzia diverse traiettorie nello stesso intervallo d'azione del processo stocastico. Il processo stocastico $(Z(t))_{t \in \{t_0, t_1, \dots\}}$ è adattato alla filtrazione $\mathbb{F} := (\mathcal{F}_t)_{t \in \{t_0, t_1, \dots\}}$ se ogni Z_t è misurabile rispetto a \mathcal{F}_t .

Un processo stocastico a tempi discreti è una Catena di Markov se per $t = 1, 2, 3, \dots, n$, si ha che

$$\begin{aligned} \mathbb{P}\{Z_{t+1} = i_{t+1} \mid Z_t = i_t \mid \dots \mid Z_1 = i_1 \mid Z_0 = i_0\} = \\ = \mathbb{P}\{Z_{t+1} = i_{t+1} \mid Z_t = i_t\}. \end{aligned} \quad (3.1)$$

Nel caso che andremo ad analizzare, la catena di Markov è stazionaria una volta definita la stagione di definizione del dato ed è detta a stati finiti perchè lo spazio degli stati è finito.

Consideriamo la matrice $\mathbb{M} = (\mathbb{M}_{i,j})_{i,j \in \Omega}$ $i, j = \text{stati}$ $\Omega = \text{spazio campionario}$.

$\mathbb{M} = \begin{pmatrix} \mathbb{M}_{00} & \mathbb{M}_{01} & \dots & \dots \\ \mathbb{M}_{10} & \mathbb{M}_{11} & \dots & \dots \\ \dots & \dots & \mathbb{M}_{ij} & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix}$ è la matrice di transizione ad un passo associata alla catena omogena.

La matrice di transizione fornisce la probabilità di una variazione degli stati dopo un singolo passo temporale, definite in Tab. 4.2. Essa gode delle seguenti proprietà:

- $\mathbb{M}_{ij} \in [0, 1]$ $\forall i, j \in \Omega$
- $\sum_{j=0}^{\infty} \mathbb{M}_{ij} = 1$ $\forall i \in \Omega$

Definiamo ora la matrice di transizione dopo n -passi $\mathbb{M}^n = (\mathbb{M}_{i,j}^{(n)})_{i,j \in \Omega}$.

Le componenti sono

$$\mathbb{M}_{i,j}^{(n)} = \mathbb{P}\{Z_{n+m} = j \mid Z_n = i\} = \mathbb{P}\{Z_n = j \mid Z_0 = i\} \quad \forall n \geq 2 \quad (3.2)$$

Otteniamo la relazione nota come equazione di Chapman - Kolmogorov:

$$M_{i,j}^{n+m} = \sum_{k=0} M_{i,k}^n M_{k,j}^m$$

Questa relazione afferma che partendo da uno stato iniziale i , il processo si troverà nello stato j in $(n+m)$ transizioni temporali lungo la traiettoria che passa dallo stato k all' n -esima transizione temporale.

Nel nostro caso, si parla di catene di Markov a due stati, nuvoloso o soleggiato. Possiamo sfruttare questa informazione per poter esprimere tutti i casi in esame, conoscendo solo la probabilità di cielo coperto.

Capitolo 4

Caratterizzazione dell'irraggiamento

In una data località, sia $G(t)$ l'irraggiamento al tempo ($t = 1, \dots, 365$). $G(t)$ sarà la somma di una componente stagionale deterministica, $S(t)$, e di un processo stocastico $Z(t)$:

$$G(t) = S(t) + Z(t) \quad (4.1)$$

$G, S, Z : D \rightarrow \mathbb{R}, \quad D = [1, \dots, 365] \subset \mathbb{N}$.

Applicheremo questo tipo di modello ai dati raccolti nella città di Amburgo.

Località	Latitudine	Longitudine
Amburgo	53,4361	9,6311

Tabella 4.1: Localizzazione di Amburgo

I parametri individuati si basano sulla raccolta dati effettuata da 01-01-2019 a 31-12-2019.

Parametro	Stima	Parametro	Stima
a_0	0,2003	$\mu_{1,s}$	-1,0316
a_1	0,5993	$\mu_{2,s}$	0,6738
β_2	0,4270	$\nu_{1,s}^2$	0,3767
β_1	0,2259	$\nu_{2,s}^2$	0,2575
β_2	0,0605	q_s	0,3951
c_0	$1,8529 \times 10^4$	$\mu_{1,w}$	-0,7220
c_1	$-1,662 \times 10^4$	$\mu_{2,w}$	0,9349
c_2	$0,3330 \times 10^4$	$\nu_{1,w}^2$	0,1321
ω_0	0,1257	$\nu_{2,w}^2$	0,5729
ω_1	0,8342	q_w	0,5642
ω_2	0,0402		

Tabella 4.2: Parametri del modello

Poiché l'anno considerato è interno al set su cui si basa l'analisi in [2], non effettueremo nuovamente la stima dei parametri.

In tutta l'analisi, il tempo $t = 1, \dots, 365$, è una variabile discreta che rappresenta il giorno nell'anno solare, valutata alle ore 12.00 UTC.

4.1 Componente stagionale

La principale componente dell'elemento deterministico stagionale $S(t)$ è la radiazione extraterrestre $\Lambda(t)$, $\Lambda : D \rightarrow \mathbb{R}$, $D = [1, \dots, 365] \subset \mathbb{N}$. Essa è descritta come:

$$\Lambda(t) = K \left(1 + 0.033 \cos \left(\frac{360t}{365} \right) \right) \cos(\theta(t)), \quad (4.2)$$

dove il parametro K è una costante solare, definita $K = 1367(W/m^2)$. La localizzazione del punto studiato viene inserita nella funzione coseno, dove si utilizzano l'angolo di zenit solare, latitudine e longitudine, definiti nella Tab. 4.1. La caratterizzazione è meglio definita in Appendice A.

Alla radiazione extraterrestre è necessario aggiungere l'influenza delle condizioni atmosferiche per ottenere dei valori compatibili con i dati registrati. A partire dalla proposta di Hottel del 1976, [4], è possibile apportare delle modifiche in modo tale che la radiazione stagionale sia definita tramite:

$$S(t) = \Lambda(t) \left(a_0 + a_1 e^{\frac{-a_2}{\cos(\theta(t))}} \right), \quad (4.3)$$

dove i parametri $a_i, i = 0, 1, 2$ sono costanti di cielo sereno. In questo modello si considerano tutta una serie di fattori che possono influenzare il cielo, come vicinanza al mare, pulizia dell'aria, umidità. Così è possibile evidenziare la differenza tra radiazione extraterrestre e radiazione registrata, mettendo in risalto la componente stocastica della seconda.

Il modello parametrico di *clear sky* è un modello deterministico per la radiazione sotto l'ipotesi di assenza di nubi. In questo modo otterremo la base su cui agiscono i processi stocastici.

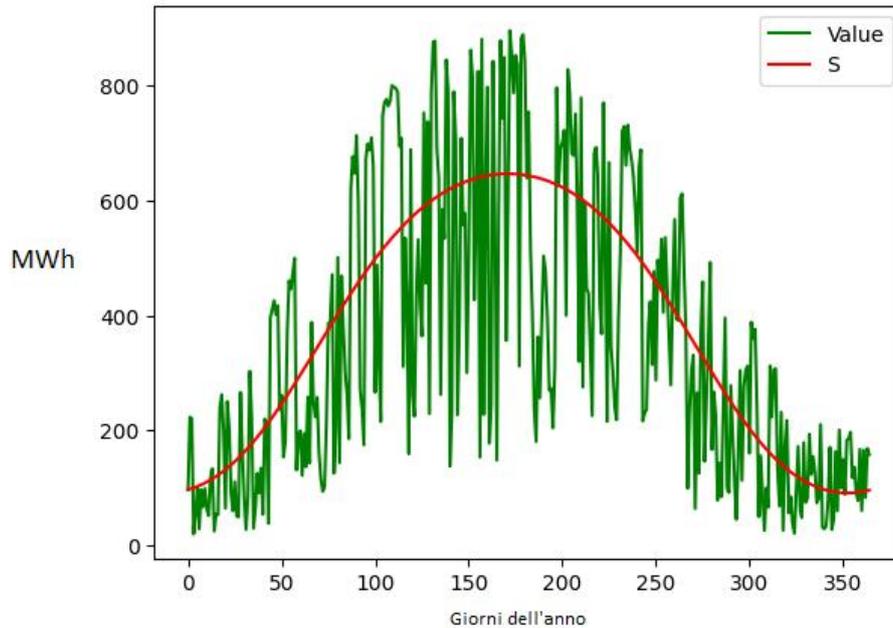


Figura 4.1: Irraggiamento nella componente stagionale $S(t)$ (rosso) ed irraggiamento totale registrato $G(t)$ (verde)

La figura precedente ci conferma che sia possibile individuare una funzione deterministica stagionale $S(t)$, sulla quale agiscono le fluttuazioni comportate dall'azione degli elementi stocastici.

Il modello di *clear sky* è calibrato sui dati di radiazione sotto le condizioni di *clear sky*.

4.2 Componente stocastica

La componente stocastica $Z(t)$ è modellizzabile come un processo autoregressivo $AR(p)$. Si usa un $AR(p)$ perchè è possibile unire un fenomeno di autoregressione, così da avere una memoria di tutti gli

avvenimenti precedenti, e un *white noise*.

Un fatto importante da sottolineare è che $p = 1$. La conseguenza di ciò è che non si manifestano andamenti ciclici, cioè non è possibile individuare un valore medio.

$$Z(t) = \sum_{i=1}^p \beta_i Z(t-i) + u(t) \quad (4.4)$$

con coefficienti costanti β_i , $i = 1, \dots, d$, e con $u(t) = \sigma(t)e(t)$ *white noise* del processo,

Definizione 5. *Il white noise è un processo stocastico che possiede momenti almeno fino al second'ordine; esso è costante nel tempo (quindi il processo è stazionario).*

Posso definire l'elemento t-esimo del *white noise*, α_t , tale che

$$\begin{aligned} \mathbb{E}[\alpha_t] &= 0 \\ \text{Var}[\alpha_t] &= \mathbb{E}[\alpha_t^2] = \sigma^2 \\ \gamma_k &= 0 \quad |k| > 0 \quad \text{covarianza} \end{aligned}$$

Esso è quindi un processo composto di un numero infinito di variabili casuali a media zero e varianza costante; queste variabili casuali, inoltre, sono tutte scorrelate l'una all'altra.

La varianza di $u(t)$ è eteroschedastica, poichè contiene sia gli effetti stagionali (in inverno si hanno maggiori nuvole) che quelli ARCH.

Definizione 6. *Un set di dati si definisce eteroschedastico se presenta varianze differenti.*

Un approfondimento sui processi di autocorrelazione è presentato in Appendice C.

4.2.1 Volatilità condizionata

Definiamo $\sigma(t)$ come la volatilità condizionata variabile nel tempo. Essa è condizionata ad \mathcal{F}_t .

La volatilità variabile nel tempo si riferisce alle fluttuazioni della volatilità in diversi periodi di tempo.

La varianza condizionata è

$$\sigma^2(t) = \sigma_S^2(t) \sigma_G^2(t). \quad (4.5)$$

$\sigma, \sigma_S, \sigma_G : D \rightarrow \mathbb{R}$, $D = [1, \dots, 365] \subset \mathbb{N}$.

La componente stagionale, $\sigma_S^2(t)$, è data da una serie di Fourier troncata al terzo termine [3]:

$$\sigma_S^2(t) = c_0 + c_1 \cos\left(\frac{2\pi t}{365}\right) + c_2 \sin\left(\frac{2\pi t}{365}\right) \quad (4.6)$$

ed è molto utilizzata per descrivere la volatilità stagionale dei modelli, con applicazioni alla temperatura, al vento e ai prezzi dell'energia. Si potrebbe considerare una serie con più termini, ma il contributo dopo il terzo è eliminabile perchè non significativo.

La componente autoregressiva GARCH $\sigma_G^2(t)$ prende la forma standard di un processo GARCH($p = 1$, $q = 1$), con i residui $v(t) = \frac{u(t)}{\sigma_S(t)}$ normalizzato dalla componente stagionale

$$\sigma_G^2(t) = \omega_0 + \omega_1 \sigma_G^2(t-1) + \omega_2 v^2(t-1). \quad (4.7)$$

4.2.2 Driving noise process

Caratterizziamo ora il disturbo $e(t)$. L'asimmetria dei dati sull'irradiamento ha un andamento stagionale. Le definizioni del periodo estivo ed invernale, provengono da osservazioni empiriche.

La distribuzione dei dati della radiazione solare suggerisce di descrivere gli stati di cielo limpido o nuvoloso mediante una distribuzione bimodale asimmetrica e stagionale:

$$e(t) = d(t) \varepsilon_S(t) + (1 - d(t)) \varepsilon_W(t) \quad (4.8)$$

$e : D \longrightarrow \mathbb{R}$, $D = [1, \dots, 365] \subset \mathbb{N}$
con la funzione indicatrice

$$d(t) = \begin{cases} 1 & \text{se } Summer(S) \\ 0 & \text{se } Winter(W) \end{cases}$$

Le $\{\varepsilon_k(t)\}_{k \in \{S, W\}}$, processi gaussiani, $\varepsilon_k : D \longrightarrow \mathbb{R}$, $D = [1, \dots, 365] \subset \mathbb{N}$. Le densità di probabilità per ε_k , $k \in \{S, W\}$, sono date da

$$f_k(x) = q_k f_{1,k}(x) + (1 - q_k) f_{2,k}(x) \quad x \in \mathbb{R}, \quad k \in \{S, W\} \quad (4.9)$$

le $f_{m,k}(x)$ sono le densità di variabili casuali gaussiane con media $\mu_{m,k}$ e varianza $\nu_{m,k}^2$ per $m = 1, 2$ e $k \in \{S, W\}$.

Il parametro q_k è la probabilità a priori di essere nel primo stato $m = 1$ nella stagione k .

Usiamo la notazione $GM(\mu_1, \mu_2, \nu_1^2, \nu_2^2, q)$ per denotare le miscele delle distribuzioni gaussiane di 2 componenti.

In questo modo si ha

$$\varepsilon_k(t) \sim GM(\mu_{1,k}, \mu_{2,k}, \nu_{1,k}^2, \nu_{2,k}^2, q_k) \quad k \in \{S, W\}. \quad (4.10)$$

Definisco quindi i momenti delle distribuzioni Gaussiane relative a $\varepsilon_k(t)$ come:

$\mu_k(t) = \mathbb{E}[\varepsilon_k(t)]$	Media
$\Sigma_k^2(t) = \text{Var}[\varepsilon_k(t)]$	Varianza
$\Omega_k(t) = \mathbb{E}[(\varepsilon_k(t) - \mathbb{E}[\varepsilon_k(t)])^3]$	Terzo Momento Centrato

che si traducono nel caso considerato in:

$$\begin{aligned} \mu_k(t) &= q_k \mu_{1,k} + (1 - q_k) \mu_{2,k} \\ \Sigma_k^2(t) &= q_k (\nu_{1,k}^2 + \lambda_{1,k}^2) + (1 - q_k) (\nu_{2,k}^2 + \lambda_{2,k}^2) \\ \Omega_k(t) &= q_k (3\lambda_{1,k} \nu_{1,k}^2 + \lambda_{1,k}^3) + (1 - q_k) (3\lambda_{2,k} \nu_{2,k}^2 + \lambda_{2,k}^3) \end{aligned}$$

con $\lambda_{m,k} = \mu_{m,k} - \mu_k$ per $m = 1, 2$. L'asimmetria del terzo momento di $\varepsilon_k(t)$ è data da $\varepsilon_k(t) = \frac{\Omega_k}{\Sigma_k^3}$.

Considerando l'indipendenza delle 2 distribuzioni $\varepsilon_k(t)$, è possibile specificare ulteriormente i momenti di $e(t)$, riprendendo la descrizione fatta per i processi di miscele gaussiane $\varepsilon_k(t)$:

$$\begin{aligned} \mathbb{E}[e(t)] &= d(t) \mu_S + (1 - d(t)) \mu_W \\ \text{Var}[e(t)] &= d(t) \Sigma_S^2 + (1 - d(t)) \Sigma_W^2 \\ \mathbb{E}[(e(t) - \mathbb{E}[e(t)])^3] &= d(t) \Omega_S + (1 - d(t)) \Omega_W \end{aligned}$$

Le variabili $\varepsilon_S(t)$, $\varepsilon_W(t)$ seguono distribuzioni di miscele gaussiane con 2 componenti, in modo da avere le rappresentazioni stocastiche [2],

$$\varepsilon_k = B_k Y_{1,k} + (1 - B_k) Y_{2,k} \quad k \in \{S, W\} \quad (4.11)$$

nella quale B_k è una variabile aleatoria che segue la distribuzione di Bernoulli con parametro $q_k \in (0, 1)$, cioè $\mathbb{P}(B_k = 1) = q_k$, e le componenti $Y_{i,k}$, $i = \{1, 2\}$, sono le variabili casuali gaussiane.

Si caratterizzano le componenti gaussiane come le condizioni, rispettivamente, di cielo nuvoloso, Y_1 , e cielo sereno, Y_2 . Definisco, quindi, q_k la probabilità a priori di avere cielo nuvoloso, mentre $(1 - q_k)$ la probabilità complementare di osservare cielo limpido, poichè ho solo 2 stati possibili per K .

In questo modo è possibile prevedere che i residui normalizzati di $e(t)$, avranno 2 picchi di diversa intensità.

Per il periodo invernale la distribuzione ha una *skewness* positiva ed il picco a sinistra più intenso di quello a destra. Questo indica che, nel periodo invernale, lo stato di nuvolosità è dominante ed ha una

maggior probabilità di manifestarsi, Fig.4.2.

Viceversa, nel periodo estivo, la *skewness* è negativa ed il picco più grande è quello di destra, con la probabilità che si sposta verso lo stato di cielo sereno, Fig.4.3.

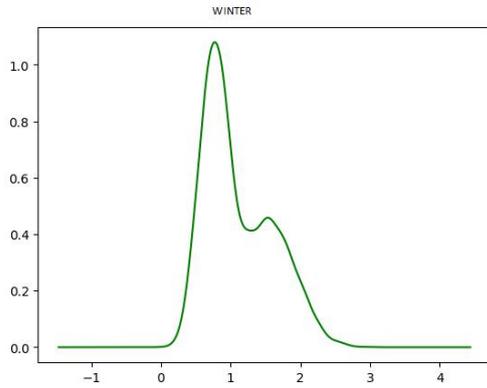


Figura 4.2: Esempio residui periodo invernale

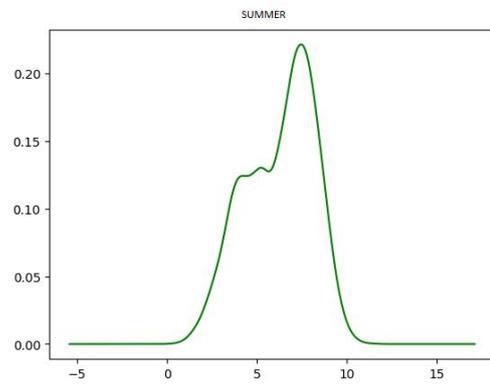


Figura 4.3: Esempio residui periodo estivo

4.2.3 Irraggiamento previsto

L'irraggiamento previsto, tramite questo modello, al tempo $t+1$ basato sull'informazione \mathcal{F}_t al tempo t , è data dal valore atteso condizionato $\hat{G}(t+1) = \mathbb{E}[G(t+1) | \mathcal{F}_t]$.

Sfruttando le proprietà sul *white noise*, si ottiene che

$$\begin{aligned} \hat{G}(t+1) &= S(t+1) + \sum_{i=1}^d \beta_i Z(t+1-i) + \sigma(t+1)e(t+1) \\ &= S(t+1) + \sum_{i=1}^d \beta_i Z(t+1-i). \end{aligned} \quad (4.12)$$

Infatti $\mathbb{E}[\sigma(t+1)e(t+1) | \mathcal{F}_t] = 0$ poichè $e(t+1)$ è indipendente da \mathcal{F}_t .

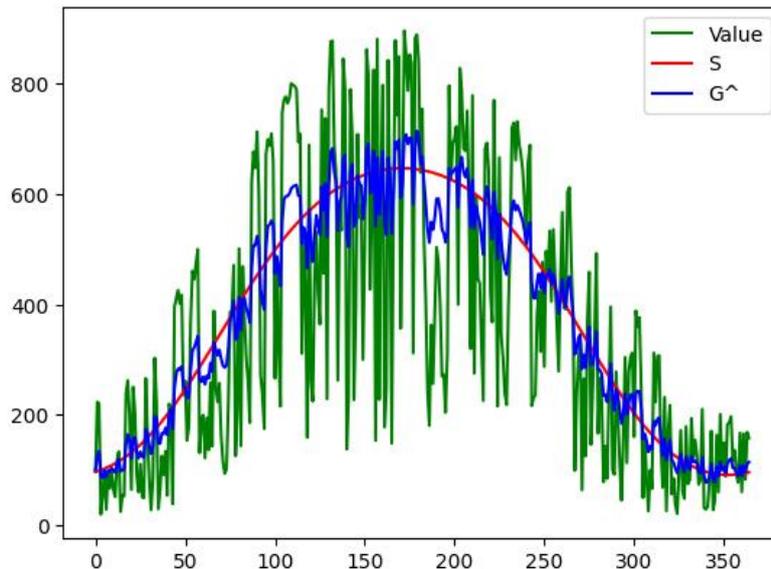


Figura 4.4: Plot dei dati registrati (verde), dei dati previsti dalla simulazione col modello (blu) e della componente stagionale (rosso)

Come possiamo vedere, il grafico precedente rappresenta l'andamento dei valori raccolti d'irraggiamento.

Osservando il picco nei giorni centrali dell'anno ed i minimi nelle regioni più esterne, possiamo osservare come la stagionalità sia rispettata. Una conferma della bontà del modello descrittivo ci giunge dall'osservare che la periodicità di $T = 366$ giorni.

La funzione d'irraggiamento predetto, $\widehat{G}(t)$, segue l'andamento della funzione stagionale deterministica $S(t)$, ma considera anche i contributi della serie autoregressiva.

Possiamo infine vedere come la rimozione del *white noise* agisca riducendo i contributi stocastici.

Capitolo 5

Applicazione al mercato dell'energia solare del modello statistico

Il modello ipotizzato per la descrizione dell'irraggiamento solare può essere sfruttato per fare delle previsioni sui consumi a breve termine, con riferimento al consumo al più settimanale, oppure è possibile andare a fare delle previsioni a lungo termine di consumo. Minore è la distanza temporale della previsione, maggiore sarà la corrispondenza tra consumi reali e consumi previsti.

Affinando la metodologia appena presentata si potrebbe giungere ad un'ottimizzazione del consumo di energia proveniente da fonti non rinnovabili.

L'energia solare prodotta dalla cella fotovoltaica per metro quadro, $h(G(t), T(t))$, è definita in funzione dell'irraggiamento solare G e della temperatura operativa, T , attraverso l'espressione

$$h(G, T) = \kappa_1 G [1 - \kappa_2 (T - T_{rif})] \quad (5.1)$$

l'espressione è in funzione del giorno t , valutato alle 12.00 UTC.

È necessario ora introdurre il mezzo che permette il passaggio tra irraggiamento solare e corrente elettrica: il pannello fotovoltaico.

Ipotizziamo di avere un irraggiamento solare di 1000 W/m^2 , di avere un parco solare di 50000 m^2 che produca 10 MWh di elettricità all'ora nella situazione di temperatura costante di 25°C .

Definiamo ora i parametri materiali delle celle fotovoltaiche:

- $\kappa_1 \in \mathbb{R}_{>0}$ è definito come l'efficienza di trasformazione dell'energia solare in energia elettrica alla temperatura di riferimento $T_{rif} = 25^\circ\text{C}$. Questo valore oscilla tra il 10% e il 20% per le celle prodotte per uso commerciale in silicio, con punte del 32.5% in laboratorio.
- κ_2 si definisce coefficiente di temperatura ed è determinato dalla tecnologia della cella. I suoi valori tipici sono compresi tra 0.0035 e 0.005.

Nell'analisi i parametri saranno definiti con $\kappa_1 = 0.2$ e $\kappa_2 = 0.005$.

Riprendendo le ipotesi fatte del parco solare, definiamo la produzione di energia (in MWh) come

$$V(t) = 0.05 \cdot h(G(t), T(t)), \quad V : D \rightarrow \mathbb{R}, \quad D = [1, \dots, 365] \subset \mathbb{N}$$

La temperatura, considerata un parametro deterministico, ha un contributo ignorabile, visti i valori tipici. Sarà come mettersi nella situazione ideale di lavoro a 25°C

La funzione di ripartizione della produzione energetica $V(t+1)$, condizionata al tempo t , è

$$F_t(x) = \mathbb{P}_t(V(t+1) \leq x) = \mathbb{P}_t(G(t+1) \cdot H(t+1) \leq x) \quad (5.2)$$

avendo definito $H(t) = 0.05\kappa_1 [1 - \kappa_2 (T(t) - T_{rif})]$ e $\mathbb{P}_t(V(t+1)) = \mathbb{P}(V(t+1) | \mathcal{F}_t)$.

Applicando quindi le definizioni di stagioni, il modello è definito come:

$$F_t(x) = q_k F_{1,k}(y_t(x)) + (1 - q_k) F_{2,k}(y_t(x)) \quad k = \begin{cases} S & \text{se } d(t+1) = 1 \\ W & \text{altrimenti} \end{cases} \quad (5.3)$$

$F : D \rightarrow \mathbb{R}, \quad D = [1, \dots, 365] \subset \mathbb{N}$,
nel quale $y_t(x)$ è il punto

$$y_t(x) = \frac{1}{\sigma(t+1)} \left(\frac{x}{H(t+1)} - S(t+1) - \sum_{j=1}^p \beta_j Z(t+1-j) \right) \quad (5.4)$$

e $F_{m,k}(x)$ sono le funzioni di ripartizione gaussiana corrispondenti alle componenti per $m = 1$ e $m = 2$.

5.0.1 Production-at-risk P@R

Andremo ora a definire il rischio di produzione (P@R).

Definizione 7. *Definiamo il P@R (Production-at-Risk) come un quantile della distribuzione per il volume di produzione del giorno prima.*

Applicheremo questo modello per il parco solare descritto precedentemente.

Per un valore dato $\alpha \in (0, 1)$, il P@R del giorno precedente al tempo t , definito come $\lambda(t, \alpha)$, soddisfa la relazione:

$$F_t(\lambda(t, \alpha)) = \mathbb{P}_t(V(t+1) \leq \lambda(t, \alpha)) = \alpha \quad (5.5)$$

Il valore di P@R, $\lambda(t, \alpha)$, deve essere ottenuto numericamente dall'Eq.(5.5). Ciò può essere fatto tramite simulazioni o attraverso algoritmi di approssimazione numerica.

Scegliamo di porre $\alpha = 0,05$, in modo da studiare il P@R al 5%.

Per la soluzione abbiamo compiuto simulazioni numeriche sulle variabili aleatorie $Y_{m,k}$, contenute nei processi gaussiani $\varepsilon_k(t)$, Eq.(4.11), tali per cui, ottenuto un valore per il giorno t , per l'elaborazione del giorno $t+1$ ci si basi su quel valore, anzichè compiere nuovamente le simulazioni in maniera iterata. Questa scelta è dettata dall'impossibilità di avere un calcolatore con un potenza di calcolo illimitata.

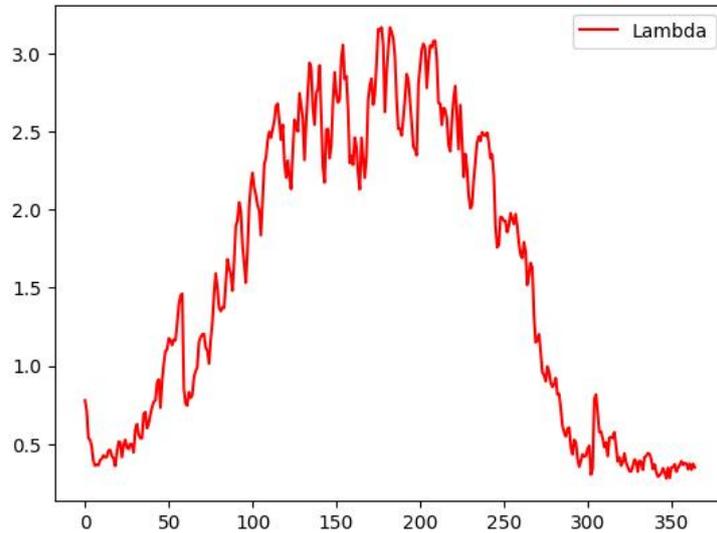


Figura 5.1: Plot delle λ ottenute dalla simulazione nell'anno

Come si vede dal grafico, il plot segue l'andamento stagionale previsto nel passare dei giorni.

Il plot di λ , cioè il valore per ciascun t del P@R è sempre positivo, come ci aspettavamo. Esso rappresenta la P@R del giorno precedente. La P@R ha una chiara dipendenza dalla volatilità e dalla forma della distribuzione del valore d'irraggiamento solare.

I primi giorni hanno un plot delle λ che potrebbe sembrare privo di senso e in disaccordo con quanto sembra suggerire l'andamento degli ultimi giorni dell'anno. Il grafico assume quella forma poichè nei

primi giorni l'influenza della serie autoregressiva è solo parzialmente corretta in quanto mancano i dati provenienti dall'anno precedente. Abbiamo scelto di non analizzarli per poter evidenziare l'importanza della serie ARCH e per avere coerenza con l'analisi sull'anno. Come previsto, già dopo $t = 5$, l'andamento diventa regolare ed il contributo parziale viene assorbito da quello dei giorni successivi.

Capitolo 6

Proposta del modello statistico

A partire dal punto considerato, abbiamo preso 4 punti ad una distanza di circa 11 km. Poichè le distanze sono molto ridotte, abbiamo ipotizzato che i parametri, individuati in Tab. 1, siano costanti. Sempre sfruttando la distanza ridotta, possiamo approssimare gli effetti dovuti alla curvatura terrestre.

Località	Latitudine	Longitudine
Nord	53,5361	9,6311
Sud	53,3361	9,6311
Est	53,4361	9,7311
Ovest	53,4361	9,5311

A partire da questi 4 punti, abbiamo, similmente al caso analizzato, ricavato in una prima fase i differenti processi stocastici: $Z_i(t) = G_i(t) - S(t)$ $i = N, S, E, O$

In questo modo riusciamo ad ottenere la funzione di previsione dell'irraggiamento sui 4 punti $\hat{G}(t)$, tramite la ricostruzione della funzione.

Sfruttando la semplificazione di aver i coefficienti costanti, l'analisi si concentrerà sui processi stocastici.

Suddividiamo ora l'anno nelle 2 stagioni ed andiamo ad analizzare le diverse località, in modo da ricavare medie e varianze (inverno: $t = 1, \dots, 59 \vee 304, \dots, 365$; estate: $t = 60, \dots, 303$):

Nord	Media	Varianza	Sud	Media	Varianza
Estate	-45.8449	39212.1020	Estate	-34.3586	40532.6436
Inverno	6.61055	9019.7745	Inverno	8.97304	8430.4371
Est	Media	Varianza	Ovest	Media	Varianza
Estate	-44.1051	39449.9037	Estate	-41.9982	40801.5879
Inverno	7.9136	8827.4799	Inverno	9.0963	8771.595

In questo modo abbiamo ottenuto un array di 4 valori per punto.

Sfruttando il fatto di essere su distanze molto piccole, passiamo a considerare i valori ottenuti come se fossero gli stessi su tutti i punti di una circonferenza di raggio uguale alla distanza dal centro.

Pensiamo, quindi, come se ci fosse una distribuzione centrata in 0 che decresce in maniera uniforme. Abbiamo in questo modo, tramite una semplificazione che fa perdere di granularità sul modello iniziale, ottenuto un'analisi valida sulle brevi distanze.

Creiamo ora le misture stagionali. Ci aspettiamo di ottenere delle distribuzioni quasi normali.

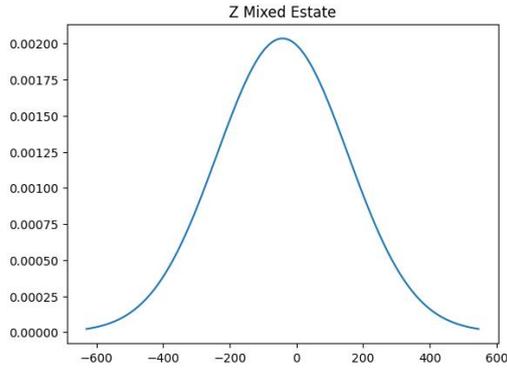


Figura 6.1: Media: -41.5767, Varianza: 38407.2449

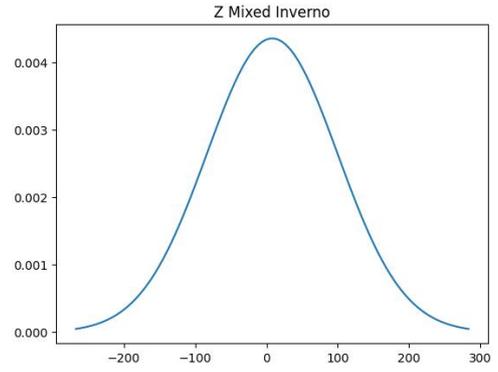


Figura 6.2: Media: 8.1484, Varianza: 8410.9705

Riusciamo in questo modo a ricomporre la funzione di irraggiamento, necessaria per determinare la P@R.

Per evitare il cambio di stagione in modo repentino, abbiamo introdotto una funzione che in base al giorno, considera i contributi delle stagioni come una mistura.

Immaginando che i solstizi siano gli unici giorni in cui ho solo il contributo della stagione di riferimento, giornalmente ho somme di percentuali di stagioni. Abbiamo ipotizzato che il passaggio da una stagione all'altra sia lineare.

Definiamo il coefficiente estivo con $\xi_S(t) = 1 - \left| \frac{t-182,5}{182,5} \right|$. Definiamo il coefficiente invernale con $\xi_W(t) = 1 - \xi_S(t)$. Per definizione abbiamo $\xi_k(t) \in [0, 1] \quad k \in \{S, W\}$

Per ovviare ai fattori negativi, che si manifestano nei giorni più invernali, cioè quelli più vicini al solstizio invernale, abbiamo mappato la componente aleatoria rispetto al valore stagionale. In questo modo abbiamo decrescite differenti, relazionate al giorno t :

$$\gamma_k(t) = \frac{S_k(t)}{\widehat{S_k(t)}}$$

dove γ è il fattore di mappatura, $S(t)$ il valore stagionale del giorno e $\widehat{S(t)}$ è il valore stagionale medio della stagione.

Otteniamo così valori sempre positivi, altrimenti avremmo dei risultati senza senso.

Definiamo, infine, il quantile al 5% della componente aleatoria $Z(t)$ delle 2 stagioni $\rho_k(t) \quad k \in \{S, W\}$. Il risultato che otteniamo è l'I@R (Irradiation-at-Risk),

Definizione 8. *I@R (Irradiation-at-Risk) è l'irraggiamento previsto nel quale si ha nel 95% dei casi un irraggiamento superiore a quello registrato.*

$$\text{I@R5}(t) = S(t) + \rho_S(t) \gamma_S(t) \xi_S(t) + \rho_W(t) \gamma_W(t) \xi_W(t) \quad (6.1)$$

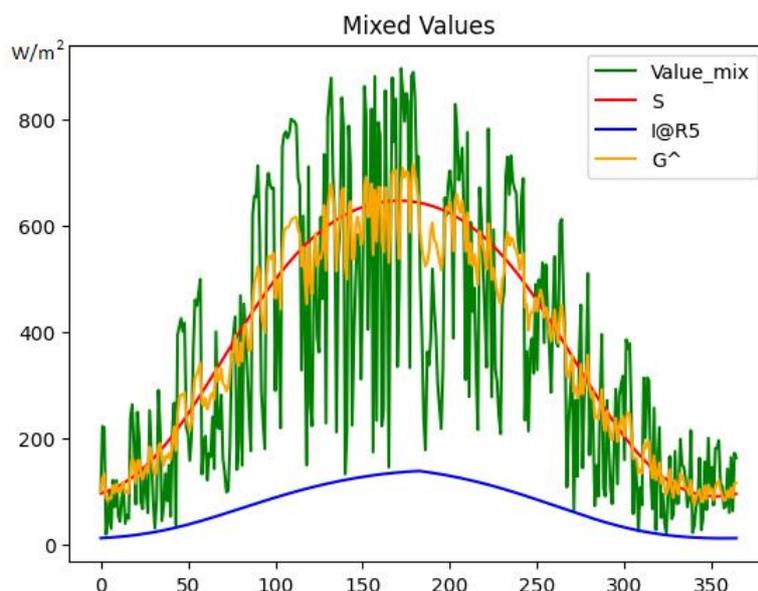


Figura 6.3: Plot dei valori di irraggiamento della mistura (verde), dell'irraggiamento previsto (giallo), della componente stagionale (rosso) e dell'I@R (blu)

In linea con quanto osservato da [2], notiamo che anche con questa modellizzazione le fluttuazioni sono maggiori nel periodo estivo.

Tramite un fattore moltiplicativo, dovuto al parco solare precedentemente descritto, otteniamo la P@R. Si pone ora la necessità di confronto delle P@R.

Mentre quella iniziale si basa su processi aleatori frutto di misture gaussiane, la seconda si basa su Z gaussiana.

Non sapendo di che distribuzione si tratta, abbiamo effettuato un riordino dei valori, per poi effettuare il taglio ad $\alpha = 5\%$. In questo modo evito il problema di sapere a priori la distribuzione.

Ora è possibile confrontare le P@R ipotizzate dai differenti modelli.

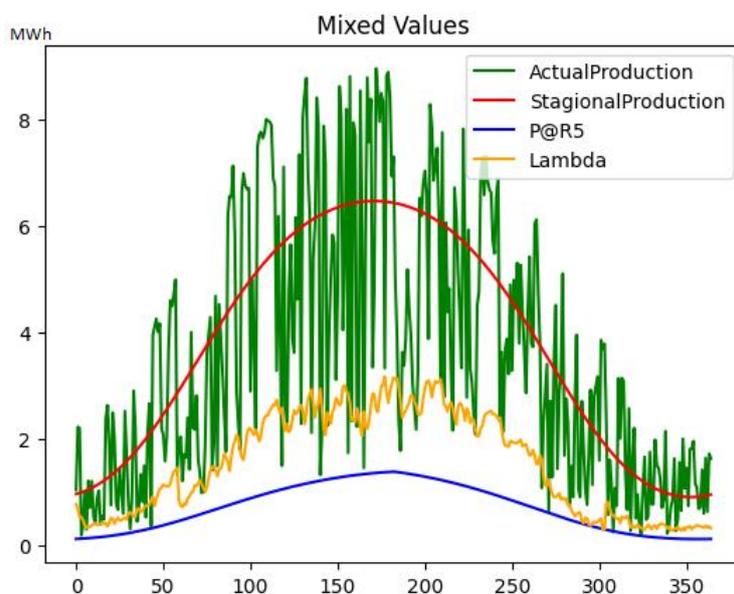


Figura 6.4: Plot dei valori di produzione della mistura a partire da $G(t)$ (verde), della produzione della componente stagionale (rosso), P@R proposta in [2] e P@R proposta qui sviluppata ed applicata alla media dei 4 punti di calibrazione (blu)

In questo grafico si osserva come, a seguito delle fluttuazioni dovute ai processi stocastici, la P@R proposta dagli autori risponda in modo coerente. Ciò non si ritrova nel modello alternativo nel calcolo della P@R.

D'altra parte, la nostra proposta ci conferma, nelle brevi distanze, che il parallelismo tra l'applicazione puntuale e l'analisi sulle distanze è possibile.

Andando a confrontare le $\lambda(t, \alpha)$, osserviamo le differenziazioni con cui sono state pensate, ciò si riflettono sul risultato.

Il plot blu parte da considerazioni esclusivamente statistiche quindi, a prezzo di avere una minor precisione puntuale, abbiamo ottenuto una maggiore stabilità. Questa stabilità ha validità per ogni anno, a patto di considerare i parametri di partenza sufficientemente stabili negli anni.

D'altra parte, la $\lambda(t, \alpha)$ proposta in giallo, risponde meglio alle fluttuazioni che vengono registrate. La caratteristica principale è quella di essere un modello basato sulle simulazioni. In questo modo, però, la specificità dai parametri permette l'analisi solo su un anno specifico, valida a posteriori.

Capitolo 7

Conclusioni

In questo lavoro di tesi, siamo partiti dalla verifica del modello proposto da [2] per la previsione dell'irraggiamento solare come somma di componenti stagionali (deterministiche) e di processi stocastici. Nell'analisi abbiamo modellizzato il tempo attraverso una variabile discreta. Un possibile sviluppo ulteriore potrebbe essere quello di un'analisi a variabile continua, in modo da non descrivere più solo il picco d'irraggiamento, ma l'irraggiamento totale giornaliero.

Tramite l'analisi effettuata siamo anche stati in grado di poter definire il rischio di produzione.

Ci siamo chiesti quanto questa proposta fosse legata alla definizione del punto analizzato (Amburgo) e quanto si potesse effettuare una analisi a partire da punti equidistanti dal centro. Il risultato ottenuto è quello di poter affermare una ottima compatibilità con il modello. A livello produttivo, possiamo quindi effettuare la definizione dei parametri specifici per punti di interesse e poi, invece di farlo nuovamente per ogni punto, possiamo su brevi distanze applicare i risultati ottenuti, senza la perdita eccessiva di compatibilità.

Infine abbiamo proposto un metodo alternativo per la definizione di P@R.

Delle componenti interessanti che si potrebbero prestare ad approfondimenti, sono il fattore di mappatura, da noi proposto, e la definizione di miscela di stagioni. Nel caso in cui si riuscissero a definire anche le stagioni di passaggio dall'estate all'inverno, si acquisirebbe una caratterizzazione ulteriore. In questo modo potremmo descrivere l'irraggiamento come miscela di più di 2 variabili aleatorie.

Un caratteristica che rende questo modello non applicabile ovunque, in accordo con quello di partenza in [2], è la semplificazione di soleggiato/nuvoloso come un tempo atmosferico unico. In questo modo, l'applicazione è particolarmente efficace nelle regioni in cui la differenza tra irraggiamento in condizione di sole e in condizione di nuvole, è bassa.

La perdita di granularità sul modello è stata necessaria per la non disponibilità del programma utilizzato dagli autori.

Come detto, il nostro modello ha una base statistica. In questo modo esso ha una maggiore generalizzazione ed applicabilità rispetto a quanto proposto in chiave simulativa.

Tramite quanto siamo stati in grado di proporre, possiamo osservare che non ammettiamo situazioni in cui la produzione energetica sia inferiore al valore di soglia, come si vede in Fig.6.4 con I@R5.

Quindi, se è vero che il modello di partenza è più vicino alla soglia prodotta, è anche vero che ammette situazioni in cui è prevista una sovrastima della produzione energetica.

Il nostro modello è meno vicino alla produzione effettiva, ma non ammette situazioni in cui è presente una sovrastima dell'energia prodotta.

In questo modo, siamo sempre in grado di fornire almeno il minimo di energia prevista. A livello finanziario, siamo così in grado di poter mettere un tetto massimo al costo dell'energia solare. Tutte le oscillazioni prodotte dai processi stocastici non potranno far aumentare il costo dell'energia, ma solo confermarlo o diminuirlo.

Possibili sviluppi futuri risiedono nella capacità di maggiore definizione delle variabili aleatorie. Ampliando la descrizione avremo la possibilità di definire con maggiore precisione l'irraggiamento e, conseguentemente, la produzione di energia solare.

Appendice A

Descrizione dell'angolo solare orario

La radiazione extraterrestre è definita mediante l'utilizzo dell'angolo di zenit alla coordinata selezionata. La funzione $\theta(n, t)$ è implicitamente determinata dall'angolo di zenit, ottenuto da

$$\cos(\theta(n, t)) = \cos(lat) \cos(\omega(n, t)) \cos(\delta(n)) + \sin(lat) \sin(\delta(n)) \quad (A.1)$$

dove $\delta(n)$ è l'inclinazione terrestre, mentre $\omega(t)$ è angolo dell'ora solare.

Definisco $\omega(n, t) = 15^\circ(LST(t) - 12)$ con $LST(t)$ *local solar time*. Quest'ultimo è ottenuto dal *local time* (LT) e dalla *time correction* (TC) $LST = LS + \frac{TC}{60}$.

La correzione temporale TC è calcolata come

$$TC = 4(LSTM - lon) + E(n)$$

$LSTM = 15^\circ U$, U è l'offset rispetto all'ora UTC, lon è la longitudine ed $E(n)$ è definita funzione del tempo, ottenuta da

$$E(n) = 229.2(0.000075 + 0.001868 \cos(B(n)) - 0.032077 \sin(B(n)) - 0.014615 \cos(2B(n)) - 0.04089 \sin(2B(n)))$$

con $B(n) = (n - 1) \frac{360}{365}$, n è il numero del giorno nell'anno. Questa variabile può essere considerata come continua.

Per quanto riguarda l'inclinazione $\delta(n)$, essa viene definita direttamente come

$$\delta(n) = 23.45^\circ \sin\left(\frac{360(284 + n)}{365}\right)$$

Appendice B

Teorema di Kolmogorov

Consideriamo la σ -algebra $G \in \mathcal{F}$ e la funzione $f : \Omega \rightarrow \mathbb{R}$ tale che $f \in \mathbf{L}^1(\Omega, \mathcal{F}, \mathbb{P})$.

$\mathbb{E}[f|G]$ tale funzione è detta valore atteso condizionato di f rispetto alla σ -algebra G .

Il valore atteso condizionato di una variabile aleatoria corrisponde alla migliore stima della variabile, nota l'informazione contenuta in G .

Nel caso in cui G sia banale, si ha $\mathbb{E}[f|G] = \mathbb{E}[f]$. Questa situazione è quella nella quale si ha l'informazione minima possibile, cioè uno stato misto. La stima cercata deve essere misurabile rispetto a G , ma è costante, poichè banale.

Viceversa, nel caso in cui ci sia una seconda σ -algebra $F = G$, sono nella situazione di massima informazione possibile, cioè uno stato puro. In questo caso si ha $\mathbb{E}[f|F] = f$.

Seguono ora alcune importanti proprietà del valore atteso condizionato:

- Se f è G -misurabile, $\mathbb{E}[f|G] = f$
- Se f e G sono indipendenti, $\mathbb{E}[f|F] = \mathbb{E}[f]$
- $\mathbb{E}[\mathbb{E}[f|G]] = \mathbb{E}[f]$

Appendice C

Processi di autocorrelazione

Un processo autoregressivo di ordine p , definito come $AR(p)$, per una variabile osservata y_t prende la forma del tipo:

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + u_t \quad (C.1)$$

dove u_t è il termine di disturbo.

Il *white noise* u_t che soddisfa $u_t = \xi + \alpha_1 u_{t-1}^2 + \alpha_2 y_{t-2}^2 + \dots + \alpha_m y_{t-m}^2 + w_t$ è descritto come un processo autoregressivo condizionale eteroschedastico di ordine m ($ARCH(m)$).

Un processo ARCH ha le seguenti caratteristiche:

- $X_t = \mathbb{E}[X_t | \mathcal{F}_{t-1}] + \varepsilon_t$
- $\mathbb{E}[X_t | \mathcal{F}_{t-1}] = f(\mathcal{F}_{t-1})$
- $\text{Var}[X_t | \mathcal{F}_{t-1}] = \mathbb{E}[\varepsilon_t^2 | \mathcal{F}_{t-1}] = \sigma_t^2$

da cui segue che la media e la varianza condizionata dipendono dalla filtrazione al tempo $t-1$.

Il modello ARCH viene introdotto per lo studio di situazioni in cui si alternano momenti con volatilità differenti. Un ambito classico di applicazione è quello finanziario.

La generalizzazione del modello ARCH è il GARCH.

La famiglia di modelli GARCH(m, q) è definita come

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2 \quad (C.2)$$

Nella famiglia dei modelli GARCH, il più utilizzato è GARCH(1,1), cioè $\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2$ con tutti i coefficienti maggiori di 0.

Appendice D

Il processo gaussiano

Un processo gaussiano $\{X_t\}_{t \in T}$ indicizzata da un set T , è una famiglia di variabili randomiche reali X_t , tutte definite sullo stesso spazio di probabilità, tale che per ogni sottoinsieme finito $F \subset T$ il vettore causale $X_F = \{X_t\}_{t \in F}$ ha una distribuzione gaussiana; se queste distribuzioni finito dimensionali sono tutte non degeneri, allora il processo gaussiano è detto non degenero.

D.1 Le catene di Markov omogenee a tempo discreto

Definiamo processo stocastico a tempo e stati discreti una successione di variabili aleatorie $X_0, X_1, \dots, X_n, \dots$ dove ciascuna X_n è una variabile aleatoria discreta a valori in un insieme S , detto *spazio degli stati*. Uno spazio degli stati numerabile S , che può essere finito o infinito ($S = \{1, 2, \dots, n\}$). Penseremo l'indice $n \in \mathbb{Z}^+$ come il *tempo* e chiameremo *stati* i valori di X_n .

Una Catena di Markov omogenea a tempo discreto è un processo stocastico discreto definito tramite:

1. Uno spazio degli stati S .
2. Una probabilità iniziale definita sugli stati del sistema al tempo $n = 0$:

$$p_i^{(0)} = \mathbb{P}\{X_0 = i\} \quad \forall i \in S.$$

3. La probabilità di osservare al tempo n il sistema nello stato i viene definita tramite $p_i^{(n)}$. Si indica il vettore p^n di componenti $p_i^{(n)}$ ed è la distribuzione della variabile aleatoria X_n , tale che $\forall n \quad \sum_{i \in \mathbb{Z}} p_i^{(n)} = 1$.
4. Un insieme di probabilità di transizione p_{ij} , definite per ogni coppia di possibili stati $\forall i, j \in S$. La probabilità p_{ij} esprime la probabilità con la quale lo stato all'istante successivo $n+1$ assumerà il valore $X_{n+1} = j$, noto che il valore corrente dello stato è $X_n = i$:

$$p_{ij} = \mathbb{P}\{X_n = i \mid X_{n+1} = j\} \quad \forall i, j \in S.$$

Nel caso di catena finita, n è il numero degli stati della catena, posso rappresentare le probabilità di transizione tramite una matrice quadrata M , di dimensione $n \times n$, detta *matrice di transizione* della catena:

$$\begin{bmatrix} p_{11} & \cdots & p_{1n} \\ \cdots & \cdots & \cdots \\ p_{n1} & \cdots & p_{nn} \end{bmatrix}$$

dove l'elemento p_{ij} della matrice rappresenta la probabilità condizionata che la catena, al passo successivo, salti allo stato i , noto che essa si trova attualmente nello stato j .

Proposizione 1. Se (X_n) è una catena di Markov omogenea con probabilità di transizione p_{ij} e distribuzione iniziale $p_i^{(0)} = \mathbb{P}\{X_n = i\}$, allora

$$\mathbb{P}\{X_0 = i_0, \dots, X_n = i_n\} = p_i^{(0)} \cdot p_{i_0 i_1} \cdots p_{i_{n-1} i_n}.$$

Proposizione 2. Se (X_n) è una catena di Markov omogenea con probabilità di transizione p_{ij} e distribuzione iniziale $p^{(0)}$, allora

$$p_i^{(n)} = \sum_{i_0, \dots, i_{n-1}} p_i^{(0)} \cdot p_{i_0 i_1} \cdots p_{i_{n-1} i}.$$

Proposizione 3. Se (X_n) è una catena di Markov omogenea con probabilità di transizione p_{ij} e distribuzione iniziale $p^{(0)}$, allora

$$p_j^{(n)} = \sum_i p_i^{(0)} (\mathbb{M}^n)_{ij},$$

dove $(\mathbb{M}^n)_{ij} = \sum_{i_0, \dots, i_{n-1}} p_{i i_1} \cdots p_{i_{n-1} i_j}$

Bibliografia

- [1] Fraunhofer-Institut für Solare Energiesysteme ISE Energy. Energy-charts. <https://energy-charts.info/>.
- [2] R. Green K. Larsson and F. E. Benth. A stochastic time-series model for solar irradiation. 2021.
- [3] Jūratė Šaltytė Benth and Fred Espen Benth. A critical view on temperature modelling for application in weather derivatives markets. *Energy Economics*, 34(2):592–602, 2012.
- [4] James Douglas Hamilton. *Time Series Analysis*. Princeton University Press, 1994.
- [5] H. C. Hottel. A simple model for estimating the transmittance of direct solar radiation through clear atmospheres. *Solar Energy*, 18, 1976.
- [6] Thomas Mikosch. *Elementary stochastic calculus with finance in view*. World scientific, 1998.
- [7] Tim Bollerslev, Robert F Engle, and Daniel B Nelson. Arch models. *Handbook of econometrics*, 4:2959–3038, 1994.
- [8] Sheldon M Ross, John J Kelly, Roger J Sullivan, William James Perry, Donald Mercer, Ruth M Davis, Thomas Dell Washburn, Earl V Sager, Joseph B Boyce, and Vincent L Bristow. *Stochastic processes*, volume 2. Wiley New York, 1996.
- [9] Iñaki Arto, Iñigo Capellán-Pérez, Rosa Lago, Gorika Bueno, and Roberto Bermejo. The energy requirements of a developed world. *Energy for Sustainable Development*, 33:1–13, 08 2016.