Indice

Elenco degli acronimi utilizzati

AR	Auto-Regressive
ARIMA	Auto-Regressive Integrated Moving Average
ARMA	Auto-Regressive and Moving Average
ASAI	Average Service Availability Index
ASUI	Average Service Unavailability Index
BN	Bayesian Network
BNT	Bayes Net Toolbox
CAIDI	Customer Average Interruption Duration Index
CCDF	Complementary Cumulative Distribution Function
CDF	Cumulative Distribution Function
COPT	Capacity Outage Probability Table
DAG	Directed Acyclic Graph
EENS	Expected Energy Not Supplied
ENS	Energy Not Supplied
EPNS	Expected Power Not Supplied
F&D	Frequency and Duration
FOR	Forced Outage Rate
HL	Hierarchical Level
HL-I	Hierarchical Level-I
HL-II	Hierarchical Level-II
HL-III	Hierarchical Level-III
i.i.d.	Indipendenti e Identicamente Distribuite
IEEE RTS	IEEE Realiability Test System
KDE	Kernel Density Estimation
LDC	Load Duration Curve
LOEE	Loss of Energy Expectation
LOLE	Loss of Load Expectation
LOLF	Loss of Load Frequency
LOLP	Loss of Load Probability
MA	Moving Average
MCS	Monte Carlo Simulation
MI	Mutua Informazione
MISE	Mean Integrated Square Error
MLE	Maximum Likelihood Estimation
MTBF	Mean Time Between Faults
MTTF	Mean Time to Failure
MTTR	Mean Time to Repair
PDF	Probability Density Function
PNS	Power Not Supplied
SAIDI	System Average Interruption Duration Index
SAIFI	System Average Interruption Frequency Index

1-Introduzione

Gli effetti economici e sociali dell'interruzione del servizio elettrico hanno impatti significativi sia sulle società di fornitura di energia elettrica e sia sugli utenti finali che ne usufruiscono. Il sistema elettrico è infatti vulnerabile ad anomalie del sistema, quali ad esempio: guasti sui componenti, errori dei dispositivi di protezione o dei sistemi di comunicazione e controllo, perturbazioni esterne, quali ad esempio sovratensioni atmosferiche , e gli errori operativi umani. Pertanto, fare in modo che il sistema elettrico sia affidabile è un requisito estremamente importante per la progettazione e il funzionamento dei sistemi elettrici di potenza.

L'*affidabilità* di un sistema elettrico indica la sua capacità di soddisfare il fabbisogno di energia elettrica richiesto dagli utenti finali con una ragionevole garanzia di continuità e qualità nella fornitura del servizio. Il concetto di affidabilità può essere suddiviso nei due aspetti fondamentali dell'*adeguatezza* e *sicurezza* del sistema elettrico. L'*adeguatezza* si riferisce all'esistenza di un numero sufficiente di strutture adatte a soddisfare la domanda dei carichi elettrici e i vincoli operativi del sistema. Tali strutture includono quelle necessarie per la generazione dell'energia elettrica e quelle per la trasmissione e la distribuzione richieste per trasportare e consegnare l'energia ai clienti. La valutazione dell'adeguatezza è dunque associata a condizioni statiche e non include la dinamica del sistema e la risposta a perturbazioni transitorie: in tale caso i diversi stati del sistema sono valutati senza prendere in considerazione possibili instabilità che possono essere introdotte da guasti sui componenti del sistema.

D'altra parte, il concetto di *sicurezza* è riferito all'abilità del sistema elettrico di rispondere ai disturbi dinamici e transitori che si possono verificare. Pertanto la valutazione della sicurezza è associata alla reazione del sistema a qualsiasi perturbazione alla quale può essere soggetto. Si può ad esempio considerare una perdita improvvisa della generazione o della capacità di trasmissione che può portare a instabilità della frequenza o della tensione. L'analisi della sicurezza può essere ulteriormente suddivisa definendo la *sicurezza dinamica* e la *sicurezza statica*. La valutazione della sicurezza dinamica consiste nel determinare se oscillazioni della frequenza conseguenti a un guasto o a un'interruzione possono causare una perdita di sincronismo tra generatori, mentre l'obiettivo dell'analisi della sicurezza statica è determinare se, a seguito del verificarsi di una contingenza, esista un nuovo punto di funzionamento stazionario in cui il sistema elettrico perturbato si stabilizzerà dopo che le oscillazioni si saranno smorzate.

La maggior parte delle tecniche attualmente disponibili per valutare l'affidabilità dei sistemi elettrici si riferiscono alla valutazione dell'adeguatezza. La capacità di valutare la sicurezza nei sistemi elettrici, invece, è ancora molto limitata principalmente a causa della complessità associata alla modellazione del comportamento dinamico del sistema. La maggior parte degli indici di affidabilità sono infatti indici di adeguatezza e non indici di sicurezza.

Come è noto, il sistema elettrico globale è composto dalle strutture di generazione, trasmissione e distribuzione. Calcolare l'affidabilità del sistema elettrico, considerandolo come unica entità, risulterebbe

complesso dal punto di vista computazionale e i risultati potrebbero inoltre essere di difficile interpretazione. Pertanto la valutazione dell'affidabilità viene solitamente effettuata suddividendo il sistema elettrico di potenza in gruppi funzionali definendo dei livelli gerarchici (Hierarchical Level, HL), in modo da poter evidenziare l'influenza di ogni sottosistema sull'affidabilità del sistema complessivo [Billinton e Allan, 1996].

Si individuano tre livelli gerarchici:

- il sottosistema di generazione (HL-I);
- il sottosistema comprendente generazione e trasmissione (HL-II);
- il sottosistema composto da generazione, trasmissione e distribuzione (HL-III).

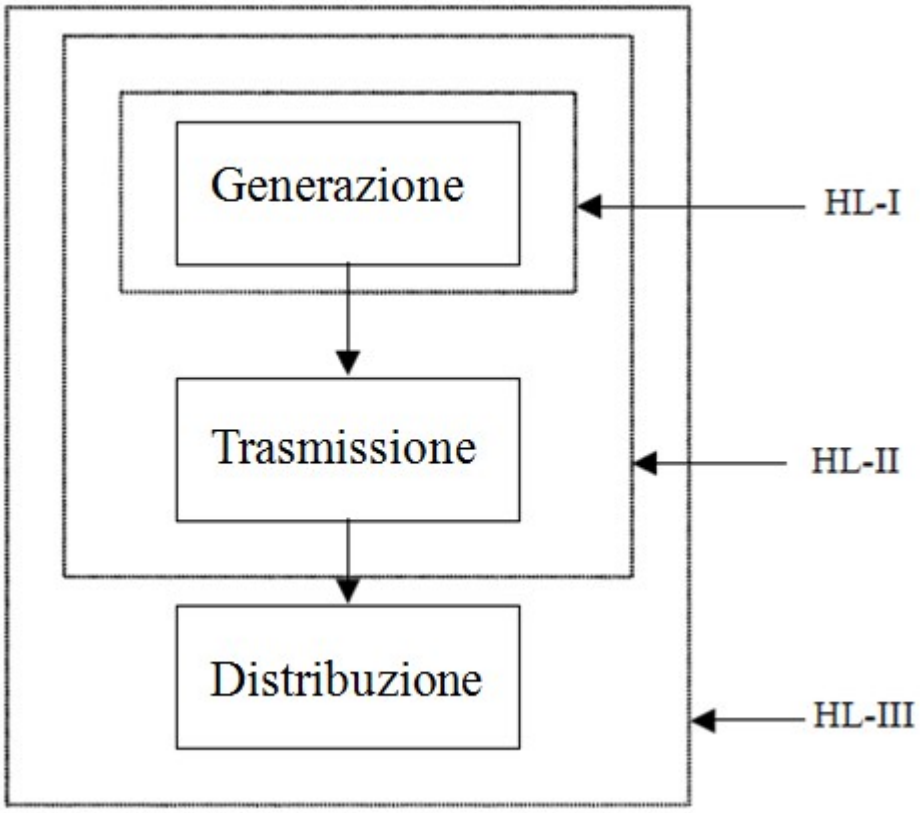

Figura 1-Livelli gerarchici del sistema elettrico.

Il primo livello (HL-I) rappresenta tutti i generatori del sistema elettrico connessi alla rete di trasmissione. Essi vengono visti come un'unica struttura in grado di poter soddisfare o meno la domanda di energia elettrica richiesta dai carichi. Il secondo livello (HL-II) rappresenta il sottosistema composto da generazione e trasmissione ed identifica pertanto la capacità, da parte di tale sistema aggregato di produrre un'adeguata energia elettrica e di renderla disponibile nei nodi di consegna alla rete di distribuzione, rappresentati dalle cabine primarie, ovvero le sottostazioni di trasformazione AT/MT. Il terzo livello (HL-III) identifica il sistema elettrico complessivo che dunque include, oltre ai sistemi di generazione e trasmissione, anche il

sistema di distribuzione considerando quindi la capacità di soddisfare le richieste di energia di ogni utente [Nicolosi, 2011].

Come accennato in precedenza, valutare l'affidabilità del sistema elettrico al livello gerarchico HL-III è difficilmente praticabile, a causa della complessità del problema: in tale ambito viene dunque eseguita una valutazione del solo sistema di distribuzione. Tale sistema si interfaccia infatti con quello di trasmissione tramite le cabine primarie, che si possono considerare come i carichi del livello HL-II, pertanto l'affidabilità del sottosistema HL-II può essere visto come un dato di input per la valutazione dell'affidabilità del sistema di distribuzione [Allan e Billinton, 2000].

2-Metodi per la valutazione dell'affidabilità

I criteri di valutazione dell'affidabilità di un sistema elettrico possono essere di due tipi: deterministici e stocastici. I metodi deterministici sono usati tipicamente per misurare l'adeguatezza dei sistemi di generazione. Un approccio di questo tipo consiste ad esempio nel determinare la minima capacità di generazione, che può essere calcolata come la somma della domanda elettrica di base prevista e del carico di picco previsto. Con tale metodo è possibile anche determinare la riserva operativa, pari alla differenza tra la capacità di generazione installata e la massima domanda elettrica prevista. Il criterio deterministico più comunemente impiegato, in particolare nei sistemi di trasmissione, è il criterio di sicurezza N-1,secondo il quale il fuori servizio di un qualsiasi componente del sistema non comporta una violazione dei vincoli operativi del sistema complessivo. I criteri deterministici sono semplici da implementare ma non sono particolarmente adatti alla valutazione dell'affidabilità dei sistemi elettrici attuali. Tale tipologia di approccio può infatti portare in molti casi a soluzioni costose senza un'apparente giustificazione, in quanto questi metodi di valutazione non considerano la natura probabilistica dei sistemi elettrici, dovuta al fatto che esistono varie fonti di incertezza associate agli eventi che si possono verificare in tali sistemi, quali ad esempio: l'aleatorietà dei guasti, ovvero non è possibile dire con esattezza quando e in quale punto si verificherà un guasto sui componenti elettrici della rete, l'aleatorietà dei tempi di ripristino, ovvero non si può sapere con certezza quando un certo guasto verrà riparato e di conseguenza verrà ripristinato il servizio, variazioni della domanda, variazioni della generazione di energia elettrica da fonti rinnovabili, ad esempio negli impianti eolici, a causa della variazione della velocità del vento, dato che talvolta non se ne può avere disponibilità a causa delle condizioni climatiche oppure la sua velocità non è sufficiente per la produzione di energia, nei sistemi fotovoltaici, a causa della variazione dell'irradianza solare, anche in tal caso per le condizioni climatiche , negli impianti idroelettrici, a causa della variazione delle portate affluenti. Per tali ragioni i metodi deterministici non sono molto impiegati nei casi pratici.

Viene dunque generalmente preferita un'analisi dell'affidabilità impostata su un approccio probabilistico. Con i metodi stocastici è difatti possibile prendere in considerazione le tipologie di incertezza incorporandole in appositi modelli probabilistici che si vanno a definire per la valutazione. Il modello di riferimento usato per componenti elettrici riparabili (ad esempio generatori, trasformatori, linee) è quello di Markov a due

stati, riportato in Fig. 2: lo stato 0 indica la condizione di normale funzionamento, nello stato 1 il componente è guasto. Si definiscono inoltre λ e μ rispettivamente come il tasso di guasto e il tasso di riparazione, impiegati per definire le transizioni da uno stato all'altro [Billinton e Bollinger, 1968].

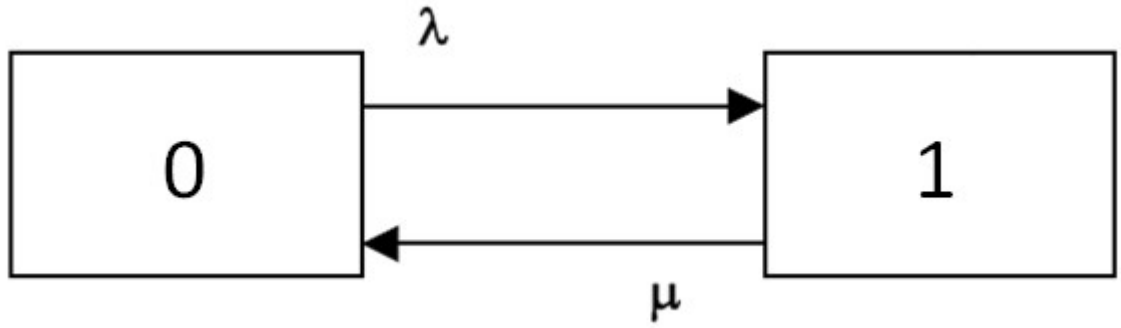

Figura 2-Modello di Markov a due stati.

L'andamento del tasso di guasto e di riparazione nel tempo è tipicamente dato dalla curva a "vasca", mostrata in Fig. 3. In tale andamento si possono individuare tre zone singolari:

- una zona detta di "early failure", ovvero dei guasti precoci, dovuti a difetti di fabbricazione, in cui il tasso di guasto è decrescente,
- una zona in cui il tasso di guasto è costante nel tempo: tale intervallo corrisponde alla vita utile del componente,
- una zona detta di "wear out failure", cioè dei guasti dovuti all'invecchiamento e all'usura del componente, in cui il tasso di guasto è crescente.

Per il calcolo dell'affidabilità si assumono i tassi di guasto e di riparazione costanti nel tempo, riferiti alla vita utile del componente.

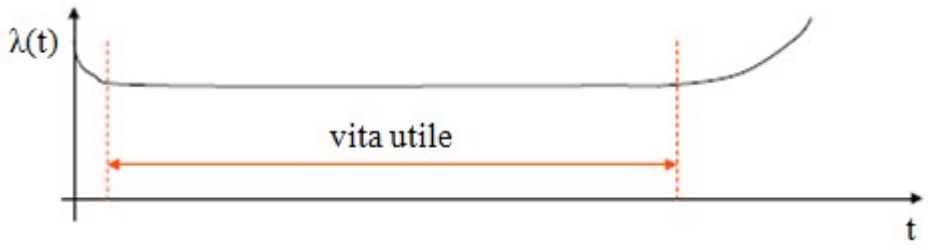

Figura 3-Curva a "vasca" per componenti riparabili.

Si definiscono dunque i seguenti termini:

$\mathbb{P}_0(t)$: probabilità che il componente sia funzionante all'istante t.

$\mathbb{P}_1(t)$: probabilità che il componente sia guasto all'istante t.

Considerando un intervallo di tempo infinitesimo dt e assumendo che la probabilità che due o più eventi che si verificano in tale intervallo sia trascurabile sia trascurabile si ha che:

$$\mathbb{P}_0(t+dt) = \mathbb{P}_0(t)(1-\lambda dt) + \mathbb{P}_1(t)\mu dt \tag{2.1}$$

$$\mathbb{P}_1(t+dt) = \mathbb{P}_1(t)(1-\mu dt) + \mathbb{P}_0(t)\lambda dt \tag{2.2}$$

dove λdt rappresenta la probabilità che il componente, funzionante all'istante t, si guasti nel successivo intervallo dt, mentre μdt indica la probabilità che il componente, guasto all'istante t, sia riparato nel successivo intervallo dt.

Si calcolano dunque le derivate delle probabilità:

$$\frac{d\mathbb{P}_0}{dt} = -\lambda\mathbb{P}_0(t) + \mu\mathbb{P}_1(t) \tag{2.3}$$

$$\frac{d\mathbb{P}_1}{dt} = \lambda\mathbb{P}_0(t) - \mu\mathbb{P}_1(t) \tag{2.4}$$

A regime, essendo le derivate pari a zero, si avrà:

$$-\lambda\mathbb{P}_{0\infty} + \mu\mathbb{P}_{1\infty} = 0 \tag{2.5}$$

$$\lambda\mathbb{P}_{0\infty} - \mu\mathbb{P}_{1\infty} = 0 \tag{2.6}$$

Essendo $\mathbb{P}_{0\infty} + \mathbb{P}_{1\infty} = 1$, si ricava che:

$$\mathbb{P}_{0\infty} = \frac{\mu}{\lambda+\mu} \tag{2.7}$$

$$\mathbb{P}_{1\infty} = \frac{\lambda}{\lambda+\mu} \tag{2.8}$$

Si definisce dunque il coefficiente di disponibilità ρ come la probabilità del componente in questione di essere funzionante ed è dunque dato dalla probabilità a regime per lo stato di normale funzionamento:

$$\rho = \frac{\mu}{\lambda+\mu} \tag{2.9}$$

Alternativamente, il coefficiente di disponibilità può essere definito nella maniera seguente:

$$\rho = \frac{T_f}{T_f+T_R} \tag{2.10}$$

dove T_f è il Mean Time To Failure (*MTTF*) , ovvero l'intervallo medio di tempo intercorrente tra l'istante in cui il componente inizia o riprende a funzionare e l'istante in cui avviene un guasto; T_R è il Mean Time To Repair (*MTTR*) e rappresenta l'intervallo medio di tempo che intercorre tra l'istante in cui il componente è guasto e l'istante in cui esso viene riparato e ne viene dunque ripristinato il funzionamento. Avendo

considerato i tassi di guasto e di riparazione costanti nel tempo, le dinamiche dei fenomeni di guasto e di riparazione seguono una distribuzione esponenziale, pertanto si ha che:

$$MTTF = \frac{1}{\lambda} \tag{2.11}$$

$$MTTR = \frac{1}{\mu} \tag{2.12}$$

Il coefficiente di indisponibilità del componente, definito anche Forced Outage Rate (*FOR*), vale:

$$FOR = 1 - \rho = \frac{\lambda}{\lambda+\mu} = \frac{T_R}{T_f+T_R} \tag{2.13}$$

Si definisce inoltre Mean Time Between Failures (*MTBF*), come l'intervallo di tempo che intercorre tra due guasti successivi, pari a:

$$MTBF = MTTF + MTTR \tag{2.14}$$

Tali coefficienti risultano particolarmente adatti per definire un modello probabilistico delle unità di generazione del sistema elettrico. Il modello del carico è invece dato dalla sua funzione di distribuzione cumulativa (*CDF*, Cumulative Distribution Function), ottenuta dalla curva di durata oraria o giornaliera del carico.

3-Indici di affidabilità

3.1-Sistemi di generazione

Combinando il modello di generazione con il modello di carico, risulta possibile effettuare una valutazione quantitativa dell'affidabilità attraverso la definizione di opportuni indici. In particolare per la valutazione dell'adeguatezza della capacità di generazione, si vanno a definire i seguenti indicatori [Billinton e Allan, 1996]:

Loss of load expectation (*LOLE*): è definito come il numero di ore o giorni nel periodo di valutazione (di solito un anno) in cui il carico di picco orario o giornaliero eccede la capacità di generazione disponibile. Rappresenta dunque il numero di ore o giorni in cui potrebbe verificarsi una carenza di potenza prodotta dai generatori ed è matematicamente espresso nel modo seguente:

$$LOLE = \sum_{i=1}^{n} \mathbb{P}_i(L_i > C_i) \tag{3.1}$$

dove n è il numero di ore o giorni, C_i indica la potenza generata disponibile nell'ora o nel giorno i-esimo, L_i è il valore di potenza di picco del carico previsto per l'ora o per il giorno i-esimo, mentre $\mathbb{P}_i(L_i > C_i)$ è la probabilità di perdita di carico nell'ora o nel giorno i-esimo.

Loss of load probability (*LOLP*): rappresenta la probabilità che la capacità di generazione disponibile sia insufficiente a soddisfare la domanda di carico nel periodo di valutazione T. Si ottiene direttamente dal LOLE ed è pari a:

$$LOLP = \frac{LOLE}{T} \tag{3.2}$$

Loss of load frequency (*LOLF*): indica il numero di volte, durante il periodo di valutazione, in cui è previsto che si verifichi una perdita di carico.

Loss of Energy expectation (*LOEE*): misura la quantità di energia che si prevede che non potrà essere fornita a causa dell'indisponibilità di generazione nel periodo considerato. Viene definito come:

$$LOEE = \sum_{k=1}^{n} \mathbb{P}_k E_k \tag{3.3}$$

dove $\mathbb{P}_k$ rappresenta la probabilità associata al fuori servizio della capacità di generazione C_k nell'ora o nel giorno k-esimo, mentre E_k è l'energia non fornita nell'ora o nel giorno k-esimo a seguito della riduzione di capacità C_k.

Tali indici possono essere utilizzati anche per il calcolo dell'adeguatezza dei sistemi di generazione distribuita connessi alla rete di distribuzione.

3.2-Sistemi di distribuzione

Per quanto concerne i sistemi di distribuzione si introducono indicatori di affidabilità per valutarne la continuità di servizio in tali sistemi. Si fa tipicamente una distinzione tra *indici di affidabilità locali*, riferiti a un singolo punto di carico, e *indici di affidabilità globali*, riferiti al sistema complessivo [Carpaneto e Chicco,2004]. Si definisce punto di carico l'insieme equivalente dei carichi di un sistema di distribuzione alimentato da un punto di consegna in MT. Il calcolo degli indici di affidabilità viene effettuato tenendo conto di un certo numero di variabili aleatorie, cosicché anch'essi risultano essere variabili casuali. Le variabili aleatorie che vengono prese in considerazione sono:

- f: frequenza delle interruzioni;
- d: durata delle interruzioni;
- n: numero di occorrenze dei guasti;
- τ_i: tempo di ripristino di un'occorrenza i-esima di guasto.

Chiaramente si ha che $f = n$; si definisce inoltre $\tau = \mathbb{E}\{\tau_i\}$ il valoreatteso del tempo di ripristino, mentre la durata delle interruzioni viene definita come $d = \sum_{i=1}^{n} \tau_i$, ovvero la somma casuale dei tempi di ripristino nelle n occorrenze dei guasti.

3.2.1-Indici di affidabilità locali

Gli indici di affidabilità locali che si vanno a definire sono valori attesi delle variabili casuali descritte precedentemente, calcolati nel periodo di osservazione T per più occorrenze dello stesso tipo di guasto, avente tasso λ. Essi sono i seguenti:

$\mathbb{E}\{n\} = \lambda T$: rappresenta il valore atteso del numero di occorrenze dei guasti nel punto di carico considerato.

$\mathbb{E}\{f\} = \mathbb{E}\{n\} = \lambda T$: misura la frequenza delle interruzioni nel punto di carico in esame.

$\mathbb{E}\{d\} = \mathbb{E}\{n\}\mathbb{E}\{\tau_i\} = \lambda T \tau$: indica il valore atteso della durata delle interruzioni che si verificano nel punto di carico considerato.

Si definisce inoltre il coefficiente di indisponibilità U, valutato sul periodo T, come la probabilità che il punto di carico non sia alimentato e vale:

$$U = \frac{\mathbb{E}\{d\}}{T} \tag{3.4}$$

Il coefficiente di disponibilità A rappresenta invece la probabilità che il punto di carico sia alimentato e vale:

$$A = 1 - U \tag{3.5}$$

3.2.2-Indici di affidabilità globali

Gli indici di affidabilità globali vengono calcolati su un numero N di punti di carico costituenti il sistema di distribuzione preso in esame. Gli indici più usati sono i seguenti:

System Average Interruption Frequency Index (*SAIFI*): definisce una media pesata della frequenza delle interruzioni per ogni utente del sistema ed è dato dalla seguente espressione:

$$SAIFI = \frac{\sum_{k=1}^{N} p_k f_k}{\sum_{k=1}^{N} p_k} \tag{3.6}$$

dove f_k è la frequenza delle interruzioni calcolata per il k-esimo punto di carico, mentre p_k è il peso con il quale si effettua la media, e può rappresentare la potenza complessivamente assorbita dal k-esimo punto di carico o il numero di utenti dello stesso.

System Average Interruption Duration Index (*SAIDI*): rappresenta la durata media delle interruzioni per ciascun utente del sistema e risulta pari a:

$$SAIDI = \frac{\sum_{k=1}^{N} p_k d_k}{\sum_{k=1}^{N} p_k} \tag{3.7}$$

dove d_k è la durata delle interruzioni del k-esimo punto di carico.

Customer Average Interruption Duration Index (*CAIDI*): è un indicatore composto, dato dal rapporto tra la durata media e la frequenza media delle interruzioni per utente del sistema. È un valore che quantifica la durata di ciascuna interruzione:

$$CAIDI = \frac{SAIDI}{SAIFI} = \frac{\sum_{k=1}^{N} p_k d_k}{\sum_{k=1}^{N} p_k f_k} \tag{3.8}$$

Average Service Availability Index (*ASAI*): è dato dal rapporto tra il numero di ore in cui l'utente è effettivamente alimentato e le ore totali di alimentazione richieste dallo stesso, che si suppone siano le ore totali di un anno (8760 h):

$$ASAI = \frac{\sum_{k=1}^{N} p_k \cdot 8760 - \sum_{k=1}^{N} p_k d_k}{\sum_{k=1}^{N} p_k \cdot 8760} \tag{3.9}$$

Si definisce l'indice ASUI (Average Service Unavailability Index) come il complementare dell'*ASAI*:

$$ASUI = 1 - ASAI \tag{3.10}$$

Energy Not Supplied (*ENS*), detta anche energia non servita: rappresenta la somma delle energie dei carichi non alimentati; viene calcolata dalle rispettive potenze non servite e dalle rispettive durate di guasto.

Sia C_k la potenza consegnata al k-esimo punto di carico durante il normale funzionamento nell'intervallo di tempo T; l'occorrenza di un guasto porta a una sequenza di fasi di ripristino, con interventi eseguiti in telecomando dal centro di controllo o manualmente sul luogo del guasto al fine di ripristinare il servizio. Si assume che il ripristino del servizio dopo un guasto f includa un numero H_f di fasi di ripristino indipendenti. Si introduce la variabile binaria $b_{fk}^{(h)} = 1$ se il punto di carico k non è alimentato durante la fase h di ripristino del servizio dopo il guasto f, altrimenti $b_{fk}^{(h)} = 0$. La potenza non servita (Power Not Supplied, *PNS*) sul sistema complessivo durante la fase di ripristino $h = 1, \dots, H_f$ dopo il guasto f è:

$$P_f^{(h)} = \sum_{k=1}^{N} C_k b_{fk}^{(h)} \tag{3.11}$$

Si introducono ora l'insieme Θ_k contenente i guasti per i quali il k-esimo punto di carico non è alimentato e la variabile $\tau_f^{(h)}$ che rappresenta il valore atteso del tempo di ripristino distribuito esponenzialmente per la fase $h = 1, \dots, H_f$ e per il guasto $f = 1, \dots, F$, con tasso di guasto λ_f. Il valore atteso dell'energia non servita al k-esimo punto di carico viene dunque calcolata considerando tutte le fasi di ripristino nelle quali il k-esimo punto di carico non è alimentato:

$$ENS_k = C_k \sum_{f \in \Theta_k} \sum_{h=1}^{H_f} \lambda_f b_{fk}^{(h)} \tau_f^{(h)} \tag{3.12}$$

L'energia non servita complessiva considerando gli N punti di carico del sistema è dunque pari a:

$$ENS = \sum_{k=1}^{N} ENS_k \tag{3.13}$$

4-Metodi probabilistici

La maggior parte delle tecniche probabilistiche sviluppate per la valutazione dell'affidabilità possono essere suddivise in due tipologie generali: analitiche e simulative [Allan e Billinton, 2000]. Il *metodo analitico* descrive il comportamento del sistema attraverso modelli matematici e successivamente effettua il calcolo numerico dei valori medi degli indici di affidabilità richiesti dall'analisi. Le tecniche analitiche restituiscono sempre gli stessi risultati numerici per lo stesso sistema in esame, e dunque per lo stesso modello e per lo stesso insieme dei dati di input. Tale tipo di approccio veniva adottato in passato soprattutto per il suo basso onere computazionale. Tuttavia, se vengono considerati sistemi particolarmente complessi, diventa necessario definire numerose supposizioni e semplificazioni per ridurre la complessità del problema e produrre un modello analitico trattabile del sistema. Pertanto è possibile che tali metodologie forniscano risultati privi di significato ai fini della valutazione.

D'altra parte il *metodo simulativo*, detto anche metodo Monte Carlo, stima gli indici di affidabilità simulando i processi reali del sistema attraverso un campionamento casuale dei casi possibili. Generalmente il metodo Monte Carlo richiede un maggior sforzo computazionale rispetto alle tecniche analitiche, tuttavia esso permette di considerare tutti gli aspetti e le eventualità connesse alla pianificazione e all'esercizio dei sistemi elettrici, superando così le difficoltà inerenti all'impiego dei metodi analitici.

4.1-Tecniche analitiche

Scegliendo questo metodo, il modello del generatore può essere rappresentato definendo la cosiddetta Capacity Outage Probability Table (*COPT*). Il calcolo della *COPT* consiste nell'enumerazione di tutti gli stati del sistema e le loro probabilità di verificarsi, essendo ogni stato rappresentato dalla sua capacità fuori servizio. È possibile usare un algoritmo ricorsivo per definire la *COPT* come un vettore di livelli di capacità con le loro associate probabilità di esistenza [Billinton e Allan, 1996]. In questo algoritmo, gli stati di tutte le unità di generazione del sistema in esame vengono inseriti nella *COPT* uno alla volta in un processo sequenziale fino a quando la *COPT* è completamente definita. Ciascuna probabilità associata a un dato livello di capacità C_i rappresenta la probabilità cumulativa $\mathbb{P}(X > C_i)$ di avere una capacità fuori servizio X maggiore o uguale di C_i.

L'unità di generazione può essere rappresentata con il modello di Markov a due stati, come è stato detto in precedenza. In tal modo, l'espressione della probabilità cumulativa per un dato livello di capacità C_i dopo aver aggiunto un'unità di generazione con un dato valore di *FOR* è la seguente:

$$\mathbb{P}^A(X) = (1 - FOR)\mathbb{P}^B(X) + (FOR)\mathbb{P}^B(X - C_i) \tag{4.1}$$

dove $\mathbb{P}^B(X)$ e $\mathbb{P}^A(X)$ sono le probabilità cumulative di avere una capacità fuori servizio X rispettivamente prima e dopo l'aggiunta di un'unità di generazione. $\mathbb{P}^B(X)$ è inizialmente assunta pari a 1 per $X \leq 0$ e pari a 0 per $X > 0$. Dopo aver ottenuto tutti i valori della *COPT*, questa viene combinata con il modello del carico

in modo da valutare gli indici di affidabilità. Il metodo per combinare i differenti stati nel modello della generazione con la curva di durata del carico viene mostrato in Fig. 4: da essa si può notare infatti come livello di capacità fuori servizio calcolato nella *COPT* che non permette di soddisfare la domanda nell'intervallo di tempo t_k contribuirà al *LOLE* del sistema.

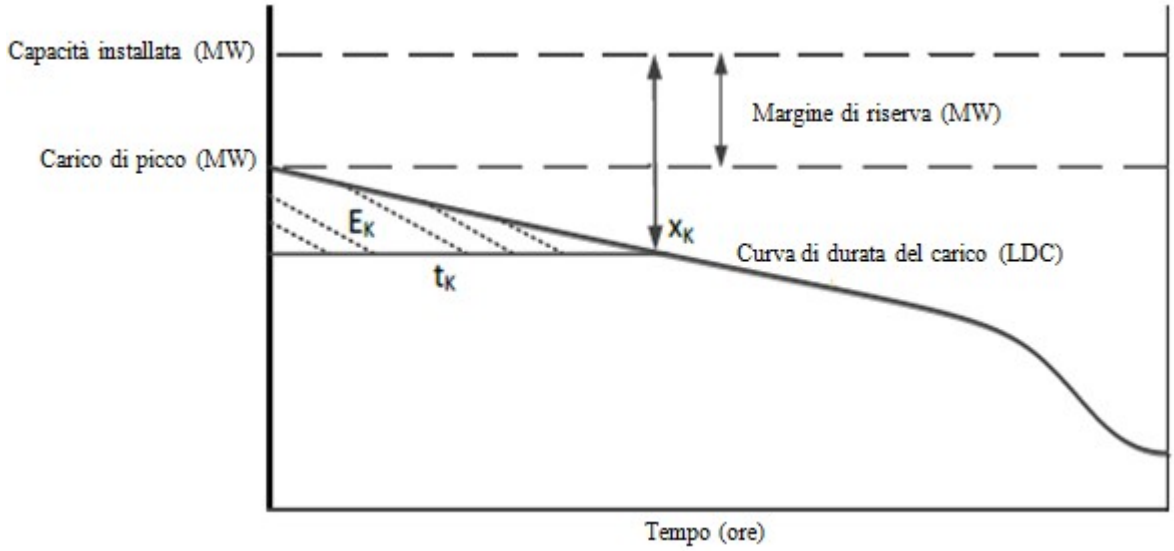

Figura 4-Relazione tra capacità, carico e riserva.

Il calcolo del *LOLE* può essere effettuato moltiplicando le probabilità, calcolate nella *COPT*, che si possa verificare un fuori servizio di una capacità di generazione con i rispettivi intervalli di tempo nei quali accade tale evento. L'indice *LOLP* si determina con lo stesso metodo usato per il *LOLE* ma escludendo gli intervallo di tempo associati al fuori servizio dei livelli di capacità. Si ottengono in tal modo le seguenti espressioni:

$$LOLE = \sum_{k=1}^{n} t_k \mathbb{P}_k(X > C_{tot} - L) \tag{4.2}$$

$$LOLP = \sum_{k=1}^{n} \mathbb{P}_k(X > C_{tot} - L) \tag{4.3}$$

dove n rappresenta il numero dei fuori servizio che si verificano, t_k è l'intervallo di tempo nel quale si ha la perdita di carico dovuta alle interruzioni dell'alimentazione, C_{tot} è la capacità di generazione complessiva del sistema, L è il valore del livello di carico.

Questo metodo consente inoltre il calcolo degli indici basati sull'energia, come il *LOEE*. Difatti l'area sottesa dalla curva di durata del carico rappresenta l'energia consumata nel periodo di studio. Pertanto, così come è possibile calcolare il numero di ore o giorni nei quali si avranno dei carichi non alimentati, risulta anche possibile ricavare i valori previsti di energia non fornita. Il *LOEE* del sistema può essere calcolato con l'espressione seguente:

$$LOEE = \sum_{k=1}^{n} t_k (X - (C_{tot} - L)) \mathbb{P}_k(X > C_{tot} - L) \tag{4.4}$$

I criteri analitici permettono inoltre di valutare gli indici che misurano la continuità di servizio; in tale caso è richiesta la conoscenza dei tassi di guasto e di riparazione che definiscono la transizione tra i due stati del modello markoviano. Come per la valutazione dell'adeguatezza della capacità di generazione, questi indici di affidabilità combinando il modello di generazione costruito ricorsivamente con il modello del carico. Gli indici così calcolati permettono di considerare l'incertezza sulla previsione del carico.

Le tecniche analitiche rappresentano dunque una scelta efficiente per la valutazione dell'affidabilità di sistemi relativamente piccoli e in presenza di generazione convenzionale; tuttavia se il sistema in questione risulta essere complesso o molto esteso e se si è in presenza di impianti di generazione variabile, come l'eolico e il fotovoltaico, queste metodologie non sono appropriate a causa di tale complessità del sistema.

4.2-Metodo Monte Carlo

Come accennato in precedenza, le tecniche analitiche sono state a lungo impiegate in passato per valutare gli indici di affidabilità grazie ai loro tempi computazionali relativamente brevi e al fatto che non sono richiesti notevoli risorse di calcolo. Lo sviluppo e la sempre più crescente disponibilità di calcolatori efficienti e con maggiore potenza di calcolo hanno creato le opportunità per analizzare molti problemi usando metodi di simulazione. Negli ultimi decenni si è avuto dunque un crescente interesse nell'utilizzare il metodo di simulazione Monte Carlo nell'analisi quantitativa dell'affidabilità dei sistemi elettrici. Sebbene il metodo Monte Carlo non sia un concetto nuovo, in quanto le sue applicazioni esistono da almeno 60 anni [Kahn e Harris, 1951], la disponibilità di strumenti di calcolo più potenti hanno ora reso tale metodologia un'alternativa preferibile per molti problemi di affidabilità. Le tecniche simulative possono essere anche impiegate in situazioni dove ci sono poche informazioni disponibili. In aggiunta, esse risultano utili per suddividere un sistema complesso in più sottosistemi, ciascuno dei quali può essere modellato ed analizzato separatamente.

I metodi di simulazione Monte Carlo possono essere classificati in due tipologie generali, indicate come *metodi non sequenziali* e *sequenziali*. Con l'approccio non sequenziale gli stati del sistema sono campionati in maniera casuale senza tener conto della cronologia degli eventi nel funzionamento del sistema. Pertanto tale metodo non è in grado di modellare eventi sequenziali e correlazioni temporali. D'altra parte, con l'approccio sequenziale gli stati di tutti i componenti del sistema sono campionati in accordo con le loro distribuzioni di probabilità e il ciclo di funzionamento del sistema è ottenuto combinando i cicli operativi di tutti i componenti; per tale ragione questa tecnica permette di tenere conto di elementi tempo-varianti come le curve di durata dei carichi, il vento, le portate affluenti in un impianto idroelettrico e le correlazioni statistiche tra di essi.

Le principali tecniche di simulazione che si possono implementare servendosi del metodo Monte Carlo per l'analisi dell'affidabilità dei sistemi elettrici sono le seguenti tre [Billinton e Li, 1994]:

- campionamento dello stato (state sampling);

- campionamento della transizione di stato (state transition sampling);
- campionamento della durata dello stato (state duration sampling).

Le prime due appartengono alla categoria delle simulazioni non sequenziali, mentre la terza è una procedura di simulazione sequenziale.

4.2.1-State sampling

Usando questo approccio [Billinton e Li, 1991], [Pereira et al.,1992], gli stati di tutti i componenti del sistema sono campionati in modo che, combinandoli, si ottenga una sequenza non cronologica degli stati operativi del sistema. La procedura di base di campionamento viene condotta generando numeri casuali e assumendo che il comportamento di ogni componente possa essere descritto da una distribuzione uniforme nell'intervallo [0,1]. Ciascun campione degli stati del sistema è selezionato casualmente e indipendentemente dai campioni precedenti e successivi. Il maggior vantaggio di questa metodologia è che richiede tempi computazionali relativamente brevi con requisiti di memoria non troppo stringenti, essendo un processo di calcolo relativamente semplice. Esso ha però lo svantaggio di non poter essere usato per calcolare indici di frequenza e durata, in quanto non è in grado di riconoscere l'impatto delle transizioni di stato dovute ai guasti e le variabilità associate a un modello del carico cronologico.

4.2.2-State transition sampling

Il metodo di state transition sampling [Billinton e Li, 1993], [Billinton e Sankarakrishnan, 1994] è focalizzato sulle transizioni degli stati del sistema complessivo anziché dei singoli componenti. Una sequenza delle transizioni degli stati può essere ottenuta con un numero ampio di campioni da cui poi risulta possibile valutare la probabilità di ciascun stato del sistema. Con tale approccio risulta possibile effettuare il calcolo di indici di frequenza e durata, avendo creato una catena delle transizioni di stato. Lo state transition sampling in generale non comporta il campionamento delle funzioni di distribuzione della durata degli stati dei componenti e nemmeno la conoscenza di informazioni cronologiche come richiesto nell'approccio sequenziale. Lo svantaggio di tale metodo è che si può applicare soltanto a componenti caratterizzati da durate degli stati distribuite esponenzialmente, cosa che non sempre risulta appropriata nell'analisi dell'affidabilità.

4.2.3-State duration sampling

Lo state duration sampling è un processo di simulazione sequenziale [Billinton e Li, 1994], ed è basato sul campionamento delle distribuzioni di probabilità della durata degli stati dei componenti del sistema. In questo approccio, vengono in prima luogo simulati i processi cronologici di transizione degli stati per tutti i componenti. Successivamente viene creato il processo di transizione degli stati del sistema complessivo combinando quelli dei singoli componenti. Il termine "simulazione sequenziale" viene spesso usato per indicare la tecnica con cui la storia di un sistema è simulata attraverso passi temporali discreti [Rubinstein,

1981]. Il termine può anche essere definito in senso ingegneristico [Ubeda e Allan, 1992], secondo il quale ogni evento che accade entro un particolare passo temporale si considera come se si fosse verificato alla fine di quel dato intervallo temporale, e gli stati sono aggiornati di conseguenza. In genere si considera sufficientemente adeguato per la valutazione dell'affidabilità un passo temporale di un'ora in quanto il numero di variazioni che si verificano in tale periodo è generalmente piccolo.

Nella simulazione sequenziale, ciascun campione relativo a uno stato del sistema è correlato all'insieme degli stati precedente, data la dipendenza storica. Si crea dunque un'evoluzione sequenziale nel tempo del comportamento del sistema che consente di valutare un'ampia gamma di indici di affidabilità. I fattori casuali che influenzano gli stati dei generatori e dei carichi in sistemi storicamente dipendenti e gli scenari operativi richiesti possono essere incorporati utilizzando la simulazione Monte Carlo sequenziale. L'approccio di simulazione sequenziale risulta dunque utile quando il sistema da analizzare è *dipendente dagli eventi passati*, ossia lo stato del sistema in qualunque momento è parzialmente determinato dalla sua evoluzione temporale. Il maggiore svantaggio della simulazione sequenziale è dovuto al fatto che richiede *maggiori tempi computazionali* rispetto ai metodi non sequenziali in quanto è necessario generare variabili casuali che seguono una data distribuzione di probabilità per ogni componente e memorizzare le informazioni sui processi cronologici di transizioni di stato di tutti i componenti per un lungo arco di tempo.

4.3-Metodologia di base della simulazione sequenziale

L'approccio di simulazione sequenziale è basato sul campionamento delle distribuzioni di probabilità delle durate degli stati dei componenti. In una rappresentazione del componente a due stati, tali funzioni sono le distribuzioni della durata dello stato di normale funzionamento e dello stato di guasto e generalmente si assume che siano esponenziali. Il metodo di simulazione sequenziale può essere riassunto nei seguenti passaggi:

Step 1: viene specificato lo stato iniziale di ciascun componente. Generalmente si assume che tutti i componenti siano inizialmente nello stato di normale funzionamento.

Step 2: la durata dello stato attuale di ogni componente viene campionata dalla sua distribuzione di probabilità. Ad esempio una variabile aleatoria T distribuita esponenzialmente, ha la seguente funzione di densità di probabilità:

$$f_T(t) = \lambda e^{-\lambda} \tag{4.5}$$

dove $\frac{1}{\lambda}$ è il valor medio della distribuzione. La funzione di distribuzione cumulativa è:

$$F_T(t) = 1 - e^{-\lambda t} \tag{4.6}$$

A questo punto per ricavare la variabile casuale T si ricorre al metodo dell'inversione, che si basa sul fatto che se una variabile aleatoria continua X ha una funzione di distribuzione cumulativa monotona crescente F_X,

allora la variabile casuale $Y = F_X(X)$ segue una distribuzione uniforme nell'intervallo $[0,1]$. Nell'esempio considerato si ha dunque che la variabile casuale T è data dalla seguente espressione [Rubinstein, 1981]:

$$T = -\frac{1}{\lambda} ln(1 - U) \tag{4.7}$$

dove U è un numero casuale avente distribuzione uniforme nell'intervallo $[0,1]$ottenuto attraverso un generatore di numeri pseudo-casuali, tipicamente con l'ausilio di appositi software di calcolo. Siccome il termine $1 - U$ è distribuito uniformemente nella stessa maniera di U nell'intervallo $[0,1]$, si ottiene che:

$$T = -\frac{1}{\lambda} ln(U) \tag{4.8}$$

Se lo stato attuale è lo stato di normale funzionamento, λ è il tasso di guasto del componente, pari a $\frac{1}{MTTF}$; altrimenti se lo stato attuale è quello di guasto, λ indica il tasso di riparazione del componente, uguale a $\frac{1}{MTTR}$.

Step 3: lo Step 2 è ripetuto nell'intervallo di tempo di simulazione considerato, tipicamente un anno, e vengono memorizzati i valori campionati di ciascuna durata degli stati per tutti i componenti. I processi cronologici di transizione di stato nel dato arco temporale di tutti i componenti vengono dunque combinati per creare il processo cronologico di transizione di stato del sistema complessivo.

Step 4: l'analisi del sistema è condotta per ciascun suo differente stato in modo da ottenere la funzione dell'indice di affidabilità $\Phi(S)$, dove S è lo stato del sistema. Il valore atteso della funzione dell'indice $\Phi(S)$ è indicato come $\mathbb{E}(\Phi)$. Tale valore è dato dalla seguente espressione:

$$\mathbb{E}(\Phi) = \sum_{S \in G} \Phi(S) \mathbb{P}(S) \tag{4.9}$$

dove G è l'insieme degli stati del sistema e $\mathbb{P}(S)$ è la probabilità di ciascun stato del sistema. Sostituendo la probabilità $\mathbb{P}(S)$ con la rispettiva frequenza relativa risulta che:

$$\mathbb{E}(\Phi) = \sum_{S \in G} \Phi(S) \frac{n(S)}{N} \tag{4.10}$$

dove N è il numero totale di campioni e $n(S)$ è il numero di occorrenze dello stato S. $\Phi(S)$ può essere ottenuta da un'appropriata analisi del sistema. Ad esempio, per determinare la probabilità di riduzione del carico del sistema, la funzione $\Phi(S)$ è definita come [Pereira e Pinto, 1992]:

$$\Phi(S) = \begin{cases} 1 \; se \; esiste \; una \; riduzione \; di \; carico \; associata \; allo \; stato \; S \\ 0 \; se \; non \; esiste \; una \; riduzione \; di \; carico \end{cases} \tag{4.11}$$

Le equazioni precedenti sono associate all'approccio dello state sampling (simulazione non sequenziale). Quando viene usata la tecnica di simulazione sequenziale, il concetto utilizzato per stimare il valore atteso dell'indice può essere espresso come segue:

$$\mathbb{E}(\Phi) = \frac{\sum_{i=1}^{NS}\left(\sum_{j=1}^{n_i(S)} \Phi(S_{j,i})\right)}{NS} \tag{4.12}$$

dove $n_i(S)$ è il numero di occorrenze dello stato S nell'anno i, $\Phi(S_{j,i})$ è la funzione dell'indice corrispondente alla j-esima occorrenza nell'anno i, NS è il numero degli anni della simulazione.

4.4-Convergenza della simulazione e criterio d'arresto

Il metodo Monte Carlo crea un processo fluttuante di convergenza e non c'è garanzia che un numero aggiuntivo di campioni possa condurre a un errore più piccolo. Risulta comunque vero che i limiti di errore diminuiscono all'aumentare del numero di campioni. Tuttavia non è ragionevole eseguire la simulazione per un numero estremamente ampio di campioni in quanto sarebbe richiesto un esteso tempo computazionale. È necessario dunque trovare un compromesso tra l'accuratezza richiesta e il tempo di calcolo. Lo scopo di un criterio di arresto è consentire che la simulazione continui ad essere effettuata fino a quando l'indice di affidabilità raggiunge un dato grado di accuratezza. Il parametro di base usato nel criterio di arresto è il coefficiente di variazione ed è derivato nella maniera espressa qui di seguito.

Un parametro fondamentale nella procedura di valutazione dell'affidabilità è il valore atteso di un dato indice di affidabilità. Pertanto le caratteristiche più importanti della simulazione Monte Carlo per l'analisi dell'affidabilità possono essere espresse da un punto di vista di valori medi [Billinton e Li, 1994].

Sia X l'indice di affidabilità da stimare. Nella simulazione sequenziale, il numero di campioni è pari al numero degli anni di simulazione. Il valore atteso dell'indice di affidabilità X è dato da:

$$\mathbb{E}(X) = \frac{1}{N}\sum_{i=1}^{N} x_i \tag{4.13}$$

dove x_i è il valore osservato di X nell'anno i, mentre N è il numero degli anni di simulazione.

La varianza dell'indice di affidabilità X è:

$$Var(X) = \frac{1}{N-1}\sum_{i=1}^{N}\left(x_i - \mathbb{E}(X)\right)^2 \tag{4.14}$$

L'incertezza sulla stima dell'indice può essere misurata attraverso la varianza del valore atteso:

$$Var\left(\mathbb{E}(X)\right) = \frac{Var(X)}{N} \tag{4.15}$$

La deviazione standard del valore atteso di X è pari a:

$$\sigma\left(\mathbb{E}(X)\right) = \sqrt{Var\left(\mathbb{E}(X)\right)} = \sqrt{\frac{Var(X)}{N}} \tag{4.16}$$

Il livello di accuratezza di una simulazione Monte Carlo può essere espresso dal coefficiente di variazione β che è definito nel modo seguente:

$$\beta = \frac{\sigma(\mathbb{E}(X))}{\mathbb{E}(X)} \tag{4.17}$$

Tale coefficiente può essere riespresso come:

$$\beta = \frac{1}{\mathbb{E}(X)}\sqrt{\frac{Var(X)}{N}} = \frac{\sigma(X)}{\mathbb{E}(X)\sqrt{N}} \tag{4.18}$$

dove $\sigma(X) = \sqrt{Var(X)}$. La simulazione può essere terminata quando viene raggiunto uno specifico valore di coefficiente di variazione. Il criterio di arresto che tipicamente si sceglie stabilisce che il processo iterativo termina quando il coefficiente di variazione è inferiore a una soglia prefissata:

$$\frac{\sigma(X)}{\mathbb{E}(X)\sqrt{N}} < \varepsilon \tag{4.19}$$

dove ε è il valore di tolleranza specificata, ad esempio 0.05.

Come mostrato nelle equazioni precedenti, il valore di β diminuisce all'aumentare del numero di anni di simulazione. È importante dunque notare che il livello di accuratezza specificato di una simulazione Monte Carlo è direttamente legato al numero di campioni della simulazione e non dipende dalla dimensione del sistema analizzato. Per tale motivo le tecniche di simulazione Monte Carlo sono molto adatte nell'analizzare sistemi grandi con caratteristiche complesse. È inoltre rilevante notare che lo sforzo computazionale è influenzato dal valore dell'indice da stimare, ovvero più il sistema è affidabile, più risulta difficoltoso stimare i valori.

4.5-Procedura di base del metodo Monte Carlo non sequenziale

Il metodo dello state sampling è usato per simulare un approccio non sequenziale. Nella tecnica dello state sampling, gli stati di tutti i componenti sono campionati in modo tale da ottenere un insieme non cronologico degli stati del sistema complessivo. La probabilità che si possa verificare un guasto su un componente è dato dal suo *FOR*. Tutti i componenti sono campionati estraendo numeri casuali U da una distribuzione uniforme nell'intervallo [0,1]. Per illustrare questa procedura, sia lo stato di un sistema con m componenti rappresentato dal vettore $S = (S_1, S_2, \dots, S_k, \dots, S_m)$, dove S_k denota lo stato del k-esimo componente. L'insieme S degli m componenti include gli stati di ogni elemento del sistema (generatori, linee, trasformatori, ecc.). Sia FOR_k il Forced Outage Rate del k-esimo componente. Lo stato S_k del k-esimo componente viene determinato nel modo seguente:

$$S_k = \begin{cases} 0\ (stato\ normale)\ se\ U \geq FOR_k \\ 1\ \ (stato\ di\ guasto)\ se\ U < FOR_k \end{cases} \tag{4.20}$$

Una volta che tutti gli stati dei componenti sono stati campionati, essi vengono combinati e in tal modo si può determinare lo stato del sistema globale. I passi seguiti nell'utilizzo della tecnica dello state sampling possono essere riassunti nella maniera seguente:

1. Si simula uno stato del sistema: unità di generazione, linee, trasformatori sono campionati confrontando i numeri casuali estratti con i *FOR* di questi componenti.

2. Se il sistema è in uno stato di normale funzionamento, allora non c'è riduzione di carico; in questo caso si ritorna al primo passaggio per effettuare il campionamento successivo e determinare un nuovo stato del sistema. Se nello stato campionato vi sono dei guasti, può essere necessaria una riduzione del carico; in tale caso si effettuerà un load flow per verificare se si è in questa situazione.

3. Se si hanno violazioni di vincoli, come ad esempio sovraccarichi di linee, potranno essere necessarie azioni correttive per alleviare i vincoli, come il ridispacciamento delle unità di generazione, riduzione di carichi.

4. Gli indici di affidabilità sono calcolati e aggiornati ad ogni iterazione. I passaggi 1-3 sono ripetuti fino a quando il coefficiente di variazione β è inferiore a una tolleranza specificata.

5-Applicazione di un modello statistico per elementi tempo-varianti

La simulazione Monte Carlo non sequenziale ha costi computazionali molto più bassi rispetto al metodo sequenziale, tuttavia rispetto a quest'ultimo la sua abilità nel rappresentare elementi tempo-varianti è molto limitata, principalmente per il fatto che, convenzionalmente, la dipendenza statistica tra le variabili casuali non viene considerata. In tale contesto, questa tesi propone l'applicazione di un nuovo modello stocastico [Borges e Dias, 2016] per la rappresentazione di elementi tempo-varianti in grado di preservare la dipendenza statistica tra di essi ed applicabile per la valutazione dell'affidabilità per mezzo della simulazione Monte Carlo non sequenziale. Il principale obiettivo di questo modello è di ottenere l'accuratezza della simulazione sequenziale e al tempo stesso oneri computazionali dell'ordine di quelli richiesti da una simulazione non sequenziale, anche quando vengono rappresentati molti elementi tempo-varianti. Il modello è basato sulla combinazione di tre attributi: caratterizzazione non parametrica delle variabili casuali, rappresentazione delle correlazioni statistiche non lineari tra le variabili, e il trattamento di problemi di elevata dimensionalità.

I metodi di valutazione dell'affidabilità basati sulla rappresentazione cronologica richiedono una conoscenza a priori delle serie temporali che descrivono gli elementi tempo-varianti. Pertanto, i modelli impiegati per costruire tali serie temporali sono di importanza critica tanto quanto il metodo stesso della simulazione in quanto questi modelli sono responsabili dell'ottenimento di casi realistici per analisi predittive coerenti.

Un approccio comune consiste nell'usare dati provenienti da serie storiche, assumendo che queste serie rappresentino processi stocastici stazionari e che i casi da analizzare possano essere definiti riproducendo i dati nella cronologia di funzionamento del sistema. Queste serie sono poi combinate con i modelli stocastici dei componenti del sistema, producendo in tal modo gli stati del sistema che saranno valutati sulla loro adeguatezza.

L'impiego di casi costruiti dalla ripetizione dei dati storici risulta semplice da comprendere e da implementare nella simulazione stocastica. Uno dei suoi principali vantaggi è che, in presenza di molti elementi tempo-varianti, gli effetti delle correlazioni tra le variabili saranno catturati dalla simulazione. Inoltre, sarà mantenuta ogni altra caratteristica statistica presente nelle osservazioni storiche.

Un'alternativa alla simulazione di dati storici è la simulazione di casi sintetici ottenuti impiegando modelli stocastici per le serie temporali. I modelli più comunemente utilizzati per questo tipo di approccio sono quelli autoregressivi (*AR*) e i loro derivati, ad esempio il modello regressivo a media mobile (*ARMA*) e il modello regressivo integrato a media mobile (*ARIMA*), che sono ampiamente utilizzati nella rappresentazione delle serie temporali di portate affluenti [Hipel eMcLeod, 1994], della domanda di carico e della generazione eolica[Billinton et al., 2009]. Le ragioni più comuni che spingono ad impiegare questi modelli nella rappresentazione delle serie temporali invece di applicare direttamente le serie storiche sono le seguenti:

1. La necessità di filtrare le serie storiche per rimuovere tendenze e nonstazionarietà, per il trattamento di errori e di dati mancanti, e per effettuare lo smoothing delle serie. I modelli stocastici hanno lo scopo di mantenere le principali caratteristiche originarie e di eliminare i problemi presenti nelle serie storiche.

2. L'esigenza di rappresentare una più ampia variabilità nei casi simulati, che accade quando i dati storici disponibili sono limitati. La generazione di casi sintetici per mezzo dell'applicazione di modelli stocastici viene usata per completare i dati storici.

Tuttavia, usando i modelli delle serie temporali si introducono alcune difficoltà aggiuntive. Un approccio comunemente utilizzato nell'analisi delle serie temporali è l'approccio parametrico, il quale richiede che le variabili casuali siano classificate in una famiglia di distribuzioni di probabilità, e successivamente lo sviluppo e gli adattamenti al modello vengono effettuati basandosi su questa assunzione. Per quanto riguarda le variabili casuali che sono rappresentate nello studio dei sistemi elettrici, le famiglie di distribuzioni che sono più ampiamente adottate sono, ad esempio, la distribuzione log-normale per rappresentare gli afflussi da corsi d'acqua, e la distribuzione di Weibull per rappresentare la velocità del vento. Uno svantaggio di questo tipo di approccio è che solitamente la stessa distribuzione non è il modo migliore per caratterizzare tutte le variabili dello stesso tipo nel sistema.

Un altro problema è che i modelli stocastici delle serie temporali generalmente riflettono la dipendenza statistica tra loro utilizzando correlazioni lineari. Una correlazione lineare non è sufficientemente accurata per spiegare la relazione tra tutte le variabili aleatorie a causa della presenza di relazioni non lineari tra di loro, come nel caso della direzione nonché del modulo della velocità del vento.

Inoltre, l'aumento del numero di variabili rappresentate tende a far sì che il numero di parametri del modello cresca esponenzialmente. A rigor di termini, un modello auto regressivo multivariato deve tenere in considerazione le correlazioni incrociate tra le variabili; cioè deve rappresentare la dipendenza di ogni

variabile in un determinato istante di tempo in relazione non solo a sé stessa, ma anche alle altre variabili in momenti precedenti. La complessità di un problema dunque aumenta esponenzialmente in rapporto alle sue dimensioni.

Va inoltre rilevata la difficoltà di rappresentare la dipendenza statistica tra elementi tempo-varianti con diversi livelli di discretizzazione. Un esempio è dato dalla rappresentazione della correlazione tra la generazione eolica di energia elettrica e gli afflussi da corsi d'acqua. L'incertezza intrinseca delle portate affluenti ha dinamiche molto più lente di quelle della generazione eolica in quanto gli afflussi sono misurati e modellati settimanalmente o mensilmente, mentre la produzione di energia da fonti eoliche richiede una rappresentazione oraria o persino una misurazione ad intervalli di 10 minuti.

Pertanto, lo scopo del modello applicato in questa tesi è quello di fornire un modo per la rappresentazione delle serie storiche che aggiri tutte queste restrizioni e che sia efficiente da un punto di vista computazionale.

Il metodo Monte Carlo non sequenziale convenzionale assume che gli stati dei componenti siano statisticamente indipendenti. Pertanto, in tale caso la probabilità di occorrenza dell'i-esimo stato di un sistema con m componenti rappresentato da $S = (S_1, S_2, \dots, S_k, \dots, S_m)$ è data da:

$$\mathbb{P}(S^i) = \prod_{k=1}^{m} \mathbb{P}(S_k) \tag{5.1}$$

dove $\mathbb{P}(S_k)$ è la probabilità di occorrenza associata al k-esimo componente. Tuttavia, l'introduzione di variabili statisticamente dipendenti nel problema implica che debbano essere tenute in conto le probabilità congiunte nell'equazione precedente, arrivando dunque ad ottenere la seguente espressione:

$$\mathbb{P}(S^i) = \mathbb{P}(v_1, v_2, \dots, v_w) \prod_{k=1}^{m} \mathbb{P}(S_k) \tag{5.2}$$

dove $v_1, v_2, \dots, v_w$ indicano le W variabili statisticamente dipendenti del sistema e $\mathbb{P}(v_1, v_2, \dots, v_w)$ rappresenta la distribuzione di probabilità congiunta che caratterizza queste variabili. Queste probabilità congiunte impediscono un calcolo accurato degli indici di affidabilità se si sceglie di usare la simulazione Monte Carlo non sequenziale convenzionale. Pertanto, per applicare il metodo Monte Carlo non sequenziale nella valutazione dell'affidabilità di un sistema con elementi dipendenti statisticamente, la simulazione deve essere modificata per incorporare l'informazione contenuta nelle probabilità congiunte di questi elementi.

Un approccio possibile consiste nella rappresentazione esplicita delle probabilità congiunte delle variabili statisticamente dipendenti, oppure nell'esplicita rappresentazione della funzione $\mathbb{P}(v_1, v_2, \dots, v_w)$. Un caso tipico per l'applicazione di questo approccio è dato quando le variabili considerate sono continue, soddisfano le condizioni di normalità, e possono essere congiuntamente caratterizzate da una singola funzione di densità di probabilità normale multivariata.

Sebbene questa procedura è semplice, solitamente è molto improbabile che tutte le variabili casuali che rappresentano gli elementi tempo-varianti soddisfino le condizioni di normalità di una distribuzione normale

multivariata. Per superare questa limitazione, si possono applicare alle variabili aleatorie originali funzioni di trasformazione in modo da ottenere nuove variabili che possono essere caratterizzate singolarmente come variabili casuali normalmente distribuite, come mostrato in Fig. 5.

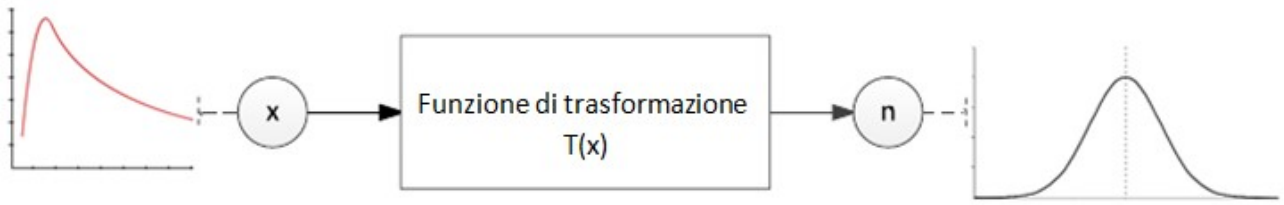

Figura 5-Trasformazione delle variabili originali in nuove, distribuite normalmente.

Successivamente viene effettuato lo state sampling secondo le densità di probabilità di queste variabili normali trasformate. Alla fine del processo, si applicano funzioni di trasformazione inverse, in maniera tale da produrre campioni con le stesse distribuzioni delle variabili originali (Fig. 6).

Figura 6-Generazione dei campioni delle variabili originali da quelle trasformate.

Questa procedura viene abbastanza comunemente utilizzata per stimare modelli stocastici autoregressivi per la rappresentazione di afflussi, in cui si assume tipicamente che le variabili casuali seguano una distribuzione di probabilità log-normale. In tale caso, ciò implica che la funzione logaritmica possa essere usata come funzione di trasformazione diretta, e di conseguenza si può usare la funzione esponenziale come funzione di trasformazione inversa.

Un altro possibile approccio è la decomposizione delle distribuzioni di probabilità congiunte in funzioni di distribuzione di probabilità condizionata applicando la regola della catena:

$$\mathbb{P}(v_1, v_2, \dots, v_w) = \mathbb{P}(v_1)\mathbb{P}(v_2|v_1)\mathbb{P}(v_3|v_2, v_1) \dots \mathbb{P}(v_w|v_{w-1}, \dots, v_2, v_1) \qquad (5.3)$$

Pertanto, le funzioni di distribuzione di probabilità condizionata sono costruite per i termini $\mathbb{P}(v_2|v_1)$,$\mathbb{P}(v_3|v_2, v_1)$, etc., che possono essere usati nella simulazione Monte Carlo non sequenziale per generare i campioni delle variabili statisticamente dipendenti. Sebbene questo approccio diminuisce la

complessità delle funzioni di distribuzione di probabilità rappresentate, esso è limitato alla rappresentazione di un numero piccolo di variabili, poiché il numero di parametri delle funzioni di distribuzione di probabilità condizionata cresce esponenzialmente all'aumentare del numero di variabili da cui esso dipende.

L'obiettivo del modello che qui ci si propone di applicare è dunque quello di offrire una nuova alternativa per la caratterizzazione delle funzioni di densità di probabilità congiunta di tali variabili casuali dipendenti statisticamente, adottando un approccio non parametrico. Il metodo basato sulla Kernel Density Estimation (*KDE*) è una buona alternativa per la stima non parametrica di funzioni di densità di probabilità di variabili continue ed è stato adottato nel modello proposto per il trattamento degli elementi tempo-varianti.

Un'iniziale limitazione all'applicazione del metodo *KDE* era dovuta al fatto che era stato dapprima sviluppato per rappresentare le funzioni di densità di probabilità di una singola variabile casuale, e la funzione che deve essere stimata è la funzione di densità di probabilità congiunta, corrispondente al caso multidimensionale. La conservazione delle relazioni statistiche non lineari tra più variabili aleatorie è direttamente collegata alla conservazione dell'*Informazione Mutua* condivisa da queste variabili. L'Informazione Mutua è un concetto derivante dalla teoria dell'informazione (Information Theory, IT), che è sensibile a qualsiasi tipo di dipendenza tra le variabili, ed è pari a zero soltanto per due variabili casuali che sono strettamente indipendenti.

In tal modo, l'approccio adottato nel modello proposto permette di far sì che la funzione di densità di probabilità congiunta da stimare possa essere scomposta in due componenti. Il primo componente descrive le variabili indipendentemente usando stime, basate sul *KDE*, delle funzioni di densità di probabilità marginale, e il secondo componente codifica la dipendenza statistica tra queste variabili tramite la conservazione dell'informazione mutua condivisa da esse.

Per i sistemi di grandi dimensioni, la rappresentazione dell'informazione mutua tra tutte le variabili rimane problematica in quanto sussiste una complessità crescente all'aumentare della dimensionalità. Per superare questo problema, viene utilizzata una rappresentazione della dipendenza statistica tra le variabili basata su reti bayesiane, in modo da evitare la crescita esponenziale della complessità del modello all'aumentare del numero di variabili considerate. Le reti bayesiane rappresentano dei modelli derivati dalla combinazione della teoria della probabilità con la teoria dei grafi e possono rappresentare in maniera efficiente e compatta le distribuzioni di probabilità congiunta di variabili di n dimensioni. Pertanto, esse costituiscono una valida alternativa per rappresentare le distribuzioni di probabilità congiunta di un ampio numero di variabili casuali.

6-Modelli autoregressivi

L'analisi statistica di serie temporali è il passo preliminare per definire criteri di modellazione che possano consentire la generazione di serie sintetiche anche di notevole estensione. In tale ambito, si possono utilizzare modelli autoregressivi (*AR*), a media mobile (*MA*) ed autoregressivi a media mobile (*ARMA*).

I modelli autoregressivi sono stati ampiamente utilizzati in idrologia per modellare serie annuali e serie con componenti stagionali significative. Nel primo caso i parametri di modello possono essere assunti costanti; nel secondo caso i parametri sono in genere stagionalizzati.

Nel processo autoregressivo, il valore assunto al tempo i-esimo da una variabile x è esprimibile come combinazione lineare dei valori assunti in un numero finito di intervalli precedenti e di un rumore additivo, indicato con ε. In un processo autoregressivo di ordine p la variabile è espressa come combinazione dei valori assunti in p intervalli precedenti.

Si introduce la variabile trasformata definita come la differenza tra la variabile casuale x e la sua media $\mu(x)$:

$$z = x - \mu(x) \tag{6.1}$$

Se il processo è stazionario, la media $\mu(x)$ è costante. Il modello autoregressivo di ordine p, che si indica comunemente come AR(p), è espresso dalla seguente relazione lineare:

$$z_i = \phi_1 z_{i-1} + \phi_2 z_{i-2} + \cdots + \phi_p z_{i-p} + \varepsilon_i \tag{6.2}$$

dove $\phi_1, \phi_2, \dots, \phi_p$ sono i parametri del modello $AR(p)$ i cui valori sono costanti nel tempo. In un processo stazionario, in genere la ε si assume distribuita secondo una legge normale con media nulla e varianza $\sigma^2(\varepsilon)$, ovvero le caratteristiche del rumore bianco.

Il processo è quindi complessivamente individuato da $p + 2$ parametri: $\mu(x), \phi_1, \phi_2, \dots, \phi_p, \sigma^2(\varepsilon)$. Indicando con B l'operatore lineare definito per mezzo della relazione:

$$Bz_i = z_{i-1} \tag{6.3}$$

Dalla quale deriva l'espressione valida per qualunque valore di p:

$$B^p z_i = z_{i-p} \tag{6.4}$$

Introducendo inoltre l'operatore autoregressivo espresso dalla seguente relazione:

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p \tag{6.5}$$

Il processo autoregressivo si può rappresentare secondo la seguente espressione compatta:

$$\phi(B) z_i = \varepsilon_i \tag{6.6}$$

Nel processo di media mobile il valore assunto al tempo i-esimo dalla variabile trasformata $z = x - \mu(x)$ è una combinazione lineare dei valori assunti dal rumore bianco ε al tempo considerato e ad un numero finito di intervalli precedenti. Il processo $MA(q)$, di media mobile di ordine q, è rappresentato dalla relazione:

$$z_i = \varepsilon_i - \theta_1 \varepsilon_{i-1} - \theta_2 \varepsilon_{i-2} - \cdots - \theta_q \varepsilon_{i-q} \tag{6.7}$$

Introducendo l'operatore di media mobile di ordine q:

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q \tag{6.8}$$

Si può rappresentare il processo *MA*(q) con l'espressione compatta:

$$z_i = \theta(B)\varepsilon_i \tag{6.9}$$

Il processo *MA*(q) è dunque individuato dai $q + 2$ parametri: $\mu(x), \theta_1, \theta_2, \dots, \theta_q, \sigma^2(\varepsilon)$.

Per descrivere in modo soddisfacente il comportamento di una serie temporale con un processo autoregressivo di ordine p oppure con un processo di media mobile di ordine q è possibile che si debba prendere in considerazione un valore di p o di q molto alto.

Senza aumentare eccessivamente il numero dei parametri del processo può risultare vantaggioso usare un processo misto che include un processo autoregressivo di ordine p ed un processo a media mobile di ordine q.

Il processo misto *ARMA*(p,q) è rappresentato dalla relazione:

$$z_i = \phi_1 z_{i-1} + \phi_2 z_{i-2} + \cdots + \phi_p z_{i-p} + \varepsilon_i - \theta_1 \varepsilon_{i-1} - \theta_2 \varepsilon_{i-2} - \cdots - \theta_q \varepsilon_{i-q} \tag{6.10}$$

Il processo *ARMA*(p,q) è dunque individuato da $p + q + 2$ parametri: $\mu(x), \phi_1, \phi_2, \dots, \phi_p, \theta_1, \theta_2, \dots, \theta_q$, $\sigma^2(\varepsilon)$. Utilizzando gli operatori $\phi(B)$ e $\theta(B)$ il processo si può scrivere con l'espressione compatta:

$$\phi(B) z_i = \theta(B)\varepsilon_i \tag{6.11}$$

Ovvero:

$$z_i = \theta(B)\phi(B)^{-1}\varepsilon_i \tag{6.12}$$

Il processo *ARMA*(p,q) si può quindi pensare come un processo autoregressivo di ordine p nel quale il rumore è sostituito da un processo di media mobile di ordine q.

Vale inoltre osservare che i processi autoregressivi e quelli di media mobile costituiscono dei casi particolari del processo misto. Ad esempio un processo *AR*(p) equivale ad un processo *ARMA*(p,0) ed un processo *MA*(q) equivale ad un processo *ARMA*(0,q).

Poiché il generico termine z_i di un processo *ARMA*(p,q) è una funzione lineare dei p termini $z_{i-1}, z_{i-2}, \dots, z_{i-p}$ e dei q termini $\varepsilon_{i-1}, \varepsilon_{i-2}, \dots, \varepsilon_{i-q}$, il problema di decidere sulla significatività dell'introduzione di nuove variabili in una regressione lineare. Le considerazioni sulle quali si basa la scelta

sono di due tipi: da una parte si valuta la significatività dei risultati da un punto di vista puramente statistico, d'altra parte si valuta la congruenza di considerare le variabili con quanto si conosce circa i fenomeni fisici.

Uno degli esempi più semplici della famiglia *ARMA* è il processo autoregressivo del primo ordine *AR*(1). Nelle applicazioni idrologiche si utilizzano normalmente solo processi *AR*(1) stazionari.

Introducendo, al solito, la variabile trasformata $z = x - \mu(x)$ che ha media nulla ed autocovarianza coincidenti con quelle di uguale ordine della variabile originaria x (si ricorda che la varianza è una autocovarianza di ordine zero), il processo si rappresenta con l'espressione:

$$z_i = \phi z_{i-1} + \varepsilon_i \tag{6.13}$$

Nella quale ϕ rappresenta l'unico parametro autoregressivo del processo. Il rumore è caratterizzato da media nulla (rumore bianco) e da varianza data dall'espressione:

$$\sigma^2(\varepsilon) = (1 - \phi^2)\sigma^2(z) \tag{6.14}$$

Il coefficiente di autocorrelazione del primo ordine ha espressione:

$$\rho_1(z) = \phi \tag{6.15}$$

Mentre il generico coefficiente di autocorrelazione ha espressione:

$$\rho_k(z) = \phi \rho_{k-1}(z) \tag{6.16}$$

E pertanto si ricava:

$$\rho_k(z) = \phi^k \tag{6.17}$$

Il processo *ARMA*(1,1) è il più semplice di tutti i processo misti. Esso è rappresentato dalla seguente espressione:

$$z_i = \phi z_{i-1} + \varepsilon_i - \theta \varepsilon_{i-1} \tag{6.18}$$

Nella quale ϕ e θ sono rispettivamente l'unico parametro autoregressivo e l'unico parametro di media mobile. La variabile trasformata $z' = z_i - \phi z_{i-1}$ modifica il processo *ARMA*(1,1) costituito dalla variabile z, in un processo *MA*(1) costituito dalla variabile z'.

Il processo *ARMA*(1,1) permette di tener conto del fenomeno della persistenza a lungo termine molto meglio del processo autoregressivo del primo ordine, grazie al grado di libertà in più rappresentato dal parametro θ che compare in questo processo.

7-Kernel Density Estimation

Storicamente, le serie degli afflussi da corsi d'acqua sono state rappresentate da modelli parametrici autoregressivi periodici, assumendo la premessa che gli afflussi abbiano una distribuzione di probabilità che appartiene a una specifica famiglia di distribuzioni, solitamente la distribuzione log-normale, e la parametrizzazione del modello viene sviluppata in base a questa premessa.

Tuttavia non sempre la distribuzione log-normale è la più adatta a caratterizzare la distribuzione di probabilità di una serie di afflussi, e la più comune conseguenza di questa impropria caratterizzazione è l'ottenimento di un lento decadimento della coda della distribuzione, comportando una generazione di campioni con valori molto più alti di quelli delle osservazioni storiche. Pertanto alcune serie possono essere meglio rappresentate da un certo tipo di distribuzione o da un altro, e non esiste un solo tipo di distribuzione che è ideale per tutte le serie.

Un approccio alternativo per rappresentare tali serie è quello non parametrico, in cui una stima della distribuzione di probabilità viene direttamente ottenuta dalle osservazioni storiche, senza la necessità di assumere a priori una distribuzione standard delle variabili casuali. L'istogramma è la rappresentazione non parametrica più semplice. Tuttavia, l'istogramma presenta alcune limitazioni importanti, quali la dipendenza della stima con l'ampiezza dell'intervallo di discretizzazione della serie, oltre al fatto che esso non costituisce necessariamente una funzione continua.

Un metodo di stima della distribuzione di probabilità più efficace e più flessibile è il Kernel Density Estimation [Borges e Dias, 2014], che è in grado di superare le limitazioni dell'istogramma precedentemente citate e che presenta interessanti proprietà aggiuntive. Il metodo viene applicato per ottenere funzioni di trasformazione che sono poi applicate alla serie temporale per consentire l'uso di un modello autoregressivo non parametrico, indipendentemente dalla distribuzione di probabilità della serie storica. Il metodo è utilizzabile non soltanto per la rappresentazione di serie di portate affluenti, ma anche di serie temporali di altre variabili.

Data una variabile casuale X con funzione di densità di probabilità $f(x)$, risulta valida la seguente approssimazione:

$$f(x) \approx \frac{1}{2h}\mathbb{P}(x-h<X<x+h) \tag{7.1}$$

dove il parametro h è chiamato *larghezza di banda*. Una stima di questa funzione di densità è data da:

$$\hat{f}(x) = \frac{1}{2h}\frac{\sum_i 1_{[x-h,x+h]}(x_i)}{n} \tag{7.2}$$

dove x_i corrisponde all'i-esima osservazione di X, n è il numero di osservazioni e $1_A(x)$ è la funzione indicatrice, definita come:

$$1_A(x) = \begin{cases} 1 \; se \; x \in A \\ 0 \; se \; x \notin A \end{cases} \tag{7.3}$$

La stima della funzione di densità può essere riscritta come:

$$\hat{f}(x) = \frac{1}{n}\sum_{i=1}^{n} w(x - x_i, h) \tag{7.4}$$

dove $w(t, h)$ è definita come:

$$w(t, h) = \begin{cases} \frac{1}{2h} \; se \; |t| < h \\ 0 \; altrimenti \end{cases} \tag{7.5}$$

In linea più generale, si può definire la seguente espressione:

$$w(t, h) = \frac{1}{h} K\left(\frac{t}{h}\right) \tag{7.6}$$

dove K è chiamata *funzione kernel*. La funzione kernel definisce la forma della funzione di densità stimata, mentre la larghezza di banda determina quanto è regolare tale forma. Tra le funzioni kernel solitamente utilizzate, la più nota è quella gaussiana, data da:

$$K(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \tag{7.7}$$

In Fig. 7 viene mostrata una distribuzione bimodale stimata da un insieme di osservazioni, con le corrispondenti funzioni kernel gaussiane sovrapposte. La funzione di densità stimata può essere visualizzata come una combinazione di gaussiane con gli stessi valori di deviazione standard, ma con medie diverse, centrate sui valori delle osservazioni. Siccome variabili quali gli afflussi da corsi d'acqua o la velocità del vento sono definite per valori positivi, viene usato il concetto di densità con supporto limitato per garantire il corretto dominio delle variabili casuali nelle densità di probabilità.

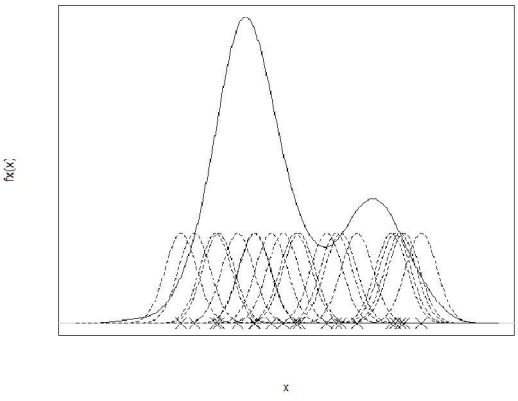

Figura 7-Funzione di distribuzione stimata con il *KDE*.

Per un insieme di osservazioni per le quali si vuole ottenere una stima della funzione di densità di probabilità, data la scelta di una certa funzione kernel, il problema diventa la determinazione della larghezza di banda. Questa determinazione è critica, perché se non è effettuata correttamente la stima porta ad ottenere una funzione eccessivamente affetta da irregolarità (*undersmoothed*) oppure eccessivamente "liscia" (*oversmoothed*). In Fig.8 vengono riportate due funzioni di distribuzione di probabilità ottenute per lo stesso insieme di dati usato per ottenere la distribuzione presentata in Fig.8. Il grafico in alto corrisponde ad una stima effettuata utilizzando una larghezza di banda molto più bassa rispetto al parametro ideale, mentre il grafico in basso corrisponde ad una stima ottenuta usando una larghezza di banda molto più alta rispetto a quella ideale. Si può notare che un valore basso di larghezza di banda porta ad una distribuzione rumorosa e poco regolare, mentre una larghezza di banda elevata porta ad una distribuzione eccessivamente "liscia", tale per cui non risulta più possibile identificare la caratteristica bimodale della distribuzione.

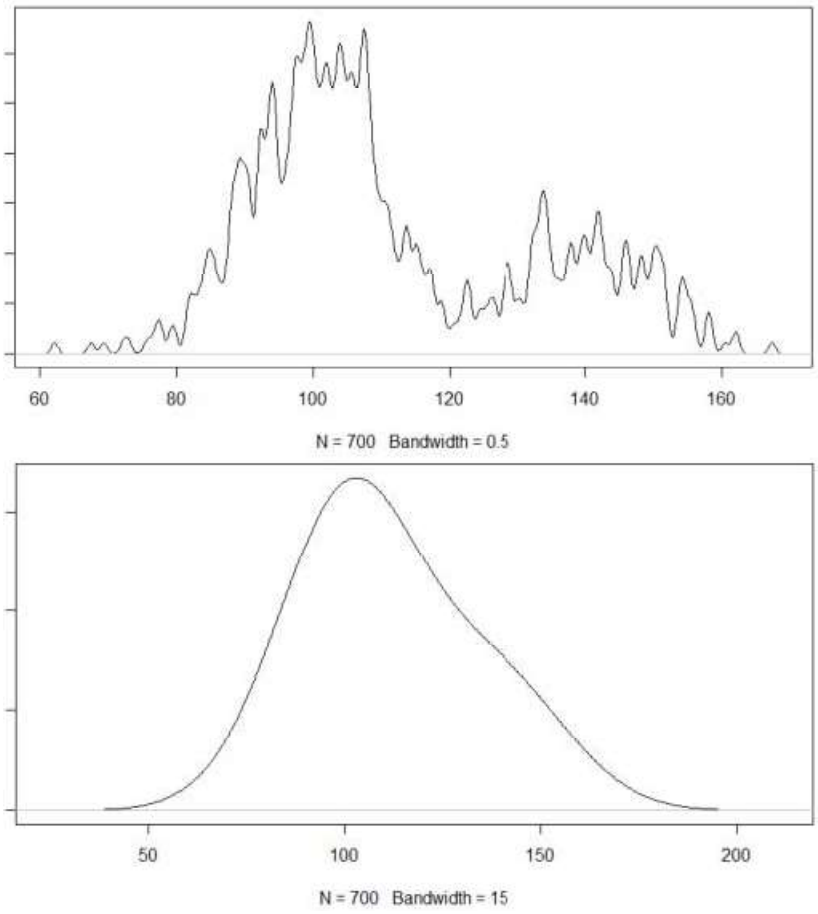

Figura 8-stima undersmoothed (in alto) e oversmoothed (in basso).

La larghezza di banda è dunque il parametro di smoothing. La scelta della larghezza di banda è essenzialmente un problema di ottimizzazione, in cui si intende minimizzare uno o più errori di misura della stima. Uno degli errori di misura più importanti è l'errore quadratico medio integrato (Mean Integrated Square Error, *MISE*), definito nella maniera seguente [Sheater e Jones, 1991]:

$$MISE = \mathbb{E} \int \left(\hat{f}_n(x,h) - f(x) \right)^2 dx \qquad (7.8)$$

dove $f(x)$ è la funzione di densità reale, ignota, $\hat{f}_n(x,h)$ è la funzione di densità stimata basandosi su n campioni della variabile casuale x.

La larghezza di banda viene scelta in modo tale che essa diminuisca all'aumentare del numero di campioni e in maniera tale per cui il *MISE* tenda a zero. Quando il valore di h è troppo piccolo, esso causa numerose oscillazioni nei valori della funzione stimata $\hat{f}_n(x,h)$ intorno ai punti di massimo e di minimo (punti critici), mentre quando il suo valore è troppo grande, la funzione ottenuta tende a sovrastimare i punti critici, ed in tale modo diventa eccessivamente liscia e regolare, non riflettendo le proprietà dei campioni empirici reali.

8-Informazione mutua

La mutua informazione (Mutual Information, *MI*) di due variabili casuali è una quantità che misura la mutua dipendenza tra le due variabili. Formalmente, l'informazione mutua di due variabili casuali discrete X e Y può essere definita nel modo seguente [Kraskov et al., 2004]:

$$I(X,Y) = \sum_{y\in Y}\sum_{x\in X} p(x,y)\, log_b\left(\frac{p(x,y)}{p_X(x)p_Y(y)}\right) \tag{8.1}$$

dove $p(x,y)$ è la funzione di distribuzione di probabilità congiunta di X e Y, mentre $p_X(x)$ e $p_Y(y)$ sono le funzioni di distribuzione di probabilità marginale rispettivamente di X e Y, che possono essere calcolate in base alle seguenti espressioni:

$$p_X(x) = \sum_{y\in Y}\mathbb{P}(X=x,Y=y) = \sum_{y\in Y}\mathbb{P}(X=x|Y=y)\mathbb{P}(Y=y) \tag{8.2}$$

$$p_y(y) = \sum_{x\in X}\mathbb{P}(X=x,Y=y) = \sum_{x\in X}\mathbb{P}(Y=y|X=x)\mathbb{P}(X=x) \tag{8.3}$$

dove $\mathbb{P}(X=x,Y=y)$ rappresenta la probabilità congiunta di X e Y, mentre $\mathbb{P}(X=x|Y=y)$ è la probabilità condizionata di X dato Y e $\mathbb{P}(Y=y|X=x)$ è la probabilità condizionata di Y dato X.

La base del logaritmo determina l'unità con la quale la mutua informazione viene misurata. In particolare, ponendo la base due l'informazione mutua è misurata in bit; utilizzando invece il logaritmo naturale, l'unità di misura è il nat; si ha che $1\ bit = \ln 2\ nat = 0.69315\ nat$.

Nel caso di variabili continue, l'espressione dell'informazione mutua diventa la seguente:

$$I(X,Y) = \int_Y \int_X p(x,y)\log_b\left(\frac{p(x,y)}{p_X(x)p_Y(y)}\right)dxdy \tag{8.4}$$

dove in questo caso $p(x,y)$ indica la funzione di densità di probabilità congiunta di X e Y, mentre $p_X(x)$ e $p_Y(y)$ rappresentano ora le funzioni di densità di probabilità marginale rispettivamente di X e Y, esprimibili come:

$$p_X(x) = \int p(x,y)dy = \int p_{X|Y}(x|y)p_Y(y)dy \tag{8.5}$$

$$p_Y(y) = \int p(x,y)dx = \int p_{Y|X}(y|x)p_X(x)dx \tag{8.6}$$

dove $p_{X|Y}(x|y)$ è la funzione di densità di probabilità condizionata di X dato Y e $p_{Y|X}(y|x)$ rappresenta la funzione di densità di probabilità condizionata di Y dato X.

L'informazione mutua quantifica la dipendenza tra la distribuzione congiunta di X e Y e quale sarebbe la distribuzione congiunta se X e Y fossero indipendenti; infatti $I(X,Y) = 0$ se e solo se X e Y sono variabili casuali indipendenti. Questo è facile da verificare nella maniera seguente: se X e Y sono indipendenti, allora si ha che $p(x,y) = p_X(x)p_Y(y)$ e perciò:

$$\log_b\left(\frac{p(x,y)}{p_X(x)p_Y(y)}\right) = \log_b 1 = 0 \tag{8.7}$$

Inoltre la mutua informazione è non negativa, cioè $I(X,Y) \geq 0$, e simmetrica, ossia $I(X,Y) = I(Y,X)$.

L'informazione mutua è strettamente legata a concetti derivati dalla teoria dell'informazione. In tale contesto, si definisce misura dell'informazione o autoinformazione relativa alla variabile casuale X la quantità:

$$I(X) = \log_b\left(\frac{1}{\mathbb{P}(x)}\right) = -\log_b \mathbb{P}(x) \tag{8.8}$$

dove $\mathbb{P}(x)$ è la probabilità che si verifichi l'evento x. Dalla definizione, si nota come l'informazione contenuta nell'evento x è tanto più grande quanto più bassa è la sua probabilità. La base del logaritmo determina l'unità di misura del contenuto informativo, in analogia a quanto detto riguardo alla mutua informazione.

Si introduce ora il concetto di entropia, il cui studio nell'ambito della teoria dell'informazione si deve principalmente a Claude Shannon. L'entropia rappresenta una misura di incertezza di una variabile casuale ed è definita come il valore atteso dell'autoinformazione, ovvero l'informazione media contenuta in ogni evento x:

$$\mathbb{H}(X) = \mathbb{E}(I(X)) = \mathbb{E}(-\log_b \mathbb{P}(X)) \tag{8.9}$$

Se X è una variabile aleatoria discreta il valore atteso si riduce ad una media dell'informazione di ogni evento x_i pesata con la propria probabilità $\mathbb{P}(x_i)$:

$$\mathbb{H}(X) = -\sum_{i=1}^{N} \mathbb{P}(x_i)\log_b \mathbb{P}(x_i) \tag{8.10}$$

dove N è il numero degli eventi che possono accadere. Se la variabile casuale X è continua, il valore atteso si calcola attraverso un integrale:

$$\mathbb{H}(X) = -\int \mathbb{P}(x)\log_b \mathbb{P}(x)\,dx \tag{8.11}$$

Estendendo ora il concetto ad una coppia di variabili casuali, si arriva alla definizione dell'entropia congiunta. L'entropia congiunta $\mathbb{H}(X,Y)$ di una coppia di variabili casuali discrete X e Y con funzione di distribuzione di probabilità congiunta $p(x,y)$ è definita come:

$$\mathbb{H}(X,Y) = -\sum_{x\in X}\sum_{y\in Y} p(x,y)\log_b p(x,y) \tag{8.12}$$

che può anche essere espressa come:

$$\mathbb{H}(X,Y) = \mathbb{E}(-\log_b p(X,Y)) \tag{8.13}$$

Nel caso di variabili casuali continue l'espressione diventa la seguente:

$$\mathbb{H}(X,Y) = -\int_Y \int_X p(x,y)\log_b p(x,y)\,dxdy \tag{8.14}$$

Si introduce inoltre l'entropia condizionata di una variabile aleatoria data un'altra. Se $\mathbb{H}(X|Y=y)$ è l'entropia della variabile X condizionata dalla variabile Y allora l'entropia condizionata $\mathbb{H}(X|Y)$ è il risultato della media pesata di $\mathbb{H}(X|Y=y)$ su tutti i possibili valori y che la Y può assumere. Per una coppia di variabili casuali discrete X e Y l'entropia condizionata $\mathbb{H}(X|Y)$ si definisce come:

$$\mathbb{H}(X|Y) = \sum_{y\in Y}\mathbb{H}(X|Y=y)p_Y(y) =$$

$$= -\sum_{y\in Y}\sum_{x\in X} p_{X|Y}(x|y)p_Y(y)\log_b p_{X|Y}(x|y) =$$

$$= -\sum_{y\in Y}\sum_{x\in X} p(x,y)\log_b p_{X|Y}(x|y) \tag{8.15}$$

Mentre nel caso di variabili continue l'entropia condizionata $\mathbb{H}(X|Y)$ è pari a:

$$\mathbb{H}(X|Y) = \int_Y \int_X p_{X|Y}(x|y)p_Y(y)\log_b p_{X|Y}(x|y)\,dxdy =$$

$$= \int_Y \int_X p(x,y)\log_b p_{X|Y}(x|y)\,dxdy \tag{8.16}$$

In pratica, l'entropia condizionata $\mathbb{H}(X|Y)$ esprime la quantità di informazione da aggiungere per conoscere l'evento X quando già si conosce l'evento Y. In tale contesto infatti condizionare significa ridurre la necessaria quantità di informazione da fornire. Di conseguenza, l'entropia condizionata $\mathbb{H}(X|Y)$ è sempre compresa tra zero e l'entropia della variabile X:

$$0 \le \mathbb{H}(X|Y) \le \mathbb{H}(X) \tag{8.17}$$

I casi limite avvengono quando l'evento Y è completamente determinato da X e quando i due eventi sono indipendenti. La più naturale definizione di entropia congiunta e di quella condizionata è mostrata dal fatto che l'entropia di una coppia di variabili casuali è l'entropia di una più l'entropia condizionata dell'altra. Questo è provato dalla seguente relazione:

$$\mathbb{H}(X,Y) = \mathbb{H}(X) + \mathbb{H}(Y|X) \tag{8.18}$$

La mutua informazione è legata all'entropia attraverso le relazioni seguenti:

$$I(X,Y) = \mathbb{H}(X) - \mathbb{H}(X|Y) \tag{8.19}$$

In tal modo la mutua informazione $I(X,Y)$ può essere intesa come la riduzione nell'incertezza di X dovuta alla conoscenza di Y. Per simmetria, segue anche che:

$$I(X,Y) = \mathbb{H}(Y) - \mathbb{H}(Y|X) \tag{8.20}$$

Inoltre, dalle relazioni precedenti, si ricava che:

$$I(X,Y) = \mathbb{H}(X) + \mathbb{H}(Y) - \mathbb{H}(X,Y) \tag{8.21}$$

Infine, si nota che:

$$I(X,X) = \mathbb{H}(X) - \mathbb{H}(X|X) = \mathbb{H}(X) \tag{8.22}$$

Così la mutua informazione di una variabile casuale con se stessa è l'entropia della variabile casuale.

9-Reti bayesiane

Una rete bayesiana è una rappresentazioni grafica di un modello probabilistico, ovvero la riproduzione di una distribuzione di probabilità su un insieme di variabili.

Le reti bayesiane sono definite tramite la specificazione di due componenti:

Una componente qualitativa: un grafo aciclico orientato (Directed acyclic graph, *DAG*), indicato con $\mathcal{G} = (V, A)$, dove $V = \{v_1, v_2, \dots, v_n\}$ rappresentano i nodi che sono in corrispondenza biunivoca con l'insieme $\boldsymbol{X} = \{X_1, X_2, \dots, X_n\}$ delle variabili aleatorie, pertanto in tale contesto verrà utilizzato il termine nodo e variabile casuale in modo interscambiabile; $A \subset V \times V$ rappresenta l'insieme degli archi e sono coppie ordinate di elementi di V; ogni arco rappresenta la dipendenza condizionata esistente tra due nodi. Nel caso delle reti bayesiane gli archi sono orientati ($v_1 \to v_2$, o anche $X_1 \to X_2$) ovvero si ha che $(v_1, v_2) \in A \wedge (v_2, v_1) \notin A$; si parla dunque di grafi orientati.

La direzionalità di questa relazione permette di definire:

- genitori di v tutti quei nodi $u \in V$: $(u, v) \in A$, ovvero i nodi u da cui parte un arco verso v,
- figli di v tutti quei nodi $u \in V$: $(v, u) \in A$, ovvero i nodi u in cui arriva un arco da v,
- coniugi di v tutti quei nodi $u \in V$: $(u, w) \in A \wedge (v, w) \in A$, ovvero i nodi u che condividono un figlio con v,
- discendenti di v tutti quei nodi $u \in V$ per cui esiste un cammino che porta da v ad u ($v \to \cdots \to u$),

- predecessori di v tutti quei nodi $u \in V$ per cui esiste un cammino che viceversa porta da u a v $(u \to \cdots \to v)$.

Anche se due nodi non sono adiacenti (ovvero se non esiste un arco che li collega) si dicono connessi se esiste un cammino (path) che li unisce, ovvero una sequenza di nodi $v_1, v_2, \dots, v_k$ a due a due adiacenti tale che:

$$\begin{cases} (u, v_1) \in A \\ (v_i, v_{i+1}) \in A \vee (v_{i+1}, v_i) \in A, i = 1, 2, \dots, k-1 \\ (v_k, v) \in A \end{cases} \tag{9.1}$$

indipendentemente dalla direzione dei singoli archi. L'unico vincolo a questo riguardo è che nessun cammino può formare un ciclo, ossia un cammino in cui il punto di partenza coincide con quello di arrivo e tutti gli archi sono concordi; per tale motivo il grafo è detto aciclico.

Nella seguente trattazione i genitori di un nodo X_i saranno indicati con $Pa(X_i)$, i nodi figli con $Ch(X_i)$.

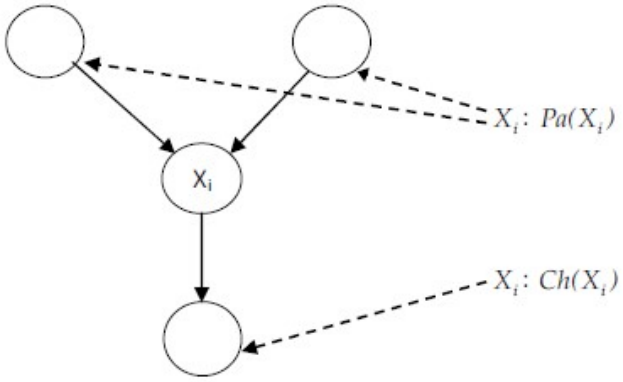

Figura 9-Esempio di rete bayesiana.

La seconda componente che definisce una rete bayesiana è quella quantitativa, rappresentata da un insieme di distribuzioni locali di probabilità, ciascuna associata ad una variabile e condizionata ad ogni configurazione dei suoi genitori. L'insieme delle distribuzioni locali di probabilità specificano la distribuzione congiunta dell'insieme di variabili. Queste distribuzioni di probabilità sono anche dette parametri della rete bayesiana. La distribuzione di ogni variabile è influenzata solamente dalle distribuzioni dei suoi diretti vicini all'interno della struttura. Quindi per ogni nodo viene definita una funzione di probabilità condizionata che quantifica gli effetti che i genitori hanno su quel dato nodo. Se un nodo non ha genitori, viene specificata una funzione di probabilità marginale.

Una rete bayesiana è in grado di esprimere le relazioni che legano le variabili aleatorie di un modello probabilistico tramite dipendenze e indipendenze condizionate, e permette quindi di trattare in modo intercambiabile nodi del grafo e variabili aleatorie.

La definizione di variabili aleatorie condizionatamente indipendenti è la seguente: siano X, Y, Z variabili casuali o sottoinsiemi di insiemi di variabili casuali $\boldsymbol{X}$, $\boldsymbol{Y}$, $\boldsymbol{Z}$. X e Y sono condizionatamente indipendenti dato Z se:

$$\mathbb{P}(X = x|Y = y, Z = z) = \mathbb{P}(X = x|Z = z) \tag{9.2}$$

per ogni valore x, y e z che le rispettive variabili aleatorie possono assumere. L'indipendenza condizionata viene indicata come $X \perp Y \mid Z$.

Si può affermare che una rete bayesiana consiste in un insieme di affermazioni di dipendenze e indipendenze condizionate che sono implicate dalla struttura della rete. In particolare, il comportamento complessivo del grafo può essere ricostruito sulla base di tre costrutti fondamentali che descrivono tutti i possibili rapporti tra due nodi non adiacenti X e Z:

Connessione seriale:

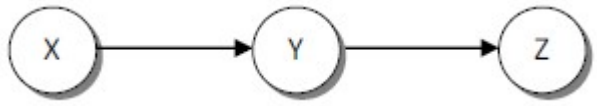

Figura 10-Connessione seriale.

La conoscenza sullo stato della variabile X influenza la conoscenza su Y che a sua volta influenza Z. Un'informazione sullo stato di Z può influenzare la conoscenza su X attraverso Y; se lo stato di Y è noto, allora il passaggio dell'informazione è bloccato e X e Z diventano indipendenti condizionatamente ad Y.

Connessione convergente:

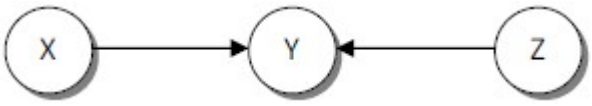

Figura 11-Connessione convergente.

Se non si hanno informazioni sullo stato di Y, eccetto quello che può essere dedotto dalla conoscenza sugli stati dei suoi genitori, X e Z risulteranno essere indipendenti e nessuna conoscenza sullo stato di uno di essi influenzerà la conoscenza sullo stato dell'altro. Se invece si ha informazione sullo stato di Y o di qualcuno tra i suoi figli, X e Z diventano dipendenti.

Connessione divergente:

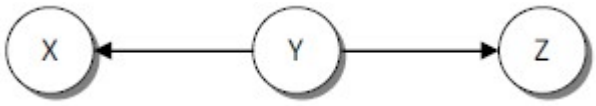

Figura 12-Connessione divergente.

La conoscenza sullo stato della variabile Y influenza la conoscenza sia su X che su Z, dato che sono suoi figli. Se si conosce lo stato assunto da Y, non si ha passaggio di informazione fra i suoi figli X e Z, che pertanto risulteranno essere indipendenti.

9.1-Caratteristiche e proprietà di una rete bayesiana

Una proprietà fondamentale di una rete bayesiana è la *d-separazione* (direction dependent separation), che è definita nel modo seguente: siano i nodi X, Y e Z tre insiemi disgiunti di variabili aleatorie in un grafo orientato aciclico $\mathcal{G}$; X e Y si dicono d-separati da Z se non esiste un cammino tra X e Y tale che:

1) ogni nodo con archi convergenti appartiene a Z o ha un discendente che appartiene a Z,
2) qualsiasi altro nodo non appartiene a Z.

La d-separazione viene indicata con $X \perp_{\mathcal{G}} Y|Z$. Se X e Y non sono d-separati allora si dicono d-connessi. Questa proprietà è considerata la regola principale per l'inferenza: si dimostra infatti che la d-separazione come regola per l'inferenza è *atomic-complete* per le distribuzioni multinomiali e multivariate. Si dice che un insieme di distribuzioni $\boldsymbol{P}$ è atomic-complete [Geiger e Pearl, 1988] per un insieme di grafi orientati aciclici $\boldsymbol{\mathcal{G}}$ se e solo se per ognuno dei grafi $g \in \boldsymbol{\mathcal{G}}$ e per ogni insieme disgiunto di nodi A, B e C esiste una distribuzione $p \in \boldsymbol{P}$ tale che $A \perp_g B|C$ se e solo se $A \perp_g B|C$ è vero in $p \in \boldsymbol{P}$.

L'applicazione della d-separazione è evidente nel caso dei tre costrutti fondamentali definiti precedentemente: nel caso della connessione seriale e divergente i nodi X e Z sono d-separati da Y per la condizione 2), mentre nella connessione convergente X e Z sono d-separati solo se la variabile Y non è istanziata, ovvero se il suo stato non è noto e non è oggetto di condizionamento. In altre parole, la presenza (o l'assenza, nel caso delle connessioni convergenti) di evidenza sullo stato di Y o di un condizionamento esplicito rispetto ai suoi possibili valori blocca il flusso di informazione tra X e Z, rendendoli indipendenti.

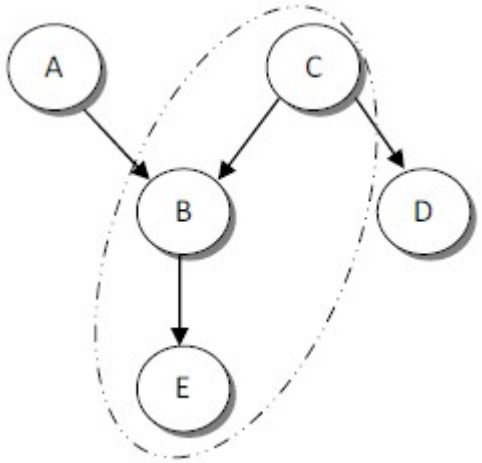

Figura 13-Esempio di d-separazione.

La Fig. 13 mostra un esempio di d-separazione: A e D sono d-connessi dato E, poiché il cammino non diretto da A a D ha solo una connessione convergente su B, che è il genitore di E. Essi risultano invece d-separati dato l'insieme $\{C, E\}$ perché C è il nodo che collega il cammino non diretto tra A e D.

La d-separazione permette inoltre di stabilire una relazione tra grafi orientati aciclici e modelli probabilistici, da cui discende la definizione formale delle reti bayesiane.

Si dice che un grafo orientato aciclico $\mathcal{G}$ è una *mappa di dipendenza* (dependency map o d-map) di un modello probabilistico P se esiste una corrispondenza biunivoca tra le X variabili aleatorie del modello e i nodi V del grafo e se per ogni possibile terna (X, Y, Z) si ha:

$$X \perp Y|Z \Rightarrow X \perp_{\mathcal{G}} Y|Z \tag{9.3}$$

Ovvero l'indipendenza condizionata implica la separazione grafica.

Viceversa $\mathcal{G}$ è una *mappa di indipendenza* (independency map o i-map) di P se:

$$X \perp_{\mathcal{G}} Y|Z \Rightarrow X \perp Y|Z \tag{9.4}$$

Ovvero la d-separazione implica la indipendenza condizionata.

Si dice inoltre che $\mathcal{G}$ è una *mappa di indipendenza minimale* (minimal independency map o minimal i-map) di P se:

- $\mathcal{G}$ è una i-map di P,
- se $\mathcal{G}' \subset \mathcal{G}$, allora $\mathcal{G}'$ non è una i-map di P.

Infine $\mathcal{G}$ è una *mappa perfetta* (perfect map) di P se:

$$X \perp Y|Z \Leftrightarrow X \perp_{\mathcal{G}} Y|Z \tag{9.5}$$

Ovvero se è sia una i-map che una d-map; in questo caso P si dice isomorfo rispetto a $\mathcal{G}$ o causale.

Da tali concetti deriva la definizione di rete bayesiana: sia P una distribuzione di probabilità su insieme di variabili X: allora un grafo orientato aciclico $\mathcal{G} = (X, A)$ è una rete bayesiana di P se e solo se $\mathcal{G}$ è una mappa di indipendenza minimale di P.

La d-separazione è anche alla base della definizione del *Markov blanket*: dato un *DAG* $\mathcal{G}$ e una variabile X_i appartenente ad un insieme di variabili $\boldsymbol{X} = \{X_1, X_2, \dots, X_n\}$ si dice Markov blanket di X_i ogni sottoinsieme $S \subset \boldsymbol{X}$ di variabili per cui:

$$X_i \perp_{\mathcal{G}} (\boldsymbol{X} \setminus \{S \cup X_i\}|S, \qquad X_i \notin S \tag{9.6}$$

Se S è minimale, ovvero nessun suo sottoinsieme è un Markov blanket, è detto Markov boundary di X_i. In ogni rete bayesiana l'unione dei seguenti nodi costituisce un Markov Blanket di X_i: i suoi genitori, i suoi figli ed i suoi coniugi.

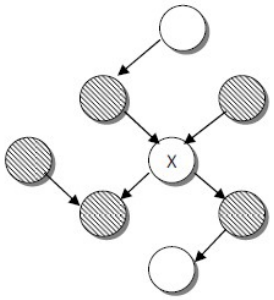

Figura 14-Esempio di Markov Blanket per un nodo X.

Per quanto si è detto, si può dunque definire una rete bayesiana come la rappresentazione della probabilità congiunta sull'insieme delle variabili aleatorie $\boldsymbol{X}$. Utilizzando la condizione di Markov si può affermare che la distribuzione di probabilità congiunta è formalizzata come prodotto di un insieme di probabilità locali, permettendo in tal modo di ridurre la complessità del problema e di facilitare la specificazione del modello probabilistico.

La definizione di condizione di Markov è la seguente: la distribuzione di probabilità congiunta $\mathbb{P}(\boldsymbol{X})$ soddisfa la condizione di Markov per un *DAG* $\mathcal{G}$ se ogni variabile X_i è condizionatamente indipendente da ogni altra variabile escludendo i figli e i genitori, dati i genitori:

$$X_i \perp \boldsymbol{X} \setminus \{Pa(X_i) \cup Ch(X_i)\} | Pa(X_i) \qquad \forall X_i \in \boldsymbol{X} \tag{9.7}$$

La probabilità congiunta di un insieme di variabili $\boldsymbol{X} = \{X_1, X_2, \dots, X_n\}$ è esprimibile, in base alla regola della catena, come:

$$\mathbb{P}(X_1, X_2, \dots, X_n) = \prod_{i=1}^{n} \mathbb{P}(X_i | \boldsymbol{X} \setminus X_i) \tag{9.8}$$

Per la condizione di Markov è possibile semplificare la formula precedente, ottenendo l'espressione seguente, caratteristica di una rete bayesiana:

$$\mathbb{P}(X_1, X_2, \dots, X_n) = \prod_{i=1}^{n} \mathbb{P}(X_i | Pa(X_i)) \tag{9.9}$$

9.2-Inferenza bayesiana

La costruzione di un modello ed il consecutivo utilizzo è l'obiettivo principale di un'analisi statistica. Il grafo ha lo scopo di caratterizzare le dipendenze e le indipendenze condizionate delle variabili aleatorie in esame. Il processo di costruzione di una rete bayesiana viene definito *apprendimento* o *learning*; in questa fase si ricerca la struttura e le probabilità associate alla rete.

L'inferenza probabilistica costituisce la successiva fase al processo di apprendimento della rete bayesiana. Uno dei principali obiettivi nell'utilizzo delle reti bayesiane riguarda la possibilità di determinare la probabilità $\mathbb{P}(X_i|e)$, ossia la probabilità a posteriori del nodo X_i data un'informazione e.

L'informazione che si ha a disposizione viene chiamata *evidenza*: la propagazione di essa consiste nell'aggiornare le distribuzioni di probabilità delle variabili aleatorie in accordo con la nuova informazione disponibile. L'inferenza probabilistica permette di utilizzare l'informazione a disposizione e calcolare le relative probabilità.

La propagazione dell'evidenza può avvenire dall'alto verso il basso, ossia può passare dai genitori ai discendenti; in questo caso l'obiettivo è calcolare le probabilità dei nodi figli dopo il passaggio dell'informazione; oppure essa può avvenire attraverso il processo inverso, in tal caso dunque l'evidenza si ha nei discendenti e passa ai genitori; chiaramente l'attenzione probabilistica è rivolta ai nodi genitori.

Esistono due tipi di evidenza: se lo stato assunto da una o più variabili aleatorie è noto si parla di *evidenza hard*; se invece non si conoscono con certezza i valori assunti da una o più variabili aleatorie, ma si possono fare delle affermazioni sul loro stato, si parla di *evidenza soft*.

La propagazione dell'evidenza avviene applicando il teorema di Bayes, la cui definizione è la seguente: sia $(H_i)_{i\geq 1}$ una partizione dello spazio campionario Ω tale che:

- $\bigcup_{i=1}^{\infty} H_i = \Omega$,
- $H_i \cap H_j = \emptyset, \quad i \neq j$,
- $\mathbb{P}(H_i) > 0, \quad i = 1, \dots, \infty$.

Sia inoltre $E \subseteq \Omega$ un evento tale che $\mathbb{P}(E) > 0$, allora per $i = 1, \dots, \infty$ si ha che:

$$\mathbb{P}(H_i|E) = \frac{\mathbb{P}(E|H_i)\mathbb{P}(H_i)}{\sum_{j=1}^{\infty} \mathbb{P}(E|H_j)\mathbb{P}(H_j)} \tag{9.10}$$

Questa tecnica risulta efficace se il numero di variabili aleatorie non è elevato e quando i valori che ogni variabile può assumere non sono molti; in caso contrario il calcolo richiesto è troppo oneroso.

Si prenda in considerazione un sottoinsieme della variabili aleatorie $O \subset X$ e un insieme di nodi Q; le probabilità marginali sono così calcolate:

$$\mathbb{P}(Q, O) = \sum_{X\setminus\{O\cup Q\}} \mathbb{P}(X) \tag{9.11}$$

$$\mathbb{P}(O) = \sum_{Q} \mathbb{P}(Q, O) \tag{9.12}$$

Inserendo l'evidenza si possono calcolare le seguenti probabilità condizionate:

$$\mathbb{P}(Q|O = o) = \frac{\mathbb{P}(Q, O=o)}{\mathbb{P}(O=o)} \tag{9.13}$$

Tenendo in considerazione la fattorizzazione delle probabilità congiunte della rete bayesiana, il problema viene riformulato nel seguente modo:

$$\mathbb{P}(Q|O=o) = \sum_{X\setminus\{O\cup Q\}} \mathbb{P}(X|O=o) = \sum_{X\setminus\{O\cup Q\}} \prod_{i=1}^{N} \mathbb{P}(X_i|Pa(X_i), O=o) \quad (9.14)$$

È chiaro che questo modo di procedere porta ad un numero elevato di fattorizzazioni che cresce all'aumentare del numero di nodi. Si consideri infatti il seguente esempio:

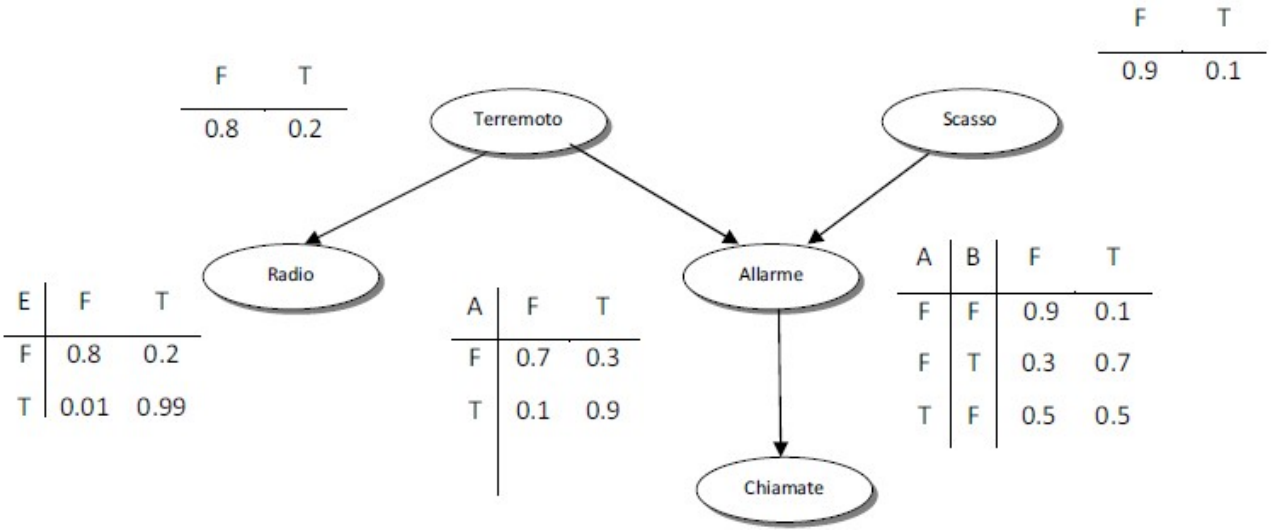

Figura 15-Inferenza.

Il grafo di Fig. 15 descrive la seguente situazione:

- Una chiamata alla stazione di polizia arriva in seguito all'attivazione un allarme;
- L'allarma può attivarsi per uno scasso o a causa di un terremoto;
- Il terremoto attiva anch'esso un allarme e la notizia della sua presenza viene trasmessa in modo da effettuare la chiamata.

Si vuole calcolare la probabilità che avvenga una chiamata se è avvenuto uno scasso; in altre parole, si vuole calcolare la probabilità che se una chiamata viene effettuata, questa sia dovuta ad uno scasso e non a causa di un terremoto. In seguito la chiamata sarà identificata dalla lettera C, lo scasso dalla B, A rappresenterà l'allarme ed infine E il terremoto.

Quindi per quanto si è appena detto si è interessati al calcolo di $\mathbb{P}(C = true|B = true)$:

$$\mathbb{P}(C = true|B = true) = \mathbb{P}(C = true|A = false)\mathbb{P}(A = false|E = false, B = true)\mathbb{P}(E = false) +$$

$$+ \mathbb{P}(C = true|A = false)\mathbb{P}(A = false|E = true, B = true)\mathbb{P}(E = true) +$$

$$+ \mathbb{P}(C = true|A = true)\mathbb{P}(A = true|E = false, B = true)\mathbb{P}(E = false) +$$

$$+ \mathbb{P}(C = true|A = true)\mathbb{P}(A = true|E = true, B = true)\mathbb{P}(E = false) = 0.7548$$

Come si può notare dall'esempio, la complessità della formula di Bayes aumenta con il crescere del dominio; ad esempio per una rete con venti variabili binomiali, dove una è osservata e in una seconda si ha l'evidenza, la sommatoria è composta da $2^{20-2} = 2^{18}$ termini. Per risolvere sono stati sviluppati molti approcci: un'idea è quello di ottimizzare le distribuzioni marginali fattorizzando le distribuzioni; in altri termini, si tratta di inserire la sommatoria più tardi possibile. Questo nell'esempio riportato si traduce nel modo seguente:

$$\mathbb{P}(C|B) = \sum_{A,E} \mathbb{P}(E)\mathbb{P}(A|E,B)\mathbb{P}(C|A) = \sum_{E} \mathbb{P}(E) \sum_{A} \mathbb{P}(A|E,B)\mathbb{P}(C|A) \qquad (9.15)$$

La complessità della marginalizzazione dipende dai termini presenti della sommatoria.

Per risolvere tali problemi sono stati sviluppati numerosi algoritmi per l'inferenza approssimata che consistono essenzialmente in una semplificazione del modello di rete in esame.

9.3-Apprendimento di reti bayesiane

L' operazione di apprendimento o learning è composta da due fasi:

- l'apprendimento della struttura del grafo, ed in particolare dei suoi archi;
- l'apprendimento dei parametri che regolano il comportamento delle distribuzioni di probabilità.

Queste due fasi sono interconnesse tra di loro, il learning della struttura spiega i dati una volta forniti i parametri adatti; l'apprendimento dei parametri è possibile se è data una struttura. Il procedimento naturale porta quindi a considerare prima l'apprendimento della struttura ed in seguito quello dei parametri.

Il learning può essere portato a termine in due modi:

- utilizzando delle conoscenze pregresse sui fenomeni in esame;
- tramite l'applicazione di opportuni algoritmi a dati osservati.

Il primo approccio è spesso dispendioso e richiede molto tempo specialmente per reti complesse, il secondo utilizza software che permettono di apprendere la struttura in modo automatico inserendo condizioni prestabilite. Quest'ultimo approccio è quello usato più frequentemente, in quanto più affidabile ed efficace, ed è quello che verrà adottato anche in questa tesi.

L'individuazione della struttura di una rete bayesiana può essere effettuata applicando a dei dati osservati due classi di metodi:

- la selezione di un modello sulla base di un punteggio che ne descriva in modo sintetico la bontà complessiva di adattamento (metodi *score-based*),
- l'individuazione delle relazioni tra le variabili utilizzando test locali (tra sottoinsiemi di $\boldsymbol{X} = \{X_1, X_2, \dots, X_n\}$ e la loro composizione attraverso gli assiomi dell'indipendenza condizionata (metodi *constraint-based*).

In entrambi i casi la definizione stessa di rete bayesiana impone alcune ipotesi per garantire la validità del modello, che derivano dalla sua condizione di i-map minimale e dal rispetto degli assiomi di indipendenza condizionata:

- *markovianità* (markov assumption): deve valere la condizione di Markov, ovvero ogni nodo deve essere indipendente da tutti i suoi non discendenti condizionatamente ai suoi genitori;
- *causalità* (causal sufficiency): non devono esistere variabili latenti; in caso contrario verrebbe meno la corrispondenza biunivoca tra i nodi del grafo e le variabili stesse. Per lo stesso motivo le variabili considerate devono essere tra loro distinte;
- *accuratezza* (faithfulness): il grafo deve essere una mappa perfetta della distribuzione di probabilità considerata; quest'ipotesi è necessaria nei metodi constraint-based, e viene spesso assunta anche per i metodi score-based.

Ai fini di questa tesi si fanno delle ipotesi ulteriori:

- le variabili aleatorie considerate sono discrete;
- non vi devono essere dati mancanti; se così fosse si rischierebbe di introdurre ulteriori dipendenze tra le variabili casuali, impedendo la fattorizzazione in probabilità locali.

9.3.1-Apprendimento della struttura

In questa tesi per il learning della struttura della rete bayesiana si farà ricorso ad uno specifico algoritmo, detto K2. L'algoritmo K2 [Cooper e Herskovits, 1992] fa parte dei metodi score-based e può essere utilizzato per una qualsiasi rete bayesiana. Come input necessita di un database dei casi che si possono verificare e di un insieme di nodi ordinati; il dover fornire un ordine dei nodi può essere un limite, che in alcuni casi può essere risolto attraverso analisi preliminari. Il risultato è chiaramente una rete che descrive la relazione tra le variabili aleatorie in esame.

Il procedimento per la ricerca delle dipendenze tra le variabili aleatorie avviene nel seguente modo: si introduce una struttura ipotetica di dipendenza e tramite l'ausilio di un calcolatore viene individuata la probabilità della distribuzione a posteriori.

Nel seguente paragrafo si utilizzerà la seguente notazione: D è il database dei casi, Z l'insieme delle variabili aleatorie, $B = (B_S, B_P)$ indica una rete bayesiana, dove B_S è la struttura e B_P è l'insieme dei parametri.

Siano B_{S_i} e B_{S_j} due differenti strutture di reti bayesiane basate però sullo stesso insieme di variabili aleatorie Z. In questa sede viene mostrato un metodo per calcolare $\mathbb{P}(B_{S_i}|D) \Big/ \mathbb{P}(B_{S_j}|D)$. Calcolando questo rapporto per coppie di strutture di reti bayesiane, risulta possibile ordinare un insieme di strutture in base alle loro

probabilità a posteriori. Per calcolare il rapporto delle probabilità a posteriori, si calcola $\mathbb{P}(B_{S_i}, D)$ e $\mathbb{P}(B_{S_j}, D)$ e si utilizza la seguente equivalenza:

$$\frac{\mathbb{P}(B_{S_i}|D)}{\mathbb{P}(B_{S_j}|D)} = \frac{\frac{\mathbb{P}(B_{S_i},D)}{D}}{\frac{\mathbb{P}(B_{S_j},D)}{D}} = \frac{\mathbb{P}(B_{S_i},D)}{\mathbb{P}(B_{S_j},D)} \tag{9.16}$$

Considerando un arbitraria struttura di rete bayesiana B_S basata su insieme di variabili Z, per procedere con il calcolo di $\mathbb{P}(B_S, D)$ si introducono dapprima quattro ipotesi:

1. Le variabili aleatorie in esame sono discrete. Questa assunzione non preclude alcuno scopo di questa tesi, poiché tutte le variabili aleatorie che qui verranno usate sono discrete. In studi in cui sono presenti variabili aleatorie continue è comunque possibile utilizzare l'algoritmo dopo un'adeguata discretizzazione delle variabili aleatorie.
2. Dato un modello di rete bayesiana, i casi sono indipendenti tra loro.
3. Non ci sono casi in cui sono presenti variabili con valori mancanti.
4. La funzione di densità $f(B_P|B_S)$ è uniforme. Ciò implica che prima di osservare il database di casi, è indifferente attribuire un qualunque valore probabilistico alla struttura della rete bayesiana.

Date queste assunzioni, vale il seguente teorema: si consideri un insieme Z di n variabili discrete, dove ogni variabile $X_i \in Z$ ha r_i possibili valori: $(v_{i1}, \dots, v_{ir_i})$. Sia D un database di m casi, dove ogni caso contiene il valore da attribuire ad ogni variabile in Z. Sia B_S la struttura di una rete bayesiana contenente proprio le variabili in Z. Ciascuna variabile X_i in B_S ha un insieme di genitori, indicato con $Pa(X_i)$. Si indichi con w_{ij} la j-esima unica istanza di $Pa(X_i)$ in D e con q_i il totale delle istanza di $Pa(X_i)$. Si definisca N_{ijk} come il numero di casi in D in cui una variabile X_i ha il valore v_{ik} mentre $Pa(X_i)$ ha valore w_{ij}. Sia:

$$N_{ij} = \sum_{k=1}^{r_i} N_{ijk} \tag{9.17}$$

Allora, date le 4 ipotesi espresse in precedenza, si ha che:

$$\mathbb{P}(B_S, D) = \mathbb{P}(B_S) \prod_{i=1}^{n} \prod_{j=1}^{q_i} \frac{(r_i-1)!}{(N_{ij}+r_i-1)!} \prod_{k=1}^{r_i} N_{ijk}! \tag{9.18}$$

È chiaro che l'efficienza del modello si valuta anche in base al tempo necessario per effettuare il calcolo dell'equazione precedente. Si dimostra che l'ordine del tempo di esecuzione dell'algoritmo è $O(mnr + t_{B_S})$, dove r è il numero massimo dei valori possibili per ciascuna variabile, dato da $r = \max_{1 \le i \le n}[r_i]$, mentre t_{B_S} indica il tempo necessario per calcolare la probabilità a priori di una struttura, supponendo noti i valori di N_{ijk}. Inoltre se risulta noto il numero massimo di genitori di ciascun nodo, indicato con u, la complessità del modello si riduce a $O(mnru + t_{B_S})$. Inoltre se r e u sono costanti e $O(t_{B_S}) = O(unr)$, tutta la complessità per il calcolo della precedente equazione si riduce ulteriormente a $O(mn)$.

L'obiettivo è ricercare la struttura, tra tutte quelle possibili che massimizza la probabilità $\mathbb{P}(B_S, D)$. Ciò si traduce nel seguente problema di massimizzazione:

$$\max_{B_S} \mathbb{P}(B_S, D) = c \prod_{i=1}^{n} \max_{Pa(X_i)} \prod_{j=1}^{q_i} \frac{(r_i-1)!}{(N_{ij}+r_i-1)!} \prod_{k=1}^{r_i} N_{ijk}! \qquad (9.19)$$

dove c indica la probabilità a priori costante $\mathbb{P}(B_S)$, per ogni struttura B_S. La complessità del calcolo può essere ridotta facendo alcune assunzioni:

- L'esistenza di un ordine per i nodi della rete,
- l'esistenza di un numero sufficientemente limitato di genitori per ogni nodo,
- $\mathbb{P}(Pa(X_i) \to X_i)$ e $\mathbb{P}(Pa(X_j) \to X_j)$ sono marginalmente indipendenti per $i \neq j$.

La seconda assunzione non può sempre essere verificata e per questo motivo è stato costruito il K2, un algoritmo euristico con una complessità polinomiale, che non richiede restrizioni sul numero di genitori di un nodo. Tale algoritmo consiste in un metodo *greedy-search*: si assume inizialmente che un nodo non abbia genitori, successivamente si aggiunge in modo incrementale quel genitore la cui aggiunta aumenta in maggiormente la probabilità della struttura risultante. Si assume, come già accennato, che esista un ordine nel dominio delle variabili e che, a priori, tutte le strutture siano ugualmente probabili.

Per il calcolo della probabilità si utilizza la seguente funzione, che rappresenta la *scoring function* dell'algoritmo:

$$g(i, Pa(X_i)) = \prod_{j=1}^{q_i} \frac{(r_i-1)!}{(N_{ij}+r_i-1)!} \prod_{k=1}^{r_i} N_{ijk}! \qquad (9.20)$$

dove i valori N_{ijk} sono calcolati rispetto ai genitori $Pa(X_i)$ di ciascuna variabile X_i e rispetto al database D che si lascia implicito. Si dimostra che $g(i, Pa(X_i))$ può essere calcolata in un tempo dell'ordine $O(mur)$, dove u è il numero massimo di genitori che a ciascun nodo è concesso di avere. Si utilizza inoltre una funzione $Pred(X_i)$ che definisce l'insieme dei nodi predecessori di X_i, nella maniera stabilita dall'ordinamento dei nodi.

In sintesi, l'algoritmo K2 è così riassumibile:

1. Si sceglie un ordinamento topologico sui nodi.
2. Per ogni nodo X_i si costruisce l'insieme dei predecessori $Pred(X_i)$, in base all'ordinamento dei nodi.
3. Si valuta ogni nodo X_j presente in $Pred(X_i)$: se X_j massimizza la scoring function $g(i, Pa(X_i))$, allora si considera X_j come genitore di X_i aggiungendolo all'insieme $Pa(X_i)$. La fase di ricerca termina dopo aver esaminato tutte le variabili in $Pred(X_i)$ o se viene raggiunta la soglia u sul numero massimo di genitori.

Ottenuta la struttura della rete, è possibile ricavare i valori numerici delle probabilità. In particolare, si considera il calcolo dei valori attesi delle probabilità.

Sia θ_{ijk} la probabilità condizionata $\mathbb{P}(X_i = v_{ik}|Pa(X_i) = w_{ij})$, ossia la probabilità che X_i abbia valore v_{ik}, con $k = 1, \dots, r_i$, dati che genitori di X_i abbiano valore w_{ij}. θ_{ijk} viene chiamata probabilità condizionata della rete. Si indicano con ξ le quattro ipotesi definite in precedenza per il calcolo di $\mathbb{P}(B_S, D)$. Il valore di $\mathbb{E}(\theta_{ijk}|D, B_S, \xi)$, che è il valore atteso di θ_{ijk} dato il database D, la struttura della rete B_S e le ipotesi ξ, è dato dalla seguente espressione:

$$\mathbb{E}\left(\theta_{ijk}\middle|D, B_S, \xi\right) = \frac{N_{ijk}+1}{N_{ij}+r_i} \tag{9.21}$$

Tale quantità viene definita stimatore Bayesiano di θ_{ijk}. Applicando un'analisi analoga alla varianza, si ricava che:

$$Var\left(\theta_{ijk}\middle|D, B_S, \xi\right) = \frac{(N_{ijk}+1)(N_{ij}+r_i-N_{ijk}-1)}{(N_{ij}+r_i)^2(N_{ij}+r_i+1)} \tag{9.22}$$

9.3.2-Apprendimento dei parametri

Per l'apprendimento dei parametri della rete, una volta nota la struttura, si può fare uso di diversi metodi. In questa tesi si è scelto di utilizzare il metodo della massima verosimiglianza [Fisher, 1922], di cui di seguito si da una breve spiegazione.

Si consideri un modello statistico parametrico con funzione di probabilità $p(x, \theta)$, dove $x = (x_1, \dots, x_n)$ è una realizzazione casuale di una variabile aleatoria discreta X con funzione di probabilità $p_0(x)$, mentre θ è il parametro del modello. La funzione di verosimiglianza per x è $L(\theta) = p(x, \theta)$, per $\theta \in \Theta$, dove Θ è lo spazio parametrico, ossia tutti i possibili valori che può assumere il parametro θ. La funzione $L(\theta)$ ha quindi la stessa forma della funzione di probabilità del modello, ma in tale caso x è fissato e θ è libero di variare. Un'ipotesi frequente è che i dati $x = (x_1, \dots, x_n)$ siano realizzazioni *i.i.d*, ossia osservazioni indipendenti e identicamente distribuite. Perciò se $p(x_i, \theta)$ è la distribuzione di probabilità marginale della singola osservazione, la funzione di verosimiglianza diventa:

$$L(\theta) = \prod_{i=1}^{n} p(x_i, \theta) \tag{9.23}$$

L'obiettivo della funzione di verosimiglianza è di ottenere il maggior numero di informazione sul vero valore del parametro θ^0. La logica dietro la funzione di verosimiglianza è la seguente: in seguito ai dati osservati, $\theta_1 \in \Theta$ è più plausibile di $\theta_2 \in \Theta$ nel modello probabilistico generatore dei dati se $L(\theta_1) > L(\theta_2)$, ossia θ_1 ha più probabilità di essere il vero valore θ^0. Se la distribuzione con parametro θ_1 è più vicina alla distribuzione empirica dei dati rispetto alla distribuzione con θ_2, allora si avrà che la verosimiglianza valutata in θ_1 è maggiore di quella valutata in θ_2.

Un metodo di confronto tra la differenza nell'evidenza empirica dei dati x a favore di θ_1 rispetto a θ_2 è il rapporto $L(\theta_1)/L(\theta_2)$, detto rapporto di verosimiglianza. I fattori che non dipendono da θ in $L(\theta)$ possono essere eliminati, dato che non cambiano il valore del rapporto di verosimiglianza. Per questo motivo, le funzioni $L(\theta)$ e $cL(\theta)$, dove $c \in \mathbb{R}^+$ è una costante, sono equivalenti.

In genere, ai fini pratici, viene utilizzata la trasformazione logaritmica di $L(\theta)$; si definisce dunque la funzione di log-verosimiglianza $l(\theta)$:

$$l(\theta) = \log L(\theta) = \sum_{i=1}^{n} p(x_i, \theta) \tag{9.24}$$

dove se $L(\theta) = 0$, $l(\theta) = -\infty$ per convenzione.

Se si hanno due differenti insiemi di dati x e y, indipendenti tra loro, che contengono entrambi dell'informazione su θ, dato che la loro funzione di probabilità congiunta è il prodotto delle due marginali, allora la verosimiglianza per θ basata su x e y sarà:

$$L(\theta, x, y) = p(y, \theta)p(x, \theta) = L(\theta, y)L(\theta, x) \tag{9.25}$$

La stima di massima verosimiglianza (maximum likelihood estimation, *MLE*) è quel valore $\hat{\theta} \in \Theta$ che massimizza $l(\theta)$ tale che $L(\hat{\theta}) \geq L(\theta)\ \forall \theta \in \Theta$. Se $\hat{\theta} = \hat{\theta}(x)$ esiste ed è unico, $\hat{\theta} = \hat{\theta}(X)$ è definito stimatore di massima verosimiglianza. Dato che il logaritmo è una funzione strettamente monotona, massimizzare $l(\theta)$ equivale a massimizzare $L(\theta)$.

Si definisce funzione punteggio:

$$l_*(\theta) = \frac{\partial l(\theta)}{\partial \theta} \tag{9.26}$$

Il punto di massima verosimiglianza va dunque cercato tra le soluzioni dell'equazione $l_*(\theta) = 0$, che prende il nome di equazione di verosimiglianza.

Per una rete bayesiana, il learning dei parametri consiste nel determinare i valori delle probabilità condizionate $\mathbb{P}(X_i|Pa(X_i)$ per tutte le variabili dell'insieme $X = \{X_1, \dots, X_n\}$. In questo paragrafo si userà $\theta_i = \mathbb{P}(X_i|Pa(X_i)$ per indicare i parametri.

Data una struttura di una rete bayesiana B_S e un database di m casi $D = \{D_1, \dots, D_m\}$ in cui ogni caso contiene il valore da attribuire ad ogni variabile in X, la funzione di verosimiglianza $L(\theta, D)$, dove in tale caso $\theta = \{\theta_1, \dots, \theta_n\}$, è pari a:

$$L(\theta, D) = \mathbb{P}(D|\theta) = \prod_{i=1}^{m} \mathbb{P}(D_i|\theta) \tag{9.27}$$

La funzione di log-verosimiglianza è:

$$l(\theta, D) = \log L(\theta, D) = \log \mathbb{P}(D|\theta) = \sum_{i=1}^{m} \mathbb{P}(D_i|\theta) \quad (9.28)$$

La stima di massima verosimiglianza $\hat{\theta}_i$ per il parametro $\theta_i = \mathbb{P}(X_i|Pa(X_i)$ risulta avere la seguente espressione:

$$\hat{\theta}_i = \frac{m(X_i, Pa(X_i))}{\sum_{X_i'=1}^{n} m(X_i', Pa(X_i))} \quad (9.29)$$

dove $m\big(X_i, Pa(X_i)\big)$ è il numero delle osservazioni della variabile X_i rispetto ai suoi genitori $Pa(X_i)$.

10-Tecniche di soluzione delle reti

La valutazione dell'affidabilità di un sistema elettrico generalmente comporta la soluzione del load flow della rete che rappresenta il sistema in esame in varie situazioni di interruzioni casuali (contingenze). A seconda dei criteri di affidabilità utilizzati e l'intento dietro questi studi, sono utilizzate varie tecniche per analizzare l'affidabilità di un sistema elettrico. Le tecniche di base utilizzate per la soluzione di una rete sono le seguenti:

- Metodo del flusso di rete,
- Load Flow in AC,
- Load Flow in DC.

La scelta di una tecnica appropriata è di primaria importanza. Il punto chiave è che la tecnica scelta deve essere in grado di soddisfare l'intento dietro gli studi da una prospettiva di gestione, pianificazione e progettazione. Una delle tecniche più semplici è quello di trattare il sistema come un modello di trasporto (modello di flusso di rete o network flow) [Ford e Fulkerson, 1962]. Questo metodo si basa sul movimento di una particolare merce da un certo numero di fonti a un certo numero di centri di domanda. Il modello di flusso di rete mantiene l'equilibrio di potenza ad ogni nodo della rete e non soddisfa la legge di Kirchhoff che può non essere adatta per il funzionamento pratico del sistema elettrico.

Tecniche approssimate di load flow quali il load flow in DC sono piuttosto semplici e veloci ma forniscono soltanto stime dei flussi di potenza di linea, senza includere qualsiasi stima delle tensioni dei nodi e i limiti di potenza reattiva delle unità di generazione. Qualora sono di interesse sia la continuità che la qualità del servizio elettrico, allora è necessario esaminare i livelli di tensione in ogni punto di carico rilevante e i limiti di potenza reattiva (Mvar) di ciascuna unità di generazione considerando inoltre l'effetto dei guasti componenti quali generatori, linee di trasmissione e trasformatori [Billinton e Allan, 1988]. Considerando un sistema di potenza come modello di trasporto o utilizzando il load flow in DC, non si ha una stima della qualità del servizio elettrico.

Se la qualità dell'alimentazione, con inclusi livelli di tensione accettabili e limiti reattivi appropriati, è un importante requisito di affidabilità, devono allora essere utilizzati metodi di load flow più accurati per

calcolare gli indici di affidabilità, quali ad esempio il metodo Newton-Raphson o Gauss-Seidel. Queste tecniche però richiedono molta memoria nei calcolatori sono computazionalmente onerosi. È possibile adottare una tecnica di load flow in AC veloce come il "load flow disaccoppiato" [Stott e Alsac, 1974], che è una modifica del metodo di load flow usando Newton-Raphson. Quanto segue è una breve descrizione delle due tecniche di soluzione di rete maggiormente utilizzate nello studio dei sistemi elettrici.

10.1-Load flow in AC

Le equazioni di base di load flow in coordinate polari sono le seguenti:

$$P_i = V_i \sum_{j=1}^{n} V_j (G_{ij} \cos \delta_{ij} + B_{ij} \sin \delta_{ij}) \quad (i = 1, \dots, n) \tag{10.1}$$

$$Q_i = V_i \sum_{j=1}^{n} V_j (G_{ij} \sin \delta_{ij} - B_{ij} \cos \delta_{ij}) \quad (i = 1, \dots, n) \tag{10.2}$$

dove P_i e Q_i sono le immissioni di potenza attiva e reattiva al nodo i; V_i e δ_i sono l'ampiezza e la fase della tensione al nodo i; $\delta_{ij} = \delta_i - \delta_j$; G_{ij} e B_{ij} sono la parte reale e immaginaria dell'elemento Y_{ij} della matrice delle ammettenze nodali; n è il numero di nodi del sistema.

Ciascun nodo ha quattro variabili $(V_i, \delta_i, P_i, Q_i)$ e pertanto n nodi hanno $4n$ variabili in totale. Le equazioni precedenti consistono dunque in un sistema di $2n$ equazioni. Per risolvere le equazioni di load flow, devono essere assegnate per ciascun nodo due delle quattro variabili. In generale, le P_i e Q_i dei nodi di carico sono note e tali nodi sono denominati nodi PQ; per i nodi di generazione, vengono specificati P_i e V_i e questi nodi vengono chiamati nodi PV; è necessario infine specificare V_i e δ_i di un nodo nel sistema per definire il bilancio di potenza dell'intero sistema, e tale nodo è detto nodo slack.

Le equazioni di load flow possono essere risolte con il metodo delle linearizzazioni successive, secondo il noto modello di Newton-Raphson. Le equazioni di load flow sono linea rizzate in modo da ottenere la seguente equazione matriciale:

$$\begin{pmatrix} \Delta P \\ \Delta Q \end{pmatrix} = \begin{pmatrix} H & N \\ J & L \end{pmatrix} \begin{pmatrix} \Delta \delta \\ \Delta V/V \end{pmatrix} \tag{10.3}$$

La matrice Jacobiana $\begin{pmatrix} H & N \\ J & L \end{pmatrix}$ è una matrice quadrata $(n + m - 1)$, dove n e m sono rispettivamente il numero di tutti i nodi e dei nodi di carico; $\Delta V/V$ indica che i suoi elementi sono $\Delta V_i/V_i$. Gli elementi della matrice Jacobiana sono calcolati nel modo seguente:

$$H_{ij} = \frac{\partial P_i}{\partial \delta_j} = V_i V_j (G_{ij} \sin \delta_{ij} - B_{ij} \cos \delta_{ij}) \tag{10.4}$$

$$H_{ii} = \frac{\partial P_i}{\partial \delta_i} = -Q_i - B_{ii} V_i^2 \tag{10.5}$$

$$N_{ij} = \frac{\partial P_i}{\partial V_j} V_j = V_i V_j (G_{ij} \cos \delta_{ij} + B_{ij} \sin \delta_{ij}) \tag{10.6}$$

$$N_{ii} = \frac{\partial P_i}{\partial V_i} V_i = P_i + G_{ii} V_i^2 \tag{10.7}$$

$$J_{ij} = \frac{\partial Q_i}{\partial \delta_j} = -N_{ij} \tag{10.8}$$

$$J_{ii} = \frac{\partial Q_i}{\partial \delta_i} = P_i - G_{ii} V_i^2 \tag{10.9}$$

$$L_{ij} = \frac{\partial Q_i}{\partial V_j} V_j = H_{ij} \tag{10.10}$$

$$L_{ii} = \frac{\partial Q_i}{\partial V_i} V_i = Q_i - B_{ii} V_i^2 \tag{10.11}$$

10.2-Load flow disaccoppiato

Le reattanze longitudinali sono normalmente molto più grandi delle resistenze longitudinali e le differenze di fase tra due nodi sono molto piccole nei sistemi elettrici. Questo porta a una prevalenza dei termini sulla diagonale della matrice Jacobiana, ovvero i valori dei blocchi matriciali N e J sono molto più piccoli rispetto a quelli di H e L. L'equazione matriciale descritta in precedenza può essere dunque disaccoppiata assumendo $N = 0$ e $J = 0$. Considerando poi le seguenti semplificazioni:

$$\left|G_{ij} \sin \delta_{ij}\right| \ll \left|B_{ij} \cos \delta_{ij}\right| \tag{10.12}$$

$$|Q_i| \ll \left|B_{ii} V_i^2\right| \tag{10.13}$$

Le equazioni disaccoppiate possono essere ulteriormente semplificate in modo da portare alle seguenti espressioni:

$$[\Delta P/V] = [B'][V \Delta \delta] \tag{10.14}$$

$$[\Delta Q/V] = [B''][\Delta V] \tag{10.15}$$

dove $[\Delta P/V]$ e $[\Delta Q/V]$ sono vettori i cui elementi sono rispettivamente $\Delta P_i/V_i$ e $\Delta Q_i/V_i$; $[V\Delta\delta]$ è un vettore i cui elementi sono $V_i \Delta \delta_i$. Gli elementi delle matrici costanti $[B']$ e $[B'']$ sono ottenuti nella maniera seguente:

$$B'_{ij} = -\frac{1}{x_{ij}} \tag{10.16}$$

$$B'_{ii} = -\sum_{j \in R_i} B'_{ij} \tag{10.17}$$

$$B''_{ij} = -\frac{x_{ij}}{r_{ij}^2 + x_{ij}^2} \tag{10.18}$$

$$B''_{ii} = -2b_{i0} - \sum_{j \in R_i} B''_{ij} \tag{10.19}$$

dove r_{ij} e x_{ij} sono rispettivamente la resistenza e la reattanza longitudinale tra nodo i e nodo j; b_{i0} è la suscettanza trasversale tra il nodo i e la terra; R_i è l'insieme dei nodi direttamente connessi al nodo i.

10.3-Load flow in DC

Le equazioni di load flow in DC [Stott, 1974] correlano la potenza attiva alle fasi delle tensioni dei nodi. I modelli basati sul load flow in DC sono ampiamente usati nella valutazione dell'adeguatezza dei sottosistemi composti da sistema di generazione e di trasmissione in quanto:

a) I più importanti indici di affidabilità sono associati alle riduzioni di potenza attiva dei carichi e il calcolo di tali indici richiede soltanto informazioni riguardante la potenza attiva.
b) I calcoli di load flow in applicazioni pratiche indicano che in molti sistemi ci sono differenze relativamente piccole (3% - 10%) tra i load flows in AC e in DC. Queste differenze sono piccole rispetto ai possibili errori dovuti alle incertezze nei dati di base per l'analisi dell'affidabilità, come i tassi di guasto dei componenti e i tempi di interruzione dell'alimentazione.
c) È necessario valutare un ampio numero di stati del sistema per garantire accuratezza nei calcoli degli indici di probabilità. I modelli basati sul load flow in DC, includendo il power flow ottimale, possono essere rapidamente calcolati.
   Si deve chiaramente comprendere che se le considerazioni riguardanti tensione e potenza reattiva sono requisiti importanti in uno studio particolare di un sistema, allora il load flow in DC non è un approccio accettabile.

Il load flow in DC è basato sulle seguenti quattro assunzioni:

1. Le resistenze longitudinali sono molto più piccole delle reattanze longitudinali. Le suscettanze longitudinali b_{ij} tra coppie di nodi i e j possono essere approssimate come:

$$b_{ij} \approx -\frac{1}{x_{ij}} \tag{10.20}$$

2. La differenza di fase delle tensioni tra due nodi di una linea è piccola e pertanto:

$$\sin \delta_{ij} \approx \delta_{ij} = \delta_i - \delta_j \tag{10.21}$$

$$\cos \delta_{ij} \approx 1 \tag{10.22}$$

3. Le ammettenze trasversali tra i nodi e la terra possono essere trascurate, dunque:

$$b_{i0} = b_{j0} = 0 \tag{10.23}$$

4. Tutte le ampiezze delle tensioni dei nodi sono considerate pari a 1 p.u.

In base alle assunzioni soprastanti, il flusso di potenza attiva in un ramo può essere calcolato come:

$$P_{ij} = \frac{\delta_i - \delta_j}{x_{ij}} \tag{10.24}$$

E pertanto i bilanci di potenza attiva nei nodi sono:

$$P_i = \sum_{j \in R_i} P_{ij} = B'_{ii}\delta_i + \sum_{j \in R_i} B'_{ij}\delta_j \quad (i = 1, \dots, n) \tag{10.25}$$

dove $B'_{ij} = -\frac{1}{x_{ij}}$ e $B'_{ii} = -\sum_{j \in R_i} B'_{ij}$. L'equazione precedente può essere espressa in forma matriciale:

$$P = [B'][\delta] \tag{10.26}$$

Se il nodo n è selezionato come nodo slack e si considera $\delta_n = 0$, $[B']$ è una matrice quadrata $(n-1)$. Corrisponde esattamente alla matrice $[B']$ definita nelle equazioni di load flow disaccoppiato.

10.4-Azioni correttive

In condizione di funzionamento normale, tutti i limiti operativi sono soddisfatti. I vincoli operativi sono descritti come segue [Billinton e Li, 1994]:

1. Vincoli sull'ampiezza della tensione: i limiti operativi sono imposti sulle ampiezze delle tensioni nei nodi, ovvero:

   $$V^{min} \leq V \leq V^{max} \tag{10.27}$$

   dove V^{min} e V^{max} rappresentano rispettivamente il limite minimo e massimo di tensione.
2. Vincoli di flusso di potenza nei rami: questi rappresentano i limiti termici sulle linee e sui trasformatori. In alcuni casi, i limiti di stabilità statica sulle linee espressi dalle differenze di fase possono essere trasformati in vincoli di flusso di potenza:

   $$|T| \leq T^{max} \tag{10.28}$$

   dove T è il flusso di potenza in un ramo e T^{max} è il limite massimo di capacità di una linea o di un trasformatore.
3. Vincoli di generazione di potenza attiva: i limiti di generazione di potenza attiva al nodo slack e ai nodi PV sono:

   $$P^{min} \leq P \leq P^{max} \tag{10.29}$$

   dove P^{min} e P^{max} rappresentano rispettivamente la minima e massima potenza attiva generata in ciascun nodo di generazione.
4. Vincoli di generazione di potenza reattiva: i limiti di generazione di potenza reattiva al nodo slack e ai nodi PV sono:

   $$Q^{min} \leq Q \leq Q^{max} \tag{10.30}$$

dove Q^{min} e Q^{max} rappresentano rispettivamente la minima e massima potenza reattiva generata in ciascun nodo di generazione.

Tutti i vincoli operativi sopra descritti devono essere soddisfatti per il normale funzionamento di un sistema elettrico. Quando un qualsiasi vincolo operativo è violato, sono necessari interventi correttivi al fine di alleviare il problema di vincolo operativo e per ripristinare il normale funzionamento del sistema. Il verificarsi di un problema di sistema può di per sé essere registrato come un evento di guasto. In molti casi, tuttavia, può essere possibile eliminare un problema del sistema adottando opportune azioni correttive. È pertanto interessante determinare se è possibile eliminare un problema del sistema impiegando una corretta azione correttiva. Non v'è alcun consenso tra le utility elettriche e organizzazioni connesse in materia di criteri uniformi di guasto e quindi tutte le organizzazioni non utilizzano la stessa tecnica fondamentale di soluzione per calcolare l'adeguatezza dei loro sistemi [Billinton e Kumar, 1985]. Le categorie di azioni correttive che possono essere impiegate sono i seguenti [Stott e Alsac, 1974]:

1. Riprogrammazione della generazione nel caso di un'insufficiente capacità del sistema.
2. Attenuazione dei sovraccarichi nelle linee.
3. Correzione delle violazioni dei limiti reattivi delle unità di generazione.
4. Correzione di un problema di tensione ai nodi e soluzione di situazioni di rete mal condizionate quando si utilizzano tecniche di load flow in AC.
5. Funzionamento in isola e frammentazione del sistema in caso di guasti alle linee di trasmissione e ai trasformatori.
6. Riduzione del carico in caso di un problema del sistema.

11-Modello multistato

Il modello del vento è l'aspetto più importante da trattare per incorporare impianti eolici negli studi di affidabilità. L'approccio abituale nella simulazione Monte Carlo non sequenziale è quello di rappresentare il generatore eolico come un modello di Markov multistato; ciascun stato del modello corrisponde ad un particolare livello di potenza generata. Il procedimento per calcolare tale modello multistato è il seguente:

1. Si considera l'andamento delle serie temporali del vento riguardanti l'impianto eolico in esame. Di solito, la serie del vento è unica per l'intero impianto, indipendentemente dal numero di turbine da cui è composto l'impianto, perché è solito considerare, per semplicità, che tutte le turbine siano interessate dallo stesso vento. Generalmente, si prende in considerazione l'andamento delle serie storiche di un anno, in modo da poter definire dati affidabili e coerenti.
2. Per ogni valore di velocità del vento estratto dalle serie storiche, si calcola la potenza generata ricavandola dalla curva del costruttore. Si ottiene in tal modo una serie di potenze che definisce la generazione del sistema eolico in un anno di riferimento.

3. Si va a calcolare sulla serie di potenze ricavata in precedenza la *funzione di sopravvivenza* (survival function, in inglese) $S(x)$, definita come il complemento a uno della funzione di distribuzione cumulativa empirica $F(x)$, dove x in tal caso indica la variabile aleatoria che rappresenta la potenza generata:

$$S(x) = 1 - F(x) \tag{11.1}$$

   Tale funzione risulta essere monotona decrescente; inoltre è una funzione discreta al pari della *CDF* empirica. Essa è anche denominata funzione di distribuzione cumulativa complementare (*CCDF*, Complementary Cumulative Distribution Function).
4. Il numero dei livelli assunti dalla funzione di sopravvivenza andrà ad indicare il numero di suddivisioni della potenza generata; i valori di potenza corrispondenti a ciascun valore della $S(x)$ definiscono i livelli della generazione eolica. Basandosi su questi valori di potenza $P_1, P_2, \dots, P_n$, si procede dunque a discretizzare la serie temporale di potenze in categorie (bins), in modo tale da ottenere una serie di $n-1$ numeri interi che andranno a rappresentare gli stati del modello multistato. Se lo stato assunto dal sistema è 1, significa che $P_1 < P < P_2$, dove P è la potenza istantanea generata, se lo stato è 2 si ha che $P_2 < P < P_3$ e così via; se infine lo stato è $n-1$ significa che $P_{n-1} < P < P_n$.
5. Si calcola la matrice di transizione $\boldsymbol{M}$, ovvero la matrice che indica le probabilità di transizione tra gli stati. Tale matrice, di dimensioni $(n-1)$ x $(n-1)$, è stocastica, ovvero la somma degli elementi su ogni riga è uguale a 1:

$$\sum_{j=1}^{n-1} m_{ij} = 1 \qquad i = 1,2, \dots, n-1 \tag{11.2}$$

   Per calcolare la matrice di transizione si procede innanzitutto a calcolare il numero di transizioni di stato che avvengono nella serie di potenze discretizzata, per ogni stato assunto.
   Inizialmente la matrice $\boldsymbol{M}$, di dimensioni $(n-1)$ x $(n-1)$, ha valore zero in ciascun suo elemento.
   Considerando un i-esimo valore di stato assunto dal sistema, compreso nell'intervallo $[1,2, \dots, n-1]$, si vanno ad individuare tutti gli istanti nei quali la serie di potenze discretizzata assume tale valore: a questo punto, per ogni instante individuato, si va a rilevare lo stato assunto nell'instante successivo. Indicando tale stato con $j \in [1,2, \dots, n-1]$, si andrà ad incrementare l'elemento (i,j) della matrice $\boldsymbol{M}$ di 1. Tale procedimento si andrà ad effettuare per tutti gli $(n-1)$ stati assumibili dal sistema.
   La matrice ottenuta non sarà necessariamente simmetrica. Se però si andassero a considerare anche le transizioni che avvengono negli istanti precedenti a quelli in esame e non solo in quelli successivi, la matrice che si ottiene è sicuramente simmetrica. Non si andrà comunque ad includere tale risultato nel modello , in quanto per ovvi motivi non ha significato considerare transizioni di stato in instanti temporali precedenti.

A questo punto, per ottenere la matrice di transizione definitiva, si vanno a dividere gli elementi di ciascuna riga per la somma dei valori della riga corrispondente. In tale modo si otterrà una matrice stocastica, come deve essere per definizione.

Per illustrare i risultati del procedimento si mostra un esempio: si considera una serie storica del vento estratta dal database del progetto KNMI HYDRA dell'Istituto Meteorologico Reale Olandese. Dai dati di una stazione di misura si è prelevata una serie di velocità medie orarie del vento in un anno, il cui andamento è riportato in Fig. 16.

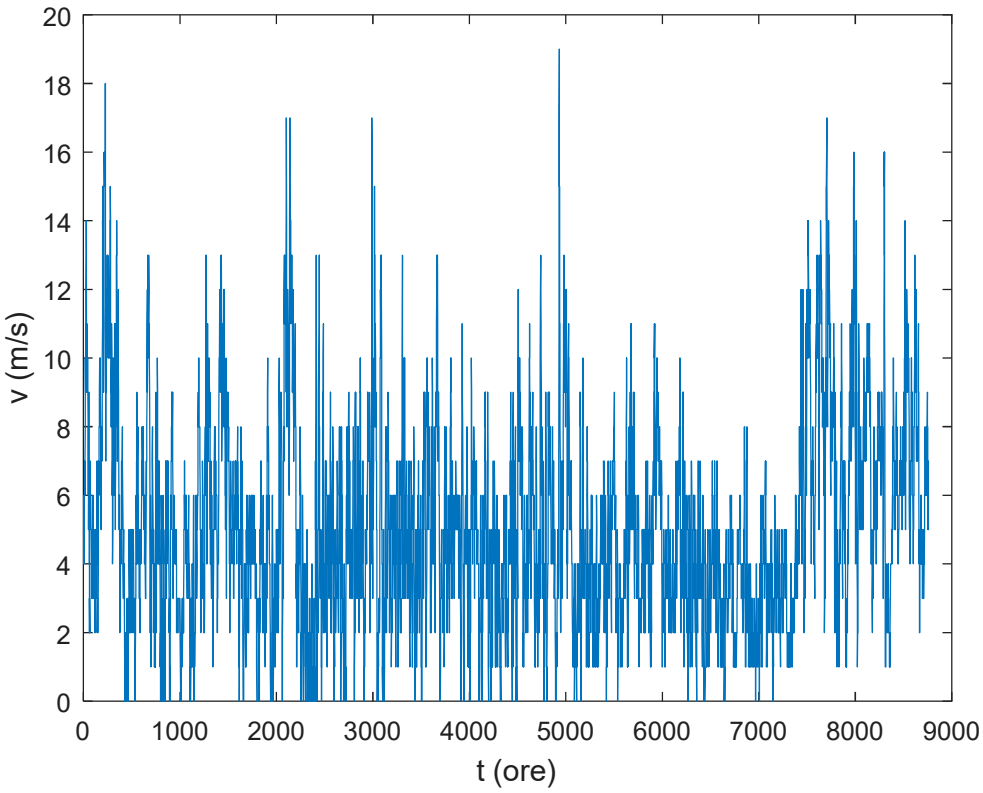

Figura 16-Andamento delle velocità medie orarie del vento.

Si considera poi una turbina eolica di potenza nominale pari a 600 kW, la cui curva caratteristica è mostrata in Fig. 17; come si può notare, la velocità di cut-in, ovvero la velocità alla quale la turbina inizia a produrre energia elettrica, è pari a circa 2 m/s, mentre la velocità di cut-out alla quale la caratteristica diventa decrescente è uguale a 22 m/s; date le velocità medie orarie del vento, si ricava dunque la serie annuale di potenza, riportata in Fig. 18.

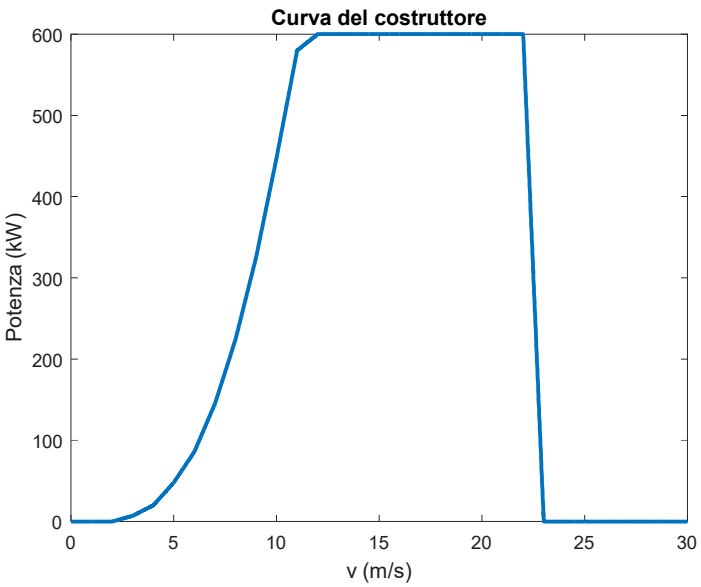

Figura 17-Curva del costruttore della turbina da 600 kW.

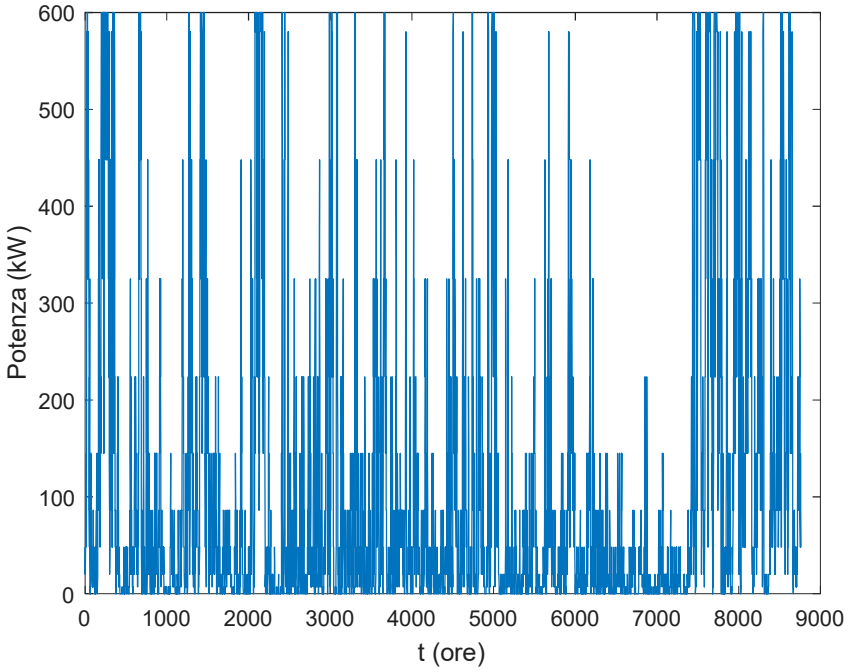

Figura 18-Andamento della potenza generata nel tempo.

Ottenuta la serie temporale di potenze, si calcola la funzione di sopravvivenza, che viene qui riportata in Fig. 19: tale funzione è definita per gli 11 seguenti valori di potenza, espressi in kW:

$$(0; 7; 20; 48; 86; 145; 224; 325; 448; 580; 600)\ \ [kW]$$

Tali valori definiscono dunque i livelli di generazione della turbina eolica considerata. È possibile a questo punto discretizzare la serie di potenze sulla base dei livelli di potenza ottenuti; in tal modo la turbina eolica sarà rappresentata da un modello a 10 stati.

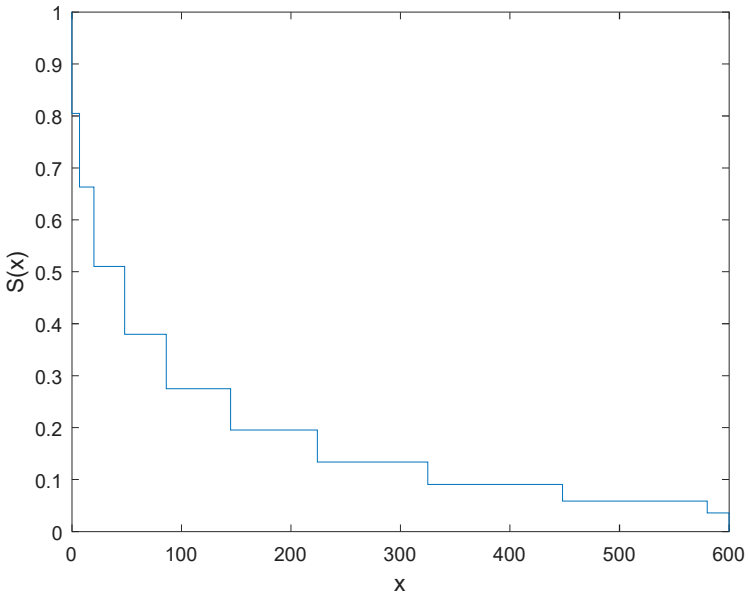

Figura 19-Funzione di sopravvivenza empirica ottenuta dalla serie temporale di potenze.

Si procede dunque al calcolo della matrice di transizione degli stati; per completezza della trattazione, si riporta in primo luogo in Tabella 1 la matrice di transizione degli stati ottenuta considerando anche le transizioni negli istanti precedenti: come ci si doveva aspettare, la matrice risulta simmetrica; gli elementi sulla diagonale indicano che si ha una permanenza del livello di stato considerato; si può notare come si hanno per la maggior parte transizioni verso livelli di stato di poco superiori a quelli considerati, ad esempio transizioni dallo stato 1 allo stato 2, dallo stato 4 allo stato 5 e così via, mentre risultano di fatto trascurabili transizioni verso livelli assai più elevati, come mostrato dagli elementi della matrice di valore molto basso o nullo. Dividendo ogni riga della matrice per la sua somma si ottengono le probabilità di transizione di stato; la matrice che ne risulta è riportata in Tabella 2.

Tabella 1-Matrice di transizione degli stati considerando transizioni reversibili nel tempo.

2790	533	77	16	2	1	0	0	1	0
533	1282	551	95	17	3	0	1	0	0
77	551	1368	569	87	17	2	1	0	1
16	95	569	1094	434	67	15	4	0	0
2	17	87	434	858	373	50	10	3	0
1	3	17	67	373	622	250	46	8	4
0	0	2	15	50	250	516	206	37	8
0	1	1	4	10	46	206	314	135	35
1	0	0	0	3	8	37	135	244	132
0	0	1	0	0	4	8	35	132	848

Tabella 2-Matrice di transizione degli stati normalizzata considerando transizioni reversibili nel tempo.

0.8158	0.1558	0.0225	0.0047	0.0006	0.0003	0	0	0.0003	0
0.2147	0.5165	0.222	0.0383	0.0068	0.0012	0	0.0004	0	0
0.0288	0.2061	0.5118	0.2129	0.0325	0.0064	0.0007	0.0004	0	0.0004
0.007	0.0414	0.248	0.4769	0.1892	0.0292	0.0065	0.0017	0	0
0.0011	0.0093	0.0474	0.2366	0.4678	0.2034	0.0273	0.0055	0.0016	0
0.0007	0.0022	0.0122	0.0482	0.2682	0.4472	0.1797	0.0331	0.0058	0.0029
0	0	0.0018	0.0138	0.0461	0.2306	0.476	0.19	0.0341	0.0074
0	0.0013	0.0013	0.0053	0.0133	0.0612	0.2739	0.4176	0.1795	0.0465
0.0018	0	0	0	0.0054	0.0143	0.0661	0.2411	0.4357	0.2357
0	0	0.001	0	0	0.0039	0.0078	0.034	0.1284	0.8249

Come si è accennato in precedenza, non avendo significato considerare la reversibilità temporale delle transizioni, il caso di interesse è quello in cui considerano unicamente le transizioni di stato negli istanti di tempo successivi; la matrice di transizione che ne deriva in questo caso è mostrata in Tabella 3: si può subito vedere come in tale situazione la matrice non risulti simmetrica. Si nota però che in alcuni casi la differenza tra il numero di transizioni dallo stato i allo stato j e il numero di transizioni dallo stato j allo stato i, ovvero nel caso opposto, sia molto piccola: per esempio risulta che le transizioni da 1 a 2 sono 264, mentre nel caso opposto, da 2 a 1 sono 269; il numero di transizioni da 2 a 3 è pari a 268 e nel caso simmetrico è uguale a 283.

Come risultava nella matrice di transizione precedente, anche in tal caso il numero di transizioni verso livelli di stato più elevati è trascurabile, ad indicare che non vi sono brusche variazioni nella potenza istantanea generata. Si nota infine come il numero di transizioni totali considerate è esattamente la metà rispetto al caso precedente: effettuando la somma di tutti gli elementi della matrice, risulta che il numero di tutte transizioni analizzate nel caso simmetrico è pari a 17518, mentre nel caso non simmetrico è uguale a 8759. Anche in tale ambito si è ricavata la matrice normalizzata indicante le probabilità di transizione, la quale è riportata in Tabella 4.

Tabella 3-Matrice di transizione degli stati considerando solo transizioni negli istanti temporali successivi.

1395	264	40	9	2	0	0	0	0	0
269	641	268	46	14	2	0	1	0	0
37	283	684	281	40	11	0	0	0	1
7	49	288	547	211	34	11	0	0	0
0	3	47	223	429	184	22	7	2	0
1	1	6	33	189	311	125	22	4	3
0	0	2	4	28	125	258	106	14	5
0	0	1	4	3	24	100	157	70	17
1	0	0	0	1	4	23	65	122	64
0	0	0	0	0	1	3	18	68	424

Tabella 4-Matrice di transizione degli stati normalizzata considerando solo transizioni negli istanti temporali successivi.

0.8158	0.1544	0.0234	0.0053	0.0012	0	0	0	0	0
0.2168	0.5165	0.216	0.0371	0.0113	0.0016	0	0.0008	0	0
0.0277	0.2117	0.5116	0.2102	0.0299	0.0082	0	0	0	0.0007
0.0061	0.0427	0.2511	0.4769	0.184	0.0296	0.0096	0	0	0
0	0.0033	0.0513	0.2432	0.4678	0.2007	0.024	0.0076	0.0022	0
0.0014	0.0014	0.0086	0.0475	0.2719	0.4475	0.1799	0.0317	0.0058	0.0043
0	0	0.0037	0.0074	0.0517	0.2306	0.476	0.1956	0.0258	0.0092
0	0	0.0027	0.0106	0.008	0.0638	0.266	0.4176	0.1862	0.0452
0.0036	0	0	0	0.0036	0.0143	0.0821	0.2321	0.4357	0.2286
0	0	0	0	0	0.0019	0.0058	0.035	0.1323	0.8249

È possibile effettuare una rappresentazione grafica tridimensionale delle matrici di transizione ottenute: ponendo sull'asse delle X le posizioni degli indici di riga delle matrice, sull'asse Y le posizioni degli indici di colonna e sull'asse delle Z il valore dell'elemento m_{ij} della matrice alla riga i e alla colonna j, si ottengono i seguenti diagrammi a steli, riportati in Fig. 20 e 21 rispettivamente per la matrice di transizione simmetrica e per la matrice non simmetrica: nonostante i grafici si riferiscano a casi differenti, essi mostrano un andamento della distribuzione dei punti molto simile tra loro, il che ribadisce le considerazioni precedenti che sottolineavano come le due matrici di transizione mostrino una distribuzione dei valori simile, riferita però a un insieme di transizioni differente.

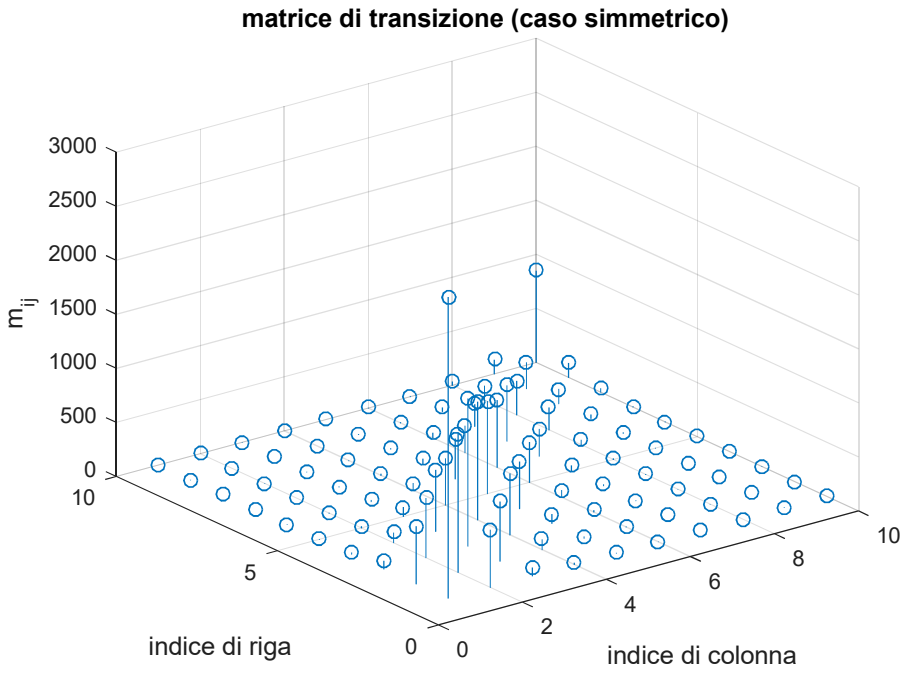

Figura 20-Distribuzione dei valori della matrice di transizione nel caso simmetrico.

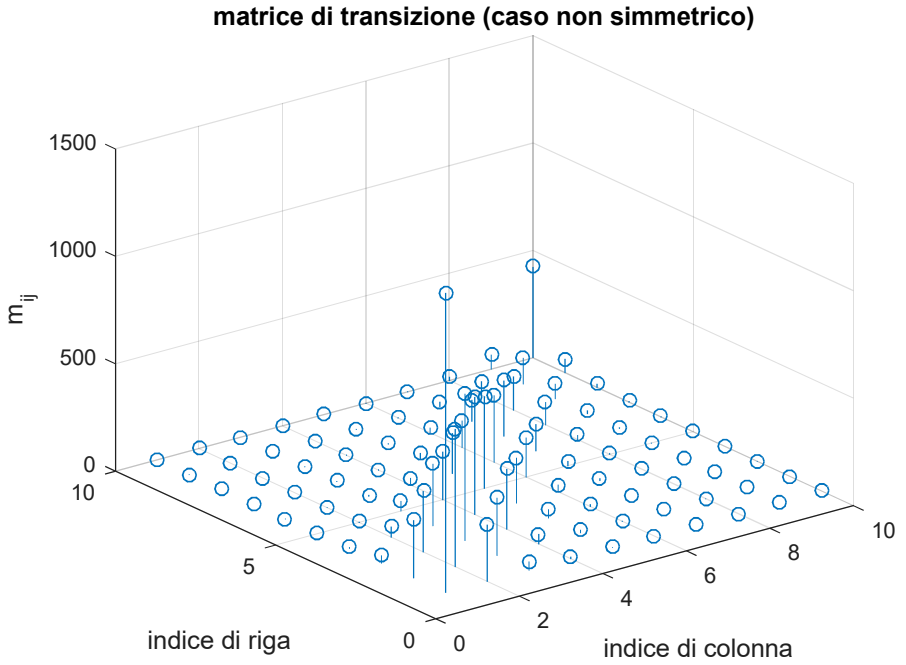

Figura 21-Distribuzione dei valori della matrice di transizione nel caso non simmetrico.

12-Stima e simulazione del modello statistico

Il modello proposto per la rappresentazione degli elementi tempo-varianti nella simulazione Monte Carlo non sequenziale, in modo tale da conservare la dipendenza statistica tra di essi, si compone di 2 fasi:

1. Stima del modello di elementi tempo-varianti.
2. Simulazione del modello attraverso lo state sampling.

Gli stadi richiesti per la fase di stima del modello degli elementi tempo-varianti sono i seguenti:

1. Trasformazione delle serie storiche: comprende la trasformazione della serie storiche, che possono avere qualsiasi distribuzione di probabilità, in serie con distribuzione uniforme. Questo comporta tre passaggi:

    a. Stima non parametrica della densità di probabilità per mezzo del *KDE*. Applicando questo metodo di stima alle serie storiche, si ottiene una stima della densità di probabilità $f_x(v_j)$ per ciascuna variabile aleatoria v_j.

    b. Calcolo delle funzioni di distribuzione cumulativa (*CDF*) e le loro rispettive funzioni inverse. La funzione di distribuzione cumulativa di una variabile casuale continua $F_x(v_j)$ è definita dall'integrale della funzione di densità di probabilità. Poiché la *CDF* è monotona crescente e biunivoca, vi è una funzione inversa della funzione di probabilità cumulativa F_{xj}^{-1} per ciascuna variabile casuale.

    c. Trasformazione delle serie storiche serie aventi distribuzione uniforme. Date le *CDF* di ciascuna variabile casuale, un nuovo insieme di serie temporali $u_1, u_2, \dots, u_w$ caratterizzate da densità di probabilità uniformi può essere ottenuto applicando la trasformazione F_x alle serie storiche. Tale trasformazione è data dalla seguente espressione:

$$\begin{cases} u_1 = F_{x1}(v_1) \\ u_2 = F_{x2}(v_2) \\ \vdots \\ u_w = F_{xw}(v_w) \end{cases} \tag{12.1}$$

2. Discretizzazione per massimizzare la mutua informazione: Questo processo ha lo scopo di rappresentare la dipendenza statistica non lineare tra le variabili casuali in modo tale da preservare la mutua informazione tra ciascuna coppia di variabili aleatorie. Il processo di discretizzazione viene applicato alle serie uniformi $u_1, u_2, \dots, u_w$ ottenute in precedenza. Come risultato, si ottengono le serie discrete di numeri interi $L_1, L_2, \dots, L_w$.

3. Rappresentazione della dipendenza statistica tra le variabili discrete: questa fase consiste nella rilevazione della dipendenza statistica tra le varie variabili casuali attraverso la stima della rete bayesiana codifica le probabilità condizionate tra di esse. La rete bayesiana è rappresentata da una struttura, dai parametri e dal database di casi che in tale ambito è rappresentato dalle serie discrete $L_1, L_2, \dots, L_w$ calcolate precedentemente. La struttura della rete è stimata attraverso l'algoritmo K2, mentre i parametri sono calcolati applicando il metodo della massima verosimiglianza.

Una volta che il modello utilizzato per rappresentare elementi tempo-varianti è stato stimato , l'algoritmo di simulazione Monte Carlo non sequenziale può essere modificato per incorporare tali elementi. Gli stadi richiesti per la fase di simulazione del modello sono i seguenti:

1. Campionamento multivariato delle variabili discrete: si tratta di ottenere campioni discreti uniformemente distribuiti utilizzando i dati ricavati dalla stima della rete bayesiana. Un campione i-esimo dell'insieme di variabili casuali discrete $L^{Si} = [L_1^{Si}\ L_2^{Si}\ \dots\ L_w^{Si}]$ descritte dalla rete bayesiana, dove l'indice Si è usato per differenziare le variabili campionate dalle rispettive osservazioni storiche, è ottenuto seguendo ricorsivamente i nodi del grafo della rete bayesiana dai nodi genitore ai nodi figlio.
2. Trasformazione delle variabili discrete in variabili continue: dall'insieme dei valori ottenuti nella fase precedente, $L^{Si} = [L_1^{Si}\ L_2^{Si}\ \dots\ L_w^{Si}]$, il campione j-esimo della variabile trasformata u_j^{Si} viene ottenuto calcolando un numero reale avente distribuzione uniforme nell'intervallo $\left[L_j^{Si}, L_j^{Si}+1\right]$. Per calcolare un numero r con distribuzione uniforme compreso in un intervallo $[a,b]$ si utilizza la seguente espressione:

$$r = (b-a)\cdot U + a \tag{12.2}$$

   dove U è un numero casuale avente distribuzione uniforme compreso nell'intervallo $[0,1]$. In tal modo si ottengono le serie uniforme sintetiche $u_1^{Si}, u_2^{Si}, \dots, u_w^{Si}$.
3. Trasformazione delle variabili uniformi in variabili aventi la distribuzione delle serie storiche originali: questo processo viene utilizzato per ottenere campioni con la stessa distribuzione delle serie storiche. I campioni variabili casuali uniformi vengono trasformati applicando la funzione di trasformazione inversa F_{xj}^{-1} a ciascuna variabile u_j^{Si} secondo le espressioni seguenti:

$$\begin{cases} v_1^{Si} = F_{x1}^{-1}(u_1^{Si}) \\ v_2^{Si} = F_{x2}^{-1}(u_2^{Si}) \\ \vdots \\ v_w^{Si} = F_{xw}^{-1}(u_w^{Si}) \end{cases} \tag{12.3}$$

13-Calcolo degli indici di affidabilità

13.1-Calcolo degli indici di adeguatezza

Per la valutazione dell'adeguatezza dei sistemi di generazione, in questa tesi si fa ricorso ad un metodo analitico, simile a quello basato sulla definizione della Capacity Outage Probability Table discusso in precedenza. Per applicare tale metodo al calcolo degli indicatori di adeguatezza, la prima priorità è selezionare un appropriato modello in grado di rappresentare il carico del sistema elettrico in esame e i generatori presenti nella rete.

Per quanto riguarda il carico di rete, è possibile definire una *curva cronologica* del carico elettrico . Una delle caratteristiche principali del carico cronologico è che esso varia con le condizioni climatiche, a seconda dell'ora del giorno, del giorno della settimana e della settimana dell'anno. Per sviluppare il modello di carico in funzione del tempo, le caratteristiche del carico orario devono essere combinate con i dati di carico di picco annuale /settimanale /giornaliero per generare modelli annuali. Tre tipi di dati caratteristici del carico possono essere adoperati per definire la curva cronologica annuale e sono i seguenti:

a) Carico giornaliero in 24 ore, espresso in percentuale del carico di picco giornaliero;
b) Carico settimanale in 7 giorni, espresso in percentuale del carico di picco settimanale;
c) Carico annuale in 52 settimane, espresso in percentuale del carico di picco annuale.

Con il profilo del carico di picco annuale e delle suddette percentuali di carico giornaliero/settimanale / annuale, il carico orario $L(t)$ per l'ora t può essere calcolato utilizzando la seguente formula:

$$L(t) = L_y \cdot P_w \cdot P_d \cdot P_h(t) \qquad (13.1)$$

dove L_y è il carico di picco annuale, espresso in MW, P_w è la percentuale del carico settimanale in termini di picco annuale, P_d è la percentuale del carico giornaliero in termini di picco settimanale, $P_h(t)$ è la percentuale del carico orario in termini di picco giornaliero.

È interessante notare come la domanda di carico oraria $L(t)$ rappresenti soltanto il carico di picco orario del sistema, dando luogo a $52 \cdot 7 \cdot 24 = 8736$ valori indipendenti per un anno.

Dalla curva cronologica del carico è possibile ricavare la *curva di durata del carico* (*LDC*, load duration curve), definita disponendo i valori della curva cronologica in ordine decrescente. La Fig. 22 mostra un esempio di curva cronologica del carico assieme alla rispettiva curva di durata.

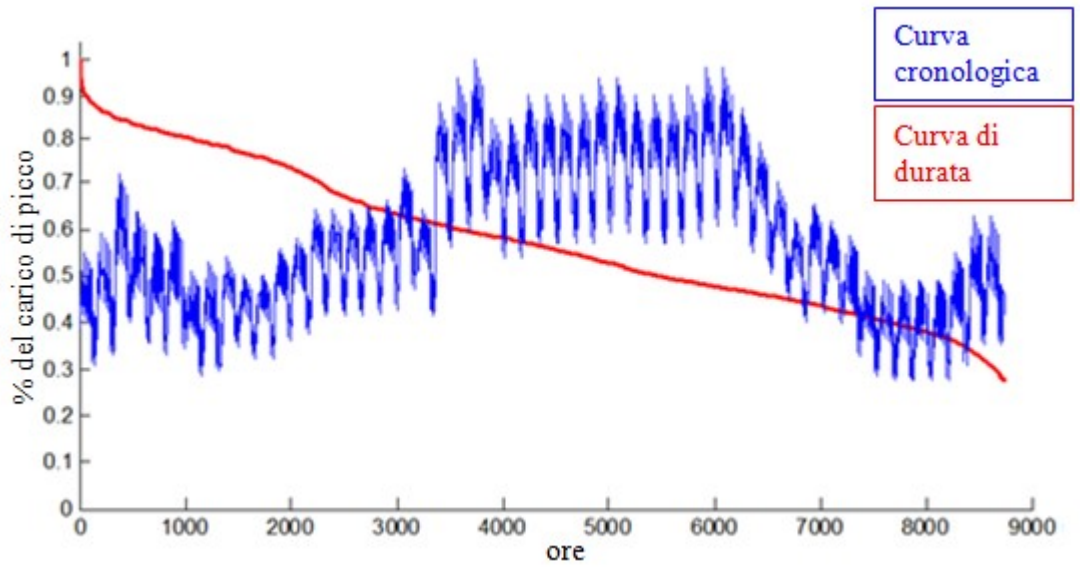

Figura 22-Esempio di curva di carico cronologica oraria e di curva di durata del carico in un anno.

Dalla curva di durata del carico è possibile ricavare la funzione di distribuzione cumulativa e da qui la rispettiva funzione di sopravvivenza. Quest'ultima sarà la funzione rappresentativa del modello di carico utilizzata per la valutazione dell'adeguatezza.

Nel modello di generazione, un generatore può essere rappresentato da un modello a due stati o da un modello multistato.

Considerando dapprima il modello a due stati, i dati caratteristici del generatore che vengono adoperati nel modello sono la sua potenza generata P_G e i coefficienti di disponibilità ρ e di indisponibilità $(1-\rho)$. Definiti questi dati si possono utilizzare due rappresentazioni alternative per il modello del generatore:

- Rappresentazione riferita alla potenza generata (Fig. 23);
- Rappresentazione con generatore ideale e un carico fittizio aggiuntivo, quest'ultimo equivalente all'indisponibilità del generatore (Fig.24).

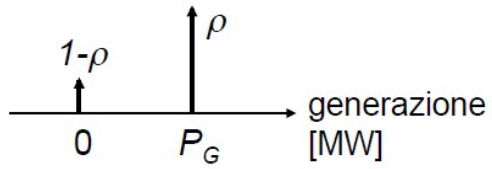

Figura 23-Modello del generatore a due stati con riferimento alla potenza generata.

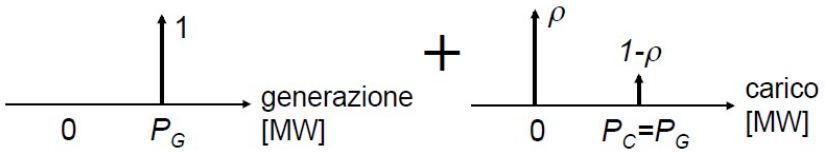

Figura 24-Modello del generatore a due stati composto da generazione ideale e carico fittizio.

Per quanto concerne il modello con più di due stati, i dati richiesti per la rappresentazione sono in questo caso i vari livelli di generazione $0, P_G^a, P_G^b, P_G^c, \ldots$ caratteristici dell'unità considerata, e i corrispettivi coefficienti $\rho_0, \rho_a, \rho_b, \rho_c, \ldots$ che indicano la probabilità che la generazione sia a quel particolare livello. Un esempio tipico di tale modello è quello a 3 stati, in cui si considera che l'unità di generazione possa funzionare erogando un valore di potenza nullo oppure una potenza ridotta P_{GR} rispetto alla capacità nominale P_{GM} oppure che funzioni alla capacità nominale P_{GM} stessa. In Fig. 25 e 26 vengono mostrate le rappresentazioni di tale modello a tre stati, rispettivamente nel caso in cui si considera unicamente il livello di potenza generata e nel caso in cui si scompone la rappresentazione in generatore ideale e carico fittizio.

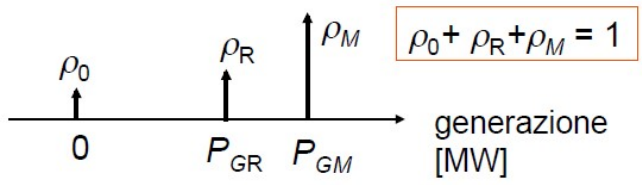

Figura 25-Modello del generatore a tre stati con riferimento ai livelli di potenza generata.

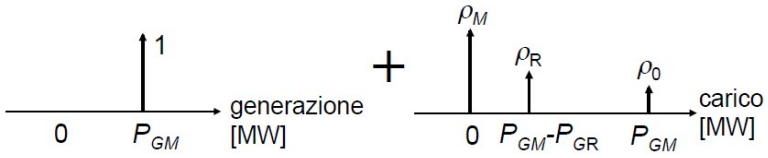

Figura 26-Modello del generatore a tre stati composto da generazione ideale e carico fittizio.

Per il calcolo degli indicatori di adeguatezza si utilizza il seguente metodo di *convoluzione*: si considera dapprima un modello di generazione a due stati; partendo dalla funzione di sopravvivenza del carico $S_0(C)$, dove C è il livello di carico espresso in MW, e dal modello di un generatore con potenza generata P_{G1} e disponibilità ρ_1 come composizione di generatore ideale e carico fittizio, si aggiunge il contributo del carico fittizio al carico iniziale, calcolando la nuova *CCDF* del carico attraverso la relazione:

$$S_1(C) = \rho_1 S_0(C) + (1-\rho_1) S_0(C - P_{G1}) \tag{13.2}$$

Tale espressione vuole indicare che il carico ha una probabilità ρ_1 di non variare e una probabilità $(1-\rho_1)$ di aumentare della quantità P_{G1}. Questa operazione si ripete per ogni contributo degli n generatori presenti nella rete in esame, sapendo che all'iterazione i-esima, in cui si aggiunge il contributo del generatore i-esimo di potenza generata P_{Gi} e disponibilità ρ_i, la $S_0(C)$ da considerare sarà la $S_{i-1}(C)$, ovvero quella ottenuta all'iterazione precedente: si può quindi definire la seguente espressione generale della convoluzione:

$$S_i(C) = \rho_i S_{i-1}(C) + (1-\rho_i) S_{i-1}(C - P_{Gi}) \tag{13.3}$$

Al termine del processo di convoluzione, si calcola la generazione totale (ideale) sommando le potenze generate $P_{G1}, P_{G2}, \dots P_{Gn}$, e si procede al calcolo degli indici di adeguatezza: il *LOLP* del sistema, espresso in p.u., viene determinato nella maniera seguente: si individua l'ascissa della funzione così ottenuta corrispondente alla potenza generata totale $P_{GTOT} = \sum_{i=1}^{n} P_{Gi}$; il valore dell'ordinata corrispondente rappresenta il *LOLP*. Per ottenere il *LOLE*, definito in ore/anno, è sufficiente, come discusso in precedenza, moltiplicare il valore del *LOLP* per il periodo di osservazione T che si è considerato nell'analisi.

Il valore atteso della potenza non servita *EPNS*, espresso in MW, si calcola nel modo seguente:

$$EPNS = \frac{1}{LOLP} \int_{D \geq 0} S(C) dC \tag{13.4}$$

dove $D = C - G$ è il deficit di produzione, ovvero la differenza tra la generazione G e il carico C del sistema. Esso è pertanto calcolabile determinando il valore dell'area sottesa dalla curva ottenuta tra l'ascissa

corrispondente alla potenza generata totale e l'ultimo valore di potenza per il quale tale curva è definita. Il valore atteso dell'energia non servita, espresso in MWh/anno, sarà semplicemente il prodotto dell'*EPNS* per il periodo di osservazione T.

Se nella rete in esame sono presenti generatori aventi modello multistato, il metodo di convoluzione è analogo al precedente, sapendo però che in questo caso si dovrà aggiungere il contributo di tutti i livelli di carichi fittizi che costituiscono il modello al carico iniziale.

Considerando la *CCDF* del carico all'iterazione $i-1$ e il generatore a tre stati definito precedentemente come esempio, la nuova *CCDF* del carico risulterà essere:

$$S_i(C) = \rho_M S_{i-1}(C) + \rho_R S_{i-1}(C - (P_{GM} - P_{GR})) + \rho_0 S_{i-1}(C - P_{GM}) \qquad (13.5)$$

Il primo termine indica il contributo del carico equivalente al generatore funzionante alla capacità massima, il secondo termine rappresenta il contributo del carico equivalente alla potenza generata ridotta, mentre l'ultimo termine indica il contributo del carico equivalente al generatore che eroga potenza nulla.

Al termine del processo, si aggiungono i livelli di capacità massima di tutti i generatori e si calcolano gli indici di adeguatezza come definito in precedenza.

Nell'applicazione del metodo di convoluzione, il numero di gradini della curva rappresentativa della *CCDF* aumenta gradualmente ad ogni iterazione in cui si aggiunge il contributo di un generatore: la posizione di tali gradini è predeterminabile in base ai valori di potenza generata e del carico iniziale, rappresentato dalla funzione di sopravvivenza empirica ricavata dalla curva di durata. Per sistemi con molti generatori e molti livelli di carico, può diventare estremamente oneroso dal punto di vista computazionale l'elaborazione del processo di convoluzione; per provvedere a tale problema, è possibile operare per incrementi di potenza uguali, predefiniti per avere tutte le potenze generate e i livelli di carico multipli dell'incremento: ad ogni iterazione si otterrà dunque una curva con un numero minore di punti che tuttavia rappresenterà un approssimazione accettabile e permetterà di eseguire i calcoli in tempi più brevi.

13.1-Calcolo degli indici F&D

Per il calcolo degli indici di affidabilità quali *SAIFI*, *SAIDI*, che hanno lo scopo di valutare la continuità di servizio nel sistema elettrico, si utilizza un algoritmo di simulazione Monte Carlo non sequenziale in grado di incorporare il modello statistico descritto precedentemente. Tale metodo si articola nei seguenti passaggi:

1. Stima del modello degli elementi tempo-varianti attraverso il modello proposto;
2. Selezione di uno stato del sistema $\underline{x}$ dipendente dalla disponibilità dei componenti, dal livello di carico, dalla disponibilità della generazione di potenza, ecc. Su tale stato:
    a. Si campionano gli stati degli elementi rappresentati da variabili aleatorie statisticamente dipendenti usando il modello proposto.

   b. Si campionano gli stati dei componenti rappresentati da variabili aleatorie indipendenti dalle rispettive distribuzioni di probabilità.
3. Si calcola il valore della funzione dell'indice $G(\underline{x})$.
4. Si aggiorna la stima di $\mathbb{E}(G)$ corrispondente al valore atteso dell'indice di affidabilità basato sul risultato ottenuto al passo 3.
5. Se il coefficiente di variazione β della stima è inferiore a una soglia prefissata, si termina la simulazione; altrimenti si ritorna al passo 2.

14-Risultati

La rete che si è utilizzata in questa tesi per testare le valutazioni di affidabilità è basata su una modifica dell'IEEE RTS (Reliability test system), in cui si sono introdotte 9 windfarm, distribuite sui nodi della rete. L'IEEE RTS originale ha 24 nodi, 32 unità di generazione per una capacità complessiva installata di 3405 MW e un carico di picco di 2850 MW. Il diagramma unifilare di tale sistema è mostrato in Fig. 27. I dati relativi all'IEEE RTS sono riportati nell'appendice A.

I dati del vento sono stati estratti dalla banca dati del progetto KNMI HYDRA dell'Istituto meteorologico Reale Olandese, che comprende oltre 20 anni di misurazioni. I dati provenienti da 9 stazioni di misura sono stati utilizzati per caratterizzare le windfarm collegate ai nodi dell'IEEE-RTS, come viene mostrato nella Tabella 5.

Tabella 5-Windfarm aggiunte ai nodi della rete IEEE RTS.

Nodo	Serie del vento per la windfarm aggiunta
1	s-210
2	s-225
7	s-235
15	s-242
16	s-248
18	s-269
21	s-275
22	s-283
23	s-380

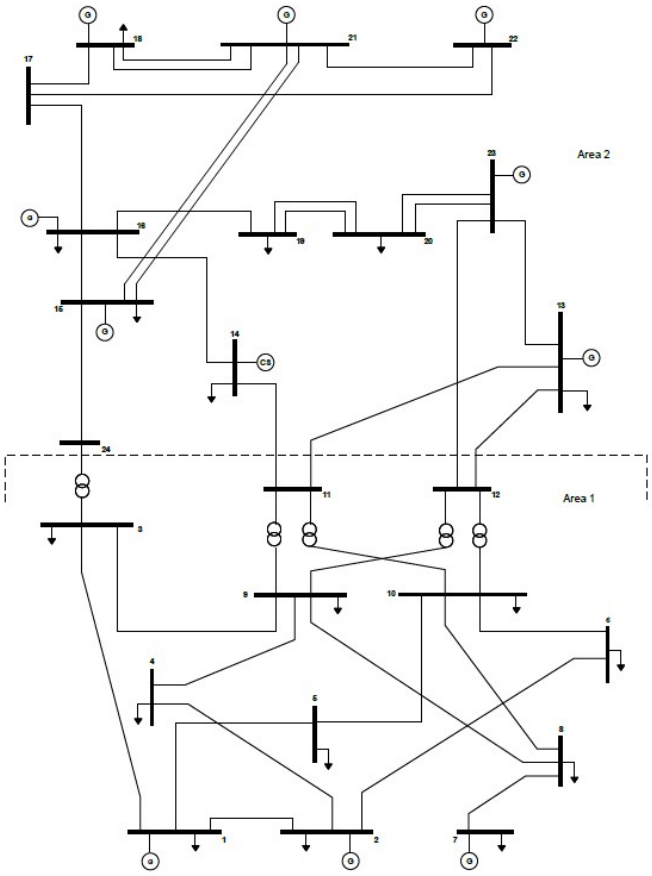

Figura 27-Schema dell'IEEE RTS.

Si è considerato che ciascuna windfarm abbia 25 unità di generazione, ciascuna di potenza nominale pari a 2 MW. Tutte le turbine hanno la stessa caratteristica di generazione di potenza, la cui curva è mostrata in Fig. 28: la velocità di cut-in è pari a 3.01 m/s, mentre la velocità di cut-out risulta essere uguale a 25.01 m/s. La capacità complessiva aggiunta dalle windfarm è dunque pari a 450 MW.

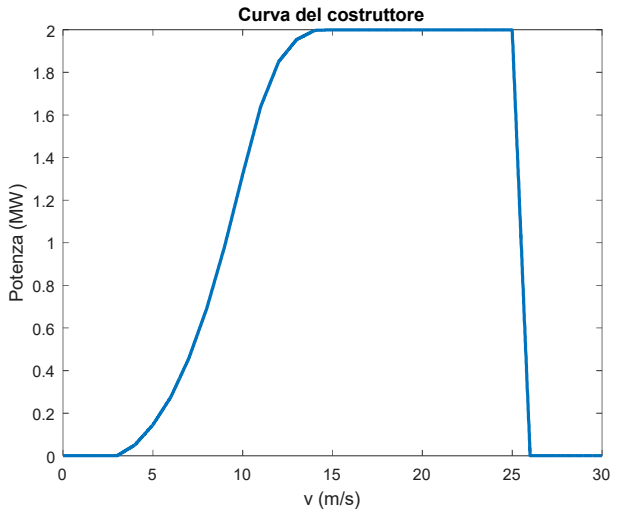

Figura 28-Curva del costruttore della turbina da 2 MW.

14.1-Indici di adeguatezza

Per il calcolo degli indici di adeguatezza si è per prima cosa provveduto a determinare i modelli del carico e dei generatori basandosi sui dati della rete test e delle serie storiche dalle quali si è determinata la generazione eolica introdotta dalle windfarm aggiuntive.

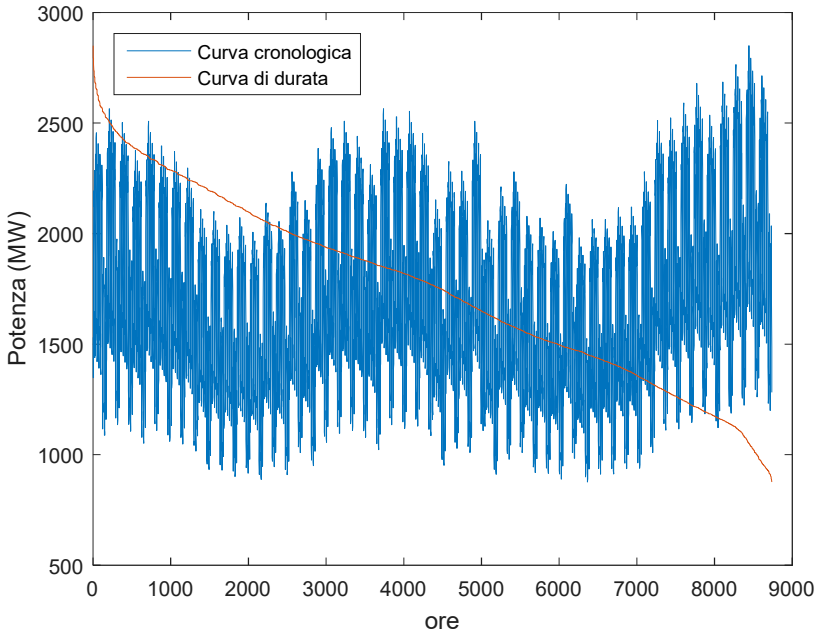

Figura 29-Curva cronologica e curva di durata del carico della rete IEEE RTS riferite ad un anno.

La Fig. 29 mostra la curva cronologica del carico dell'IEEE RTS su un anno, assieme alla rispettiva curva di durata. Il livello di carico massimo è pari a 2850 MW, mentre il carico minimo corrisponde a 876.48 MW. Dalla *LDC* della rete test si è ricavata la funzione di sopravvivenza empirica, che è riportata in Fig. 30.

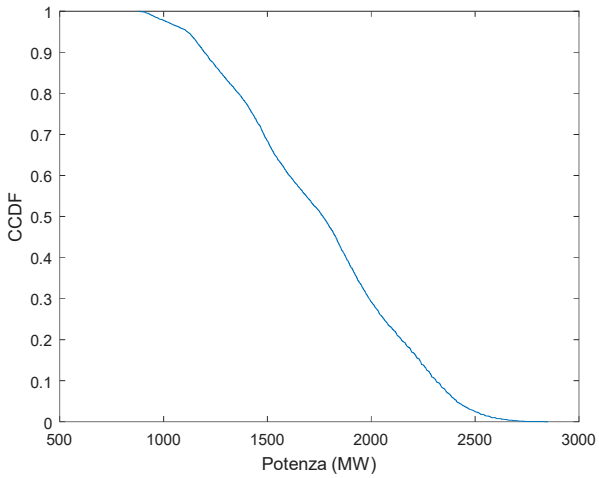

Figura 30-Funzione di sopravvivenza empirica del carico della rete IEEE RTS.

Per quanto concerne i sistemi di generazione, dai dati della rete IEEE RTS si apprende che tutti i generatori sono modellabili tramite modelli a due stati: al posto dei coefficienti di disponibilità sono stati definiti i *FOR*, che, come noto, sono i complementi a uno delle disponibilità; i dati relativi alle potenze generate e ai *FOR* sono disponibili nell'appendice A.

Per definire invece il modello delle windfarm che si sono aggiunte nella rete test, si è proceduti in questo modo: partendo dalle serie storiche, riferite ad un anno, delle velocità medie orarie del vento, si sono inseriti tali valori nella curva del costruttore rappresentativa della turbina da 2 MW che si è considerata, ottenendo così delle serie di potenze che rappresentano la generazione delle windfarm nell'anno di riferimento. Da queste serie di valori di potenza si sono ricavate le funzioni di sopravvivenza empiriche e da qui, attraverso l'operazione di derivazione, le rispettive funzioni di massa di probabilità: quest'ultime sono le funzioni che si sono adoperate per identificare il modello delle windfarm, in maniera tale da poterlo utilizzare nel metodo di convoluzione.

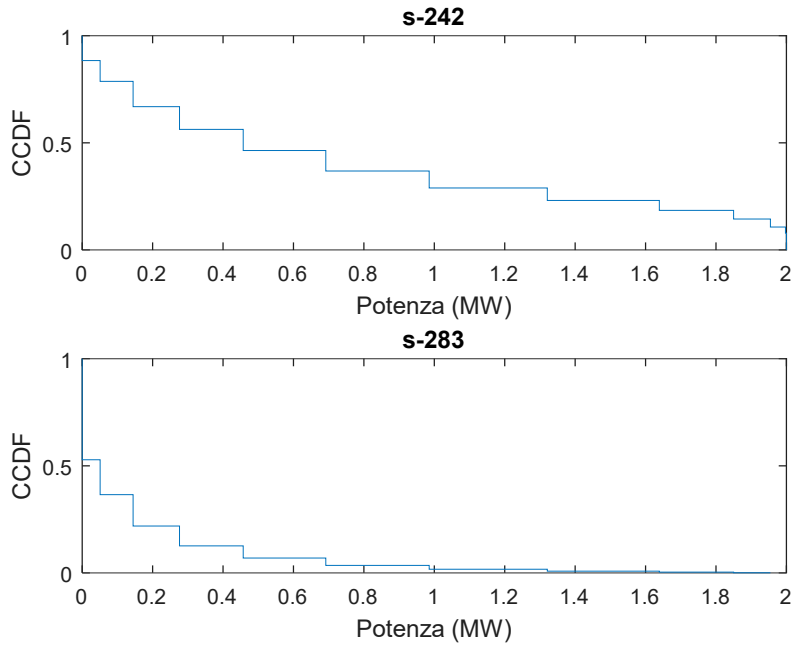

Figura 31-Funzioni di sopravvivenza empiriche.

La Fig. 31 mostra l'andamento delle funzioni di sopravvivenza empiriche riferite a due serie di potenze, una ricavata dalla serie di velocità del vento s-242, l'altra dalla serie s-283. Da queste funzioni si sono ottenute le rispettive funzioni di massa di probabilità, riportate in Fig. 32: da come si può notare, si deduce che il modello della windfarm interessata dal vento della serie s-242 è definito a 13 stati, mentre la windfarm soggetta al vento della serie s-283 è caratterizzata da un modello a 11 stati. Si può inoltre notare che nei generatori riferiti alla serie s-242 il livello di potenza massima ha una probabilità mediamente pari a quella degli altri livelli di potenza, mentre nelle unità di generazione riferite alla serie s-283 il livello di potenza massima ha una probabilità molto ridotta, mentre la probabilità più elevata è quella assunta dal livello zero, segno che questi generatori avranno una maggiore probabilità di funzionare senza erogare potenza elettrica.

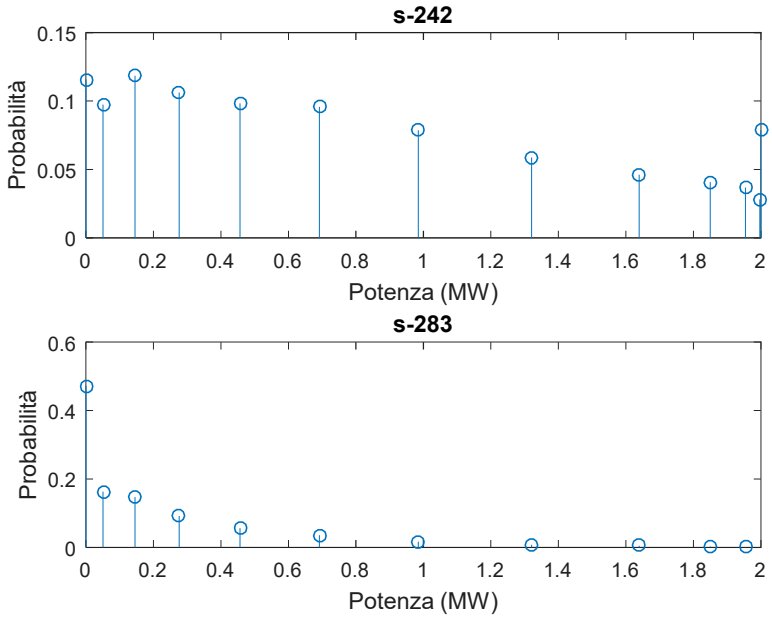

Figura 32-Funzioni di massa di probabilità.

Per stabilire una base per valutare l'impatto di incorporare le windfarm nel sistema, si sono calcolati gli indici di adeguatezza per la prima volta senza considerare la generazione di energia eolica.

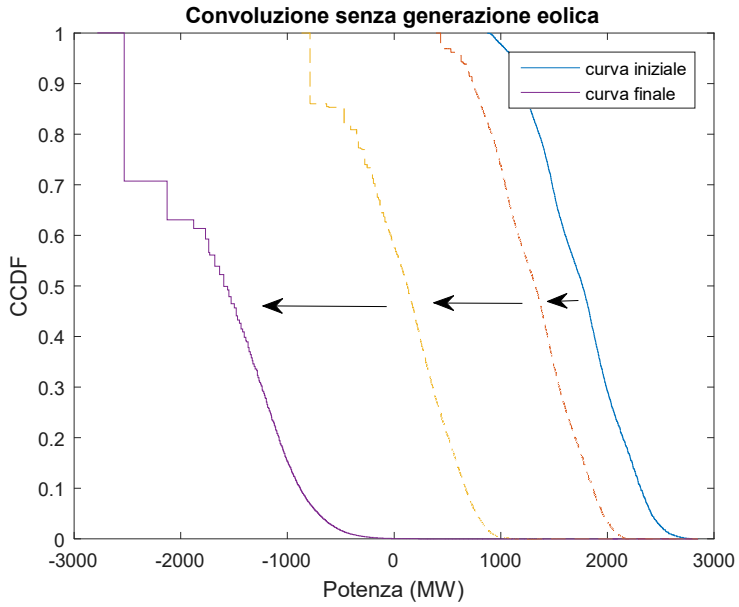

Figura 33-Convoluzione senza generazione eolica.

In Fig. 33 vengono mostrati i risultati del processo di convoluzione: come si può notare sull'asse delle potenze sono indicati valori negativi: questo perché si è effettuata la convoluzione considerando i valori delle potenze generate con segno negativo; si è infatti assunta la convenzione secondo la quale la generazione rappresenta per la rete un carico negativo. Fatta tale assunzione, si è effettuato il procedimento di convoluzione considerando direttamente il modello dei generatori riferito alla potenza generata, con i segni

scambiati per i livelli di potenza, ottenendo in tal modo un risultato equivalente a quello ottenibile con il procedimento descritto in precedenza. In questo caso però il valore del *LOLP* sarà pari al valore dell'ordinata assunta in corrispondenza del livello zero di potenza, e il calcolo della *EPNS* si effettuerà integrando la curva ottenuta al termine del processo tra 0 MW e l'ultimo valore di potenza per la quale la curva è definita.

La figura mostra dunque l'andamento della funzione di sopravvivenza ottenuta alla fine del processo di convoluzione, assieme a quello della funzione di partenza che, si ricorda, corrisponde alla *CCDF* del carico della rete test. Le curve tratteggiate rappresentano invece l'andamento della *CCDF* in stadi intermedi del processo di convoluzione: come viene indicato, a mano a mano che il processo avanza, la curva si sposta progressivamente verso sinistra, ad indicare che all'aumentare del contributo delle unità di generazione della rete, i valori degli indici di adeguatezza si riducono, e dunque la generazione contribuisce in maniera progressiva a soddisfare la domanda di energia elettrica della rete.

Dalla *CCDF* ottenuta al termine della convoluzione si sono potuti calcolare gli indici di adeguatezza, i cui valori sono i seguenti, sapendo che il periodo di osservazione T è pari a 1 anno, ovvero 8760 ore:

$$LOLP = 0.0575\%$$

$$LOLE = 5.037 \text{ ore/anno}$$

$$EPNS = 1.140 \text{ MW}$$

$$EENS = 9986.4 \text{ MWh}$$

Successivamente si è effettuato il processo di convoluzione dopo aver incorporato le 9 windfarm nella rete test. In Fig. 34 è riportato l'andamento della *CCDF* che si è ottenuta in tale caso.

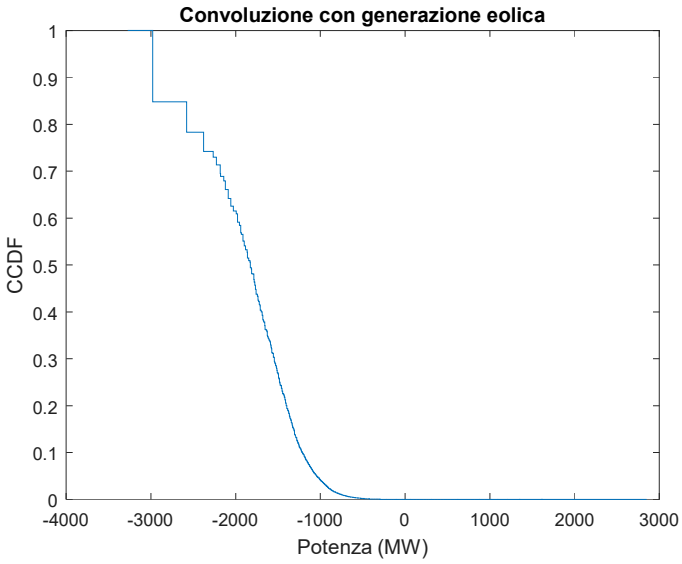

Figura 34-Convoluzione con generazione eolica.

L'andamento della *CCDF* è variato rispetto al caso precedente, in quanto, grazie all'aggiunta della generazione eolica, la capacità della rete è aumentata. Dai calcoli si ottengono i seguenti valori degli indici di adeguatezza:

$$LOLP = 0.0022\%$$

$$LOLE = 0.196 \text{ ore/anno}$$

$$EPNS = 0.823 \text{ MW}$$

$$EENS = 7209.48 \text{ MWh}$$

L'analisi degli indici di adeguatezza delle rete IEEE RTS mostra dunque che i valori degli indici diminuiscono significativamente con l'introduzione della generazione eolica. Si riporta inoltre in Fig. 35 in alto l'andamento di entrambe le *CCDF* ottenute al termine del processo di convoluzione, nel caso in cui non si considerano le windfarm aggiuntive e nel caso in cui si considerano: come si nota, la presenza della capacità di generazione eolica porta a spostare ulteriormente a sinistra la funzione. Il grafico sottostante mostra uno zoom del grafico precedente, in modo da mettere in evidenza i valori dei *LOLP* ottenuti nei due casi, in corrispondenza del livello zero di potenza.

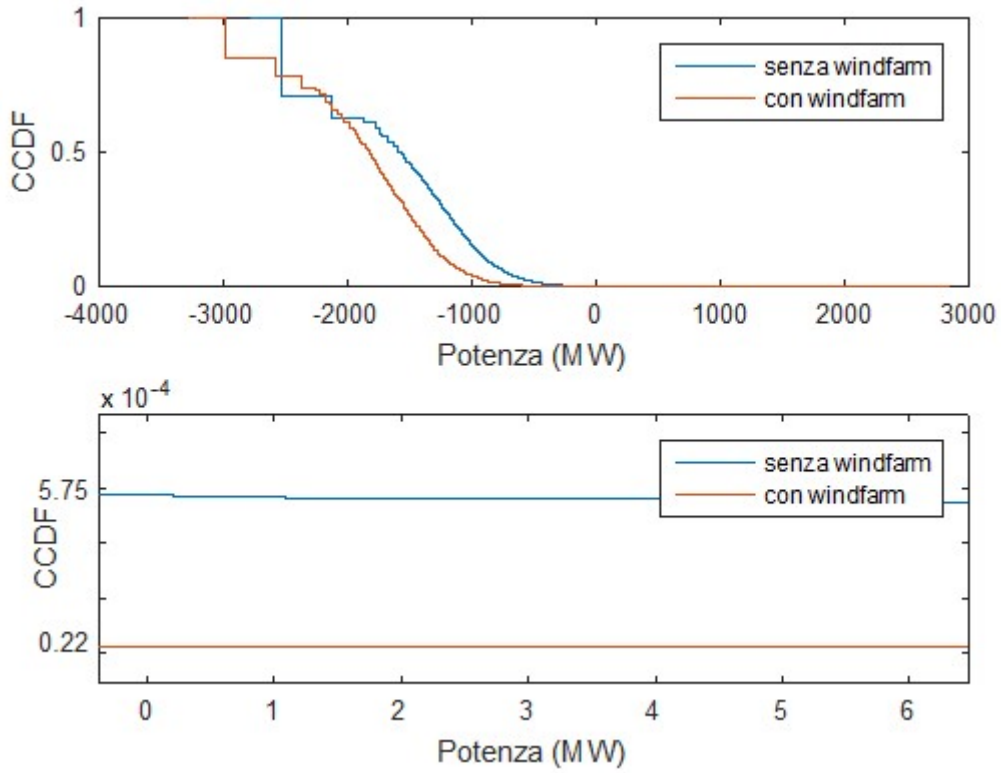

Figura 35-In alto, andamento delle *CCDF* ottenute al termine del processo di convoluzione, nel caso in cui non si sono incluse le windfarm aggiuntive e nel caso in cui si sono considerate. In basso, valori dei *LOLP* evidenziati nei due casi.

14.2-Analisi del modello statistico

Si riportano ora i risultati ottenuti dalla stima e dalla simulazione del modello statistico proposto, considerando ciascun stadio di entrambe le fasi. I dati di ingresso alla stima del modello sono le serie storiche delle velocità medie orarie del vento in un dato periodo; in questa trattazione si sono considerati 700 valori per tutte le storiche, ottenuti in un periodo corrispondente a circa un mese dell'anno. Per l'analisi dei risultati del modello si considera ora come esempio la serie s-380, il cui andamento dei valori del vento rispetto al tempo è mostrato in Fig. 36.

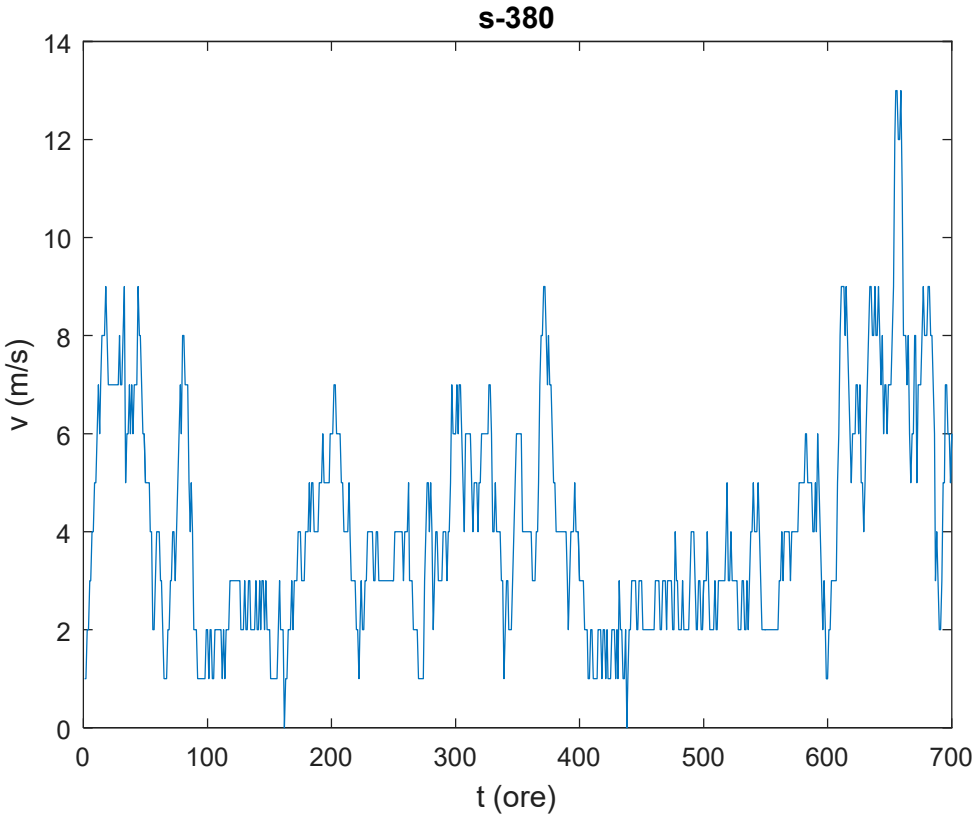

Figura 36-Andamento dei valori delle velocità medie orarie del vento della serie storica s-380.

Applicando il *KDE* a tale serie con MATLAB si è ottenuta una stima della funzione di densità di probabilità per la variabile aleatoria rappresentata in questo caso dalla velocità del vento della serie storica s-380, il cui andamento è riportato in Fig. 37. Utilizzando MATLAB, non risulta necessario integrare la *PDF* del vento per ricavare la rispettiva *CDF*, in quanto è possibile applicare il *KDE* alla serie storica modificato in tal caso per stimare direttamente la *CDF*. L'andamento della *CDF* riferita alla serie storica s-380 è mostrato in Fig. 38.

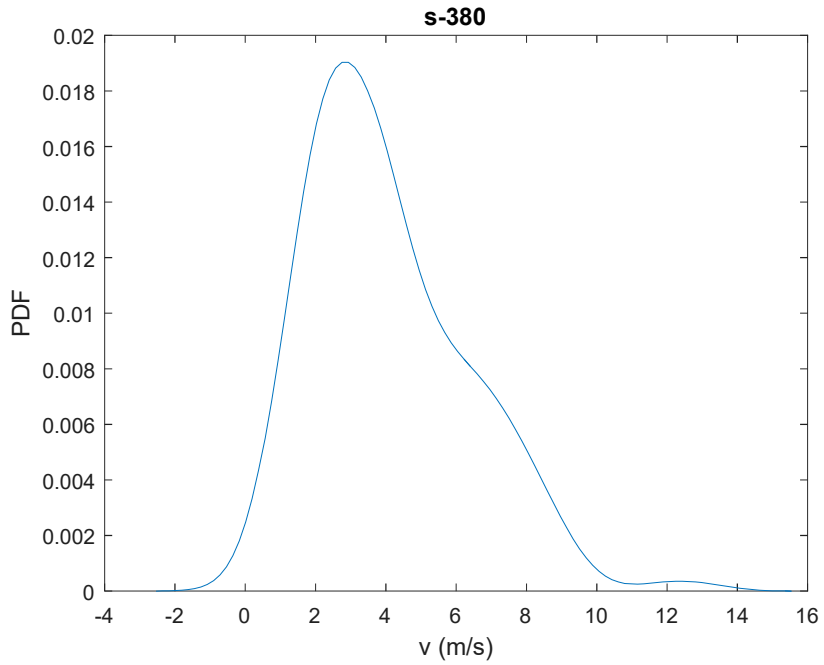

Figura 37-Andamento della PDF della serie storica s-380.

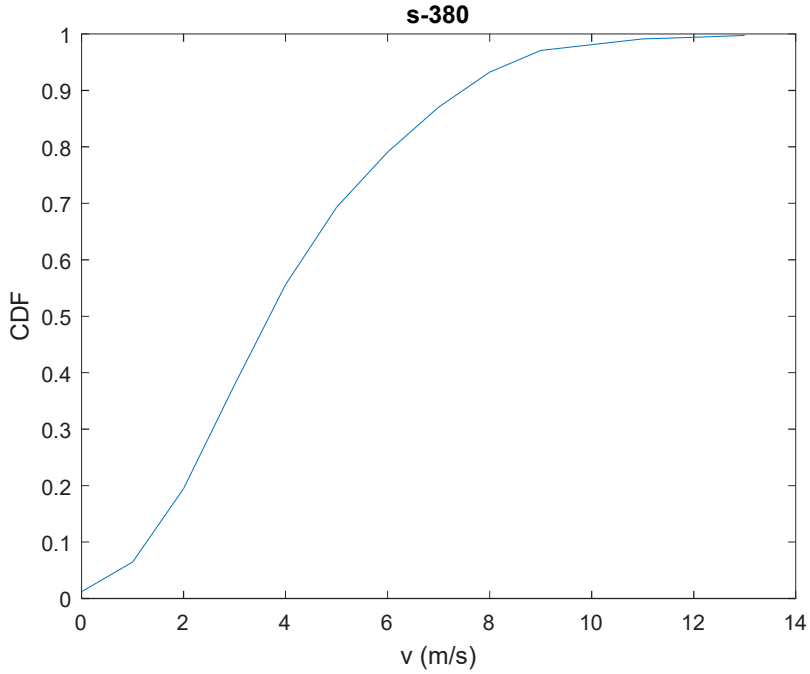

Figura 38-Andamento della CDF della serie storica s-380.

Con riferimento alle espressioni in (12.1), si è poi proceduto ad effettuare la trasformazione delle serie storiche in serie aventi distribuzione uniforme. Riferendosi alla s-380, si è ottenuta una nuova serie u_{s-38} i cui valori corrispondono di fatto a quelli della *CDF* della serie storica s-380. In Tabella 6, sono riportati i valori di tale serie uniforme, disposti per righe in ordine crescente.

Tabella 6-Valori della serie uniforme riferita alla serie storica s-380.

0.012	0.012	0.065	0.065	0.065	0.065	0.065	0.065	0.065	0.065	0.065	0.065	0.065	0.065
0.065	0.065	0.065	0.065	0.065	0.065	0.065	0.065	0.065	0.065	0.065	0.065	0.065	0.065
0.065	0.065	0.065	0.065	0.065	0.065	0.065	0.065	0.065	0.065	0.065	0.065	0.065	0.065
0.065	0.065	0.065	0.065	0.065	0.065	0.065	0.065	0.065	0.065	0.195	0.195	0.195	0.195
0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195
0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195
0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195
0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195
0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195
0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195
0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195
0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195
0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195
0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195	0.195
0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378
0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378
0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378
0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378
0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378
0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378
0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378
0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378
0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378
0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378	0.378
0.378	0.378	0.378	0.557	0.557	0.557	0.557	0.557	0.557	0.557	0.557	0.557	0.557	0.557
0.557	0.557	0.557	0.557	0.557	0.557	0.557	0.557	0.557	0.557	0.557	0.557	0.557	0.557
0.557	0.557	0.557	0.557	0.557	0.557	0.557	0.557	0.557	0.557	0.557	0.557	0.557	0.557
0.557	0.557	0.557	0.557	0.557	0.557	0.557	0.557	0.557	0.557	0.557	0.557	0.557	0.557
0.557	0.557	0.557	0.557	0.557	0.557	0.557	0.557	0.557	0.557	0.557	0.557	0.557	0.557
0.557	0.557	0.557	0.557	0.557	0.557	0.557	0.557	0.557	0.557	0.557	0.557	0.557	0.557
0.557	0.557	0.557	0.557	0.557	0.557	0.557	0.557	0.557	0.557	0.557	0.557	0.557	0.557
0.557	0.557	0.557	0.557	0.557	0.557	0.557	0.557	0.557	0.557	0.557	0.557	0.557	0.557
0.557	0.557	0.557	0.557	0.557	0.557	0.557	0.557	0.557	0.557	0.557	0.557	0.693	0.693
0.693	0.693	0.693	0.693	0.693	0.693	0.693	0.693	0.693	0.693	0.693	0.693	0.693	0.693
0.693	0.693	0.693	0.693	0.693	0.693	0.693	0.693	0.693	0.693	0.693	0.693	0.693	0.693
0.693	0.693	0.693	0.693	0.693	0.693	0.693	0.693	0.693	0.693	0.693	0.693	0.693	0.693
0.693	0.693	0.693	0.693	0.693	0.693	0.693	0.693	0.693	0.693	0.693	0.693	0.693	0.693
0.693	0.693	0.693	0.693	0.693	0.693	0.693	0.693	0.791	0.791	0.791	0.791	0.791	0.791
0.791	0.791	0.791	0.791	0.791	0.791	0.791	0.791	0.791	0.791	0.791	0.791	0.791	0.791
0.791	0.791	0.791	0.791	0.791	0.791	0.791	0.791	0.791	0.791	0.791	0.791	0.791	0.791
0.791	0.791	0.791	0.791	0.791	0.791	0.791	0.791	0.791	0.791	0.791	0.791	0.791	0.791
0.791	0.791	0.791	0.791	0.791	0.791	0.791	0.791	0.791	0.791	0.791	0.791	0.870	0.870
0.870	0.870	0.870	0.870	0.870	0.870	0.870	0.870	0.870	0.870	0.870	0.870	0.870	0.870
0.870	0.870	0.870	0.870	0.870	0.870	0.870	0.870	0.870	0.870	0.870	0.870	0.870	0.870
0.870	0.870	0.870	0.870	0.870	0.870	0.870	0.870	0.870	0.870	0.870	0.870	0.870	0.870
0.870	0.870	0.870	0.870	0.870	0.870	0.870	0.870	0.870	0.932	0.932	0.932	0.932	0.932
0.932	0.932	0.932	0.932	0.932	0.932	0.932	0.932	0.932	0.932	0.932	0.932	0.932	0.932
0.932	0.932	0.932	0.932	0.932	0.932	0.932	0.932	0.932	0.932	0.932	0.932	0.932	0.932
0.932	0.932	0.932	0.932	0.971	0.971	0.971	0.971	0.971	0.971	0.971	0.971	0.971	0.971
0.971	0.971	0.971	0.971	0.971	0.971	0.971	0.991	0.994	0.994	0.994	0.997	0.997	0.997

Il passo successivo consiste nella discretizzazione delle serie uniformi, effettuata in modo tale da massimizzare la mutua informazione condivisa tra le variabili aleatorie rappresentative delle serie storiche. Dai calcoli è risultato che la *MI* è massima discretizzando le serie uniformi in 550 bin. Si sono dunque ottenute le serie discrete $L_1, \dots, L_w$ con valori compresi tra 1 e 550. Si riporta in Tabella 7 i valori della serie

discreta L_{s-380}, ottenuta dalla serie uniforme u_{s-380}, disposti per righe in ordine crescente. I valori delle *MI*, calcolate su ogni coppia di serie discrete, sono mostrati in forma matriciale in Tabella 8: ogni elemento della tabella rappresenta il valore della mutua informazione tra la serie discreta L_i e la serie discreta L_j. Siccome la *MI* è simmetrica, è sufficiente calcolare i valori solo per la parte triangolare superiore o inferiore; inoltre non è necessario calcolare gli elementi sulla diagonale che rappresentano le auto informazioni.

Tabella 7-Valori della serie discreta riferita alla serie storica s-380.

1	1	30	30	30	30	30	30	30	30	30	30	30	30
30	30	30	30	30	30	30	30	30	30	30	30	30	30
30	30	30	30	30	30	30	30	30	30	30	30	30	30
30	30	30	30	30	30	30	30	30	30	102	102	102	102
102	102	102	102	102	102	102	102	102	102	102	102	102	102
102	102	102	102	102	102	102	102	102	102	102	102	102	102
102	102	102	102	102	102	102	102	102	102	102	102	102	102
102	102	102	102	102	102	102	102	102	102	102	102	102	102
102	102	102	102	102	102	102	102	102	102	102	102	102	102
102	102	102	102	102	102	102	102	102	102	102	102	102	102
102	102	102	102	102	102	102	102	102	102	102	102	102	102
102	102	102	102	102	102	102	102	102	102	102	102	102	102
102	102	102	102	102	102	102	102	102	102	102	102	102	102
102	102	102	102	102	102	102	102	102	102	102	102	102	102
205	205	205	205	205	205	205	205	205	205	205	205	205	205
205	205	205	205	205	205	205	205	205	205	205	205	205	205
205	205	205	205	205	205	205	205	205	205	205	205	205	205
205	205	205	205	205	205	205	205	205	205	205	205	205	205
205	205	205	205	205	205	205	205	205	205	205	205	205	205
205	205	205	205	205	205	205	205	205	205	205	205	205	205
205	205	205	205	205	205	205	205	205	205	205	205	205	205
205	205	205	205	205	205	205	205	205	205	205	205	205	205
205	205	205	205	205	205	205	205	205	205	205	205	205	205
205	205	205	205	205	205	205	205	205	205	205	205	205	205
205	205	205	304	304	304	304	304	304	304	304	304	304	304
304	304	304	304	304	304	304	304	304	304	304	304	304	304
304	304	304	304	304	304	304	304	304	304	304	304	304	304
304	304	304	304	304	304	304	304	304	304	304	304	304	304
304	304	304	304	304	304	304	304	304	304	304	304	304	304
304	304	304	304	304	304	304	304	304	304	304	304	304	304
304	304	304	304	304	304	304	304	304	304	304	304	304	304
304	304	304	304	304	304	304	304	304	304	304	304	304	304
304	304	304	304	304	304	304	304	304	304	304	304	380	380
380	380	380	380	380	380	380	380	380	380	380	380	380	380
380	380	380	380	380	380	380	380	380	380	380	380	380	380
380	380	380	380	380	380	380	380	380	380	380	380	380	380
380	380	380	380	380	380	380	380	380	380	380	380	380	380
380	380	380	380	380	380	380	380	435	435	435	435	435	435
435	435	435	435	435	435	435	435	435	435	435	435	435	435
435	435	435	435	435	435	435	435	435	435	435	435	435	435
435	435	435	435	435	435	435	435	435	435	435	435	435	435
435	435	435	435	435	435	435	435	435	435	435	435	480	480
480	480	480	480	480	480	480	480	480	480	480	480	480	480
480	480	480	480	480	480	480	480	480	480	480	480	480	480
480	480	480	480	480	480	480	480	480	480	480	480	480	480
480	480	480	480	480	480	480	480	480	514	514	514	514	514
514	514	514	514	514	514	514	514	514	514	514	514	514	514
514	514	514	514	514	514	514	514	514	514	514	514	514	514
514	514	514	514	536	536	536	536	536	536	536	536	536	536
536	536	536	536	536	536	536	547	549	549	549	550	550	550

Tabella 8-Matrice delle mutue informazioni

0	7.086	7.1163	7.1048	7.112	7.1098	7.1243	7.1006	7.1024
0	0	7.0935	7.0955	7.0798	7.1038	7.0857	7.1123	7.1064
0	0	0	7.1029	7.0831	7.1034	7.1119	7.1028	7.1063
0	0	0	0	7.0975	7.0619	7.109	7.1189	7.0976
0	0	0	0	0	7.115	7.1049	7.1097	7.1118
0	0	0	0	0	0	7.1083	7.1076	7.1073
0	0	0	0	0	0	0	7.1181	7.1243
0	0	0	0	0	0	0	0	7.1163
0	0	0	0	0	0	0	0	0

Successivamente si è effettuato il procedimento per ottenere la stima della rete Bayesiana tra le variabili aleatorie discrete rappresentate dalle serie discrete ottenute nello step precedente. A tale scopo si è utilizzato il Bayes Net Toolbox (BNT), un pacchetto open-source di funzioni di MATLAB per l'inferenza e il learning di modelli di reti bayesiane. Servendosi del BNT, si è effettuato dapprima il learning della struttura della rete Bayesiana, utilizzando l'algoritmo K2; definendo come dataset i valori delle serie discrete $L_1, L_2, \dots, L_w$, supponendo come ordine dei nodi della rete Bayesiana l'elenco dei nodi della rete IEEE RTS così come è stato definito in Tabella 5 e ammettendo al massimo 2 genitori per ogni nodo della rete Bayesiana, sì è ottenuto il grafo della rete, mostrato in Fig. 39. La corrispondenza tra i nodi della rete Bayesiana e le serie storiche è mostrata in Tabella 9: come si può notare dalla struttura del grafo, si ha una dipendenza statistica tra i nodi 1, 2 e 3, come ci si poteva aspettare in quanto le windfarm rappresentate da queste serie storiche si trovano in nodi della rete IEEE RTS vicini; tale terzetto di nodi della rete Bayesiana risulta scollegato dai restanti nodi, i cui corrispondenti nodi nella rete elettrica sono infatti distanti dai nodi 1,2,3 e pertanto si assume che non vi sia dipendenza stocastica tra il gruppo di nodi 1,2,3 e il restante. Si notano invece dipendenze statistiche proprio in quest'ultimo gruppo di nodi, dal nodo 4 al nodo 9, nell'ordine che si è considerato per applicare la stima; il nodo 5 è genitore sia del nodo 7 che del nodo 9.

Tabella 9-Corrispondenza tra nodi della rete Bayesiana e le serie storiche.

Nodo della rete Bayesiana	Serie storica
1	s-210
2	s-225
3	s-235
4	s-242
5	s-248
6	s-269
7	s-275
8	s-283
9	s-380

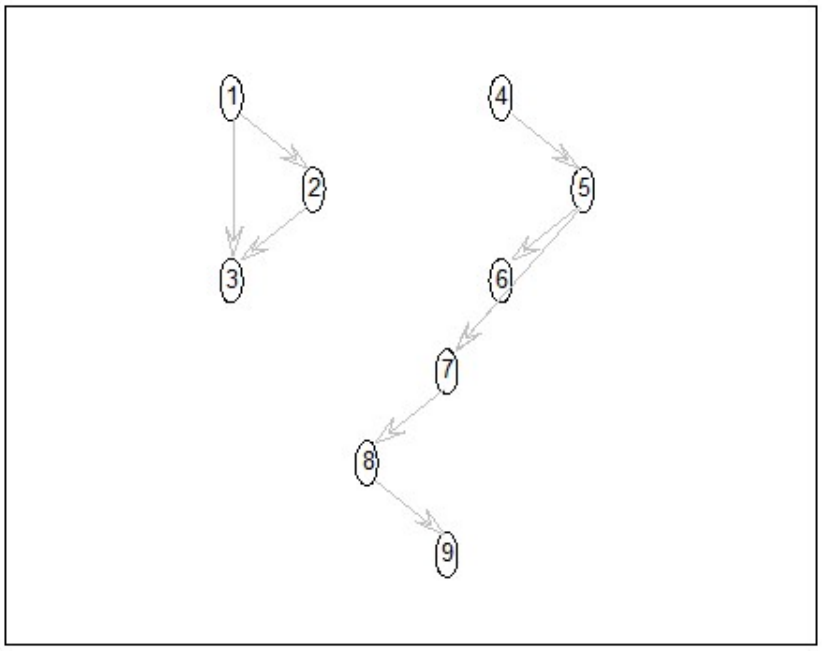

Figura 39-Grafo della rete bayesiana.

Per l'apprendimento dei parametri si è applicato, utilizzando ancora il BNT, il metodo della massima somiglianza. Con la stima della rete Bayesiana, si è completato il processo di stima del modello statistico e si è dunque passati ad effettuarne la simulazione: effettuando il campionamento multivariato dalla rete Bayesiana così ottenuta, si sono ricavate serie sintetiche di valori discreti uniformi $L_1^{Si}, L_2^{Si}, \dots, L_w^{Si}$. In Tabella 10 sono mostrati i valori della serie L_{s-380}^{Si} riferita alla serie s-380, disposti per righe nell'ordine in cui si sono ottenuti dal campionamento.

Lo step successivo consiste nella trasformazione delle variabili discrete in continue: dati i valori $L_1^{Si}, L_2^{Si}, \dots, L_w^{Si}$ ottenuti in precedenza si sono ottenute serie uniformi continue $u_1^{Si}, u_2^{Si}, \dots, u_w^{Si}$. In Tabella 11 sono riportati i valori della serie u_{s-380}^{Si} riferita alla serie s-380, disposti per righe nell'ordine in cui si sono ricavati.

Infine si è realizzata la trasformazione delle serie uniforme continue sintetiche in serie aventi la distribuzione delle serie storiche, utilizzando le formule in (12.3). Per fare ciò, si è applicato il *KDE* su MATLAB, in modo tale da stimare la *CDF* inversa delle serie storiche. Dati i valori $u_1^{Si}, u_2^{Si}, \dots, u_w^{Si}$ ricavati in precedenza, si sono dunque ottenuti serie sintetiche di valori rappresentativi delle velocità del vento. In Fig. 40 è mostrato l'andamento dei valori delle velocità del vento della serie sintetica riferita alla serie s-380.

Tabella 10-Valori della serie sintetica discreta riferita alla serie s-380.

82	10	16	267	191	6	255	110	394	410	433	481	62	205
297	267	7	198	290	183	94	510	284	436	495	362	443	311
231	33	471	381	516	77	94	259	449	258	162	473	86	67
57	281	218	265	222	525	430	114	210	506	512	72	302	509
543	466	110	209	353	318	72	12	202	312	204	433	350	490
350	333	89	140	527	442	402	8	453	520	432	202	69	549
436	343	179	205	8	302	267	438	159	70	310	142	516	209
160	12	40	519	45	56	344	33	285	540	238	18	38	279
383	261	410	235	265	127	540	131	81	205	331	170	126	11
259	520	385	486	253	444	451	395	416	48	19	419	484	453
181	159	293	430	176	546	207	46	247	278	233	490	276	8
520	71	35	369	9	259	387	430	186	110	315	456	380	448
420	218	92	390	358	392	282	141	307	260	225	367	291	26
427	177	313	142	254	179	505	55	402	103	103	270	546	11
110	216	70	56	34	361	400	313	295	282	338	291	466	170
126	11	338	431	398	167	83	1	87	152	491	88	416	173
165	395	134	281	398	113	69	274	234	247	182	234	36	311
38	470	18	34	282	266	483	313	410	315	537	395	238	378
415	391	460	294	480	181	259	456	479	265	154	104	546	548
355	125	410	38	267	318	123	247	142	6	456	477	426	472
432	167	321	155	402	501	210	441	34	472	416	168	154	16
469	435	543	234	7	48	331	173	241	539	234	2	409	330
69	34	512	103	525	255	168	471	342	19	442	77	531	168
8	377	237	60	353	272	405	453	387	535	188	210	72	472
258	167	142	34	48	195	209	303	520	255	478	5	300	169
278	466	319	267	46	81	361	11	216	48	176	395	127	272
152	8	204	255	247	191	110	114	415	60	472	202	127	318
410	343	409	133	6	158	10	42	479	512	202	438	413	26
365	200	378	479	162	436	198	436	183	385	480	430	477	341
535	34	481	381	437	98	11	256	540	272	243	492	175	430
270	96	397	104	481	383	140	315	521	21	481	491	105	207
258	381	159	167	350	491	471	365	11	225	369	407	471	260
331	151	518	135	436	72	493	536	267	405	257	240	394	250
34	472	104	154	362	168	278	284	228	534	2	87	8	519
494	7	183	430	212	182	313	437	456	327	440	131	396	186
262	255	60	261	388	487	471	48	449	476	48	534	425	104
38	357	189	353	211	86	6	423	506	448	449	501	222	306
247	226	392	535	28	204	315	395	364	392	6	487	258	105
126	386	407	348	469	306	282	126	317	340	175	10	343	113
267	311	441	269	523	85	516	397	94	381	246	98	207	141
520	148	477	535	321	523	512	377	450	501	526	221	453	172
285	176	365	153	409	113	221	460	267	12	477	191	56	11
402	75	516	340	167	281	315	531	64	357	255	436	313	207
284	207	449	460	460	204	340	344	279	240	369	198	272	387
270	429	302	35	273	340	177	11	398	160	450	455	259	495
47	293	160	348	101	491	464	480	72	110	212	470	397	543
77	372	218	469	57	282	205	252	264	405	344	473	340	353
34	271	533	493	370	526	260	392	321	69	279	272	278	460
493	402	280	103	282	346	5	210	313	449	443	338	261	272
435	350	210	87	470	40	511	246	431	320	146	338	416	417

Tabella 11-Valori della serie sintetica uniforme riferita alla serie s-380.

0.158	0.029	0.039	0.490	0.352	0.021	0.468	0.207	0.716	0.746	0.787	0.872	0.123	0.379
0.543	0.489	0.024	0.366	0.531	0.339	0.180	0.925	0.520	0.792	0.898	0.659	0.805	0.568
0.425	0.070	0.855	0.693	0.935	0.149	0.180	0.476	0.816	0.473	0.301	0.859	0.165	0.131
0.112	0.515	0.401	0.486	0.408	0.951	0.781	0.215	0.388	0.917	0.928	0.139	0.552	0.923
0.984	0.846	0.207	0.386	0.644	0.581	0.141	0.032	0.373	0.570	0.376	0.786	0.638	0.890
0.638	0.608	0.170	0.262	0.955	0.802	0.731	0.025	0.823	0.943	0.784	0.374	0.135	0.994
0.792	0.625	0.332	0.378	0.025	0.552	0.489	0.797	0.296	0.137	0.567	0.266	0.936	0.385
0.297	0.032	0.083	0.942	0.091	0.111	0.627	0.070	0.522	0.979	0.438	0.042	0.078	0.512
0.697	0.479	0.746	0.433	0.486	0.239	0.979	0.245	0.157	0.378	0.603	0.315	0.236	0.031
0.475	0.942	0.700	0.882	0.464	0.806	0.819	0.718	0.756	0.097	0.045	0.762	0.877	0.823
0.335	0.296	0.535	0.781	0.326	0.990	0.381	0.093	0.453	0.509	0.428	0.889	0.505	0.024
0.942	0.137	0.073	0.673	0.027	0.476	0.705	0.781	0.345	0.208	0.576	0.829	0.692	0.814
0.763	0.402	0.176	0.710	0.652	0.714	0.516	0.264	0.560	0.477	0.413	0.668	0.532	0.058
0.777	0.329	0.571	0.265	0.467	0.331	0.915	0.109	0.731	0.196	0.195	0.495	0.990	0.030
0.208	0.398	0.136	0.111	0.071	0.658	0.727	0.572	0.540	0.516	0.617	0.533	0.847	0.316
0.237	0.031	0.616	0.783	0.723	0.310	0.160	0.013	0.166	0.283	0.890	0.169	0.756	0.321
0.307	0.718	0.251	0.514	0.724	0.213	0.135	0.503	0.431	0.453	0.337	0.430	0.075	0.569
0.080	0.853	0.043	0.072	0.516	0.488	0.876	0.572	0.746	0.575	0.973	0.719	0.437	0.689
0.754	0.711	0.836	0.537	0.871	0.335	0.474	0.828	0.869	0.485	0.286	0.197	0.989	0.992
0.647	0.235	0.745	0.078	0.489	0.581	0.231	0.453	0.265	0.022	0.828	0.866	0.775	0.857
0.785	0.310	0.586	0.289	0.732	0.909	0.387	0.800	0.073	0.857	0.756	0.313	0.286	0.040
0.851	0.790	0.985	0.430	0.023	0.097	0.604	0.321	0.442	0.977	0.430	0.015	0.744	0.603
0.134	0.073	0.929	0.196	0.951	0.468	0.311	0.856	0.624	0.045	0.803	0.149	0.963	0.312
0.025	0.686	0.435	0.119	0.643	0.498	0.737	0.822	0.705	0.969	0.347	0.387	0.141	0.857
0.473	0.310	0.266	0.071	0.097	0.360	0.385	0.553	0.943	0.468	0.867	0.021	0.548	0.314
0.509	0.846	0.582	0.490	0.094	0.156	0.658	0.030	0.399	0.097	0.326	0.720	0.239	0.498
0.283	0.025	0.377	0.468	0.454	0.352	0.209	0.216	0.754	0.119	0.856	0.374	0.238	0.581
0.745	0.625	0.744	0.249	0.021	0.294	0.029	0.086	0.870	0.928	0.373	0.797	0.752	0.058
0.666	0.370	0.688	0.869	0.302	0.792	0.366	0.791	0.339	0.702	0.872	0.782	0.866	0.622
0.970	0.072	0.874	0.693	0.793	0.186	0.030	0.469	0.979	0.498	0.446	0.893	0.325	0.782
0.494	0.183	0.723	0.198	0.873	0.697	0.262	0.576	0.944	0.049	0.872	0.891	0.200	0.382
0.473	0.694	0.295	0.310	0.638	0.890	0.856	0.664	0.030	0.414	0.673	0.739	0.855	0.477
0.604	0.281	0.940	0.253	0.793	0.140	0.895	0.972	0.488	0.737	0.472	0.441	0.717	0.459
0.072	0.857	0.197	0.286	0.659	0.312	0.510	0.521	0.419	0.968	0.015	0.166	0.026	0.942
0.895	0.023	0.340	0.782	0.391	0.336	0.572	0.794	0.829	0.596	0.799	0.245	0.720	0.344
0.480	0.468	0.119	0.479	0.706	0.883	0.855	0.096	0.815	0.865	0.097	0.967	0.773	0.197
0.079	0.651	0.350	0.644	0.390	0.165	0.022	0.769	0.918	0.814	0.816	0.908	0.409	0.559
0.453	0.416	0.713	0.970	0.062	0.377	0.576	0.719	0.662	0.713	0.021	0.884	0.473	0.198
0.237	0.703	0.740	0.635	0.850	0.560	0.516	0.237	0.580	0.619	0.324	0.028	0.626	0.213
0.489	0.569	0.801	0.492	0.948	0.164	0.935	0.722	0.180	0.694	0.451	0.186	0.381	0.263
0.942	0.276	0.866	0.970	0.587	0.948	0.928	0.686	0.816	0.908	0.953	0.407	0.823	0.318
0.522	0.326	0.666	0.285	0.744	0.214	0.408	0.835	0.489	0.032	0.866	0.354	0.112	0.030
0.731	0.146	0.936	0.620	0.309	0.514	0.575	0.962	0.125	0.651	0.467	0.792	0.571	0.381
0.520	0.382	0.816	0.835	0.835	0.376	0.620	0.627	0.511	0.442	0.671	0.365	0.498	0.705
0.495	0.779	0.552	0.073	0.501	0.619	0.329	0.031	0.724	0.297	0.817	0.827	0.475	0.897
0.096	0.536	0.297	0.635	0.192	0.891	0.843	0.871	0.140	0.208	0.390	0.854	0.721	0.985
0.150	0.678	0.401	0.851	0.113	0.517	0.379	0.462	0.484	0.737	0.628	0.858	0.621	0.643
0.072	0.497	0.967	0.894	0.673	0.954	0.477	0.713	0.585	0.135	0.510	0.499	0.510	0.834
0.894	0.732	0.512	0.195	0.516	0.631	0.020	0.388	0.571	0.815	0.806	0.616	0.478	0.499
0.790	0.638	0.388	0.166	0.853	0.083	0.926	0.451	0.782	0.584	0.272	0.617	0.756	0.758

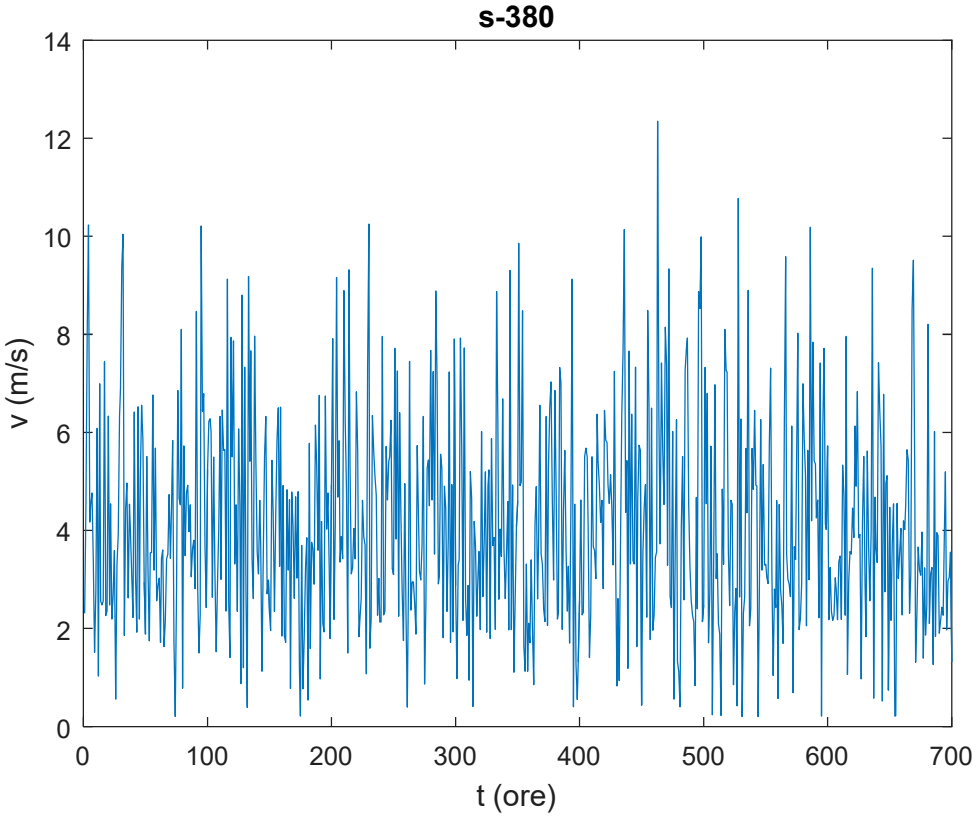

Figura 40-Andamento delle velocità del vento ricavate dalla serie sintetica riferita alla serie s-380.

Da questa serie sintetica si sono potute stimare la *PDF* e la *CDF*, i cui andamenti sono mostrati rispettivamente in Fig. 41 e 42.

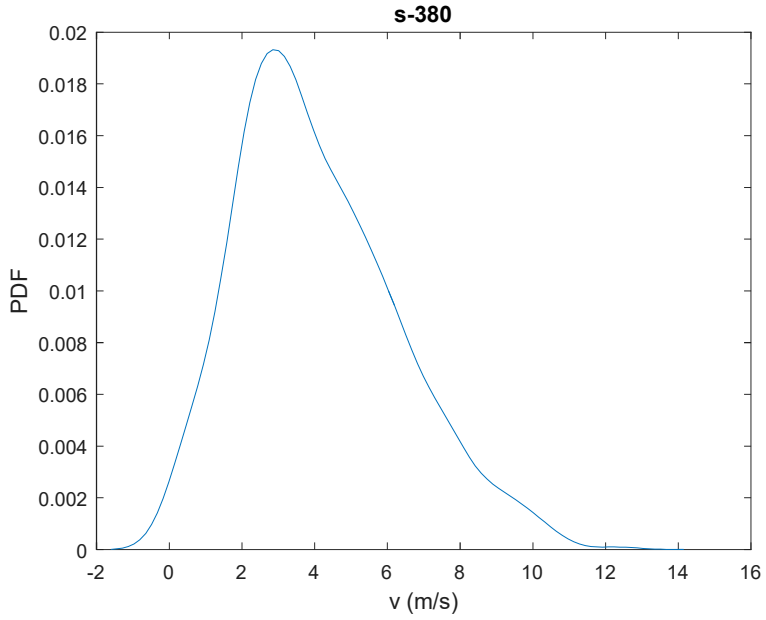

Figura 41-Andamento della PDF della serie sintetica riferita alla serie s-380.

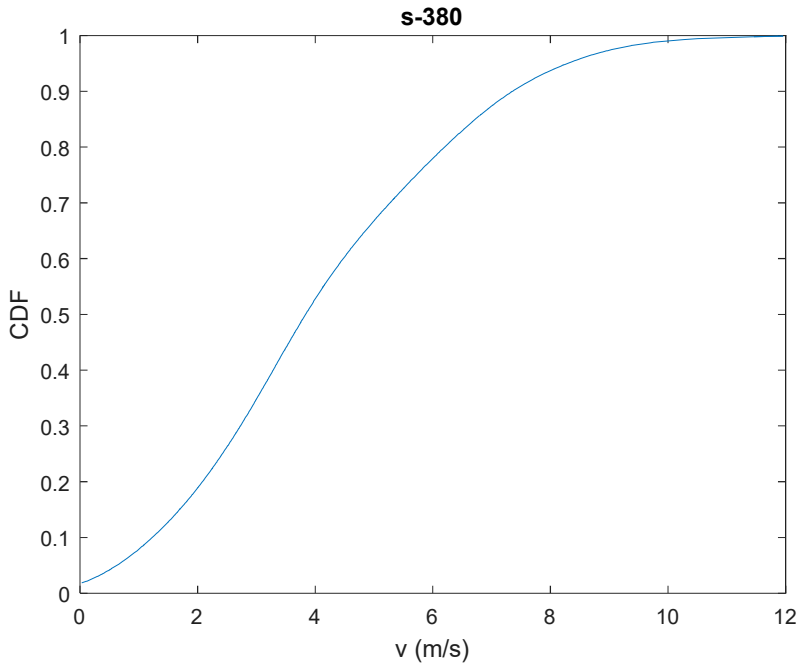

Figura 42-Andamento della CDF della serie sintetica riferita alla serie s-380.

Dai risultati del modello, un aspetto importante da analizzare è dunque la qualità dei campioni delle variabili casuali generate dal modello tenendo conto della conservazione delle caratteristiche statistiche delle serie storiche. A titolo di esempio, le Figure 43 e 44 mostrano le funzioni di densità di probabilità per le variabili casuali corrispondenti alla serie di velocità del vento s-225, s-235, s-242 e s-248 e confronta le serie storiche e sintetiche ottenute dal modello.

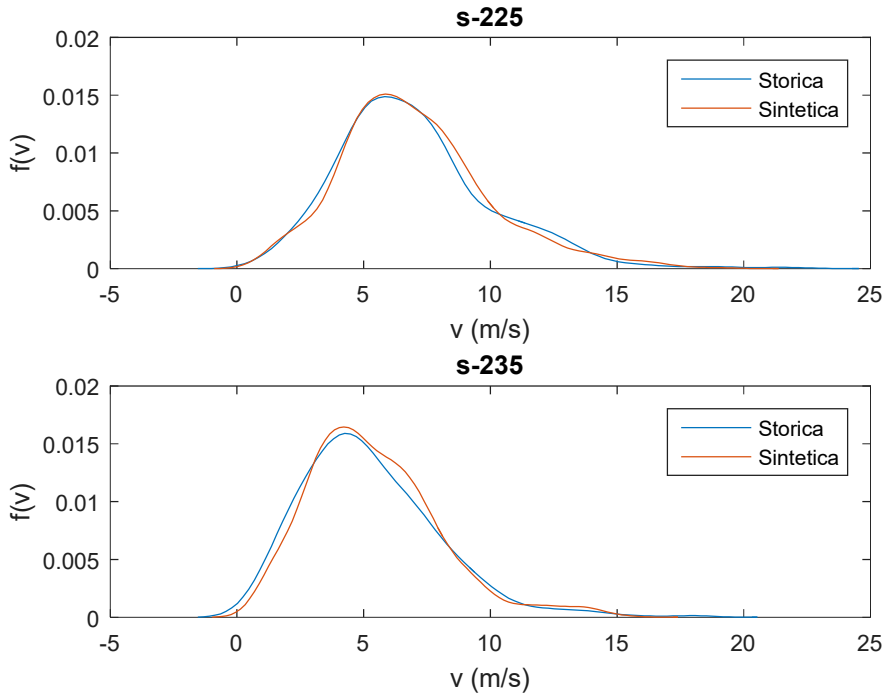

Figura 43-Funzioni di densità di probabilità ottenute dalle serie storiche e sintetiche delle serie s-225 e s-235.

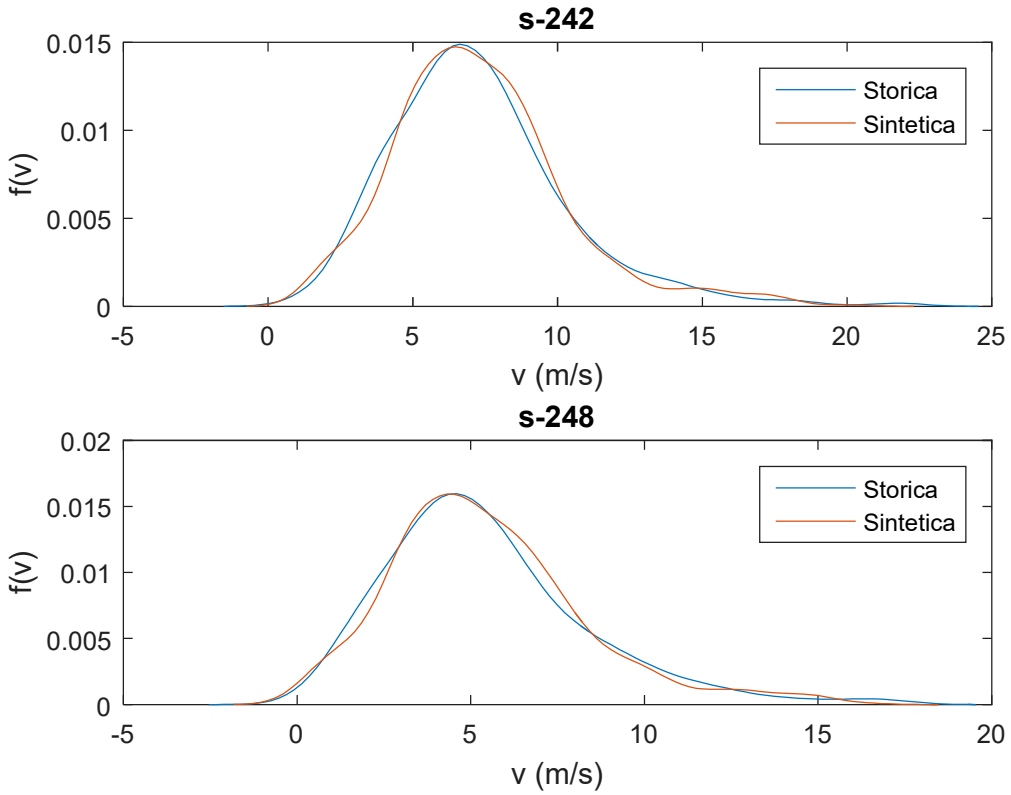

Figura 44-Funzioni di densità di probabilità ottenute dalle serie storiche e sintetiche delle serie s-242 e s-248.

Come si può vedere, le funzioni di densità di probabilità delle serie sintetiche sono molto simili a quelle delle serie storiche; si nota però come le funzioni delle serie sintetiche siano leggermente sottostimate sul fronte di salita, mentre sono lievemente sovrastimate sul fronte di discesa; tali errori risultano comunque trascurabili nell'andamento globale delle curve.

Per un'analisi quantitativa, le densità delle serie sintetiche possono essere confrontate con i dati storici applicando il test statistico di Mann-Whitney-Wilcoxon [Hollander et al., 2014], che viene utilizzato per verificare l'ipotesi di due gruppi di campioni appartenenti alla stessa distribuzione, con un certo livello di significatività. I valori *p* per le quattro serie analizzate in precedenza sono mostrati nella Tabella 12. I valori *p* ottenuti superano il livello di significatività α di 0.01 per tutte le serie e solo una serie, s-235, è al di sotto del livello di significatività di 0.05. Quindi, quantitativamente, si può affermare che le densità della serie sintetiche e storiche sono compatibili tra di loro.

Tabella 12-Valori p ottenuti dal test di Mann-Whitney-Wilcoxon.

Serie	Valore *p*
s-225	0.12
s-235	0.026
s-242	0.056
s-248	0.19

15-Conclusioni

L'affidabilità dei sistemi di generazione è un aspetto importante nella pianificazione per il futuro sviluppo dei sistemi elettrici. Le metodologie di valutazione dell'affidabilità per i sistemi di generazione convenzionali che hanno coinvolto grandi e affidabili unità basate su combustibili fossili sono state sviluppate e utilizzate per molti decenni. La maggior parte delle utility del mondo hanno operato con successo con l'uso di questi metodi. In seguito allo sviluppo ed alla crescente espansione dei sistemi di generazione basati su fonti rinnovabili, la natura intermittente di tali risorse ha avuto un impatto significativo sulla valutazione dell'affidabilità del sistema elettrico. Quindi, sviluppare modelli appropriati in modo da includere capacità di generazione, quali l'eolico o il fotovoltaico, nella valutazione generale dell'adeguatezza, è un notevole interesse che molto spesso è presente in attività di ricerca scientifica.

Lo scopo di questa tesi è stato mostrare un modello statistico per la rappresentazione di elementi tempo-varianti che può essere applicato per la valutazione dell'affidabilità impiegando il metodo Monte Carlo non sequenziale. Il modello è stato sviluppato applicando un metodo di stima non parametrica di funzioni di densità di probabilità di variabili continue, il Kernel Density Estimation; per misurare le dipendenze statistiche non lineari tra le serie tempo-varianti si è utilizzata la mutua informazione, una grandezza derivante dalla teoria dell'informazione. Si è inoltre definita una rete Bayesiana in modo da fornire una rappresentazione grafica delle probabilità condizionate tra gli elementi.

Per valutare l'accuratezza, l'efficienza e la flessibilità della rappresentazione proposta, sono state eseguite simulazioni sul sistema IEEE RTS, modificato per incorporare 9 windfarm, tutte soggette a differenti serie del vento. Queste simulazioni hanno dimostrato che il modello può riprodurre accuratamente le proprietà statistiche delle serie storiche e conservare l'influenza di qualsiasi tipo di correlazione tra le serie.

Il modello può essere applicato direttamente anche, ad esempio, per rappresentare le correlazioni tra la generazione di energia eolica e le portate affluenti in un impianto idroelettrico, o rappresentare qualsiasi altro tipo di generazione intermittente di energia elettrica o di carico variabile nel tempo, su sistemi di grandi dimensioni.

Appendice A

IEEE Reliability Test System

L'IEEE Reliability Test System (RTS) è stato sviluppato dal sottocomitato per l'applicazione dei metodi di probabilità nella IEEE Power Engineering Society Questo sistema è costituito da 24 nodi collegati da 38 linee.

Dati del sistema di trasmissione

Il sistema di trasmissione è a due livelli di tensione, 138 kV e 230 kV. L'area a 230 kV è nella parte superiore della rete, con sottostazioni 230/138 kV presenti ai nodi 11, 12 e 24. Le posizioni delle unità di generazione sono elencate nella Tabella 13. Nel sistema sono presenti dispositivi di correzione della tensione al nodo 14 (condensatore sincrono) e al nodo 6 (reattore di potenza). La Tabella 14 mostra i limiti reattivi in Mvar delle unità di generazione del sistema. I dati relativi alle lunghezze della linee di trasmissione, dei tassi di guasto delle linee in regime permanente e transitorio e del tempo medio di durata dei guasti in regime permanente sono riportati nella Tabella 15. I dati del carico nei nodi negli istanti in cui si verificano i picchi sono mostrati nella Tabella 16.

Tabella 13-Posizioni delle unità di generazione nei nodi della rete.

Nodo	Unità 1	Unità 2	Unità 3	Unità 4	Unità 5	Unità 6
	MW	MW	MW	MW	MW	MW
1	20	20	76	76		
2	20	20	76	76		
7	100	100	100			
13	197	197	197			
15	12	12	12	12	12	155
16	155					
18	400					
21	400					
22	50	50	50	50	50	50
23	155	155	350			

Tabella 14-Limiti reattivi delle unità di generazione.

Taglia	Mvar	
MW	Minimi	Massimi
12	0	6
20	0	10
50	-10	16
76	-25	30
100	0	60
155	-50	80
197	0	80
350	-25	150
400	-50	200

Tabella 15-Dati delle linee di trasmissione.

			Permanente		Transitorio
Da	A	Lunghezza	Tasso di guasto	Durata di guasto	Tasso di guasto
Nodo	Nodo	miglia	1/anno	ore	1/anno
1	2	3	0.24	16	0.0
1	3	55	0.51	10	2.9
1	5	22	0.33	10	1.2
2	4	33	0.39	10	1.7
2	6	50	0.48	10	2.6
3	9	31	0.38	10	1.6

			Permanente		Transitorio
Da	A	Lunghezza	Tasso di guasto	Durata di guasto	Tasso di guasto
Nodo	Nodo	miglia	1/anno	ore	1/anno
3	24	0	0.02	768	0.0
4	9	27	0.36	10	1.4
5	10	23	0.34	10	1.2
6	10	16	0.33	35	0.0
7	8	16	0.30	10	0.8
8	9	43	0.44	10	2.3
8	10	43	0.44	10	2.3
9	11	0	0.02	768	0.0
9	12	0	0.02	768	0.0
10	11	0	0.02	768	0.0
10	12	0	0.02	768	0.0
11	13	33	0.40	11	0.8
11	14	29	0.39	11	0.7
12	13	33	0.40	11	0.8
12	23	67	0.52	11	1.6
13	23	60	0.49	11	1.5
14	16	27	0.38	11	0.7
15	16	12	0.33	11	0.3
15	21	34	0.41	11	0.8
15	21	34	0.41	11	0.8
15	24	36	0.41	11	0.9
16	17	18	0.35	11	0.4
16	19	16	0.34	11	0.4
17	18	10	0.32	11	0.2
17	22	73	0.54	11	1.8
18	21	18	0.35	11	0.4
18	21	18	0.35	11	0.4
19	20	27.5	0.38	11	0.7
19	20	27.5	0.38	11	0.7
20	23	15	0.34	11	0.4
20	23	15	0.34	11	0.4
21	22	47	0.45	11	1.2

Tabella 16-Dati dei carichi nei nodi.

	Carico		Carico nel nodo
Nodo	MW	Mvar	% del carico di sistema
1	108	22	3.8
2	97	20	3.4
3	180	37	6.3
4	74	15	2.6
5	71	14	2.5
6	136	28	4.8
7	125	25	4.4
8	171	35	6
9	175	36	6.1
10	195	40	6.8
13	265	54	9.3
14	194	39	6.8
15	317	64	11.1
16	100	20	3.5
18	333	68	11.7
19	181	37	6.4
20	128	26	4.5
Totale	2850	580	100

Dati del carico

Il carico annuale di picco per questo sistema è di 2850 MW. La Tabella 17 presenta i carichi di picco settimanali in percentuale del carico di picco annuale. Si assume che la settimana 1 sia il dato per la prima settimana di gennaio e che sia abbia un carico di picco invernale per questo sistema. La Tabella 18 elenca il carico di picco giornaliero su un ciclo di 7 giorni, in percentuale del valore di carico di picco settimanale. In generale, si assume che avvenga lo stesso ciclo di carico di 7 giorni per tutto l'anno. I dati nelle Tabelle 17 e 18 insieme al carico di picco annuale descrivono un profilo cronologico giornaliero in 52 × 7 = 364 giorni, con il lunedì assunto come il primo giorno dell'anno. La Tabella 19 fornisce i dati del carico orario nei giorni feriali e nel fine settimana per diverse stagioni. Il modello di carico annuale in 364 × 24 = 8736 ore per il sistema è costituito combinando la Tabella 17, 18 e 19 con il valore di carico di picco annuale di 2850 MW.

Tabella 17-Carico di picco settimanale, in percentuale del carico di picco annuale.

Settimana	Carico di picco	Settimana	Carico di picco
1	86.2	27	75.5
2	90	28	81.6
3	87.8	29	80.1
4	83.4	30	88
5	88	31	72.2
6	84.1	32	77.6
7	83.2	33	80
8	80.6	34	72.9
9	74	35	72.6
10	73.7	36	70.5
11	71.5	37	78
12	72.7	38	69.5
13	70.4	39	72.4
14	75	40	72.4
15	72.1	41	74.3
16	80	42	74.4
17	75.4	43	80
18	83.7	44	88.1
19	87	45	88.5
20	88	46	90.9
21	85.6	47	94
22	81.1	48	89
23	90	49	94.2
24	88.7	50	97
25	89.6	51	100
26	86.1	52	95.2

Tabella 18-Carico di picco giornaliero, in percentuale del carico di picco settimanale.

Giorno	Carico di picco
Lunedì	93
Martedì	100
Mercoledì	98
Giovedì	96
Venerdì	94
Sabato	77
Domenica	75

Tabella 19- Carico di picco orario, in percentuale del carico di picco giornaliero.

	Settimane		Settimane		Settimane	
	invernali		estive		Primavera/autunno	
	1-8 & 44-52		18-30		9-17 & 31-43	
Ora	Wkdy	Wknd	Wkdy	Wknd	Wkdy	Wknd
00-01	67	78	64	74	63	75
01-02	63	72	60	70	62	73
02-03	60	68	58	66	60	69
03-04	59	66	56	65	58	66
04-05	59	64	56	64	59	65
05-06	60	65	58	62	65	65
06-07	74	66	64	62	72	68
07-08	86	70	76	66	85	74
08-09	95	80	87	81	95	83
09-10	96	88	95	86	99	89
10-11	96	90	99	91	100	92
11-12	95	91	100	93	99	94
12-13	95	90	99	93	93	91
13-14	95	88	100	92	92	90
14-15	93	87	100	91	90	90
15-16	94	87	97	91	88	86
16-17	99	91	96	92	90	85
17-18	100	100	96	94	92	88
18-19	100	99	93	95	96	92
19-20	96	97	92	95	98	100
20-21	91	94	92	100	96	97
21-22	83	92	93	93	90	95
22-23	73	87	87	88	80	90
23-24	63	81	72	80	70	85
Wkdy = giorni feriali, Wknd = fine settimana						

Dati dei generatori

In Tabella 20 sono riportati i dati dei generatori per quanto riguarda le potenze nominali e il numero di unità, assieme ai dati di affidabilità.

Tabella 20-Dati delle unità di generazione.

Taglia dell'unità	Numero di unità	*FOR*	*MTTF*	*MTTR*	Manutenzioni programmate
MW			ore	ore	settimane/anno
12	5	0.02	2940	60	2
20	4	0.1	450	50	2
50	6	0.01	1980	20	2
76	4	0.02	1960	40	3
100	3	0.04	1200	50	3
155	4	0.04	960	40	4
197	3	0.05	950	50	4
350	1	0.08	1150	100	5
400	2	0.12	1100	150	6

Bibliografia

R. Billinton e R. N. Allan, "Reliability Evaluation of Power Systems, Plenum Press, 1996.

R. Nicolosi, "Valutazione di affidabilità e adeguatezza per la pianificazione di sistemi di distribuzione multi-microrete. Dispacciamento ottimo delle risorse nell'esercizio di microreti autonome", Tesi di Dottorato, Università di Catania, 2011.

R. N. Allan e R. Billinton, "Probabilistic Assessment of Power Systems", Proceedings of the IEEE, Vol.88 (2), February 2000.

R. Billinton e K. E. Bollinger, "Trasmission System Reliability Evaluation Using Markov Processes", *IEEE Transactions on Power Apparatus and Systems*, Vol.87 (2), pp. 538-547, February 1968.

E. Carpaneto e G. Chicco, "Evaluation of the Probability Density Functions of Distribution System Reliability Indices With a Characteristic Functions-Based Approach", *IEEE Transactions on Power Systems*, Vol. 19 (2), pp. 724-734, May 2004.

H. Kahn e T. E. Harris, "Estimation of particle transmission by random sampling", *National Bureau of Standards Applied Mathematics*, pp. 27-30, 1951.

R. Billinton e W. Li, "Hybrid Approach For Reliability Evaluation of Composite Generation and Transmission Systems Using Monte Carlo Simulation and Enumeration Technique", *IEEE Proceedings-C*, Vol. 138 (3), pp. 233-241, May 1991.

M. V. F. Pereira, M. E. P. Maceira, G. C. Oliveira e L. M. V. G. Pinto, "Combining Analytical Models and Monte Carlo Techniques in Probabilistic Power System Analysis", *IEEE Transactions on Power Systems*, Vol. 7 (1), pp. 265-272, February 1992.

R. Billinton e W. Li, "A System State Transition Sampling Method for Composite System Reliability Evaluation", *IEEE Transactions on Power Systems*, Vol. 8 (3), pp. 761-771, August 1993.

R. Billinton e A. Sankarakrishnan, "A System State Transition Sampling Technique for Reliability Evaluation", *Reliability Engineering and System Safety*, Vol. 44, pp. 131-134, 1994.

R. Billinton e W. Li, "Reliability Assessment of Electrical Power Systems UsingMonte Carlo Methods", Plenum Publishing, 1994.

R. Y. Rubinstein, "Simulation and the Monte Carlo Methods", Wiley, 1981.

R. Ubeda e R. N. Allan, "Sequential Simulation Applied to Composite System Reliability Evaluation", *IEEE Proceedings-C*, Vol. 139 (2), March 1992.

M. V. F. Pereira e L. M. V. G. Pinto, "A New Computational Tool for Composite Reliability Evaluation", *IEEE Transactions on Power Systems*, Vol. 7 (1), pp. 258-264, February 1992.

K. W. Hipel e A. I. McLeod, "Time Series Modelling of Water Resources and Environmental System", Elsevier, Amsterdam, 1994.

R. Billinton, Y. Gao e R. Karki, "Composite System Adequacy Assessment Incorporating Large-Scale Wind Energy Conversion Systems Considering Wind Speed Correlation", *IEEE Transactions on Power Systems*, Vol. 24 (3), pp. 1374-1382, August 2009.

C. L. T. Borges e J. A. S. Dias, "A Model to Represent Correlated Time Series in Reliability Evaluation by Non-Sequential Monte Carlo Simulation", *IEEE Transactions on Power Systems*, 2016, DOI: 10.1109/TPWRS.2016.2585619.

C. L. T. Borges e J. A. S. Dias, "A Non Parametric Stochastic Model for River Inflows Based on Kernel Density Estimation", *Proceedings of 13th International Conference on Probabilistic Methods Applied to Power Systems*, pp. 100-107, Durham, June 2014.

S. J. Sheather and M. C. Jones, "A reliable data-based bandwidth selection method for kernel density estimation", *Journal of the Royal Statistical Society, Series B*, Vol. 53, 683–690, 1991.

A. Kraskov, H. Stogbauer e P. Grassberguer, "Estimating Mutual Information", *Physical Review E*, Vol 69 (6), pp. 66-138, 2004.

D. Geiger e J. Pearl, "On the logic of influence diagrams", *Proceedings of 4th workshop on uncertainty in AI*, pp. 136-147, Minneapolis, 1988.

G. F. Cooper, E. H. Herskovits, "A Bayesian Method for the Induction of Probabilistic Networks from Data", *Machine Learning*, Vol. 9 (4), pp. 309-347, 1992.

R.A. Fisher, "On the mathematical foundations of theorical statistics", Philosophical Transactions of the Royal Society of London Series A,Vol. 222, pp. 309-368, 1922.

L. R. Ford e D. R. Fulkerson, "Flow in Networks", Princeton University Press, Princeton, New Jersey, 1962.

R. Billinton and R. N. Allan, "Reliability Assessment of Large Electric Power Systems", Kluwer Academic Publishers, 1988.

B. Stott e O. Alsac, "Fast Decoupled Load Flow", *IEEE Transactions on Power Apparatus and Systems*, Vol. PAS-93, pp. 859-869, May 1974.

B. Stott, "Review of Load Flow Calculation Methods", *Proceedings of the IEEE*, Vol. 62, pp. 916-929, July 1974.

R. Billinton e S. Kumar, "Adequacy Evaluation of a Composite Power System – A Comparative Study for Existing Programs", *CEA Transactions, Engineering and Operating Division*, Vol. 24, Part 3, Paper No. 85-SP-141, pp. 1-14, March 1985.

M. Hollander, D. A. Wolfe, E. Chicken, "Nonparametric Statistical Methods", John Wiley & Sons, 2014.

Istituto Meteorologico Reale Olandese, disponibile a www.knmi.nl.

IEEE RTS Task Force of APM Subcommittee, "IEEE Reliability Test System", IEEE PAS, vol. 98(6), pp. 2047-2054, Nov 1979.