

POLITECNICO DI MILANO

Facoltà di Ingegneria Industriale

Corso di Laurea in
Ingegneria Energetica



Reti neurali per la previsione di energia da fotovoltaico.
Analisi di sensitività

Relatore: Prof.ssa Sonia LEVA

Correlatore Ing. Emanuele OGLIARI

Tesi di Laurea di:

Edoardo PORTACCIO Matr: 771117

Anno Accademico 2012-2013

Indice

1	La situazione energetica europea ed italiana	11
1.1	Il contesto europeo	11
1.1.1	Le fonti rinnovabili	13
1.1.2	L'evoluzione del fotovoltaico in Europa	15
1.2	Il contesto italiano	16
1.2.1	Le fonti rinnovabili: il fotovoltaico	18
1.3	Energy Roadmap 2050	21
2	Direttive italiane in materia di previsione della produzione	25
2.1	Documento per la consultazione 35/2012/R/efr[8]	25
2.1.1	La previsione dell'energia immessa: situazione precedente al DCO	26
2.1.2	La previsione dell'energia immessa: orientamenti dell'Autorità	27
2.1.3	Il ruolo del GSE	28
2.2	Deliberazione 281/2012/R/efr[9]	29
2.3	Deliberazione 462/2013/R/eel[10]	30
3	Le reti neurali	33
3.1	Modelli di previsione dell'energia	33
3.2	Circuiti neurali biologici	34
3.3	Reti neurali artificiali	36
3.3.1	Funzione di attivazione	36
3.3.2	Architetture	37
3.3.3	Gli algoritmi di apprendimento	38
3.4	Confronto tra metodi analitici e tecniche di soft-computing	40
4	Definizioni di errore	43
5	Pre-processing	47
5.1	Validazione dei dati	47
5.2	Criteri di ammissione dati	51

Indice

6 Processing	55
6.1 Analisi di sensitività	56
6.1.1 Variazione configurazione di rete, numero iterazioni e giorni di training	56
6.1.2 Variazioni dataset	61
6.1.2.1 Riduzione del numero di input	61
6.1.2.2 Sostituzione valori mancanti	62
6.1.2.3 Incremento giorni dataset	63
6.1.2.4 Riduzione giorni dataset	64
6.1.2.5 Incremento soglia minima ammissione irraggiamento orario previsto	65
7 Post-processing	71
7.1 Configurazione di rete ottimale	71
7.2 Analisi errore	74
7.3 Confronto tra la precisione dei dati meteorologici e la previsione della rete neurale	78
8 Conclusioni	81

Elenco delle figure

1.1	Evoluzione della produzione e della potenza per fonte nell'UE 27	12
1.2	Mix energetico UE 15	12
1.3	Evoluzione della quota di energia rinnovabile sul consumo finale lordo	13
1.4	Obiettivo 2020: grado di raggiungimento nel 2010 nei paesi dell'UE 27	14
1.5	Evoluzione europea della potenza installata di impianti fotovoltaici (MW)	15
1.6	Evoluzione della potenza cumulata installata in Europa di impianti fotovoltaici	16
1.7	Potenza efficiente degli impianti elettrici di generazione in Italia al 31 dicembre 2012 [GW]	17
1.8	Evoluzione della potenza e numero di impianti fotovoltaici in Italia	19
1.9	Produzione degli impianti fotovoltaici in Italia	19
1.10	Distribuzione regionale della produzione nel 2012	20
1.11	Produzione italiana di energia elettrica nei primi 7 mesi del 2013	21
1.12	Strategia per la riduzione delle emissioni di CO2 in Europa	22
3.1	Struttura schematica del neurone biologico	35
3.2	Struttura del neurone artificiale	37
3.3	Schema rete multistrato	38
5.1	Schema pre-processing	47
5.2	Confronto tra irraggiamento teorico, irraggiamento previsto e irraggiamento misurato	48
5.3	Confronto dati misurati, suddivisione quartoraria e suddivisioni orarie	50
6.1	Schema rete neurale	55
6.2	Andamento dell'errore in funzione del numero di iterazioni	59
6.3	Andamento dell'errore in funzione della configurazione di rete	60
6.4	Andamento dell'errore in funzione dei giorni di addestramento	60
6.5	Grafico riassuntivo simulazione migliore #1	67
6.6	Grafico riassuntivo simulazione migliore #2	68
6.7	Grafico riassuntivo simulazione migliore #3	69
7.1	Schema post-processing	71

Elenco delle figure

7.2	Andamento dell'errore percentuale riferito alla potenza misurata, al variare del numero di giorni di addestramento	72
7.3	Andamento dell'errore in ordine decrescente	75
7.4	Dettaglio errore percentuale riferito alla potenza misurata	76
7.5	Dettaglio errore percentuale riferito alla potenza prevista	77
7.6	Dettaglio errore percentuale riferito alla potenza prevista e alla potenza misurata	77
7.7	Confronto tra irraggiamento previsto e irraggiamento misurato con il relativo errore	79

Elenco delle tabelle

1.1	Composizione del mix medio nazionale	18
5.1	Errore giornaliero suddivisioni “30-15” e “45-30”	49
5.2	Esempio affidabilità (Dettaglio 20/07/2012)	52
5.3	Esempio calcolo val_m (Dettaglio 20/07/2012)	53
5.4	Esempio calcolo $val_{m,h}$ (Dettaglio 20/07/2012)	53
5.5	Esempio dati input a seguito delle modifiche (dettaglio 20/07/2012)	54
6.1	Tabella errori configurazione “11-9”	58
6.2	Tabella errori configurazione “9-7”	58
6.3	Tabella errori configurazione “8-6”	58
6.4	Simulazioni dataset ridotto	62
6.5	Simulazioni con dati mancanti sostituiti	62
6.6	Simulazioni giorni dataset incrementati	63
6.7	Simulazioni giorni dataset ridotti	64
6.8	Simulazioni soglia minima irraggiamento previsto incrementata	65
7.1	Andamento degli errori in funzione del numero di giorni di addestramento	73
7.2	Confronto errori per due diversi giorni di inizio addestramento	74
7.3	Errore previsione con la configurazione ottimale di rete scelta	74
7.4	Errore dati previsionali irraggiamento	78
7.5	Errori relativi all’irraggiamento previsto rispetto all’irraggiamento misurato	80
7.6	Errori relativi alla previsione della rete neurale	80

Introduzione

La rapida penetrazione delle fonti rinnovabili non programmabili nel sistema elettrico nazionale, avvenuta in particolar modo negli ultimi anni, richiede una maggiore partecipazione di tutti gli operatori del settore alla gestione in sicurezza del sistema elettrico. In particolare è necessaria una maggiore responsabilizzazione degli utenti del dispacciamento di impianti alimentati da fonti rinnovabili non programmabili in relazione alla efficiente previsione dell'energia elettrica immessa nella rete.

Lo scopo del presente lavoro di tesi è la messa a punto e la successiva validazione di un sistema di previsione basato su tecniche di soft computing, in particolare su reti neurali artificiali, per la previsione del profilo di produzione di energia su base oraria di un impianto fotovoltaico.

La prima parte è dedicata alla descrizione del contesto energetico europeo ed italiano nel quale si inserisce il presente lavoro.

Nel capitolo 2 vengono descritte le direttive dell'Autorità per l'energia elettrica e il gas in tema di previsione della produzione di energia ed i corrispettivi stabiliti per la valorizzazione degli sbilanciamenti imputabili alle unità di produzione.

Il capitolo 3 descrive gli strumenti attualmente disponibili per la previsione dell'energia distinguendo tra modelli fisici, statistici e ibridi. Una descrizione più approfondita è riservata alle reti neurali artificiali che costituiscono lo strumento di previsione utilizzato in questo lavoro.

Nel quarto capitolo vengono riportate e descritte le definizioni d'errore utilizzate per l'analisi dell'accuratezza delle previsioni.

L'ultima parte è dedicata alle fasi che riguardano la messa a punto e la successiva validazione del sistema di previsione basato sulle reti neurali artificiali e il calcolo dell'errore commesso nel prevedere l'energia immessa in rete.

La prima fase, detta "pre-processing", consiste nell'allineamento temporale tra le grandezze misurate sull'impianto e i parametri meteorologici e nell'individuazione di eventuali dati corrotti o incongruenti. La seconda fase, detta "processing", descrive l'analisi di sensitività effettuata sui principali parametri che caratterizzano la rete neurale allo scopo di determinarne la configurazione migliore. In ultimo, alla luce delle definizioni di errore elencate, è stata valutata la precisione dello strumento di previsione anche in relazione all'entità dell'errore delle previsioni meteorologiche.

1 La situazione energetica europea ed italiana

Il costante aumento registrato negli ultimi anni della produzione di energia da fonti non programmabili, ha accentuato la necessità di una corretta previsione dell'energia immessa in rete al fine di garantire il corretto bilanciamento tra la potenza richiesta e la potenza immessa nella rete.

L'intrinseca variabilità della produzione da fonti rinnovabili, fortemente dipendente dalle condizioni meteorologiche, pone evidenti problematiche associate ai costi di dispacciamento, alla disponibilità di riserve ed alla gestione ottimale della rete.

Il presente capitolo si pone l'obiettivo di dare una breve descrizione della situazione energetica europea ed italiana e di come i modelli previsionali della produzione di energia da fonti rinnovabili possono essere d'aiuto in un'ottica di ottimizzazione della gestione della rete elettrica.

1.1 Il contesto europeo

Una delle principali criticità dell'Unione Europea è rappresentata dalla dipendenza dall'estero per gli approvvigionamenti di fonti primarie di energia, le quali molto spesso provengono da Paesi con elevato profilo di rischio geopolitico. Questa dipendenza riguarda soprattutto l'intero settore energetico (industria, servizi, trasporti etc.) ma in misura inferiore il settore elettrico, dove il ricorso all'energia nucleare o al carbone riesce a ridurre il problema.

I Paesi europei possono contare su combinazioni di fonti utilizzate per la produzione di energia elettrica molto diverse tra loro.

In figura 1.1 si evidenzia l'evoluzione del mix energetico in Europa dal 2000 al 2010. E' possibile notare un forte ricorso alla generazione termoelettrica, in particolare con l'utilizzo di carbone (25%) e gas naturale (22%). L'apporto di petrolio per la generazione elettrica è limitato e nel 2010 si è attestato attorno al 2%. Le fonti rinnovabili assommano a circa il 20% (sommando il contributo di idroelettrico, fotovoltaico, eolico e biomassa).

1 La situazione energetica europea ed italiana

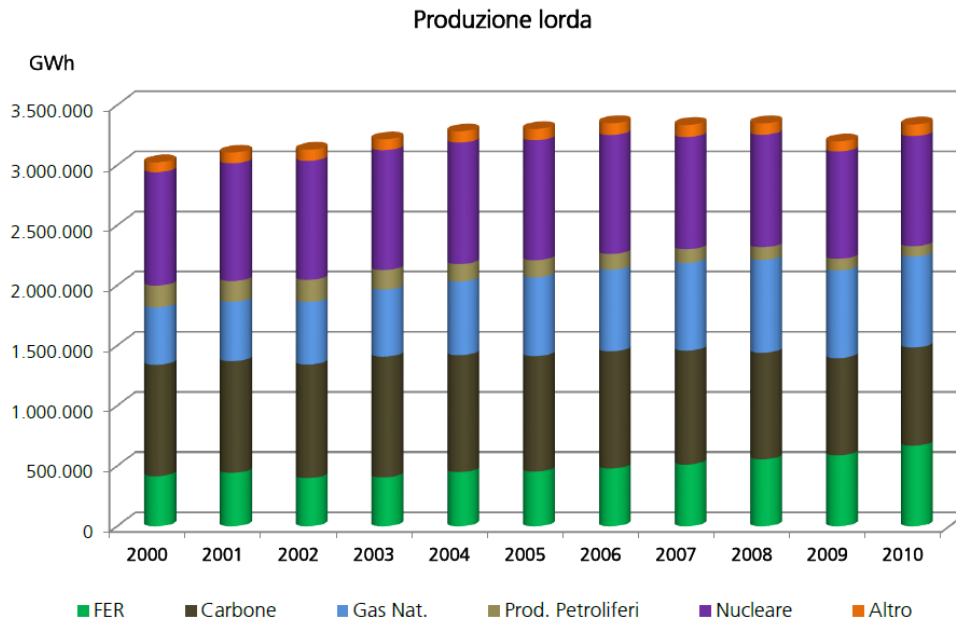


Figura 1.1: Evoluzione della produzione e della potenza per fonte nell'UE 27

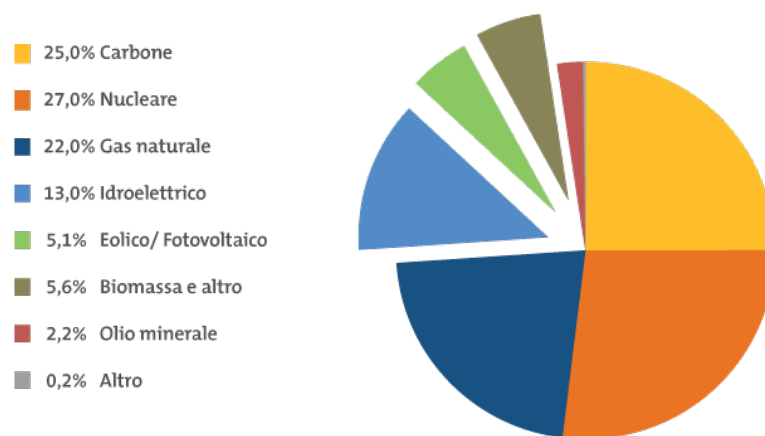


Figura 1.2: Mix energetico UE 15

La Germania utilizza largamente il carbone, che possiede in abbondanza, per la generazione elettrica. Il Paese sta però attraversando una fase di forti cambiamenti dovuti a una incisiva spinta ecologica e in cui l'attenzione per l'ambiente viene vissuta come un'opportunità per il futuro. Negli ultimi anni gli sforzi si sono con-

1.1 Il contesto europeo

centrati maggiormente nel settore del fotovoltaico per il quale sono stati stanziati generosi incentivi finanziari, tra i più alti in Europa. Lo sviluppo della produzione da questa fonte è stato positivamente accompagnato da uno sviluppo dell'industria collegata ad essa.

In Francia circa l'80% dell'energia elettrica è prodotta dalle centrali nucleari, mentre solo in minima parte il Paese ricorre alla generazione termoelettrica. La generazione da fonti rinnovabili riguarda soprattutto lo sfruttamento delle biomasse solide, seguito dall'idroelettrico (le altre fonti rinnovabili sono quasi inesistenti).

La Spagna si è affidata a un mix energetico piuttosto bilanciato tra le varie fonti; tuttavia, negli ultimi anni sono state stanziare risorse finanziarie al fine di incrementare il contributo delle rinnovabili con particolare riferimento all'eolico e al fotovoltaico.

1.1.1 Le fonti rinnovabili

La comunità Europea si è imposta per il 2020 obiettivi sfidanti da raggiungere attraverso una strategia di sviluppo sostenibile. L'efficacia e l'esito dipendono anche dal cambiamento delineato per il settore energetico: al 2020 l'Europa deve coprire i suoi consumi finali con il 20% di produzione da fonti rinnovabili, aumentare del 20% l'efficienza energetica e diminuire del 20% le emissioni di gas serra.

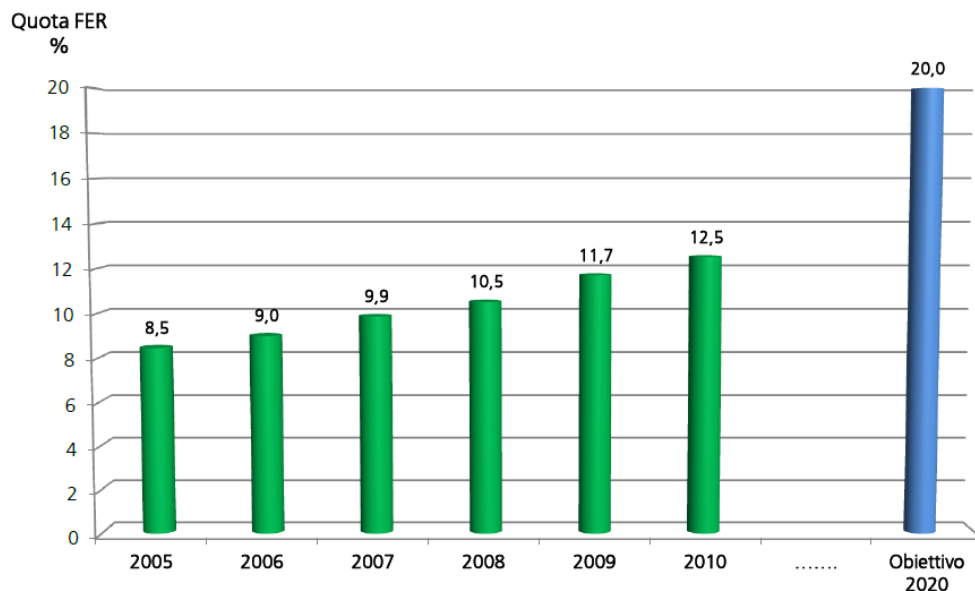


Figura 1.3: Evoluzione della quota di energia rinnovabile sul consumo finale lordo

1 La situazione energetica europea ed italiana

Queste misure costituiscono parti importanti del pacchetto di azioni necessarie per rispettare il protocollo di Kyoto della convenzione quadro delle Nazioni Unite sui cambiamenti climatici e gli ulteriori impegni assunti a livello comunitario e internazionale per la riduzione delle emissioni di gas a effetto serra oltre il 2012.

Le situazioni di partenza, le possibilità di sviluppo dell'energia da fonti rinnovabili e il mix energetico variano da uno Stato membro all'altro. Occorre pertanto tradurre l'obiettivo complessivo comunitario del 20% in obiettivi individuali per ogni Stato membro, procedendo ad un'allocatione giusta e adeguata che tenga conto della diversa situazione di partenza e delle possibilità degli Stati membri, ivi compreso il livello attuale dell'energia da fonti rinnovabili e il mix energetico. A questo scopo, occorre ripartire l'aumento totale richiesto dell'uso dell'energia da fonti rinnovabili tra gli Stati membri sulla base di un aumento uguale della quota di ogni Stato ponderato in funzione del rispettivo PIL, modulato in modo da tenere conto della loro situazione di partenza, ed effettuando i calcoli in termini di consumo finale lordo di energia, tenuto conto dell'impegno precedentemente profuso dagli Stati membri in merito all'uso dell'energia da fonti rinnovabili.

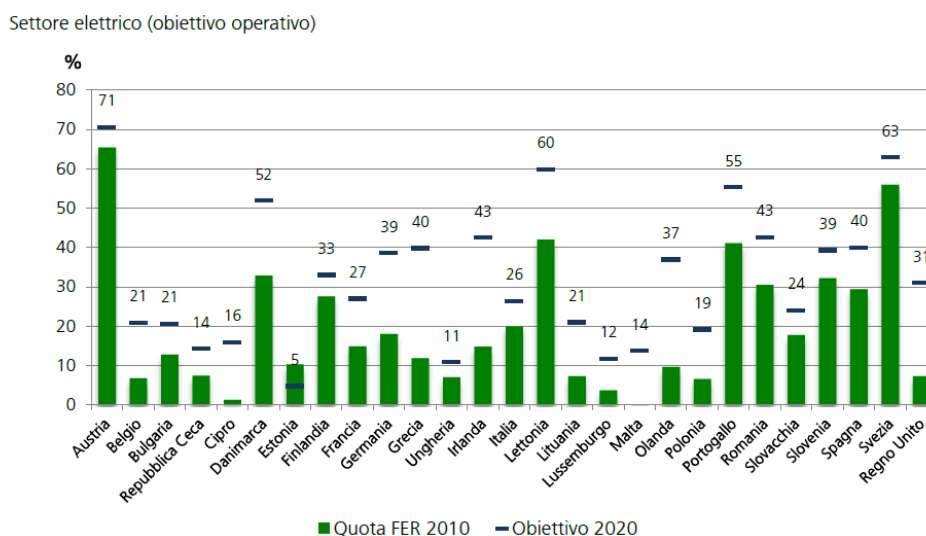


Figura 1.4: Obiettivo 2020: grado di raggiungimento nel 2010 nei paesi dell'UE 27

Ciascun Stato membro ha dovuto adottare dei piani di azione nazionale per le energie rinnovabili che permettessero loro di avanzare verso il conseguimento dei loro obiettivi obbligatori finali. Il punto di partenza è stato fissato all'anno 2005, trattandosi dell'anno più recente in cui si dispone di dati affidabili sulle quote di energia da fonti rinnovabili [1].

1.1.2 L'evoluzione del fotovoltaico in Europa

Lo sviluppo e l'evoluzione in Europa di un mercato dell'energia fotovoltaica è principalmente il risultato dell'impegno di pochi Paesi, con la Germania che, anno dopo anno, si è imposta come leader nello sviluppo e nel supporto economico di questa tecnologia.

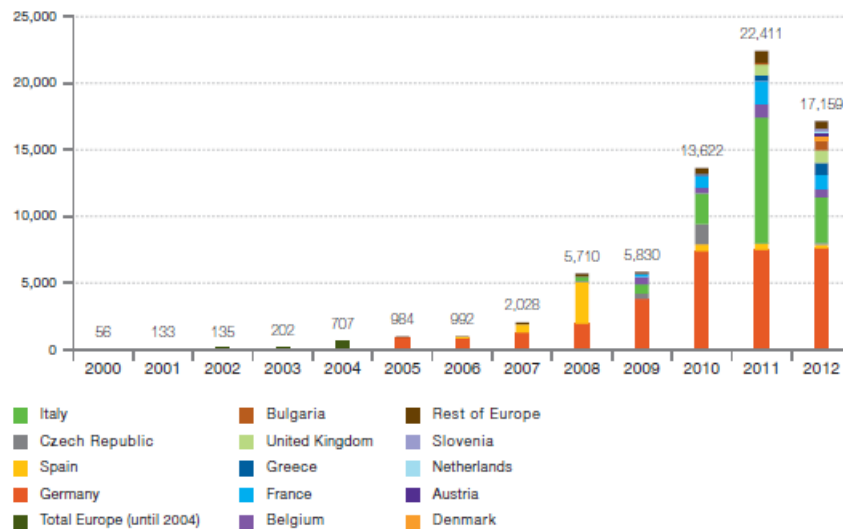


Figura 1.5: Evoluzione europea della potenza installata di impianti fotovoltaici (MW)

Dopo il boom di installazioni in Spagna nel 2008, La Germania ha avuto un deciso incremento nel 2009 della potenza installata, con il resto dell'Europa che mostrava una crescita limitata. Questa situazione è parzialmente imputabile alla crisi economica ma anche a una stabilizzazione del mercato conseguente alla rapida crescita dell'anno precedente.

Una maggiore crescita ritornò nel 2010, con la Germania che ha avuto un numero di installazioni mai visto prima e con Italia e Repubblica Ceca che iniziano a far registrare numeri importanti con una potenza installata intorno ai 3,8 GW. Come successo in Spagna, anche la Repubblica Ceca ha registrato un significativo rallentamento dell'installazione di nuovi impianti nell'anno successivo al boom, soprattutto a causa delle pressioni dei produttori di energia da fonti convenzionali e della preoccupazione del regolatore in seguito alla rapida crescita del mercato.

Nel 2011 Italia e Germania confermano la posizione di leader nel settore con la Francia che si affaccia per la prima volta con decisione nel mercato del fotovoltaico.

Il 2012 è l'anno con maggiori attivazioni di impianti in Germania e assieme all'Italia permette all'Europa di arrivare a una potenza installata nell'anno di 19,2

1 La situazione energetica europea ed italiana

GW. Dietro questi due Paesi, il resto del mercato è rappresentato dal contributo di Gran Bretagna, Grecia, Bulgaria e Belgio.

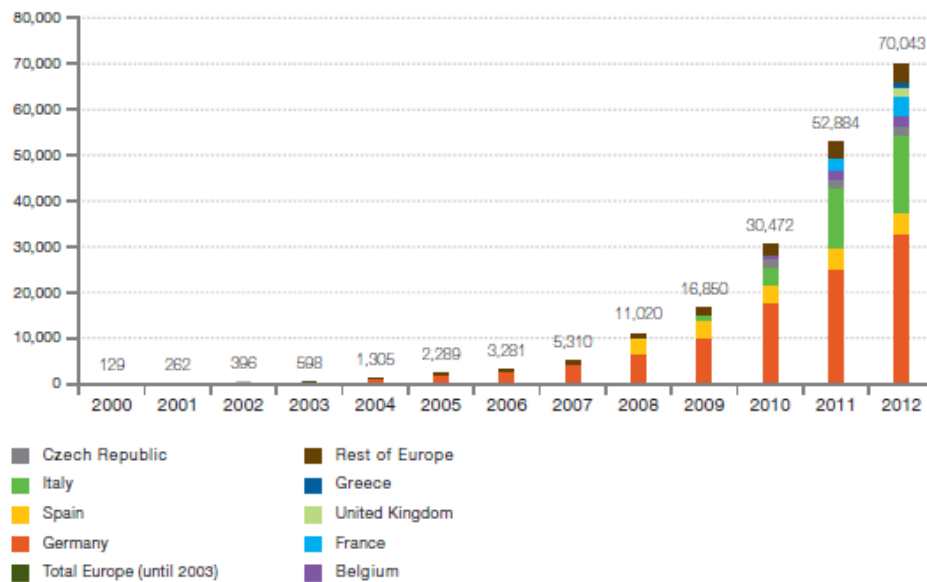


Figura 1.6: Evoluzione della potenza cumulata installata in Europa di impianti fotovoltaici

La suddivisione geografica del mercato del fotovoltaico in Europa è collegato alla velocità di sviluppo negli anni della tecnologia. La Germania è stata protagonista della maggiore crescita negli ultimi dieci anni e rappresenta chiaramente il mercato del fotovoltaico più sviluppato. Vi sono però altri Paesi come Repubblica Ceca, Italia, Grecia e Belgio che nonostante un inizio ritardato, hanno raggiunto velocemente buoni livelli di sviluppo.

Dal punto di vista delle prospettive del fotovoltaico lo sviluppo dell'Europa è senza rivali. USA e Giappone, inizialmente pionieri del settore, sono ora anni indietro all'Europa in termini di penetrazione del fotovoltaico. A parte l'Australia, il resto del mondo ha ancora uno straordinario potenziale non sfruttato [2].

1.2 Il contesto italiano

In Italia la produzione di energia elettrica è basata in gran parte sull'utilizzo di fonti energetiche non rinnovabili (gas naturale, carbone e petrolio in gran parte importati dall'estero) e in misura minore sulle fonti rinnovabili (energia geotermica, solare,

eolica e biomasse). Una parte del fabbisogno elettrico viene coperto con l'acquisto di energia dall'estero.

In termini di potenza installata, nel 2012 la potenza efficiente netta di generazione ha raggiunto i 124.234 MW, con un incremento di 5.791 MW, pari a +4,9% rispetto al 2011. Il maggiore incremento si è avuto nel parco fotovoltaico, passato da 12.773 MW a 16.420 MW, con una crescita del 28,6% rispetto al 2011 [3].

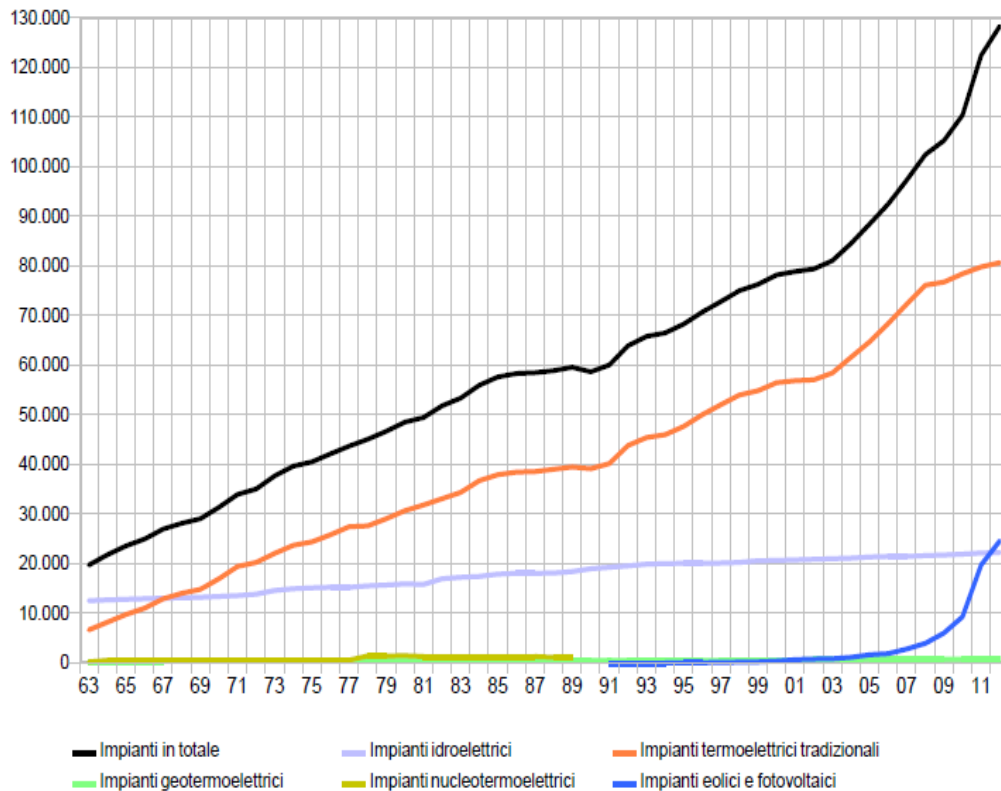


Figura 1.7: Potenza efficiente degli impianti elettrici di generazione in Italia al 31 dicembre 2012 [GW]

1 La situazione energetica europea ed italiana

Composizione del mix medio nazionale utilizzato per la produzione dell'energia elettrica immessa nel sistema elettrico italiano nel 2011		Composizione del mix medio nazionale utilizzato per la produzione dell'energia elettrica immessa nel sistema elettrico italiano nel 2012	
Fonti primarie utilizzate	[%]	Fonti primarie utilizzate	[%]
Fonti rinnovabili	35.2	Fonti rinnovabili	29.8
Carbone	14.9	Carbone	18.5
Gas naturale	40.6	Gas naturale	39.8
Prodotti petroliferi	1.3	Prodotti petroliferi	1.32
Nucleare	1.9	Nucleare	4.8
Altre fonti	6.1	Altre fonti	5.8

Tabella 1.1: Composizione del mix medio nazionale

Il GSE (Gestore Servizi Energetici), nella determinazione del mix energetico per gli anni 2011-2012, precisa che la diminuzione del dato relativo alla produzione da fonte rinnovabile è dovuta principalmente al mancato contributo di Garanzie di Origine Estere, in relazione all'energia elettrica importata [4].

1.2.1 Le fonti rinnovabili: il fotovoltaico

I primi sistemi solari fotovoltaici per la generazione di energia elettrica risalgono ai primi anni '80, dunque il fotovoltaico è piuttosto giovane se si considerano i normali tempi di apprendimento e diffusione di una nuova tecnologia energetica. In questo breve periodo vi è stato uno sviluppo rapidissimo dovuto in particolar modo ai generosi incentivi stanziati. I principali vantaggi che questa tecnologia offre sono:

- l'energia solare è una fonte percepita come particolarmente pulita;
- l'impatto è praticamente nullo in fase di esercizio;
- gli impianti hanno un'elevata affidabilità e ridotti costi di manutenzione;
- risulta ideale per realizzare un nuovo modello di generazione distribuita.

Negli ultimi anni si è assistito a una rapida crescita del numero e della potenza degli impianti fotovoltaici. Il numero di impianti passa dai 32.018 del 2008 ai 330.196 del 2011 (il numero di impianti raddoppia di anno in anno).

La potenza è cresciuta più che proporzionalmente rispetto alla numerosità, in quanto sono entrati in esercizio impianti di dimensioni più grandi (questa tendenza è particolarmente evidente fino al 2011).

1.2 Il contesto italiano

Il numero più elevato di impianti si riscontra al Nord, in particolare in Lombardia e in Veneto, ma è in Puglia che si registra la potenza installata maggiore con 2.499 MW installati e la dimensione media maggiore con 73 kW.

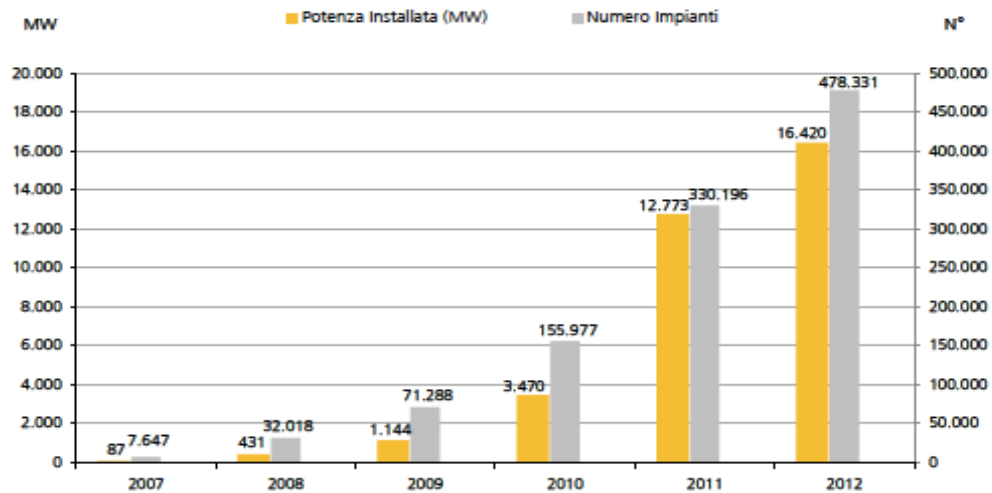


Figura 1.8: Evoluzione della potenza e numero di impianti fotovoltaici in Italia

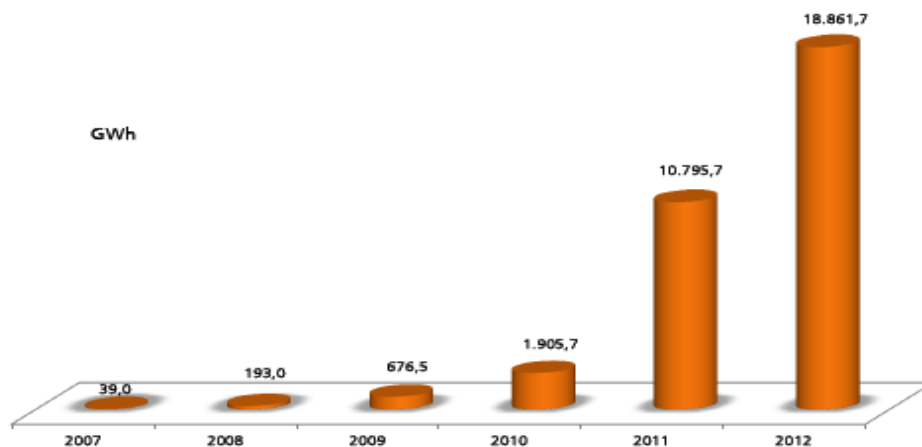


Figura 1.9: Produzione degli impianti fotovoltaici in Italia

1 La situazione energetica europea ed italiana

Nel 2012 la produzione degli impianti fotovoltaici in Italia ha raggiunto 18.862 GWh con un incremento del 75% rispetto all'anno precedente. In soli sei anni la produzione è aumentata di circa 485 volte. Oggi tra le fonti rinnovabili, la fonte solare è seconda solamente alla fonte idraulica.

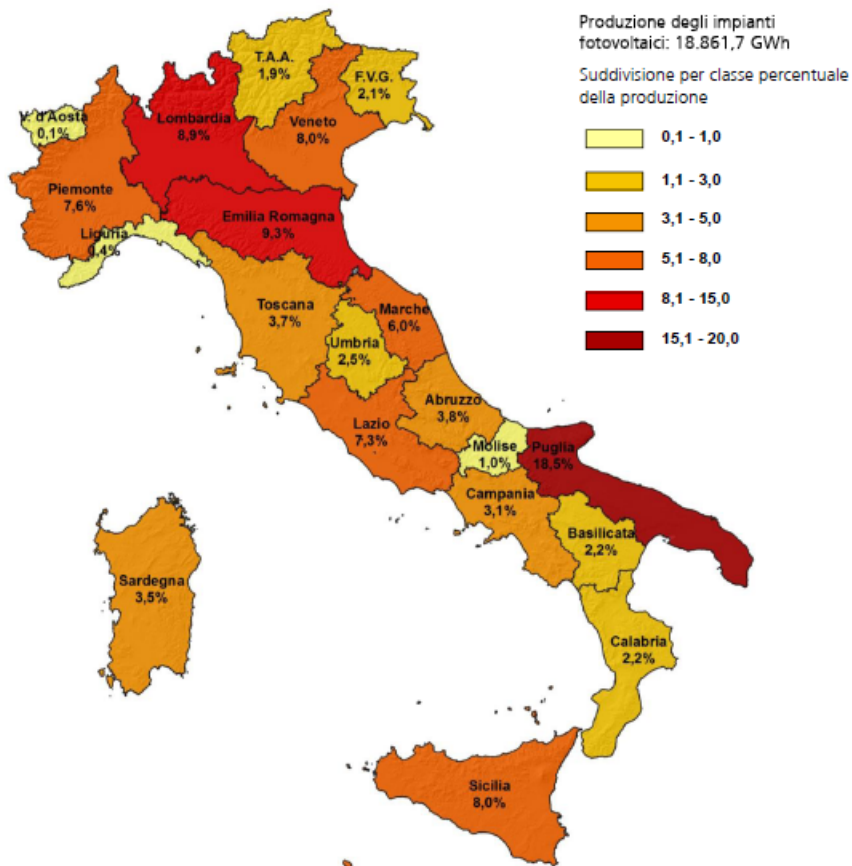


Figura 1.10: Distribuzione regionale della produzione nel 2012

In Puglia sono stati prodotti 3.491 GWh, circa il 19% del totale nazionale. Segue a distanza l'Emilia Romagna, per la quale la produzione è stata pari a 1.758 GWh, il 9% del totale. La Lombardia e la Sicilia hanno prodotto rispettivamente 1.681 GWh e 1.512 GWh, contribuendo rispettivamente con il 9% e l'8% del valore nazionale.

La provincia di Lecce è storicamente la provincia in cui si produce più energia dagli impianti fotovoltaici. Nel 2012 rappresenta il 5% dei 18.862 GWh prodotti a

livello nazionale pari a 942 GWh. Altre Province che si sono distinte per l'entità della produzione sono Brindisi e Foggia seguite da Cuneo e Viterbo.

Nei primi 7 mesi del 2013 la produzione di energia elettrica da fonti rinnovabili è stata pari al [5]:

- 35,9% della domanda italiana;
- 40,5% della produzione italiana;

e la suddivisione per fonti è rappresentata in figura 1.11.

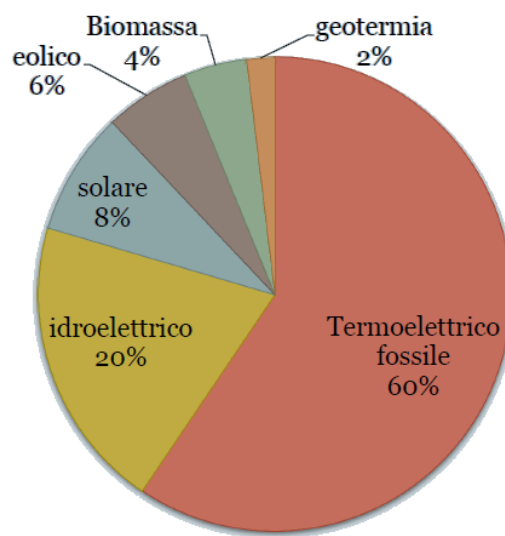


Figura 1.11: Produzione italiana di energia elettrica nei primi 7 mesi del 2013

I dati presentati mostrano un radicale cambiamento in Italia nel settore dell'energia, ed in particolare:

- il numero di impianti di produzione elettrica passa in un decennio da 1 migliaio a 550.000;
- si passa gradualmente da un sistema centralizzato a un sistema misto di generazione;
- un numero crescente di famiglie e imprese ormai controlla la produzione di energia.

1.3 Energy Roadmap 2050

Il Consiglio Europeo ha approvato l'obiettivo dell'Unione Europea di ridurre entro il 2050 le emissioni di gas a effetto serra dell'80-95% rispetto ai livelli del 1990 per

1 La situazione energetica europea ed italiana

scongiurare pericolosi cambiamenti climatici. Per raggiungere tale obiettivo sono state fissate delle tappe intermedie:

- riduzione del 40% entro il 2030;
- riduzione del 60% entro il 2040.

Per fare ciò è necessario uno sforzo considerevole di de-carbonizzazione da parte della generazione elettrica e un incremento dell'efficienza energetica nel settore residenziale e commerciale. Per il 2020 l'obiettivo iniziale di riduzione delle emissioni del 20% è già raggiungibile con le misure finora adottate, in particolare raggiungendo le quote di rinnovabili e proseguendo sul risparmio energetico.

Le politiche del pacchetto 20-20-20, tuttavia, sono solo sufficienti a raggiungere una riduzione delle emissioni interne del 30% entro il 2030 e del 40% entro il 2050, dunque queste politiche andranno rafforzate ed integrate.

Dal punto di vista degli investimenti, nei prossimi 40 anni l'Unione Europea si deve impegnare a investire annualmente l'1.5% del Pil, pari a 270 miliardi di euro. I benefici che si manifesteranno non saranno solo legati alla diminuzione delle emissioni: i Paesi potranno godere della creazione di nuovi posti di lavoro, del mantenimento degli occupati attuali, dell'abbassamento dell'inquinamento atmosferico e dei costi sanitari ad esso connessi.



Figura 1.12: Strategia per la riduzione delle emissioni di CO₂ in Europa

Il 15 Dicembre 2011 la Commissione Europea ha presentato la sua Energy Roadmap 2050. Le modalità prospettate per il raggiungimento degli obiettivi sono quattro:

- efficienza energetica, con effetti sulla domanda di energia;
- fonti rinnovabili;
- tecnologie di cattura e stoccaggio della CO_2 ;
- ricorso al nucleare, che l'Unione Europea non pone come obbligo agli Stati membri, ma continua a sostenere come alternativa sicura e sostenibile rispetto alle fonti tradizionali.

Le modalità appena descritte sono state combinate per ottenere 5 possibili scenari, che hanno lo stesso obiettivo di ottenere la riduzione delle emissioni al 2050 dell'85%, di aumentare la sicurezza degli approvvigionamenti e la competitività del settore energetico.

Questi scenari possono essere così riassunti:

1. Elevata efficienza energetica, con una diminuzione del 41% al 2050 della domanda di energia rispetto al periodo 2005-2006;
2. Diversificazione delle tecnologie di produzione, in cui tutte le tecnologie a basse emissioni di carbonio competono liberamente senza alcuna specifica misura di sostegno;
3. Quota elevata di fonti rinnovabili, aumentando il sostegno a queste tecnologie al fine di raggiungere un'incidenza del 97% delle rinnovabili sulla generazione elettrica totale e una quota del 75% sui consumi finali di energia;
4. Maggiore contributo dell'energia nucleare, ritardando la penetrazione delle tecnologie di cattura e stoccaggio della CO_2 e con la decarbonizzazione guidata dai prezzi del carbonio più che da spinte tecnologiche;
5. Minor contributo dell'energia nucleare, rinunciando alla costruzione di nuove centrali nucleari con una conseguente maggiore penetrazione delle tecnologie CCS portandole a un utilizzo del 32% della generazione totale.

Combinando fra loro i diversi fattori (efficienza energetica, energie rinnovabili, energia nucleare, tecnologie CCS) è possibile raggiungere gli obiettivi prefissati, indipendentemente dallo scenario scelto. Si possono individuare una serie di elementi fondamentali che hanno un impatto positivo, tra cui:

1 La situazione energetica europea ed italiana

- la decarbonizzazione del sistema energetico è fattibile sia sul piano tecnico che su quello economico con un utilizzo di risorse meno oneroso rispetto alle strategie attuali;
- indipendentemente dal mix energetico scelto, l'efficienza energetica e le fonti rinnovabili sono elementi cruciali;
- gli investimenti nelle infrastrutture energetiche portano un risparmio economico rispetto alla possibilità di rimandare questo tipo di interventi;
- il prezzo dell'energia è destinato a crescere fino al 2030 a causa del crescente costo dei combustibili fossili, dovuto alla sempre maggiore richiesta dei Paesi emergenti. Dal 2030 in poi il prezzo potrà scendere grazie agli investimenti nel miglioramento del sistema elettrico;
- Un approccio a livello europeo è certamente vantaggioso a livello economico rispetto ai singoli programmi nazionali, vista l'incidenza delle economie di scala sul settore energetico [7].

2 Direttive italiane in materia di previsione della produzione

La marcata penetrazione delle fonti rinnovabili nel sistema elettrico richiede che le reti di distribuzione si trasformino progressivamente da passive ad attive, con una partecipazione maggiore dei produttori ad una gestione efficace, efficiente e sicura del sistema elettrico. Al fine di consentire un aumento sostenibile dell'energia immessa in rete da fonti rinnovabili, è necessario promuovere una maggiore responsabilizzazione dei produttori in relazione alla previsione dell'energia immessa, evitando che i costi di sbilanciamento connessi a un'errata previsione della stessa ricadano sui soli consumatori.

Il presente capitolo si pone l'obiettivo di descrivere in che modo le Autorità competenti in tema di energia affrontano il problema della corretta previsione della produzione da fonti rinnovabili non programmabili e i provvedimenti presi in tal senso.

2.1 Documento per la consultazione 35/2012/R/efr[8]

Negli ultimi anni, per effetto dell'incentivazione della produzione da fonti rinnovabili, si sta assistendo a un incremento della capacità di produzione di energia, principalmente di tipo non programmabile. Tale sviluppo riguarda in particolare gli impianti eolici connessi alla rete di trasmissione in alta tensione e gli impianti fotovoltaici connessi alla rete di distribuzione in media e bassa tensione. E' auspicabile che negli anni aumenti la partecipazione di questi impianti alla gestione della rete elettrica.

L'Autorità per l'energia elettrica e il gas (di seguito: l'Autorità) prevede che l'aumento della generazione da fonti non programmabili potrà essere causa di criticità:

- sui sistemi di difesa;
- sull'approvvigionamento di risorse per il servizio di dispacciamento.

L'Autorità evidenzia anche l'esigenza di:

- ampliare l'intervallo di frequenza di funzionamento degli impianti di generazione distribuita, allineandoli con gli impianti attivi sulla rete di trasmissione

2 *Direttive italiane in materia di previsione della produzione*

nazionale, al fine di diminuire il rischio di “effetto domino” in caso di criticità sulla rete;

- valutare di dare la possibilità a Terna di ridurre la generazione distribuita, ad iniziare da quella connessa in media tensione, al fine di ricostruire la riserva nel caso in cui le altre azioni perseguibili risultino impraticabili;
- promuovere una maggiore responsabilizzazione dei produttori di energia da fonti non programmabili in relazione a una corretta previsione dell'energia immessa in rete;
- valutare una generale revisione della disciplina del dispacciamento in relazione ai rapidi mutamenti in atto;
- prevedere che Terna valuti, con cadenza periodica, la massima penetrazione della generazione da fonte rinnovabile intermittente.

La definizione di “non programmabilità” delle fonti rinnovabili consiste non tanto nell'imprevedibilità dell'energia prodotta e immessa in rete, quanto nell'impossibilità di effettuare azioni di regolazione sulla base di un programma efficiente predefinito.

2.1.1 La previsione dell'energia immessa: situazione precedente al DCO

L'articolo 11, comma4, del decreto legislativo n. 79/99 prevede che le unità di produzione alimentate da fonti rinnovabili hanno diritto alla priorità di dispacciamento a pari prezzo di offerta, compatibilmente con la sicurezza della rete.

- per le unità di produzione alimentate da fonti rinnovabili programmabili è previsto un corrispettivo di sbilanciamento in base agli effettivi costi che esso comporta sul sistema;
- per le unità di produzione alimentate da fonti rinnovabili non programmabili non vengono attribuiti i costi indotti sul sistema che pertanto vengono socializzati.

Tale approccio non è di tipo cost reflective e si traduce quindi in un ulteriore incentivo per le fonti rinnovabili non programmabili, escludendo inoltre i produttori dal rendersi parte attiva nell'interazione tra impianto e rete.

La deliberazione ARG/elt 5/10 prevede per le unità di produzione rilevanti (potenza nominale superiore o uguale a 10 MVA) strumenti incentivanti per la corretta previsione della produzione da fonti rinnovabili non programmabili. In particolare

se la differenza (in valore assoluto) tra l'energia immessa in rete E_{imm} e l'energia prevista E_{prog} , per una determinata ora h , è inferiore al prodotto tra l'energia immessa e il parametro S_{rif} , allora si prevede un corrispettivo per la corretta previsione (CCP) definito come segue:

$$CCP = premio \cdot (S_{rif} \cdot E_{imm} - |E_{imm} - E_{prog}|) \quad (2.1)$$

dove:

- premio è il corrispettivo unitario per la corretta programmazione, pari a 3 €/MWh;
- S_{rif} è la soglia di riferimento, pari a 0,3 per il 2010, 0,2 per il 2011 e 0,15 per il 2012 e gli anni successivi;
- E_{imm} è l'energia elettrica immessa in rete nell'ora h ;
- E_{prog} è l'energia prevista che dovrebbe essere immessa nell'ora h .

Con riferimento alle unità di produzione non rilevanti alimentate da fonti rinnovabili non programmabili l'onere della previsione non ricade sui produttori. L'assenza di penalizzazioni in caso di sbilanciamenti contribuisce a:

1. sottostimare la produzione da fonti rinnovabili non programmabili che il GSE (Gestore dei Servizi Energetici) deve offrire a prezzo nullo sul Mercato del Giorno Prima (MGP);
2. sottostimare la produzione da fonti rinnovabili non programmabili che gli utenti del dispacciamento diversi dal GSE offrono direttamente su MGP o su MI (Mercato Intraday);
3. incrementare l'errore di previsione di Terna dei fabbisogni orari zionali residui ai fini dell'approvvigionamento di risorse nella fase di programmazione del mercato dei servizi di dispacciamento (MSD).

L'effetto della sottostima della produzione da fonti rinnovabili non programmabili è quello di non favorire un'offerta concorrenziale sui mercati, ritardando una probabile riduzione dei prezzi dell'energia. L'errore della previsione dei fabbisogni orari zionali residui, invece, induce Terna a non stimare correttamente la reali necessità di risorse per la regolazione.

2.1.2 La previsione dell'energia immessa: orientamenti dell'Autorità

L'Autorità stabilisce che il prezzo per la valorizzazione degli sbilanciamenti per ciascun punto di dispacciamento deve essere pari:

2 Direttive italiane in materia di previsione della produzione

- se lo sbilanciamento aggregato zonale è positivo, al minimo tra:
 - il prezzo medio delle offerte di acquisto accettate nel MSD, ponderato per le relative quantità, nel medesimo periodo rilevante, nella macrozona cui il punto di dispacciamento appartiene;
 - il prezzo di valorizzazione delle offerte accettate nel MGP nel medesimo periodo rilevante e nella zona in cui è localizzato il punto di dispacciamento (prezzo zonale orario);
- se lo sbilanciamento aggregato zonale è negativo, al massimo tra:
 - il prezzo medio delle offerte di acquisto accettate nel MSD, ponderato per le relative quantità, nel medesimo periodo rilevante, nella macrozona cui il punto di dispacciamento appartiene;
 - il prezzo di valorizzazione delle offerte accettate nel MGP nel medesimo periodo rilevante e nella zona in cui è localizzato il punto di dispacciamento (prezzo zonale orario).

Uno sbilanciamento dell'unità di produzione di segno opposto rispetto allo sbilanciamento aggregato zonale comporta uno sgravio per il sistema elettrico e si configura dunque come un maggior ricavo (o minor onere) per l'unità di produzione. Al contrario nelle ore in cui lo sbilanciamento dell'unità di produzione è di segno uguale allo sbilanciamento aggregato zonale si ha un aggravio per il sistema elettrico che si traduce in un minor ricavo (o maggior onere) per l'unità di produzione.

L'Autorità ritiene inoltre opportuno prevedere che, nel caso di produzione alimentata da fonti rinnovabili non programmabili, i prezzi di sbilanciamento previsti su base oraria si applichino esclusivamente alla quota di sbilanciamento che, nella medesima ora, eccede l'8% del programma vincolante modificato e corretto del punto di dispacciamento. Per la restante quota si applica il prezzo zonale orario. Inoltre nel caso in cui lo sbilanciamento effettivo è inferiore al 15% del programma vincolante, Terna provveda ad erogare un corrispettivo per la corretta previsione calcolato con la formula 2.1, con una soglia di riferimento S_{rif} pari a 0,15.

2.1.3 Il ruolo del GSE

Il GSE attualmente svolge:

1. un ruolo attivo nella previsione delle immissioni in rete dell'energia elettrica prodotta da fonti rinnovabili non programmabili:
 - a) effettua una previsione delle immissioni di energia da unità di produzione da fonti rinnovabili non programmabili e la rende disponibile a Terna per la gestione in sicurezza del sistema elettrico;

- b) redige i programmi di immissione di unità di produzione non rilevanti;
 - c) redige i programmi di immissione di unità di produzione in regime cip 6/92;
2. un ruolo passivo nella previsione della produzione da fonti rinnovabili di energia elettrica da fonti rinnovabili non programmabili in regime di ritiro dedicato

Il GSE già svolge dunque un'attività di previsione della produzione da fonti rinnovabili non programmabili e i costi sostenuti vengono remunerati dalla collettività, tramite la componente tariffaria A3.

Al fine di valorizzare le competenze acquisite nell'ambito della previsione della produzione si ritiene opportuno che il GSE possa svolgere tale attività, a pagamento, per i produttori che ne fanno richiesta.

2.2 Deliberazione 281/2012/R/efr[9]

La deliberazione trattata riguarda la revisione del servizio di dispacciamento dell'energia elettrica per le unità di produzione alimentate da fonti rinnovabili non programmabili. E' datata 5 luglio 2012 e fa seguito al documento per la consultazione 35/2012/R/efr sopra descritto.

In merito al documento di consultazione sopra citato, gli utenti interessati hanno espresso i seguenti pareri:

- è necessario un congruo anticipo rispetto alla data di prima applicazione della nuova regolazione;
- la necessità di rinegoziare i contratti di forniture in essere suggerisce l'inopportunità di applicare la nuova regolazione fino al 2013;
- è necessario un periodo di almeno 6-12 mesi prima della piena applicazione della nuova regolazione, per permettere agli operatori di sperimentare le tecniche di previsione dell'energia immessa in rete e per raccogliere dati storici;
- la franchigia per lo sbilanciamento deve essere innalzata fino a valori compresi tra il 15% e il 35%;
- la franchigia deve essere differenziata per fonte;
- la franchigia deve essere riferita alla potenza installata;
- gli utenti del dispacciamento possono aggregare i programmi di immissione e di prelievo;

2 *Direttive italiane in materia di previsione della produzione*

- i programmi di immissione devono essere comunicati entro il giorno n-1.

Tenuto conto dei suggerimenti pervenuti l'Autorità ha ritenuto opportuno:

- definire un transitorio iniziale al fine di garantire la necessaria gradualità nella gestione degli impianti, fermo restando la necessità di pervenire al più presto a una situazione a regime di tipo cost reflective;
- prevedere una franchigia più alta nel periodo transitorio, ed in particolare di valorizzare gli sbilanciamenti effettivi solo per la quota che eccede:
 - il 20% del programma vincolante modificato e corretto del punto di dispacciamento per il periodo compreso tra l'1 gennaio 2013 e il 30 giugno 2013;
 - il 10% del programma vincolante modificato e corretto del punto di dispacciamento per il periodo compreso tra l'1 luglio 2013 e il 31 dicembre 2013;

Per la restante quota il corrispettivo di sbilanciamento è pari al prezzo di valorizzazione dell'energia nella zona di competenza

- la disciplina transitoria entri in vigore dall'1 gennaio 2013 al fine di garantire i necessari tempi di adeguamento degli operatori interessati;
- prevedere che il GSE abbia piena autonomia nel presentare le offerte delle unità di produzione per cui è utente del dispacciamento, sia nel MGP che nel MI;
- abrogare il premio per la corretta previsione, visto l'innalzamento delle franchigie relative all'applicazione della valorizzazione degli sbilanciamenti.

2.3 **Deliberazione 462/2013/R/eel[10]**

A seguito del ricorso presentato da ANEV (Associazione nazionale energia del vento) e da alcune aziende associate, il Tar Lombardia ha annullato:

- la deliberazione 281/2012/R/efr, avente a oggetto "Revisione del servizio di dispacciamento dell'energia elettrica per le unità di produzione di energia elettrica alimentate da fonti rinnovabili";
- la deliberazione 493/2012/R/efr, recante "Approvazione delle modalità per l'attribuzione dei corrispettivi di sbilanciamento e dei corrispettivi a copertura dei costi amministrativi da attribuire ai produttori in regime di ritiro dedicato e di tariffa fissa onnicomprensiva";

2.3 Deliberazione 462/2013/R/eel[10]

- le relative “Regole Tecniche per il trasferimento delle partite economiche relative ai corrispettivi di sbilanciamento e alle offerte accettate sul mercato infragiornaliero”.

A seguito di tali sentenze l’Autorità ha presentato ricorso al Consiglio di Stato, con richiesta di sospensiva dei provvedimenti presi.

La deliberazione 462/2013/R/eel fa seguito alle sentenze del Consiglio di Stato e considerato che:

- la direttiva 2009/72/CE impone che le produzioni da fonti rinnovabili non programmabili non possono essere esentate dal contribuire alla sicurezza del sistema elettrico ed alla partecipazione ai corrispettivi di sbilanciamento;
- lo sviluppo delle rinnovabili deve essere compatibile con una gestione sicura della rete elettrica;
- la corretta applicazione dei corrispettivi di sbilanciamento consente di attribuire agli sbilanciamenti il proprio valore di mercato dando un segnale economico ai produttori di energia elettrica da fonti rinnovabili non programmabili al fine di promuovere la corretta previsione dell’energia immessa nella rete elettrica;
- con le delibere precedenti non si penalizzano le fonti rinnovabili non programmabili, bensì si prevede un trattamento differente e più favorevole rispetto a quello applicato alle fonti programmabili;

l’Autorità ha deciso che:

- è necessario dare tempestiva esecuzione alle ordinanze del Consiglio di Stato al fine di salvaguardare la sicurezza del sistema elettrico nazionale;
- è confermato l’annullamento della diminuzione delle franchigie previste nella deliberazione 281/2012/R/efr in quanto diminutive della effettiva differenziazione tra fonti rinnovabili programmabili e non programmabili.

3 Le reti neurali

Dopo aver analizzato il contesto e le prospettive future nel quale si colloca il presente lavoro, è utile analizzare gli strumenti che verranno utilizzati nel processo di previsione della produzione di energia elettrica di un impianto fotovoltaico.

Sotto il nome soft computing può essere raggruppata una serie di tecniche che si propongono di affrontare problemi complessi per i quali non esistono soluzioni analitiche. In termini generali si può dire che le tecniche di soft computing imitano i processi di apprendimento dell'intelligenza umana e si differenziano dai metodi tradizionali, per i quali sono necessarie relazioni matematiche chiuse. Fra le principali tecniche di soft computing esistenti sono state utilizzate le reti neurali artificiali. Sono tre le caratteristiche principali di questo strumento che ne suggeriscono la scelta: robustezza, flessibilità e capacità di generalizzazione.

Le reti neurali sono robuste poichè riescono a fornire una risposta anche nel caso in cui vengano eliminate una o più connessioni. Sono flessibili perchè si adattano a molte finalità diverse. Inoltre le reti neurali hanno una forte capacità di generalizzazione poichè, anche con un numero ristretto di esempi per l'addestramento, riescono a fornire una risposta adeguata a dei dati in ingresso mai visti prima.

3.1 Modelli di previsione dell'energia

I modelli di previsione dell'energia ad oggi esistenti possono essere classificati in tre categorie: fisici, statistici e ibridi [19].

Modelli fisici Sotto il nome di modelli fisici sono racchiusi quei metodi di previsione che descrivono la capacità dell'impianto di convertire i parametri metereologici in ingresso in un valore di potenza elettrica. Una volta stabilita l'influenza delle condizioni atmosferiche e delle coordinate geografiche in cui l'impianto è situato, il modello non necessita di dati storici di produzione. Frequentemente questi modelli sono stati utilizzati per gli impianti eolici, ma possono anche trovare applicazione nel caso di impianti fotovoltaici.

Un esempio di modello semplificato può essere descritto dalla seguente equazione:

$$P_E = e \cdot G \cdot A$$

3 Le reti neurali

dove P_E è l'energia elettrica prodotta, G è l'irraggiamento previsto, e è l'efficienza di conversione del pannello solare e A è la sua superficie. I limiti di questo tipo di modelli sono principalmente la forte dipendenza dalla precisione delle previsioni meteorologiche e dalle caratteristiche dell'impianto.

Modelli statistici

Naive Prediction. L'approccio più lineare per determinare i valori futuri di una serie storica è quello che si basa sull'assunzione che l'energia prodotta in un certo periodo P_{p+1} è pari a quella osservata nel periodo P_p . Questo metodo è denominato naive o persistent prediction. Nonostante i forti limiti di precisione questo metodo è comunemente utilizzato perchè facile da implementare.

Similar-Days Model. Questo metodo consente di fare una previsione utilizzando i dati storici di giornate tra loro simili. Sono modelli molto utilizzati nella previsione del carico, dove la dipendenza dalle condizioni atmosferiche è inferiore rispetto all'influenza dei modelli di consumo derivanti dalle serie storiche. Risultano utili nel caso non siano disponibili previsioni meteo.

Reti neurali artificiali. L'utilizzo delle reti neurali rappresenta una valida alternativa all'utilizzo di metodi lineari di previsione. Mediante dei dati storici di esempio la rete è in grado di imparare a riconoscere e trattare dati in ingresso mai visti prima. Nel prosieguo del capitolo verranno descritte in dettaglio le caratteristiche di questo strumento.

Modelli ibridi Una combinazione di uno o più metodi prima citati dà origine ai modelli ibridi. L'utilizzo di questi modelli offre la possibilità di prendere i vantaggi di ognuno dei metodi che la compongono.

3.2 Circuiti neurali biologici

Siccome le reti neurali artificiali aspirano ad essere un modello semplificato dei circuiti biologici è bene dare di quest'ultimi una breve descrizione.

Il sistema nervoso è organizzato in regioni e moduli, ciascuno dei quali è caratterizzato dalla presenza di tre elementi costitutivi: i neuroni principali, i neuroni intrinseci e le fibre nervose. Le fibre nervose trasportano i segnali a entrambi i tipi di neuroni attraverso sinapsi posizionate sull'albero dendritico o sul soma dei neuroni. I neuroni principali possiedono un lungo assone che trasmette il segnale ad altri moduli, mentre i neuroni intrinseci trasmettono il segnale soltanto all'interno del modulo in cui si trovano.

3.2 Circuiti neurali biologici

Un neurone è solitamente composto da un albero dendritico, un soma e un assonone ed è caratterizzato da un alto grado di convergenza e divergenza: i segnali provenienti da migliaia di altri moduli convergono verso il soma del neurone, il quale a sua volta diverge attraverso migliaia di terminali verso altri neuroni.

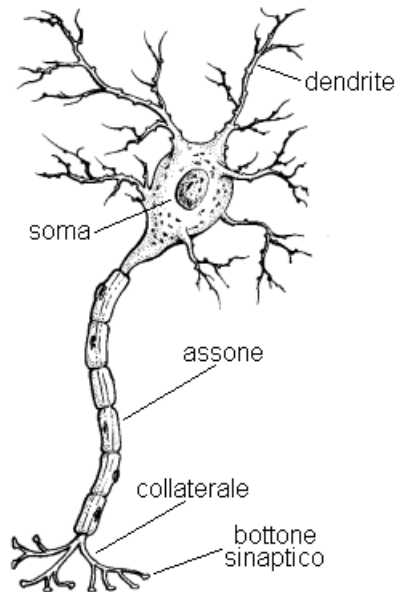


Figura 3.1: Struttura schematica del neurone biologico

Le sinapsi svolgono un ruolo fondamentale in questa fitta rete di comunicazione poichè regolano l'entità e l'effetto dell'informazione sul neurone ricevente. Le sinapsi possiedono tre importanti proprietà:

- sono puntiformi;
- trasmettono il segnale in una sola direzione;
- utilizzano neurotrasmettitori chimici.

Dal neurone presinaptico arriva un segnale che causa l'emissione di alcune sostanze chimiche (neurotrasmettitori) che agiscono sulla membrana postsinaptica permettendo l'assorbimento di ioni (positivi o negativi). In sostanza viene convertito un segnale elettrico in segnale chimico che viene nuovamente convertito in segnale elettrico. Le sinapsi possono essere di due tipi: eccitatorie o inibitorie. Solitamente nelle reti neurali artificiali l'effetto di queste sinapsi è dato dalla somma algebrica di contributi eccitatori e inibitori.

3 Le reti neurali

Un neurone emette un segnale quando la differenza di potenziale fra la parte interna e quella esterna raggiunge una certa soglia. Nella fase di riposo il neurone si trova in una fase di equilibrio a circa -70mV. La ricezione del segnale provoca una depolarizzazione verso valori positivi di potenziale, in grado di provocare una scarica. Successivamente si ha una fase di iperpolarizzazione che riporta il potenziale a valori negativi fino ad arrivare alla situazione iniziale di riposo [11].

3.3 Reti neurali artificiali

Le reti neurali artificiali sono una notevole semplificazione dei circuiti neurali biologici descritti prima. Un neurone artificiale è rappresentato da:

- un insieme di sinapsi che corrispondono ai terminali di altri neuroni
- una soglia
- una funzione di attivazione

L'effetto di un segnale x sul neurone postsinaptico è semplicemente uguale al prodotto $w \cdot x$, dove w è il peso attribuito alla sinapsi corrispondente. Il potenziale di attivazione è dato dalla somma algebrica dei prodotti di tutti i segnali di ingresso x_j e i valori dei pesi corrispondenti w_{ij}

$$A_i = \sum_j^N w_{ij} \cdot x_j \quad (3.1)$$

a cui va sottratto il valore soglia s_i del neurone

$$y_i = f(A_i) = f\left(\sum_j^N w_{ij} \cdot x_j - s_i\right) \quad (3.2)$$

dove il peso w_{ij} di ciascun neurone può assumere valori continui sia positivi che negativi.

3.3.1 Funzione di attivazione

La funzione di attivazione determina il tipo di risposta che un neurone è in grado di emettere. In letteratura esistono varie formulazioni e le più comuni sono definite come segue:

- la funzione di attivazione assume valore 1 se $A > s$, 0 altrimenti
- la funzione di attivazione assume valore 1 se $A > s$, -1 altrimenti

- la funzione di attivazione è del tipo $f(A) = kA$

Le funzioni continue hanno il vantaggio di poter trasmettere dei segnali di intensità graduale, e ciò li fa assomigliare di più ai neuroni biologici. Vi è inoltre una famiglia di funzioni continue non-lineari tra cui una delle più utilizzate è la funzione sigmoide:

$$f(A) = \frac{1}{1 + e^{-kA}} \quad (3.3)$$

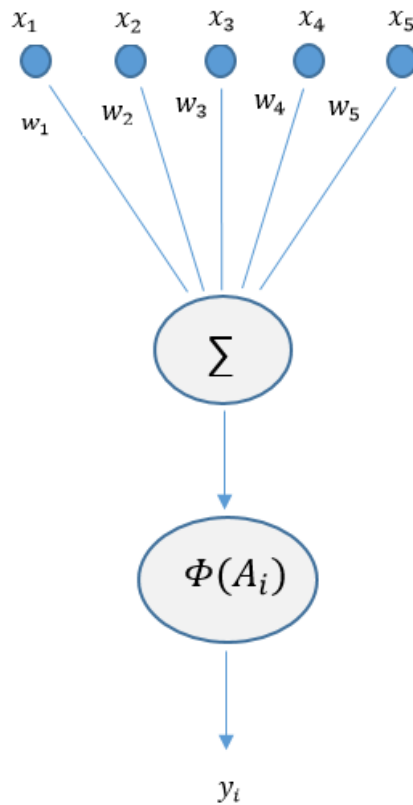


Figura 3.2: Struttura del neurone artificiale

3.3.2 Architetture

Le diverse architetture di rete si distinguono in base al numero di neuroni in ingresso e uscita, dal numero di strati di neuroni e dal tipo di connessioni tra essi. Per una prima suddivisione si possono dividere le reti in due grandi famiglie:

3 Le reti neurali

- le reti auto-associative
- le reti etero-associative

Le reti auto-associative sono composte da un unico strato di neuroni interamente connessi fra di loro. Ciascun neurone riceve quindi dei segnali sia dall'ambiente esterno che dagli altri neuroni, dunque la risposta potrà variare anche in presenza di un segnale esterno costante. Queste reti vengono impiegate soprattutto nella memorizzazione di pattern. Nelle reti etero-associative i nodi di ingresso che ricevono il segnale dall'ambiente esterno sono distinti dai nodi in uscita che forniscono la risposta della rete. Non vi sono collegamenti tra neuroni dello stesso strato.

All'interno delle reti etero-associative si possono fare ulteriori distinzioni. In alcuni casi un unico strato di neuroni non è sufficiente per apprendere correttamente la relazione tra l'input e l'output: in questi casi sono necessarie le reti multistrato (o MLP, Multi-Layer Perceptron) che possiedono più strati di neuroni.

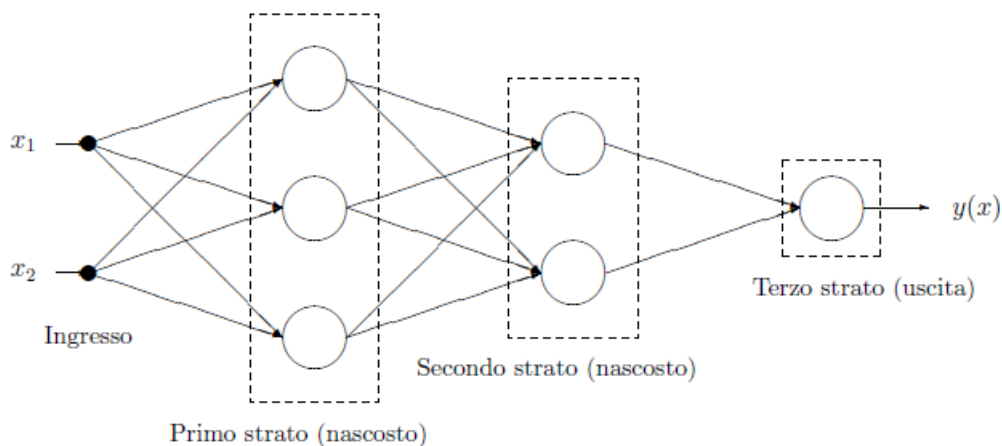


Figura 3.3: Schema rete multistrato

3.3.3 Gli algoritmi di apprendimento

Affinchè una rete neurale funzioni è necessario che i parametri che la definiscono (architettura, funzione di attivazione, pesi sinaptici) vengano determinati in modo corretto. L'architettura di rete e la funzione di attivazione possono essere stabilite a priori in base alle conoscenze e agli obiettivi dello sperimentatore, ma per trovare dei valori corretti dei pesi sinaptici è necessario un processo di addestramento.

Le procedure che consentono di determinare questi fattori vengono dette algoritmi di apprendimento e in generale, partendo da un valore casuale, modificano

i parametri in modo tale che l'output previsto approssimi sempre meglio l'output richiesto.

Con il termine apprendimento supervisionato si indica un tipo di apprendimento caratterizzato dalla presenza di un "insegnante esterno" che fornisce informazioni alla rete neurale. Questa informazione risulta utile alla rete per valutare la bontà della risposta ed eventualmente apportare le modifiche ai pesi per migliorarla.

Percettroni semplici L'idea di modificare i pesi sinaptici confrontando la risposta della rete e la risposta desiderata fu proposta da Rosenblatt. I percettroni di Rosenblatt erano delle reti neurali composte da due strati: le sinapsi del primo strato erano fissate in modo casuale e non potevano essere modificate, le sinapsi del secondo strato erano invece sottoposte ad apprendimento. Data una rete neurale con N nodi d'ingresso con un vettore d'ingresso x e una risposta desiderata t , l'output della rete è dato da

$$y = \begin{cases} 1 & \text{se } \sum_{i=0}^N w_i x_i > 0 \\ -1 & \text{altrimenti} \end{cases}$$

questo valore viene confrontato con l'uscita desiderata e se i valori sono uguali i pesi non vengono modificati, altrimenti si agisce sul peso in base alla risposta corretta. Man mano che l'addestramento procede le modifiche dei pesi diventano sempre più rare perchè un minor numero di ingressi provoca delle risposte scorrette.

La regola delta A differenza dei percettroni la regola delta non utilizza una funzione di attivazione a gradino, bensì una funzione continua e differenziabile. Il vantaggio di avere una funzione differenziabile è che è possibile calcolare il differenziale dell'errore rispetto ai pesi sinaptici, cioè è possibile valutare la variazione dell'errore al variare dei pesi.

Dato che si è interessati a ridurre il valore dell'errore E , ovvero a minimizzare lo scarto quadratico medio tra la risposta desiderata e la risposta prevista, è necessario modificare i pesi nella direzione opposta al gradiente di E :

$$\Delta w_{ij} = -\frac{\partial E}{\partial w_{ij}}$$

e dato che la variazione dei pesi è data dalla differenza tra la risposta desiderata e la risposta della rete si può scrivere che

$$\Delta w_{ij} = a (t_i - y_i) x_j$$

dove a è il tasso di apprendimento che modifica il valore del peso.

3 Le reti neurali

Rispetto ai perceptron semplici la regola delta opera con funzioni di attivazione continue. La condizione per il raggiungimento di una soluzione è l'indipendenza lineare dei dati in ingresso. Questa condizione è sufficiente ma non necessaria. In alcuni casi, infatti, questo algoritmo di apprendimento può portare la rete a una soluzione anche nel caso di dati in ingresso linearmente dipendenti, ma vi è una forte dipendenza dal modo in cui viene presentato il vettore di uscita.

Back-propagation L'algoritmo di back-propagation è un'estensione della regola delta ed è uno dei più diffusi algoritmi di apprendimento per la sua semplicità, efficacia e potenza. Fino al momento della formulazione di questo algoritmo di apprendimento non si riuscivano a sfruttare le grosse potenzialità delle reti multistrato con uno o più strati nascosti tra ingresso e uscita.

Il principio di funzionamento è semplicemente la discesa del gradiente dell'errore tramite la modifica dei pesi sinaptici. La funzione di attivazione utilizzata è solitamente una funzione sigmoideale o logistica, anche se altre funzioni possono essere usate. L'errore viene calcolato come differenza tra la risposta della rete e la risposta desiderata. La differenza sta nel calcolo del termine delta per le unità nascoste: è l'errore delle unità di output trasportato all'indietro dalle connessioni sinaptiche e moltiplicato per la derivata prima della funzione di output di ciascuna unità nascosta.

Questo metodo è ricorsivo e può essere applicato ad un qualsiasi numero di strati nascosti di neuroni.

3.4 Confronto tra metodi analitici e tecniche di soft-computing

Nelle precedenti sezioni sono state descritte le caratteristiche delle reti neurali artificiali e le principali tecniche di addestramento. Per affrontare il problema della previsione della produzione di un impianto da fonte rinnovabile non programmabile, in [12] sono state messe a confronto due diverse soluzioni:

- metodo analitico (APVF)
- rete neurale multistrato (MPVF)

Il metodo analitico fornisce una previsione della produzione oraria di un impianto fotovoltaico come somma di due componenti: la prima è data dalla produzione in condizioni "clear sky" e la seconda è un fattore di attenuazione della potenza. I valori di quest'ultima componente sono calcolati basandosi sui dati meteorologici previsti. Il metodo di calcolo può essere riassunto con i seguenti punti:

1. Si ha a disposizione l'irraggiamento previsto I_t

3.4 Confronto tra metodi analitici e tecniche di soft-computing

2. Si calcola il valore dell'irraggiamento extraterrestre I_0
3. Per ogni ora si calcola il valore dell'irraggiamento in condizioni clear sky I_c
4. Si calcola il fattore di attenuazione dell'irraggiamento K_t dove $K_t = \frac{I_t}{I_c}$
5. Mediante una regressione polinomiale si ottiene il valore della potenza in condizioni clear sky P_c con $P_c = f(I_c)$
6. Partendo dal valore di K_t , mediante una seconda regressione polinomiale, si ottiene il valore del coefficiente di attenuazione della potenza K_p
7. Il valore della potenza oraria prevista si ottiene dal prodotto di P_c per il fattore di attenuazione K_p

Il modello MPVF prevede l'utilizzo di una rete neurale che fornisce la produzione oraria di un impianto fotovoltaico a partire dai dati meteorologici in ingresso. Gli input in ingresso sono in parte riferiti a dati meteorologici previsionali e in parte rappresentano il momento dell'anno e della giornata al quale si riferiscono. Il modello utilizza 11 variabili d'ingresso e 7 neuroni in entrambi gli strati nascosti di neuroni.

Per valutare la precisione dei due modelli è stata utilizzata la seguente definizione di errore:

$$DEMAE_M = 100 \cdot \frac{1}{M} \sum \frac{|rde - fde|}{rde}$$

dove $DEMAE_M$ è l'errore medio giornaliero della previsione del modello, rde è l'energia realmente prodotta dall'impianto fotovoltaico, fde è l'energia prevista e M è il numero di giorni di test utilizzati.

Viene inoltre definito un modello di riferimento puramente numerico che prevede la potenza oraria dell'impianto a partire dalla potenza prevista in condizioni clear sky corretta di un fattore ottenuto facendo il rapporto tra l'energia effettivamente prodotta in un anno e l'energia in condizioni clear sky.

Lo studio evidenzia che entrambi i modelli presi in considerazione forniscono delle previsioni più accurate rispetto al modello di riferimento, con valori di $DEMAE$ pari al 25,4% nel caso APVF e 26,6% nel caso MPVF (il modello di riferimento fa registrare un valore di errore pari a 33,1%).

4 Definizioni di errore

Come conseguenza dell'elevata distribuzione e penetrazione mondiale della generazione di potenza da fonte solare, la previsione della produzione è diventata una variabile cruciale nella pianificazione e nell'integrazione di queste nuove forme di produzione. Gli operatori del settore (utilities, produttori e gestore del sistema di trasmissione dell'energia) richiedono una buona previsione della potenza prodotta per svolgere al meglio le proprie attività. E' necessario dunque definire delle metodologie comuni per valutare la qualità delle previsioni [15].

Il riferimento iniziale è l'errore orario e_h [8], definito come la differenza tra la potenza prodotta nell'ora $P_{m,h}$ e la potenza prevista $P_{p,h}$:

$$e_h = P_{m,h} - P_{p,h} \quad (4.1)$$

che può assumere valori sia positivi che negativi.

Partendo da questa definizione di base si possono introdurre ulteriori definizioni:

- Errore orario assoluto $e_{h,abs}$ [16] che è il valore assoluto della definizione precedente:

$$e_{h,abs} = |e_h| \quad (4.2)$$

- Errore giornaliero e_d , calcolato come sommatoria estesa alle 24 ore di e_h :

$$e_d = \sum_h e_h \quad (4.3)$$

- Errore giornaliero assoluto $e_{d,abs}$, calcolato come sommatoria estesa alle 24 ore dell'errore assoluto $e_{h,abs}$:

$$e_{d,abs} = \sum_1^h |e_h| \quad (4.4)$$

- Errore percentuale $e_{\%,r}$, riferito alla potenza nominale dell'impianto fotovoltaico (P_r):

$$e_{pu,r} = \frac{|e_h|}{P_r} \quad (4.5)$$

4 Definizioni di errore

$$e_{\%,r} = \frac{|e_h|}{P_r} \cdot 100 \quad (4.6)$$

- Errore percentuale $e_{\%,p}$, riferito alla potenza oraria prevista ($P_{p,h}$):

$$e_{pu,p} = \frac{|e_h|}{P_{p,h}} \quad (4.7)$$

$$e_{\%,p} = \frac{|e_h|}{P_{p,h}} \cdot 100 \quad (4.8)$$

- Errore minimo $e_{min,p}$, che è il minimo valore dell'errore percentuale $e_{pu,p}$:

$$e_{min,p} = \min(e_{pu,p}) \quad (4.9)$$

- Errore massimo $e_{max,p}$, che è il massimo valore dell'errore percentuale $e_{pu,p}$:

$$e_{max,p} = \max(e_{pu,p}) \quad (4.10)$$

- Errore medio $e_{avg,p}$, che è il valore medio dell'errore percentuale $e_{pu,p}$:

$$e_{avg,p} = \frac{\sum_1^N e_{pu,p}}{N} \quad (4.11)$$

dove N è il numero di ore.

- Errore quadratico medio $RMSE_p$:

$$RMSE_p = \sqrt{\frac{\sum_1^N e_{pu,p}^2}{N}} \quad (4.12)$$

- Errore percentuale giornaliero $e_{\%,d,p}$, riferito alla potenza oraria prevista ($P_{p,h}$):

$$e_{\%,d,p} = \sum_h \frac{|e_h|}{P_{p,h}} \quad (4.13)$$

- Errore percentuale $e_{\%,m}$ [17], riferito alla potenza oraria prodotta ($P_{m,h}$)

$$e_{pu,m} = \frac{|e_h|}{P_{m,h}} \quad (4.14)$$

$$e_{\%,m} = \frac{|e_h|}{P_{m,h}} \cdot 100 \quad (4.15)$$

- Errore minimo $e_{min,m}$, che è il minimo valore dell'errore percentuale $e_{pu,m}$

$$e_{min,m} = \min(e_{pu,m}) \quad (4.16)$$

- Errore massimo $e_{max,m}$, che è il massimo valore dell'errore percentuale $e_{pu,m}$

$$e_{max,m} = \max(e_{pu,m}) \quad (4.17)$$

- Errore medio $e_{avg,m}$, che è il valore medio dell'errore percentuale $e_{pu,m}$:

$$e_{avg,m} = \frac{\sum_1^N e_{pu,m}^2}{N} \quad (4.18)$$

- Errore quadratico medio $RMSE_m$:

$$RMSE_m = \sqrt{\frac{\sum_1^N e_{pu,m}^2}{N}} \quad (4.19)$$

- Errore assoluto medio MAE (se percentuale è chiamato $MAPE$):

$$MAE = \frac{1}{N} \cdot \sum_{h=1}^N |P_{m,h} - P_{p,h}| \quad (4.20)$$

- Errore assoluto medio normalizzato $NMAE_{\%}$ [18], basato sulla capacità netta dell'impianto (C):

$$NMAE_{\%} = \frac{1}{N} \cdot \sum_{h=1}^N \frac{|P_{m,h} - P_{p,h}|}{C} \cdot 100 \quad (4.21)$$

dove C è la "capacità netta dell'impianto".

- se C è pari al massimo valore di $P_{m,h}$:

4 Definizioni di errore

$$NMAE_{m\%} = \frac{1}{N} \cdot \sum_{h=1}^N \frac{|P_{m,h} - P_{p,h}|}{\max(P_{m,h})} \cdot 100 \quad (4.22)$$

- se C è pari al massimo valore di $P_{p,h}$:

$$NMAE_{p\%} = \frac{1}{N} \cdot \sum_{h=1}^N \frac{|P_{m,h} - P_{p,h}|}{\max(P_{p,h})} \cdot 100 \quad (4.23)$$

- se C è pari alla potenza nominale dell'impianto (P_r):

$$NMAE_{r\%} = \frac{1}{N} \cdot \sum_{h=1}^N \frac{|P_{m,h} - P_{p,h}|}{P_r} \cdot 100 \quad (4.24)$$

- Media ponderata dell'errore assoluto $WMAE_{\%}$, basato sull'energia totale prodotta:

$$WMAE_{\%} = \frac{\sum_{h=1}^N |P_{m,h} - P_{p,h}|}{\sum_{h=1}^N P_{m,h}} \cdot 100 \quad (4.25)$$

- Errore quadratico medio normalizzato $nRMSE$ [19], basato sulla massima potenza misurata $P_{m,h}$:

$$nRMSE = \frac{\sqrt{\frac{\sum_{h=1}^N |P_{m,h} - P_{p,h}|^2}{N}}}{\max(P_{m,h})} \cdot 100 \quad (4.26)$$

5 Pre-processing

L'obiettivo di questo capitolo è analizzare la fase di preparazione del set di dati che verrà successivamente utilizzato nella fase di addestramento della rete neurale. Questa fase preliminare può apparire secondaria rispetto alle fasi successive ma, come vedremo in seguito, risulta di fondamentale importanza. La buona riuscita della fase di addestramento della rete è direttamente legata alla qualità dei dati forniti in input.

Di seguito verranno elencate le operazioni di modifica apportate ai dati in input e i criteri di ammissione dei dati alla fase successiva.

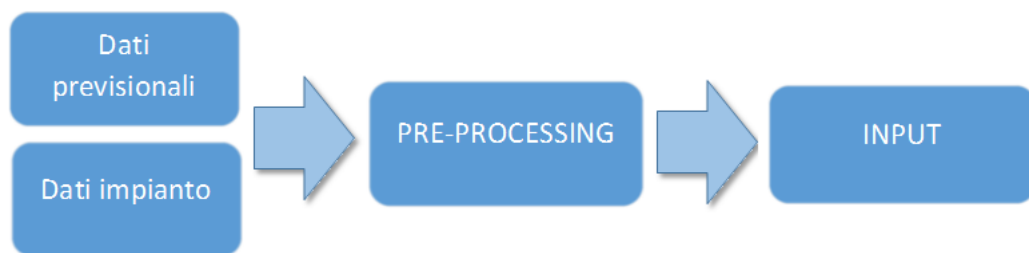


Figura 5.1: Schema pre-processing

5.1 Validazione dei dati

Allo scopo di valutare la possibilità di utilizzare i dati per la previsione, ne è stata effettuata una prima analisi di coerenza considerando i seguenti aspetti:

- Allineamento del fuso orario dei dati previsionali con quelli misurati
- Mancanza di dati a causa di:
 - passaggio dall'ora solare all'ora legale e viceversa
 - errata comunicazione del valore da parte dei sensori e dell'inverter.

Per quanto riguarda il primo punto si è scelto di allineare i dati al fuso orario UTC. I dati mancanti sono stati aggiunti ed è stato loro assegnato il valore 0.

Una volta eliminati gli errori di corrispondenza dei dati, si pone il problema di verificare che essi abbiano una consistenza dal punto di vista fisico. Dopo aver

5 Pre-processing

calcolato l'irraggiamento teorico previsto per la località in esame si confronta l'andamento teorico dell'irraggiamento con l'irraggiamento previsto e quello misurato, come mostrato in figura 5.2. E' possibile fare due considerazioni che confermano la bontà dei dati:

- Gli orari di alba e tramonto coincidono
- I dati previsti e quelli misurati non superano mai l'irraggiamento teorico

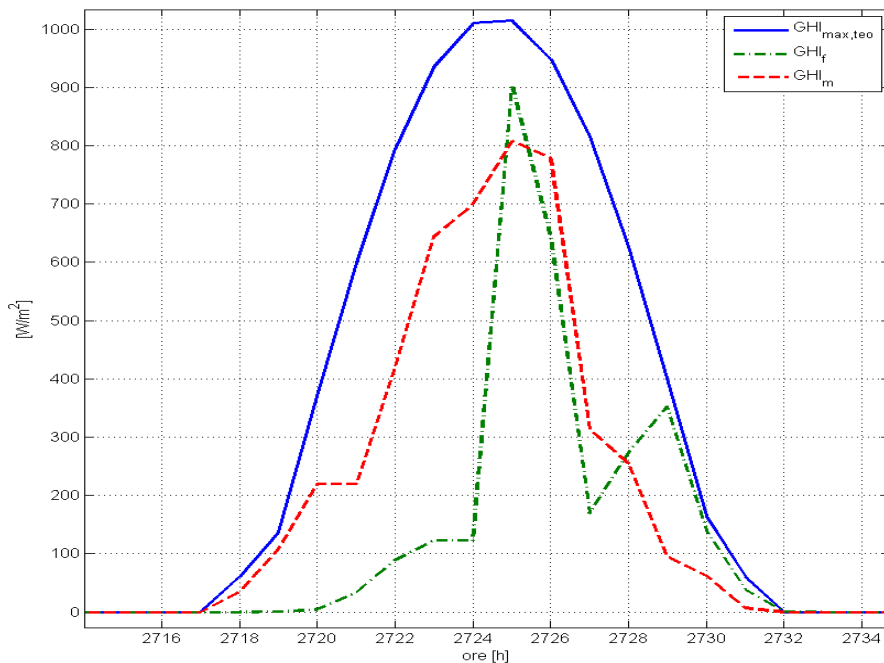


Figura 5.2: Confronto tra irraggiamento teorico, irraggiamento previsto e irraggiamento misurato

Una volta verificato che i dati sono coerenti, si pone il problema della diversa suddivisione temporale tra i dati previsionali e quelli misurati. I dati reali vengono misurati su base oraria, invece i dati misurati vengono forniti su base quattoraria (4 campioni per ogni ora). Bisogna dunque trovare il miglior metodo per riportare i dati dalla base quattoraria a quella oraria. Sono state valutate due possibili soluzioni:

5.1 Validazione dei dati

- media oraria utilizzando i dati della mezz'ora precedente e del quarto d'ora successivo (esempio media ore 9:00 con i 4 campioni dalle 8:30 alle 09:15). Per comodità verrà individuata col nome "30-15".
- media oraria utilizzando i dati del quarto d'ora precedente e della mezz'ora successiva (esempio media ore 9:00 con i 4 campioni dalle 8:45 alle 09:30). Per comodità verrà individuata col nome "45-30".

Per valutare quale delle due soluzioni scegliere si calcola l'errore medio giornaliero definito come segue:

$$e_{\%,avg,m} = \frac{1}{N} \sum_1^N \frac{|GHImq - GHImh|}{GHImq} \cdot 100 \quad (5.1)$$

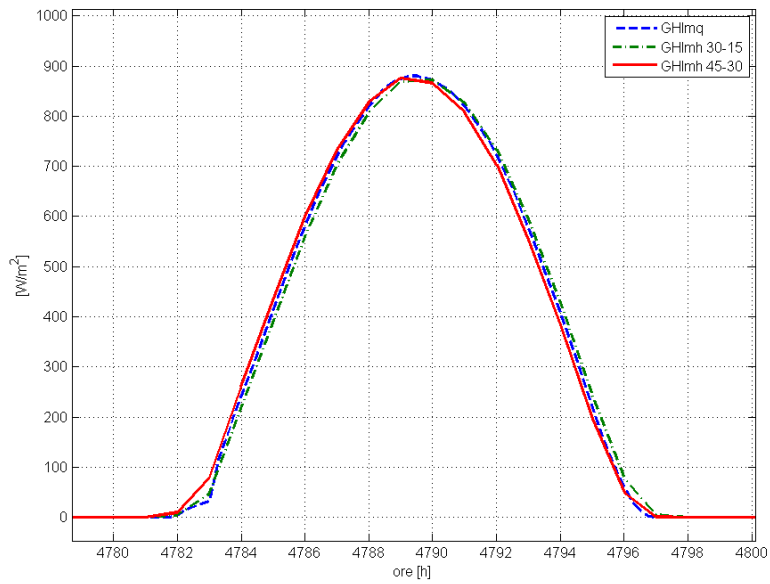
dove $GHImq$ è il dato relativo all'ora precisa (ad esempio il campione registrato alle 08:00, 09:00 etc.) e $GHImh$ è il valore dell'irraggiamento ottenuto passando dalla suddivisione quartoraria a quella oraria. L'errore è stato calcolato considerando solo le ore in cui l'irraggiamento è diverso da 0. Sono state considerate due giornate rappresentative del campione (il 18/07/2012 per il calcolo su una giornata serena e il 05/04/2012 per il calcolo su una giornata nuvolosa) ed è stata applicata la formula 5.1 sia alla suddivisione "30-15" che a quella "45-30".

GIORNO	$e_{\%,avg,m}$ "30-15"	$e_{\%,avg,m}$ "45-30"
05/04/2012	22.3%	25.2%
18/07/2012	8.9%	23.2%

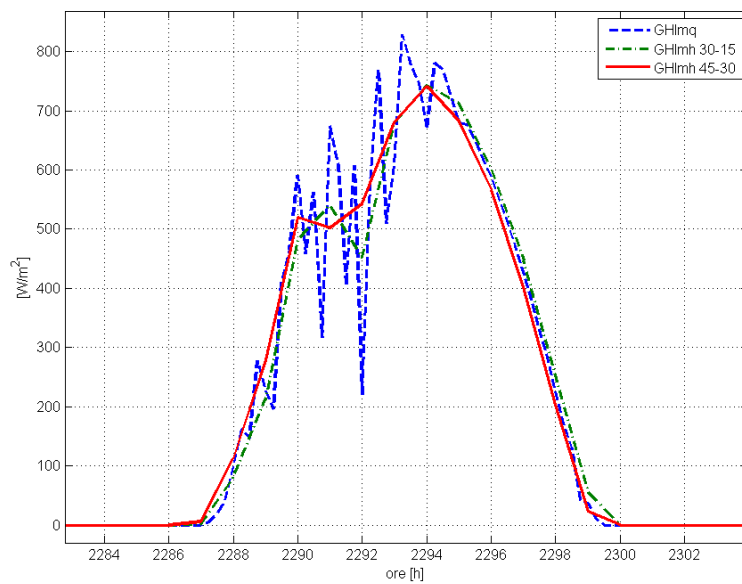
Tabella 5.1: Errore giornaliero suddivisioni "30-15" e "45-30"

In tabella 5.1 sono stati riportati i risultati ottenuti che suggeriscono la scelta della suddivisione "30-15" come migliore.

5 Pre-processing



(a) giornata serena



(b) giornata nuvolosa

Figura 5.3: Confronto dati misurati, suddivisione quartoraria e suddivisioni orarie

5.2 Criteri di ammissione dati

Per quanto riguarda l'addestramento della rete neurale, nonché l'analisi degli errori sottesi alla previsione, i dati storici acquisiti devono essere sempre validati. L'utilizzo infatti di dati non affidabili implicherebbe in generale un aumento dell'errore commesso nella previsione. Per dati non affidabili o incongruenti si intendono:

- mancanza dati per problemi di trasmissione degli stessi;
- dati fisicamente errati;
- potenza prodotta nulla causata da periodi di fermo impianto per manutenzione, malfunzionamenti o guasti dell'impianto;
- irraggiamento misurato diverso da zero nel periodo tra tramonto e alba.

Il metodo proposto consiste nel confrontare i valori quartorari dell'irraggiamento teorico con i valori di irraggiamento e potenza misurati. Definiamo aff_{ghim} e aff_{pwm} rispettivamente l'affidabilità relativa all'irraggiamento misurato e l'affidabilità relativa alla potenza misurata. I valori dell'affidabilità vengono definiti come segue:

- aff_{ghim} assume valore:
 - 0 se il valore dell'irraggiamento teorico è nullo o prossimo a 0
 - 0 se il valore dell'irraggiamento misurato è nullo ed il valore dell'irraggiamento teorico è diverso da 0
 - 1 se entrambi i valori di irraggiamento sono diversi da 0.
- aff_{pwm} assume valore:
 - 0 se il valore dell'irraggiamento teorico è nullo
 - 0 se il valore della potenza misurata è nullo e il valore dell'irraggiamento teorico è diverso da 0
 - 1 se irraggiamento teorico e potenza misurata sono entrambi diversi da 0.

In tabella 5.2 si riporta un esempio di calcolo dell'affidabilità per il giorno 20/07/2012 dalle 04:30 alle 08:15

5 Pre-processing

ORA	GHI_{teo}	GHI_m	P_m	aff_{ghim}	aff_{pwm}
04:30	0	0	0	0	0
04:45	0	0	0	0	0
05:00	0	0	0	0	0
05:15	6.9e-134	0	0	0	0
05:30	1.4	0	0	0	0
05:45	2.8	0	0	0	0
06:00	5.5	6.38	0	1	0
06:15	66.3	20.7	80	1	1
06:30	158.1	36.7	2000	1	1
06:45	249.7	53.7	6000	1	1
07:00	333.4	65.4	8160	1	1
07:15	408.7	61.9	7200	1	1
07:30	476.6	134.3	18960	1	1
07:45	538	151.4	21680	1	1
08:00	593.8	221.8	34160	1	1
08:15	644.7	244.9	38880	1	1

Tabella 5.2: Esempio affidabilità (Dettaglio 20/07/2012)

dove GHI_{teo} è l'irraggiamento teorico, GHI_m è l'irraggiamento misurato e P_m è la potenza misurata.

A questo punto è opportuno definire un unico parametro che possa contenere sia le informazioni dell'affidabilità sulla potenza misurata che dell'irraggiamento misurato. Definiamo val_m come segue:

$$val_m = aff_{ghim} \cdot aff_{pwm} \quad (5.2)$$

che assume valore 0 o 1.

5.2 Criteri di ammissione dati

ORA	$GHI_{max,teo}$	GHI_m	P_m	aff_{ghim}	aff_{pwm}	val_m
04:30	0	0	0	0	0	0
04:45	0	0	0	0	0	0
05:00	0	0	0	0	0	0
05:15	6.9e-134	0	0	0	0	0
05:30	1.4	0	0	0	0	0
05:45	2.8	0	0	0	0	0
06:00	5.5	6.38	0	1	0	0
06:15	66.3	20.7	80	1	1	1
06:30	158.1	36.7	2000	1	1	1
06:45	249.7	53.7	6000	1	1	1
07:00	333.4	65.4	8160	1	1	1
07:15	408.7	61.9	7200	1	1	1
07:30	476.6	134.3	18960	1	1	1
07:45	538	151.4	21680	1	1	1
08:00	593.8	221.8	34160	1	1	1
08:15	644.7	244.9	38880	1	1	1

Tabella 5.3: Esempio calcolo val_m (Dettaglio 20/07/2012)

Anche in questo caso è necessario passare dalla suddivisione quartoraria a quella oraria. Utilizziamo la suddivisione “30-15” per ottenere i valori orari di affidabilità $aff_{ghim,h}$ e $aff_{pwm,h}$, potenza misurata $P_{m,h}$, irraggiamento teorico $GHI_{teo,h}$, irraggiamento misurato $GHI_{m,h}$ e parametro di validità $val_{m,h}$.

ORA	$GHI_{teo,h}$	$GHI_{m,h}$	$P_{m,h}$	$aff_{ghim,h}$	$aff_{pwm,h}$	$val_{m,h}$
05:00	1.72e-134	0	0	0	0	0
06:00	18.925	6.77	20	0.5	0.25	0.25
07:00	287.5	54.4	5840	1	1	1
08:00	563.3	188.1	28420	1	1	1

Tabella 5.4: Esempio calcolo $val_{m,h}$ (Dettaglio 20/07/2012)

In precedenza è stato evidenziato che valori molto bassi di irraggiamento portano a commettere errori in percentuale molto alti ma che in termini assoluti sono irrilevanti. E’ per questo motivo che si decide di stabilire delle soglie al di sotto delle quali i dati non vengono presi in considerazione. val_{mh} viene posto pari a 0 se:

- l’ irraggiamento orario previsto GHI_{fh} è inferiore a $10 \frac{W}{m^2}$

5 Pre-processing

- la potenza oraria misurata $P_{m,h}$ è inferiore a $1kW$
- il suo valore è inferiore a 1.

In ultimo i valori di irraggiamento teorico $GHI_{teo,h}$ inferiori a 1 vengono posti pari a 0.

ORA	$GHI_{teo,h}$	GHI_{mh}	$P_{m,h}$	$aff_{ghim,h}$	$aff_{pwm,h}$	$val_{m,h}$
05:00	0	0	0	0	0	0
06:00	18.925	6.77	20	0.5	0.25	0
07:00	287.5	54.4	5840	1	1	1
08:00	563.3	188.1	28420	1	1	1

Tabella 5.5: Esempio dati input a seguito delle modifiche (dettaglio 20/07/2012)

Conclusa la fase di pre-processing il risultato è un set di dati che combina i dati meteorologici previsionali con i dati dell'impianto misurati relativi all'anno 2012. E' composto da 8784 righe e 10 colonne e fornisce le seguenti informazioni:

- Data $dayh$
- Ora $hourh$
- Irraggiamento orario previsto su superficie GHI_{fh}
- Temperatura ambiente oraria prevista ta_{fh}
- Precipitazioni orarie previste mm_{fh}
- Velocità del vento oraria prevista ws_{fh}
- Direzione del vento oraria prevista wdf_{fh}
- Potenza oraria misurata $P_{m,h}$
- Affidabilità oraria val_{mh}
- Irraggiamento teorico orario GHI_{th}

6 Processing

In questo capitolo verranno descritte le fasi di addestramento e test della rete neurale e l'analisi di sensitività dei principali parametri che la caratterizzano. La rete è composta da due strati nascosti di neuroni i quali hanno tipicamente molti ingressi e una sola uscita. Ogni ingresso ha associato un peso che determina la conducibilità del canale d'ingresso. L'attivazione del neurone è una funzione della somma pesata degli ingressi.

Il metodo più usato per addestrare una rete neurale consiste nel presentare in ingresso alla rete un insieme di esempi (training set). La risposta fornita dalla rete viene confrontata con la risposta desiderata, si valuta la differenza tra le due e, in base a tale differenza, si aggiustano i pesi. Questo processo viene ripetuto sull'intero training set finché le uscite della rete producono un errore al di sotto di una soglia prestabilita. Dopo l'addestramento la rete viene controllata testandone il comportamento su un insieme di dati, detto test set, costituito da esempi non utilizzati durante la fase di training. La fase di test ha quindi lo scopo di valutare la capacità di generalizzazione della rete neurale. Diremo che la rete ha "imparato" se è in grado di fornire risposte anche per ingressi che non le sono stati mai presentati durante la fase di addestramento.

In questo lavoro è stata utilizzata una rete neurale del tipo MLP (Multi-Layer Perceptron) con procedura di addestramento di tipo EBP (Error Back Propagation).

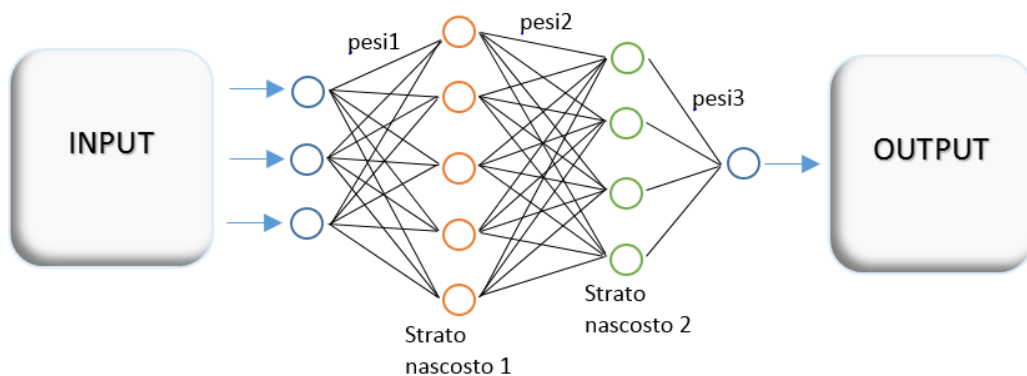


Figura 6.1: Schema rete neurale

Ovviamente le prestazioni di una rete neurale dipendono fortemente dall'insieme di esempi scelti per l'addestramento. Tali esempi devono quindi essere rappresentativi della realtà in cui la rete neurale verrà utilizzata[13].

6.1 Analisi di sensitività

Allo scopo di indagare la dipendenza dell'errore dai parametri che definiscono la rete viene effettuata un'analisi di sensitività. I parametri presi in considerazione sono:

1. Configurazione della rete
2. Numero di iterazioni
3. Numero giorni di training
4. Numero di giorni del dataset

6.1.1 Variazione configurazione di rete, numero iterazioni e giorni di training

La variazione della configurazione di rete consiste nel variare il numero di neuroni degli strati nascosti. Il numero di neuroni determina la complessità della rete e in base al caso applicativo, può risultare più conveniente utilizzare una rete più o meno complessa. Il criterio cui affidarsi nella scelta è quello della minimizzazione del rischio di overlearning che si corre quando si decide di inserire un numero eccessivo di neuroni, i quali disegnano perfettamente il pattern della serie storica ma non sono in grado di generare una previsione affidabile. Viceversa, il rischio di assegnare un numero basso di neuroni è quello di avere una scarsa capacità di apprendimento. Le formule proposte in letteratura sono varie:

$$n_{hl} = 2 \cdot n_{input} + 1 \quad (6.1)$$

$$n_{hl} = 2 \cdot n_{input} \quad (6.2)$$

$$n_{hl} = n_{input} \quad (6.3)$$

$$n_{hl} = \frac{n_{input} + n_{output}}{2} + \sqrt{n_{training}} \quad (6.4)$$

dove n_{hl} è il numero di neuroni nascosti, n_{input} è il numero di input, n_{output} è il numero di output e $n_{training}$ è il numero di osservazioni contenute nel training set.

I risultati empirici dimostrano come nessuna di queste regole risulta generalizzabile ad ogni problema previsionale [14].

Dopo aver fatto dei veloci tentativi allo scopo di escludere le configurazioni con evidenti problemi di stabilità si è scelto di focalizzare l'attenzione sulle seguenti configurazioni di rete:

- configurazione “11-9” composta da:
 - 11 neuroni nel primo strato nascosto
 - 9 neuroni nel secondo strato nascosto
- configurazione “9-7” composta da:
 - 9 neuroni nel primo strato nascosto
 - 7 neuroni nel secondo strato nascosto
- configurazione “8-6” composta da:
 - 8 neuroni nel primo strato nascosto
 - 6 neuroni nel secondo strato nascosto

Il numero di giorni di training determinano il numero di esempi da fornire alla rete per “imparare”. I valori scelti sono:

- 120 giorni
- 150 giorni
- 180 giorni

Infine sono stati scelti i seguenti valori di iterazioni:

- 1000
- 2000
- 3000
- 5000
- 10000

I risultati delle simulazioni effettuate vengono riassunti nelle tabelle 6.1, 6.2 e 6.3. Per le tre configurazioni di rete scelte e al variare del numero di iterazioni e dei giorni di addestramento è stato riportato in tabella l'errore commesso (media ponderata dell'errore assoluto basata sull'energia prodotta), definito nell'equazione 4.25.

6 Processing

GIORNI TRAINING	NUMERO ITERAZIONI				
	1000	2000	3000	5000	10000
120	34.6	33.8	34.2	32.2	36.1
150	37.3	32.5	31	29.7	31.8
180	30.7	29.3	29.7	30.3	31

Tabella 6.1: Tabella errori configurazione “11-9”

GIORNI TRAINING	NUMERO ITERAZIONI				
	1000	2000	3000	5000	10000
120	32.8	36	34.2	34.6	40.8
150	38.8	31.1	33.7	32.8	32.7
180	32.8	32.8	30.6	32.9	32.6

Tabella 6.2: Tabella errori configurazione “9-7”

GIORNI TRAINING	NUMERO ITERAZIONI				
	1000	2000	3000	5000	10000
120	33	36.1	35.1	36.5	36.6
150	36.9	30.6	33.2	36.4	35.6
180	32.9	31.2	32.4	33.7	37

Tabella 6.3: Tabella errori configurazione “8-6”

Nei prossimi capitoli verranno commentati e analizzati i risultati appena riassunti in tabella e i risultati delle prossime simulazioni. Di seguito si vuole fare una prima analisi sull’andamento dell’errore in funzione dei parametri della rete considerati.

1. Andamento dell’errore in funzione del numero di iterazioni. In figura 6.2 sono riportati gli errori commessi considerando le tre configurazioni di rete scelte e 180 giorni di addestramento
2. Andamento dell’errore in funzione della configurazione di rete. In figura 6.3 sono riportati gli errori commessi considerando i tre possibili valori dei giorni di addestramento e 2000 iterazioni
3. Andamento dell’errore in funzione del numero di giorni di addestramento. In figura 6.4 è riportato l’errore commesso per le tre configurazioni di rete scelte e 3000 iterazioni

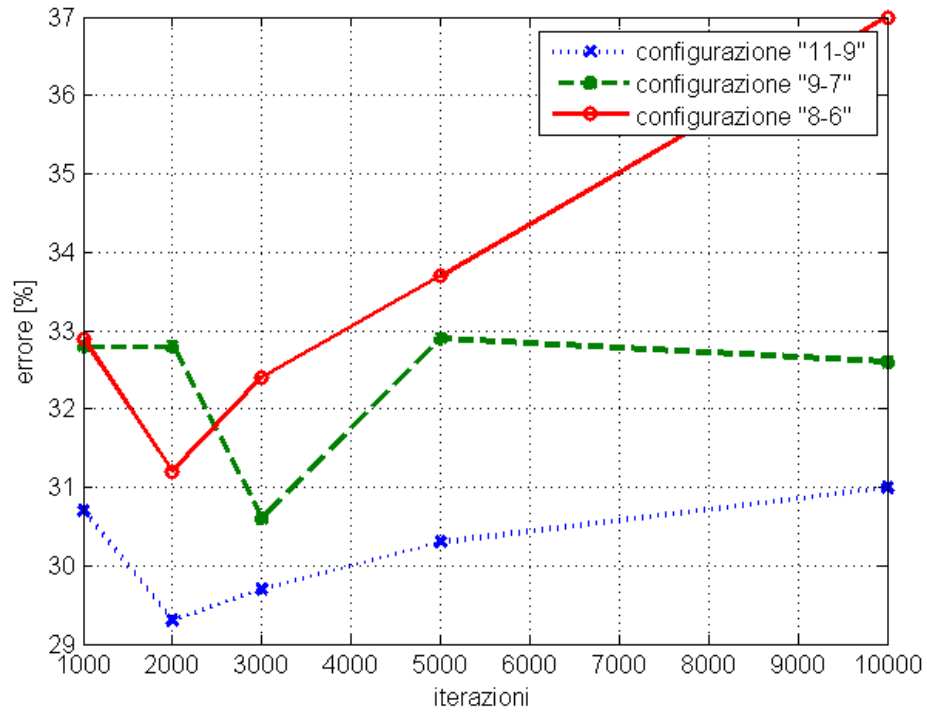


Figura 6.2: Andamento dell'errore in funzione del numero di iterazioni

Valutando l'andamento dell'errore in funzione del numero di iterazioni, per le tre configurazioni della rete neurale considerate, si nota come l'errore scende rapidamente per valori intorno a 2000-3000 iterazioni per poi riprendere a crescere per valori più elevati.

6 Processing

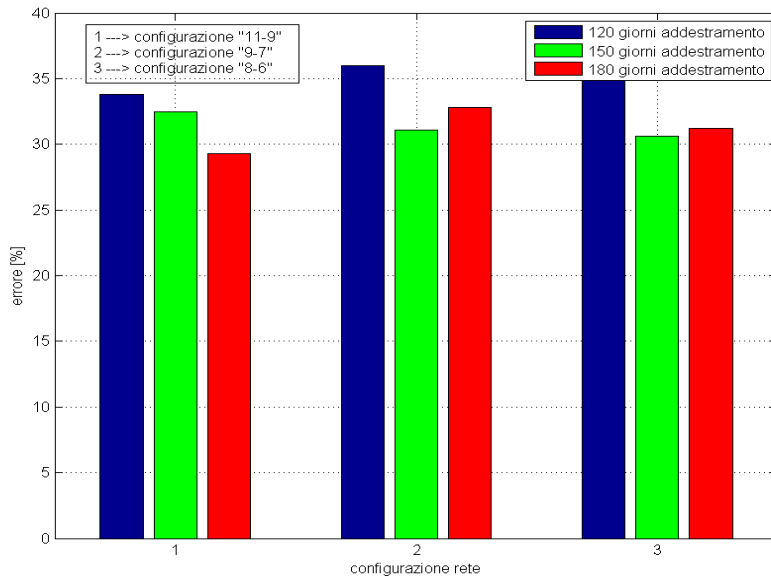


Figura 6.3: Andamento dell'errore in funzione della configurazione di rete

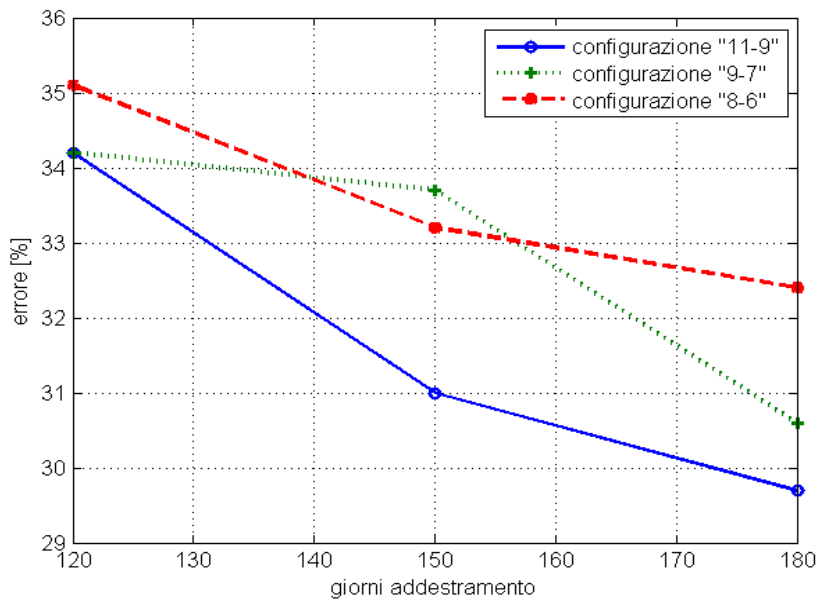


Figura 6.4: Andamento dell'errore in funzione dei giorni di addestramento

La figura 6.3 mostra l'andamento dell'errore in relazione alla configurazione di rete scelta. Il grafico non evidenzia un particolare andamento dell'errore in relazione alla complessità della rete. Verranno dunque valutati singolarmente i casi e verrà scelta la configurazione di rete che, nello specifico caso, consente di ottenere i risultati migliori.

La figura 6.4, evidenzia invece la diminuzione dell'errore all'aumentare dei giorni di addestramento della rete neurale, indipendentemente dalla configurazione di rete scelta.

6.1.2 Variazioni dataset

Le simulazioni descritte nella sezione 6.1.1 non hanno portato a risultati pienamente soddisfacenti. Si procede dunque tentando di modificare il dataset in modo da ottenere previsioni migliori. Di seguito vengono elencate e motivate le scelte fatte con i relativi risultati.

6.1.2.1 Riduzione del numero di input

Nelle precedenti simulazioni sono state utilizzate tutte le informazioni a disposizione relative alle previsioni meteorologiche. Ipotizzando che alcune di queste informazioni (e.g. l'intensità del vento) possa aver influenzato negativamente la risposta della rete neurale, si è pensato di eliminare alcune informazioni e di fornire in input alla rete un set di dati costruito come segue:

- Data
- Ora
- Irraggiamento orario previsto
- Potenza oraria misurata
- affidabilità oraria
- irraggiamento teorico orario

Le configurazioni scelte e i risultati ottenuti (errore calcolato con la formula 4.25) sono riassunti nella tabella 6.4

6 Processing

1° strato	2° strato	iterazioni	1° giorno add.	giorni add.	$WMAE\%$
8	7	5000	60	210	32.1
9	6	2000	60	210	32.5
9	6	5000	60	210	33.8
9	7	2000	1	210	39.6
10	7	2000	1	210	36.5

Tabella 6.4: Simulazioni dataset ridotto

L'analisi dei risultati ottenuti mostra che la riduzione delle informazioni in input alla rete neurale non ha portato miglioramenti in termini di riduzione dell'errore nella previsione dell'energia prodotta dall'impianto.

6.1.2.2 Sostituzione valori mancanti

Analizzando i dati misurati dell'impianto fotovoltaico si è notato che nei giorni compresi tra il 28/10/2012 e il 12/12/2012 il sensore d'irraggiamento non ha comunicato i valori. Per come è stato costruito il dataset, e in particolare la colonna affidabilità, questi dati non vengono utilizzati ai fini dell'apprendimento. Per evitare di non utilizzare una parte consistente dei dati, si è pensato di sostituire i valori mancanti con i corrispondenti giorni relativi al 2011.

Sono state effettuate alcune simulazioni e i dati vengono riportati di seguito in tabella.

1° strato	2° strato	iterazioni	1° giorno training.	giorni training.	$WMAE\%$
9	7	3000	1	180	32.6
9	7	5000	1	150	37.3
10	6	10000	1	150	34.3
10	8	3000	1	150	33.2
11	8	5000	1	150	33.1
11	8	5000	1	180	31.2
11	9	3000	1	210	30.2
11	9	2000	1	210	33.2
11	9	4000	1	210	33.2
11	9	5000	1	210	31.7

Tabella 6.5: Simulazioni con dati mancanti sostituiti

La modifica del dataset non ha permesso di ottenere risultati soddisfacenti ed in particolare il superamento della soglia di errore del 30%.

6.1.2.3 Incremento giorni dataset

Fino ad ora sono stati utilizzati i dati relativi al 2012. Si procede col tentativo di aumentare i dati in input alla rete neurale aggiungendo anche le previsioni relative al periodo che va dal 01/01/2013 al 23/03/2013. In questo modo si hanno abbastanza dati a disposizione per passare da alcuni mesi di addestramento a un intero anno (366 giorni del 2012).

La scelta è stata fatta pensando di incrementare fortemente i giorni dedicati al training della rete neurale in modo da migliorarne l'apprendimento e quindi le previsioni. Le configurazioni scelte e i risultati ottenuti (errore calcolato con la formula 4.25) sono riassunti nella tabella 6.6

1° strato	2° strato	iterazioni	1° giorno training.	giorni training.	$WMAE_{\%}$
9	7	2000	1	366	32.3
9	7	3000	1	366	32.6
9	7	5000	1	366	37.3
10	6	10000	1	366	34.3
10	8	3000	1	366	33.2
11	8	2000	1	366	33.2
11	8	3000	1	366	33.4
11	8	5000	1	366	33.1
11	8	5000	1	366	31.2
11	9	3000	1	366	30.2
11	9	2000	1	366	33.2
11	9	4000	1	366	33.2
11	9	5000	1	366	31.7
12	9	2000	1	366	33.2
12	9	3000	1	366	33.1
12	10	5000	1	366	33.1
13	10	3000	1	366	37.2
13	10	4000	1	366	32.8
13	10	8000	1	366	30.5
13	10	10000	1	366	30.5
13	10	20000	1	366	32.6
14	10	5000	1	366	34.9
14	11	5000	1	366	31.1
15	12	5000	1	366	40.7
15	13	5000	1	366	32.1

Tabella 6.6: Simulazioni giorni dataset incrementati

6 Processing

Nonostante il deciso aumento dei giorni di addestramento non sono stati riscontrate particolari riduzioni dell'errore.

6.1.2.4 Riduzione giorni dataset

Come descritto nella sezione 6.1.2.2 per l'anno 2012 vi è un periodo in cui mancano dei valori. Al contrario di quanto fatto prima, invece di sostituire tali valori, si è proceduto con la riduzione del dataset. Vengono dunque utilizzati i dati relativi ai primi 300 giorni dell'anno 2012. Nella tabella 6.7 vengono riportati i risultati delle simulazioni fatte.

1° strato	2° strato	iterazioni	1° giorno training.	giorni training.	$WMAE_{\%}$
9	7	2000	1	210	26
9	7	2000	1	240	24.7
9	7	2000	1	270	26.6
9	7	2000	60	210	35.4
9	7	2000	60	240	32.2
9	7	2000	60	270	30.2
9	7	3000	1	120	42.4
9	7	3000	1	180	34.9
9	7	3000	1	210	26.6
9	7	3000	60	180	34.2
9	7	3000	60	210	33.4
9	7	3000	60	240	31.1
9	7	5000	1	120	52.7
9	7	5000	1	150	45.3
9	7	5000	1	180	34.2
9	7	5000	1	210	29.1
9	7	5000	1	240	26.8
9	7	5000	1	270	29.3
9	7	5000	1	298	28.3
9	7	5000	60	120	32.4
9	7	5000	60	150	29.4
9	7	5000	60	180	29.8
9	7	5000	60	210	28.6
9	7	5000	60	240	30.2

Tabella 6.7: Simulazioni giorni dataset ridotti

La riduzione del numero di campioni oggetto di previsione ha portato, con alcune configurazioni di rete, a un notevole abbassamento dell'errore, che in alcuni casi si

attesta su valori abbondantemente al di sotto del 30%.

6.1.2.5 Incremento soglia minima ammissione irraggiamento orario previsto

Nei criteri di ammissione descritti nella sezione 5.2 è stata introdotta una soglia sull'irraggiamento orario previsto che prevedeva di escludere i valori al di sotto di $10 \frac{W}{m^2}$ ponendo a 0 il valore corrispondente del parametro $valmh$, che definisce quali dati vengono presi in considerazione nella fase di addestramento e test della rete neurale, e quali no. L'incremento della soglia utile ($50 \frac{W}{m^2}$) per l'ammissione del dato ha l'obiettivo di escludere ulteriormente le ore in cui l'irraggiamento è molto basso e non vi è trasformazione dell'energia da parte dell'impianto fotovoltaico. L'errore assoluto che si commette nella previsione è molto basso ma non lo è in termini relativi. Si procede dunque con le simulazioni ed in tabella 6.8 vengono riportati i risultati ottenuti.

1° strato	2° strato	iterazioni	1° giorno training.	giorni training.	WMAE%
9	7	2000	1	210	26.6
9	7	2000	1	240	25.3
9	7	2000	1	270	23.9
9	7	2000	60	240	28.4
9	7	3000	30	210	27.9
9	7	3000	90	210	28.5
9	7	4000	30	240	30.1
9	7	4000	60	210	26.9
9	7	4000	90	210	25.5
9	7	5000	90	210	26.1
9	7	8000	90	210	29
10	6	3000	90	210	28.5
11	9	2000	90	210	28.6
11	9	5000	90	210	27.5

Tabella 6.8: Simulazioni soglia minima irraggiamento previsto incrementata

L'ultima modifica del dataset proposta ha portato a una stabilizzazione dell'errore su valori soddisfacenti e nella maggior parte dei casi ampiamente al di sotto del 30%.

La campagna di simulazioni condotta ha indagato il comportamento della rete neurale in 96 diverse configurazioni. Di seguito vengono riportati gli output restituiti dalla rete neurale addestrata e testata con le configurazioni che, da una prima analisi, hanno mostrato un migliore comportamento in relazione all'entità dell'errore commesso. Le figure 6.5, 6.6 e 6.7 riassumono bene le scelte fatte per ciò che

6 Processing

riguarda i parametri di rete scelti. Per le tre le simulazioni sono state apportate le modifiche descritte nella sezione 6.1.2.5, ed in particolare si ricorda che è stata innalzata a $50 \frac{W}{m^2}$ la soglia minima di accettabilità dell'irraggiamento misurato.

E' importante aggiungere che:

- La simulazione in figura 6.5 prevede:
 - 70% dei dati disponibili utilizzati per la fase di addestramento della rete
 - 30% dei dati disponibili utilizzati per la fase di test della rete
- La simulazione in figura 6.6 prevede:
 - 80% dei dati disponibili utilizzati per la fase di addestramento della rete
 - 20% dei dati disponibili utilizzati per la fase di test della rete
- La simulazione in figura 6.7 prevede:
 - 90% dei dati disponibili utilizzati per la fase di addestramento della rete
 - 10% dei dati disponibili utilizzati per la fase di test della rete

Per le tre simulazioni il grafico in figura rappresenta l'andamento della potenza oraria prevista (verde), della potenza oraria misurata (blu) e dell'errore (rosso) relativo alla migliore simulazione tra le 10 effettuate per ciascuna configurazione di rete. Sono state effettuate più simulazioni indipendenti con la medesima configurazione poichè, essendo i metodi neurali dei metodi stocastici, a parità di tutte le condizioni, si ottengono risultati differenti ad ogni trial della simulazione. Nelle figure si riportano i risultati relativi alla migliore simulazione (best) e i valori relativi alla media delle simulazioni.

6.1 Analisi di sensitività

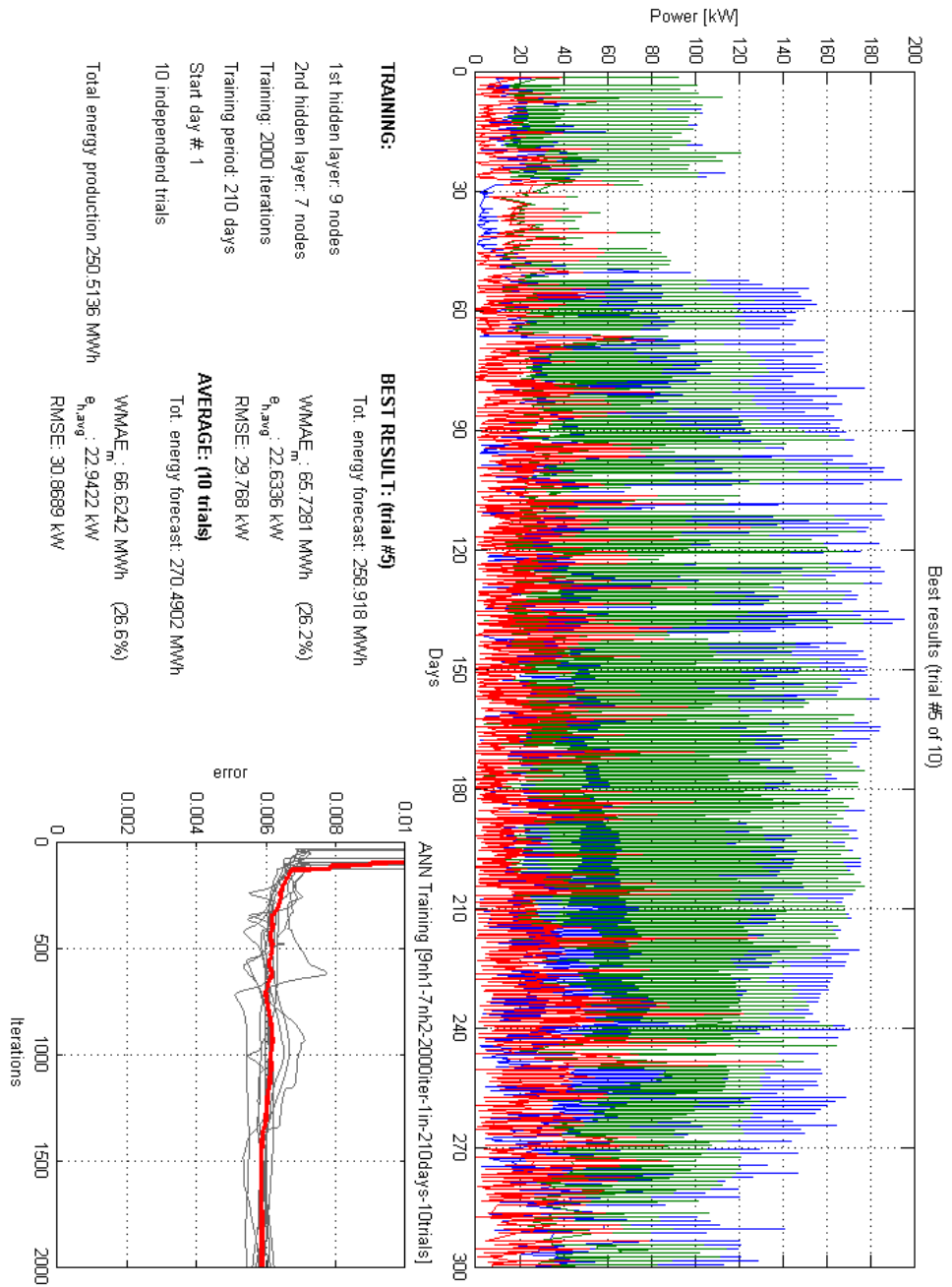


Figura 6.5: Grafico riassuntivo simulazione migliore #1

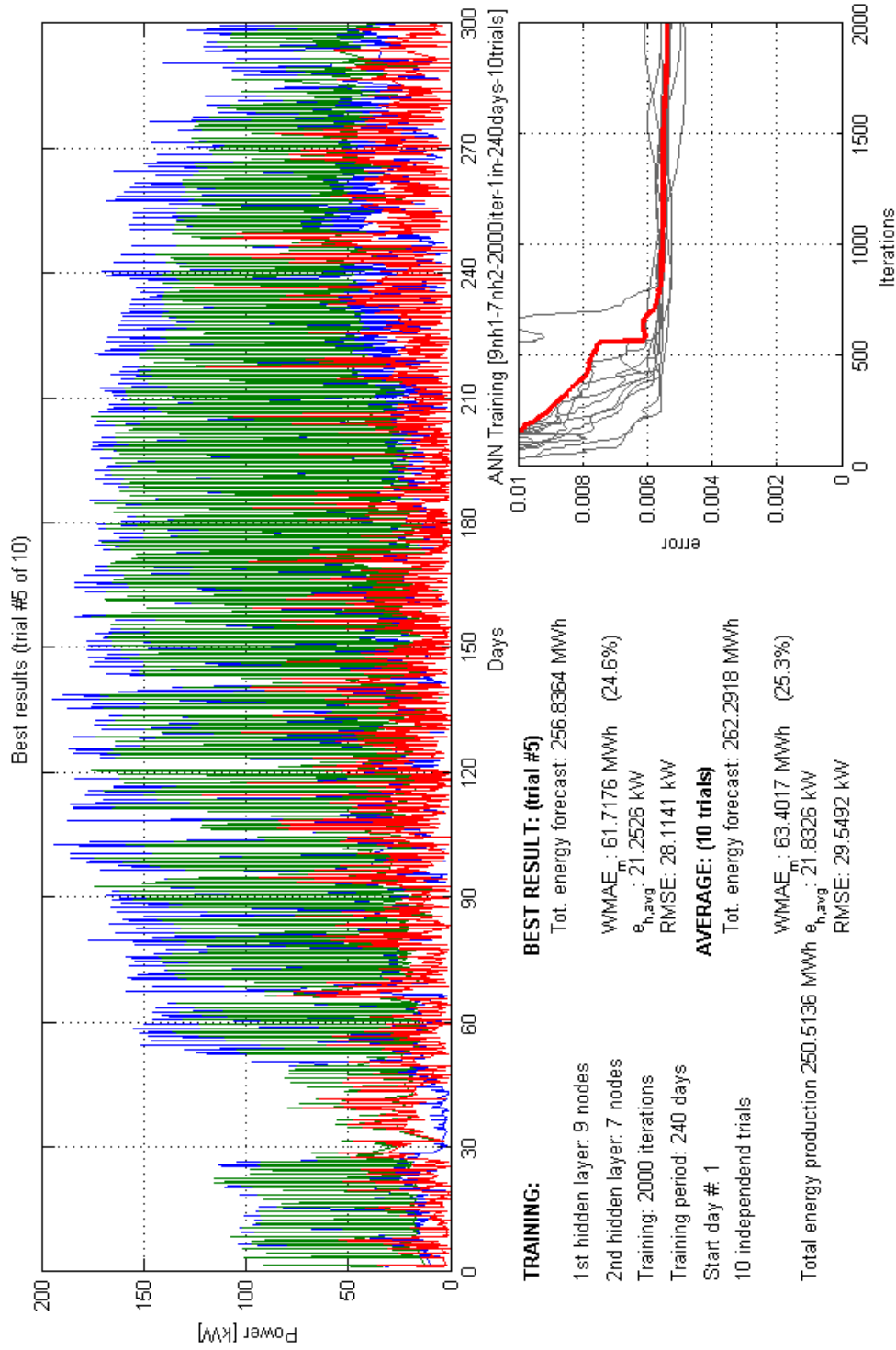


Figura 6.6: Grafico riassuntivo simulazione migliore #2

6.1 Analisi di sensitività

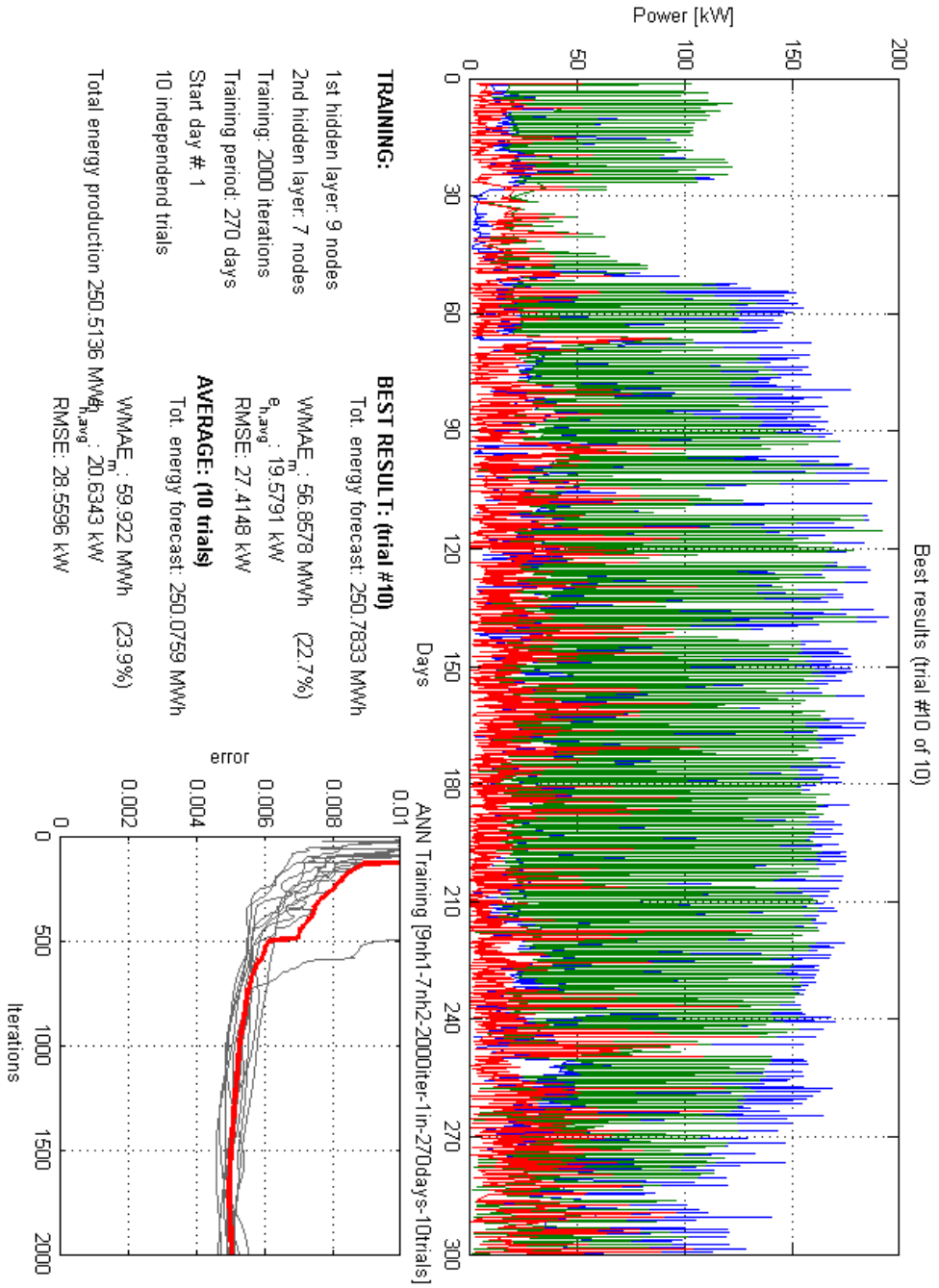


Figura 6.7: Grafico riassuntivo simulazione migliore #3

7 Post-processing

Dopo aver effettuato una consistente campagna di simulazioni variando la configurazione di rete e apportando modifiche al dataset, si ha a disposizione una quantità rilevante di dati da analizzare. La ricerca della migliore configurazione di rete è direttamente legata ai parametri che vengono scelti per valutare la bontà dei risultati ottenuti. L'obiettivo di questo capitolo è quello di confrontare le simulazioni al fine di determinare quale configurazione di rete restituisce i risultati migliori.

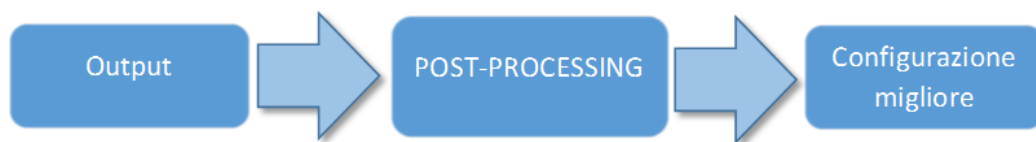


Figura 7.1: Schema post-processing

7.1 Configurazione di rete ottimale

Le definizioni d'errore precedentemente elencate consentono di dare una valutazione oggettiva sulla qualità delle previsioni della potenza immessa nella rete elettrica, ottenute utilizzando la rete neurale.

Con il precedente capitolo è stata effettuata un'analisi di sensitività sui principali parametri che descrivono la rete neurale. Le previsioni migliori e più stabili dal punto di vista dell'errore sono state ottenute tutte con la medesima configurazione di rete:

- 9 neuroni nel primo strato nascosto
- 7 neuroni nel secondo strato nascosto
- 2000 iterazioni

Restano ora da definire gli ultimi due parametri che sono:

- numero di giorni di addestramento
- primo giorno di addestramento

7 Post-processing

Per stabilire il numero di giorni di addestramento ottimale sono state fatte le simulazioni con la rete neurale sopra descritta, variando il numero di giorni di addestramento. I valori del numero di giorni di addestramento analizzati sono:

- 180
- 210
- 240
- 270

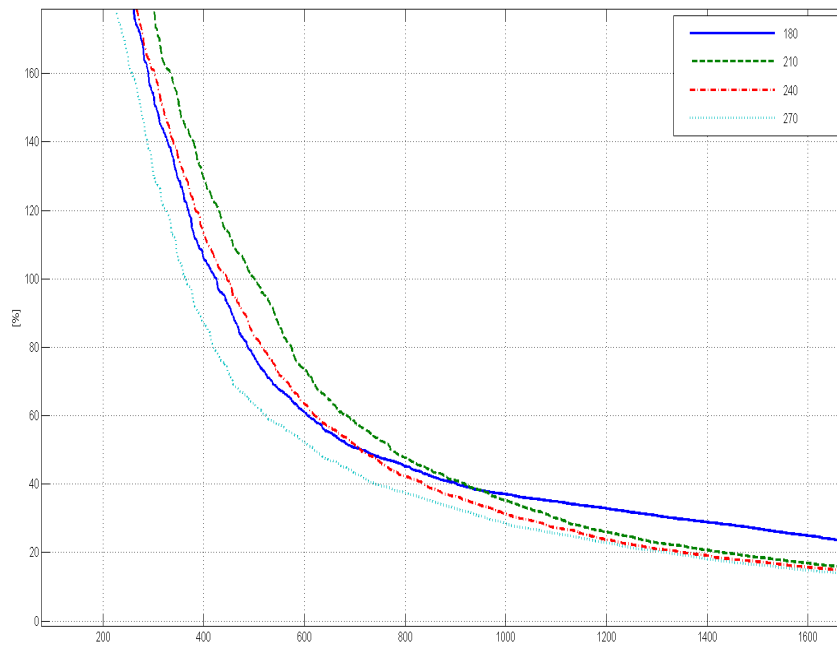


Figura 7.2: Andamento dell'errore percentuale riferito alla potenza misurata, al variare del numero di giorni di addestramento

7.1 Configurazione di rete ottimale

errori [%]	giorni di addestramento			
	180	210	240	270
$e_{\%,max,m}$	2897	3686	3674	3771
$e_{\%,min,m}$	4.26e-02	2.03e-02	1.2e-02	2.07e-02
$e_{\%,max,p}$	291	183	359	329
$e_{\%,min,p}$	4.26e-02	2.03e-02	1.2e-02	2.07e-0.2
$e_{\%,avg,m}$	73.7	80.6	74.5	65.4
$e_{\%,avg,p}$	37.7	29.3	29	29.1
$NMAE_{m\%}$	13.3	11.3	10.8	9.9
$NMAE_{p\%}$	15.6	12.5	11.7	11
$NMAE_{r\%}$	9.1	7.7	7.4	6.8
$WMAE_{\%,m}$	30	25.6	24.3	22.5
$WMAE_{\%,p}$	32.7	23.7	23.2	22.6

Tabella 7.1: Andamento degli errori in funzione del numero di giorni di addestramento

Sono state prese in considerazione 11 definizioni di errore e i risultati sono riportati in tabella 7.1. Appare evidente che la scelta migliore coincida con il numero massimo di giorni di addestramento scelto. In realtà è bene fare ulteriori considerazioni. Considerando che non sempre si è in possesso di un numero elevato di dati storici di produzione dell'impianto fotovoltaico (necessari per l'addestramento della rete neurale) è bene trovare un compromesso tra la necessità di ottenere delle previsioni il più possibile affidabili e la necessità di non dover disporre di una quantità di dati misurati elevata.

La scelta ricade dunque su 210 giorni di addestramento che può risultare un valido compromesso tra le necessità prima descritte.

L'ultimo parametro da impostare è la scelta del primo giorno di addestramento. Sono state valutate due possibilità:

- 1° giorno del dataset (01/01/2012)
- 92° giorno del dataset (01/04/2012)

Le definizioni di errore utilizzate portano a scegliere 1 come valore per il giorno iniziale di addestramento. Probabilmente, facendo iniziare la fase di addestramento l'1 gennaio, la rete neurale ha a disposizione una maggiore variabilità delle condizioni meteorologiche che le consentono di generalizzare maggiormente il problema, riuscendo a fornire previsioni più accurate. In tabella 7.2 vengono riportati i risultati ottenuti.

7 Post-processing

errori [%]	primo giorno addestramento	
	1°	92°
$e_{\%,max,m}$	3686	3346
$e_{\%,min,m}$	2.03e-02	5.53e-03
$e_{\%,max,p}$	183	793
$e_{\%,min,p}$	2.03e-02	5.53e-03
$e_{\%,avg,m}$	80.6	73.7
$e_{\%,avg,p}$	29.3	37.7
$NMAE_{m\%}$	11.3	13.3
$NMAE_{p\%}$	12.5	15.6
$NMAE_{r\%}$	7.7	9.1
$WMAE_{\%,m}$	25.6	30
$WMAE_{\%,p}$	23.7	32.7

Tabella 7.2: Confronto errori per due diversi giorni di inizio addestramento

Le ultime scelte fatte consentono di definire la configurazione di rete che nel caso in oggetto è risultata la più performante:

- 9 neuroni nel primo strato nascosto
- 7 neuroni nel secondo strato nascosto
- 2000 iterazioni
- 210 giorni di addestramento
- 1° giorno per l'inizio dell'addestramento

Con la configurazione di rete scelta sono stati ottenuti i seguenti risultati:

$e_{\%,avg,m}$	$e_{\%,avg,p}$	$e_{\%,max,m}$	$e_{\%,max,p}$	$e_{\%,min,m}$	$e_{\%,min,p}$
80.6	29.3	3686	183	2.03e-02	2.03e-02
$NMAE_{m\%}$	$NMAE_{p\%}$	$NMAE_{r\%}$	$WMAE_{\%,m}$	$WMAE_{\%,p}$	
11.3	12.5	7.7	25.6	23.7	

Tabella 7.3: Errore previsione con la configurazione ottimale di rete scelta

7.2 Analisi errore

Per chiarire meglio l'entità degli errori di previsione commessi può risultare utile il grafico in figura 7.3. E' stato calcolato l'errore percentuale riferito alla potenza

prevista $e_{\%,p}$ e l'errore percentuale riferito alla potenza misurata $e_{\%,m}$ e i valori sono stati ordinati in ordine decrescente.

Per un ristretto numero di campioni si commette un errore di previsione della potenza immessa in rete decisamente elevato. Nel calcolo dell'errore medio questi campioni andranno a incidere fortemente dato che, ad esempio, il massimo errore percentuale riferito alla potenza misurata si attesta sul 3800% (Per visualizzare in modo chiaro l'andamento dell'errore, in figura 7.3 non è stato possibile visualizzare l'errore massimo poichè sensibilmente distante dagli altri valori).

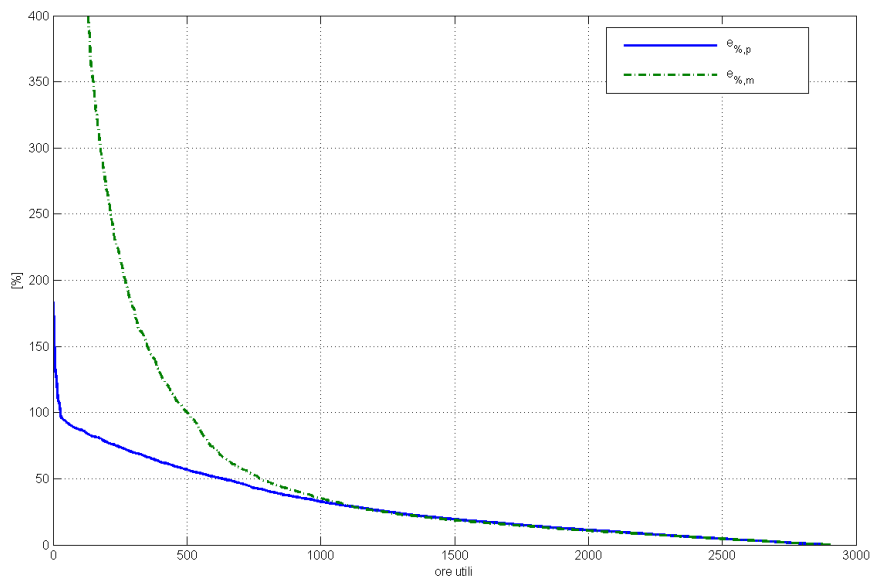


Figura 7.3: Andamento dell'errore in ordine decrescente

E' interessante calcolare il numero di campioni per i quali la rete neurale ha dato una previsione di potenza accettabile dal punto di vista dell'errore. Poniamo dunque due soglie di $e_{\%,m}$ pari a 20 % e 10%. La scelta del primo valore coincide, non a caso, con la franchigia introdotta nella Deliberazione 281/2012/R/efr dell'Autorità per l'energia elettrica e il gas, entro la quale non vengono valorizzati gli sbilanciamenti sulla rete elettrica. E' stato calcolato che:

- il numero di campioni orari con errore inferiore al 20% è pari a 1473 (51% dei campioni utili)
- il numero di campioni orari con errore inferiore al 10% è pari a 853 (29% dei campioni utili)

7 Post-processing

Dopo aver valutato l'andamento medio degli errori è opportuno fare delle brevi considerazioni su casi particolari in cui l'errore raggiunge picchi anomali. In figura 7.4 si evidenzia il caso dell'errore percentuale riferito alla potenza misurata. Il buon andamento dell'errore, che registra valori molto bassi, è interrotto da un picco anomalo che coincide con una brusca diminuzione della potenza misurata. Nel caso descritto in dettaglio si può notare che per il campione 1082 la rete neurale è stata in grado di prevedere la diminuzione di potenza, ma avendone sovrastimato il reale valore è stato comunque commesso un errore di previsione decisamente elevato.

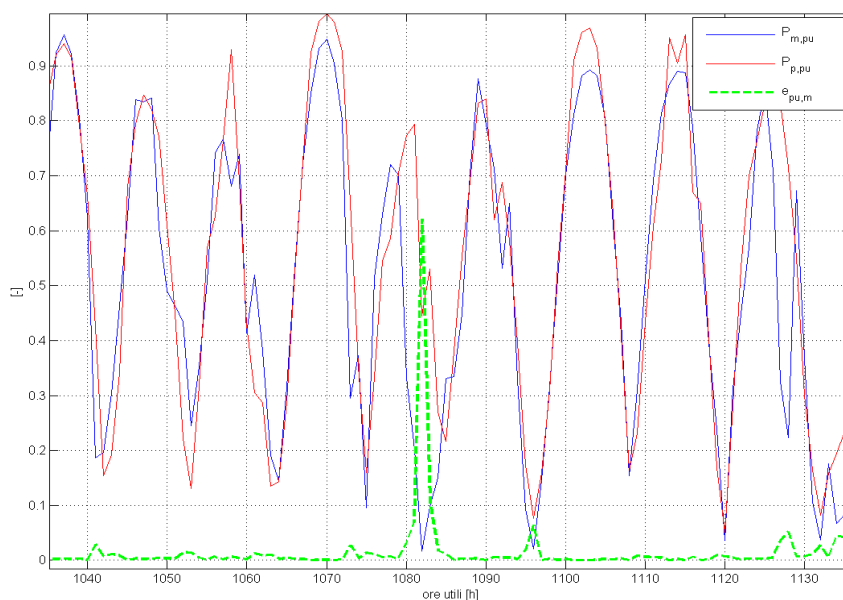


Figura 7.4: Dettaglio errore percentuale riferito alla potenza misurata

In figura 7.6 viene inoltre mostrato contemporaneamente sia l'andamento dell'errore percentuale riferito alla potenza misurata che quello riferito alla potenza prevista. Il grafico mostra che nelle giornate particolarmente nuvolose entrambi gli errori presentano picchi frequenti. Ciò è dovuto al fatto che generalmente la rete neurale riesce a prevedere nell'arco della giornata variazioni consistenti della potenza immessa nella rete elettrica, ma molto spesso, essendoci delle brusche variazioni, la rete neurale posticipa o anticipa queste variazioni. Può capitare dunque che durante una giornata, la rete neurale preveda che la produzione non sarà regolare, ma dato che i picchi minimi e massimi della potenza prevista e della potenza misurata non coincidono, si rilevano valori di errore percentuale orario molto elevati.

7.2 Analisi errore

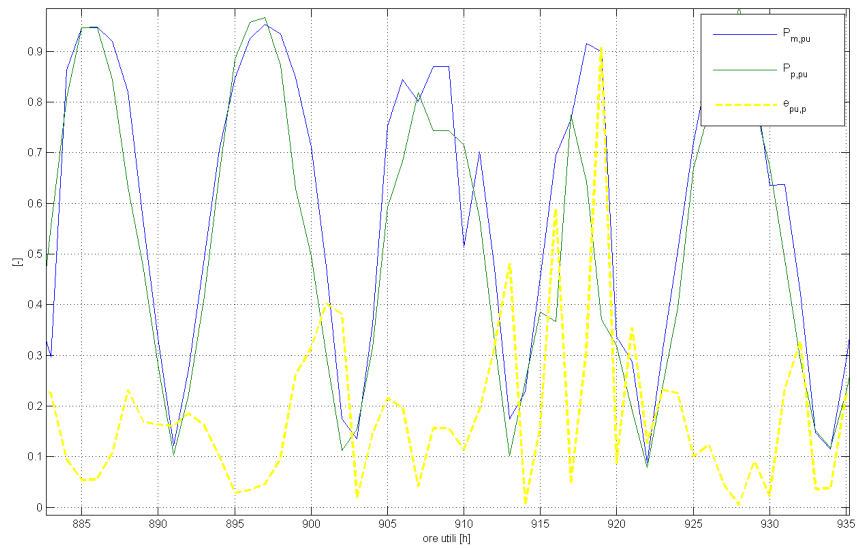


Figura 7.5: Dettaglio errore percentuale riferito alla potenza prevista

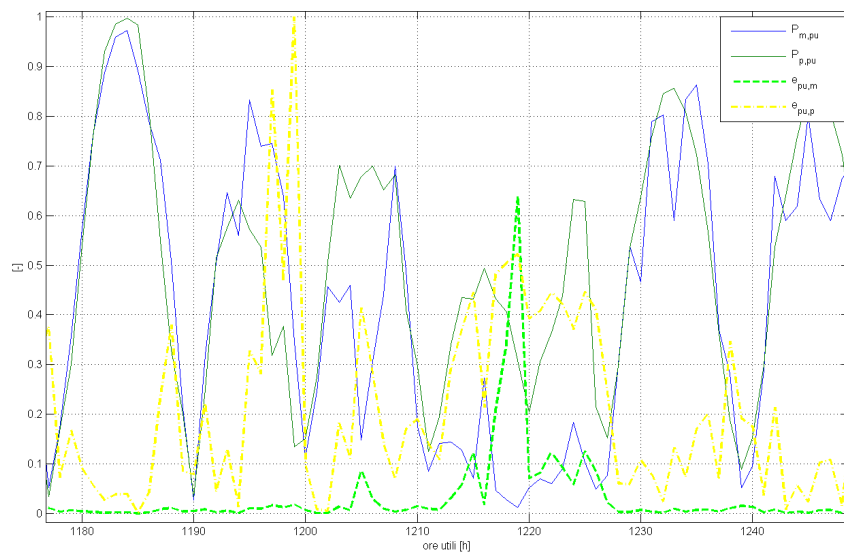


Figura 7.6: Dettaglio errore percentuale riferito alla potenza prevista e alla potenza misurata

7.3 Confronto tra la precisione dei dati meteorologici e la previsione della rete neurale

La previsione fornita dalla rete neurale è fortemente dipendente dalla qualità delle previsioni meteorologiche utilizzate per la fase di addestramento ed in particolare modo dal dato sull'irraggiamento previsto. Per valutarne la precisione viene utilizzata la seguente definizione di errore:

- errore medio percentuale riferito all'irraggiamento previsto $e_{\%,avg,f}$:

$$e_{\%,avg,f} = \frac{1}{N} \sum_1^N \frac{|GHI_{fh} - GHI_{mh}|}{GHI_{fh}} \cdot 100 \quad (7.1)$$

dove N sono le ore in cui l'irraggiamento previsto è diverso da 0. Solo per queste ore è stato calcolato l'errore (questa scelta serve ad evitare di includere nella media anche i dati riferiti alla notte, per i quali siamo certi di commettere un errore nullo. Si avrebbe una sottostima dell'errore). Per $e_{\%,avg,f}$ si prevede una variante:

$$e_{\%,avg,f10} = \frac{1}{N} \sum_1^N \frac{|GHI_{fh} - GHI_{mh}|}{GHI_{fh}} \cdot 100 \quad (7.2)$$

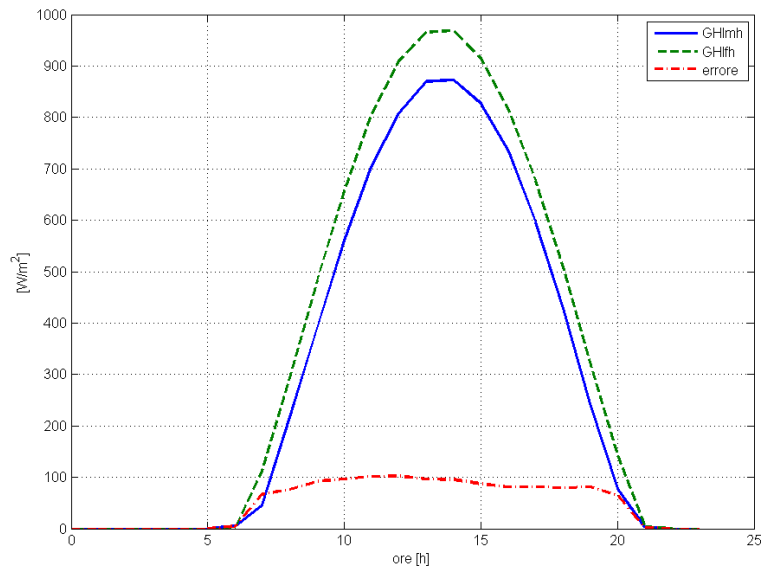
La 7.2 è stata utilizzata per non considerare, nel calcolo dell'errore, le ore in cui l'irraggiamento è basso (inferiore a $10 \frac{W}{m^2}$). In queste ore capita spesso di commettere degli errori relativi elevati ma che in realtà, in termini assoluti, sono bassi. Utilizzando le due definizioni di errore precedentemente descritte, applicate a due giornate campione (una giornata serena ed una nuvolosa), si ottengono i valori riportati in tabella:

GIORNO	$e_{\%,avg,f}$ [%]	$e_{\%,avg,f10}$ [%]
05/04/2012	91.3	75.2
18/07/2012	21.6	20

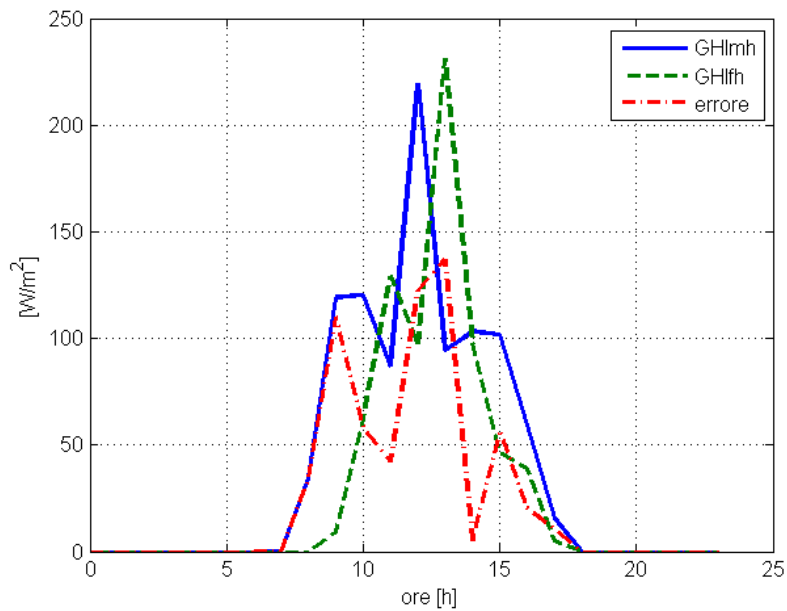
Tabella 7.4: Errore dati previsionali irraggiamento

La differenza tra l'irraggiamento previsto e quello misurato è particolarmente rilevante nella giornata nuvolosa.

7.3 Confronto tra la precisione dei dati meteorologici e la previsione della rete neurale



(a) giornata serena



(b) giornata nuvolosa

Figura 7.7: Confronto tra irraggiamento previsto e irraggiamento misurato con il relativo errore

7 Post-processing

In ultimo si vuole confrontare l'errore delle previsioni meteorologiche, ed in particolare il dato sull'irraggiamento, con l'errore commesso dalla rete neurale nella previsione dell'energia prodotta. Si vuole dunque valutare il comportamento della rete neurale a fronte di un certo errore delle previsioni meteorologiche.

Per fare ciò si definiscono due errori che valutano lo scostamento tra l'irraggiamento previsto e l'irraggiamento misurato:

- Errore assoluto medio normalizzato $NMAE_{f\%}$, basato sul massimo valore di irraggiamento misurato:

$$NMAE_{f\%} = \frac{1}{N} \cdot \sum_{h=1}^N \frac{|GHImh - GHIfh|}{\max(GHImh)} \cdot 100 \quad (7.3)$$

- Media ponderata dell'errore assoluto $WMAE_{f\%}$, basato sulla somma dell'irraggiamento misurato:

$$WMAE_{f\%} = \frac{\sum_{h=1}^N |GHImh - GHIfh|}{\sum_{h=1}^N GHImh} \cdot 100 \quad (7.4)$$

Si applicano le due definizioni di errore 7.3 e 7.4 per le sole ore in cui l'irraggiamento previsto e l'irraggiamento misurato sono diversi da 0 e i risultati sono riportati in tabella 7.5.

$NMAE_{f\%}$	$WMAE_{f\%}$
13.1	37.7

Tabella 7.5: Errori relativi all'irraggiamento previsto rispetto all'irraggiamento misurato

$NMAE_{m\%}$	$NMAE_{p\%}$	$NMAE_{r\%}$	$WMAE_{m\%}$	$WMAE_{p\%}$
11.3	12.4	7.74	25.6	23.7

Tabella 7.6: Errori relativi alla previsione della rete neurale

Grazie all'utilizzo di diversi parametri ambientali e alla caratteristica di profilo Clear Sky (irraggiamento teorico) l'errore finale commesso sulla previsione di potenza risulta inferiore a quello commesso con riferimento all'irraggiamento fornito dal servizio meteo.

8 Conclusioni

In questo lavoro è stata valutata l'utilità e l'affidabilità delle tecniche di soft computing applicate alla previsione della produzione di un impianto alimentato da fonti rinnovabili non programmabili e in particolare il caso di un impianto fotovoltaico. L'utilizzo di queste tecniche di previsione può essere esteso anche ad altri tipi di impianti a fonte rinnovabile caratterizzati dalla stessa difficoltà di programmare la produzione. Si possono trarre le seguenti considerazioni.

La prima riguarda l'importanza della fase di pre-processing. Questa fase, onerosa in termini di tempo di realizzazione, consiste nella preparazione dei dati ai fini dell'apprendimento della rete neurale artificiale. Oltre a trovare ed eliminare eventuali dati non corretti da un punto di vista fisico, è necessario un allineamento temporale delle previsioni meteorologiche con i dati misurati sull'impianto. Fornire in input alla rete neurale dei dati coerenti è una base necessaria per la buona riuscita dell'attività di forecasting dell'energia da fonte rinnovabile.

I risultati confermano poi che vi è una relazione diretta tra il numero di giorni di addestramento e la precisione della previsione della rete neurale: all'aumentare dei giorni disponibili per la fase di apprendimento, l'errore commesso dalla rete decresce notevolmente. La disponibilità di dati storici riferiti alla produzione dell'impianto fotovoltaico non è però scontata, come ad esempio nel caso di impianti di recente installazione: il numero di giorni di addestramento scelto nel presente lavoro è stato dunque il risultato di un compromesso tra la necessità di ottenere buoni risultati e il ricorso a un numero non troppo ampio di dati storici.

Inoltre, confrontando la bontà delle previsioni della rete neurale con la qualità delle previsioni meteorologiche si nota che la rete neurale riesce a fornire una risposta affidabile nonostante la non elevata qualità dei dati meteorologici (in particolare il dato sull'irraggiamento). Ciò avviene grazie alla capacità della rete neurale di utilizzare tutti i dati in input a disposizione e di attribuire ad essi il giusto peso.

Infine lo strumento previsionale qui presentato potrà essere utile agli operatori del settore energetico in vista degli obblighi imposti dall'Autorità di prevedere l'energia immessa nella rete elettrica nazionale. Per un consistente numero di campioni analizzati, l'errore di previsione della rete neurale è infatti al di sotto delle soglie stabilite nelle ultime deliberazioni dell'Autorità.

Gli sviluppi futuri riguardano l'affinamento della tecnica di previsione con ricorso a una finestra temporale di apprendimento più ampia, l'utilizzo di dati meteorolo-

8 Conclusioni

gici più affidabili e l'estensione della tecnica di previsione ad impianti che utilizzano fonti primarie differenti.

Bibliografia

- [1] Direttiva 2009/28/CE del Parlamento europeo e del Consiglio del 23 aprile 2009
- [2] Global market outlook for photovoltaics 2013-2017, Epia, 2013
- [3] Dati statistici sull'energia elettrica in Italia, www.terna.it, 2012
- [4] Fuel Mix Disclosure: determinazione del mix energetico per gli anni 2011-2012, www.gse.it, Giugno 2013
- [5] G. Silvestrini, Rided, i primi 10 anni e il prossimo decennio, Rided, 20 settembre 2013
- [6] Rapporto statistico 2012 Solare Fotovoltaico, GSE, 2013
- [7] Rapporto Energia e Ambiente - L'analisi 2009-2010, Enea, 2012
- [8] Documento per la consultazione 35/2012/R/efr, Orientamenti in materia di regolazione del servizio di dispacciamento da applicarsi alle unità di produzione di energia elettrica, con particolare riferimento a quelle non programmabili, AEEG, 9 Febbraio 2012
- [9] Deliberazione 281/2012/R/efr, Revisione del servizio di dispacciamento dell'energia elettrica per le unità di produzione di energia elettrica alimentate da fonti rinnovabili non programmabili, AEEG, 5 luglio 2012
- [10] Deliberazione 462/2013/R/eel, Disposizioni in materia di dispacciamento delle fonti rinnovabili non programmabili in seguito alle ordinanze del Consiglio di Stato, sezione sesta, nn.3566,3567 e 3568 dell'11 settembre 2013, 17 ottobre 2013
- [11] D. Floreano, C. Mattiussi, Manuale sulle reti neurali, il Mulino, ISBN 88-15 08504-1, 2002
- [12] C. Monteiroa, L. A. Fernandez-Jimenezb, I. J. Ramirez-Rosadoc, A. Muñoz Jimenezb, P. M. Lara-Santillanb, Short-term forecasting models for photovoltaic plants: analytical versus soft-computing techniques, Hindawi mathematical problems in engineering, 2013

Bibliografia

- [13] B. Lazzerini, Introduzione alle Reti Neurali
- [14] C. Gallo, Reti Neurali Artificiali: Teoria ed Applicazioni, 2007
- [15] V. Kostylev, A. Pavloski. Solar power forecasting performance-towards industry standards. In 1st Int. Workshop on the integration of solar power into power systems, Aarhus, Denmark, 2011
- [16] E. Ogliari, F. Grimaccia, S. Leva, M. Mussetta, Hybrid predictive models for accurate forecasting in PV systems. *Energies*, 2013.
- [17] Relazione tecnica alla deliberazione 5 luglio 2012 281/2012/R/efr
- [18] Documento per la consultazione 35/2012/R/efr, Oneri di sbilanciamento fonti rinnovabili non programmabili – Analisi risultati simulazioni, 22 giugno 2012
- [19] R.Ulbricht, U. Fischer, W. Lehner, H. Donker, First Steps Towards a Systematical Optimized Strategy for Solar Energy Supply Forecasting, European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases, Praga, 23-27 Settembre 2013