

65° CONGRESSO NAZIONALE ATI

ANALISI STATISTICA DEI REGIMI DI VENTO E DI PRODUCIBILITÀ DI UN IMPIANTO EOLICO SITO IN SICILIA. STIMA DI PREVISIONE DELLA PRODUCIBILITÀ EOLICA ATTRAVERSO RETI NEURALI

Marco Beccali¹, Simona Culotta¹, Josè Massimiliano Galletto^{*1}, Antonino Macaione²

Email: galletto@dream.unipa.it

¹Dipartimento di Ricerche Energetiche ed Ambientali (DREAM), Università degli studi di Palermo
Viale delle scienze, Edificio 9, 90128 Palermo

²Asja Ambiente Italia S.p.A.

SOMMARIO

Tra le fonti energetiche rinnovabili quella generata dal vento ha avuto un forte crescita legata ai costi accessibili della tecnologia ed alla alta remuneratività degli investimenti. L'aumento di produzione di energia eolica, annoverata fra le fonti non programmabili, rende però necessario lo sviluppo di strumenti finalizzati alla previsione della produttività di un parco eolico. L'analisi che verrà illustrata si basa sulle misurazioni delle grandezze caratteristiche del vento, quali intensità e direzione, rilevate nel corso di una campagna anemometrica e sulle registrazioni della potenza effettiva prodotta in un parco eolico in Sicilia. Dopo una prima fase in cui i dati sono stati sottoposti ad un accurato pretrattamento attraverso il quale sono stati filtrati gli outliers ed eliminati quelli registrati in maniera erronea dai data loggers si è svolta l'analisi statistica. Attraverso reti neurali, è stata effettuata la previsione della producibilità di breve periodo (*short forecasting*) e su diverse scale temporali utilizzando anche dati provenienti da stazioni anemometriche non strettamente collegate al sito in esame.

INTRODUZIONE

Tra le risorse energetiche rinnovabili, l'energia del vento diventa sempre più attraente così in questo ultimo decennio la costruzione dei siti eolici è aumentata nei luoghi dove le condizioni climatiche, orografiche ed ambientali ne hanno permesso lo sfruttamento.

Nella realizzazione di un impianto eolico il fattore determinante è la scelta del sito di installazione; da ciò dipende la scelta della tipologia di turbina da installare, la produzione dell'impianto e quindi la validità del progetto.

Una buona conoscenza delle caratteristiche dei regimi dei venti è il pre-requisito per un'efficiente pianificazione e per l'attuazione di qualsiasi progetto di energia eolica [1,2].

In un futuro più o meno prossimo, la presenza di una elevata potenza eolica nella rete elettrica siciliana causerà dei problemi, se non affrontata con i giusti tempi e strumenti, al network management. In previsione di ciò il Gestore Servizi Energetici (GSE) ha elaborato una procedura tecnica per il miglioramento della prevedibilità delle immissioni dell'energia elettrica prodotta da impianti alimentati da fonti rinnovabili non programmabili basata su modelli neurali e fisici [3].

Le strategie che vengono affrontate nel presente studio non

riguardano interventi infrastrutturali di potenziamento della capacità di trasmissione della rete, o di collegamento dell'isola con la parte continentale dell'Italia, bensì mirano ad essere complementari ai tools di gestione dei sistemi di potenza.

E' in questo ambito che si collocano le potenzialità degli algoritmi predittivi basati sulle cosiddette reti neurali [4]. Sfruttando la capacità di adattamento e apprendimento delle reti neurali è possibile predire degli eventi a partire da altri apparentemente non correlati.

Questo può avvenire attraverso la creazione di sistemi che a partire dalla raccolta dati di variabili ambientali (velocità e direzione media del vento, temperatura, pressione, umidità) siano capaci di stimare la producibilità di breve periodo [5] di centrali eoliche al fine di sviluppare modelli di gestione della rete elettrica più evoluti ed affidabili. Per valutare le potenzialità di questi strumenti si è sviluppato un caso studio in cui l'utilizzo di reti neurali è finalizzato alla previsione della producibilità dell'impianto in esame mediante la correlazione dei dati storici di produzione e dei dati di vento provenienti da reti private e pubbliche (S.I.A.S. – Servizio Informativo Agrometeorologico Siciliano ed CNMCA - Centro Nazionale di Meteorologia e Climatologia Aeronautica).

L'utilizzo di dati non strettamente collegati (per posizione e altezza di rilevamento) alla località indagata rappresenta una

novità nel panorama del settore eolico. È infatti noto come, soprattutto nelle fasi di *micrositing*, i dati di vento, necessari per la stima della producibilità, vengano raccolti nelle aree più prossime a quella d'impianto ed ad altezze comparabili con quelle degli aerogeneratori (tipicamente 50 m).

ESPERIENZE DI PREVISIONE TEMPORALI CON RETI NEURALI

In letteratura non sono numerosi gli esempi di previsione di dati di vento e rari gli studi di producibilità di impianti eolici mediante reti neurali.

Ad esempio Lopez in [6] suggerisce un metodo per la stima della velocità media annua del vento in un sito ad orografia complessa, utilizzando una rete MLP (*Multy Layer Perceptron*) utilizzando dati di 5 stazioni anemometriche (4 di input ed 1 di output o di target) con scansione temporale ogni 10'.

Uno dei problemi osservati in questo lavoro riguarda la fase di apprendimento ed è rappresentato dall'*overfitting* che si manifesta quando l'errore osservato nel training set è inferiore a quello osservato quando nuovi dati vengono introdotti nella rete.

In questo studio è stata condotta una analisi di dati di vento confrontando la velocità media, la varianza, il coefficiente di variazione e la direzione prevalente.

Per la scelta del numero di stazioni vicine sono state utilizzate tre diverse opzioni:

- le stazioni erano ai vertici di un triangolo;
- le stazioni erano posizionate lungo la direzione prevalente;
- le stazioni che avevano una buona correlazione lineare con la stazione target;

Per ognuna di queste scelte sono state create diverse reti neurali con 2 input (velocità) o 3 (2 velocità e una direzione). Il metodo dimostra che le reti in cui non c'è la direzione prevalente come input (o come criterio di valutazione) hanno l'errore più alto.

Mabel [4] invece ha effettuato uno studio, per la stima dell'energia mensile di 7 wind farm. I dati meteorologici usati erano relativi all'altezza del mozzo delle turbine e non a 10 metri.

Nello stesso periodo sono stati raccolti:

- l'energia mensile generata - $E_{MM}(\text{target})$;
- le ore di produzione medie mensili - $H_{MM}(\text{input})$;
- pressione, temperatura, umidità relativa (Φ_{MM}), velocità media mensile (v_{MM});

Per pressione e temperatura sono state rilevate basse variazioni e quindi non sono state considerate come input, mentre v_{MM} e H_{MM} hanno dimostrato avere una buona dipendenza con l'energia prodotta, così come, ma più bassa, per Φ_{MM} .

Nella valutazione preliminare dei dati, gli autori segnalano anche la presenza di apparenti anomalie, come la presenza di alti valori di v_{MM} e bassi valori di E_{MM} . Tali eventi possono essere causati dalla non disponibilità temporanea della rete oppure ad operazioni di fermo impianto per manutenzione. L'errore quadratico medio in fase di training è risultato pari a 0,007, e 0,065 in fase di test mentre l'errore assoluto medio è risultato del 5,8 % in entrambe le fasi.

Uno altro studio è stato condotto per la località di La Venta

in Messico da Cadenas e Rivera [5] per la stima della velocità media oraria a partire da dati di vento della stazione anemometrica presente nella località esaminata.

In [5] si sottolinea l'importanza della ricerca dell'architettura di rete migliore attraverso la variazione del numero di neuroni e di layer nascosti nella rete. Le discrepanze fra il dato reale e quello stimato sono state valutate attraverso gli indici Errore Quadratico Medio (MSE - Mean Squared Error) e Errore Assoluto Medio (MAE - Mean Absolut Error) risultati rispettivamente compresi fra 0,016-0,020 e 0,0399-0,0435.

Ancora sull'argomento in [7] Shuhui et altri effettuano una valutazione comparativa dell'energia prodotta, su scala mensile, per la wind farm (12 turbine da 500 kW) di Fort Davis (Texas) attraverso reti neurali e attraverso modelli tradizionali basati sull'uso della velocità media, della direzione media e della curva di potenza delle turbine.

In [7] gli aspetti fondamentali sono rappresentati da:

- La necessità di effettuare un confronto, tra il dato di vento rilevato e la curva di potenza della turbina per valutare il posizionamento dei punti di lavoro (tanto più si addenseranno intorno alla curva di potenza tanto più saranno correlati, diversamente potrebbe essere necessario utilizzare altri input - come direzione, temperatura, umidità - o indagare sulle cause di difformità);
- L'utilizzo di una rete neurale per ogni aerogeneratore giustificato dalla presenza nei dati di forti variazioni dei valori di producibilità per ogni turbina (a parità di vento registrato);
- I dati ingresso alla rete non vengono forniti tal quali ma riscalati attraverso una funzione chiamata di compressione, diversa per velocità e direzione, di cui però non viene fatta menzione sulla natura e sulle modalità di creazione;
- La differenza fra il dato misurato e quello stimato tramite rete neurale, su scala mensile, è variabile fra lo 0,30 % e il 4,06 % utilizzando dati registrati ogni 10';
- Nessuna menzione viene fatta sulle previsioni di brevissimo periodo (10'-1 h).

In sintesi si può affermare che pochi sono gli studi effettuati sulla previsione di breve periodo (*short forecast*), limitati ai dati di velocità del vento utilizzando come dato di ingresso la serie storica della stessa stazione, e solo in alcuni studi velocità rilevate in stazioni remote.

CASO STUDIO: PREVISIONE DELLA PRODUCIBILITÀ DI UN IMPIANTO EOLICO IN SICILIA

Descrizione del sito

Il presente lavoro è stato svolto utilizzando i dati di produzione anemoelettrica di una wind farm nella provincia di Trapani. L'impianto (il cui layout è rappresentato in figura 1) è costituito da 11 aerogeneratori (entrati in funzione il 28 aprile 2005) da 850 kW ciascuno per una potenza totale di 9,25 MW.

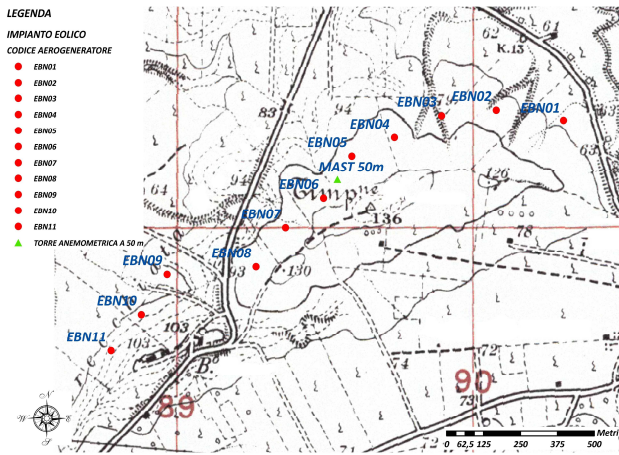


Figura 1– Layout dell’impianto eolico su cartografia

In questo studio sono stati inoltre utilizzati parametri di vento rilevati dalla stazione anemometrica del CNMCA di Trapani Birgi e da quelle del SIAS di Mazara del Vallo, Trapani Fulgatore e Castelvetro ricadenti nell’area del parco eolico o al più nei territori contermini il cui posizionamento è visibile in figura 2.

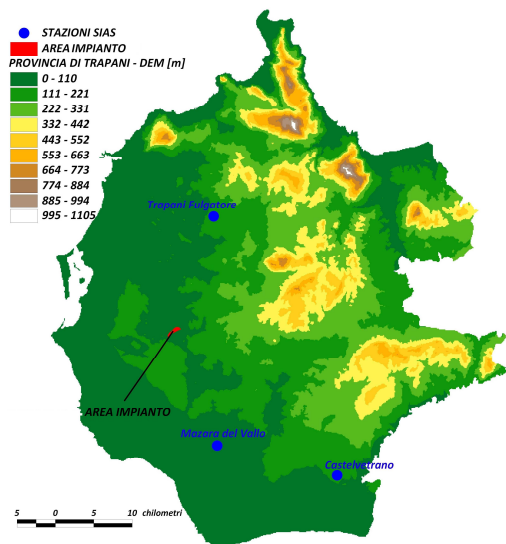


Figura 2 – Orografia della provincia di Trapani, area dell’impianto e posizione delle stazioni SIAS

Dalla figura 2 è possibile inoltre valutare la singolarità di tale area rispetto al complesso regionale prevalentemente ad orografia complessa. Difatti la provincia di Trapani presenta una vasta area ad orografia piana (con altitudini comprese fra 0 e 250 m.s.l.m) in cui ricadono le stazioni anemometriche e l’impianto oggetto dell’analisi. Esso infatti si sviluppa su un crinale collinare con un picco di altitudine pari a 136 m, mentre le aree circostanti sono tutte inferiori a tale quota.

Interrogando l’Atlante eolico l’area d’impianto presenta un fattore di carico intorno alle 2000 ore anno.

Analisi dei dati di input

Al fine di realizzare un modello di previsione per il parco eolico oggetto di questo lavoro si è fatto uso di reti neurali artificiali il cui limite però è spesso rappresentato dalla qualità

delle serie storiche utilizzate.

I dati (mediati su intervalli di 10 minuti) sono stati rilevati in loco dal gestore del sito e consistono in:

- velocità del vento in m/s rilevate da due sensori posti a 50 m campionati dal 04/2005 al 7/3/2008), da due sensori a 30 m e da un sensore a 10 m ciascuno dei quali campionati dall’ 8/7/2002 al 15/6/2006
- direzione del vento rilevata da un sensore posto alla quota di 30 m e da due sensori a 50 m; per un periodo complessivo compreso tra luglio 2002 e marzo 2008;
- dati di potenza prodotta (10⁷) da ogni turbina a partire dal 28 aprile 2005 fino a marzo 2008;

Gli altri dati utilizzati sono quelli della stazione anemometrica, velocità e direzione, del CNMCA di Trapani Birgi. Nel prosieguo dello studio i dati provenienti dall’anemometro a 30 metri non verranno utilizzati poiché antecedenti alla realizzazione dell’impianto e quindi non correlabili con dati di reali di potenza.

Vengono inoltre utilizzati i dati di alcune stazioni della rete SIAS precedentemente citate.

Nella tabella che segue viene riportato un prospetto riepilogativo dei dati anemometrici disponibili per il caso studio con le principali caratteristiche.

Tabella 1 – Prospetto di riepilogo dei dati di vento disponibili per il caso studio

STAZIONE	CODICE	RETE	ALTEZZA ANEMOMETRO [m]	PERIODI DI DISPONIBILI	SCANSIONE DI ACQUISIZIONE
MAZARA DEL VALLO	M	SIAS	10	2002-2007	1 h
CASTELVETRANO	C	SIAS	10	2002-2007	1 h
TRAPANI FULGATORE	TF	SIAS	10	2002-2007	1 h
TRAPANI BIRGI	T	CNMCA	10	2000-2006	1 h
MAST 50 SUD	V1	PRIVATA	50	Maggio 2005 - Dicembre 2007	10'
MAST 50 NORD	V2	PRIVATA	50	Maggio 2005 - Dicembre 2007	10'

Tali dati sono stati sottoposti ad un accurato pretrattamento attraverso il quale sono stati filtrati gli outliers ed eliminati quelli registrati in maniera errorea dal data logger, o comunque caratterizzati da simboli non processabili matematicamente, sono stati inseriti all’interno di un apposito database. Suddividendo le serie temporali di dati in andamenti giornalieri, mensili ed annuali si sono effettuate le seguenti analisi statistiche preliminari:

1. l’analisi di copertura dei dati e di calma, mensile ed annuale, per le grandezze rilevate;
2. i valori medi mensili ed annuali per le serie di velocità del vento e per la potenza prodotta per l’intero impianto;
3. l’andamento medio giornaliero della velocità del vento per ogni anno di registrazione;
4. l’andamento medio mensile della potenza resa da ogni singola turbina per il periodo complessivo di rilevazione;
5. i diagrammi polari della direzione del vento per ogni anno di rilevazione;
6. le curve della distribuzione di Weibull per le serie annuali di dati di velocità del vento disponibili;
7. i diagrammi polari della direzione del vento per classi di velocità per ogni anno di misurazione;
8. la distribuzione media oraria del vento per classi di velocità, per ogni anno di dati disponibile;
9. la distribuzione mensile del vento suddivisa per classi di velocità seguendo i parametri caratteristici delle turbine
10. la distribuzione annuale della potenza resa per classi

di potenza;

- la verifica degli andamenti della potenza resa in base alla variazione della risorsa eolica.

Da tale analisi è risultata evidente la necessità di scartare i dati di produzione rilevati nell'anno 2008 in quanto incompleto e quelli relativi al 2007 a causa di una scarsa copertura di dati dovuta o ad un malfunzionamento di alcune turbine o molto più probabilmente a qualche anomalia del sistema di acquisizione dati (vedi figura 3).

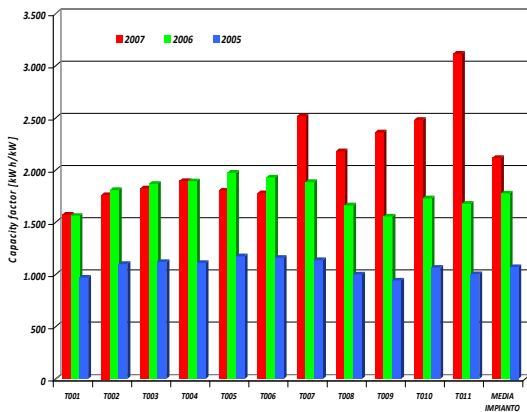


Figura 3 – Capacity factor delle turbine dell’impianto eolico per gli anni 2005,2006 e 2007

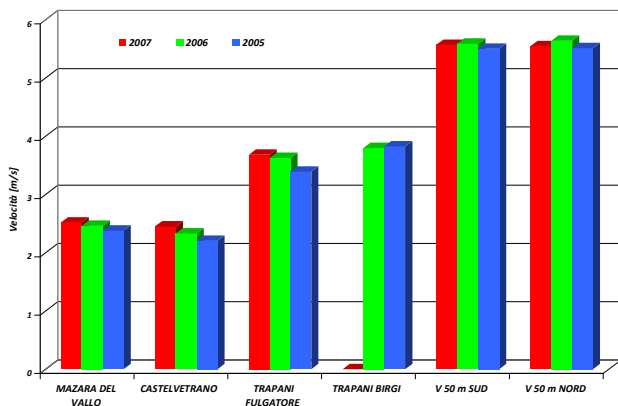


Figura 4 – Velocità medie rilevate dai sensori SIAS/CNMCA/PRIVATE negli anni 2005,2006 e 2007

In particolare sono state calcolate le velocità medie del vento per le stazioni anemometriche ed il capacity factor per ogni singola turbina.

Le figure 3 e 4 mostrano come il dato relativo alla velocità media sia pressoché confermato nei vari anni e per le varie stazioni, mentre sul dato di producibilità (che per il 2005 è parziale poiché non compendia tutto l’anno), l’andamento è simile fra il 2005 e il 2006 in tutte le turbine e dissimile fra questi e il 2007 nelle turbine T007, T008, T009, T010, T011.

ARCHITETTURA DELLA RETE NEURALE

Terminata la fase di analisi dei dati di producibilità e di vento disponibili per l’area d’impianto il lavoro si è concentrato sulla scelta del modello neurale da utilizzare e sulle caratteristiche dello stesso in funzione dei dati

disponibili.

I passaggi chiave [8] della procedura descritta nei paragrafi successivi sono rappresentati in figura 5.

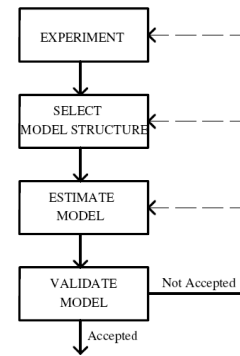


Figura 5 – Procedura di creazione e valutazione di sistemi neurali [8]

Esistono in letteratura delle relazioni empiriche che legano il numero di neuroni al numero di dati disponibili, per la fase di training, e al numero di input ed output [9]:

Siano:

- H = numero di neuroni degli strati nascosti;
- N = dimensione dello strato di input;
- M = dimensione dello strato di output;
- T = dimensione del training set.

Il numero di neuroni ed il numero di pesi (per reti ad un solo strato) sono dati dalle equazioni 1 e 2:

$$H \approx \frac{T}{5(N + M)} \quad (1)$$

$$\text{NumeroPesi} = [(N + M) \times N + H + M \times H] \quad (2)$$

Tali prime elaborazioni sono necessarie per valutare se l’architettura di rete si presta alle analisi in funzione dei dati disponibili per l’addestramento.

Esiste infatti una regola empirica che indica la dimensione dell’insieme di addestramento della rete neurale in circa 10 volte il numero dei pesi derivante dalle relazioni sopra riportate [9].

Considerando che i dati disponibili sono, per ogni anno, 8760 e poiché il numero di input, per come è stata condotta l’analisi, è variabile (da un minimo di 3 ad un massimo di 6) si è deciso di utilizzare delle reti con 100 neuroni e con un numero di iterazioni, in fase di addestramento, pari a 100.

L’utilizzo di reti con numero di neuroni, e quindi di pesi, più alto è stato testato, in questa primissima fase, solo per alcune architetture e non ha condotto a miglioramenti nei risultati così come l’incremento di iterazioni, da 100 a 500. Entrambe le opzioni comportano invece alti tempi di elaborazione.

Limitando il campo ai cosiddetti MLP in questi lavoro sono stati sperimentate 2 tipologie di modelli:

- Modelli che utilizzano il segnale di ingresso e le uscite precedenti;
- Modelli che utilizzano il segnale di ingresso, le previsioni delle uscite precedenti e l’errore di

predizione;

Appartengono alla prima famiglia i modelli NNARX (Neural Network Autoregressive eXogenous signal) ed alla seconda i modelli NNARMAX (Neural Network Autoregressive Moving Average eXogenous signal).

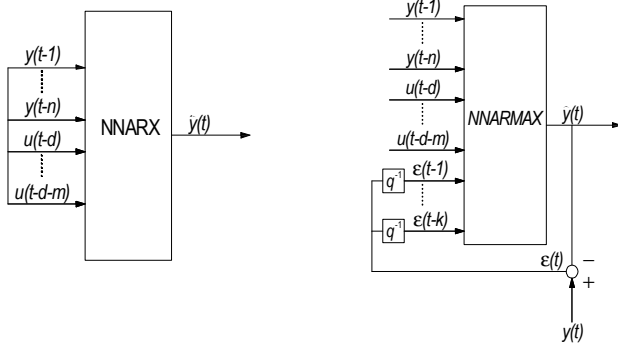


Figura 6 – Schema di funzionamento di una rete NNARX e di una rete NNARMAX

Fase di addestramento

Dall'osservazione combinata dei dati di vento e producibilità si è deciso di utilizzare come set di addestramento i dati relativi all'anno 2006.

Il modello di rete inizialmente utilizzato è stato il NARX le cui caratteristiche vengono riportate in figura 6

Con tale architettura sono state effettuate le elaborazioni come combinazione dei dati di input disponibili, variando il passo temporale da 2 a 5 e cioè variando il numero di step temporali da cui dipende l'output (la potenza erogata dalle turbine).

Ad esempio una rete con passo temporale 5 utilizza i 5 input precedenti, nel nostro caso i dati delle cinque ore precedenti, per stimare l'output della sesta ora. Nelle tabelle 2 e 3 sono riportati prospetti riepilogativi per reti con 100 neuroni, input variabili e time step 2/5.

Nella fase di addestramento è stato utilizzato come indice di performance delle reti il NSSE (Normalized Sum Squared Error) definito dall'Eq.3. Ovviamente performance migliori implicano bassi valori di questo indice. Le architetture più performanti sono evidenziate in tabella.

$$NSSE = \frac{\sum_{k=1}^n (Y_{REALE} - Y_{PREDIZIONE})^2}{\sum_{k=1}^n (Y_{REALE} - \bar{Y}_{REALE})^2} \quad (3)$$

Dove:

- Y_{REALE} è il valore di potenza della centrale eolica;
- $Y_{PREDIZIONE}$ è il valore in uscita dalla rete neurale dopo l'addestramento;
- \bar{Y}_{REALE} è la media dei valori di potenza dell'insieme di addestramento;
- n è numero di valori di potenza in uscita (8760 meno il time step);

Tabella 2 – Prospetto riepilogativo delle simulazioni su

modello NARX – time step 2

FASE DI ADDESTRAMENTO	RETE CON 100 NEURONI E 2 PASSI TEMPORALI	
NOME RETE	INPUT	NSSE
MAZARA	M-V1	0,072
CASTELVETRANO	C-V1	0,079
TRAPANI FULGATORE	TF-V1	0,079
VELOCITA' DIREZIONE 50 METRI	D-V1	0,093
SIAS	M-C-TF	0,054
SIASL	M-C-TF-V1	0,051
SIAS2L	M-C-TF-V1-V2	0,049
SIAS2L	M-C-TF-T-V1-V2	0,045

Tabella 3 – Prospetto riepilogativo delle simulazioni su modello NARX – time step 5

FASE DI ADDESTRAMENTO	RETE CON 100 NEURONI E 5 PASSI TEMPORALI	
NOME RETE	INPUT	NSSE
MAZARA	M-V1	0,042
CASTELVETRANO	C-V1	0,048
TRAPANI FULGATORE	TF-V1	0,042
VELOCITA' DIREZIONE 50 METRI	D-V1	0,067
SIAS	M-C-TF	0,029
SIASL	M-C-TF-V1	0,023
SIAS2L	M-C-TF-V1-V2	0,020
SIAS2L	M-C-TF-T-V1-V2	-

I risultati mostrano che al crescere degli step temporali, per la stessa rete, diminuisce l'NSSE. All'interno di reti con pari step temporali l'utilizzo di un maggior numero di input fornisce risultati migliori mentre si registra una scarsa influenza se non addirittura peggiorativa dell'utilizzo di dati di direzione come input. Addirittura le reti SIAS (che non hanno all'interno nessuna informazione contestualizzata al parco eolico forniscono risultati migliori delle reti D-V1)

Si nota anche che i risultati della rete che utilizza come input i dati di velocità e direzione a 50 metri (D-V1) sono meno performanti rispetto a reti con pari step temporali, ciò indica che l'utilizzo del dato di direzione, anche se a 50 metri e misurato all'interno del parco eolico, non comporta miglioramenti nell'addestramento della rete e ancor più nella stima della producibilità.

Lo stesso lavoro è stato svolto utilizzando un modello di rete neurale differente, il NNARMAX, che differisce dal modello NNARX in quanto utilizza gli errori come feedback per l'addestramento della rete.

Con il modello NNARMAX1 si sono effettuate simulazioni solo per le reti SIASL, SIAS2L e SIAS2L, rispettivamente con 4, 5 e 6 input in quanto hanno dimostrato avere prestazioni migliori.

I valori dell'NSSE riscontrati in fase di apprendimento di reti NNARMAX1 sono leggermente migliori delle omologhe NNARX con step temporale 2, il contrario si è ottenuto invece con uno step temporale 3. I risultati della fase di addestramento non permettono di individuare un modello migliore, fra i due indagati, stante che l'ordine di grandezza dell'NSSE è pressoché identico. Le simulazioni condotte con NNARMAX1, per le potenze di calcolo richieste, non superano i tre passi temporali.

Tabella 4 – Prospetto riepilogativo delle simulazioni su modello NNARMAX1 – time step 2

FASE DI ADDESTRAMENTO	RETE CON 100 NEURONI E 2 PASSI TEMPORALI	
NOME RETE	INPUT	NSSE
SIASL	M-C-TF-V1	0,050
SIAS2L	M-C-TF-V1-V2	0,046
SIAS2L	M-C-TF-T-V1-V2	0,043

Tabella 5 – Prospetto riepilogativo delle simulazioni su modello NNARMAX1 – time step 3

FASE DI ADDESTRAMENTO RETE CON 100 NEURONI E 3 PASSI TEMPORALI		
NOME RETE	INPUT	NSSE
SIASL	M-C-TF-V1	0,039
SIAS2L	M-C-TF-V1-V2	0,038
SIAS2L	M-C-TF-T-V1-V2	0,033

La valutazione delle performance delle reti neurali, in fase di addestramento, è completata dall'esame comparativo fra l'output predetto ed il reale realizzato mediante grafici e distribuzione degli errori (figure 7-10).

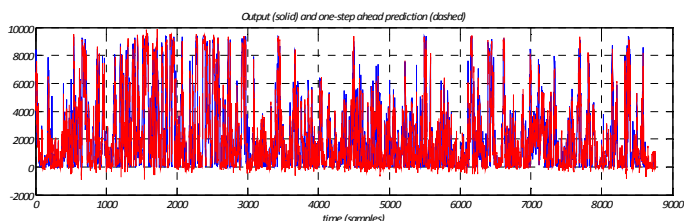


Figura 7 – Confronto Output-Predizione rete D-V1 con 2 time step (NNARX)

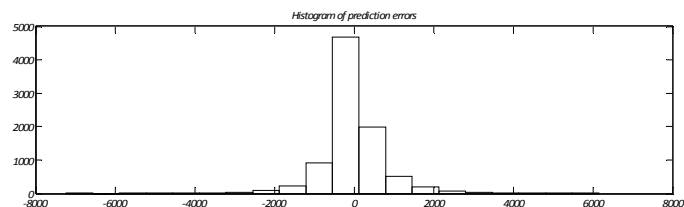


Figura 8 – Istogramma degli errori per rete Narx D-V1 con 3 time step (NNARX)

Si noti come l'istogramma degli errori sia pressoché simmetricamente distribuito attorno al valore nullo e presenti basse frequenze per valori piuttosto elevati dell'errore. Già in questa fase di addestramento (reti con 3 input di velocità) ci si accorge che la rete fornisce stime con valori negativi soprattutto in vicinanza di eventi di produzione reale nulla o abbassamenti repentini delle velocità del vento. Tale fenomeno è dovuto alla funzione di trasferimento, iperbolica, che prevede anche la presenza di valori in uscita negativi [9].

Per reti con numero di input superiore a 3 le indicazioni fornite dagli istogrammi degli errori sono per lo più le stesse. Si riscontra infatti un addensamento degli errori attorno al valore nullo, una distribuzione simmetrica e basse frequenze di accadimento per alti valori dell'errore.

Permane anche per reti in cui è presente il dato di velocità a 50 metri (da 4 input in poi) la presenza di valori negativi nel output di addestramento per eventi di produzione reale a scarsa incidenza produttiva (sotto al soglia del 20 % della potenza nominale dell'impianto) o in presenza di valori reali nulli.

Dall'analisi si nota come tutte le reti si dimostrino pressoché equivalenti nella predizione della producibilità oraria. Le differenze si possono cogliere solo con l'ausilio di un indice aggregato come l'NSSE. La gestione degli eventi con produzione nulla appare, per tutte le simulazioni effettuate, deficitaria e causata da un lato dalla funzione di trasformazione degli output che prevede valori negativi e

dall'altro alla presenza nel training set di eventi anemometrici non nulli ed eventi di produzione nulla imputabili a guasti o manutenzione non ordinaria, fenomeno riscontrato anche in alcuni studi che risulta falsante per l'addestramento delle reti [4].

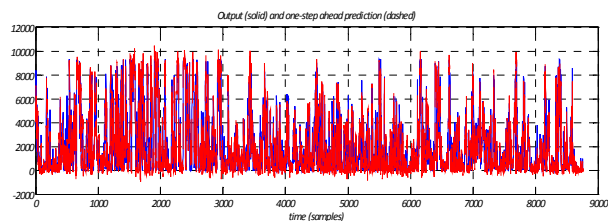


Figura 9 – Confronto Output-Predizione rete SIAS2L con 2 time step (NNARX)

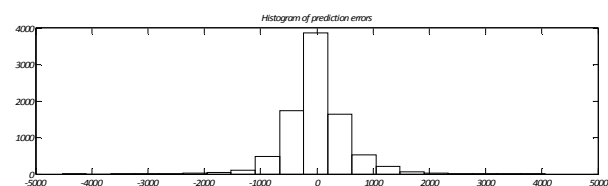


Figura 10 – Istogramma degli errori rete SIAS2L con 2 time step (NNARX)

Fase di validazione

Poiché i dati relativi all'anno 2006 sono stati utilizzati per la fase di addestramento e quelli del 2007, come detto, presentano anomalie per alcune turbine si è deciso di utilizzare per la fase di test un sottoinsieme di dei dati (velocità e produzione) rilevati nel 2005.

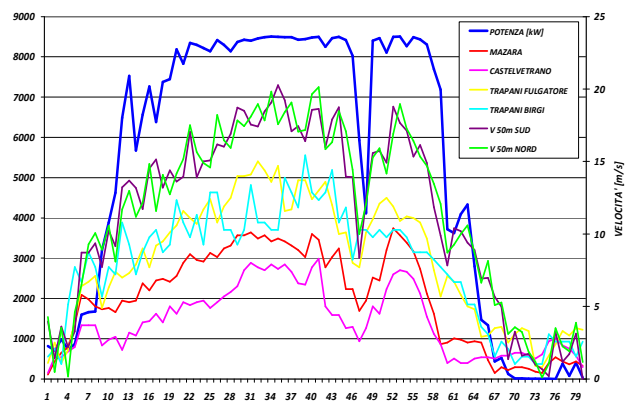


Figura 11– Andamento dei valori di velocità e produzione eolica per il set di validazione 2005

L'insieme di test è stato scelto in modo tale che gli output reali di produzione, all'interno della serie temporale, non siano mai nulli e riescano per un certo periodo ad avvicinarsi alla potenza nominale dell'impianto.

Tale scelta è stata necessaria al fine di valutare il comportamento delle reti in assenza di dati nulli e soprattutto verificare le performance predittive su eventi con alta produzione di energia eolica poiché questi potrebbero generare sovraccarichi alla rete e quindi instabilità.

Nella figura 11 viene riportato l'andamento della

produzione eolica e dei valori di velocità di vento rilevati nel periodo che va dalle ore 8.00 del 16/12/2005 alle ore 15.00 del 19/12/2005 per un totale di 80 eventi orari.

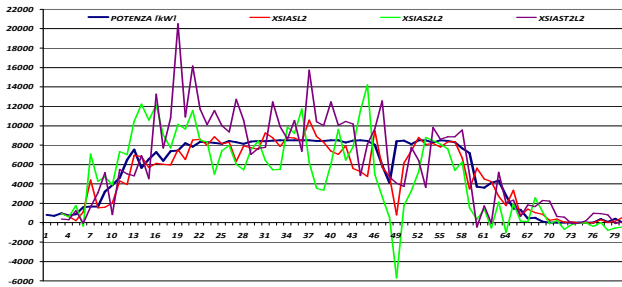


Figura 12 – Confronto output di rete neurali con 2 time step (NNARMAX1) e produzione reale

I risultati (di cui un esempio è riportato in figura 12) evidenziano una certa deficienza della rete nello stimare la producibilità e sono in contrasto con quelli ottenuti nella fase di addestramento; per tale motivo si sono effettuate alcune indagini precauzionali.

Dapprima, per testare la validità dei modelli di previsione di wind farm tramite dati di vento, si è costruita una rete che utilizzasse i dati anemometrici alla stessa altezza del mozzo delle turbine.

La rete così costituita prende il nome di 50 VV, e su essa è stata effettuata la fase di addestramento con i dati dell'anno 2006. Osservato quindi che i risultati non miglioravano al variare del numero e del tipo di input si è deciso di variare il numero di neuroni di ogni rete.

E' infatti possibile che reti con architetture complesse imparino a descrivere bene i fenomeni di addestramento e offrano risultati insoddisfacenti quando vengono testati su un insieme di dati nuovo ricadendo nel cosiddetto *overfitting* [9].

Si tratta quindi di affinare i modelli in modo tale da trovare quell'architettura neurale capace di minimizzare l'errore di stima.

Si è deciso pertanto di ripetere le operazioni di addestramento per reti neurali con 10 neuroni. La diminuzione della complessità della rete potrebbe portare ad un errore più alto in fase di training e ad una migliore stima in fase di validazione. L'operazione di ri-addestramento è stata condotta solo su reti neurali con un numero di input superiore a 3.

Riaddestramento (reti con 10 neuroni)

Come sopra descritto in questa fase si è deciso di riaddestrare le reti neurali con numero di input superiore a 3 passando da 100 a 10 neuroni. Tale modifica comporta anche la possibilità di valutare, per reti con numero alto di input i risultati con time step superiori a 4, ma per quanto osservato nelle precedenti fasi in questa analisi il time step sarà inferiore o uguale a 3.

Anche in questo caso per il modello NARX al crescere del numero di input l'NSSE decresce, per entrambi i time step. Per il modello NNARMAX1 non si possono fare generalizzazioni. Rimane pressoché invariato l'ordine di grandezza dell'NSSE per entrambi i modelli (a 10 neuroni), mentre risulta più alto al confronto con reti omologhe a 100 neuroni (per la minore complessità).

Rivalidazione (reti con 10 neuroni)

Terminata la nuova fase di addestramento si è proseguito con la validazione delle reti sullo stesso insieme di test utilizzato per le reti con 100 neuroni.

Nella figura 13 sono riportati i risultati del ri-testing per le reti a 10 neuroni omologhe della figura 12.

Si nota innanzitutto un netto miglioramento nella previsione non soltanto nel seguire il dato reale di produzione ma anche, e soprattutto, nell'intervallo di valori predetti, in altre parole le oscillazioni attorno al valore reale diminuiscono.

Tale effetto è vero per tutte le reti indagate e risulta necessario il ricorso al calcolo di alcuni indici prestazionali per effettuare una valutazione comparativa rigorosa.

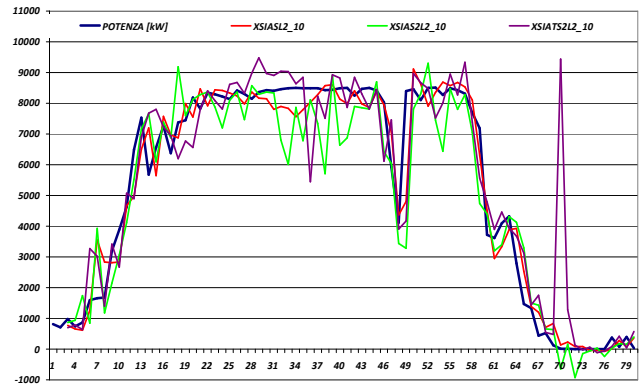


Figura 13 – Confronto previsione – reale in reti NNARMAX1 (10 NEURONI) con input crescente e time step 2

Per il calcolo di tali indici si è deciso di normalizzare il valore assoluto degli scarti delle previsioni sul valore reale stimato attraverso la seguente formula:

$$E_{NORMALIZZATO} = \frac{|Y_{REALE} - Y_{PREDIZIONE}|}{Y_{REALE}} \quad (4)$$

Per questo indice di prestazione sono state valutate le seguenti grandezze:

- Massimo;
- Minimo;
- Somma;
- Media;
- Varianza;
- Scarto quadratico medio;

Nelle figure 14 e 15 vengono rappresentati i valori di media e scarto quadratico medio degli errori assoluti per i due modelli. È evidente come il modello NNARX offra, al confronto, i risultati migliori e mostri una diminuzione dei valori di media e varianza al passaggio da 100 a 10 neuroni.

Nel modello NNARMAX1 l'effetto di diminuzione degli errori al passaggio dall'architettura complessa ad una più semplice non è generalizzabile. Pur tuttavia l'errore assoluto medio è piuttosto elevato per tutte le architetture indagate.

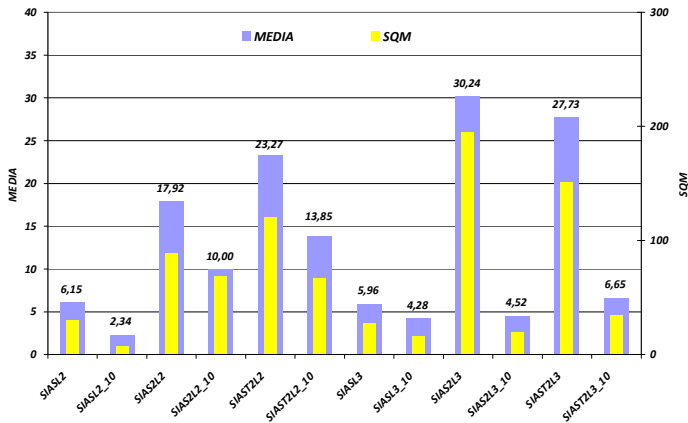


Figura 14 – Confronto media-scarto quadratico medio per modelli NNARX (10 e 100 neuroni)

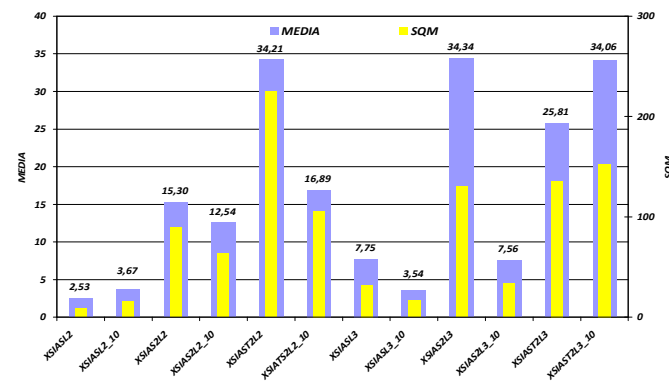


Figura 15 – Confronto media-scarto quadratico medio per modelli NNARMAX1 (10 e 100 neuroni)

PERFORMANCE DELLE RETI A SCANSIONE ORARIA PER EVENTI CON ALTA PRODUZIONE EOLICA

Le evidenze numeriche dei precedenti paragrafi non sembrano giustificare il ricorso alle reti neurali come tools di gestione dei sistemi di potenza in cui sono inseriti impianti eolici.

Tuttavia si è osservato come sia nella fase di addestramento che di test (per reti con 100 e 10 neuroni) la presenza di valori negativi dell'output della rete a fronte di valori reali della potenza erogata dalla wind farm vicini allo zero incrementi notevolmente gli indici prestazionali delle reti. Si è quindi provveduto a circoscrivere questi eventi (dati reali inferiori al 18 % della potenza nominale dell'impianto) nell'insieme di test e alla valutazione prestazionale per la rimanente parte.

In particolare si è ritenuto necessario eliminare dall'insieme di test i dati dalle ore 8 alle ore 14 del 16/12/2005 e dalle ore 4.00 alle ore 15.00 del 19/12/2005 in cui appunto si ha un dato di producibilità prossimi allo zero o scarsa incidenza produttiva. La figura 16 evidenzia quanto sopra detto ed inoltre consente di osservare come, nell'insieme di validazione, non si raggiunga mai la potenza nominale (100%) della wind farm. Fenomeno quest'ultimo in contrasto con i dati di velocità registrati dagli anemometri a 50 metri (figura 11) ben sopra la soglia delle velocità nominale (16 m/s) delle turbine V52.

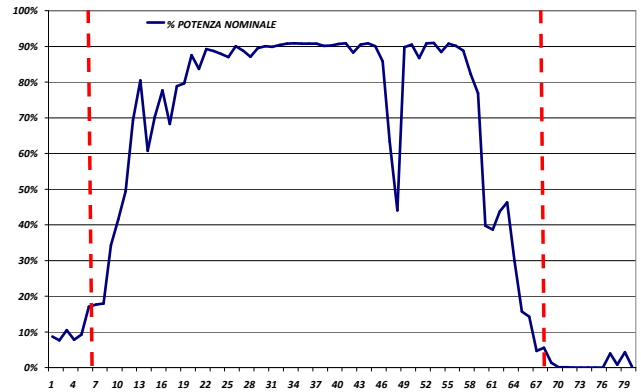


Figura 16 – Insieme di validazione: valori reali in % rispetto alla potenza nominale dell'impianto eolico

I risultati delle rianalisi prestazionali (con modello NNARX) vengono riassunti nella figura 17 e confermano quanto osservato sia in sede di training sia in sede di validazione.

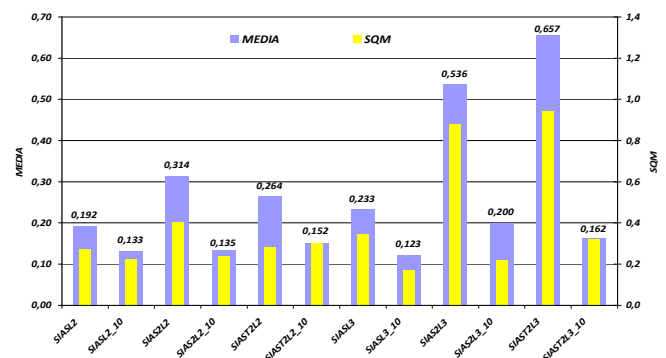


Figura 17 – Risultati reti NARX esclusi dati a scarsa incidenza produttiva

L'errore medio, infatti, si riduce (in tutte le architetture simulate su rete NNARX) mostrando un netto miglioramento nel passaggio da architettura complessa (100 neuroni) ad una architettura più semplice (10 neuroni) sia sul dato errore medio assoluto sia sul dato scarto quadratico medio.

L'errore medio assoluto raggiunge un minimo del 14 % per la rete SIASL3_10 che (come nelle valutazioni precedenti) conferma essere la più performante.

Anche su reti NNARMAX (figura 18) si osserva un miglioramento ma più contenuto rispetto al modello precedente. L'errore medio raggiunge un minimo intorno al 12,7 % (per la rete XSIASL3_10), ed anche in questo caso la diminuzione della complessità della rete ha un effetto migliorativo nella predizione. Rispetto alle reti NARX si osserva anche uno scarto quadratico medio più alto.

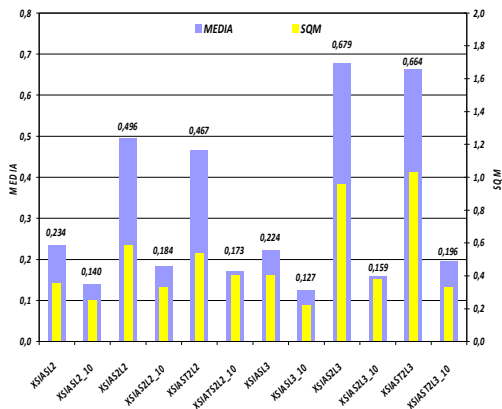


Figura 18 –Risultati reti NARMAX esclusi dati a scarsa incidenza produttiva

Il non considerare dati a scarsa incidenza produttiva (come detto i dati eliminati dal test set variano da un minimo dello 0,005 % ad un massimo del 17,80 % rispetto alla potenza nominale dell’impianto) è giustificato dal fatto che l’obiettivo del lavoro è quello di valutare le condizioni di instabilità causate da un’alta penetrazione di energia eolica. Per ovvi motivi (curva di potenza sigmoideale in testa) tale circostanze possono verificarsi solo per eventi anemologici piuttosto elevati.

Considerando poi che nell’insieme di validazione “ristretto” non si raggiunge mai la potenza di targa dell’impianto i risultati ottenuti appaiono ancor più confortanti (poiché superata la velocità di 16 m/s le macchine dovrebbero produrre tutte a potenza nominale).

I risultati ottenuti, nell’insieme di validazione ristretto, assumono ancor più rilievo se si osserva la figura 19. In essa vengono riportati i punti di lavoro degli aerogeneratori (coppie potenza reale – dato di vento misurato a 50 m) nell’insieme di validazione (punti in blu ed in rosso) e i punti di lavoro delle macchine eoliche considerando la sola curva di potenza della turbina per le stesse velocità registrate all’altezza del mozzo (punti in verde).

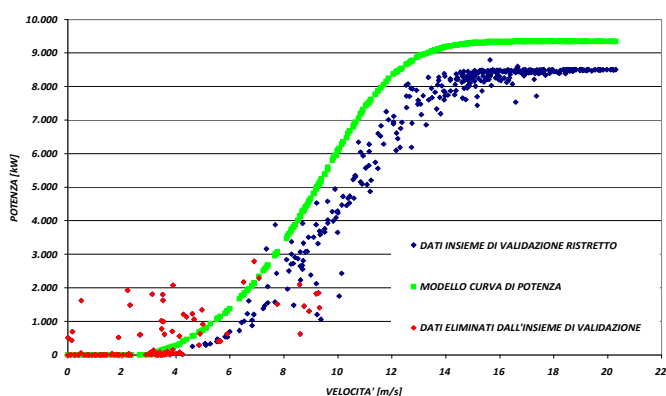


Figura 19 – Punti di lavoro reale e teorico per la wind farm nell’insieme di validazione

Si noterà come i valori reali registrati siano sensibilmente inferiori rispetto al dato teorico (modello curva di potenza). In particolare si può affermare che nell’insieme di validazione non tutte le turbine funzionano e la capacità predittiva della rete viene ancor più amplificata se si considera che le

informazioni in input al modello sono relative solo ai dati di velocità del vento e non alla produzione di energia o ad un segnale on/off per ogni turbina. Questo significa che le reti neurali sono capaci di percepire, dalle informazioni acquisite in fase di addestramento, quanto la wind farm nel suo insieme è capace di produrre anche in assenza di informazioni sul reale funzionamento di ogni turbina.

Sempre dalla figura 18 si può notare come i valori eliminati dall’insieme di validazione (punti in rosso) ricadano in aree distanti dal valore teorico estrapolabile dalla curva di potenza. Si riscontra infatti la presenza, nell’intervallo 0-4 m/s, di eventi a potenza reale non nulla (con picchi di 2 MW pari a circa il 15 % della potenza nominale dell’impianto) a fronte di un valore teorico pari a 0.

E’ evidente che la capacità predittiva delle reti neurali, come di qualsiasi altro strumento, cade in difetto in presenza di valori fuori scala rispetto a quelli teorici e ciò giustifica la scelta di eliminare dal test set i dati con potenza reale inferiore al 20 % e ricadenti in aree scarsamente correlate con gli output teorici delle macchine.

La presenza di eventi fuori scala dovrebbe essere ridotta utilizzando dei protocolli di acquisizione e validazione dei dati registrati dalle turbine e dagli anemometri.

VALUTAZIONI SU DIVERSE SCALE TEMPORALI RETI NEURALI CON SCANSIONE OGNI 10’

I risultati ottenuti delle simulazioni su scansione oraria hanno dimostrato la capacità predittiva delle reti neurali per la producibilità eolica, seppur con i problemi sopra descritti, che potrebbero essere limitati attraverso:

- l’adozione di un protocollo di acquisizione e controllo dati di producibilità;
- l’adozione di un protocollo di acquisizione e controllo dei dati di vento a varie altezze;
- la programmazione delle operazioni di manutenzione ordinaria nelle stesse giornate di ogni anno o mese;
- la ricerca dell’architettura neurale migliore al variare di numero neuroni e numero di iterazioni.

Tali considerazioni nascono dall’osservazione dei risultati, dai valori di potenza nell’insieme di validazione (un massimo del 90 %) e della valutazione della figura 20 in cui sono rappresentati i punti di lavoro degli aerogeneratori correlati alla velocità registrata a 50 m ogni 10’ (anno 2006 –insieme di addestramento).

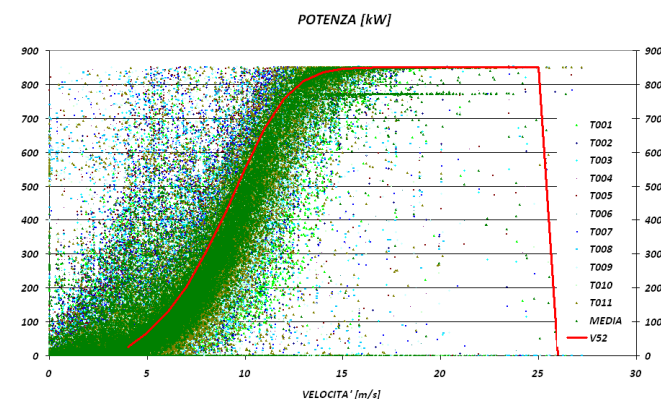


Figura 20 – Punti di lavoro degli aerogeneratori nell’impianto eolico in esame per l’anno 2006 –Set d’addestramento

Le reti neurali (10 neuroni), sempre con modello NNARX e NNARMAX, hanno solo 2 input provenienti dagli anemometri a 50 metri, posizionati all'interno della wind farm (poiché questi sono gli unici presenti nel database con scansione ogni 10'). I risultati, sullo stesso insieme di test delle simulazioni a scansione oraria, sono rappresentati nelle figure successive.

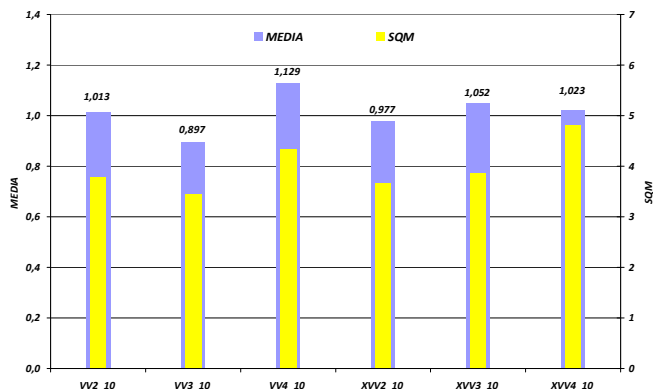


Figura 21 – Confronto media-scarto quadratico medio per modelli NNARX e NNARMAX (scansione 10')

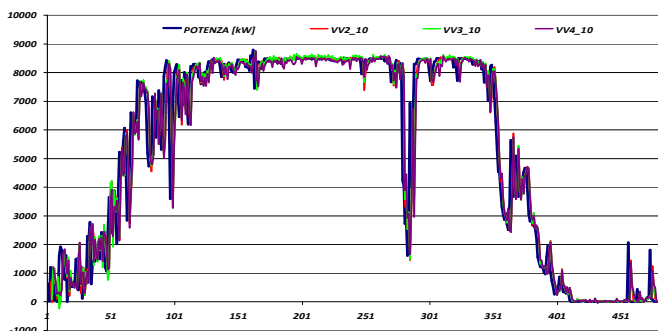


Figura 22 – Confronto previsione – reale in reti NNARX con 2 input e time step crescente (scansione 10')

Dalla figura 21 si nota che pur rimanendo sostenuto l'errore medio normalizzato (il valore più basso è di 0,897) il dato di varianza, rispetto al caso orario, diminuisce notevolmente e questo indica un miglior adattamento della stima al valore reale.

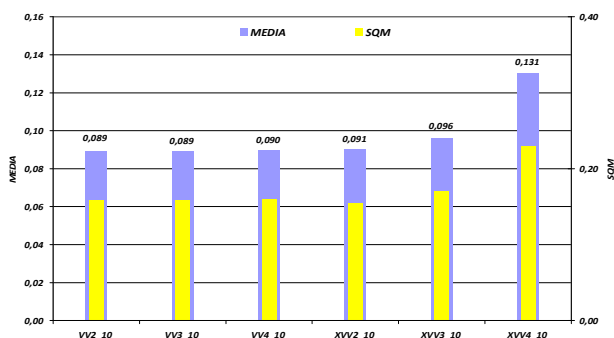


Figura 23–Risultati reti con scansione ogni 10' esclusi dati scarsa incidenza produttiva

Anche per reti con scansione 10' sono stati valutati gli indici errore medio assoluto e scarto quadratico medio dell'errore medio assoluto nell'insieme "ristretto" di

validazione individuato come sopra descritto (figura 23).

Appare evidente il miglioramento delle prestazioni delle reti a 10' rispetto alle omologhe a scansione oraria seppur addestrate con un numero inferiore di input (2 –anemometri interni al campo eolico).

L'errore assoluto medio varia da un minimo di 0,089 ad un massimo di 0,131 con uno scarto quadratico medio contenuto sotto 0,25. I risultati ottenuti pur essendo migliori di quelli a scansione oraria non sono direttamente comparabili per le differenze dovute al numero di input utilizzati in questa fase.

CONCLUSIONI

Lo studio svolto permette di individuare una serie di fattori chiave per l'implementazione di sistemi di previsione della producibilità eolica basati su reti neurali.

Innanzitutto, come dimostrato dalla letteratura [4,7], lo strumento rete neurale è capace di predire, con un certo errore, la producibilità di una turbina, di una wind farm o di un insieme di esse utilizzando come input i dati provenienti da reti anemometriche remote e da quelle presenti nelle immediate vicinanze degli impianti anche se rare sono gli esempi di previsioni a scala temporale diversa da quella mensile.

L'inserimento, come input della rete neurale, di altri parametri meteo climatici, quali ad esempio la temperatura misurata all'altezza del mozzo, potrebbe portare ad un affinamento dei risultati, mentre come si è visto in questo lavoro l'utilizzo di dati di direzione non fornisce miglioramenti nella stima.

È evidente, dal lavoro, come l'individuazione del modello migliore per la descrizione del fenomeno sia relativo all'architettura neurale (100 neuroni piuttosto che 10 o 20 e così via) e che, non esistendo dei veri e propri metodi di ottimizzazione ante training delle reti neurali, esse vadano ottimizzate in itinere, e cioè dopo la fase di validazione, verificando se si ricade in *overfitting* o meno.

Nel caso in cui le reti ricadano in *overfitting*, pur utilizzando gli stessi modelli occorre cambiare l'architettura del sistema (nel nostro caso si è passati da 100 a 10 neuroni) e, come qui verificato, i risultati potrebbero contraddirsi con le evidenze numeriche di reti con architetture più complesse.

In particolare, ad esempio, utilizzando reti neurali con 100 neuroni, in fase di training, si è osservata una migliore risposta del modello NNARMAX1 rispetto al NNARX mentre in fase di validazione si è avuto il risultato opposto. Tuttavia la differenza fra i due modelli non permette di distinguere univocamente quale risultato più performante (gli indici di performance rimangono infatti dello stesso ordine di grandezza per reti omologhe).

Riducendo il numero di neuroni, e quindi la complessità della rete neurale, ci si è accorti che il modello NNARX è sicuramente quello che meglio riesce a descrivere gli eventi presentando per tutti i casi, confrontati con le reti a 100 neuroni di pari input, un errore assoluto medio inferiore (variabile da 2,34 a 13,85).

Non si è avuto lo stesso risultato, invece, per reti del tipo NARMAX1, cioè non è verificato sempre che al diminuire della complessità della rete, con questo modello, si riduca l'errore assoluto medio (variabile da 3,54 a 34,06). I valori riscontrati, non su base comparativa ma assoluta, non permettono di annoverare fino a questa fase del lavoro le reti

neurali come strumenti idonei al network management.

Tuttavia è da sottolineare comunque che la presenza o l'assenza nei dati di input e di output di sottoinsiemi falsanti (ad esempio eventi con dati di vento non nulli e di producibilità nulla) può contribuire a deteriorare o migliorare la stima.

Diventa quindi elemento essenziale per lo sviluppo di tali sistemi la validazione innanzitutto dei dati di ingresso (vento) e di uscita (produzione) che debbono, perlomeno per gli input provenienti da stazioni anemometriche presso gli impianti, essere legate secondo il più possibile alla curva di potenza delle macchine e non legate ad eventi esterni quali: operazioni di manutenzione ordinaria, straordinaria, interruzioni della linea elettrica.

Ovviamente un piano di manutenzione ordinario ripetuto nei vari anni, negli stessi periodi o addirittura giorni, permetterebbe alla rete di riconoscere questi eventi e quindi di minimizzare gli errori finali.

Si è osservata sia in fase di training che di test la presenza, nelle risposte delle reti, di valori negativi. Questi eventi si sono verificati in presenza di repentine diminuzioni delle velocità di input o in presenza di dati vento nullo o ancora in presenza di dati di produzione nulla (soprattutto fase di addestramento).

Considerati gli scopi del presente, orientati alla valutazione di sistemi di potenza fortemente penetrati da eolico la cui instabilità può essere causata solo da eventi anemometrici superiori alla soglia nominale delle turbine, e osservato che i valori dell'errore medio assoluto e dello scarto quadratico medio dell'errore medio assoluto vengono fortemente distorti proprio da questo sotto insieme di dati si è provveduto ad investigare un insieme di validazione "ristretto" in cui la producibilità minima considerata rimane comunque superiore al 5 % rispetto a quella nominale.

Su tale insieme i valori di performance delle reti migliorano sensibilmente passando per l'errore assoluto medio a valori compresi fra 0,123 e 0,200 per reti NNARX (10 neuroni) e a valori compresi fra 0,127 e 0,196 per reti NNARMAX1 (10 neuroni). Per entrambe si è rilevata la diminuzione (più sensibile per modello NARX) dello scarto quadratico medio.

Un altro elemento che influenza il risultato finale è rappresentato dalla scansione temporale di previsione.

Utilizzando dati a scansione di 10', la rete riesce a ridurre l'errore di stima su entrambi gli indici di prestazione. Il risultato è ancora più confortante osservando gli indici di prestazione del sottoinsieme di validazione "ristretto".

In [10] vengono distinte tre diverse scale temporali rispetto alle quali collocare le attività degli operatori di sistema:

- *unit commitment*: con orizzonte temporale da 1 giorno a 1 settimana e incrementi di 1 ora;
- *load following*: con orizzonte temporale di 1 ora e incrementi di 5-10 minuti;
- *controllo della frequenza*: con orizzonte temporale da 1 minuto a 1 ora e incrementi di 1-5 secondi;

Per quanto visto le reti neurali potrebbero essere utilizzate nei casi 1 e 2.

Ovviamente all'aumentare della scala temporale di stima la presenza di eventi falsanti, come sopra descritti, in genere diminuisce e questo potrebbe portare ad un miglioramento

sensibile della previsione.

Nell'uno e nell'altro caso comunque è da ricercare il modello migliore per la descrizione del fenomeno così come dell'architettura di rete.

RIFERIMENTI BIBLIOGRAFICI

- [1] Sathyajith Mathew , K.P. Pandey , Anil Kumar.V (2002). Analysis of wind regimes for energy estimation. *Renewable Energy* 25 (2002) 381–399.
- [2] Yoreley Cancino-Solorzano , Jorge Xiberta-Bernat (2009). Statistical analysis of wind power in the region of Veracruz (Mexico). *Renewable Energy* 34 (2009) 1628–1634.
- [3] GSE – Gestore Servizi Elettrici - Procedura tecnica per il miglioramento della prevedibilità delle immissioni dell'energia elettrica prodotta da impianti alimentati da fonti rinnovabili non programmabili
http://www.gse.it/GSE%20Informa/Documents/Procedura%20tecnica%20per%20ARG-elt%2004-10%20documento%20di%20consultazione%2010022010_V1.pdf
- [4] M. Carolin Mabel, E. Fernandez, Analysis of wind power generation and prediction using ANN: A case study, *Renewable Energy* 33, 986–992, 2008;
- [5] Erasmo Cadenas, Wilfrido Rivera, Short term wind speed forecasting in La Venta, Oaxaca, Mexico, using artificial neural networks, *Renewable Energy* 34, 274–278, 2009;
- [6] P. Lopez, R. Velo, F. Maseda, Effect of direction on wind speed estimation in complex terrain using neural networks, *Renewable Energy* 33, 2266–2272, 2008
- [7] Shuhui Li, Member, IEEE, Donald C. Wunsch, Senior Member, IEEE, Edgar A. O'Hair, and Michael G. Giesselmann, Senior Member, IEEE, Using Neural Networks to Estimate Wind Turbine Power Generation, *IEEE TRANSACTIONS ON ENERGY CONVERSION*, VOL. 16, NO. 3, SEPTEMBER 2001
- [8] Magnus Nørgaard, Neural Network Based, System Identification, TOOLBOX, version 2, Department of Automation, Department of Mathematical Modelling, Technical University of Denmark, 2000
- [9] Antonino Marvuglia Utilizzo di reti neurali artificiali a supporto della pianificazione energetica, Tesi di Dottorato, Dottorato in Fisica Tecnica Ambientale, 2007;
- [10] Georgilakis P.S. Technical challenges associated with the integration of wind power into power systems. *Renewable and Sustainable Energy Reviews* 12, 852-863, 2008

SUMMARY

Wind energy technology had a strong growth as widely used for electricity production. This dramatic rise requires the development of new tools for grid management as well as for the assesment wind farm short and medium term productivity. This paper presents a study based on measures of wind parameters, such as intensity and direction, related to productivity data of a real wind farm. After a first step of pre-processing of data a statistical analysis has been done. The model of correlation input-output is based on the use of artificial neural networks. Authors have developed a model for the short-term forecasting and for different time scales. Also wind data not strictly related to the site have been used