



SAPIENZA  
UNIVERSITÀ DI ROMA

# ***Previsione di brevissimo termine della produzione elettrica da impianti eolici***

MASTER DI 2° LIVELLO IN  
“CALCOLO SCIENTIFICO”  
ANNO ACCADEMICO 2014/2015

Università di Roma La Sapienza  
Facoltà di Scienze Matematiche Fisiche e Naturali

Candidato Claudio Pregagnoli  
Matricola 644081

Relatore  
Prof. Camillo Cammarota



## Indice

<b>1.</b>	<b>Introduzione .....</b>	<b>6</b>
<b>2.</b>	<b>La produzione eolica .....</b>	<b>8</b>
2.1.	Principi di funzionamento.....	8
2.2.	L'impianto eolico.....	10
<b>3.</b>	<b>La previsione a brevissimo termine di produzione eolica .....</b>	<b>12</b>
3.1.	Caratteristiche delle serie storiche utilizzate .....	12
3.2.	Data Cleaning.....	15
3.3.	Approccio utilizzato.....	18
3.3.1.	Schema generale .....	18
3.3.2.	Processo impiegato.....	18
3.3.3.	Metriche di errore considerate .....	19
3.3.4.	I metodi di analisi e previsione utilizzati .....	21
3.3.4.1.	Richiami sul metodo della trasformata Wavelet.....	21
3.3.4.2.	Applicazione al caso in esame .....	26
3.3.4.3.	Cenni sui modelli ARIMA .....	27
3.3.4.4.	Cenni sul filtro di Kalman .....	28
3.3.4.5.	Cenni sulle reti neurali.....	31
3.3.4.6.	Scomposizione wavelet e modello arima.....	35
3.3.4.7.	Scomposizione wavelet e filtro di Kalman .....	40
3.3.4.8.	Scomposizione wavelet e rete neurale .....	45
3.3.4.9.	Persistenza .....	50
3.4.	Considerazioni.....	51
<b>4.</b>	<b>Riferimenti e Bibliografia .....</b>	<b>53</b>



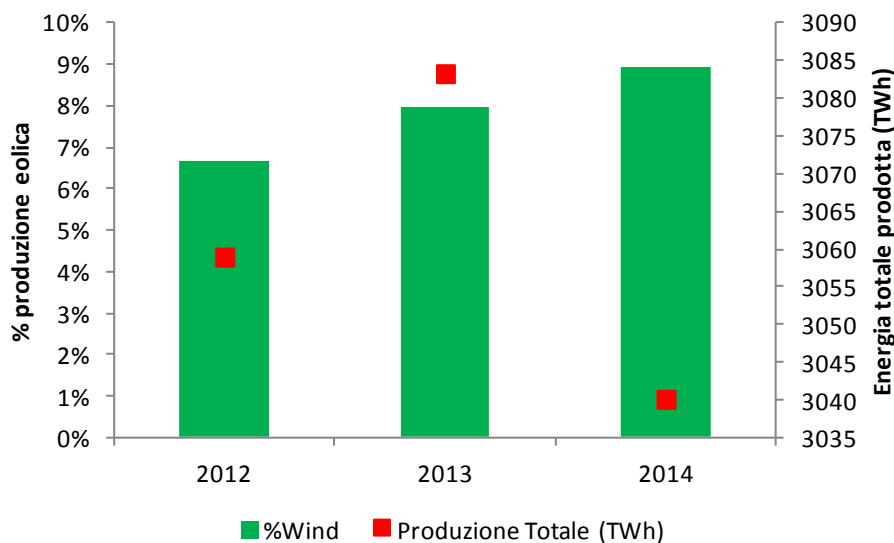
## Indice delle figure

Figura 1 - Quota di produzione eolica rispetto produzione elettrica totale europea .....	6
Figura 2 - Curva di potenza teorica.....	9
Figura 3 - Esempio di layout di impianto eolico.....	10
Figura 4 - Velocità del vento ed energia prodotta da una singola WTG.....	11
Figura 6 - Andamento della velocità del vento e dell'energia prodotta.....	12
Figura 7 - Distribuzione della velocità del vento per singola WTG .....	13
Figura 8 - Distribuzione energia prodotta per singola WTG .....	13
Figura 9 - Boxplot velocità del vento per WTG .....	14
Figura 10 - Boxplot produzione per WTG .....	14
Figura 11 - Curva di potenza dati non processati .....	16
Figura 12 - Curva di potenza dati processati .....	16
Figura 13 - Curva di potenza interpolata.....	17
Figura 14 - Schema generale di processo .....	18
Figura 15 - NMAE vs MAPE .....	20
Figura 16 - Diagramma di flusso dell'algorithmo piramidale .....	22
Figura 17 - Esempi di wavelet: Haar, Daubechies D(4), D(8); Least Asymmetric LA(8).....	25
Figura 18 - Scomposizione Wavelet MODWT Daubechies ordine 4 (D4).....	26
Figura 20 - Pesi di un collegamento neurale.....	31
Figura 19 - Elemento di una rete neurale artificiale .....	31
Figura 21 - Rete MLP feed forward a 2 strati .....	32
Figura 22 - Ricerca del minimo attraverso la discesa gradiente .....	33
Figura 23 - Ricerca del minimo attraverso il gradiente: minimo relativo.....	34
Figura 24 - Errore sulla velocità del vento per WTG (Wavelet + Arima) .....	35
Figura 25 - Errore sull'energia prodotta per WTG (Wavelet + Arima) .....	35
Figura 26 - Esempio di confronto curve velocità del vento reale e prevista (WTG 1- lag 1) .....	36
Figura 27 - Wavelet + Arima, energia prevista vs energia reale LAG 1 (WTG 1 -16) .....	36
Figura 28 - Wavelet + Arima, velocità del vento prevista vs reale LAG 1 (WTG 1 -16).....	36
Figura 29 - Distribuzione errori su tutte le WTG per lag temporale di previsione (Wavelet + Arima) .....	37
Figura 30 - Distribuzione degli errori per impianto (Wavelet – Arima).....	37
Figura 31 - Curve orarie previste e reale di impianto vs NMAE Lag temporale 1 (Wavelet + Arima).....	38
Figura 32 - Curve orarie previste e reale di impianto vs NMAE Lag temporale 2 (Wavelet + Arima).....	38
Figura 33 - Curve orarie previste e reale di impianto vs NMAE Lag temporale 3 (Wavelet + Arima).....	39
Figura 34 - Errore sulla velocità del vento per WTG (Wavelet + Kalman).....	40
Figura 35 - Errore sull'energia prodotta per WTG (Wavelet + Kalman) .....	40
Figura 36 - Wavelet + Kalman, Previsione vs velocità del vento reale LAG 1 (WTG 1 -16).....	41
Figura 37 - Wavelet + Kalman, Previsione vs energia reale LAG 1 (WTG 1 -16) .....	41
Figura 38 - Esempio di confronto curve velocità del vento reale e prevista (WTG 1- lag 1) .....	42
Figura 39 - Distribuzione errori su tutte le WTG per lag temporale di previsione (Wavelet + Kalman) .....	42
Figura 40 - Distribuzione degli errori per impianto (Wavelet – Kalman).....	42
Figura 42 - Curve orarie previste e reale di impianto vs NMAE Lag temporale 2 (Wavelet + Kalman).....	43
Figura 41 - Curve orarie previste e reale di impianto vs NMAE Lag temporale 1 (Wavelet + Kalman).....	43
Figura 43 - Curve orarie previste e reale di impianto vs NMAE Lag temporale 3 (Wavelet + Kalman).....	44
Figura 44 - Errore sulla velocità del vento per WTG (Wavelet + rete neurale).....	45
Figura 45 - Errore sull'energia prodotta per WTG (Wavelet + rete neurale).....	45
Figura 46 - Wavelet + rete neurale, Previsione vs energia prodotta reale LAG 1 (WTG 1 -16) .....	46
Figura 47 - Wavelet + rete neurale Previsione vs velocità del vento reale LAG 1 (WTG 1 -16).....	46
Figura 48 - Esempio di confronto curve velocità del vento reale e prevista (WTG 1- lag 1) .....	47
Figura 49 - Distribuzione errori su tutte le WTG per lag temporale di previsione (Wavelet + rete neurale) .....	47
Figura 50 - Distribuzione degli errori per impianto .....	47
Figura 51 - Curve orarie previste e reale di impianto vs NMAE Lag temporale 1 (Wavelet + rete neurale) .....	48
Figura 52 - Curve orarie previste e reale di impianto vs NMAE Lag temporale 2 (Wavelet + rete neurale).....	48
Figura 53 - Curve orarie previste e reale di impianto vs NMAE Lag temporale 3 (Wavelet + rete neurale).....	49
Figura 55 - Errori persistenza .....	50
Figura 56 - Istogramma errori sulla velocità del vento per WTG .....	50
Figura 57 - Istogramma errori sulla energia per WTG .....	50
Figura 58 – Persistenza Tabella errori.....	50

## 1. Introduzione

La produzione di energia elettrica da fonte rinnovabile ha assunto negli ultimi anni un'importanza sempre maggiore, a fine 2014 circa il 38% della energia elettrica prodotta in Italia proveniva da una fonte rinnovabile (idrica, geotermica, eolica, solare). A livello europeo inoltre, la quota di produzione eolica ha superato lo 8% del totale (vedi Figura [1]).

Figura 1 - Quota di produzione eolica rispetto produzione elettrica totale europea



L'estrema variabilità e quindi la difficoltà di previsione di una produzione elettrica di questo tipo, introducono una complicazione in relazione alla sicurezza del sistema elettrico in quanto, come noto, l'energia elettrica deve essere prodotta nel momento in cui serve in una quantità istantaneamente pari al carico richiesto sulla rete, ogni variazione di produzione o di carico deve essere immediatamente compensata. Per questo motivo gli operatori della rete (TSO) stanno introducendo modalità di esercizio di questi impianti che vanno nella direzione di un controllo sempre maggiore.

Esiste inoltre una caratteristica propria di questo tipo di produzione in relazione al mercato elettrico in quanto la produzione rinnovabile di tipo non programmabile, per la quale cioè non è possibile intervenire sulla risorsa primaria, gode della cosiddetta "priorità di dispacciamento". Questa consiste nel fatto che tale produzione oraria è offerta sul mercato a prezzo zero, quindi implicitamente accettata. Tale produzione non impatta direttamente sul prezzo orario ma ne determina comunque in maniera indiretta il valore "occupando" volume.

Tuttavia gli scenari regolatori stanno cambiando a livello europeo e quindi anche nei mercati nazionali, la quota importante di produzione rinnovabile ed i progressi tecnologici, fanno sì che sia sempre più richiesta una stima molto accurata di quanta energia "intermittente" sarà disponibile in orizzonti temporali di breve e brevissimo termine (qualche ora in avanti).

### Lavori collegati

Nella letteratura tecnica sono stati proposti diversi metodi per la previsione della energia eolica, sia fisici [2] sia statistici [3]. I metodi fisici possono essere vantaggiosi nella previsione di lungo termine, mentre i metodi statistici vanno meglio per il breve termine. I modelli statistici convenzionali sono in genere quelli autoregressivi (AR) e soprattutto gli ARIMA. Spesso può essere utilizzato il metodo della persistenza (proiettare in avanti l'ultimo dato reale disponibile) questo metodo, per il breve e brevissimo termine, rappresenta una utile ed economica approssimazione, rappresenta quindi un benchmark per il confronto [4].

---

Negli ultimi anni sono stati proposti diversi metodi che hanno stimolato attenzione da parte dei ricercatori, tra questi alcuni riportano l'utilizzo di una scomposizione wavelet associata ad un metodo di previsione tradizionale o con rete neurale.

In [5] ad esempio, è proposto un metodo ibrido che vede l'utilizzo di una trasformata wavelet ed una "cascata" di reti neurali per la previsione dei prezzi del mercato elettrico. La serie dei prezzi viene scomposta attraverso una wavelet e, per ciascuna componente, viene addestrata una rete neurale per la previsione. Alle previsioni così ottenute sono poi applicate le trasformazioni inverse per ottenere la serie dei prezzi prevista. Anche in [6] è proposto un metodo per la previsione dei prezzi del mercato elettrico, questo metodo associa la trasformata wavelet e modelli ARIMA. In particolare la scomposizione wavelet fornisce delle componenti più regolari che vengono utilizzate per stimare uno o più modelli ARIMA di previsione, le serie risultanti sono poi, anche qui, utilizzate nella trasformata inversa per ricostruire il comportamento futuro della serie dei prezzi. Infine in [8] viene proposta un'associazione tra rete neurale e trasformata wavelet per la previsione di energia eolica. Anche in questo lavoro si utilizza la trasformata wavelet per scomporre il segnale in parti più facilmente trattabili, una rete neurale prevede il comportamento futuro ed infine, viene utilizzata la trasformata inversa per ricostruire la serie prevista.

Lo scopo di questo lavoro è quindi cercare un metodo di previsione della produzione di un impianto eolico in un orizzonte temporale di brevissimo termine, utilizzando per questo solo dati disponibili nei sistemi di automazione degli impianti, senza fare ricorso quindi a dati di previsione meteorologica esterni. In questo contesto perciò si cerca di stimare dei modelli autoregressivi in grado di fornire una stima di quanto sarà la produzione nelle ore immediatamente successive (1-3).

Come già anticipato, le serie storiche di produzione di un impianto eolico sono caratterizzate da un'estrema variabilità, già nell'intervallo di pochi minuti ci possono essere brusche variazioni di vento e quindi repentine variazioni di produzione, considerando anche il fatto che l'energia elettrica prodotta non è legata alla velocità del vento in modo lineare. Qui si è perciò tentato un approccio che utilizzasse in maniera combinata metodologie proprie di analisi dei segnali, in grado cioè di determinarne le caratteristiche principali e metodologie più o meno classiche di previsione. In particolare è stata usata l'analisi wavelet per la scomposizione del segnale in associazione con modelli ARIMA, Reti Neurali e filtri di Kalman.

Il lavoro è organizzato principalmente in:

- una parte introduttiva che descrive brevemente i principi di funzionamento di un impianto eolico,
- una descrizione delle serie temporali utilizzate e dei criteri di pulizia del dato utilizzati
- una descrizione sintetica delle metodologie utilizzate
- una presentazione e discussione dei risultati

## 2. La produzione eolica

### 2.1. Principi di funzionamento

Le turbine eoliche (o aerogeneratori) convertono direttamente l'energia cinetica del vento in energia meccanica al fine di produrre energia elettrica [9]. La potenza cinetica  $P_D$ , resa disponibile da una prefissata portata massica d'aria  $m$  che si muove a velocità  $u$ , è data da:

$$P_D = \frac{1}{2} \cdot m \cdot u^2$$

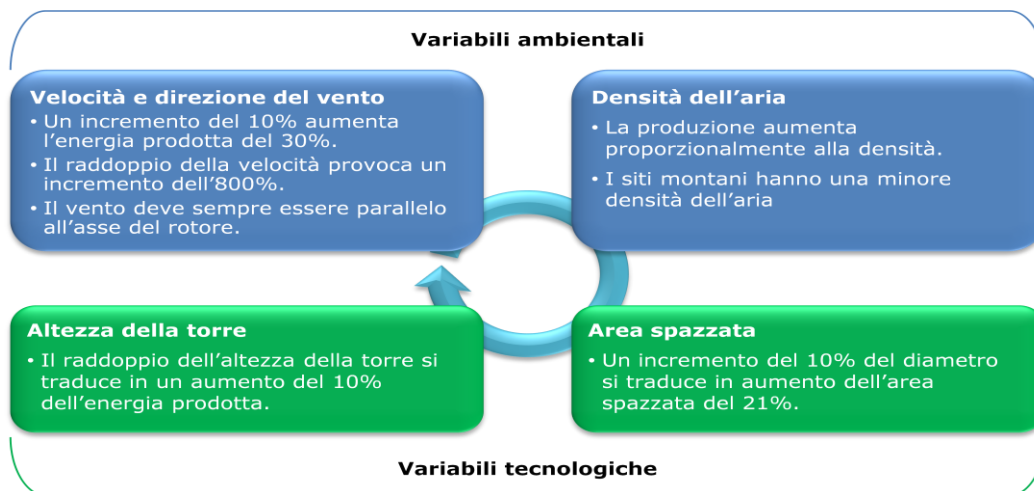
La portata massica è funzione della sezione di passaggio  $S$ , della sua velocità istantanea  $u$ , e della densità dell'aria  $\rho$ .

Pertanto si può scrivere:

$$P_D = \frac{1}{2} \cdot u^2 \cdot (\rho \cdot S \cdot u) = \frac{1}{2} \cdot \rho \cdot S \cdot u^3$$

Tale relazione evidenzia come la potenza eolica disponibile dipenda dal cubo della velocità, dalla sezione di passaggio (ovvero dall'area spazzata dal rotore, quindi funzione del quadrato del diametro) e dalla densità dell'aria (e quindi dalla sua temperatura, pressione ed umidità).

Per avere un'idea degli ordini di grandezza dell'impatto che hanno sulla produzione delle diverse variabili in gioco, si tenga conto che:



L'altezza della torre ed il diametro delle pale sono variabili soltanto in fase di progettazione del sito, mentre sono considerate costanti nel processo di previsione della produzione.

Quindi se si deve prevedere la produzione di energia elettrica da fonte eolica in fase di esercizio, è sufficiente prendere in considerazione le seguenti tre variabili ambientali:

- velocità del vento;
- direzione del vento;
- densità dell'aria.



Non tutta la potenza disponibile però può essere utilizzata. La massima potenza  $P_D$  estraibile da una determinata portata massica è fisicamente limitata da quello che si può considerare l'analogo del rendimento di Carnot per i cicli termodinamici: il limite di Betz. In base ad esso è possibile comprendere come la Potenza Effettiva  $P_E$  si riduca al valore:

$$P_E = P_D \cdot C_P = 0,594 \cdot P_D$$

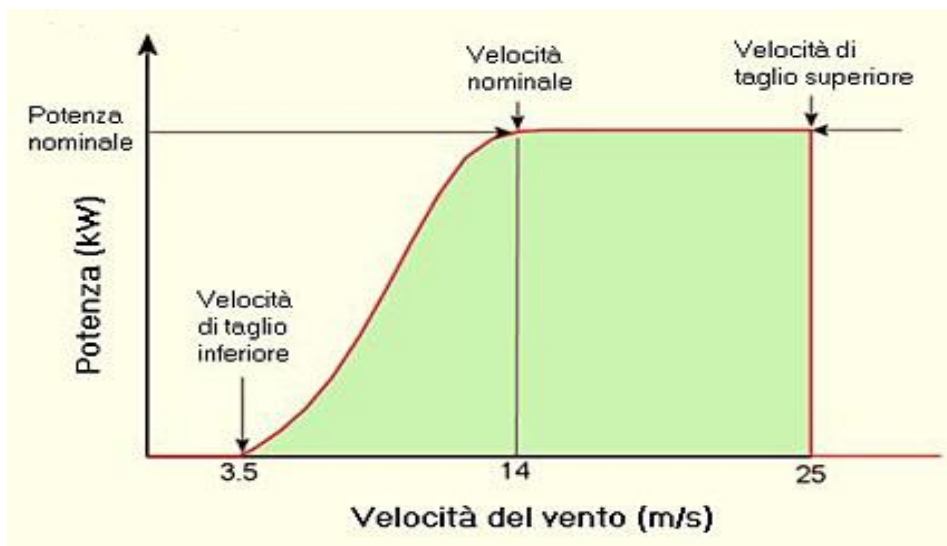
Oltre a tale limite fisico vanno considerati dei limiti tecnici legati alla velocità del vento. Infatti il range di funzionamento di una turbina eolica è legato a:

- Velocità di spunto (o di *cut-in*): è la velocità minima del vento (3-5 m/s a seconda del modello) in grado di avviare la rotazione del rotore;
- Velocità nominale: è la minima velocità (10-15 m/s) che permette alla turbina di fornire la potenza di progetto (appunto la potenza nominale);
- Velocità limite (o di *cut-off*): è il valore massimo della velocità del vento (25-30 m/s) che la turbina può sopportare senza subire danni. Oltre tale valore l'aerogeneratore si pone automaticamente fuori servizio per motivi di sicurezza.

In base a quanto detto finora ed in base ad altri fattori dipendenti dallo specifico modello di aerogeneratore utilizzato, i produttori di turbine eoliche costruiscono la cosiddetta **Curva di Potenza**.

Tale curva mostra come varia la produzione rispetto alla velocità del vento. È possibile distinguere le velocità di cut-in, nominale e di cut-off precedentemente esaminate.

Figura 2 - Curva di potenza teorica



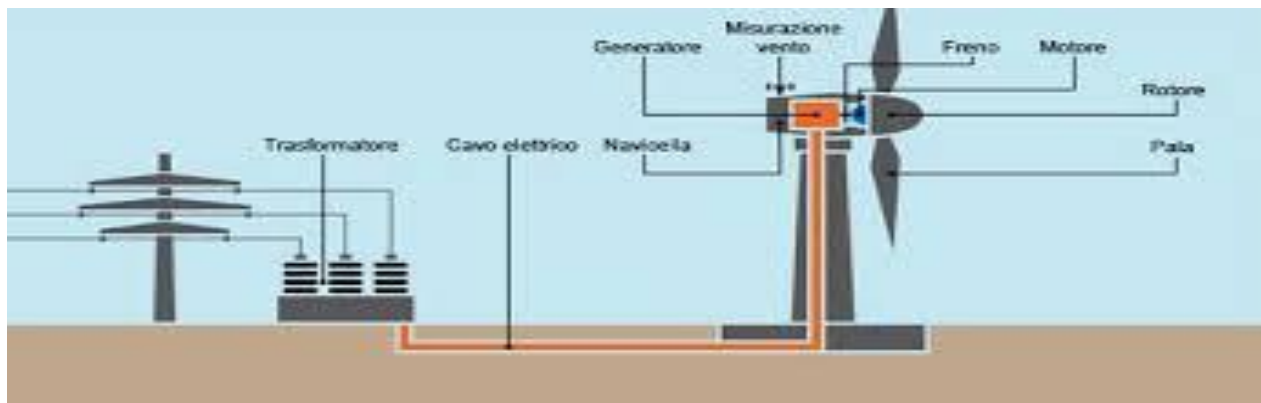
Anche la direzione del vento deve essere presa in considerazione per la valutazione della producibilità di una wind farm. A livello di singola WTG (Wind Turbine Generator) la direzione del vento è un parametro trascurabile in quanto ogni turbina è dotata di un meccanismo d'imbardata o di una banderuola che ne orientano l'asse del rotore in posizione parallela alla direzione del vento. A livello di wind farm, invece, la direzione del vento può determinare l'allineamento di due o più WTG, generando un disturbo aerodinamico e diminuzione della velocità, con conseguente abbattimento della producibilità delle torri presenti sulle file successive. La turbolenza di scia può addirittura creare stress meccanici sulla turbina, oltre che un ritardo nello spunto (ad esempio, se la prima macchina raggiunge i 4 m/s di cut-in, la successiva potrebbe avere un

vento a 3,5 m/s, insufficiente all'avvio della turbina). A tutte queste perdite di potenza legate alla componente vento devono essere aggiunte le perdite meccaniche nelle macchine rotanti e le elettriche nei circuiti dell'alternatore e nella linea di trasmissione.

## 2.2. L'impianto eolico

Un impianto eolico, come accennato prima, è costituito da un certo numero (anche consistente) di WTG distribuite su una porzione sufficientemente ampia di terreno. Semplificando, la produzione elettrica di ogni WTG viene convogliata verso un unico punto dove avviene l'immissione nella rete elettrica. In corrispondenza di questo punto viene installato un apparato di misura (contatore) che rileva l'energia effettiva che transita verso la rete. Dal punto di vista del mercato elettrico, è questa la grandezza di interesse. Su ciascuna WTG però sono installati degli apparati di misura che rilevano, tra le altre grandezze, la velocità, la direzione del vento e l'energia prodotta dall'alternatore ad intervalli di tempo molto piccoli (inferiori al secondo). Tali grandezze sono necessarie al sistema della WTG per effettuare il controllo in tempo reale. Queste grandezze sono poi inviate ad un sistema centrale per la loro memorizzazione come media sui dieci minuti.

**Figura 3 - Esempio di layout di impianto eolico**

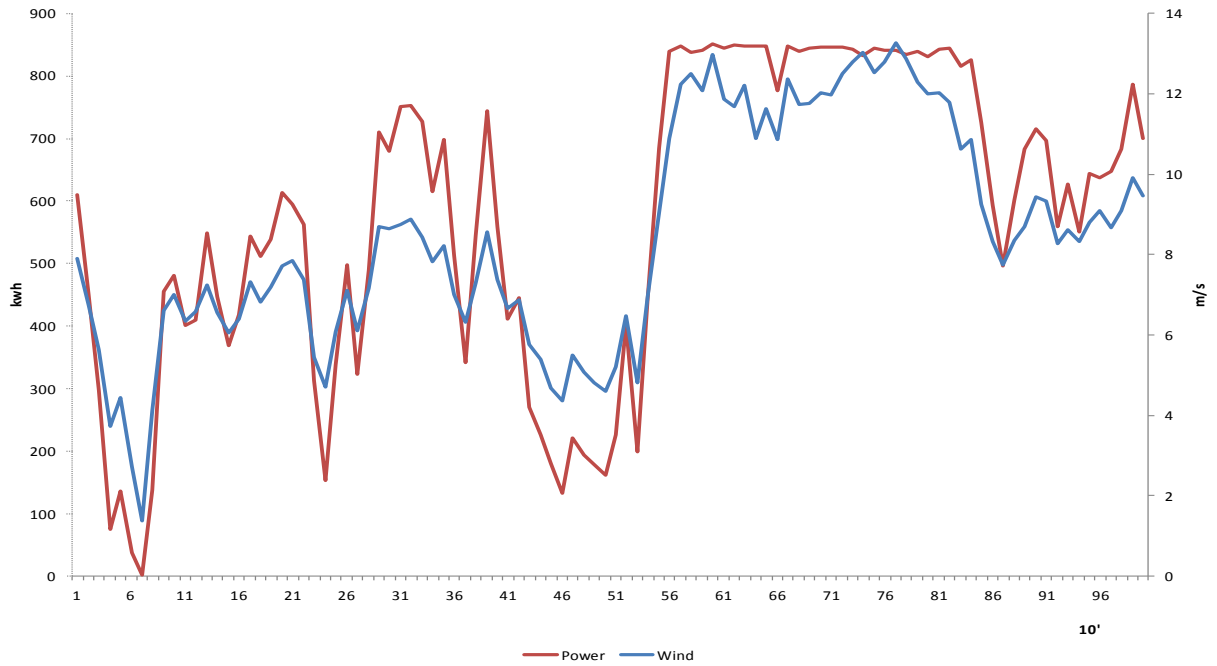


In corrispondenza dell'alternatore posto in ogni WTG quindi viene effettuata la misura di energia prodotta, sulla sommità della torre invece è presente l'anemometro che misura la velocità del vento. Tale misura non è molto precisa perché lo strumento è posto spesso in una posizione che subisce qualche interferenza da parte di altre componenti dell'impianto (come ad esempio le pale). Tuttavia sono grandezze sicuramente disponibili in ogni impianto e comunque sono coerenti tra loro (Fig. 4).

Le serie temporali utilizzate in questo lavoro provengono da queste misurazioni. Nel punto di connessione con la rete invece, è presente il contatore ufficiale che misura l'energia oraria netta immessa.

Nel grafico in figura 4 vengono riportate a titolo di esempio, la curva di velocità del vento e la curva di produzione ai 10 minuti (su scale diverse) di una singola WTG. Si nota che, quando la velocità del vento supera un determinato valore, la corrispondente curva di produzione elettrica si "appiattisce", questa è già un'indicazione di come sia preferibile lavorare sulla velocità del vento anziché direttamente sulla curva di produzione.

Figura 4 - Velocità del vento ed energia prodotta da una singola WTG



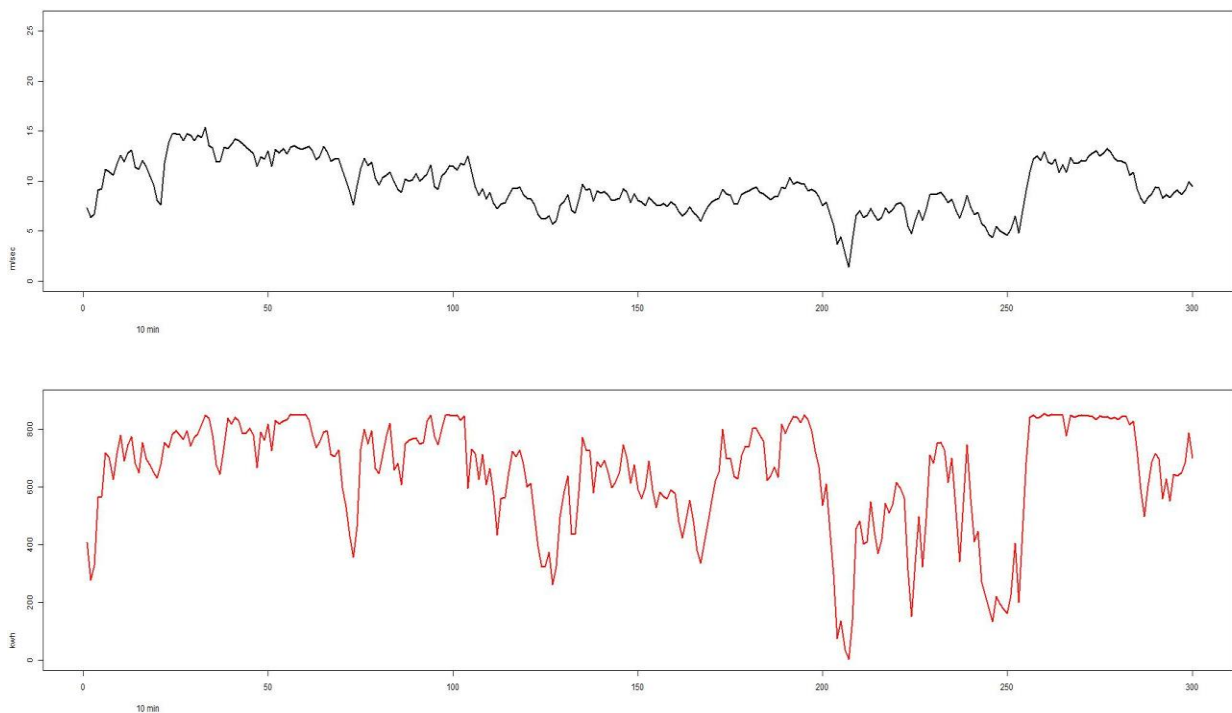
### 3. La previsione a brevissimo termine di produzione eolica

#### 3.1. Caratteristiche delle serie storiche utilizzate

In questo lavoro saranno prese in considerazione le serie temporali relative alla velocità del vento ed alla energia prodotta in un anno dalle WTG di un impianto eolico con potenza nominale complessiva pari a circa 27MW. L'impianto si compone di 32 WTG ciascuna di potenza nominale pari a 850kW.

I dati costituenti le serie sono stati acquisiti dal sistema di controllo e memorizzati come media ai **10** minuti, ogni serie quindi contiene **52554** campioni. Nei grafici seguenti si mostra l'andamento tipico delle serie e le loro distribuzioni.

Figura 5 - Andamento della velocità del vento e dell'energia prodotta



Dai grafici è possibile vedere come sia variabile nel tempo questa serie e soprattutto il differente comportamento della serie di velocità del vento rispetto alla serie di produzione a causa del legame non lineare che c'è tra le due grandezze.

Figura 6 - Distribuzione della velocità del vento per singola WTG

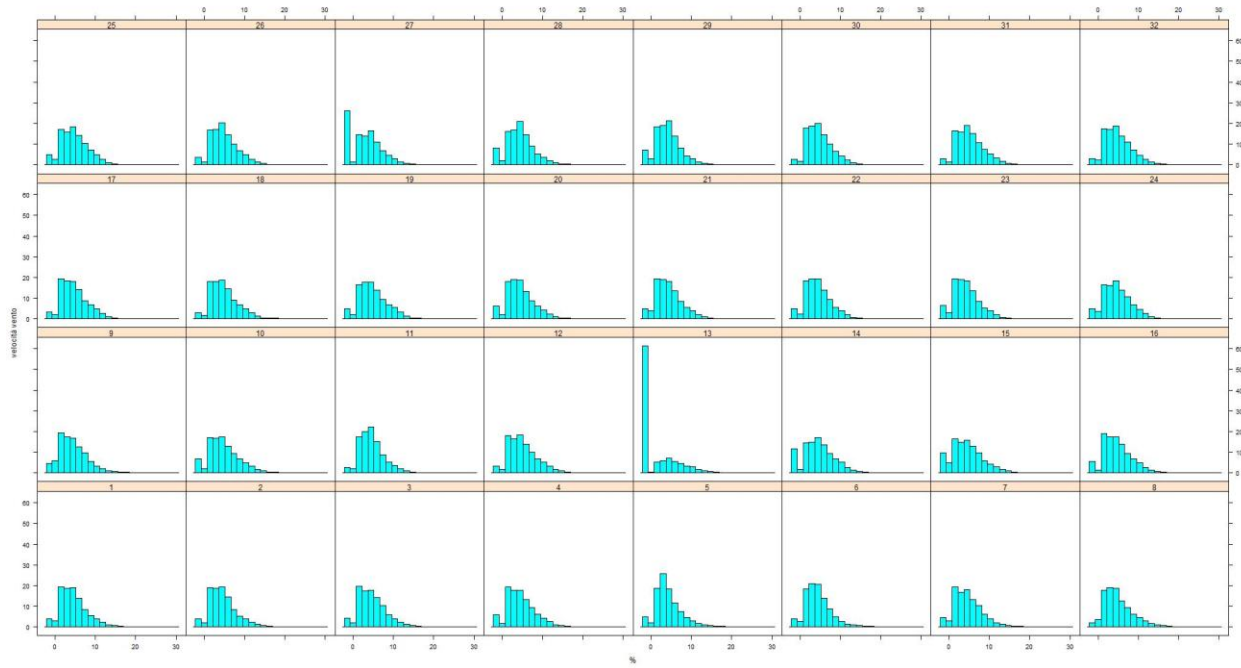


Figura 7 - Distribuzione energia prodotta per singola WTG

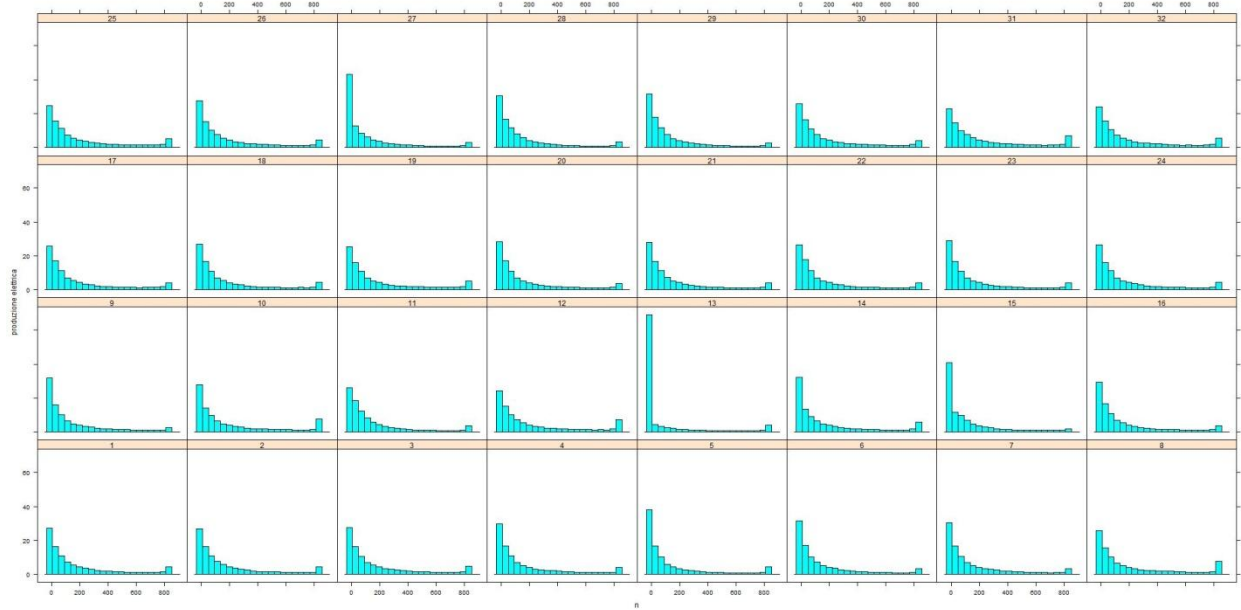


Figura 8 - Boxplot velocità del vento per WTG

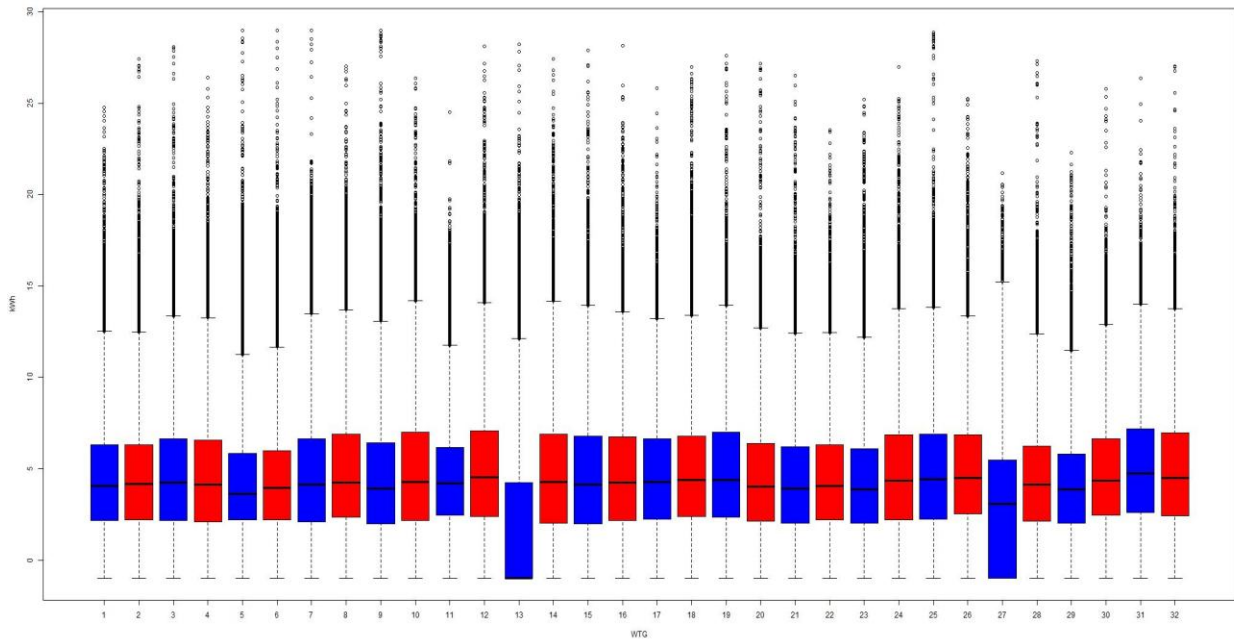
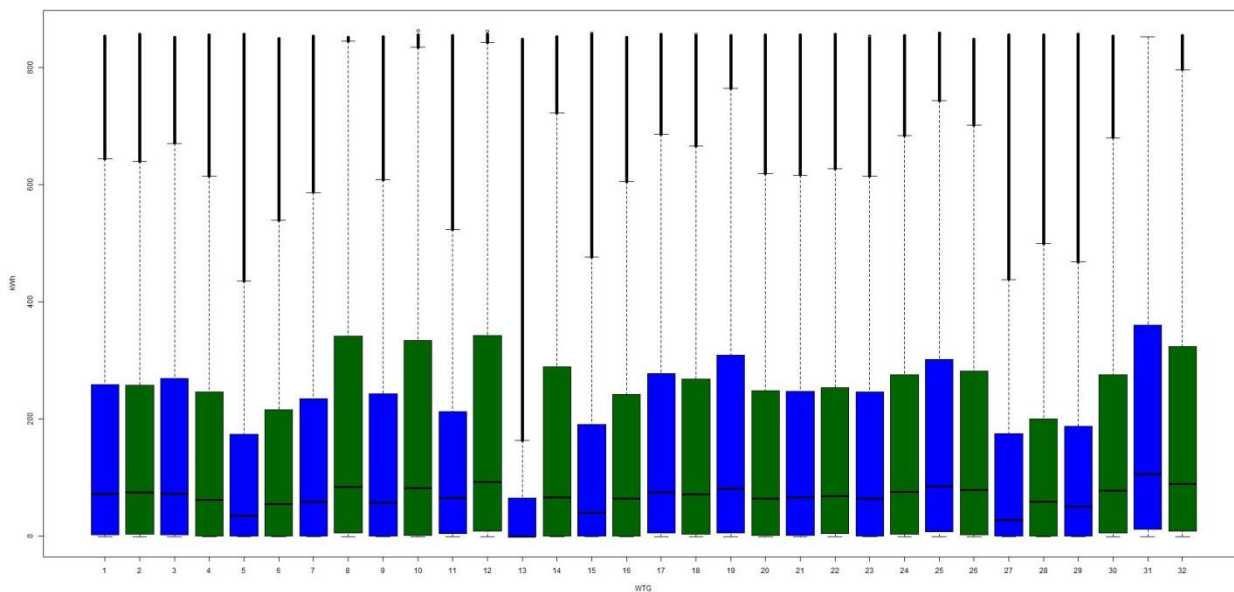


Figura 9 - Boxplot produzioni per WTG



Questo tipo di rappresentazione è molto sintetica ma fornisce, a colpo d'occhio, un'idea delle principali grandezze statistiche delle serie in esame. In particolare si notano le caratteristiche proprie della produzione eolica (almeno in Italia) cioè una produzione sbilanciata verso i valori bassi ed un'alta varianza.

Si nota innanzitutto che una WTG (la numero 13) presenta una distribuzione molto più sbilanciata verso lo zero rispetto alle altre, questo è probabilmente dovuto a malfunzionamenti della WTG stessa o del sistema di acquisizione

E' noto inoltre, che la velocità del vento presenta una distribuzione (teorica) di tipo Weibull che sembra sufficientemente rispettata a meno degli errori presenti sui dati.

Dai grafici sembra evidente come la serie relativa alla velocità del vento sia più regolare di quella relativa alla energia prodotta, questa è una conseguenza della forte non linearità che lega le due grandezze (cfr.

---

curva di potenza). Tale evidenza suggerisce di lavorare soprattutto sulla velocità del vento riconducendo poi questa all'energia. Per fare questo però è necessario costruire la curva di potenza reale di ogni WTG. Un altro aspetto importante è la qualità del dato che non è mai molto elevata, nel paragrafo successivo saranno descritti i criteri di data clearing adottati.

### **3.2. Data Cleaning**

I dati prodotti dal sistema di controllo di un impianto sono praticamente sempre disponibili, non è necessario installare una strumentazione specifica con relativo sistema di acquisizione, la disponibilità di tali dati quindi non comporta costi aggiuntivi, per contro però non presentano una qualità particolarmente buona. Il principale difetto riscontrato, è la mancata lettura o mancata trasmissione del dato al sistema di raccolta, questo può manifestarsi con una segnalazione di errore nella variabile interessata o con l'impostazione di un valore impossibile. Queste situazioni sono facilmente riconoscibili e quando si incontrano, è necessario scartare l'occorrenza.

Dal momento che siamo interessati alla costruzione della curva di potenza, è necessario scartare tutte e due le variabili anche nel caso in cui una fosse valida.

Un caso più difficile da evidenziare si ha quando il sistema di acquisizione si blocca ma la variabile rimane impostata sull'ultimo valore valido, in questo caso si ha una sequenza ripetuta di uno stesso numero con tutte le sue cifre decimali. Una sequenza del genere è altamente improbabile che sia reale, questa caratteristica rende la situazione riconoscibile anche se in modo più complesso.

Abbiamo visto che la produzione di energia elettrica è consentita alla macchina solo nel suo range di funzionamento, per evitare di dover trattare dati inutilmente si scartano i dati di vento al di sotto della velocità di cut-in ed al di sopra della velocità di cut-off

Quando invece i dati superano i controlli precedenti, è necessario sapere se in quel momento l'impianto era in pieno esercizio oppure no. E' possibile di rilevare tale situazione verificando che, a fronte di una velocità del vento all'interno del range consentito la variabile di energia prodotta sia maggiore di zero, altrimenti si scarta.

Infine, superati tutti i controlli di sopra si applica un filtro di tipo statistico, suddividendo l'intero range di energia presente in intervalli sufficientemente piccoli. Su ciascun intervallo poi si calcola la mediana e la deviazione standard, scartando i dati che cadono al di fuori di un intervallo di ampiezza  $+ 0 - 2$  volte la deviazione standard.

Questi criteri sono applicati sulle serie di ogni WTG, una volta effettuata la pulizia si può mettere in relazione la velocità del vento con l'energia prodotta, interpolando attraverso una curva SPLINE è possibile determinare la curva di potenza reale di ciascuna macchina. Di seguito sono mostrati tre grafici scatterplot, il primo presenta i dati "grezzi" e si vede come siano dispersi. Il secondo lo stesso scatterplot con i dati puliti. Il terzo la curva di potenza interpolata.

Figura 10 - Curva di potenza dati non processati

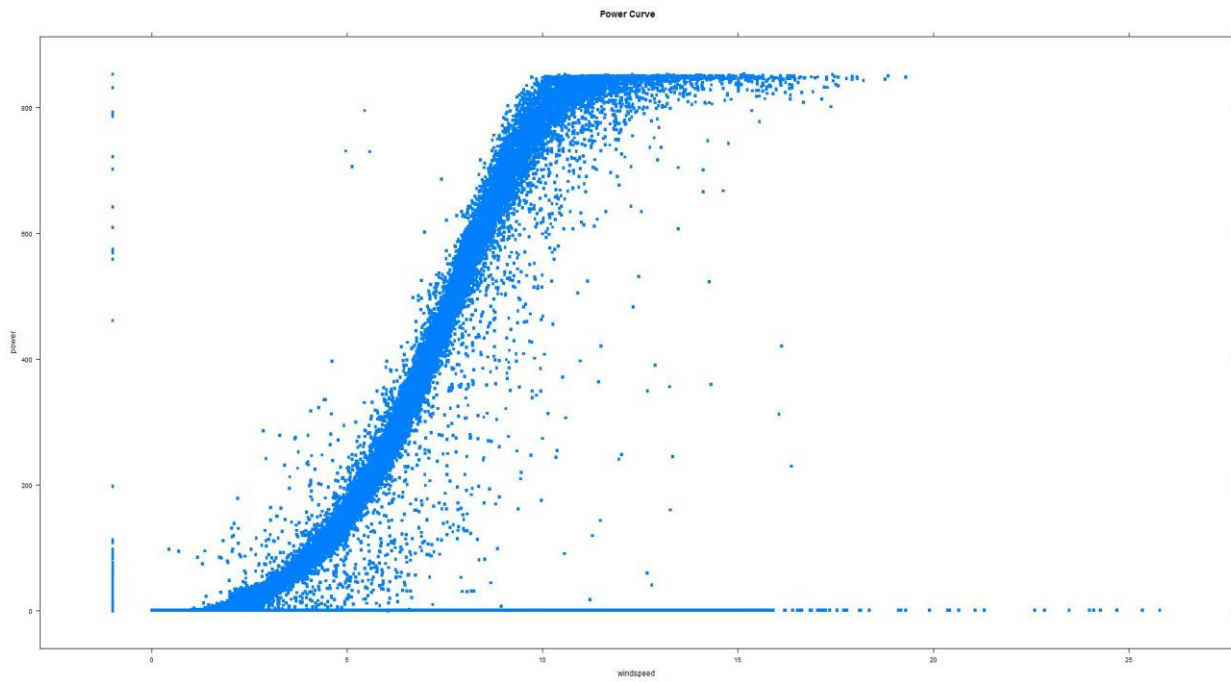


Figura 11 - Curva di potenza dati processati

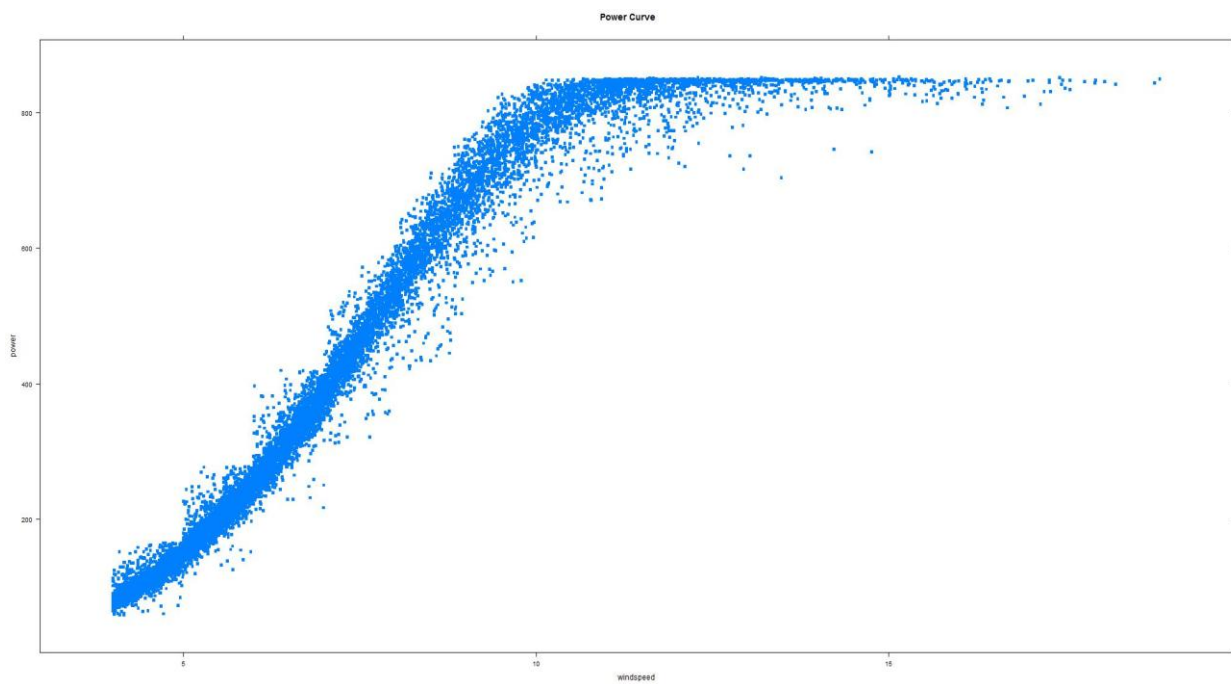
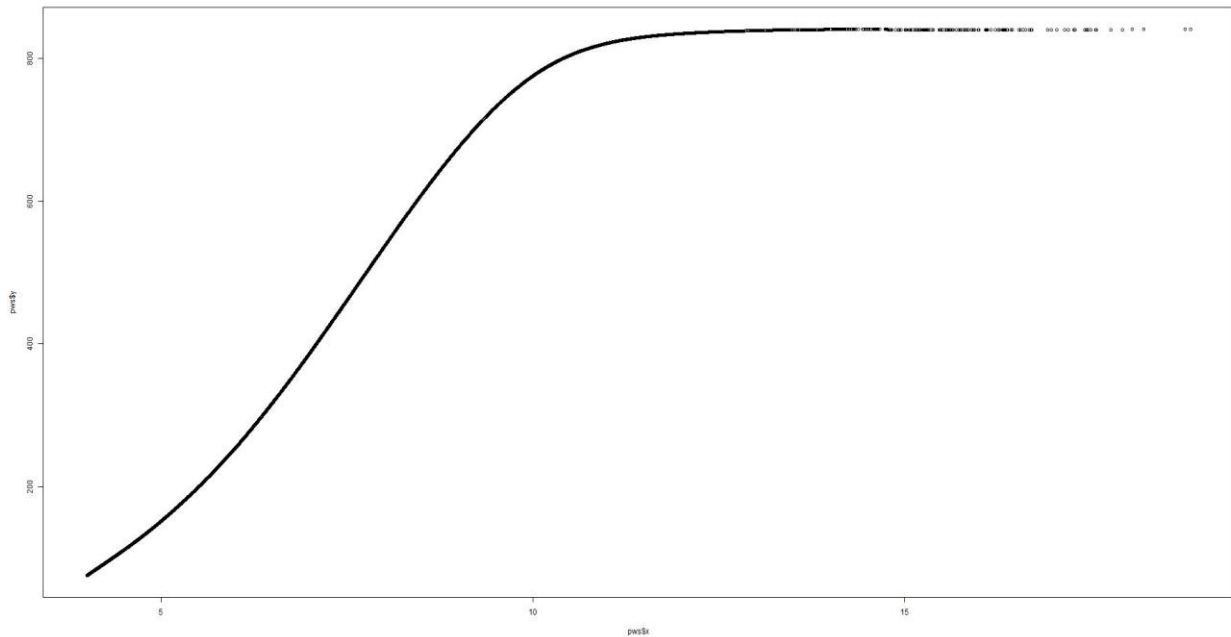




Figura 12 - Curva di potenza interpolata



I criteri descritti prima valgono per la costruzione della curva di potenza ma per la stima ed il test dei modelli previsionali è sufficiente prendere in considerazione la sola serie di vento, applicando quindi solo i criteri relativi alla attendibilità della misura. Dal momento però che siamo interessati alla produzione complessiva dell'impianto, sarà necessario prendere in considerazione le porzioni di serie in comune a tutte le WTG e che presentano continuità, senza cioè "periodi di buco". I valori in comune tra tutte le WTG quindi sono divise in blocchi di serie continue, utilizzando un blocco per la stima ed uno per la verifica.

### 3.3. Approccio utilizzato

#### 3.3.1. Schema generale

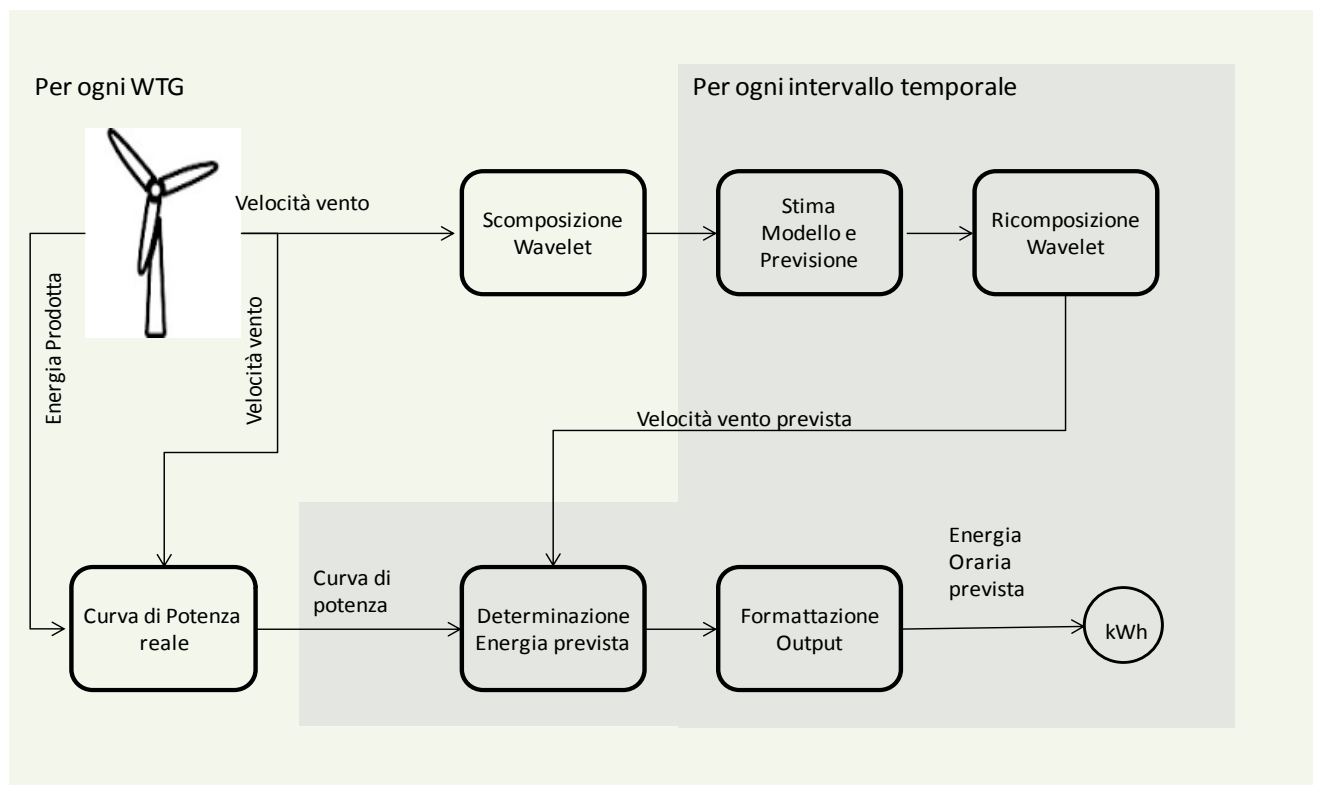
Le caratteristiche delle serie a disposizione, hanno suggerito il tentativo di utilizzo di un approccio misto. Una prima elaborazione in grado di rendere la serie meno irregolare ed una successiva previsione. In questo senso si sono utilizzate le wavelet in quanto, attraverso la sua componente di “approssimazione”, fornisce dei risultati più regolari. Successivamente, la previsione vera e propria, è stata tentata utilizzando tre modelli:

- Rete neurale
- ARIMA
- Filtro di Kalman

#### 3.3.2. Processo impiegato

Di seguito uno schema che descrive il processo generale:

Figura 13 - Schema generale di processo



Tale processo prevede quindi:

1. acquisizione delle serie temporali di misura della energia prodotta e della velocità del vento misurate nella WTG
2. processamento dei dati per escludere outliers e tutti i dati non attendibili
3. determinazione della curva di potenza reale della macchina
4. individuazione delle porzioni continue di serie in comune con tutte le WTG

5. suddivisione del campione in una parte da utilizzare per la stima del modello ed in una parte per la sua validazione
6. per ogni WTG:
  - a. scomposizione in wavelet del campione di stima
  - b. per ogni intervallo temporale di previsione:
    - i. stima del modello
    - ii. previsione per l'orizzonte temporale voluto
    - iii. acquisizione del nuovo dato orario reale

### 3.3.3. Metriche di errore considerate

Per la valutazione delle prestazioni dei modelli è stata utilizzata una porzione di serie temporale diversa da quella impiegata per la stima dei modelli. Le prestazioni sono state calcolate considerando due tipologie di errore.

- NMAE: Normalized Mean Average Error per la produzione di energia elettrica

$$\frac{1}{NC} \cdot \frac{\sum_{h=1}^H |Em_h - Pv_h|}{H}$$

Dove:

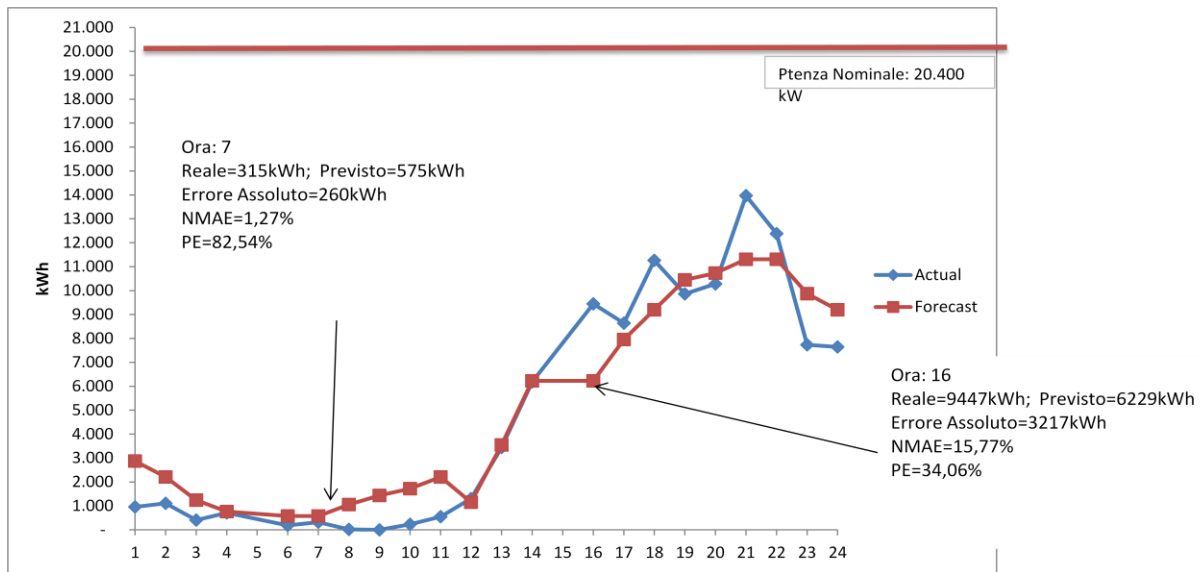
- $Em_h$  è l'energia prodotta nell'ora h (kWh)
- $Pv_h$  è l'energia prevista per l'ora h (kWh)
- NC potenza nominale della macchina o dell'impianto (kW)
- H numero di ore nel periodo in esame
  
- MAPE: Mean Average Percentage Error per la velocità del vento

$$\frac{1}{H} \cdot \sum_{h=1}^H \frac{|Vm_h - Vp_h|}{Vm_h}$$

- $Vm_h$  è la velocità del vento misurata nell'ora h (m/s)
- $Vp_h$  è la velocità del vento prevista per l'ora h (m/s)

Nel caso della produzione è preferibile l'utilizzo dell'errore NMAE a causa della non linearità del fenomeno, il tradizionale MAPE rischierebbe di non valutare correttamente l'entità dell'errore (vedi figura 15)

Figura 14 - NMAE vs MAPE



Nel grafico si vede che nell'ora 7 (di bassa produzione) l'errore percentuale risulta più grande che nell'ora 16 (di produzione più alta) ma in realtà meno significativo. L'errore NMAE invece fornisce un'indicazione più corretta dell'entità dell'errore.

A livello di impianto sono utilizzati criteri analoghi con la differenza che le energie sono calcolate come somma delle energie delle singole WTG. In questo caso non viene calcolato l'errore sulla velocità del vento.

$$Em_h^T = \sum_{i=1}^{NWTG} Em_h^i$$

$$Pv_h^T = \sum_{i=1}^{NWTG} Pv_h^i$$

$$\frac{1}{NC} \cdot \frac{\sum_{h=1}^H |Em_h^T - Pv_h^T|}{H}$$

- $Em_h^i$  è l'energia prodotta nell'ora h dalla WTG i (kWh)
- $Pv_h^i$  è l'energia prevista per l'ora h dalla WTG i (kWh)
- $Em_h^T$  è l'energia totale prodotta nell'ora h dall'impianto (kWh)
- $Pv_h^T$  è l'energia totale prevista per l'ora h per tutto l'impianto (kWh)
- NC potenza nominale dell'impianto (kW)
- H numero di ore nel periodo in esame
- NWTG numero di WTG nell'impianto

### 3.3.4. I metodi di analisi e previsione utilizzati

#### 3.3.4.1. Richiami sul metodo della trasformata Wavelet

In generale l'obiettivo dell'analisi basata sulle wavelet è quello di convertire le informazioni di un segnale in componenti che possono essere manipolate, immagazzinate, trasmesse, analizzate o usate per ricostruire il segnale originale, con l'importante caratteristica che un numero limitato di componenti può contenere informazioni rilevanti.

Tra i possibili utilizzi nei più svariati settori, la trasformata wavelet, nella forma continua o discreta, è particolarmente adatta all'analisi delle serie storiche caratterizzate da non-linearità. Tale tecnica consente di analizzare ed esplorare dettagli di dati che altre metodologie d'analisi non permettono di esaminare, quali le tendenze, le discontinuità o i punti di rottura.

Un'area importante d'applicazione consiste nell'utilizzo della trasformata wavelet nel preprocessing dei dati. L'idea è quella di usare le wavelet per decomporre i dati in sottoinsiemi che siano più facili da modellare infatti, l'utilizzo della trasformata wavelet può consentire di ridurre il contenuto di rumore presente nel segnale. Le componenti decomposte del segnale trasformato possono essere utilizzate come ingressi "semplificati" da fornire ad esempio ad una rete neurale oppure a modelli auto regressivi.

Una trasformata wavelet continua CWT può essere definita come:

$$CVT_f(a, b) = \int_{-\infty}^{+\infty} f(t)\Psi_{a,b}^*(t)dt$$

con:

$$\Psi_{a,b}^*(t) = \frac{1}{\sqrt{a}} \Psi\left(\frac{t-b}{a}\right)$$

La funzione  $\psi(t)$  è la wavelet madre, ed i parametri  $a$  e  $b$  sono detti parametri di scala ( $a$ ) e di traslazione ( $b$ ). Una wavelet è una forma d'onda limitata nel tempo a valore medio nullo. La trasformata wavelet scompone il segnale in versioni scalate e traslate della wavelet madre. Una volta che è stato fissato il fattore di scala  $a$ , i coefficienti della trasformata wavelet rappresentano la correlazione tra le wavelet e il segnale di origine. Per valori grandi di  $a$ , si ottengono versioni dilatate (lunghe e a bassa frequenza) della wavelet madre, mentre per piccoli valori di  $a$  avremo wavelet contratte (corte ed ad alta frequenza). Il fattore di scala è l'elemento che rende questo tipo d'analisi flessibile: a valori bassi di scala corrispondono wavelet più compresse, che ben si adattano alla rappresentazione di rapidi cambiamenti del segnale, mentre fattori di scala alti allungano le onde rendendole simili a segnali di bassa frequenza. Il parametro  $b$  è il fattore che determina la traslazione della wavelet nel tempo. La trasformata wavelet continua calcola i coefficienti wavelet ad ogni possibile scala, lungo tutto l'intervallo di tempo del segnale. Ciò comporta una mole di dati notevole, ridondante ai fini della descrizione del segnale, esistono quindi alcune difficoltà con la CWT che la rendono più adatta allo studio delle funzioni piuttosto che dei segnali o serie temporali.

E' possibile semplificare la trasformata wavelet imponendo che i parametri assumano valori discreti, in questo caso si parla di trasformata wavelet discreta (DWT) che si basa sugli stessi principi della CWT ma è più facilmente trattabile in applicazioni pratiche.

La differenza fondamentale tra CWT e DWT è che quest'ultima utilizza solo un limitato numero di versioni dilatate e traslate della wavelet madre per decomporre il segnale originale.

#### Discrete Wavelet Transform (DWT)

Supponiamo di avere la serie:

$$\mathbf{X} = (X_0, \dots, X_{N-1}) = (X_t) \quad t = 0, \dots, N-1 ; \quad N = 2^n \quad (4.1.1)$$

Questa è una serie diadica, cioè presenta una scala del tipo  $2^n$ , quindi se il segnale originale ha  $T$  osservazioni, il più grande numero di scale per la DWT è uguale ad un intero  $J$  tale che  $J = \log_2 T$  ciò significa che non è possibile applicare direttamente la DWT se la serie originale non presenta tale caratteristica. Sarà quindi necessario togliere (o aggiungere) osservazioni.

La DWT si basa su due filtri discreti:

$$\mathbf{h}_l = (h_0, \dots, h_{L-1})$$

$$\mathbf{g}_l = (g_0, \dots, g_{L-1})$$

Con le seguenti proprietà:

$$\sum_{l=0}^{L-1} h_l = 0; \quad \sum_{l=0}^{L-1} h_l^2 = 1; \quad \sum_{l=0}^{L-1} h_l h_{l+2n} = 0 \quad \text{per tutti gli interi } n \neq 0 \quad (4.1.2)$$

$$\sum_{l=0}^{L-1} g_l = \sqrt{2}; \quad \sum_{l=0}^{L-1} g_l^2 = 1; \quad \sum_{l=0}^{L-1} g_l g_{l+2n} = 0 \quad \text{per tutti gli interi } n \neq 0 \quad (4.1.3)$$

$$g_l = (-1)^{L+1} h_{L-1-l}; \quad \sum g_l h_{l+2n} = 0 \quad (4.1.4)$$

Si può quindi definire l'algoritmo piramidale, calcolando:

$$W_{1,t} = \sum_{l=0}^{L-1} h_l X_{2t+1-l \bmod N} \quad (4.1.5)$$

$$V_{1,t} = \sum_{l=0}^{L-1} g_l X_{2t+1-l \bmod N} \quad (4.1.6)$$

La prima si comporta come un filtro passa alto, la seconda come un filtro passa basso

$$W_{2,t} = \sum_{l=0}^{L-1} h_l V_{1,2t+1-l \bmod N_1} \quad (4.1.7)$$

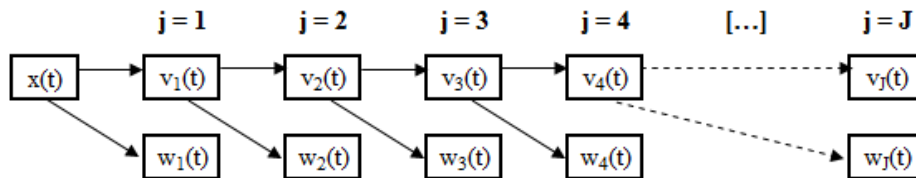
$$V_{2,t} = \sum_{l=0}^{L-1} g_l V_{1,2t+1-l \bmod N_1} \quad (4.1.8)$$

$$W_{j,t} = \sum_{l=0}^{L-1} h_l V_{j-1,2t+1-l \bmod N_{j-1}} \quad (4.1.9)$$

$$V_{j,t} = \sum_{l=0}^{L-1} g_l V_{j-1,2t+1-l \bmod N_{j-1}} \quad (4.1.10)$$

Con  $t = 0, \dots, N_{j-1}$   $N_j = \frac{N}{2^j}$

**Figura 15 - Diagramma di flusso dell'algoritmo piramidale**



L'applicazione ripetuta dell'algoritmo piramidale (vedi figura precedente) fornisce una serie per ogni  $J_0$  con le seguenti proprietà:

$$\mathbf{W} = (\mathbf{W}_1, \mathbf{W}_2, \dots, \mathbf{W}_{J_0}, \mathbf{V}_{J_0}) = \mathbb{W}\mathbf{X} \quad (4.1.11)$$

Con  $\mathbb{W}$  matrice ortogonale, cioè  $\mathbb{W}^{-1} = \mathbb{W}^T$   
e

$$\|\mathbb{W}\|^2 = \|\mathbf{X}\|^2; \quad \mathbf{X} = \mathbb{W}^{-1}\mathbf{W} \quad (4.1.12)$$

Quindi

$$||\mathbf{X}||^2 = \sum_{j=1}^{J_0} ||\mathbf{W}_j||^2 + ||\mathbf{V}_{J_0}||^2 \quad (4.1.13)$$

con  $\mathbf{W}_{j,t}$  = coefficienti wavelet,  $\mathbf{V}_{j,t}$  = coefficiente di scaling

Maximal Overlap Discrete Wavelet Transform (MODWT)

La DWT standard presenta alcuni inconvenienti. Primo richiede una serie di lunghezza diattica (cioè  $N=2^J$ ) secondo non è invariante per translazione, cioè se si sposta la serie un periodo più avanti, i coefficienti sono diversi. Terzo, questo può introdurre sfasamenti nei coefficienti wavelet, ad esempio i picchi potrebbero non essere allineati. Per risolvere questi problemi è stata introdotta la MODWT. Questo tipo di trasformata può gestire qualsiasi dimensione del campione, ha una maggiore risoluzione a scale grossolane, in confronto alla DWT ed è invariante per translazione. In realtà la principale differenza tra DWT e MODWT sta nel fatto che la MODWT considera tutte le translazioni intere quindi, ad ogni frequenza, presenta una risoluzione completa della serie. Qualsiasi sia la scala considerata, la lunghezza dei vettori dei coefficienti wavelet e scaling sarà uguale alla lunghezza della serie originale. Nella MODWT quindi, al primo livello di decomposizione i coefficienti wavelet e scaling si ottengono come:

$$\tilde{w}_{1,t} = \sum_{l=0}^{L-1} h_l x_{t-l \bmod N} \quad (4.1.14)$$

$$\tilde{v}_{1,t} = \sum_{l=0}^{L-1} g_l x_{t-l \bmod N} \quad (4.1.15)$$

Anche in questo caso si può applicare l'algoritmo piramidale, ottenendo:

$$\tilde{w}_{j,t} = \sum_{l=0}^{L-1} h_l \tilde{v}_{j-1,t-2^{j-1}l \bmod N} \quad (4.1.16)$$

$$\tilde{v}_{j,t} = \sum_{l=0}^{L-1} g_l \tilde{v}_{j-1,t-2^{j-1}l \bmod N} \quad (4.1.17)$$

Multiresolution Analysis (MRA)

La Multiresolution analysis può essere utilizzata per ricostruire la serie temporale originale partendo dai vettori di scaling e wavelet. Per ottenere questo è necessario applicare l'inversa della MODWT su  $\tilde{v}_{j,t}$  e  $\tilde{w}_{j,t}$  con  $j = 1, \dots, J$  utilizzando l'algoritmo piramidale.

Al primo passaggio, si applica l'inversa della MODWT ai coefficienti di scaling e wavelet alla scala J, per ripristinare il coefficiente di scaling a J-1

$$\tilde{v}_{j-1,t} = \sum_{l=0}^{L-1} h_l \tilde{w}_{j,t+2^{j-1}l \bmod N} + \sum_{l=0}^{L-1} g_l \tilde{v}_{j,t+2^{j-1}l \bmod N} \quad (4.1.18)$$

Possiamo riscrivere la parte a destra dell'equazione precedente come:

$$\tilde{v}_{j-1,t} = \tilde{B}_j \tilde{w}_j + \tilde{A}_j \tilde{v}_j \quad (4.1.19)$$

Estendiamo questo risultato ad ogni scala j, per esempio alla scala j=1 si ha:

$$\tilde{v}_{1,t} = \tilde{B}_2 \tilde{w}_2 + \tilde{A}_2 \tilde{v}_2 = \tilde{B}_2 \tilde{w}_2 + \tilde{A}_2 (\tilde{B}_3 \tilde{w}_3 + \tilde{A}_3 \tilde{v}_3) = \tilde{B}_2 \tilde{w}_2 + \tilde{A}_2 \tilde{B}_3 \tilde{w}_3 + \dots + \tilde{A}_2 \dots \tilde{A}_{j-1} \tilde{A}_j \tilde{v}_j$$

Ricordando

$$x = \tilde{B}_1 \tilde{w}_1 + \tilde{A}_1 \tilde{v}_1$$

Si ottiene:

$$x = \tilde{B}_1 \tilde{w}_1 + \tilde{A}_1 \tilde{B}_2 \tilde{w}_2 + \tilde{A}_1 \tilde{A}_2 \tilde{B}_3 \tilde{w}_3 + \dots + \tilde{A}_1 \dots \tilde{A}_{j-1} \tilde{A}_j \tilde{v}_j \quad (4.1.20)$$

Ponendo

$$D_j = \tilde{A}_1 \dots \tilde{A}_{j-1} \tilde{B}_j \tilde{w}_j \quad (4.1.21)$$

e

$$S_j = \tilde{A}_1 \dots \tilde{A}_{j-1} \tilde{A}_j \tilde{v}_j \quad (4.1.22)$$

Possiamo ricostruire la serie originale come:

$$x = S_j + D_j + \dots + D_1 \quad (4.1.23)$$

Questa ricostruzione è nota come *Multiresolution Analysis* (MRA) gli elementi  $S_j$  sono collegati ai coefficienti di scaling alla massima scala e quindi rappresentano le componenti *smooth* di  $x$ . Gli elementi  $D_j$  sono i coefficienti di *dettaglio* di  $x$  alla scala  $j$ .

#### Scelta del filtro Wavelet (esempi di funzione)

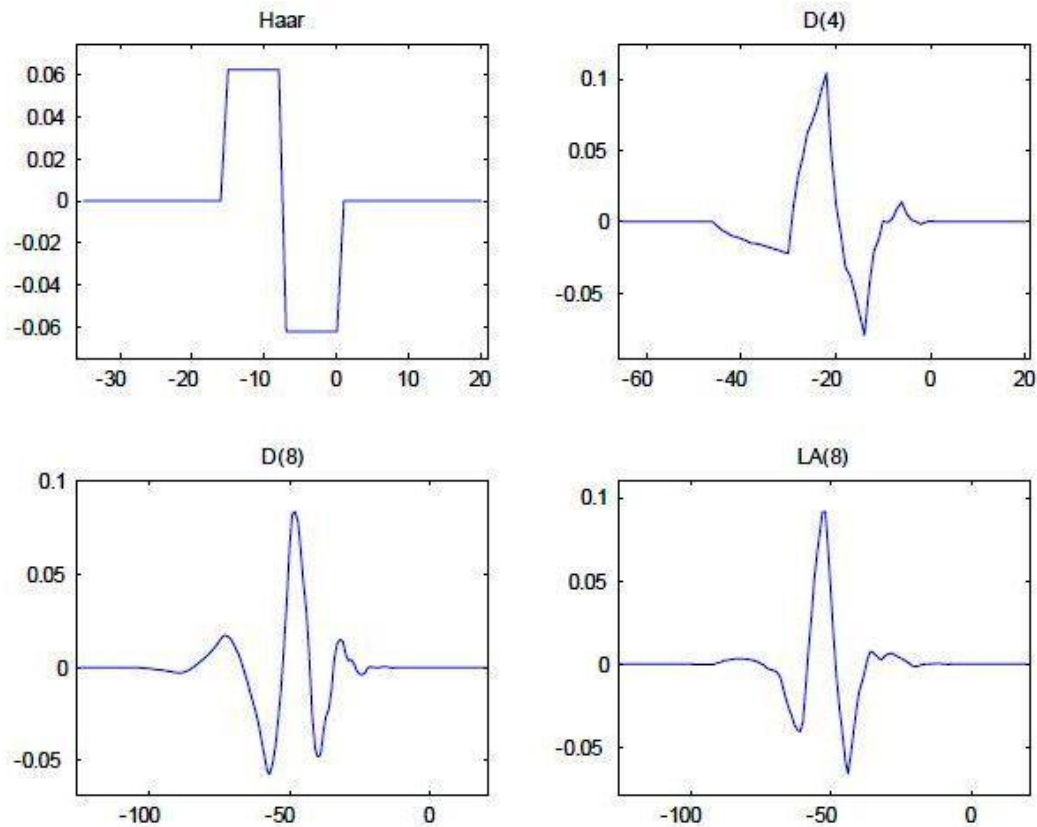
Esistono molti differenti filtri, ognuno adatto per specifici scopi di analisi, tali filtri differiscono nelle loro proprietà e nelle loro abilità di essere coerenti con le caratteristiche della serie in esame. Inoltre è necessario anche fissare la lunghezza del filtro, a causa delle condizioni al contorno, filtri lunghi sono adatti per lunghe serie. Il filtro più semplice è la wavelet Haar, che è sostanzialmente una differenza e media di ordine 2. In finanza ad esempio, molti hanno lavorato con le wavelet Daubechies (indicata come “D”) “Least Asymetrics” (“LA”) che sono filtri con lunghezza da 4 a 8. Alcuni criteri generali per la scelta del filtro possono essere:

- Simmetria : sono interessanti perché assicurano che non c’è sfasamento nella serie di output, purtroppo molte wavelet sono asimmetriche. Tale requisito comunque non ha molta importanza se si usa la MODWT.
- Ortogonalità: questa proprietà si riferisce al fatto che i coefficienti wavelet e di scaling contengono informazioni differenti. Questa è una caratteristica importante perché assicura che la decomposizione preserverà l’energia (cioè la varianza) della serie originale. Le Daubechies e Least-Asymmetric hanno questa caratteristica, i coefficienti di scaling e quelli wavelet sono ortogonali per costruzione.
- Smoothness: è misurato dal numero di derivate continue delle funzioni base. Haar è quella meno smooth. La scelta di questa caratteristica dipende anche dal grado di “smoothness” della serie originale. La Haar infatti è indicata per processi che presentano salti
- Numero di momenti nulli: questa caratteristica ha implicazione diretta nell’abilità che ha la wavelet di dare conto del comportamento del segnale. Ad esempio, se un segnale ha una struttura polinomiale o se può essere approssimato da un polinomio di ordine  $q$  allora la trasformata wavelet, per essere in grado di cogliere questa struttura, deve avere  $q$  momenti nulli. Daubechies ha un numero di momenti nulli pari alla metà della lunghezza del filtro. Haar e D(8) hanno rispettivamente 1 e 4 momenti nulli.



Esempi di filtri wavelet.

Figura 16 - Esempi di wavelet: Haar, Daubechies D(4), D(8); Least Asymmetric LA(8)



La figura riporta i coefficienti per le wavelet Haar, D(4), D(8), LA(8) per livello  $j=4$ . Si può osservare la struttura molto semplice della Haar. Se confrontiamo quest'ultima con D(4) e D(8) diventa chiaro più lungo è un filtro e più è smooth. LA(8) sembra meno asimmetrica della D(8) anche se è lontana dalla simmetria.

### 3.3.4.2. Applicazione al caso in esame

Per le considerazioni svolte nel paragrafo precedente, è stata adottata la cosiddetta: “Maximum Overlap Discrete Wavelet Transform” (MODWT) che come noto fornisce come risultato serie tutte con la stessa lunghezza. La trasformata Wavelet converte la serie originale in un set di serie costitutive. Queste serie presentano un comportamento migliore (varianza più stabile) di quella originale e quindi può essere prevista più accuratamente, la ragione di questo è l’effetto di filtro. Già in [7] e[8], per applicazioni simili sono stati utilizzati i filtri Daubechies di ordine 4, dopo diversi tentativi anche qui è stata confermata tale scelta. In questo caso si hanno i seguenti coefficienti di scaling:

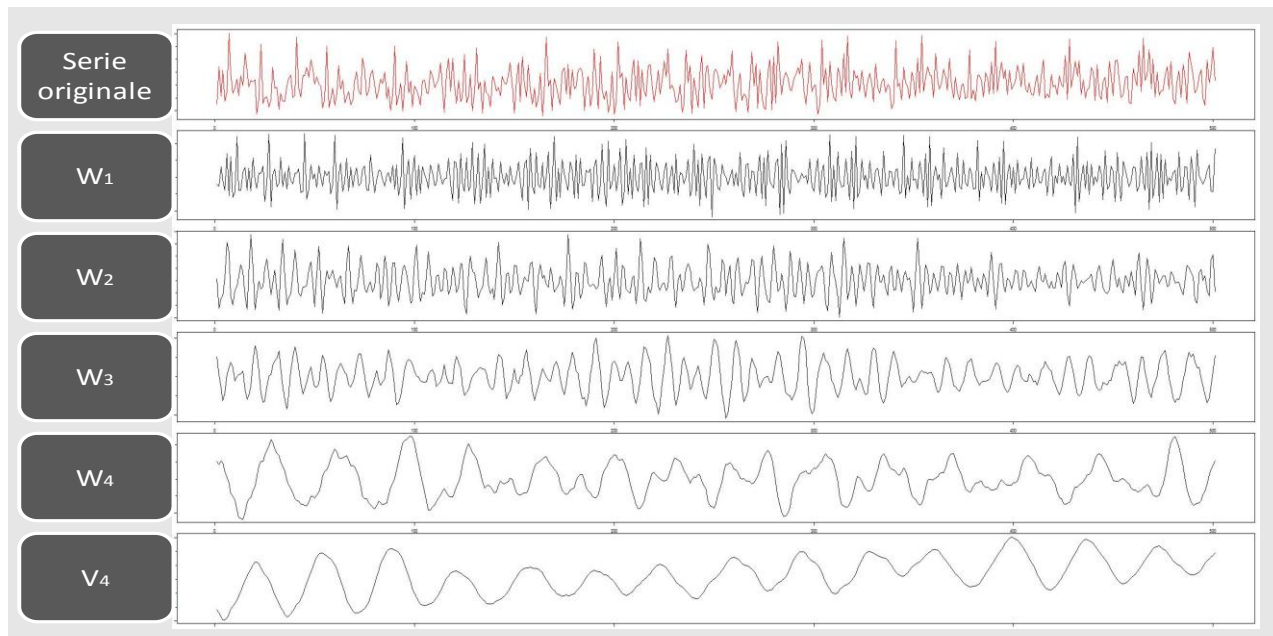
$$g_0 = \frac{1+\sqrt{3}}{4\sqrt{2}} ; g_1 = \frac{3+\sqrt{3}}{4\sqrt{2}} ; g_2 = \frac{3-\sqrt{3}}{4\sqrt{2}} ; g_3 = \frac{1-\sqrt{3}}{4\sqrt{2}}$$

Ed i coefficienti della funzione wavelet

$$h_0 = \frac{1-\sqrt{3}}{4\sqrt{2}} ; h_1 = \frac{-3+\sqrt{3}}{4\sqrt{2}} ; h_2 = \frac{3+\sqrt{3}}{4\sqrt{2}} ; h_3 = \frac{-1-\sqrt{3}}{4\sqrt{2}}$$

L’utilizzo dei coefficienti precedenti con le equazioni (4.1.16) (4.1.17) applicate al segnale in esame forniscono il grafico in figura

Figura 17 - Scomposizione Wavelet MODWT Daubechies ordine 4 (D4)



I grafici della figura precedente mostrano una tipica scomposizione in wavelet di ordine 4, l’ultima serie rappresenta quella più rilevante, la cosiddetta approssimazione che ha un andamento più regolare. Questa scomposizione ha rappresentato il punto di partenza per ciascuno dei metodi di previsione utilizzati successivamente. In ogni caso infatti si è applicato un modello di previsione all’ultima componente della scomposizione wavelet (la più rilevante in termini di contenuto informativo del segnale originale). Una volta effettuata la previsione si è ricostruita la serie originale (prevista) utilizzando la trasformazione inversa della MODWT applicata alle prime n-1 serie originali più quella prevista. Cioè:

$$x = S_j^p + D_j + \dots + D_1 \tag{4.2.24}$$

Con:

$$D_j = \tilde{A}_1 \dots \tilde{A}_{j-1} \tilde{B}_j \tilde{w}_j \text{ non modificata rispetto alla scomposizione precedente} \tag{4.2.25}$$

$$S_j^p = \tilde{A}_1 \dots \tilde{A}_{j-1} \tilde{A}_j \tilde{v}_j^p \text{ proveniente dal modello di previsione} \tag{4.2.26}$$

### 3.3.4.3. Cenni sui modelli ARIMA

Un modello di regressione ha, in generale, la forma [10]:

$$Y = b_0 + b_1X_1 + b_2X_2 + \dots + b_pX_p + e \quad (4.3.1)$$

Con  $X_1, \dots, X_p$  variabili esplicative,  $b_0, \dots, b_p$  coefficienti di regressione lineare ed  $e$  termine di errore. I termini  $X_i$  possono rappresentare qualsiasi fattore correlato con la grandezza da stimare  $Y$ . Supponiamo che le  $X_i$  siano definite come:  $X_1 = Y_{t-1}$ ,  $X_2 = Y_{t-2}$ ,  $X_p = Y_{t-p}$  l'equazione (4.3.1) diventa:

$$Y = b_0 + b_1Y_{t-1} + b_2Y_{t-2} + \dots + b_pY_{t-p} + e \quad (4.3.2)$$

Questa è ancora un'equazione di regressione con la differenza che le variabili esplicative sono diventate valori precedenti della variabile da stimare. Tale equazione quindi prende il nome di auto regressione di ordine  $p$  [AR( $p$ )]. Le differenze con un modello di regressione generale sono sostanzialmente due:

1. Nell'autoregressione l'assunzione di base dell'indipendenza dei termini di residui può essere facilmente violata, dal momento che le "variabili esplicative" a destra nell'equazione 4.3.2 hanno usualmente una relazione di dipendenza.
2. La determinazione dei passati valori di  $Y_t$  da includere nell'equazione non è necessariamente semplice

Proprio come è possibile una regressione rispetto a valori passati di una serie, esiste un modello che usa gli errori passati come variabili esplicative.

$$Y = \theta_0 + \theta_1e_{t-1} + \theta_2e_{t-2} + \dots + \theta_qe_{t-q} + e_t \quad (4.3.3)$$

Questa equazione è chiamata Media Mobile (Moving Average, MA). La media mobile è intesa sugli errori  $e_t$ . Le due equazioni possono essere combinate per dare luogo ad un modello di serie chiamato ARMA( $p, q$ ) valido però per serie stazionarie. Introducendo anche l'operazione di differenza nella serie è possibile dar luogo ai cosiddetti modelli ARIMA dove la lettera I sta per Integrazione. Questi modelli sono indicati sinteticamente come ARIMA( $p, d, q$ ), dove:

- AR:  $p$  = ordine della parte auto regressiva
- I:  $d$  = ordine della prima differenza coinvolta
- MA:  $q$  = ordine della parte a media mobile.

Ad esempio il rumore bianco può essere classificato come ARIMA(0,0,0) perché non c'è nessun aspetto AR ( $Y_t$  non dipende da  $Y_{t-1}$ ) non c'è nessuna differenza coinvolta e non c'è nessun termine MA ( $Y_t$  non dipende da  $e_{t-1}$ ). Mentre invece ad esempio un cammino casuale può essere classificato come ARIMA(0,1,0) è presente solo una differenza ( $Y_t - Y_{t-1} = e$ ).

Introducendo l'operatore "backward"  $B$  tale che:

$$BY_t = Y_{t-1} \quad (4.3.4)$$

In generale un modello ARMA di ordine  $p$  e  $q$ , ARMA( $p, q$ ) cioè ARIMA( $p, 0, q$ ) può essere scritto come:

$$Y_t = c + \Phi_1 Y_{t-1} + \dots + \Phi_p Y_{t-p} + e_t - \theta_1 e_{t-1} - \dots - \theta_q e_{t-q} \quad (4.3.5)$$

Oppure

$$(1 - \Phi_1 B - \dots - \Phi_p B^p) Y_t = c + (1 - \theta_1 B - \dots - \theta_q B^q) e_{t-q} \quad (4.3.6)$$

Ad esempio, un modello ARIMA(1,1,1) sarebbe

$$\underbrace{(1 - \Phi_1 B)}_{AR(1)} \underbrace{(1 - B)}_{\text{differenza 1}} Y_t = c + \underbrace{(1 - \theta_1 B)}_{MA(1)} e_t \quad (4.3.7)$$

Una volta analizzata la serie ed ipotizzato un modello, il problema è quindi la stima dei parametri questa può essere ottenuta attraverso il metodo dei minimi quadrati oppure, in modo iterativo, con il metodo di massima verosimiglianza.

Nel caso in esame, per ogni WTG sono stati costruiti blocchi congruenti di serie temporale di velocità del vento (cfr. par 3.2). Uno di questi blocchi è stato scelto per una prima stima del modello ARIMA più adatto, salvando i parametri su un vettore (un elemento per WTG). In realtà, come era da attendersi, non si sono riscontrate molte differenze, il modello più frequente era un ARIMA(2,1,2). Per la successiva verifica del modello, è stato utilizzato un blocco diverso da quello impiegato nella stima. Il processo (come illustrato in Fig. 14) consiste nello stimare il modello ad ogni iterazione, simulando quindi la lettura di nuovi dati reali ogni ora, effettuando poi la previsione ai tre lag temporali.

#### 3.3.4.4. Cenni sul filtro di Kalman

Questo metodo è stato introdotto già negli anni 50 nel contesto relativo ai sistemi di controllo dove si ha la necessità di descrivere lo stato di un sistema ed eventualmente la sua evoluzione, effettuando la misurazione diretta di alcune variabili ma facendo inferenza su altre. Tali variabili sono dette “di stato” e sono affette da rumore [12]. Il filtro di Kalman quindi, cerca una soluzione al problema generale della stima dello stato  $\mathbf{x}_k$  di un processo a tempo discreto governato dall’equazione stocastica alle differenze (Modello dinamico lineare)

$$\begin{aligned} \mathbf{x}_k &= \mathbf{A}\mathbf{x}_{k-1} + \mathbf{B}\mathbf{u}_{k-1} + \mathbf{w}_{k-1} && \text{(equazione di stato)} \\ \mathbf{z}_k &= \mathbf{H}\mathbf{x}_k + \mathbf{v}_k && \text{(equazione di misura)} \end{aligned}$$

Lo stato  $\mathbf{x}_k$  del sistema al tempo  $k$  dipende dallo stato al tempo  $k-1$  ma in realtà, non è direttamente determinabile, lo è solo una sua osservazione (misura)  $\mathbf{z}_k$  nota a meno di un termine di rumore.

Le variabili aleatorie  $\mathbf{w}_k$  e  $\mathbf{v}_k$  rappresentano rispettivamente il rumore del processo e della misura. Esse sono assunte essere indipendenti (tra loro) e normalmente distribuite:  $\mathbf{w} \sim \mathcal{N}(0, \mathbf{Q})$  e  $\mathbf{v} \sim \mathcal{N}(0, \mathbf{R})$

$\mathbf{Q}$  è la matrice di covarianza del rumore del processo,  $\mathbf{R}$  è la matrice di covarianza del rumore di misura. La matrice  $n \times n$   $\mathbf{A}$  nell’equazione alle differenze, lega lo stato al tempo precedente  $k-1$  allo stato corrente al tempo  $k$  in assenza di altro rumore. La matrice  $l \times n$   $\mathbf{B}$  lega l’input di controllo allo stato  $\mathbf{x}$ . La matrice  $m \times n$   $\mathbf{H}$  lega la variabile di stato alla sua misura.

#### *Le origini computazionali del filtro di Kalman*

Il filtro di Kalman viene anche definito metodo *previsione-correzione*. Nella fase di previsione, viene prodotta una stima dello stato corrente del sistema e della sua incertezza. Quando è disponibile la misurazione dello stato corrente effettivo, affetta da rumore, la stima precedente viene corretta tenendo conto dell’incertezza.

Definiamo  $\mathbf{x}_k^-$  la stima a priori dello stato al tempo  $k$ , data la conoscenza del processo prima del passo  $k$  e  $\mathbf{x}_k^+$  la stima a posteriori dello stesso stato dopo la misura  $\mathbf{z}_k$

Possiamo definire una stima degli errori “a priori” e “a posteriori” come

$$\begin{aligned} \mathbf{e}_k^- &= \mathbf{x}_k - \mathbf{x}_k^- \\ \mathbf{e}_k^+ &= \mathbf{x}_k - \mathbf{x}_k^+ \end{aligned}$$

quindi le stime delle covarianze a priori e posteriori sono:

$$\mathbf{P}_k^- = \text{cov}(\mathbf{x}_k - \mathbf{x}_k^-) \text{ e } \mathbf{P}_k^+ = \text{cov}(\mathbf{x}_k - \mathbf{x}_k^+) \quad (4.4.1)$$

Per derivare le equazioni del filtro di Kalman, si cerca una equazione che calcoli una stima dello stato a posteriori come combinazione di una stima a priori ed una differenza pesata fra le misure attuali e la previsione di queste misure.

$$\mathbf{x}_k^+ = \mathbf{x}_k^- + \mathbf{K}(z_k - \mathbf{H}\mathbf{x}_k^-) \quad (4.4.2)$$

il termine  $(z_k - \mathbf{H}\mathbf{x}_k^-)$  è chiamato misura dell'innovazione o termine residuo. Il residuo riflette la discrepanza tra le misure previste e quelle effettive. Un residuo nullo significherebbe completo accordo tra le due grandezze. La matrice  $n \times m$   $\mathbf{K}$  è scelta come fattore di "guadagno" o di "combinazione" che minimizza l'errore dello stato calcolato a posteriori sullo stato effettivo  $(\mathbf{x}_k - \mathbf{x}_k^+)$  cioè del valore atteso della sua norma al quadrato:

$$\mathbf{E}(\|\mathbf{x}_k - \mathbf{x}_k^+\|^2)$$

Minimizzare questa quantità corrisponde a minimizzare la traccia della matrice di covarianza:

$$\mathbf{P}_k^+ = \text{cov}(\mathbf{x}_k - \mathbf{x}_k^+)$$

Questa minimizzazione può essere calcolata sostituendo la (4.4.2) nella equazione precedente e sviluppando il calcolo della covarianza tenendo conto che vale:

$$\text{cov}(A\mathbf{x} + \mathbf{b}) = A \text{cov}(\mathbf{x}) A^T + \text{cov}(\mathbf{b})$$

Infine derivando rispetto a  $\mathbf{K}$  e uguagliando a zero.

Un risultato per  $\mathbf{K}$  può essere espresso come:

$$\mathbf{K}_k = \mathbf{P}_k^- \mathbf{H}^T (\mathbf{H} \mathbf{P}_k^- \mathbf{H}^T + \mathbf{R})^{-1} \quad (4.4.3)$$

Guardando alla (4.4.3) possiamo osservare che:

$$\lim_{P_k^- \rightarrow 0} K_k = 0 \rightarrow x_k^+ = x_k^-$$

quindi, se la stima a priori della covarianza dell'errore va a zero,  $\mathbf{K}$  è tale che i residui non contano quindi la stima a posteriori corrisponde a quella a priori, sostanzialmente la misura non ha impatto.

Se invece  $\mathbf{R}$  va a zero:

$$\lim_{R_k \rightarrow 0} K_k = \mathbf{H}^{-1}$$

non c'è incertezza sulla misura il filtro si affida su questa piuttosto che sulla stima a priori

Un altro modo di pensare a  $\mathbf{K}$  è che quando  $\mathbf{R}$  tende a zero, la misura reale di  $z_k$  è sempre più "affidabile" mentre la misura prevista  $\mathbf{H}\mathbf{x}_k^-$  lo è sempre meno. Inoltre se la covarianza dell'errore a priori  $\mathbf{P}_k^-$  tende a zero la misura reale di  $z_k$  è sempre meno affidabile e la misura prevista  $\mathbf{H}\mathbf{x}_k^-$  lo è sempre di più.

#### *Algoritmo per il filtro di Kalman discreto*

Il filtro di Kalman stima un processo usando una forma di feedback: il filtro stima lo stato del processo ad un certo tempo e successivamente ottiene dei feedback in termini di misure (affette da rumore). Per questo le equazioni ricadono in due categorie: equazione di aggiornamento temporale ed equazione di aggiornamento della misura. Le prime sono responsabili nel proiettare in avanti (nel tempo) lo stato corrente e la stima della covarianza dell'errore per ottenere la stima a priori del successivo passo temporale. La seconda equazione è responsabile del feedback, cioè nell'incorporare una nuova misura nella stima a priori per ottenere una stima a posteriori.

Le equazioni di aggiornamento temporale possono essere anche pensate come *equazioni di previsione* ("predictor") mentre le equazioni di aggiornamento della misura come *equazioni di correzione*. Infatti l'algoritmo finale è detto di "*previsione-correzione*" (predictor-corrector).

---

Le equazioni di aggiornamento temporale (Discrete Kalman Filter time update equations)

$$\mathbf{x}_k^- = \mathbf{A}\mathbf{x}_{k-1}^- + \mathbf{B}\mathbf{u}_{k-1} \quad (4.4.4)$$

$$\mathbf{P}_k^- = \mathbf{A}\mathbf{P}_{k-1}\mathbf{A}^T + \mathbf{Q} \quad (4.4.5)$$

Si vede come l'equazione proietta in avanti di un passo (da k-1 a k), lo stato e la stima della covarianza. A e B vengono dalla equazione di stato mentre la Q dalla equazione della misura.

Equazioni di aggiornamento delle misure

$$\mathbf{K}_k = \mathbf{P}_k^- \mathbf{H}^T (\mathbf{H}\mathbf{P}_k^- \mathbf{H}^T + \mathbf{R})^{-1} \quad (4.4.6)$$

$$\mathbf{x}_k = \mathbf{x}_k^- + \mathbf{K}_k (\mathbf{z}_k - \mathbf{H}\mathbf{x}_k^-) \quad (4.4.7)$$

$$\mathbf{P}_k = (\mathbf{I} - \mathbf{K}_k \mathbf{H}) \mathbf{P}_k^- \quad (4.4.8)$$

Il primo passo del processo di aggiornamento della misura consiste nel calcolare il guadagno  $\mathbf{K}_k$ . Da notare che la (4.4.6) è la stessa equazione (4.4.3). Il passo successivo è ottenere la misura  $\mathbf{z}_k$  per generare una stima a posteriori incorporando le misure (anche qui la 4.4.7 è semplicemente la 4.4.2). Il passo finale è la stima della covarianza dell'errore a posteriori utilizzando la 4.4.8.

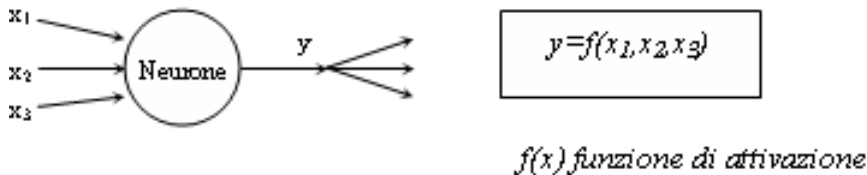
Dopo ogni coppia di aggiornamento temporale ed aggiornamento della misura, il processo viene ripetuto con la precedente stima "a posteriori" usata per proiettare in avanti una nuova stima a priori. Questa natura ricorsiva è il vero vantaggio del filtro di Kalman.

Nel caso in esame, per ogni WTG sono stati costruiti blocchi congruenti di serie temporale di velocità del vento (cfr. par 3.2). Selezionato il blocco più numeroso, questo è stato utilizzato per il test. Il processo (come illustrato in Fig. 14) consiste nello stimare il modello ad ogni iterazione, simulando quindi la lettura di nuovi dati reali ogni ora, effettuando poi la previsione ai tre lag temporali. A questo scopo è stata utilizzata la routine "StructTS" di R che stima il modello state space con filtro di Kalman partendo direttamente dalla serie storica fornita.

### 3.3.4.5. Cenni sulle reti neurali

Una rete neurale artificiale è costituita da un insieme d'unità elementari d'elaborazione interconnesse [13]. L'unità fondamentale delle reti neurali, il neurone artificiale, riceve informazioni numeriche attraverso un certo numero di nodi d'ingresso e fornisce una risposta. La risposta del neurone artificiale è definita da una funzione scalare, detta "funzione di attivazione", che ha come argomento una combinazione lineare degli ingressi del neurone.

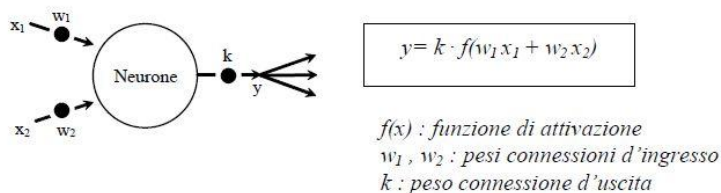
Figura 18 - Elemento di una rete neurale artificiale



Possono essere utilizzate diverse funzioni di attivazione anche in funzione dell'output desiderato, comunque in genere di valore assoluto minore di 1. A seconda se sono necessari valori negativi oppure no, ad esempio, può essere utilizzata una funzione sigmoideale (tra 0 e 1) oppure una combinazione di tangenti iperboliche (tra -1 e 1). Nel nostro caso, essendo possibili valori negativi si è utilizzata la tangente iperbolica.

Le connessioni della rete neurale collegano tra loro i neuroni e sono caratterizzate da un peso. In pratica una connessione non esegue il semplice collegamento tra due unità, ma moltiplica anche il dato che la attraversa per il proprio peso. Possiamo rappresentare una generica relazione ingresso/uscita del neurone come nella figura seguente.

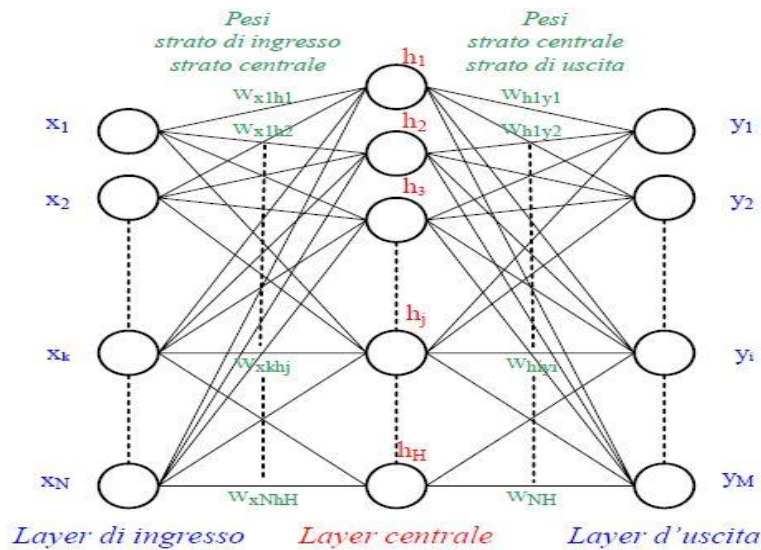
Figura 19 - Pesì di un collegamento neuronale



Ogni neurone ha generalmente in ingresso anche uno scalare detto soglia (o bias), che può essere visto come il peso di un ingresso unitario. L'architettura della rete è il modo in cui i neuroni sono organizzati. L'architettura più utilizzata è la "MLP feed forward". MLP sta per "MultiLayer Perceptron" e indica che i neuroni sono organizzati in strati, i neuroni di uno strato possono condividere gli ingressi ma non sono connessi tra loro.

Nella figura seguente è mostrata una rete a più strati. Il primo strato è costituito dai nodi d'ingresso, attraverso i quali vengono passati alla rete i dati da elaborare. Il primo strato è collegato al successivo attraverso una serie di connessioni "pesate". Lo strato centrale è detto nascosto (hidden layer). Dallo strato nascosto le informazioni elaborate sono trasferite alle uscite attraverso le connessioni pesate d'uscita. Gli strati nascosti possono essere più di uno

Figura 20 - Rete MLP feed forward a 2 strati



Una rete che non presenta connessioni retroattive, vale a dire connessioni che portano dati a neuroni appartenenti a strati precedenti, sono dette del tipo feedforward. Nelle reti feedforward le informazioni viaggiano in un'unica direzione, dallo strato d'ingresso a quello d'uscita. La rete rappresentata in figura è detta a 2 strati in quanto si considerano gli strati interessati da connessioni pesate in ingresso, in questo caso quello centrale e quello d'uscita. La determinazione dei pesi avviene in quella che viene generalmente definito "addestramento" della rete neurale.

Comunemente le reti neurali sono addestrate in modo tale che ad una particolare configurazione di ingressi corrisponda una specifica configurazione di uscite. In questi casi per addestrare la rete vengono utilizzate molte coppie di vettori di ingressi e di uscite. Questo tipo di addestramento è detto "supervisionato". L'algoritmo di addestramento è usualmente del tipo "backpropagation", di cui ci sono però molte varianti. Il termine "backpropagation" indica che i pesi vengono aggiornati durante l'addestramento nella direzione in cui l'errore della rete neurale diminuisce più rapidamente: il negativo del gradiente.

Dato un insieme di dati d'addestramento del tipo:

$$\begin{bmatrix} x_1^1 \\ x_2^1 \\ \vdots \\ x_k^1 \\ \vdots \\ x_N^1 \end{bmatrix} \rightarrow \begin{bmatrix} T_1^1 \\ T_2^1 \\ \vdots \\ T_i^1 \\ \vdots \\ T_M^1 \end{bmatrix} \Rightarrow \begin{bmatrix} x_1^s \\ x_2^s \\ \vdots \\ x_k^s \\ \vdots \\ x_N^s \end{bmatrix} \rightarrow \begin{bmatrix} T_1^s \\ T_2^s \\ \vdots \\ T_i^s \\ \vdots \\ T_M^s \end{bmatrix} \Rightarrow \begin{bmatrix} x_1^{Ns} \\ x_2^{Ns} \\ \vdots \\ x_k^{Ns} \\ \vdots \\ x_N^{Ns} \end{bmatrix} \rightarrow \begin{bmatrix} T_1^{Ns} \\ T_2^{Ns} \\ \vdots \\ T_i^{Ns} \\ \vdots \\ T_M^{Ns} \end{bmatrix}$$

dove il vettore  $(x_1^*, \dots, x_n^*)$  è l'ingresso ed il vettore  $(T_1^*, \dots, T_n^*)$  è l'uscita voluta, l'errore complessivo è definito come lo scarto quadratico medio:

$$E = \frac{1}{2} \sum_{s=1}^{Ns} \left[ \sum_{t=1}^M (y_t^s - T_t^s)^2 \right]$$



Indicando con  $W_{kj}$  il peso della connessione tra il k-esimo ingresso ed il j-esimo neurone nascosto e con  $W_{ji}$  il peso della connessione tra il j-esimo neurone nascosto e l'i-esimo neurone d'uscita, l'errore complessivo può essere considerato come funzione di tali pesi:

$$E = E(W_{kj}, W_{ji})$$

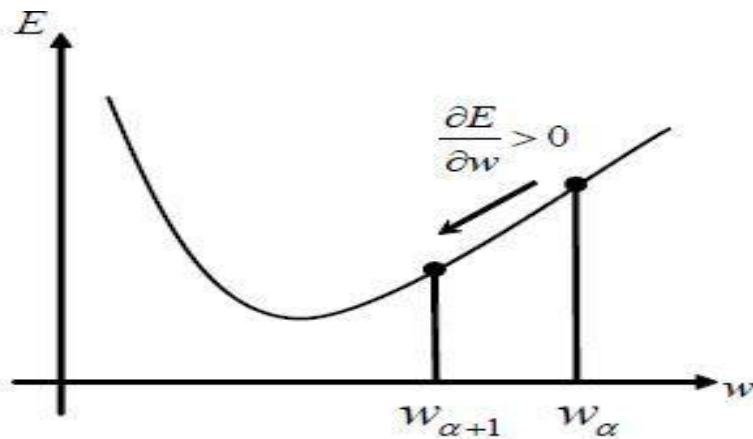
La procedura iterativa prevede il calcolo del gradiente dell'errore rispetto ai pesi della rete ad ogni iterazione, in base ad esso aggiorna il valore dei pesi stessi:

$$(W_{kj})_{\alpha+1} = (W_{kj})_{\alpha} - \left( \eta \frac{\partial E}{\partial W_{kj}} \right)_{\alpha}$$

$$(W_{ji})_{\alpha+1} = (W_{ji})_{\alpha} - \left( \eta \frac{\partial E}{\partial W_{ji}} \right)_{\alpha}$$

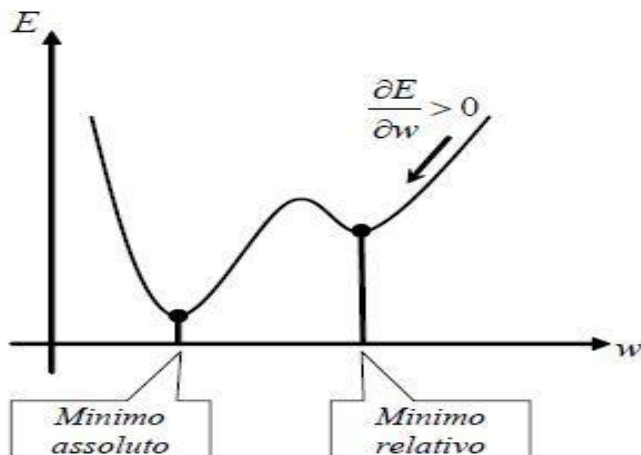
dove  $\eta$  è un parametro che determina la velocità di variazione dei pesi ad ogni iterazione. Per avere elevate velocità di convergenza dell'algoritmo è opportuno che  $\eta$  vari durante la fase d'addestramento. In particolare occorre che  $\eta$  sia elevato nei primi cicli d'addestramento (alta velocità di convergenza) e sia piccolo nelle fasi successive (migliore precisione). L'algoritmo modifica i pesi in modo da "discendere" lungo il gradiente verso il minimo della funzione errore. Il concetto è illustrato nella figura seguente per il caso di un solo peso.

Figura 21 - Ricerca del minimo attraverso la discesa gradiente



Il valore iniziale dei pesi della rete determina il punto di partenza nella ricerca del minimo della funzione. Alla prima iterazione il valore dei pesi può essere zero o può essere scelto in modo casuale. L'algoritmo utilizzato nei modelli proposti imposta il valore dei pesi inizialmente in modo casuale. L'impostazione iniziale dei pesi può influenzare il risultato dell'addestramento. Si può infatti incorrere in un minimo relativo della funzione errore globale, determinando l'interruzione del processo di addestramento, pur non avendo raggiunto il minimo assoluto (vedi figura seguente).

Figura 22 - Ricerca del minimo attraverso il gradiente: minimo relativo



La procedura detta “error back propagation with momentum” evita che l’addestramento si fermi in corrispondenza di un minimo locale del gradiente, imponendo ai pesi cambiamenti uguali alla somma di una frazione dell’ultimo cambiamento e del nuovo cambiamento suggerito dalla regola del gradiente. Il criterio che termina l’addestramento è il raggiungimento da parte dell’errore di un valore prefissato, oppure il raggiungimento di un prefissato numero d’iterazioni d’addestramento. Per evitare l’“overfitting”, cioè l’approssimazione troppo spinta dell’insieme di dati d’addestramento a discapito della generalità della rete, si può effettuare una verifica dell’errore rispetto a dati diversi da quelli utilizzati per l’addestramento, non utilizzati quindi dall’algoritmo d’aggiornamento dei pesi. Dal momento che la funzione di attivazione fornisce generalmente un risultato inferiore a 1 (in valore assoluto) è spesso necessario scalare i dati affinché tutti si riferiscano ad un unico intervallo. La parte dell’algoritmo più pesante in termini di costo computazionale è il calcolo delle derivate parziali del gradiente.

Dopo l’addestramento, la rete è pronta per l’implementazione del modello che essa rappresenta. Questa fase successiva è detta di produzione o di simulazione e prevede il passaggio degli ingressi alla rete ed il calcolo attraverso di essa delle uscite. Le informazioni sono trasmesse e filtrate dalle connessioni tra gli strati e ogni neurone concorre al calcolo attraverso la propria funzione d’attivazione. Come avviene anche per altri modelli di previsione, il test vero e proprio viene effettuato utilizzando una parte della serie storica non coinvolta nell’addestramento.

Le caratteristiche di questo modello sono tali per cui non è necessario effettuare la stima utilizzando un set di dati congruente al set che viene utilizzato per la validazione. Quindi, nel caso in esame, sono stati costruiti due insieme di dati, uno per l’addestramento per ogni WTG ed uno per la validazione.

Il set utilizzato per la validazione era costruito, come i precedenti, da blocchi congruenti di serie temporale di velocità del vento (cfr. par 3.2)

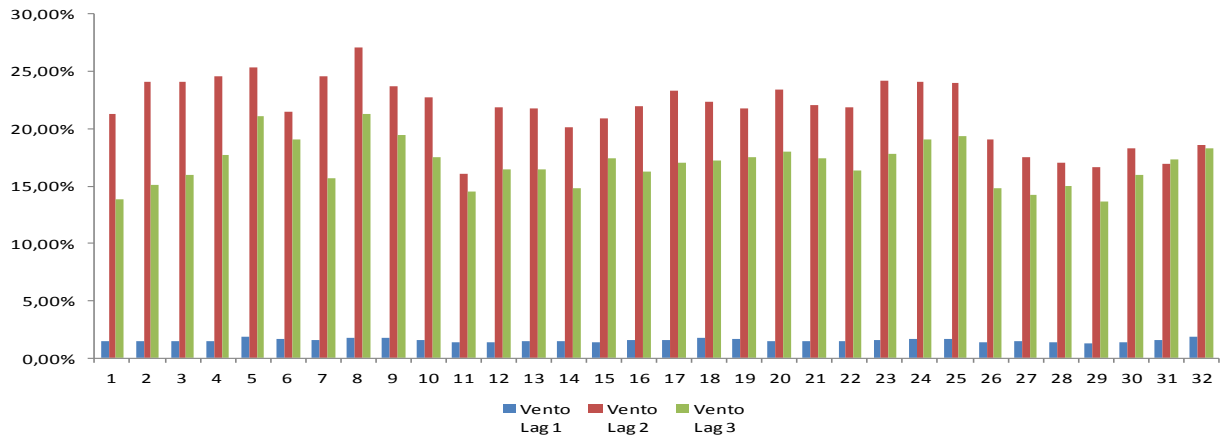
La fase di addestramento, inoltre non viene ripetuta ad ogni ciclo ma effettuata una volta per tutte (nel caso di utilizzo regolare ovviamente sarebbe necessaria un riaddestramento periodico).

E’ stata utilizzata una serie ai 10 minuti di velocità del vento, la rete neurale che ha dato i migliori risultati è stata una rete Feed-Forward con 12 elementi di input (2 ore precedenti) 8 elementi nello strato nascosto ed un elemento in uscita. L’output viene poi immesso nello strato di input per prevedere gli intervalli successivi: 6,12,18 e successivamente mediati quindi 1,2,3 ore in avanti.

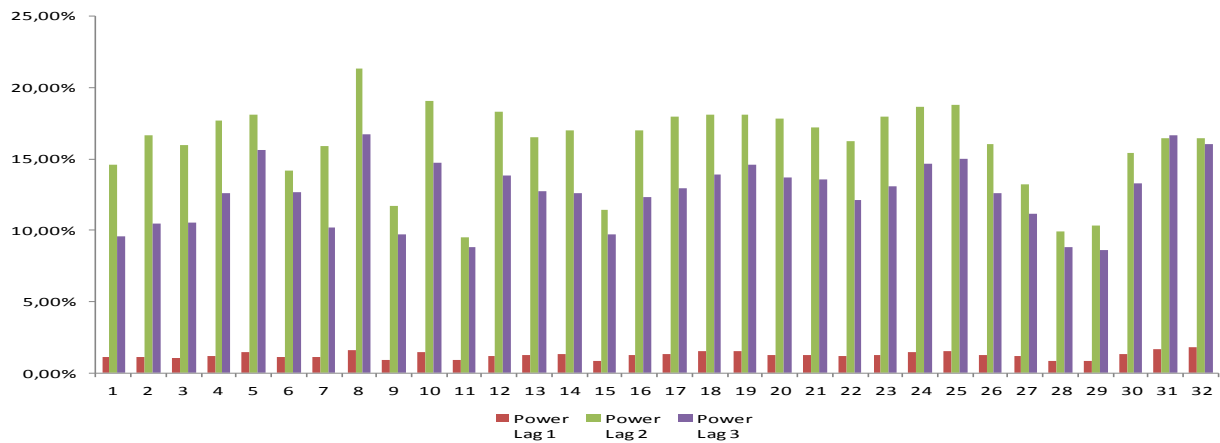
### 3.3.4.6. Scomposizione wavelet e modello arima

In questo caso è stata effettuata la scomposizione della serie utilizzando le wavelet Daubochies di D4 di ordine 2. E' stato scelto un ordine più basso perché altrimenti non si riusciva ad ottenere la convergenza nella stima del modello.

**Figura 23 - Errore sulla velocità del vento per WTG (Wavelet + Arima)**



**Figura 24 - Errore sull'energia prodotta per WTG (Wavelet + Arima)**



**Tabella 1 - Riepilogo errori (Wavelet + Arima)**

	NMAE			MAPE		
	Power Lag 1	Power Lag 2	Power Lag 3	Vento Lag 1	Vento Lag 2	Vento Lag 3
Errore medio	1,27%	16,06%	12,63%	1,56%	21,65%	16,93%

Figura 27 - Wavelet + Arima, velocità del vento prevista vs reale LAG 1 (WTG 1-16)

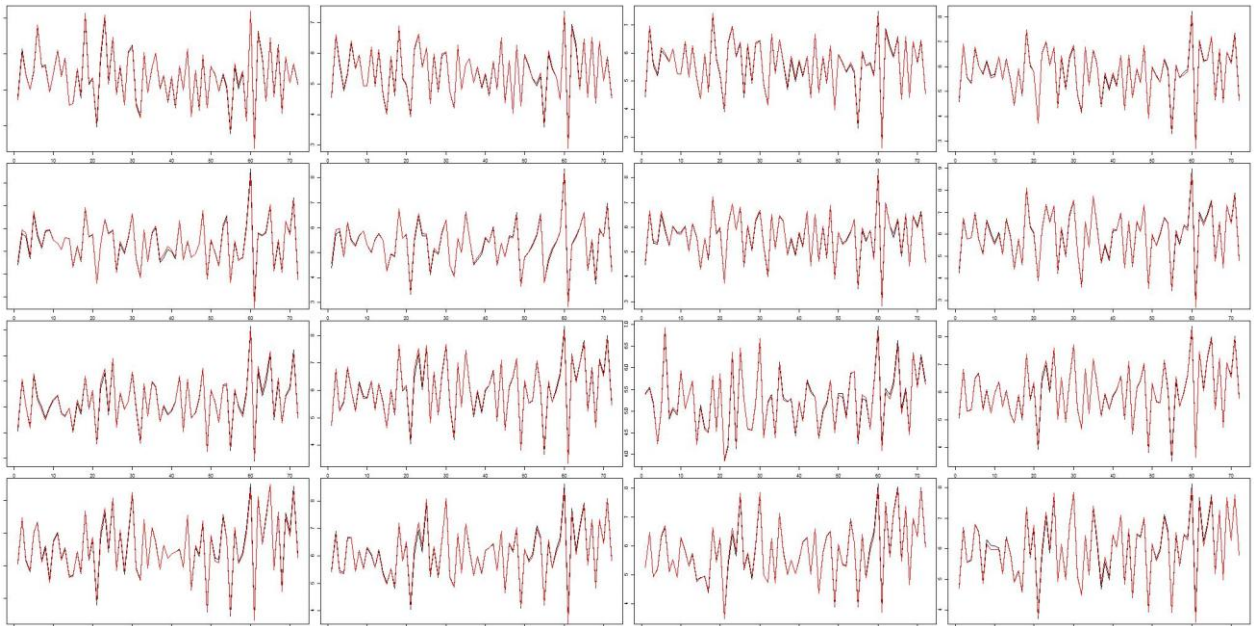


Figura 26 - Wavelet + Arima, energia prevista vs energia reale LAG 1 (WTG 1-16)

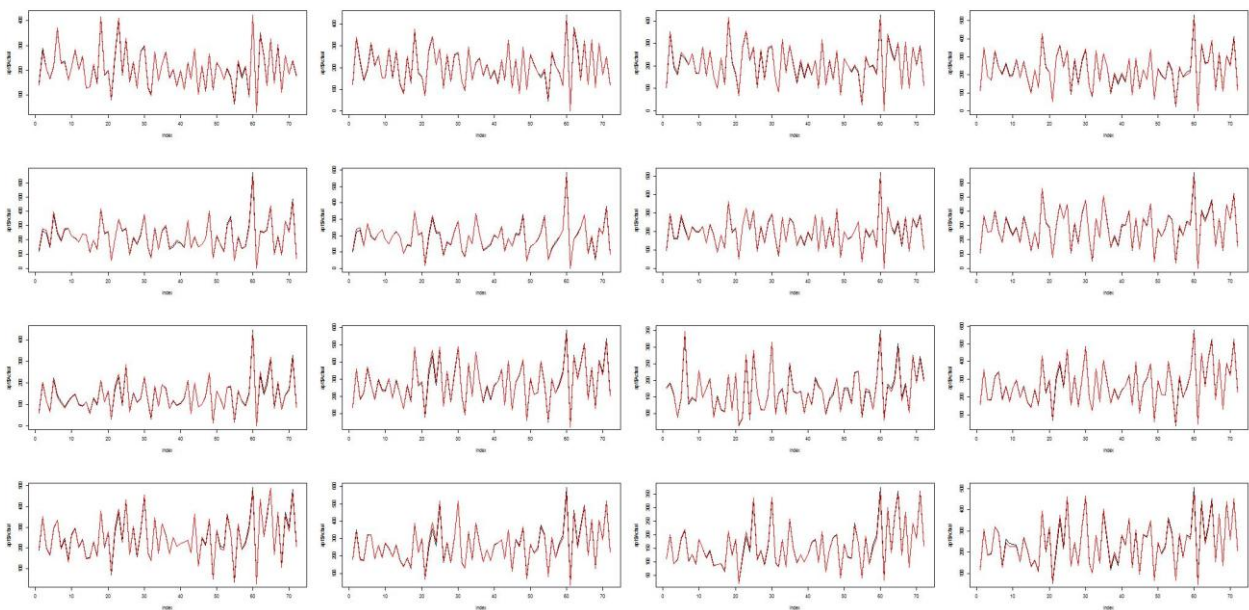
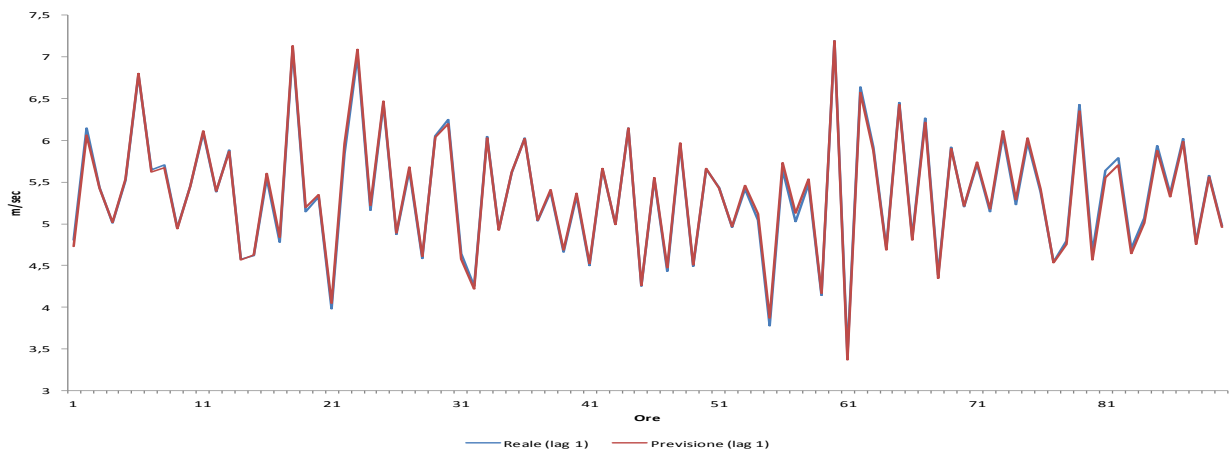
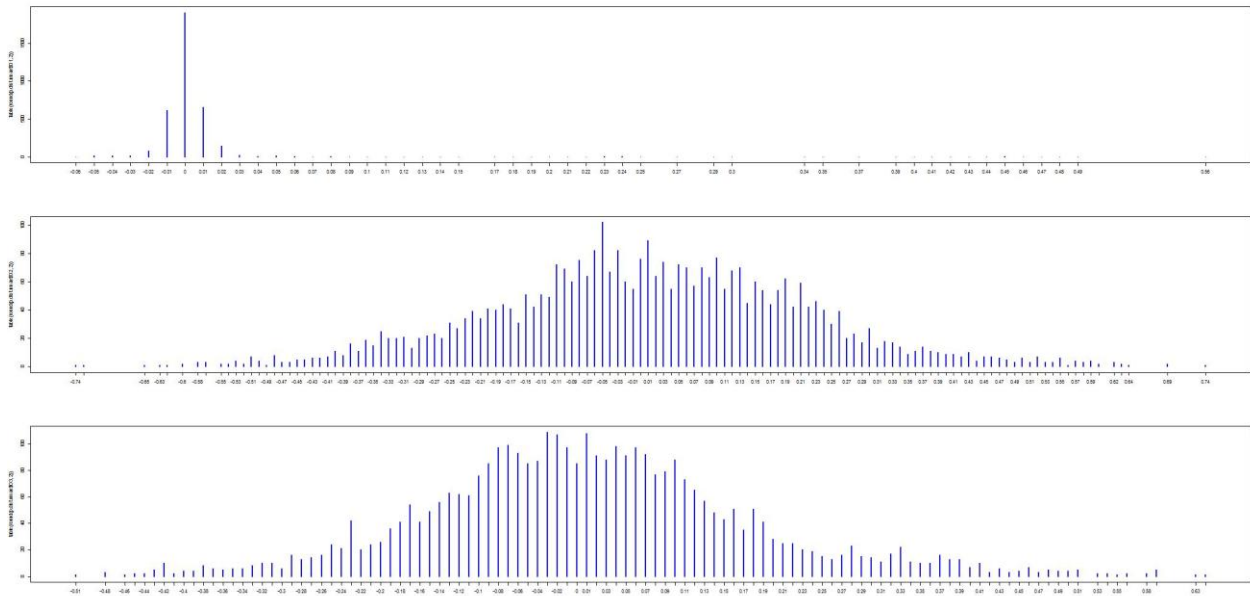


Figura 25 - Esempio di confronto curve velocità del vento reale e prevista (WTG 1- lag 1)



Nel grafico della figura precedente è riportato, a titolo di esempio, un campione della curva prevista a lag 1 e la curva effettiva per una singola turbina. Si osserva una quasi completa sovrapposizione.

Figura 28 - Distribuzione errori su tutte le WTG per lag temporale di previsione (Wavelet + Arima)



I risultati mostrano un buon comportamento del modello per l'orizzonte temporale più breve (1 ora) sia l'errore sul vento sia quello sull'energia sono molto bassi mentre sui successivi orizzonti temporali (2 e 3 ore) i risultati non sono molto buoni seppur in assoluto nella media per questo tipo di previsione. L'andamento orario segue con buon accordo quello reale, la distribuzione dell'errore è simmetrica rispetto allo 0 ed abbastanza concentrata su di esso, con una leggera predominanza di valori positivi che indicano una sovrastima del modello.

### Risultati per impianto

Figura 29 - Distribuzione degli errori per impianto (Wavelet – Arima)

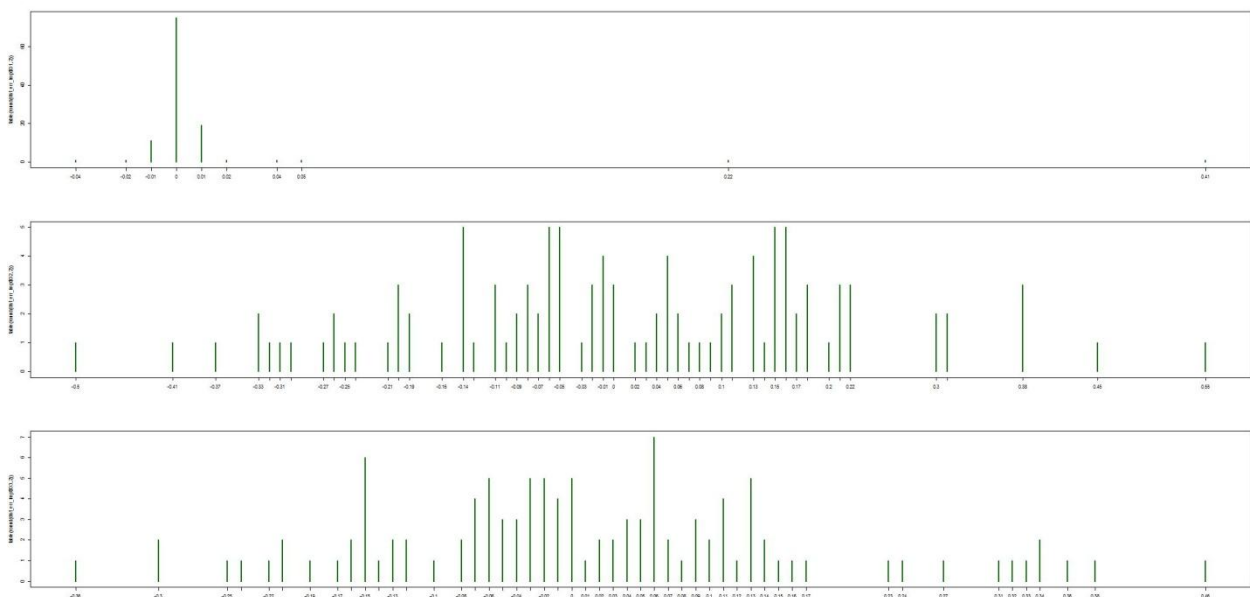


Figura 30 - Curve orarie previste e reale di impianto vs NMAE Lag temporale 1 (Wavelet + Arima)

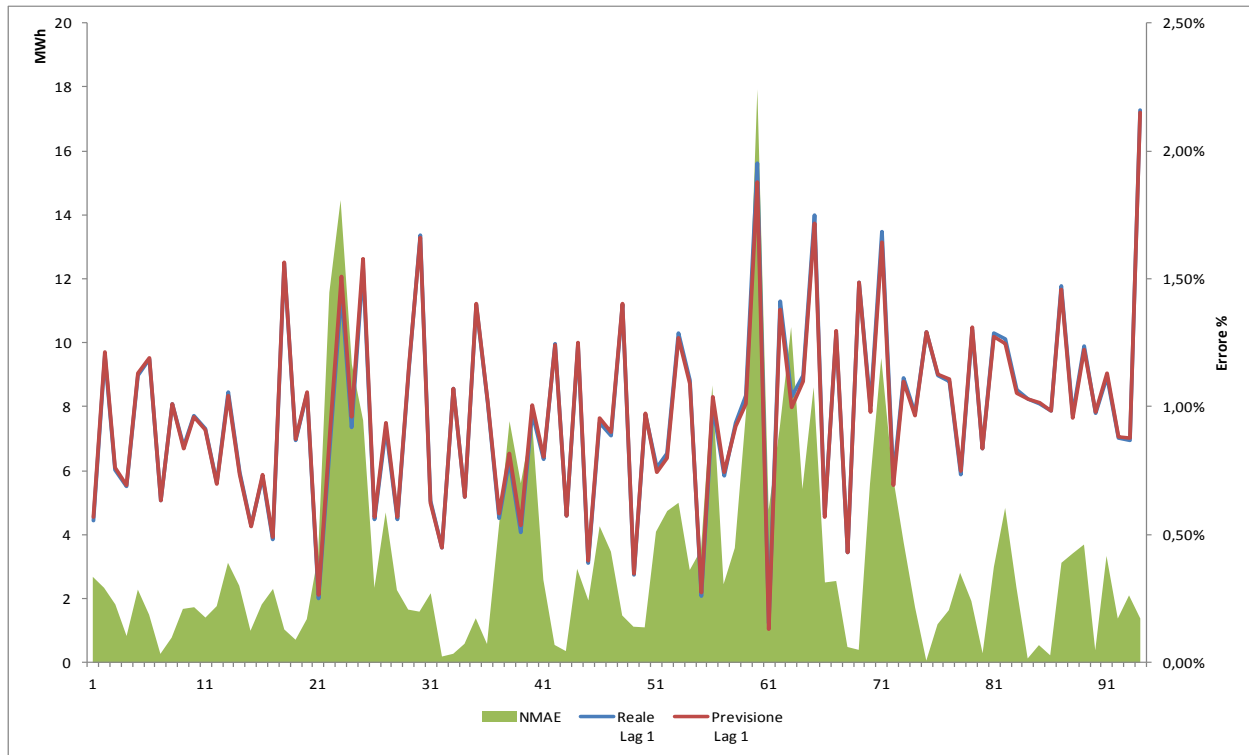
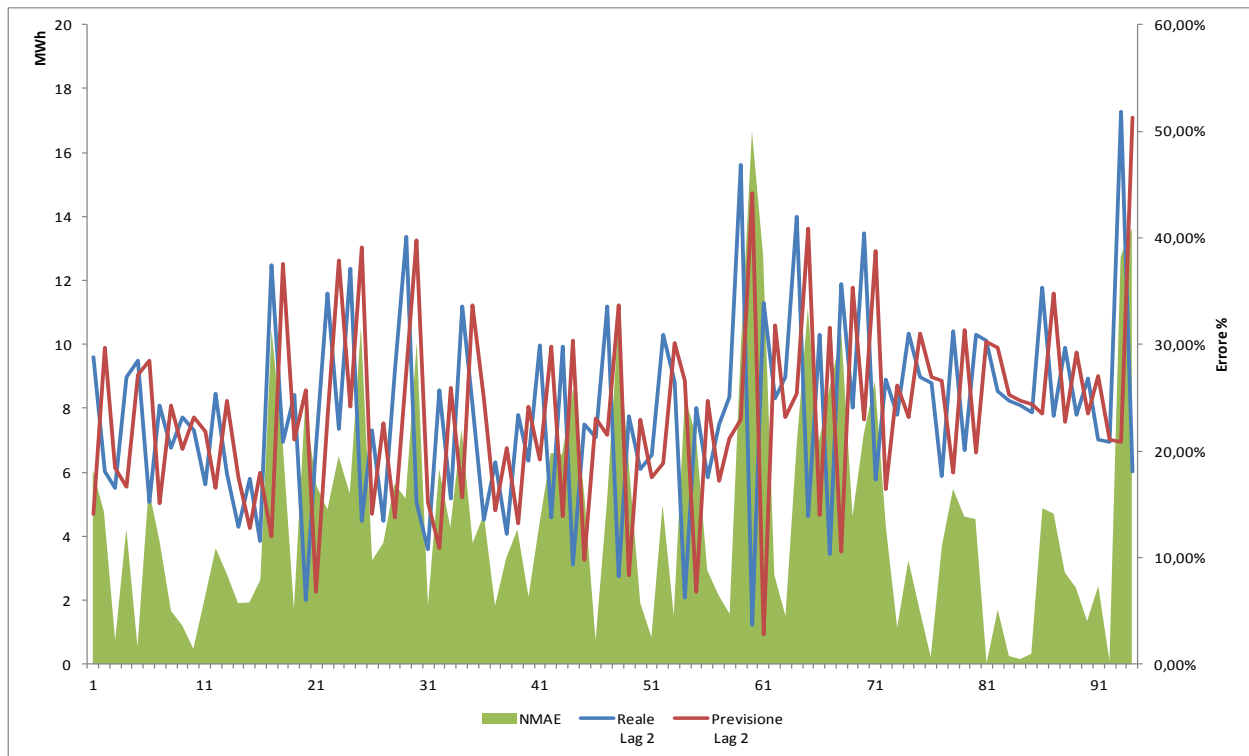
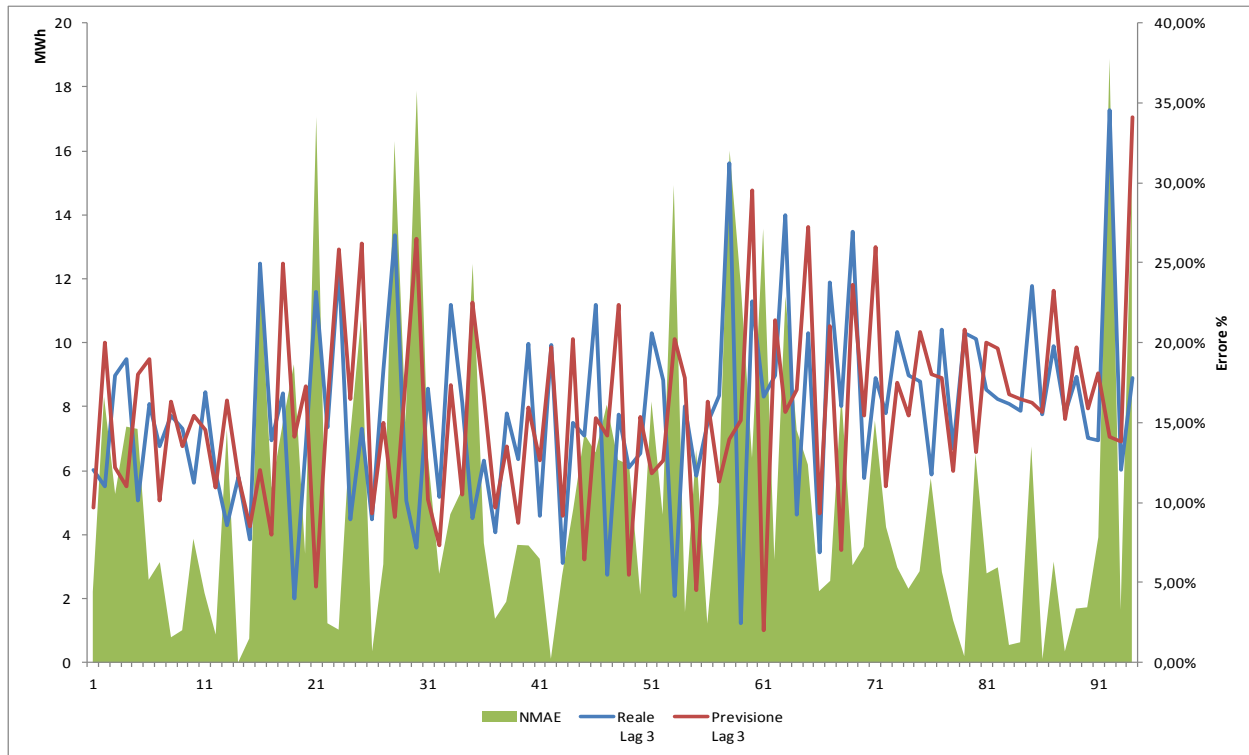


Figura 31 - Curve orarie previste e reale di impianto vs NMAE Lag temporale 2 (Wavelet + Arima)



**Figura 32 - Curve orarie previste e reale di impianto vs NMAE Lag temporale 3 (Wavelet + Arima)**



**Tabella 2 - Riepilogo errori per impianto (Wavelet + Arima)**

NMAE Lag 1	NMAE Lag 2	NMAE Lag 3
1,28%	15,04%	10,48%

I risultati considerati al livello di impianto confermano il comportamento relativo alla singola WTG, nell'orizzonte temporale 1 l'accordo con i valori reali è quasi perfetto, viene seguito bene l'andamento anche se in qualche caso non sono colti i picchi positivi o negativi. Negli orizzonti temporali successivi si nota come, in realtà, il modello effettui sempre la stessa previsione del Lag 1 che ovviamente, in questo caso, comporta un errore maggiore. Il riepilogo generale dell'errore mostra dei valori leggermente più bassi rispetto quelli calcolati come media su tutte le WTG.

### 3.3.4.7. Scomposizione wavelet e filtro di Kalman

Anche in questo caso è stata effettuata la scomposizione della serie utilizzando le wavelet Daubochies di D4 di ordine 2. E' stato scelto un ordine più basso perché altrimenti non si riusciva ad ottenere la convergenza nella stima del modello.

Figura 33 - Errore sulla velocità del vento per WTG (Wavelet + Kalman)

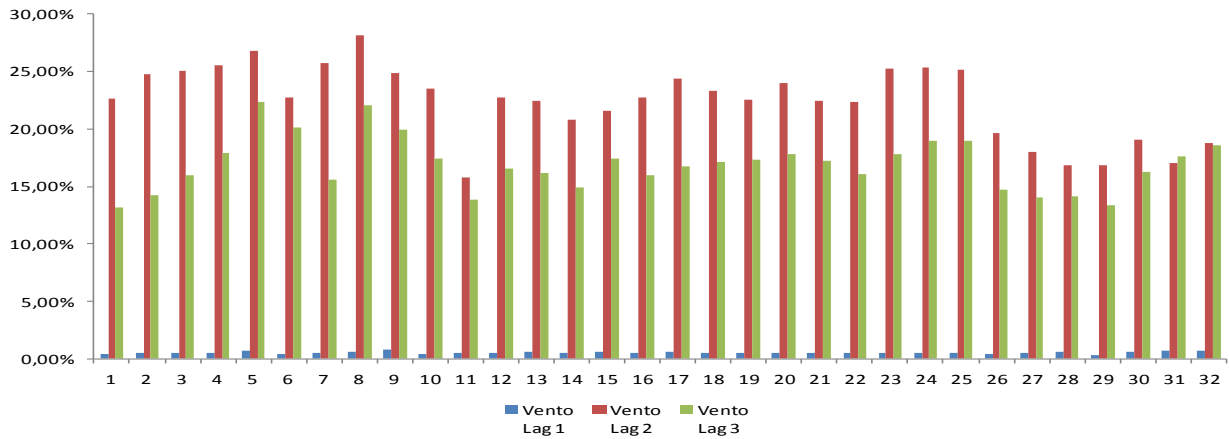


Figura 34 - Errore sull'energia prodotta per WTG (Wavelet + Kalman)

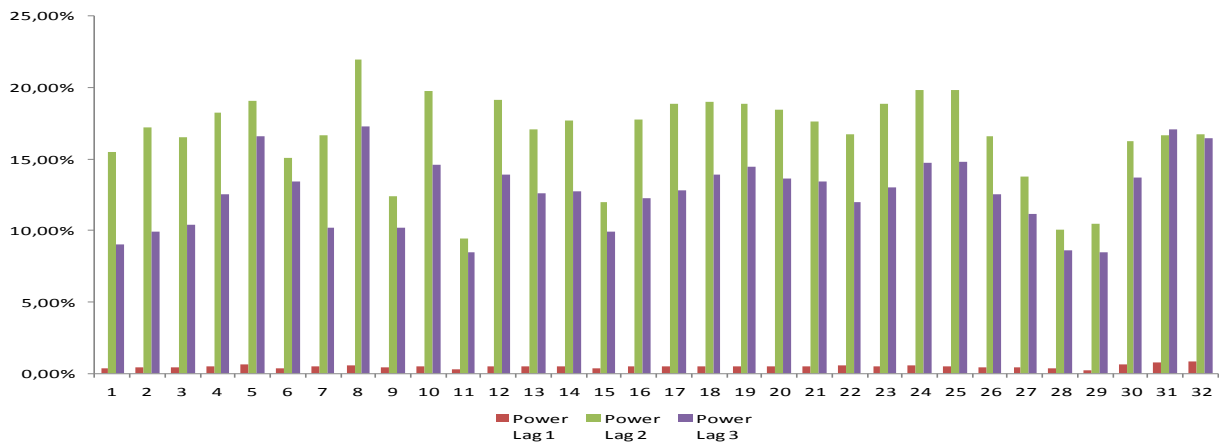


Tabella 3 - Riepilogo errori (Wavelet + Kalman)

	NMAE			MAPE		
	Power Lag 1	Power Lag 2	Power Lag 3	Vento Lag 1	Vento Lag 2	Vento Lag 3
Errore medio	0,50%	16,69%	12,66%	0,55%	22,39%	16,89%



Figura 35 - Wavelet + Kalman, Previsione vs velocità del vento reale LAG 1 (WTG 1 -16)

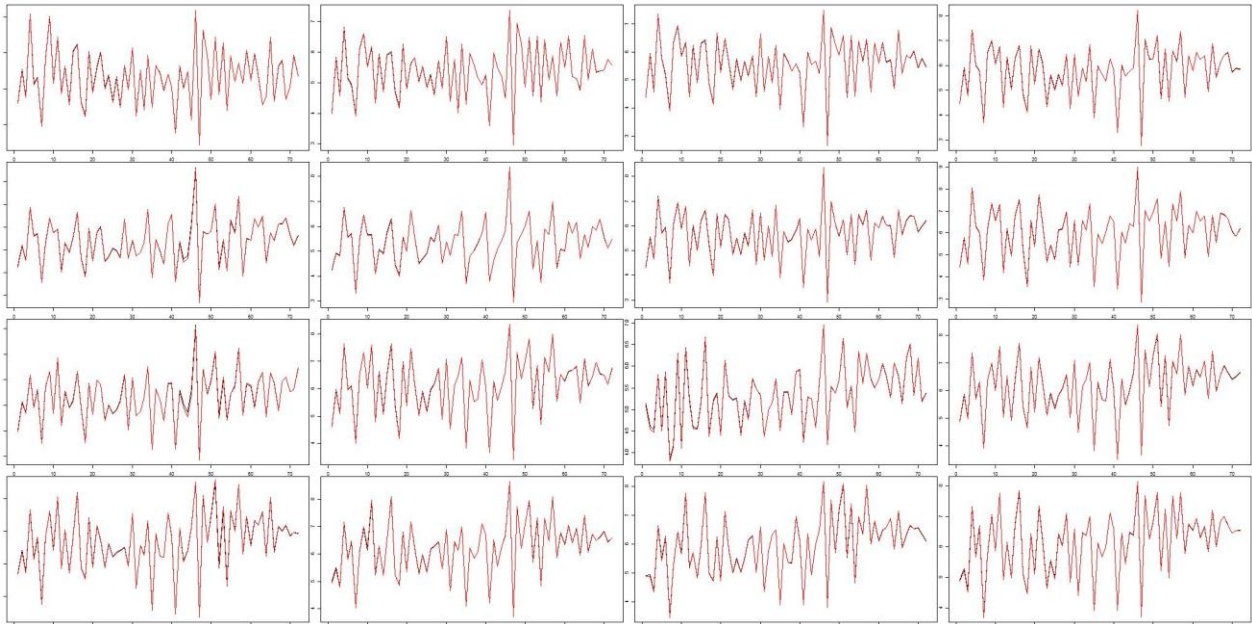


Figura 36 - Wavelet + Kalman, Previsione vs energia reale LAG 1 (WTG 1 -16)

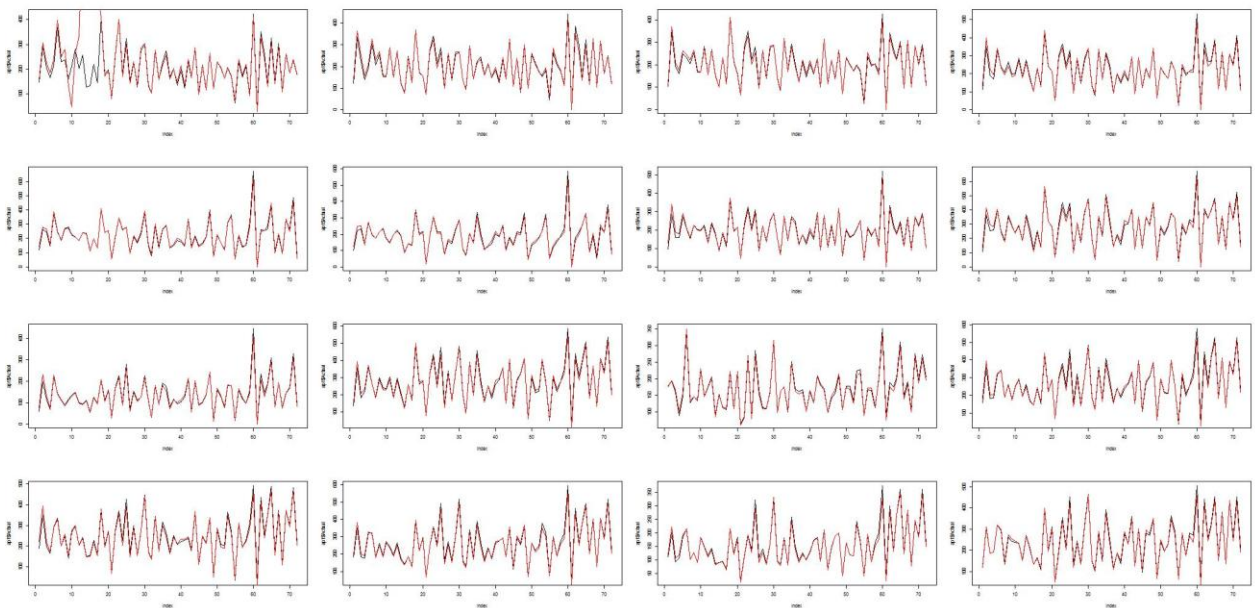


Figura 37 - Esempio di confronto curve velocità del vento reale e prevista (WTG 1- lag 1)

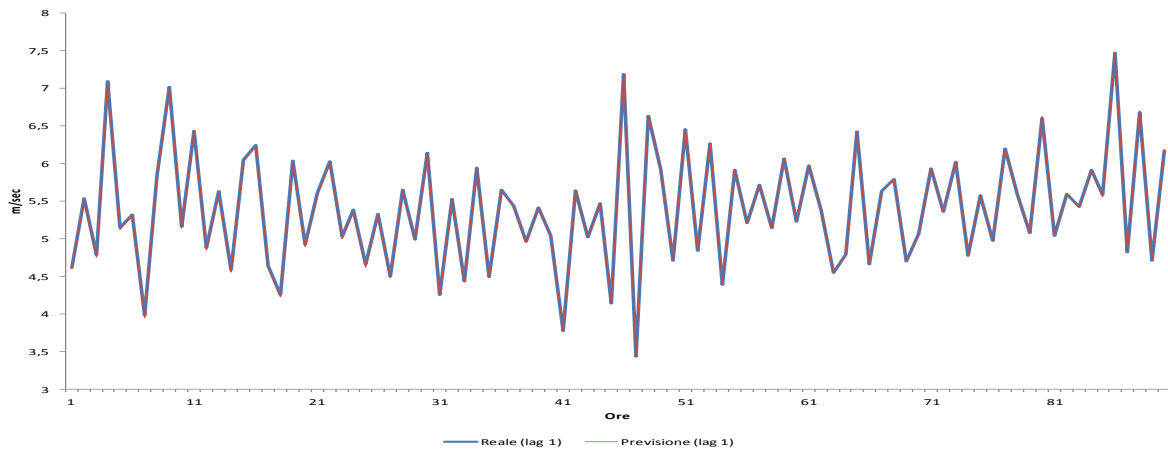
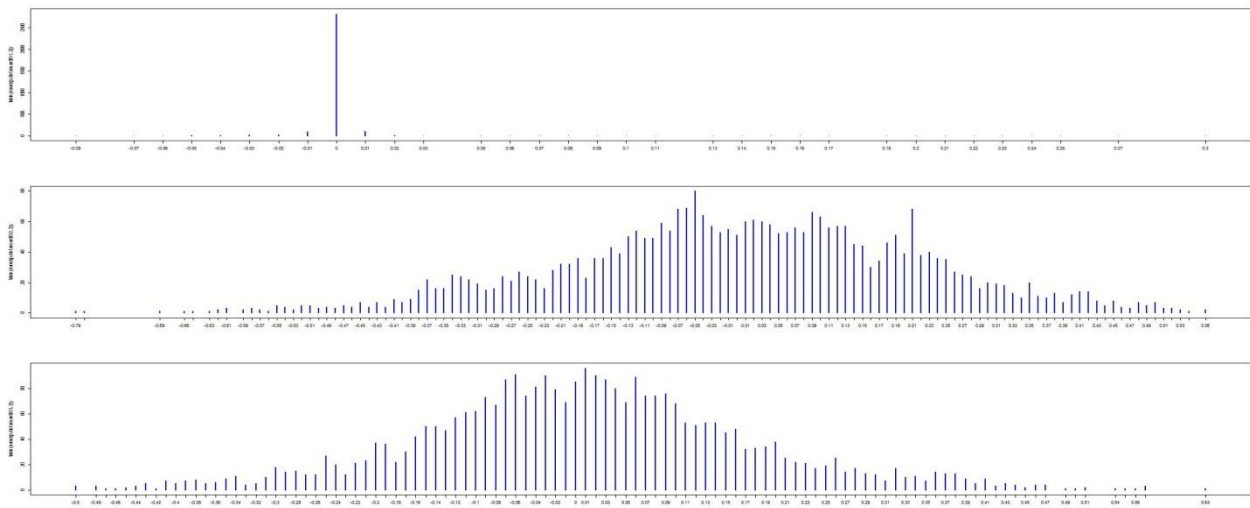


Figura 38 - Distribuzione errori su tutte le WTG per lag temporale di previsione (Wavelet + Kalman)



## Risultati per impianto

Figura 39 - Distribuzione degli errori per impianto (Wavelet – Kalman)



Figura 40 - Curve orarie previste e reale di impianto vs NMAE Lag temporale 1 (Wavelet + Kalman)

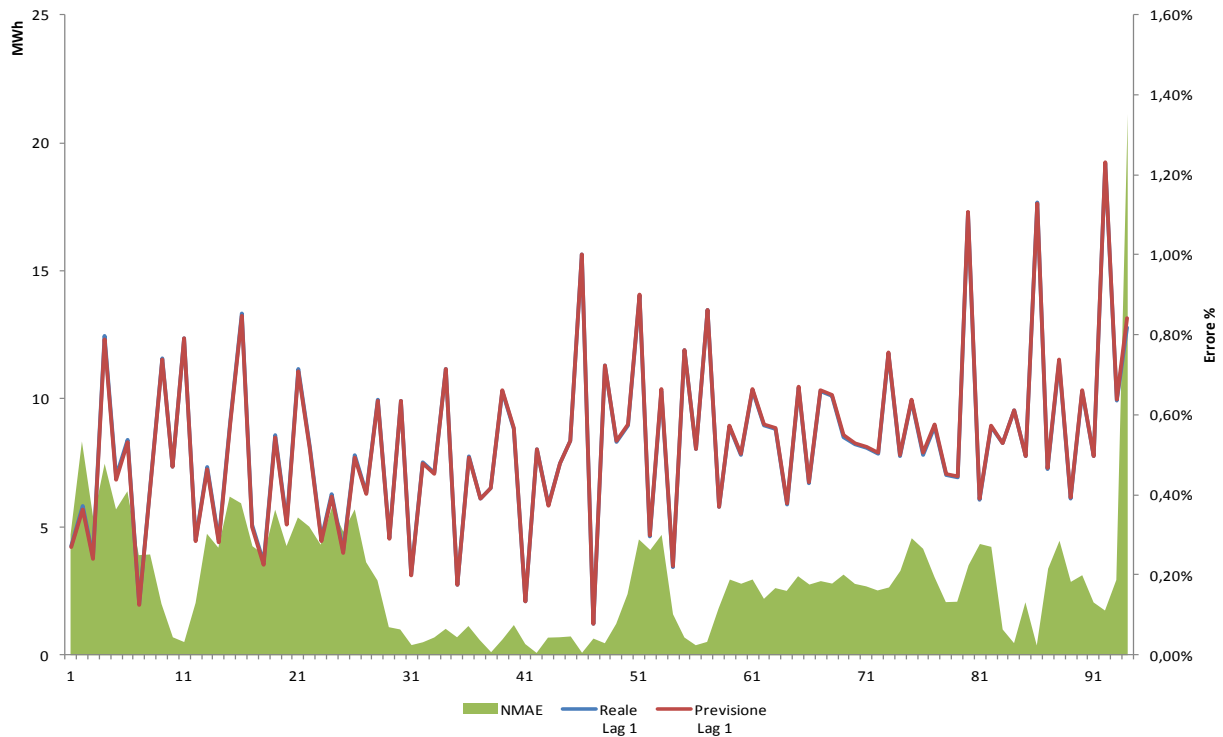
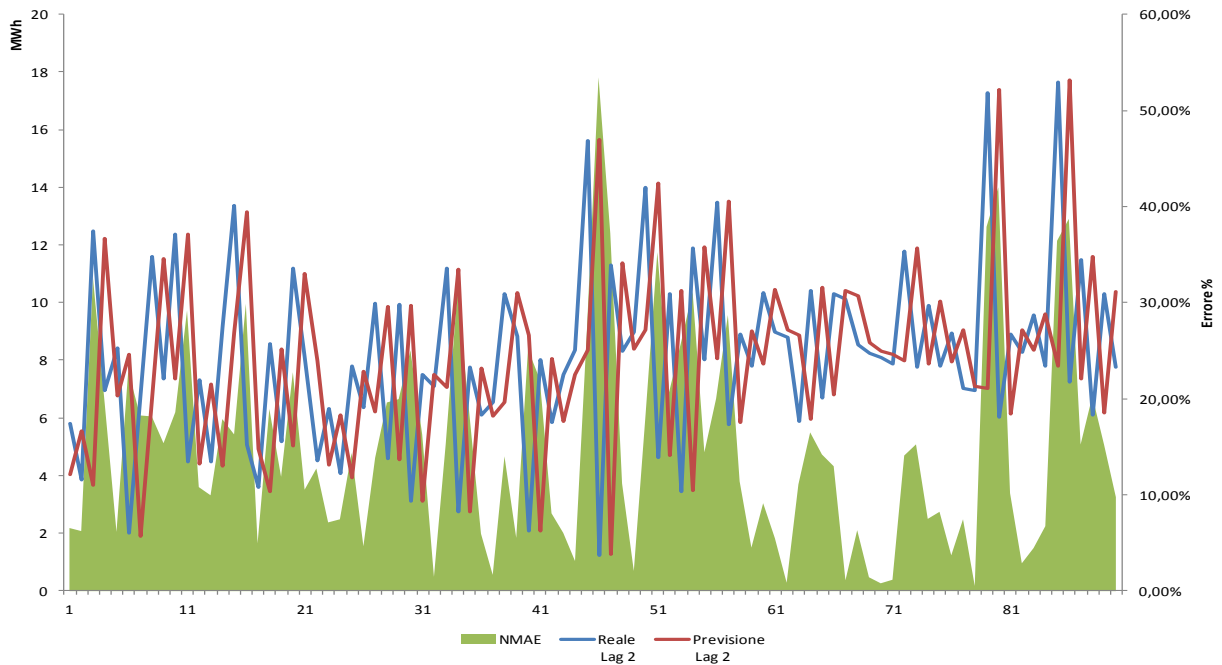
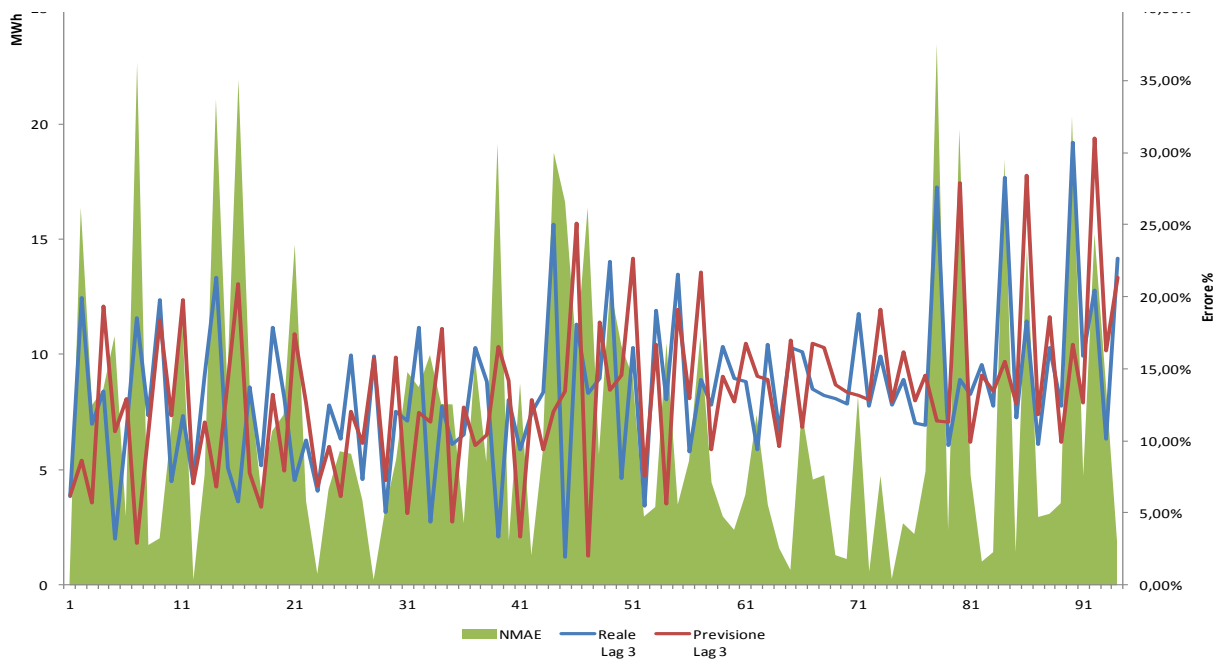


Figura 41 - Curve orarie previste e reale di impianto vs NMAE Lag temporale 2 (Wavelet + Kalman)



**Figura 42 - Curve orarie previste e reale di impianto vs NMAE Lag temporale 3 (Wavelet + Kalman)**



**Tabella 4 - Riepilogo errori per impianto (Wavelet + Kalman)**

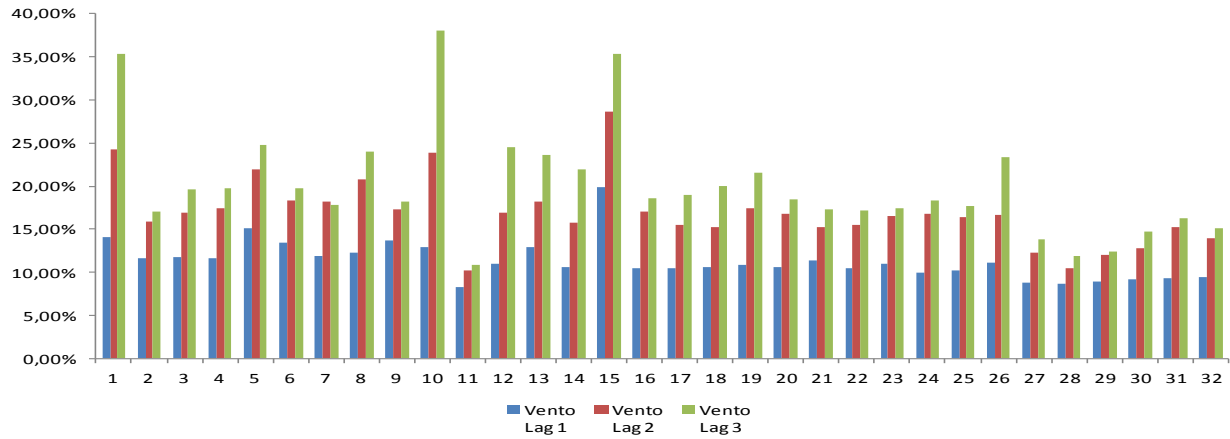
NMAE Lag 1	NMAE Lag 2	NMAE Lag 3
0,20%	15,56%	11,53%

Rispetto al precedente, questo modello ha prestazioni addirittura migliori per quanto riguarda il primo orizzonte temporale (1 ora) mentre sugli altri va un po' peggio. Questo sia nel caso di singole WTG sia nel caso di intero impianto. L'andamento orario del Lag 1 mostra un accordo completo tra la previsione e la produzione reale.

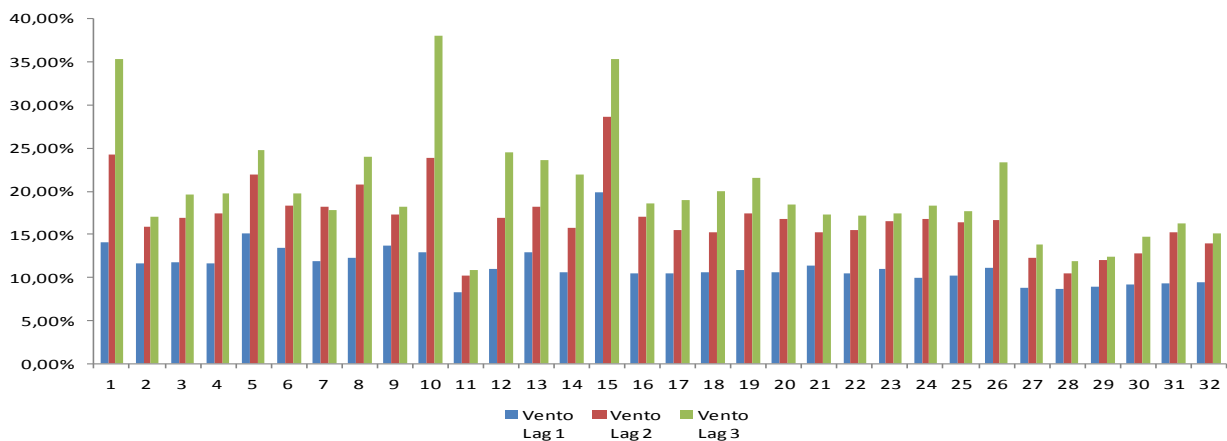
### 3.3.4.8. Scomposizione wavelet e rete neurale

In questo caso si è cercato l'utilizzo di un modello non lineare come sono le reti neurali artificiali, sempre associato alla scomposizione wavelet

**Figura 43 - Errore sulla velocità del vento per WTG (Wavelet + rete neurale)**



**Figura 44 - Errore sull'energia prodotta per WTG (Wavelet + rete neurale)**



**Tabella 5 - Riepilogo errori (Wavelet + rete neurale)**

	NMAE			MAPE		
	Power Lag 1	Power Lag 2	Power Lag 3	Vento Lag 1	Vento Lag 2	Vento Lag 3
Errore medio	7,60%	11,82%	13,18%	11,33%	16,89%	20,12%

Figura 46 - Wavelet + rete neurale Previsione vs velocità del vento reale LAG 1 (WTG 1 -16)

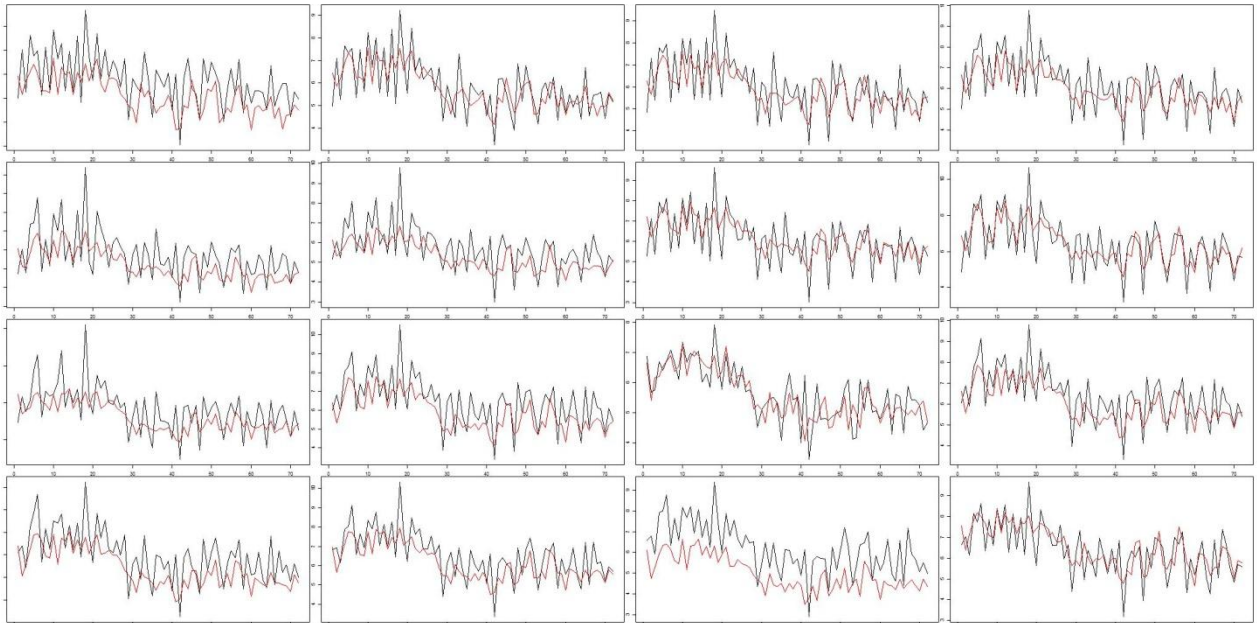


Figura 45 - Wavelet + rete neurale, Previsione vs energia prodotta reale LAG 1 (WTG 1 -16)

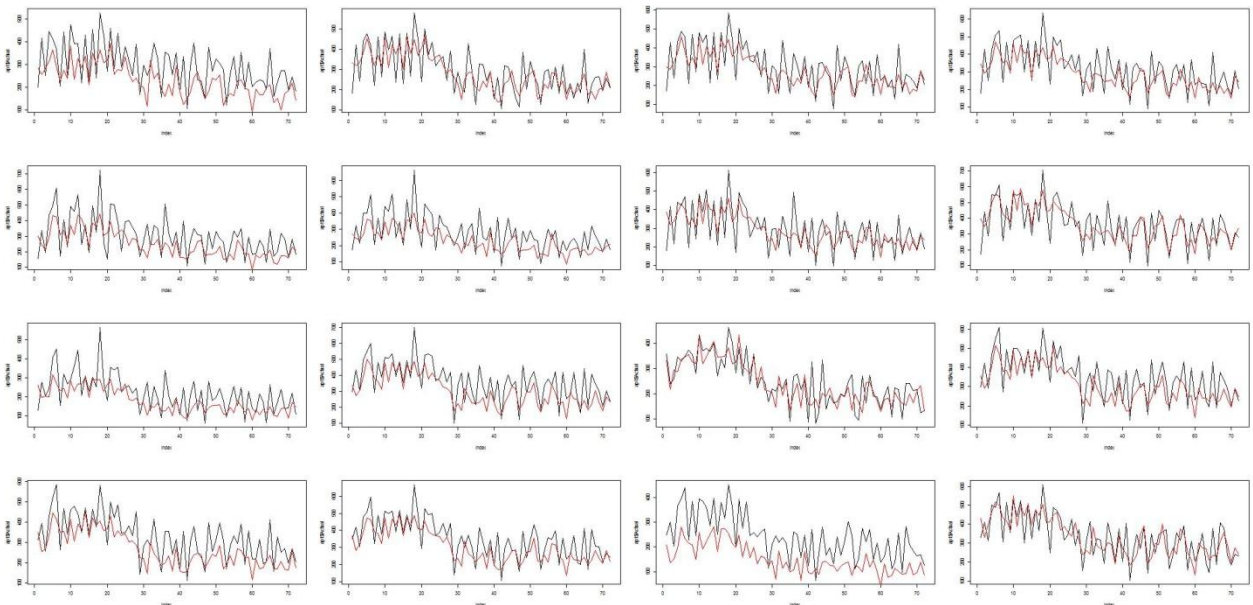


Figura 47 - Esempio di confronto curve velocità del vento reale e prevista (WTG 1- lag 1)

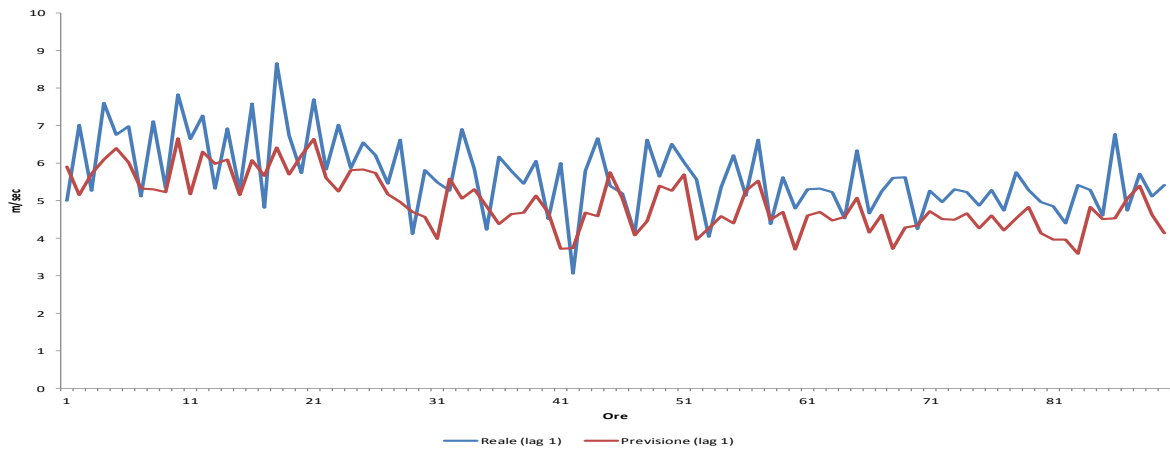
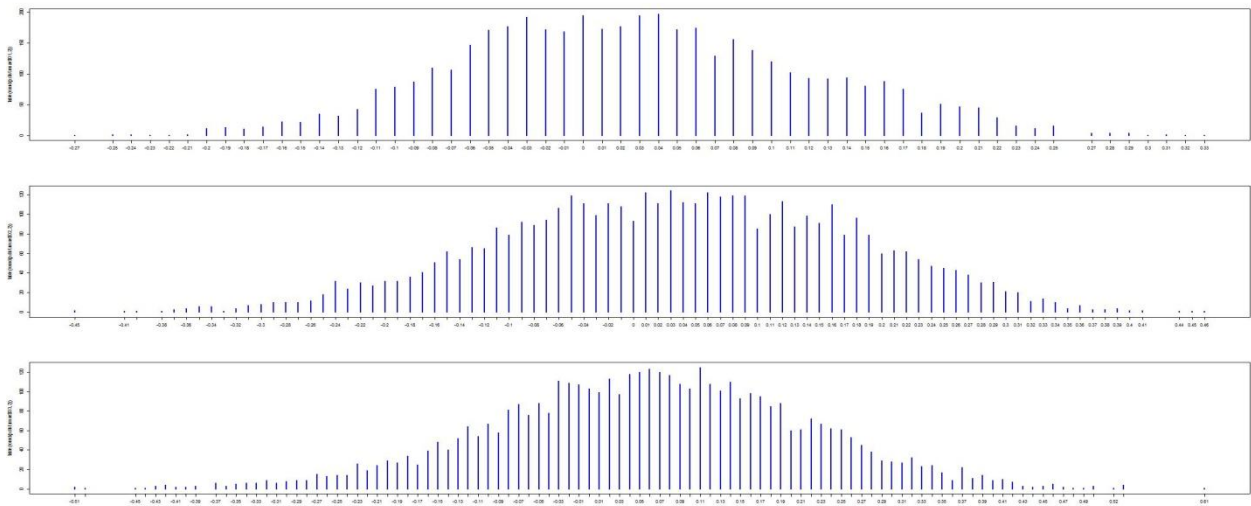


Figura 48 - Distribuzione errori su tutte le WTG per lag temporale di previsione (Wavelet + rete neurale)



## Risultati per impianto

Figura 49 - Distribuzione degli errori per impianto

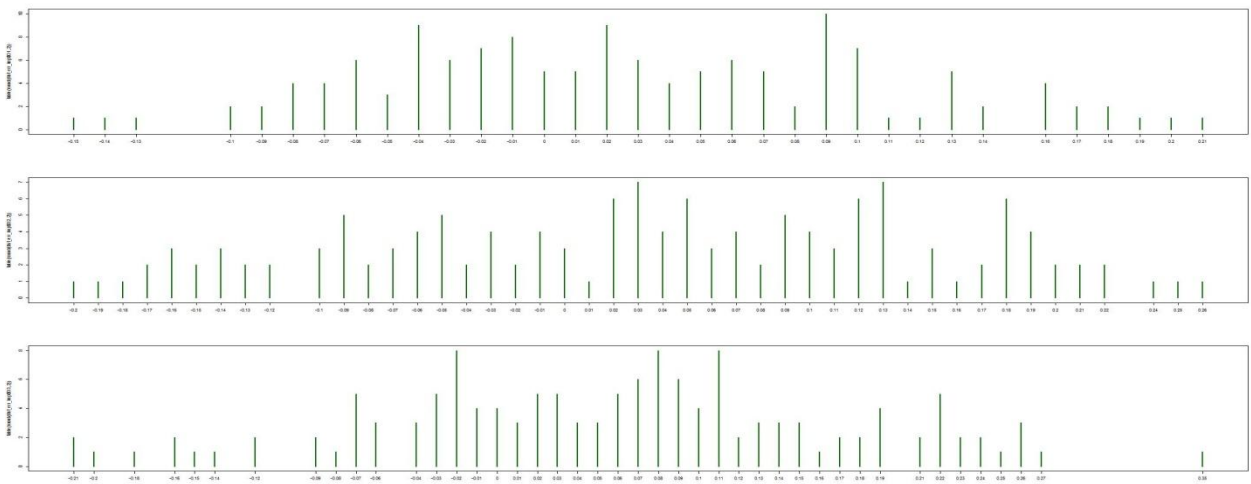


Figura 50 - Curve orarie previste e reale di impianto vs NMAE Lag temporale 1 (Wavelet + rete neurale)

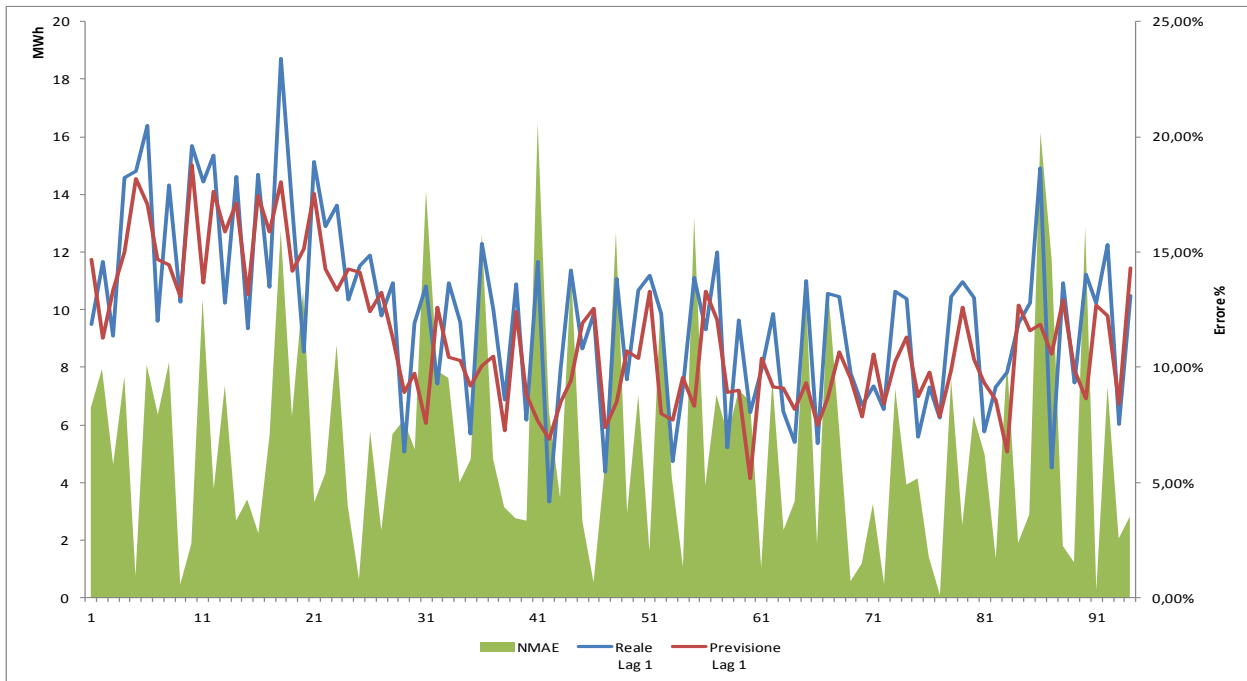
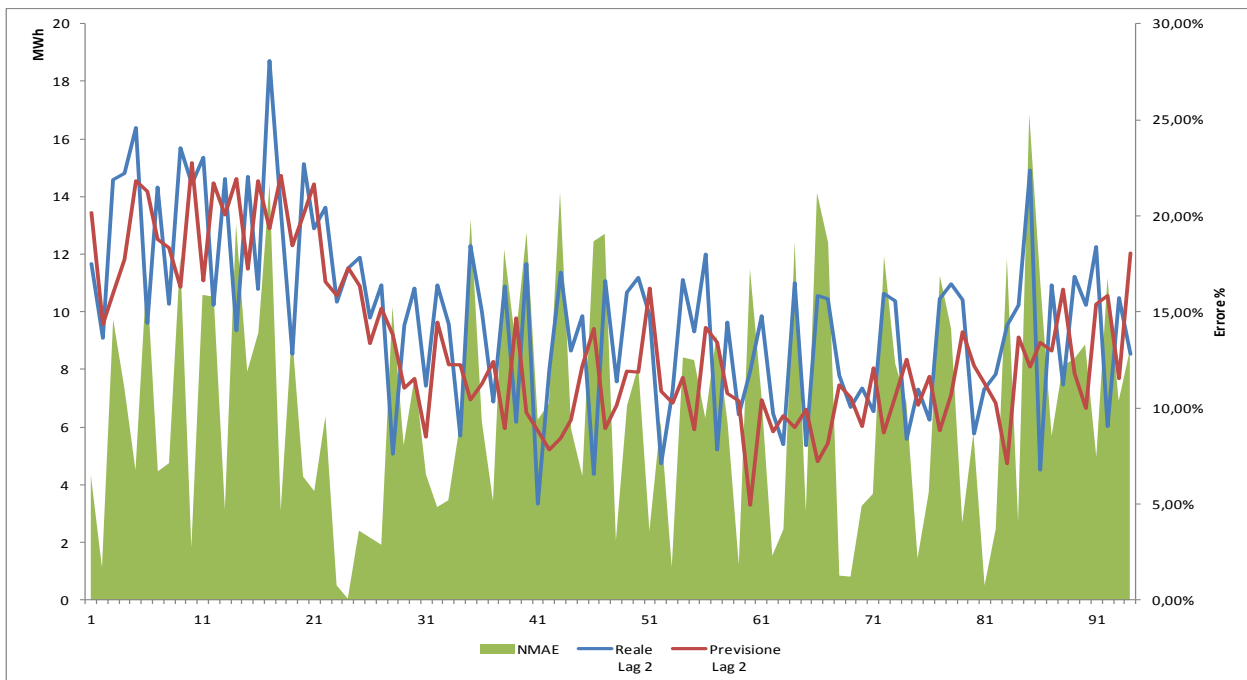
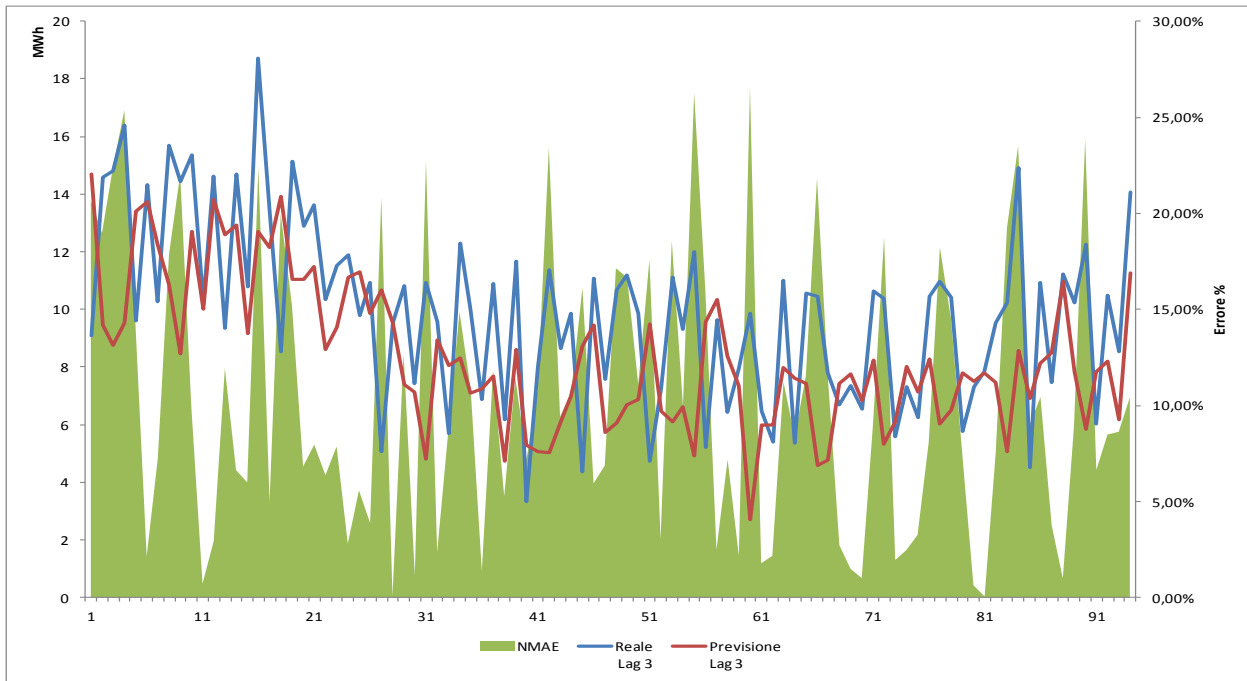


Figura 51 - Curve orarie previste e reale di impianto vs NMAE Lag temporale 2 (Wavelet + rete neurale)





**Figura 52 - Curve orarie previste e reale di impianto vs NMAE Lag temporale 3 (Wavelet + rete neurale)**



**Tabella 6 - Riepilogo errori per impianto (Wavelet + Rete Neurale)**

NMAE Lag 1	NMAE Lag 2	NMAE Lag 3
6,73%	9,94%	10,06%

In questo caso si notano delle prestazioni peggiori sul Lag 1 mentre migliorano quelle sugli altri orizzonti temporali, anche la distribuzione degli errori appare più ampia rispetto alle precedenti. L'andamento orario denota ovviamente uno scarso accordo ma si vede come il modello riesce a seguire comunque abbastanza bene il trend. Inoltre i valori più bassi di errore negli altri orizzonti temporali rende questo approccio abbastanza promettente.

### 3.3.4.9. Persistenza

Per valutare i risultati di un metodo di previsione, in genere si effettua un confronto con il metodo più semplice possibile. La persistenza consiste nell'utilizzare l'ultimo valore reale di una serie come previsione per il futuro. In questo caso viene utilizzato il valore della velocità del vento all'ora  $n$  come migliore stima delle ore  $n+1$ ,  $n+2$ ,  $n+3$ .

Figura 54 - Istogramma errori sulla velocità del vento per WTG

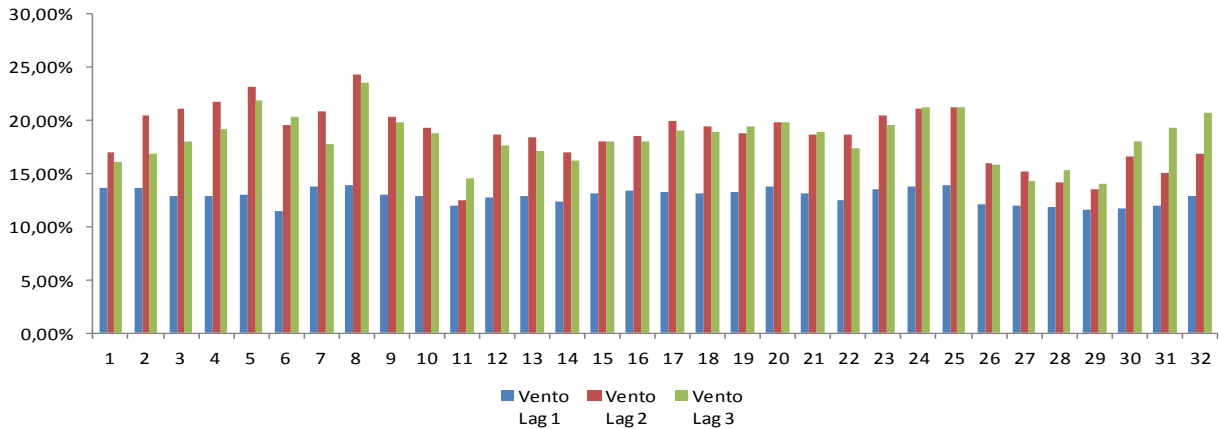


Figura 55 - Istogramma errori sulla energia per WTG

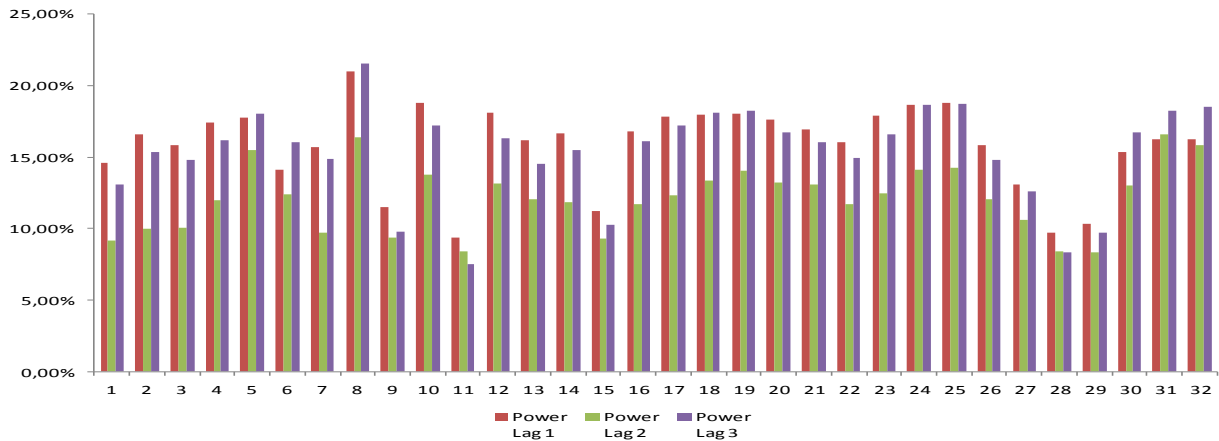


Figura 53 - Errori persistenza

	NMAE			MAPE		
	Power Lag 1	Power Lag 2	Power Lag 3	Vento Lag 1	Vento Lag 2	Vento Lag 3
Errore medio	15,89%	12,14%	15,36%	12,89%	18,65%	18,36%

### 3.4. Considerazioni

Di seguito è riportata una tabella riepilogativa degli errori calcolati nei diversi casi.

Tabella 7 - Riepilogo errori

	Power NMAE per WTG			Power NMAE Impianto			Vento MAPE per WTG		
	Lag 1	Lag 2	Lag 3	Lag 1	Lag 2	Lag 3	Lag 1	Lag 2	Lag 3
<b>Persistenza</b>	15,89%	12,14%	15,36%	14,38%	10,78%	12,98%	12,89%	18,65%	18,36%
<b>Arima</b>	1,40%	15,89%	11,91%	1,28%	15,04%	10,48%	1,74%	21,59%	16,40%
<b>Kalman</b>	0,50%	16,69%	12,66%	0,20%	15,56%	11,53%	0,55%	22,39%	16,89%
<b>Neural</b>	7,60%	11,82%	13,18%	6,73%	9,94%	10,06%	11,33%	16,89%	20,12%

Ricordiamo che il metodo scelto utilizza solo la velocità del vento misurata dall'anemometro di turbina, memorizzata dal suo sistema di automazione ed inviata al sistema che colleziona tutti i dati di impianto. Tale misura non ha mai una precisione elevata, sicuramente non paragonabile a strumenti più specifici che però raramente sono presenti in un impianto. Inoltre, sappiamo che la produzione eolica di energia elettrica è legata anche ad altre grandezze come la direzione del vento e la densità dell'aria. Grandezze che ovviamente variano nel tempo. Lo scopo qui era però verificare la fattibilità di una previsione con il minimo set di dati disponibile.

E' necessario sottolineare ancora che il tipo di errore calcolato sulla velocità del vento non è lo stesso calcolato per la produzione di energia elettrica, quindi non si deve fare un confronto quantitativo. L'ordine di grandezza dell'errore assoluto medio (MAPE) è in genere più elevato dell'errore assoluto normalizzato (NMAE)

Altro aspetto da ricordare è che la produzione di energia è legata alla velocità del vento attraverso una relazione cubica (vedi 2.1) questo ovviamente tende a dilatare l'errore commesso sul vento. Del resto però la serie di vento è più regolare di quella di produzione, quindi risulta più conveniente passare comunque per questa.

I valori di errore calcolati come media su tutte le WTG sono leggermente più alti di quelli calcolati considerando la produzione complessiva dell'impianto. Questo è dovuto al fatto che esiste un debole fenomeno di compensazione, tale fenomeno non è molto marcato perché i modelli si comportano allo stesso modo su tutte le WTG. A livello di impianto viene riportata solo la previsione di produzione in quanto non avrebbe senso determinare una velocità unica.

I modelli inoltre sono uguali per tutte le WTG e sebbene il vento non presenti differenze sostanziali all'interno del perimetro di impianto, c'è da dire che possono esistere fenomeni di interferenza tra una WTG ed un'altra come pure può verificarsi una differenza di rendimento delle WTG.

Per quanto riguarda l'intervallo di previsione di un ora (Lag 1) ogni metodo è migliore della persistenza, questo rappresenta già un buon risultato in quanto in genere nel tempo reale, difficilmente si riesce a fare meglio. Il livello di errore del metodo Kalman in questo orizzonte temporale, è inferiore al 1% sia nella previsione di vento sia nella previsione di energia, tale risultato è molto interessante anche se un orizzonte temporale così breve non è detto sia molto utile.

Sui successivi orizzonti temporali il fenomeno che emerge curiosamente è innanzitutto dovuto al fatto che, gli errori a Lag 2 sono in genere più alti rispetto a quelli a Lag 3 soprattutto per i modelli filtro di Kalman e Arima, tale fenomeno andrebbe investigato meglio, una prima ipotesi potrebbe essere legata alla relativamente scarsa dimensione del campione su cui sono effettuate le medie. Inoltre c'è da dire che non sono stati calcolati livelli di confidenza per queste previsioni, la differenza tra l'errore a Lag 2 e quello a Lag 3 potrebbe in realtà non esistere se gli intervalli di confidenza si sovrapponevano.

L'unico modello le cui prestazioni nella previsione di energia (la grandezza di vero interesse) sono sempre migliori della persistenza (anche se di poco), è quello che vede l'utilizzo delle wavelet con una rete neurale. I valori di NMAE su Lag 2 e 3 sono molto buoni, crescono con l'aumentare dell'orizzonte temporale come ci si potrebbe aspettare. Questo tipo di approccio inoltre, consente di specializzare reti neurali diverse per

---

ogni orizzonte temporale quindi potenzialmente in grado di fornire prestazioni migliori, anche questo è un aspetto che potrebbe essere approfondito. In generale l'approccio più promettente sembra essere quello misto cioè utilizzando metodi diversi per diversi orizzonti temporali.

Tutto ciò ci porta a concludere che l'utilizzo anche della sola velocità del vento, misurata dall'anemometro di WTG, è un approccio fattibile anche se da approfondire, nella previsione di brevissimo termine della produzione elettrica di un impianto eolico.

---

## 4. Riferimenti e Bibliografia

- [1] [http://www.terna.it/default/Home/SISTEMA\\_ELETTRICO/statistiche/dati\\_statistici.aspx](http://www.terna.it/default/Home/SISTEMA_ELETTRICO/statistiche/dati_statistici.aspx)
- [2] Lange M, Focken U, Physical Approach to Short-Term Wind Power Prediction. Springer-Verlag: Berlin, Heidelberg, 2006.
- [3] Liu H, Tian H-Q, Chen C, Li Y-F. A hybrid statistical method to predict wind speed and wind power. *Renew Energy* 2010;35(8):1857–61.
- [4] Ma L, Luan SY, Jiang CW, Liu HL, Zhang Y. A review on the forecasting of wind speed and generated power. *Renew Sust Energy Rev* 2009;13(4):915–20.
- [5] *N. Amjady, F. Keynia*; “Day ahead price forecasting of electricity markets by a mixed data model and hybrid forecast method”. *Electrical Power and Energy Systems* 30 (2008) 533–546
- [6] *A.J.Conejo, M. A. Plazas, R. Espínola. A. B. Molina*; “Day-Ahead Electricity Price Forecasting Using the Wavelet Transform and ARIMA Models”. *IEEE TRANSACTIONS ON POWER SYSTEMS*, VOL. 20, NO. 2, MAY 2005
- [7] *S.Lahmiri*; “Wavelet Transform, Neural Network and the prediction of S&P Price index: a comparative study of back propagation numerical algorithms”. *Business Intelligence Journal* - July, 2012 Vol.5 No.2
- [8] *J.P.S. Catalão, H.M.I. Pousinho a, V.M.F. Mende*, Short-term Wind Power Forecasting in Portugal *Renewable Energy* Volume 36, Issue 4, April 2011, Pages 1245–1251
- [9] “La produzione di energia Eolica” monografia tecnica aziendale ENEL Green Power, 2011
- [10] *C. Chatfield*, “Time Series Forecast”, Chapman & Hall, 2000.
- [11] *B.Percival, A.T. Walden*; *Wavelet Methods for time series Analysis*. Cambridge University Press 2000
- [12] *G.Welch, G Bishop*; *An Introduction to the Kalman Filter*, SIGGRAPH course-8, 2001 16
- [13] *J.Hertz, A. Krogs, R.G, Palmer*; *Introduction to the theory of neural computation*. Addison Wesley 1994
- [14] *P.S.P Cowpertwait, A.V. Metcalf*; *Introductory Time Serie with R*, Springer, 2009