



**Politecnico
di Torino**

Politecnico di Torino

Corso di Laurea Magistrale in Ingegneria Elettrica
a.a. 2021/2022
Sessione di Laurea ottobre 2022

Analisi dei metodi di ricostruzione delle curve di carico

Relatore:

Gianfranco Chicco

Candidati:

Roberta Dabbicco

Co-relatore:

Andrea Mazza

ABSTRACT

La tesi ha l'obiettivo di ricostruire, con l'ausilio di reti neurali, gli andamenti della potenza elettrica richiesta da un carico, quando incognita in alcuni periodi di tempo. I dati numerici necessari, ossia la potenza campionata e gli istanti di misura, sono stati ricavati dalla banca dati REFIT. Per la costruzione delle reti neurali, si è sfruttato il software *Matlab* e due suoi applicativi (*Regression Learner* e *Neural Net Time Series*). In essi sono presenti delle reti neurali pre-programmate che richiedono i valori di *predictor* (grandezze sempre note, impiegate per la previsione) e della potenza nei giorni antecedenti a quello incognito.

Inizialmente, sono stati utilizzati solo *predictor* "temporali", ricavati dai momenti di campionamento. Poiché i risultati così raggiunti non sono stati soddisfacenti, si sono aggiunti altri *predictor*, i valori medi della potenza, sul minuto e sull'ora, sia nel giorno incognito che in quello precedente. Inoltre, l'incremento del numero di neuroni costituenti le reti neurali e la media tra i risultati dei diversi modelli sono stati altri due tentativi condotti per giungere a ricostruzioni più accurate. Si è concluso che la singola rete *wide*, composta da 50 neuroni con sufficienti informazioni, dia un ottimo risultato; quindi, essa è stata la sola impiegata nelle fasi successive.

A seguire, sono stati sostituiti i *predictor* relativi ai valori medi nel giorno incognito con quelli di giorni precedenti, simulando la reale mancanza delle misure. Questo è stato ritenuto accettabile dallo studio delle curve di carico nei singoli giorni, nei quali si sono individuati almeno quattro andamenti che si ripetono; questi sono associabili a carichi che, in momenti diversi, conducono alla stessa variazione della potenza totale. I risultati raggiunti con il nuovo modello presentano un prevedibile peggioramento, con la cancellazione di molte delle oscillazioni reali della potenza. Si è giunti ad un miglioramento di tali risultati costruendo un andamento tipo dei carichi ripetitivi e sostituendolo nel risultato della previsione fornita dal modello *wide*, quando essa presenta un'evoluzione che richiami uno di questi carichi.

In conclusione, i modelli di *Matlab* sono accurati solo se è disponibile un adeguato numero di informazioni; tuttavia, l'apporto "umano" tratto dall'esperienza nel dominio specifico può condurre al miglioramento dei risultati.

SOMMARIO

ABSTRACT	3
INDICE DELLE FIGURE.....	9
INDICE DELLE TABELLE	15
1.....	INTRODUZIONE
.....	17
1.1 BANCHE DATI COME FONTI DEI DATI NUMERICI DI INTERESSE	18
1.1.1 Banche dati disponibili On-Line	18
1.1.2 Lacune della serie analizzata: Identificazione mese migliore	20
1.2 SOFTWARE E TECNICHE PER LA REALIZZAZIONE DEL MODELLO	24
1.2.1 Machine Learning e reti neurali.....	24
1.2.2 Regression Learner App [6].....	27
1.2.3 Neural Net Time Series App [8].....	29
2.....	ANALISI PRELIMINARE DEI LIMITI DELLE RETI NEURALI
.....	33
2.1 PREPARAZIONE DEI DATI	33
2.1.1 Predictor temporali e RMSE delle reti pre-programmate	33
2.2 RISULTATI ANALISI PRELIMINARE	36
2.2.1 Modelli costruiti.....	36
2.2.2 Confronto dei risultati.....	37
3.....	MIGLIORAMENTO DEI MODELLI
.....	47
3.1 VARIAZIONE DELLE CARATTERISTICHE DELLE RETI NEURALI	47
3.1.1 Incremento del numero di predictor	47
3.1.2 Incremento dei neuroni.....	52
3.2 CORRELAZIONE ERRORE-ISTANTE DI PREVISIONE.....	54
3.2.1 Valutazione sulla previsione del 22 gennaio 2015.....	54

3.2.2	<i>Valutazione sulla previsione del 21 marzo 2015</i>	55
3.3	SERIE TEMPORALI E NARX	57
3.3.1	<i>Definizione dei parametri delle reti narx impiegate</i>	58
3.3.2	<i>Ispezione grafica dei risultati</i>	58
3.4	MEDIA DEI RISULTATI	66
3.4.1	<i>Modelli da utilizzare nello svolgimento della media</i>	66
3.5	CONCLUSIONI DELLA FASE 1	77
4.	PERIODICITÀ E PREDICTOR MANCANTI	79
4.1	RICERCA DI PERIODICITÀ	79
4.1.1	<i>Periodicità giornaliera e settimanale</i>	79
4.1.2	<i>Confronto dei giorni precedenti</i>	81
4.1.3	<i>Adattamento dei dati</i>	85
4.1.4	<i>Metodo 1: giorno precedente</i>	86
4.1.5	<i>Metodo 2: medesimo giorno della settimana precedente</i>	89
4.1.6	<i>Metodo 3: medesimo giorno di due settimane prima</i>	92
4.2	CONCLUSIONI DELLA FASE 2	94
5.	MIGLIORAMENTI DEGLI ANDAMENTI OTTENUTI	97
5.1	TIPOLOGIA DI CARICHI RIPETITIVI	97
5.1.1	<i>Identificazione carichi ripetitivi</i>	97
5.1.2	<i>Step 1: ricerca tipologia carichi ripetitivi nei giorni del mese</i>	99
5.1.3	<i>Step 2: manipolazione dei dati</i>	101
5.1.4	<i>Step 3: Carichi di tipologia A</i>	102
5.1.5	<i>Step 4: carichi di tipologia c</i>	105
5.1.6	<i>step 5: carichi di tipologia D</i>	107
5.1.7	<i>Step 6: carichi di tipologia B</i>	110
5.2	MIGLIORAMENTO DEL RISULTATO DELLA SIMULAZIONE	111

5.2.1	<i>Caso 1: modello con predictor del 15 gennaio</i>	112
5.2.2	<i>Caso 2: modello con predictor dell'8 gennaio</i>	115
5.3	RIDUZIONE DEGLI INTERVALLI DI ANALISI	119
5.3.1	<i>Mancanza di 30 minuti nelle ore serali</i>	119
6	SIMULAZIONE FINALE	123
6.1	SCALETTA RIASSUNTIVA	123
6.2	RISULTATI DELLA PREVISIONE DEL 17 GENNAIO 2015	124
6.2.1	<i>Visualizzazione grafica dei risultati</i>	124
7	CONCLUSIONI	129
	BIBLIOGRAFIA	131
	SITOGRAFIA	131
	RINGRAZIAMENTI	133

INDICE DELLE FIGURE

Figura 1.1.2-1 Time table originale	21
Figura 1.1.2-2 Carico aggregato abitazione 1 (REFIT)	22
Figura 1.1.2-3 Numero di anomalie superiori ai 16 secondi per abitazione 1 (REFIT).....	22
Figura 1.1.2-4 Carico aggregato nel mese di gennaio 2015	24
Figura 1.2.1-1 Schema semplificato di una rete neurale	25
Figura 1.2.2-1 Regression Learner App	28
Figura 1.2.3-1 NARX anello aperto.....	31
Figura 1.2.3-2 NARX anello chiuso	31
Figura 2.1.1-1 Filtro delle misurazioni per ottenere differenti frequenze.....	34
Figura 2.2.2-1 RMSE per i modelli costruiti con campioni a 8 secondi	37
Figura 2.2.2-2 RMSE per i modelli costruiti con campioni a 16 secondi	38
Figura 2.2.2-3 RMSE per i modelli costruiti con campioni a 24 secondi	38
Figura 2.2.2-4 Previsione delle 12 ore diurne con modello a 6 giorni e $T_{\text{campionamento}}=8s$	39
Figura 2.2.2-5 Previsione delle 12 ore pomeridiane con modello a 6 giorni e $T_{\text{campionamento}}=8s$...	40
Figura 2.2.2-6 Previsione di un giorno con modello a 4 giorni e $T_{\text{campionamento}}=16s$	41
Figura 2.2.2-7 Previsione di 2 giorni con modello a 2 giorni e $T_{\text{campionamento}}=24 s$	41
Figura 2.2.2-8 Previsione di 4 giorni con modello a 6 giorni e $T_{\text{campionamento}}=8 s$	42
Figura 2.2.2-9 Previsione di 6 giorni con modello ad 1 mese e $T_{\text{campionamento}}=16 s$	42
Figura 2.2.2-10 Previsione di 2 giorni con modello ad 1 giorno e $T_{\text{campionamento}}=24 s$	43
Figura 2.2.2-11 Previsione di 6 giorni con modello a 2 giorni e $T_{\text{campionamento}}=16 s$	43
Figura 2.2.2-12 Confronto RMSE ottenuti per tutti i modelli	45
Figura 3.1.1-1 Previsione del 22 gennaio 2015 con modello ad un mese e 11 <i>predictor</i>	50
Figura 3.1.1-2 Dettaglio della porzione nel riquadro verde di Figura 3.1.1-1.....	51
Figura 3.1.1-3 Dettaglio della porzione nel riquadro rosso di Figura 3.1.1-2	51
Figura 3.1.1-4 Dettaglio delle ore notturne in Figura 3.1.1-1	52

Figura 3.1.2-1 Dettaglio della previsione del 22 gennaio con rete wide 30 neuroni	54
Figura 3.2.1-1 Dettaglio della previsione con rete <i>wide</i> , 50 neuroni e 11 <i>predictor</i>	55
Figura 3.2.2-1 Carico aggregato reale nel mese di marzo 2015	56
Figura 3.2.2-2 Previsione del 21 marzo con rete <i>wide</i> , pre-programmata	56
Figura 3.2.2-3 Dettaglio di Figura 3.2.2-2	57
Figura 3.3.2-1 Previsione ottenuta con rete NARX costituita da 30 neuroni e ritardo 4	59
Figura 3.3.2-2 Previsione ottenuta con rete NARX costituita da 30 neuroni e ritardo 8	60
Figura 3.3.2-3 Dettaglio Figura 3.3.2-1 : disturbo nella potenza prevista con rete NARX a 30 neuroni e ritardo 4	61
Figura 3.3.2-4 Dettaglio Figura 3.3.2-2: disturbo nella potenza prevista con rete NARX a 30 neuroni e ritardo 8	61
Figura 3.3.2-5 Dettaglio Figura 3.3.2-1: potenza prevista nelle ore notturne con rete NARX a 30 neuroni e ritardo 4	62
Figura 3.3.2-6 Dettaglio Figura 3.3.2-2: potenza prevista nelle ore notturne con rete NARX a 30 neuroni e ritardo 8	62
Figura 3.3.2-7 Previsione ottenuta con rete NARX costituita da 50 neuroni e ritardo 4	63
Figura 3.3.2-8 Dettaglio Figura 3.3.2-7: potenza prevista nelle ore notturne con rete NARX a 50 neuroni e ritardo 4	63
Figura 3.3.2-9 Istogramma RMSE per i modelli NARX con ritardo inferiore a 6	65
Figura 3.3.2-10 Previsione con rete NARX con 40 neuroni e ritardo 1	65
Figura 3.3.2-11 Previsione con rete NARX con 40 neuroni e ritardo 5	66
Figura 3.4.1-1 Previsione con rete neurale <i>wide</i> , 50 neuroni	67
Figura 3.4.1-2 Previsione con rete neurale NARX, 30 neuroni e ritardo 1	67
Figura 3.4.1-3 Previsione con rete neurale NARX, 30 neuroni e ritardo 2	67
Figura 3.4.1-4 Previsione con rete neurale NARX, 30 neuroni ritardo 3	68
Figura 3.4.1-5 Previsione con rete neurale NARX, 30 neuroni ritardo 4	68
Figura 3.4.1-6 Previsione con rete neurale NARX, 40 neuroni ritardo 1	68

Figura 3.4.1-7 Previsione con rete neurale NARX, 40 neuroni ritardo 2	69
Figura 3.4.1-8 Previsione con rete neurale NARX, 40 neuroni e ritardo 3.....	69
Figura 3.4.1-9 Previsione con rete neurale NARX, 40 neuroni ritardo 4	69
Figura 3.4.1-10 Previsione con rete neurale NARX, 50 neuroni e ritardo 1	70
Figura 3.4.1-11 Previsione con rete neurale NARX, 50 neuroni e ritardo 2	70
Figura 3.4.1-12 Previsione con rete neurale NARX, 50 neuroni e ritardo 3	70
Figura 3.4.1-13 Previsione con rete neurale NARX, 50 neuroni e ritardo 4	71
Figura 3.4.1-14 Ricostruzione dell'andamento della potenza aggregata del giorno 22 gennaio 2015 mediante media istante per istante.....	71
Figura 3.4.1-15 Confronto medesimi istanti della previsione con rete neurale wide e con processo di mediatura.....	72
Figura 3.4.1-16 Previsione rete neurale <i>narrow</i> , campionamento 8 secondi e 11 <i>predictor</i>	74
Figura 3.4.1-17 Previsione con rete neurale <i>medium</i> , campionamento 8 secondi e 11 <i>predictor</i>	74
Figura 3.4.1-18 Previsione con rete neurale <i>wide</i> , campionamento 8 secondi e 11 <i>predictor</i>	75
Figura 3.4.1-19 Previsione con rete neurale <i>bilayered</i> , campionamento 8 secondi e 11 <i>predictor</i>	75
Figura 3.4.1-20 Previsione con rete neurale <i>trilayered</i> , campionamento 8 secondi e 11 <i>predictor</i>	75
Figura 3.4.1-21 Confronto tra potenza aggregata reale e prevista con media dei risultati dei modelli precedenti, campionamento 8 secondi e 11 <i>predictor</i>	76
Figura 3.4.1-22 Potenza prevista ottenuta per media dei modelli precedenti filtrata.....	77
Figura 4.1.1-1 Schema di principio per identificazione <i>predictor</i> mancanti dai giorni precedenti	80
Figura 4.1.2-1 Potenza aggregata reale 22 gennaio 2015	81
Figura 4.1.2-2 Potenza aggregata reale 21 gennaio 2015	81
Figura 4.1.2-3 Potenza aggregata reale 20 gennaio 2015	82
Figura 4.1.2-4 Potenza aggregata reale 19 gennaio 2015	82
Figura 4.1.2-5 potenza aggregata reale 18 gennaio 2015	82

Figura 4.1.2-6 Potenza aggregata reale 17 gennaio 2015	83
Figura 4.1.2-7 Potenza aggregata reale 16 gennaio 2015	83
Figura 4.1.2-8 Potenza aggregata reale 15 gennaio 2015	83
Figura 4.1.2-9 Potenza aggregata reale 8 gennaio 2015	84
Figura 4.1.4-1 Confronto tra potenza aggregata realtà e previsione con <i>predictor</i> ottenuti dal 21 gennaio 2015	87
Figura 4.1.4-2 Dettaglio ore notturne della previsione in Figura 4.1.4-1.....	87
Figura 4.1.4-3 Dettaglio ore notturne della previsione in Figura 4.1.3 1	87
Figura 4.1.4-4 Dettaglio ore serali della previsione in Figura 4.1.3 1.....	87
Figura 4.1.5-1 Confronto tra potenza aggregata realtà e previsione con <i>predictor</i> ottenuti dal 15 gennaio 2015	91
Figura 4.1.5-2 Dettaglio ore notturne della previsione in Figura 4.1.5-1.....	91
Figura 4.1.5-3 Dettaglio ore notturne della previsione in Figura 4.1.4 1	91
Figura 4.1.5-4 Dettaglio ore diurne della previsione in Figura 4.1.4 1	91
Figura 4.1.5-5 Dettaglio ore serali della previsione in Figura 4.1.4 1.....	91
Figura 4.1.6-1 Confronto tra potenza aggregata realtà e previsione con <i>predictor</i> ottenuti dal 8 gennaio 2015	93
Figura 4.1.6-2 Dettaglio ore notturne della previsione in Figura 4.1.6-1.....	94
Figura 4.1.6-3 Dettaglio ore notturne della previsione in Figura 4.1.5 1	94
Figura 4.1.6-4 Dettaglio ore diurne della previsione in Figura 4.1.5 1	94
Figura 4.1.6-5 Dettaglio ore serali della previsione in Figura 4.1.5 1.....	94
Figura 5.1.1-1 Previsione dell'andamento della potenza aggregata il 22 gennaio 2015 basandosi su quanto avviene il 15 gennaio 2015	97
Figura 5.1.1-2 Tipologia di carico A - ore notturne	98
Figura 5.1.1-3 Tipologia di carico B -mattina presto	98
Figura 5.1.1-4 Tipologia di carico C – picchi ore diurne	98
Figura 5.1.1-5 Tipologia di carico D - ore serali	98

Figura 5.1.4-1 Carico di tipo A disturbato	102
Figura 5.1.4-2 Carico di tipo A "puro"	102
Figura 5.1.4-3 Sovrapposizione dei carichi di tipo A nei giorni in analisi	103
Figura 5.1.4-4 Carico di tipologia A, ricostruito	104
Figura 5.1.4-5 Sostituzione della porzione della previsione del modello con il carico di tipo A ricostruito.....	105
Figura 5.1.5-1 Ricostruzione andamento della potenza di un carico di tipo C	106
Figura 5.1.5-2 Sostituzione della porzione della previsione del modello con il carico di tipo C ricostruito.....	107
Figura 5.1.6-1 Sovrapposizione carichi di tipo D rilevati nei giorni costituenti il mese di gennaio	108
Figura 5.1.6-2 Ricostruzione del carico di tipo D	109
Figura 5.1.6-3 Sostituzione della porzione della previsione del modello con il carico di tipo D ricostruito.....	109
Figura 5.1.7-1 Ricostruzione del carico di tipo B.....	110
Figura 5.2.1-1 Confronto tra potenza aggregata misurata, quella prevista dal modello wide e la medesima migliorata con <i>predictor</i> del 15 gennaio 2015	112
Figura 5.2.1-2 Dettaglio di Figura 5.2.1-1- ore notturne.....	113
Figura 5.2.1-3 Dettaglio di Figura 5.2.1 1- ore diurne.....	114
Figura 5.2.1-4 Dettaglio della Figura 5.2.1 1 - ore serali	115
Figura 5.2.2-1 Confronto tra potenza aggregata misurata, quella prevista dal modello wide e la medesima migliorata con <i>predictor</i> dell'8 gennaio 2015	116
Figura 5.2.2-2 Dettaglio di Figura 5.2.2-1- ore notturne.....	117
Figura 5.2.2-3 Dettaglio della Figura 5.2.2-1 - ore diurne.....	118
Figura 5.2.2-4 Dettaglio di Figura 5.2.2-1- ore serali	118
Figura 5.3.1-1 Misure reali del 22 gennaio con annullamento della potenza aggregata nell'intervallo 22:00-22:30.....	120

Figura 5.3.1-2 Confronto tra potenza reale e ricostruita con l'impiego del solo carico caratteristico di tipo D	120
Figura 5.3.1-3 Intervallo con l'ipotetica mancanza di misure 21:50-22:20 il giorno 22 gennaio 2015	121
Figura 5.3.1-4 Intervallo con l'ipotetica mancanza di misure 7:20-7:50 il giorno 22 gennaio 2015	122
Figura 6.2.1-1 Previsione del 17 gennaio 2015 con modello <i>wide</i> a 50 neuroni e predictor relativi al 10 gennaio 2015	125
Figura 6.2.1-2 Confronto tra risultato del modello <i>wide</i> prima e dopo i miglioramenti apportati	125
Figura 6.2.1-3 Confronto tra potenza campionata e risultati dei modelli di previsione	126

INDICE DELLE TABELLE

Tabella 1.1.1-1 Apparecchiature monitorate nell'abitazione selezionata.....	20
Tabella 2.1.1-1 Valori di <i>RMSE</i> per differenti modelli con campionamento ad 8 secondi	35
Tabella 2.1.1-2 Valori di <i>RMSE</i> per differenti modelli con campionamento ad 16 secondi	35
Tabella 2.1.1-3 Valori di <i>RMSE</i> per differenti modelli con campionamento ad 24 secondi	35
Tabella 2.2.1-1 Intervalli temporali di misure mancanti impiegate per il test dei modelli.....	36
Tabella 3.1.1-1 Valori di <i>RMSE</i> per il modello ad un mese completo dell'insieme di 11 <i>predictor</i>	49
Tabella 3.1.1-2 Valori di <i>RMSE</i> per il modello ad un mese con 10 e 9 <i>predictor</i>	49
Tabella 3.1.2-1 Valori di <i>RMSE</i> per la <i>wide neural network</i> con differente numero di neuroni...	53
Tabella 3.3.1-1 Valori di <i>RMSE</i> per le reti NARX con differente numero di neuroni e ritardi	58
Tabella 3.3.2-1 Valori di <i>RMSE</i> previsioni del 22 gennaio 2015 con reti NARX costituite da 30 neuroni.....	59
Tabella 3.3.2-2 <i>RMSE</i> per i modelli NARX aggiuntivi con ritardi inferiori a 6	64
Tabella 5.1.2-1 Tipologia di carichi per ogni giorno del mese di gennaio 2015.....	100

1. INTRODUZIONE

La conoscenza dell'energia elettrica impiegata in un'abitazione, o in un contesto più ampio quale un quartiere, è un elemento di importanza rilevante per una migliore gestione della risorsa stessa. Infatti, da un lato, l'ente che la distribuisce può tendere ad una suddivisione più omogenea del flusso di energia che interessa le sue linee, così da garantire una fornitura costante e, contemporaneamente, preservare la vita utile dei suoi stessi impianti. Dall'altro lato, l'utilizzatore finale della medesima energia può essere guidato in un consumo più attento, derivante da una maggiore consapevolezza dei consumi correlati a ciascun dispositivo.

Con le medesime informazioni si possono raggiungere altri obiettivi, quali il monitoraggio del deterioramento dei dispositivi, la valutazione del numero di persone presenti in un ambiente e molti altri.

Per ottenere i vantaggi sopra citati bisogna disporre di dati sull'evoluzione temporale della potenza, identificata anche come *diagramma di carico* o *curva di carico*. A tal fine, in svariati paesi del mondo, sono state condotte molte campagne di misure nel corso degli anni, con rilevamenti che hanno interessato sia singoli dispositivi (frigorifero, forno, televisore, etc.) sia l'aggregazione degli stessi; entrambi questi elementi sono strettamente correlati con le abitudini giornaliere degli occupanti delle abitazioni coinvolte nelle iniziative.

I diagrammi di carico che si possono tracciare dalle misurazioni dirette dipendono dalla frequenza con cui il campionamento viene svolto: tanto la frequenza è elevata, tanto più approfondita è la conoscenza della potenza nel tempo, poiché se ne possono cogliere meglio le variazioni dovute all'avviamento o alla cessazione dell'impiego di uno o più dispositivi.

Tuttavia, spesso gli insiemi di dati non sono completi, ossia, privi di intervalli temporali in cui le misure non sono presenti per qualsivoglia motivazione. In presenza di tali mancanze, è necessario fare delle previsioni sull'evoluzione della richiesta di potenza. Dunque, l'obiettivo di questo lavoro

di tesi è la ricerca di una metodologia che consenta la ricostruzione dei diagrammi di carico, in assenza di misure dirette, a partire dalla conoscenza dei valori della potenza rilevati e tenuto conto delle peculiarità dell'evoluzione temporale del carico. Infatti, le curve di carico sono influenzate dalla stagionalità dei consumi e dalla tipologia di applicazione monitorata; se si considera un aggregato di carichi, quale può essere un insieme di abitazioni, sicuramente il diagramma di carico ha un andamento ripetitivo, dovuto alla costanza delle abitudini energetiche di ogni utilizzatore. Generalmente, si assiste ad un incremento della potenza richiesta nelle ore dei pasti, poiché corrispondenti a intervalli temporali in cui il maggior numero delle persone si trova in casa e impiega degli elettrodomestici. In questo contesto, la previsione del valore della potenza richiesta in un determinato momento è più semplice rispetto a quanto si verificherebbe nel caso di un singolo dispositivo, il cui utilizzo è assolutamente casuale e il cui consumo dipende dalle condizioni di lavoro. Tali aspetti, sono stati approfonditi nel seguito, suddividendo il lavoro in differenti fasi, ognuna sviluppata nel tentativo di apportare dei miglioramenti ai risultati ottenuti al passo precedente.

1.1 BANCHE DATI COME FONTI DEI DATI NUMERICI DI INTERESSE

1.1.1 BANCHE DATI DISPONIBILI ON-LINE

Il primo passo per la costruzione di un modello di previsione è la ricerca di fonti dalle quali attingere i valori campionati della potenza elettrica. Nel web sono liberamente disponibili svariate banche dati, le quali forniscono le misure della potenza o, equivalentemente, della corrente e della tensione. Queste grandezze, rilevate con frequenze differenti, sono correlate all'impiego contemporaneo di più dispositivi (carico aggregato) o di uno solo di essi.

Le campagne sono state condotte in modi simili, sebbene impiegando sistemi di misura differenti: per ogni applicazione si rileva la grandezza di interesse, la si trasmette ad un sistema di gestione dati e le si salva su un server.

Tra le banche dati esistenti, tre sono risultate di più semplice consultazione: *PLAID (Plug Appliance Identification Dataset)* [1], *GREEND Electrical Energy Dataset* [2] e *REFIT Electrical Load Measurements Dataset* [3].

Riguardo la banca dati *PLAID*, la campagna è stata svolta in Pennsylvania (USA) e si è basata sul campionamento a 30 kHz della tensione e della corrente per svariate apparecchiature. Considerata l'elevata frequenza di misurazione, per ogni dispositivo si sono monitorati solo pochi secondi, così da valutare i transitori di avviamento e il raggiungimento del regime. Quindi, si ritiene che i dati resi disponibili siano utili per studi di disaggregazione dei carichi e di identificazione di inserzione o disconnessione di un apparecchio, sulla base delle oscillazioni che esso genera nell'andamento della potenza complessiva. Per tale motivazione questo dataset è stato escluso.

I valori di potenza attiva forniti dalla campagna *GREEND* sono stati rilevati in Italia e Austria, con una frequenza di campionamento di 1 Hz, in otto differenti abitazioni. Per ciascuna di queste, i dati sono stati organizzati in differenti file, ognuno relativo a una giornata differente; tale aspetto rende più complesso e più lento il processo di acquisizione degli stessi, tanto da prediligere l'ultima banca dati.

La *REFIT* contiene i campionamenti della potenza attiva richiesta in ventuno abitazioni, site nel Regno Unito, per un periodo di circa due anni, compresi tra il 2013 ed il 2015.

Un periodo di osservazione così lungo è reso possibile grazie alla minore frequenza di campionamento (uno ogni otto secondi); in tal modo, si monitorano i consumi di un'abitazione e il loro cambiamento stagionale, mentre i suoi occupanti conducono le loro normali attività domestiche. Per ogni casa, si sono considerati nove dispositivi cominciando da quelli con maggiore impatto energetico (dispositivi di raffreddamento o riscaldamento) fino agli apparecchi di minor rilevanza, quali i caricatori dei telefoni o le piastre per capelli. Questi ultimi, infatti, sono utilizzati in maniera non regolare, meno frequentemente e con un consumo minore rispetto ai primi.

Questa banca dati è stata scelta, principalmente, perché mette a disposizione i valori della potenza aggregata (assente nelle altre) organizzati in un unico file per la singola abitazione. Tale aspetto, da un lato, ne limita la fruizione richiedendo l'impiego di particolari programmi ma, dall'altro, la semplifica raccogliendo tutto il necessario in un unico documento.

Tra le abitazioni osservate, molte dispongono di pannelli solari per la produzione di parte della energia elettrica richiesta; tuttavia, in alcune di esse, tale presenza determina delle interferenze nella valutazione del carico aggregato, in quanto la potenza prodotta appare come un consumo aggiuntivo piuttosto che come una generazione locale. Per tale motivazione, da questa banca dati si è considerata l'abitazione 1, nella quale l'effetto dei pannelli solari è stato rimosso e per la quale sono stati misurati i carichi riportati in Tabella 1.1.1-1, oltre a quello totale.

Tabella 1.1.1-1 Apparecchiature monitorate nell'abitazione selezionata

Carichi abitazione 1 REFIT							
Frigo	Freezer	Pozzetto	Asciugatrice	Lavatrice	Lavastoviglie	Postazione PC	Stufa

Come si può leggere negli studi di Murray, L. e V. Stankovic [3], i dati forniti dalla campagna *REFIT* sono stati sottoposti a delle correzioni seguendo il principio del “riempimento in avanti”; quindi, nel caso di assenza di misure per un intervallo di tempo breve, commisurato alla frequenza adottata per il campionamento, si è assunta la potenza pari al valore dell'ultima misura disponibile; invece, per range di durata maggiore, la potenza è stata annullata.

Una seconda correzione si basa sulla saturazione della potenza quando essa eccede un determinato valore; in questa situazione, infatti, si ipotizza un errore nella misurazione e si è imposta la potenza pari al limite massimo ammesso.

1.1.2 LACUNE DELLA SERIE ANALIZZATA: IDENTIFICAZIONE MESE MIGLIORE

Le misure sono fornite in file “.csv” organizzate in *time-table* di cui si osserva una porzione in Figura 1.1.2-1. L'estrazione dei dati contenuti in questo file ha richiesto l'impiego del software *Matlab*, a causa dell'elevato numero di campionamenti e, di conseguenza, di righe che

compongono la tabella finale. Per ogni riga, corrispondente ad un preciso istante, sono fornite l'ora e la data della misurazione (prima colonna), l'unità temporale (seconda colonna), la potenza aggregata (terza colonna) e tutte le potenze dei singoli dispositivi monitorati (Tabella 1.1.1-1).

Il consumo di ogni applicazione viene rilevato da uno strumento di misura; ogni strumento fornisce un valore numerico in un certo istante che può essere differente da quello riportato nel file; tuttavia, per avere un unico documento di più semplice lettura, sulla *time-table* è dichiarato l'istante relativo al carico complessivo, ultimo ad essere rilevato. Ciò significa che tutte le altre misure sono fatte in un momento compreso tra un campionamento del carico aggregato e il successivo; tale intervallo è dichiarato essere di otto secondi.

	Time	1 Unix	2 Aggregate	3 Appliance1	4 Appliance2	5 Appliance3	6 Appliance4	7 Appliance5	8 Appliance6	9 Appliance7	10 Appliance8	11 Appliance9
1	2013-10-09 13:06:17	1.3813e+09	523	74	0	69	0	0	0	0	0	1
2	2013-10-09 13:06:31	1.3813e+09	526	75	0	69	0	0	0	0	0	1
3	2013-10-09 13:06:46	1.3813e+09	540	74	0	68	0	0	0	0	0	1
4	2013-10-09 13:07:01	1.3813e+09	532	74	0	68	0	0	0	0	0	1
5	2013-10-09 13:07:15	1.3813e+09	540	74	0	69	0	0	0	0	0	1
6	2013-10-09 13:07:18	1.3813e+09	539	74	0	69	0	0	0	0	0	1
7	2013-10-09 13:07:30	1.3813e+09	537	74	0	69	0	0	0	0	0	1
8	2013-10-09 13:07:32	1.3813e+09	537	74	0	69	0	0	0	0	0	1
9	2013-10-09 13:07:44	1.3813e+09	548	74	0	69	0	0	0	0	0	1
10	2013-10-09 13:07:47	1.3813e+09	557	73	0	68	0	0	0	0	0	1
11	2013-10-09 13:07:59	1.3813e+09	561	74	0	69	0	0	0	0	0	1
12	2013-10-09 13:08:01	1.3813e+09	561	74	0	69	0	0	0	0	0	1
13	2013-10-09 13:08:13	1.3813e+09	558	73	0	69	0	0	0	0	0	1
14	2013-10-09 13:08:16	1.3813e+09	558	73	0	69	0	0	0	0	0	1
15	2013-10-09 13:08:17	1.3813e+09	569	73	0	69	0	0	0	0	0	1
16	2013-10-09 13:08:27	1.3813e+09	561	73	0	69	0	0	0	0	0	1
17	2013-10-09 13:08:30	1.3813e+09	549	73	0	37	0	0	0	0	0	1
18	2013-10-09 13:08:32	1.3813e+09	549	73	0	37	0	0	0	0	0	1

Figura 1.1.2-1 Time table originale

Analizzando la colonna degli istanti di campionamento e considerando solo gli intervalli di durata minore o uguale agli otto secondi previsti, il passo temporale medio che intercorre tra un campionamento e il successivo è poco superiore ai cinque secondi; quindi, ci sono misure che vengono effettuate e registrate in un tempo inferiore a quello nominale. Tuttavia, è possibile accorgersi anche della presenza di alcune anomalie, ossia, periodi più o meno lunghi in cui non ci sono misure disponibili. In dettaglio, tali anomalie sono state divise in due gruppi, sulla base del tempo trascorso tra le varie misurazioni: al primo gruppo appartengono quelle con durata inferiore a sedici secondi, doppia rispetto a quella dichiarata, mentre nel secondo gruppo rientrano tutte le altre, di durata superiore.

Questa distinzione è stata necessaria in quanto gli istanti di registrazione dei dati non sono tutti equi-distanziati di otto secondi e, quindi, in alcuni casi, due misure adiacenti possono presentare una distanza tra loro superiore a tale passo ma compresa in un tempo totale di sedici secondi, come dovrebbe essere se esse fossero effettuate nei momenti previsti. Per questo motivo, le anomalie inferiori o pari a sedici secondi sono state trascurate.

Per comprendere meglio quanto illustrato, in Figura 1.1.2-2 è riportato l'andamento temporale del carico aggregato dell'abitazione in analisi mentre, in Figura 1.1.2-3, è stato tracciato il numero delle anomalie di durata superiore a sedici secondi, espresse in numero di giorni.

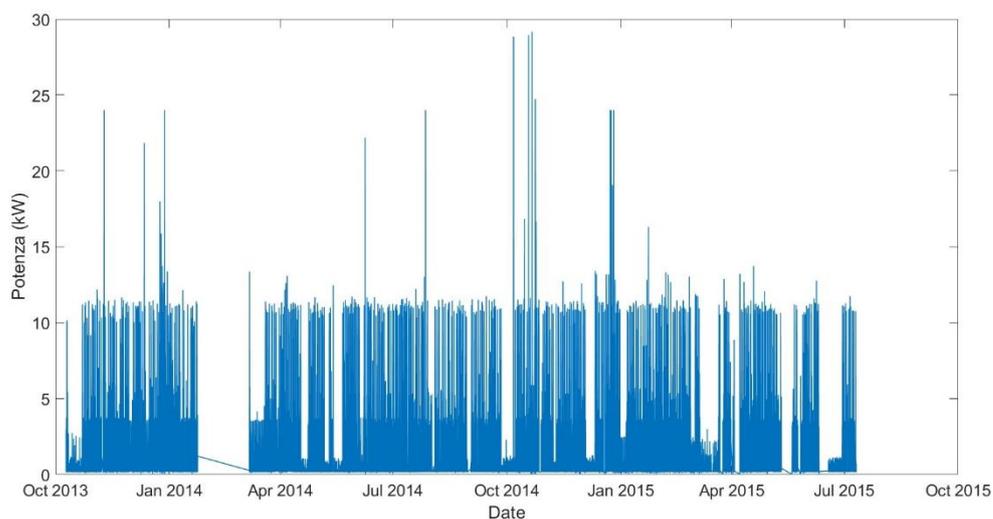


Figura 1.1.2-2 Carico aggregato abitazione 1 (REFIT)

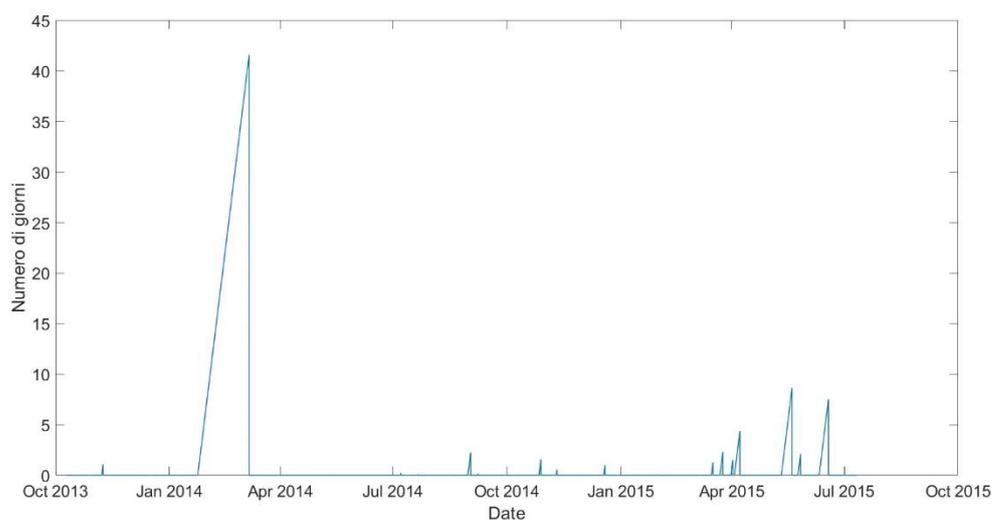


Figura 1.1.2-3 Numero di anomalie superiori ai 16 secondi per abitazione 1 (REFIT)

Osservando gli andamenti sopra riportati, si nota una mancanza di misure in corrispondenza del mese di febbraio 2014 per il quale, a causa della presenza di problematiche nella rete di comunicazione [3], non è disponibile alcun dato; tale lacuna, di circa quarantadue giorni, copre in parte anche i mesi di gennaio e marzo. Sono presenti numerose altre mancanze, molto meno evidenti e comunque di durata inferiore ai dieci giorni, soprattutto nei mesi di agosto, maggio e giugno del 2015.

Indipendentemente dalla loro durata, nei giorni in cui non ci sono misure non si conosce esattamente quale sia stato il consumo dell'abitazione; proprio per questo si vuole cercare un modello per la sua identificazione.

Tenuto conto che il tempo richiesto per lo svolgimento dei calcoli è strettamente correlato con l'ammontare dei dati da elaborare, al fine di snellire il processo di ricerca e validazione del modello, si è deciso di concentrare lo studio su un solo mese di misure; in particolare, si è scelto quello di gennaio 2015 poiché presenta il maggior numero di dati disponibili e, come conseguenza, un minor numero di anomalie.

L'andamento del carico aggregato nel mese di gennaio 2015 è riportato in Figura 1.1.2-4; esso presenta una evidente periodicità con un consumo di poco superiore ai 200 W nelle ore notturne che permette di distinguere chiaramente i trentuno giorni che compongono il mese in analisi, anche senza indicarli direttamente.

Il numero di misure mancanti in questo periodo si è ritenuto trascurabile; infatti, si sono identificate due lacune principali: la prima il sette gennaio, di circa 168 misure e la seconda, il ventinove del mese, di quasi 900 misure. Ipotizzando di avere un passo di campionamento esattamente pari a quanto dichiarato (otto secondi), in un giorno ci dovrebbero essere 10800 misure, numero di gran lunga superiore all'effettiva mancanza delle stesse in queste due giornate. Per tale motivo, queste anomalie sono state trascurate, così come tutte le altre presenti a causa della loro estensione ridotta.

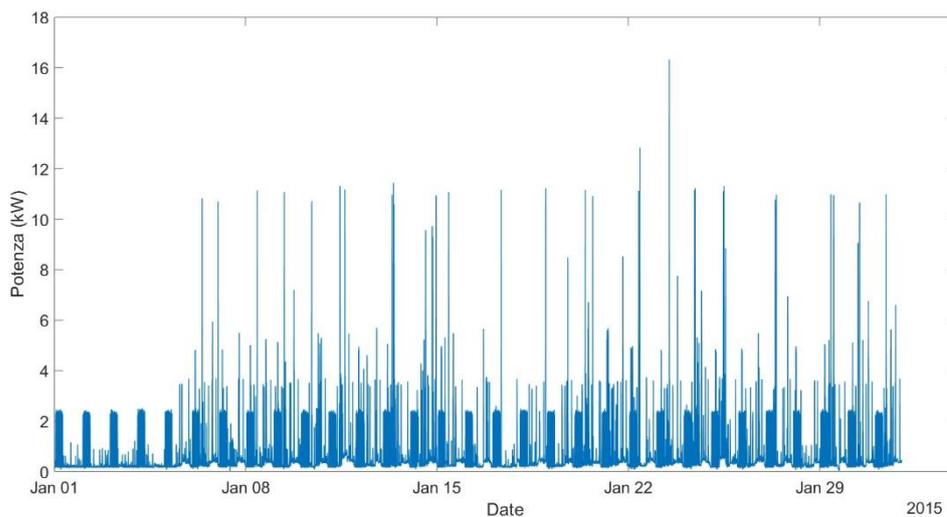


Figura 1.1.2-4 Carico aggregato nel mese di gennaio 2015

Definito l'intervallo da analizzare, nel seguito sono descritti i software e le tecniche che sono state impiegate.

1.2 SOFTWARE E TECNICHE PER LA REALIZZAZIONE DEL MODELLO

1.2.1 MACHINE LEARNING E RETI NEURALI

Il *Machine Learning* è la capacità di un software di fornire informazioni su un insieme di dati a partire da quelli disponibili, senza l'utilizzo di un modello o un'equazione predefinita. In particolare, in questo campo rientrano le tecniche di regressione, molto diffuse per la previsione dei carichi elettrici e, più in generale, di risposte continue nel tempo.

Tra esse le reti neurali sono ad oggi quelle più impiegate. Come suggerito dal loro stesso nome, esse sono nate dal tentativo di emulare il funzionamento del cervello umano, soprattutto per la risoluzione di problemi complessi con un elevato numero di dati. I 'neuroni' artificiali che lavorano in queste reti sviluppano connessioni l'un l'altro acquisendo informazioni dai dati forniti; in questo modo, essi guadagnano la possibilità di prevedere quello che potrebbe verificarsi come conseguenza ad un determinato input. Le reti neurali devono quindi essere addestrate in quanto, come i bambini, apprendono dall'esperienza per poi applicare quanto imparato in un secondo momento, in completa autonomia.

Il funzionamento di una generica rete neurale può essere descritto schematicamente come in Figura 1.2.1-1.

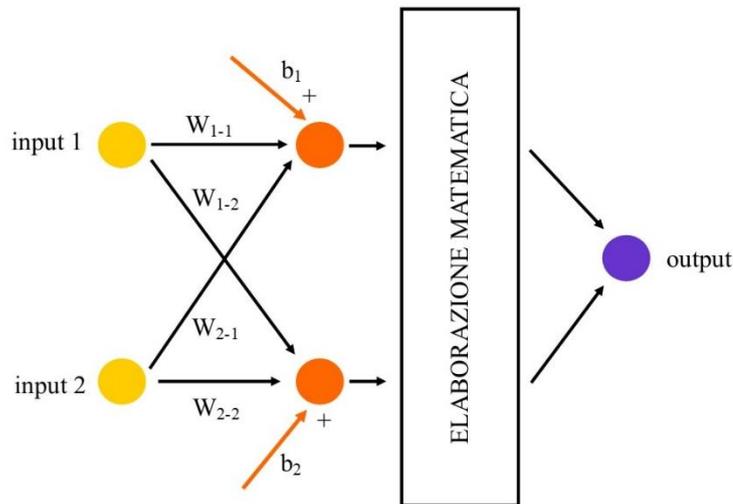


Figura 1.2.1-1 Schema semplificato di una rete neurale

Le reti neurali sono composte da strati di neuroni o “*layer*”, i quali interagiscono gli uni con gli altri, essendo ognuno di essi connesso con quello precedente.

Nel primo strato, “*input*” (giallo in Figura 1.2.1-1), vengono inserite le informazioni di partenza. Per il primo utilizzo, la rete deve essere addestrata, fornendo il valore della grandezza da stimare, ossia il carico aggregato dell’abitazione quando esso è noto, e dei parametri aggiuntivi. Questi ultimi, chiamati *predictor*, sono correlati alla grandezza analizzata e si conoscono anche quando questa non è nota, quindi, anche nell’intervallo temporale nel quale si vuole fare la previsione. Con tali informazioni, la rete viene sottoposta ad un processo iterativo di apprendimento, durante il quale ricerca delle correlazioni tra incognita e *predictor*.

Completato l’addestramento, la rete può essere impiegata per lo svolgimento della previsione; in questo caso, lo strato di input riceve come informazioni i soli *predictor* relativi agli istanti nei quali la potenza aggregata è incognita.

A valle del primo livello, è inserito un numero variabile di strati (in Figura 1.2.1-1 ne è stato rappresentato solo uno, in arancione, a titolo esemplificativo) che elaborano i dati ricevuti secondo

un preciso algoritmo: le informazioni vengono moltiplicate per i corrispondenti coefficienti ponderali (w_i) e sommate a fattori denominati "bias" (b_i).

In uscita da ciascuno dei livelli interposti tra il primo e l'ultimo, si applica uno strato per una ulteriore manipolazione dei dati prima di mandarli allo strato successivo. I processi attuati più comunemente sono l'azzeramento degli output negativi, l'applicazione della funzione tangente iperbolica o la compressione degli output tra $[-1;1]$. È altresì possibile che non avvenga nessuna elaborazione matematica, per cui il dato in uscita da uno strato viene direttamente applicato a quello successivo.

Il numero di informazioni scambiate tra uno strato e quello adiacente non è uguale per ogni strato, ma decresce avvicinandosi allo strato finale, ovvero quello di output. Quest'ultimo restituisce il valore della previsione.

Il problema principale di una qualunque rete neurale è correlato ai coefficienti ponderali e al valore dei bias da applicare in ogni strato; infatti, il loro valore iniziale è scelto in maniera del tutto casuale per poi essere affinato con il processo di addestramento: la rete modifica questi parametri valutando la differenza tra previsione e realtà, con l'intento di ridurre tale errore di valutazione; il processo di miglioramento si arresta quando viene raggiunto l'obiettivo o il numero massimo di iterazioni imposto. Per tale motivazione, come conseguenza della casualità nell'inizializzazione dei parametri sopra citati, l'addestramento ripetuto della medesima rete neurale conduce a risultati leggermente differenti ad ogni iterazione.

Il software *Matlab* è risultato di grande utilità nel lavoro di tesi condotto, poiché permette la realizzazione semplificata e guidata di reti neurali.

In dettaglio, sono stati sfruttati due applicativi del toolbox *Statistics and Machine Learning*, di seguito analizzati in maggiore dettaglio: *Regression Learner App* e *Neural Net Time Series*.

1.2.2 REGRESSION LEARNER APP [6]

Questo applicativo di *Matlab* fornisce cinque modelli di rete neurale pre-programmati, con complessità differente, ognuno caratterizzato da differenti parametri ottimizzabili al fine di ridurre l'errore della previsione che si vuole attuare. Le reti neurali disponibili nel software sono le seguenti cinque:

- *Narrow Neural Network*
- *Medium Neural Network*
- *Wide Neural Network*
- *Bilayered Neural Network*
- *Trilayered Neural Network*

Per ognuna di queste reti è possibile modificare alcune caratteristiche quali, per esempio, il numero di strati intermedi (con un massimo di tre, estendibile utilizzando l'applicazione *Deep Learning*), la loro estensione e il numero di iterazioni svolte.

La finestra di interfaccia si presenta come in Figura 1.2.2-1 e richiede tutte le informazioni necessarie per l'addestramento della rete neurale che si vuole costruire. Queste comprendono i *predictor*, dei quali è possibile selezionare anche solo una parte (riquadro viola), e la potenza aggregata in tutti i momenti del mese scelto (gennaio 2015), eccetto che per la giornata ipotizzata incognita. Quest'ultima viene selezionata in corrispondenza del riquadro ocra nella figura seguente, dal quale si seleziona la grandezza incognita tra quelle che compongono le colonne della matrice inserita come ingresso (riquadro verde). Infatti, tutti i dati vengono forniti in forma matriciale e, per quanto concerne i *predictor*, questi devono essere gli stessi che verranno utilizzati nella fase successiva di previsione.

Inserite tutte le informazioni richieste, si può selezionare il modello da analizzare e il software comincia a ricercare una correlazione tra le informazioni e costruire il modello di calcolo.

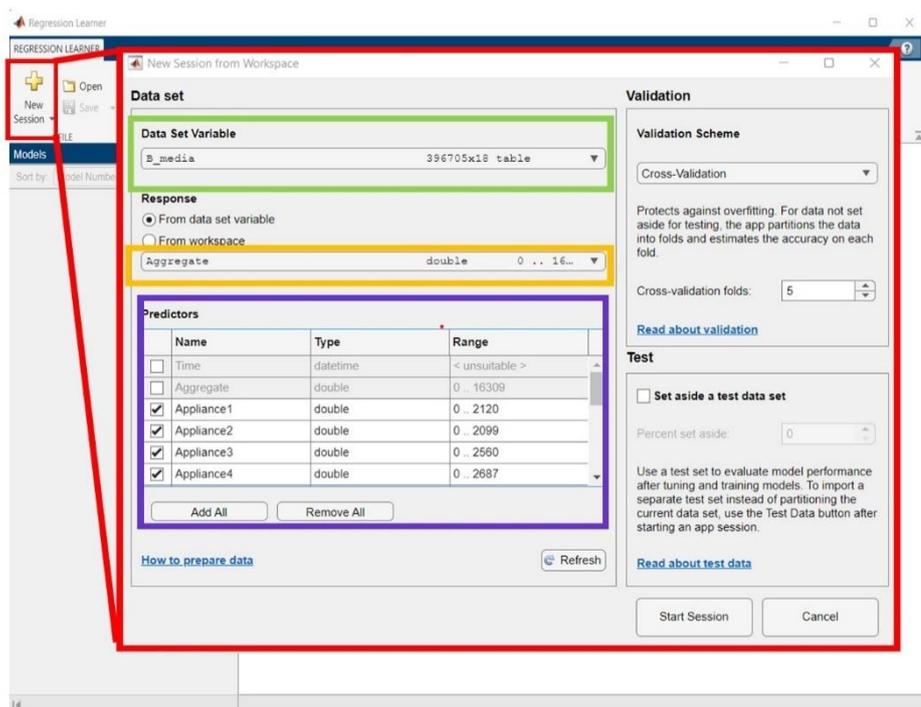


Figura 1.2.2-1 Regression Learner App

Completato l'addestramento, l'applicazione permette il confronto diretto tra i dati reali e quelli ricostruiti mediante la contemporanea visualizzazione nel medesimo piano. Questa analisi è molto importante poiché dà contezza dell'accuratezza che il modello è in grado di ottenere. Infatti, se il modello si applicasse direttamente ad un intervallo temporale in cui non ci sono misure disponibili, come per esempio febbraio 2014, non si avrebbe nessuna percezione dell'errore commesso e dell'accuratezza della ricostruzione, mancando i dati reali per il confronto. Inoltre, dal paragone diretto tra previsioni e misure, si possono modificare i parametri di ciascun modello al fine di affinarne le caratteristiche e di fare avvicinare quanto più è possibile le prime alle seconde.

Oltre alla rappresentazione grafica, l'ideatore del software indica come altro parametro di valutazione della bontà di un modello la radice dell'errore quadratico medio, nel seguito indicata con la sigla RMSE (*root-mean-square-error*): tanto minore è il suo valore, tanto migliore è ricostruzione della grandezza. L'idealità si ottiene quando l'errore quadratico medio è nullo, poiché

in questo caso si avrebbe la perfetta sovrapposizione della potenza prevista con quella realmente campionata. Matematicamente, questo parametro si definisce con la seguente espressione:

$$RMSE = \sqrt{\frac{\sum_1^n (x_{reale} - x_{previsione})^2}{n}}$$

dove:

- x_{reale} è il valore misurato della grandezza x ;
- $x_{previsione}$ è il valore della grandezza x stimato dal modello di previsione;
- n è il numero di previsioni effettuate e dei campioni impiegati.

L'errore $RMSE$ viene calcolato in maniera automatica dall'applicazione per ognuno dei modelli addestrati sul medesimo insieme di dati, evidenziando quale modello presenta il valore minore dell'errore. In questo modo si semplifica la scelta della rete neurale da adottare per le previsioni.

Scelta ed addestrata la rete, la si può esportare dall'applicazione al *workspace* di *Matlab*; in questa fase si è preferito salvare i differenti modelli addestrati in file separati in modo da poterli riutilizzare in un secondo momento, evitando la riapertura dell'applicativo. Disponibile il modello nell'ambiente di lavoro, lo si utilizza per la previsione con la funzione seguente:

$$X_{previsione} = \text{trainedModel.predictFcn}(X_{predictor})$$

dove:

- $X_{previsione}$ è il risultato elaborato dalla rete neurale;
- $X_{predictor}$ è la matrice dei *predictor* relativi all'intervallo temporale nel quale la grandezza X si vuole prevedere.

1.2.3 NEURAL NET TIME SERIES APP [8]

Questa applicazione è focalizzata sulla programmazione di una particolare famiglia di reti neurali, identificata come NARX (*Nonlinear Autoregressive Neural Networks*), ideali per l'analisi di serie temporali. In particolare, essa lavora a partire dalle evoluzioni passate della serie incognita e di una o più altre serie.

Il principio di funzionamento rimane invariato rispetto ai casi precedenti: inizialmente, si forniscono alla rete i *predictor* e la potenza aggregata negli istanti precedenti a quelli incogniti; a seguire, si inseriscono i *predictor* relativi ai momenti in cui si vuole fare la previsione e con i quali il modello restituisce il valore del calcolo svolto.

Le reti NARX lavorano seguendo un flusso unidirezionale delle informazioni, che comincia dai neuroni nello strato di input ed evolve verso lo strato di output, senza alcuna retroazione delle informazioni ottenute. Tuttavia, il software fornisce due possibili strutture e la spiegazione del loro operato: rete ad anello aperto (Figura 1.2.3-1) e rete ad anello chiuso (Figura 1.2.3-2).

Entrambe le strutture sono ottenute con due strati successivi allo strato di input: l'*hidden layer* e l'*output layer*. Nel primo strato (nascosto) agisce una funzione di ritardo (nei cerchi azzurri nelle figure) con la quale la rete immagazzina i valori passati delle serie temporali in ingresso: le $x(t)$ che rappresentano i *predictor* e la $y(t)$ che è la grandezza incognita. Il valore di questo ritardo è impostabile e da esso dipende in parte l'accuratezza della previsione.

A valle di tale funzione, l'*hidden layer* opera come uno dei generici strati intermedi costituenti una rete neurale, come spiegato nel paragrafo 1.2.1: gli ingressi vengono moltiplicati per i corrispondenti coefficienti ponderali per poi essere sommati ai *bias*. Come nei paragrafi precedenti, questi due parametri vengono inizializzati in maniera casuale, quindi l'addestramento della medesima rete con gli stessi dati di ingresso può condurre a risultati differenti tra loro, sebbene simili. Inoltre, il numero di neuroni componenti questo primo strato è selezionabile tra le opzioni della rete.

Il risultato della somma viene poi elaborato con una funzione sigmoide e fornito come input all'*output layer*. Questo strato si comporta esattamente come quello che lo precede, con l'unica differenza che il suo risultato viene mandato ad una funzione lineare.

Il valore finale ottenuto dalla sequenza di queste operazioni viene trasmesso allo strato di output, il quale elabora il risultato finale della previsione.

Come dichiarato dalla casa produttrice dell'applicazione, sebbene le due strutture siano simili, quella ad anello aperto è più accurata della seconda nella fase di addestramento, poiché gli input ricevuti sono più precisi. Infatti, osservando i quadratini verdi in Figura 1.2.3-1, l'output vero viene impiegato come ingresso (essendo noto il suo valore passato). Al contrario, nella struttura ad anello chiuso l'input relativo alla grandezza incognita è ottenuto come retroazione del valore stimato, del quale non si conosce l'esattezza. In altre parole, nell'anello chiuso, è la previsione stessa ad essere usata come ingresso per il passo successivo. Per tali motivazioni, si è proceduto costruendo in un primo momento la rete NARX ad anello aperto per l'addestramento e, in un secondo momento, la si è chiusa per procedere con la previsione nell'intervallo incognito.

I dati richiesti per il corretto funzionamento di queste reti neurali sono i medesimi di quelli impiegati per i modelli ottenuti con la *Regression Learner App*.

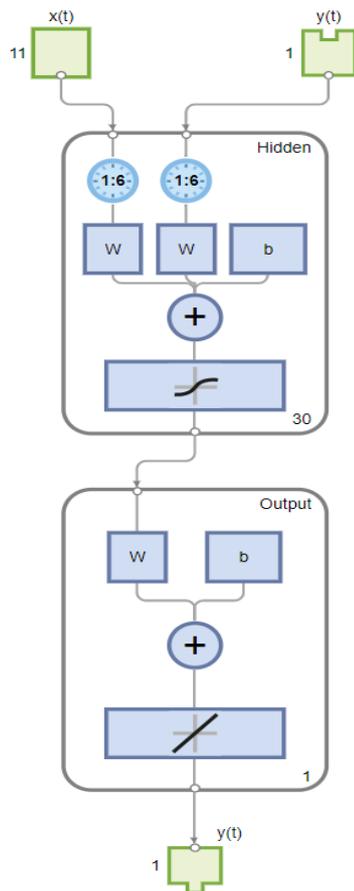


Figura 1.2.3-1 NARX anello aperto

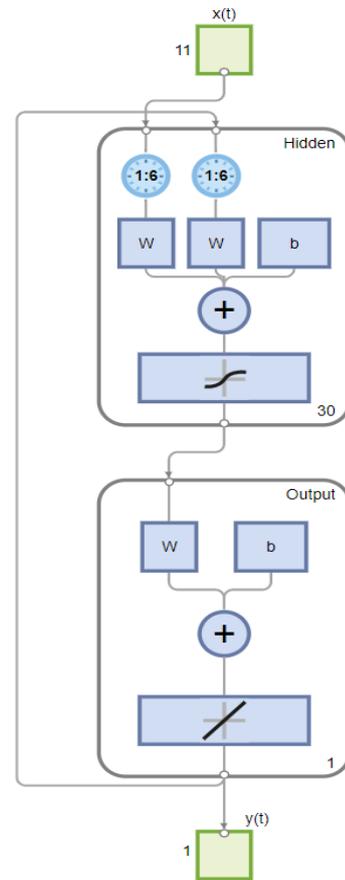


Figura 1.2.3-2 NARX anello chiuso

2. ANALISI PRELIMINARE DEI LIMITI DELLE RETI NEURALI

2.1 PREPARAZIONE DEI DATI

2.1.1 PREDICTOR TEMPORALI E RMSE DELLE RETI PRE-PROGRAMMATE

Poiché si vuole trovare una legge per la previsione del carico in funzione del tempo, in un primo momento si sono utilizzati solo *predictor* “temporali”, ossia, strettamente correlati con gli istanti in cui le misure sono state effettuate. In maggiore dettaglio, dai file forniti è possibile estrarre, per ogni misurazione, le seguenti informazioni:

- Ora
- Giorno della settimana
- Mese
- Anno
- Se il giorno indicato è feriale o festivo.

Questi *predictor* sono calcolati per tutti i momenti del mese di gennaio 2015, ma solo una parte viene impiegata nella fase di addestramento, inserendoli come colonne aggiuntive nella matrice principale di partenza insieme alla potenza campionata. In questo modo, ad ogni valore di carico aggregato viene associato questo insieme di parametri che, noti in un qualunque altro momento, permettono la valutazione dell’incognita. In particolare, i *predictor* correlati ai momenti in cui si vuole fare la previsione vengono rimossi dalla matrice principale e sono inseriti in una matrice differente, nella quale non sarà presente la grandezza in analisi.

In un primo momento si è fatto uso delle reti neurali pre-programmate presenti nell’applicazione *Regression Learner App*. Con l’intento di evidenziare i loro limiti, non sono stati modificati i loro parametri caratterizzanti, ma sono state impiegate direttamente per costruire cinque differenti modelli di previsione, ognuno basato su un numero di giorni diverso: ad un giorno, due giorni, quattro giorni, sei giorni ed un mese. Considerare intervalli minori per l’addestramento della rete

sarebbe troppo riduttivo: per prevedere l'andamento temporale di una grandezza con la migliore accuratezza possibile è bene studiarne gli andamenti per periodi lunghi, così da poterne osservare le variazioni ed eventuali periodicità nell'arco delle ventiquattro ore e dell'anno.

Si è inoltre posta l'attenzione su una eventuale dipendenza diretta tra la precisione del modello costruito e la frequenza di campionamento della potenza complessiva. Quindi, come mostrato nella illustrazione esemplificativa di Figura 2.1.1-1, dall'insieme iniziale (in viola) si sono esclusi alcuni campionamenti, immaginando di avere a disposizione delle misurazioni ogni 16 secondi (in verde) ed ogni 24 secondi (in arancione), ovvero, rispettivamente con la metà ed un terzo della frequenza iniziale.

In questo modo, per ognuno degli intervalli temporali utilizzati per la costruzione dei modelli di previsione sono state valutate tre possibili frequenze, per un totale di 15 modelli.

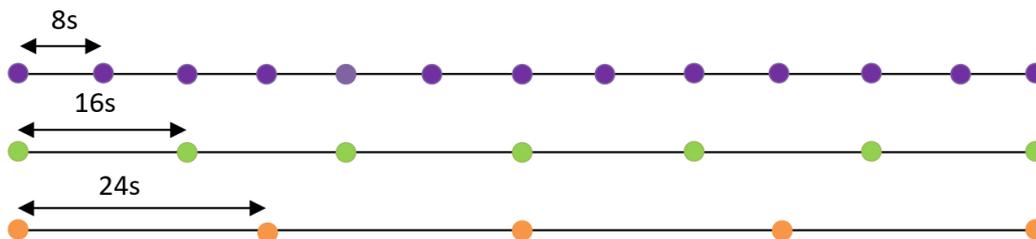


Figura 2.1.1-1 Filtro delle misurazioni per ottenere differenti frequenze

A causa della elevata quantità di dati, appare chiaro che un confronto grafico dei risultati non sarebbe attuabile né tanto meno affidabile, perché non fornirebbe alcuna informazione sull'inesattezza della previsione.

Tuttavia, si possono utilizzare i valori di *RMSE* che l'applicazione calcola in maniera automatica; questi sono stati analizzati, al fine di fare una selezione tra i quindici modelli di cui sopra. A tale scopo, sono state utilizzate tre tabelle, una per ogni frequenza di campionamento ipotizzata (Tabella 2.1.1-1, Tabella 2.1.1-2 e Tabella 2.1.1-3). Nelle righe di ciascuna tabella, il valore del parametro viene specificato per ciascun intervallo temporale e per ogni tipologia di rete neurale (indicata in tabella con la propria iniziale) impiegata per la costruzione del modello. Per facilitare la

lettura, si sono evidenziati in verde i valori minori di *RMSE* a parità di intervallo temporale impiegato.

Tabella 2.1.1-1 Valori di *RMSE* per differenti modelli con campionamento ad 8 secondi

	Campionamento 8 secondi				
	N	M	W	B	T
Modello ad 1 giorno	0.706	0.702	0.704	0.699	0.701
Modello a 2 giorni	0.771	0.752	0.748	0.762	0.754
Modello a 4 giorni	0.794	0.784	0.769	0.779	0.782
Modello a 6 giorni	0.874	0.860	0.848	0.862	0.860
Modello ad 1 mese	0.821	0.810	0.792	0.812	0.806

Tabella 2.1.1-2 Valori di *RMSE* per differenti modelli con campionamento ad 16 secondi

	Campionamento 16 secondi				
	N	M	W	B	T
Modello ad 1 giorno	0.697	0.694	0.689	0.698	0.691
Modello a 2 giorni	0.766	0.756	0.751	0.761	0.756
Modello a 4 giorni	0.796	0.781	0.771	0.783	0.778
Modello a 6 giorni	0.875	0.864	0.852	0.864	0.862
Modello ad 1 mese	0.825	0.821	0.816	0.821	0.819

Tabella 2.1.1-3 Valori di *RMSE* per differenti modelli con campionamento ad 24 secondi

	Campionamento 24 secondi				
	N	M	W	B	T
Modello ad 1 giorno	0.722	0.711	0.710	0.708	0.712
Modello a 2 giorni	0.772	0.754	0.749	0.755	0.753
Modello a 4 giorni	0.793	0.782	0.771	0.784	0.776
Modello a 6 giorni	0.876	0.863	0.849	0.870	0.857
Modello ad 1 mese	0.824	0.821	0.816	0.824	0.820

Analizzando i dati forniti in queste tabelle, si osserva che con il cambiamento della frequenza di campionamento l'errore quadratico medio cambia in maniera poco sensibile, a parità di intervallo temporale e modello.

Tuttavia, a parità di rete neurale e di frequenza di campionamento, si ha un incremento del medesimo parametro con l'aumentare dell'estensione dell'intervallo temporale, ad esclusione dei modelli ad un mese. Altresì, la rete neurale di tipo *wide* è quella che presenta l'errore minore a parità di altri parametri nella maggioranza dei casi, con l'esclusione dei modelli ad un giorno per i quali altre tipologie di reti sono più accurate.

In ogni caso, si constata che il parametro di valutazione calcolato rimane a valori relativamente elevati, più vicini all'unità che allo zero. Conseguentemente, la potenza aggregata ricostruita è ben differente da quella realmente misurata, come spiegato nel seguito.

2.2 RISULTATI ANALISI PRELIMINARE

2.2.1 MODELLI COSTRUITI

Selezionando i modelli con il minor valore di errore quadratico medio, in un primo momento si è proceduto ad applicarli a intervalli temporali di differente lunghezza, per simulare la previsione di un quantitativo crescente di possibili misure mancanti. Per gli intervalli di durata ridotta, si è preferito scegliere gli intervalli temporali nelle ore notturne poiché, come evidenziato in Figura 1.1.2-4, in esse si è potuta osservare una elevata periodicità e quindi si presume una più semplice valutazione della potenza aggregata. In particolare, gli intervalli temporali trattati sono indicati in Tabella 2.2.1-1.

Tabella 2.2.1-1 Intervalli temporali di misure mancanti impiegate per il test dei modelli

RANGE	INIZIO	FINE
10 minuti	22 gennaio 2015, 04:50:00	22 gennaio 2015, 05:00:00
30 minuti	22 gennaio 2015, 04:30:00	22 gennaio 2015, 05:00:00
1 ora	22 gennaio 2015, 04:00:00	22 gennaio 2015, 05:00:00
Mattina	22 gennaio 2015, 00:00:00	22 gennaio 2015, 12:00:00
Pomeriggio	22 gennaio 2015, 12:00:00	23 gennaio 2015, 00:00:00
1 giorno	22 gennaio 2015	
2 giorni	22 gennaio 2015	24 gennaio 2015
4 giorni	22 gennaio 2015	26 gennaio 2015
6 giorni	22 gennaio 2015	28 gennaio 2015
1 mese	gennaio 2015	

2.2.2 CONFRONTO DEI RISULTATI

Successivamente, si è confrontato il risultato ottenuto dal modello di previsione con il valore reale della potenza aggregata, basandosi sui differenti aspetti:

- valore dell'errore istantaneo, come differenza tra valore misurato e valore stimato;
- errore complessivo, come somma di tutti gli errori commessi in ogni istante;
- radice quadrata dell'errore quadratico medio;
- rappresentazione grafica sul medesimo piano delle due evoluzioni temporali.

I risultati restituiti dal calcolatore sono imprecisi ed includenti, come ci si aspettava dalle previsioni fornite nell'applicativo per la costruzione dei modelli. Per fare un confronto si è evitato di riportare tutti gli andamenti temporali, poiché richiederebbero un numero copioso di pagine senza apportare alcuna reale informazione; infatti, essendo l'imprecisione molto elevata in qualunque modello e intervallo temporale di applicazione, tali grafici sono tra essi molto simili.

Piuttosto, si è preferito effettuare un confronto a partire dai valori di *RMSE* e, solo in un secondo momento, osservare alcuni andamenti temporali esemplificativi dai quali è più semplice constatare l'errore della previsione. I valori ottenuti per questo parametro sono stati raggruppati nei diagrammi a dispersione mostrati in Figura 2.2.2-1, Figura 2.2.2-2 e Figura 2.2.2-3.

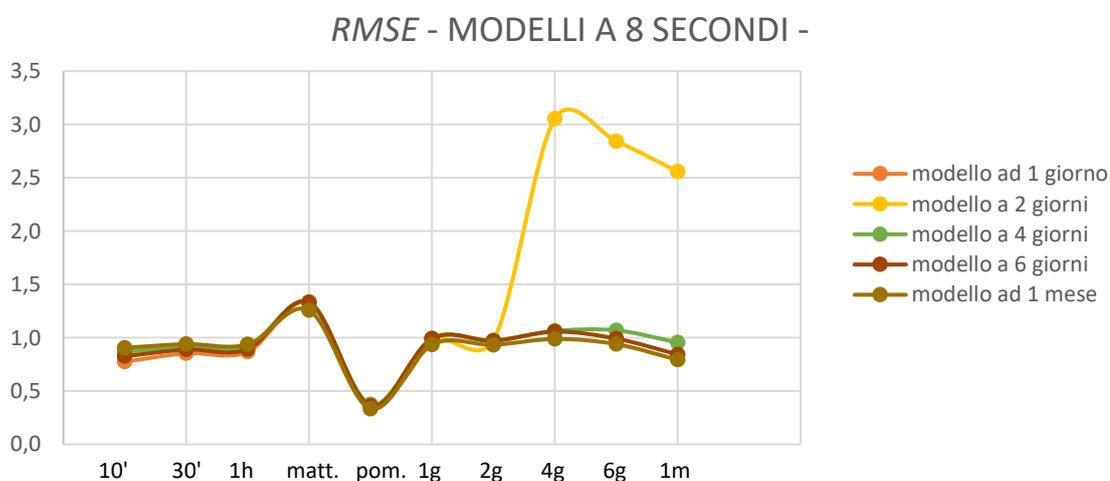


Figura 2.2.2-1 *RMSE* per i modelli costruiti con campioni a 8 secondi

RMSE - MODELLI A 16 SECONDI -

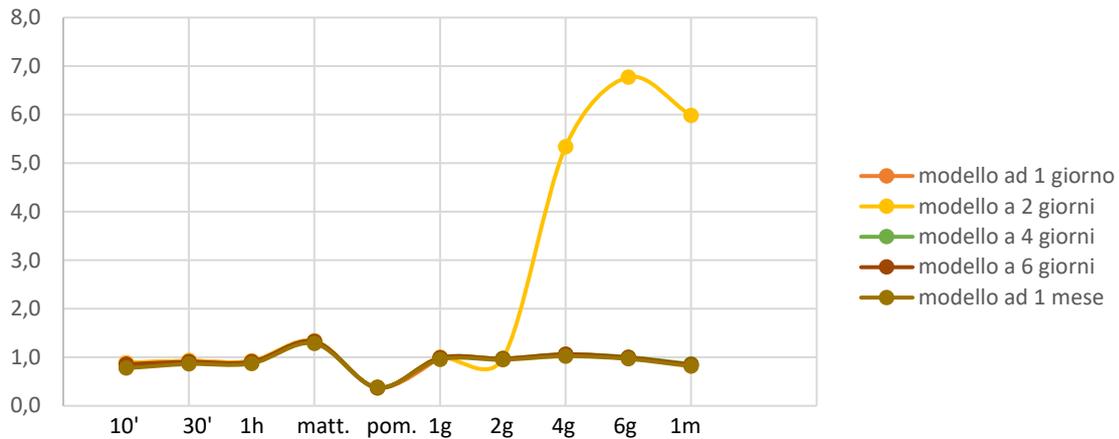


Figura 2.2.2-2 RMSE per i modelli costruiti con campioni a 16 secondi

RMSE - MODELLI A 24 SECONDI -

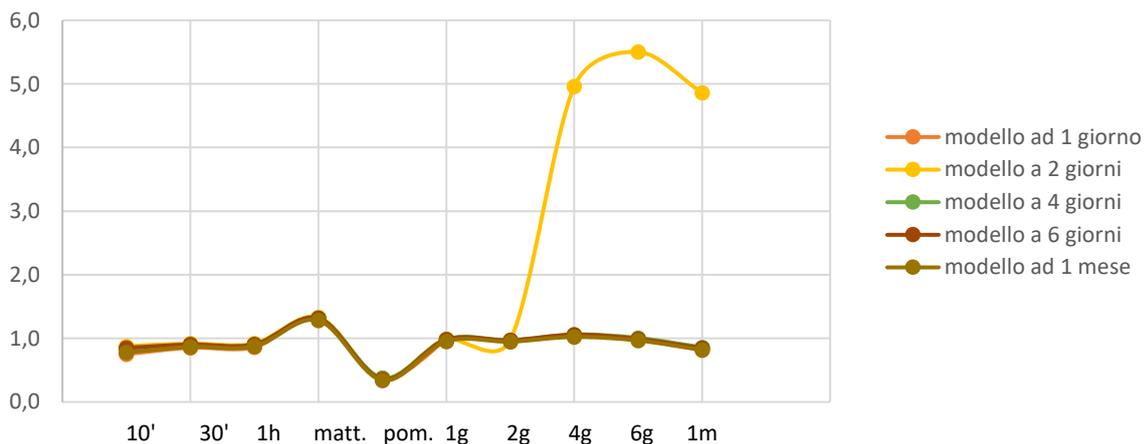


Figura 2.2.2-3 RMSE per i modelli costruiti con campioni a 24 secondi

In ogni diagramma, l'asse orizzontale è rappresentativo degli intervalli temporali impiegati per le simulazioni, mentre quello verticale identifica il valore del parametro analizzato; ad ogni modello viene inoltre associato un colore differente.

Tuttavia, non è possibile distinguere le linee associate ad ognuno di essi, se non in pochi casi, poiché le linee sono praticamente sovrapposte a causa delle differenze veramente ridotte tra i valori dell'*RMSE*. In particolare, i modelli a due giorni presentano un notevole incremento nel valore dell'errore prodotto dalla previsione quando il range temporale di applicazione supera i due giorni stessi; questo comportamento è una conseguenza del fatto che si cerca di prevedere il

comportamento della potenza aggregata in istanti nei quali la rete neurale non ha conoscenza derivante da misurazioni precedenti; come conseguenza, l'errore diverge. La stessa conclusione è valida anche per i modelli ad un giorno, per i quali la rappresentazione dell'*RMSE* è stata troncata all'intervallo di un giorno al fine di evitare l'inserimento nei grafici di punti di ordine di grandezza molto diverso dagli altri, determinando un accavallamento delle altre curve.

Tralasciando queste due particolari situazioni, relativamente ai valori assunti dal parametro analizzato, essi oscillano tra 0.33 e 1.25, toccando il massimo nella previsione delle dodici ore diurne e il minimo nel caso delle dodici ore pomeridiane. Tuttavia, questi valori estremi sono da valutarsi con una particolare attenzione perché, se da un lato è vero che uno è molto alto e l'altro è molto basso, dall'altro canto le ricostruzioni ottenute sono errate in entrambi i casi. Per approfondire tale aspetto e tenuto conto che si giunge a forme simili con l'impiego di uno qualunque dei modelli sopra citati, a titolo esemplificativo in Figura 2.2.2-4 sono stati tracciati gli andamenti temporali della potenza reale e prevista ottenuta con l'applicazione alle 12 ore diurne del 22 gennaio del modello a 6 giorni, costruito con i campionamenti ad 8 secondi; invece, in Figura 2.2.2-5, sono riportate le medesime grandezze calcolate con lo stesso modello ma nelle 12 ore pomeridiane.

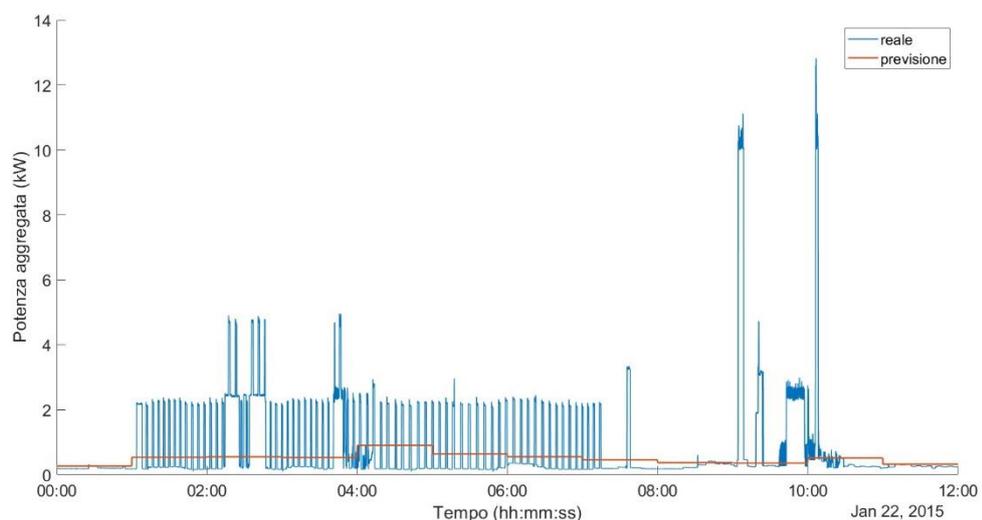


Figura 2.2.2-4 Previsione delle 12 ore diurne con modello a 6 giorni e $T_{\text{campionamento}}=8s$

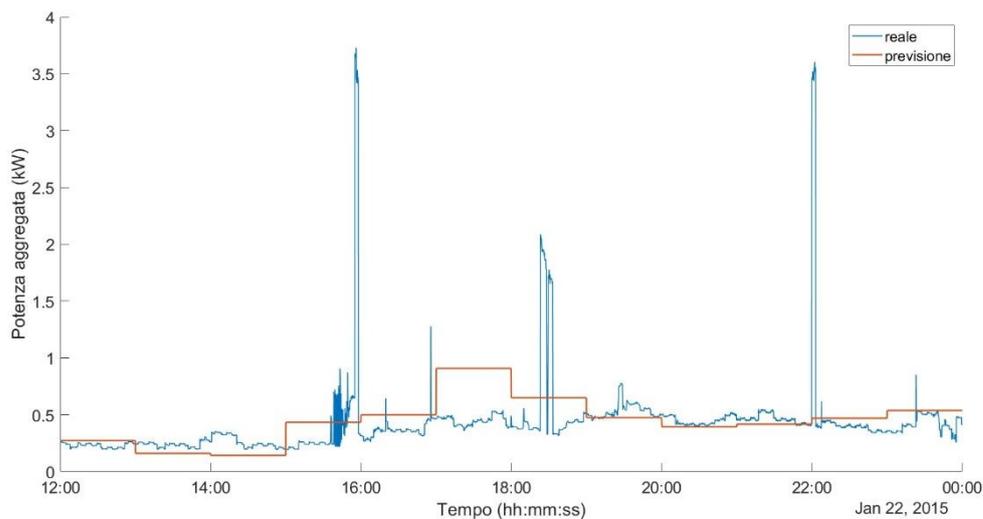


Figura 2.2.2-5 Previsione delle 12 ore pomeridiane con modello a 6 giorni e $T_{\text{campionamento}}=8s$

Nella previsione per l'intervallo tra le ore 12:00 e le ore 24:00, la potenza prevista è numericamente più vicina a quella reale, essendo quest'ultima non molto elevata nella maggior parte degli istanti, con dei picchi che non superano i 4 kW. Al contrario, per le ore comprese tra le 00:00 e le 12:00, la potenza misurata presenta dei picchi molto grandi, anche superiori ai 12 kW, mentre la previsione si mantiene a valori di molto inferiori. Questi picchi sono molto influenti e conducono ad ottenere errori numericamente molto differenti nei due casi e quindi due *RMSE* distanti tra loro, sebbene la ricostruzione non sia affidabile in entrambi i casi.

Da queste stesse figure, si osservano potenze previste con una caratteristica forma d'onda a tratti orizzontali, la quale non si adatta all'evoluzione molto variabile della grandezza reale. Quest'ultima, infatti, è correlata ai dispositivi impiegati durante la giornata all'interno dell'abitazione e, dunque, si evolve molto rapidamente ed in maniera del tutto casuale.

Alle stesse considerazioni si arriva studiando gli andamenti cui si è giunti impiegando diversi modelli per intervalli temporali di previsione variabili, ricordando nuovamente che a parità di intervallo di previsione e/o con differente numero di campioni disponibili, le forme d'onda sono simili passando da un modello all'altro, come dimostra la vicinanza dei valori di *RMSE* sopra ottenuti.

A tal proposito si sono riportati di seguito alcuni esempi.

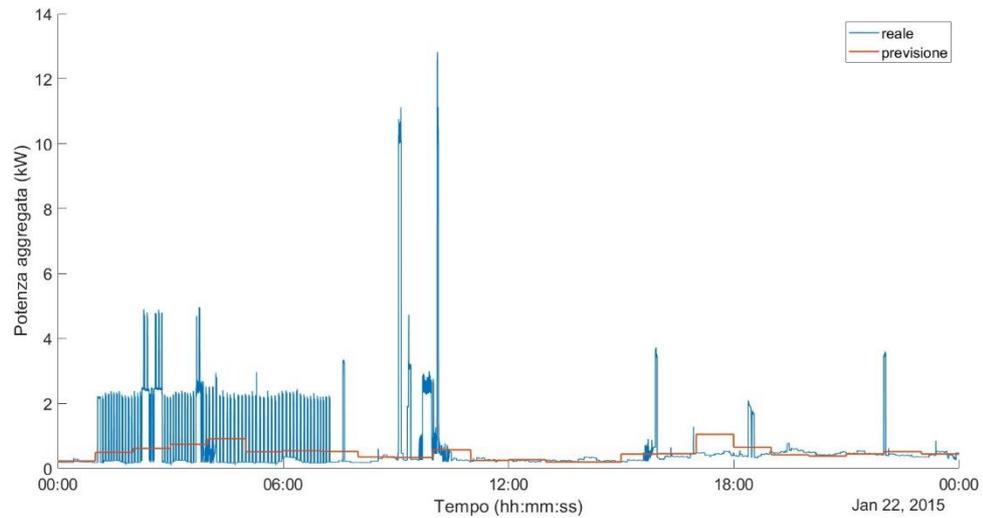


Figura 2.2.2-6 Previsione di un giorno con modello a 4 giorni e $T_{\text{campionamento}}=16s$

In Figura 2.2.2-6 è riportato il risultato della previsione della giornata del 22 gennaio 2015 con l'ausilio di un modello costruito sui quattro giorni antecedenti lo stesso: la ricostruzione in arancione della potenza aggregata non rispecchia la linea blu, che è ottenuta dai valori reali della stessa, ipotizzati ignoti. Lo stesso si può affermare per la Figura 2.2.2-7, rappresentante le medesime grandezze ottenute dal tentativo di prevedere gli avvenimenti di 48 ore con un modello costruito sul medesimo numero di informazioni.

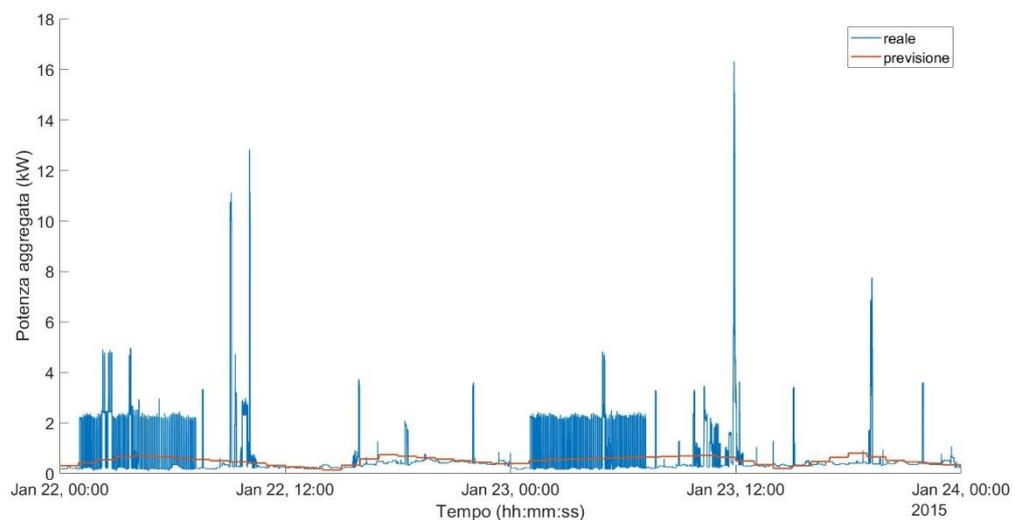


Figura 2.2.2-7 Previsione di 2 giorni con modello a 2 giorni e $T_{\text{campionamento}}=24 s$

Purtroppo, anche l'aumento del numero di informazioni fornite alla rete neurale non conduce a miglioramenti, come si deduce dall'analisi di Figura 2.2.2-8 e Figura 2.2.2-9, ove vengono rappresentate la potenza reale e prevista calcolata con un modello costruito con un insieme di dati relativo ad un numero di giorni maggiore a quelli da prevedere. In particolare, neanche il modello costruito sull'intero mese di dati è risultato essere efficace, sebbene in parte presenti gli stessi *predictor* dei sei giorni ipotizzati incogniti. Infatti, nell'addestramento sono stati rimossi solo i *predictor* del 22 gennaio e non quelli dei giorni ad esso adiacenti.

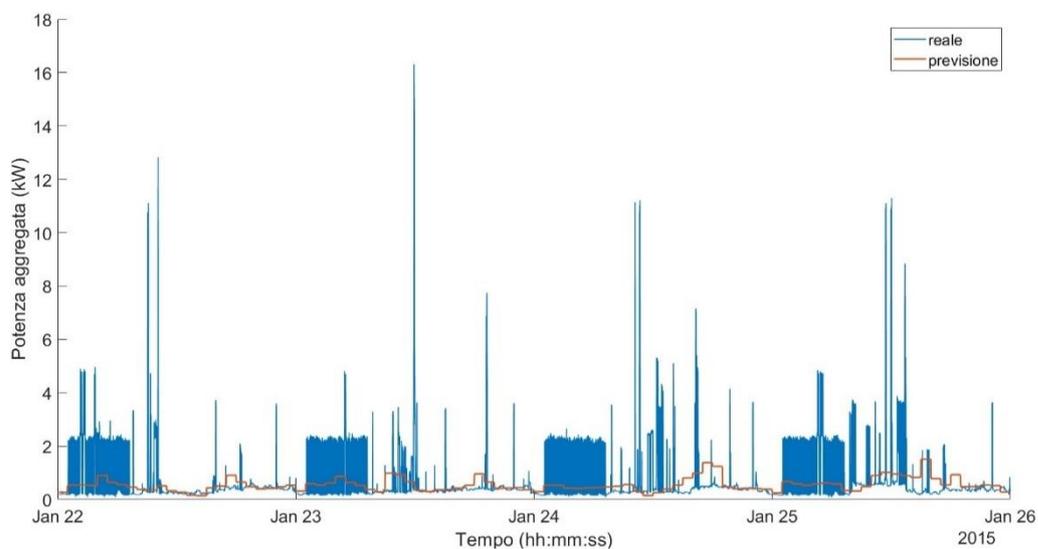


Figura 2.2.2-8 Previsione di 4 giorni con modello a 6 giorni e $T_{\text{campionamento}}=8 \text{ s}$

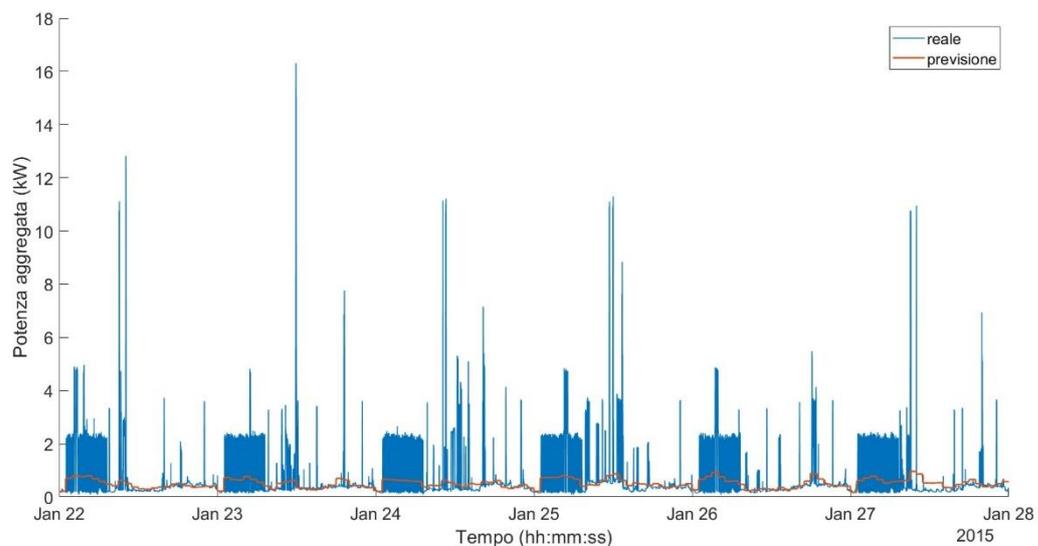


Figura 2.2.2-9 Previsione di 6 giorni con modello ad 1 mese e $T_{\text{campionamento}}=16 \text{ s}$

Questo conduce a pensare che i *predictor* impiegati non siano sufficienti o non diano il corretto apporto di informazioni: se nell'addestramento sono usati dei *predictor* che poi ricompaiono invariati nella previsione e, nonostante questo, la ricostruzione della potenza risulta scorretta, significa che la rete non ha imparato a sufficienza.

In aggiunta, sono state volutamente inserite le due seguenti immagini: la Figura 2.2.2-10 e la Figura 2.2.2-11.

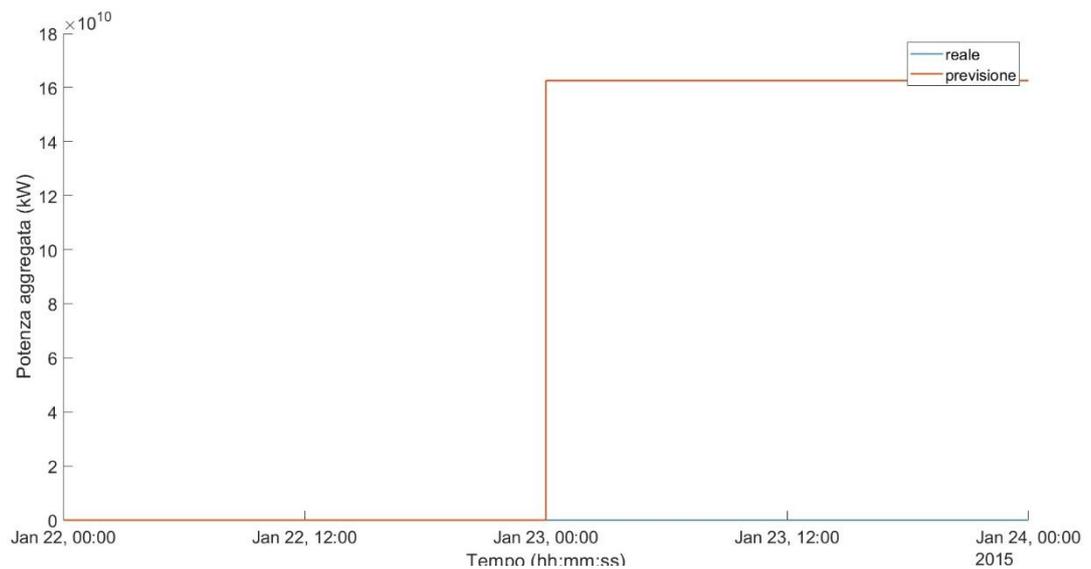


Figura 2.2.2-10 Previsione di 2 giorni con modello ad 1 giorno e $T_{\text{campionamento}}=24$ s

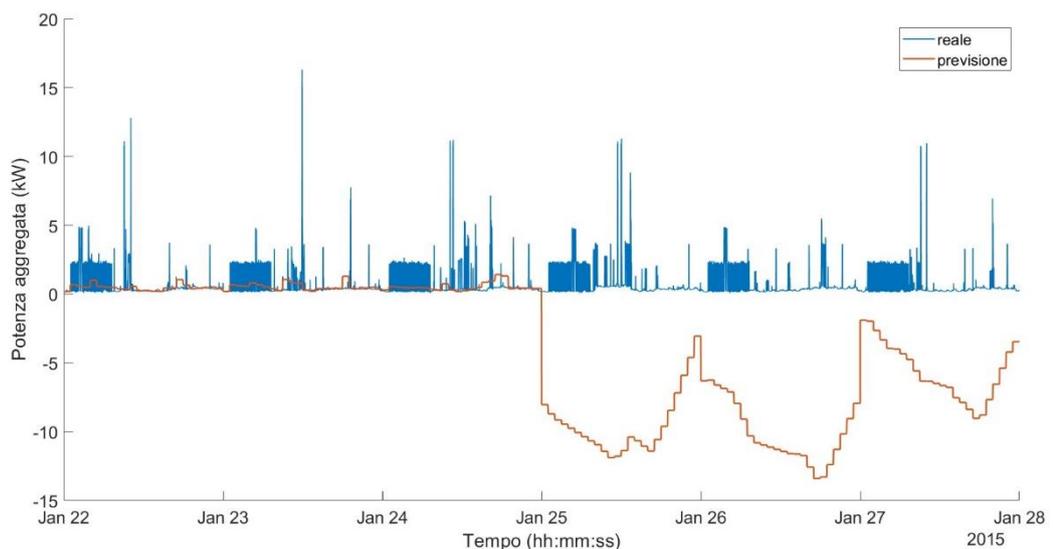


Figura 2.2.2-11 Previsione di 6 giorni con modello a 2 giorni e $T_{\text{campionamento}}=16$ s

Queste figure riportano i risultati ottenuti dei modelli ad un giorno (Figura 2.2.2-10) ed a due giorni (Figura 2.2.2-11), entrambi riferiti a due intervalli temporali nei quali l' $RMSE$ assume valori molto elevati. Nel primo caso la potenza aggregata campionata è talmente lontana da quella calcolata che la prima viene sovrastata dalla seconda, assumendo un andamento apparentemente orizzontale e costante; nel secondo caso, la previsione si sviluppa in buona parte nel semipiano negativo, in punti distanti da quelli reali e con un segno che non ha significato fisico, poiché un carico non può produrre potenza ma solo consumarla.

Il problema è nuovamente correlato alle informazioni disponibili; tuttavia, rispetto alle casistiche precedenti, nelle quali i dati non sono adeguati allo scopo preposto, in queste ultime due situazioni il problema è la mancanza stessa di dati, accompagnata dalla pretesa di voler prevedere qualcosa che si sviluppa su intervalli temporali molto più ampi rispetto a quelli forniti al software per la costruzione delle reti neurali, nei quali le potenze sono note. Dunque, l'errore è legato all'intervallo temporale utilizzato per la realizzazione del modello, ma anche a quello di applicazione.

Infine, per quanto concerne la frequenza di campionamento, sembrerebbe che la sua variazione non sia influente sul risultato della previsione, poiché i valori di $RMSE$ sono praticamente invariati nel passaggio da un numero di campioni all'altro. Questo appare chiaro se si uniscono nel medesimo piano (Figura 2.2.2-12) le curve tracciate in Figura 2.2.2-1, Figura 2.2.2-2 e Figura 2.2.2-3.

Le curve alle differenti frequenze sono tra esse sovrapposte in tutti i casi, eccetto per quelle dei modelli a due giorni, a causa delle motivazioni sopra esposte. Tuttavia, il comportamento "anomalo" di queste tre curve consente di mettere in luce delle caratteristiche, mascherate negli altri casi da un errore più o meno costante.

In primo luogo, l'errore è legato all'intervallo temporale utilizzato per la realizzazione del modello ma anche a quello di applicazione, a conferma di quanto visto precedentemente. Infatti, il modello a due giorni produce un errore di previsione molto grande quando utilizzato per intervalli temporali ampi, ma si riconduce al comportamento degli altri modelli quando tale intervallo si riduce.

In secondo luogo, l'errore è correlato alla frequenza di campionamento, assumendo il valore minimo quando essa è maggiore. In tale condizione, infatti, le informazioni disponibili alla rete neurale sono più numerose e dettagliate rispetto a quelle che si avrebbero con un numero minore di campioni, dove alcune caratteristiche e variazioni della potenza aggregata vanno perse a causa della mancanza delle misure stesse.

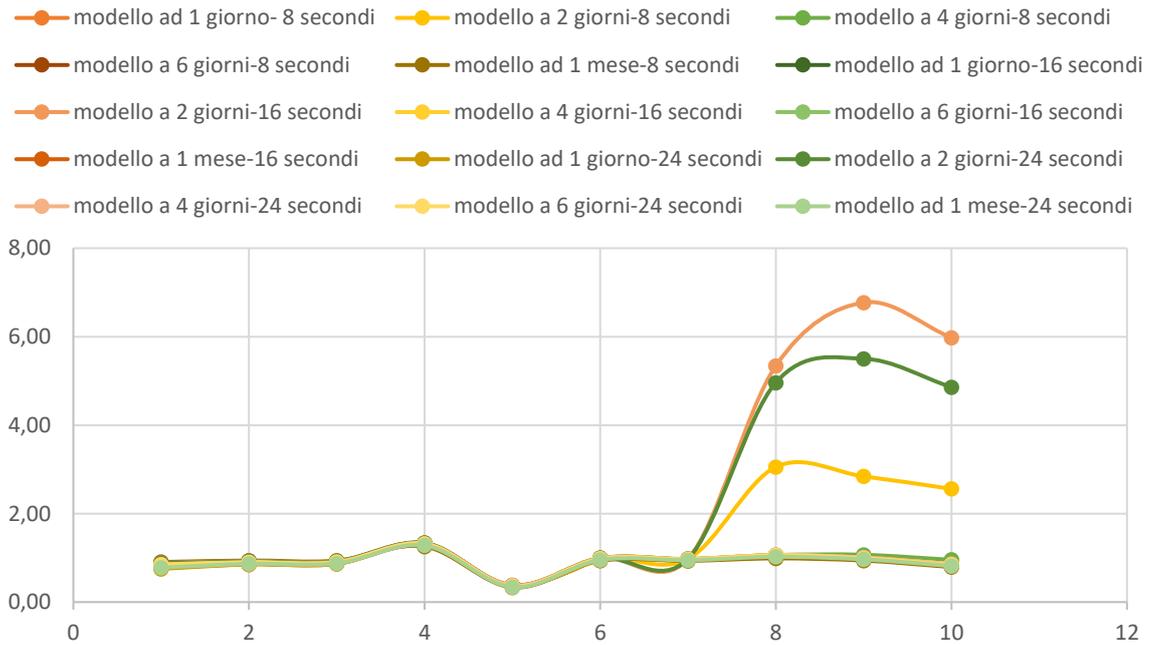


Figura 2.2.2-12 Confronto RMSE ottenuti per tutti i modelli

In conclusione, è necessario apportare delle modifiche per migliorare la qualità e la quantità di informazioni e la loro gestione da parte della rete neurale.

3. MIGLIORAMENTO DEI MODELLI

3.1 VARIAZIONE DELLE CARATTERISTICHE DELLE RETI NEURALI

Le previsioni delle reti neurali sono influenzate dai *predictor* impiegati e dal numero di neuroni componenti ogni strato. Per tali ragioni, si è proceduto a modificare separatamente ognuno di questi due elementi.

3.1.1 INCREMENTO DEL NUMERO DI PREDICTOR

Nel precedente capitolo si sono utilizzati soli cinque *predictor*, ricavati dagli istanti temporali in cui le misure sono state condotte. Tuttavia, essi risultano essere riduttivi, poco correlati in maniera diretta con la potenza aggregata in analisi e con valori invariati anche in momenti differenti della giornata. Per esempio, le ore assumono un valore compreso tra 0 e 23 sia nei giorni festivi che in quelli feriali; il valore del giorno della settimana (da 1 a 7) rimane inalterato per tutte le ventiquattro ore di cui esso si compone, così come il *flag* relativo alla festività o meno della giornata.

In aggiunta, poiché si è analizzato l'insieme di dati relativo al solo mese di gennaio 2015, è chiaro che i *predictor* relativi al mese e all'anno sono inefficaci, poiché mantengono lo stesso valore per tutte le misure disponibili, senza apportare delle informazioni aggiuntive da correlarsi con le variazioni della grandezza incognita.

Tenuto conto di quanto detto, l'anno è stato rimosso dai *predictor*, sebbene esso potrebbe diventare un elemento rilevante nelle analisi su insiemi di dati più ampi; di contro, sono stati introdotti nuovi elementi di previsione: alcuni *predictor* legati agli istanti di campionamento e, quindi, ricavati dalle informazioni sulle tempistiche riportate nel file di partenza; altri calcolati sulla base della potenza apparente misurata. In dettaglio:

- secondi;
- minuti;
- giorno del mese;

- potenza media nel minuto (P_{mean_min}), calcolata come media aritmetica tra le potenze rilevate in ciascuno dei secondi appartenente al minuto considerato;
- potenza media nell'ora (P_{mean_hours}), calcolata come media aritmetica tra le potenze rilevate durante l'ora in esame;
- potenza media nel medesimo minuto del giorno precedente ($P_{mean_min_previous_day}$), determinata come media aritmetica tra le potenze rilevate nel minuto in analisi ma del giorno precedente;
- potenza media nella medesima ora del giorno precedente ($P_{mean_hours_previous_day}$), calcolata eseguendo la media aritmetica tra le potenze rilevate nell'ora in esame ma appartenente al giorno precedente.

Questi ultimi quattro *predictor* sono maggiormente connessi con la grandezza incognita e permettono di mettere in relazione tra loro le giornate. Infatti, sebbene i valori istantanei non siano uguali o comunque i campionamenti non avvengano nel medesimo istante, può verificarsi che nel minuto e nell'ora i valori medi rimangano inalterati, o differiscano di poco, nel passaggio da un giorno all'altro. Questo fenomeno si spiega ricordando che, tralasciando particolari picchi di potenza correlati a situazioni imprevedibili, i consumi di un'abitazione si ripetono nel tempo, sulla base delle abitudini degli abitanti e della stagione dell'anno.

Il totale dei *predictor* così ottenuto sale da quattro a undici.

Introdotti questi nuovi *predictor*, si è proceduto alla creazione di un nuovo modello costruito con i dati dell'intero mese di gennaio e la totalità dei *predictor* sopra indicati, così da fare il confronto con il modello ad un mese precedentemente analizzato.

Si è scelto di tralasciare i modelli ad un giorno ed a due giorni, in quanto hanno manifestato una importante e non trascurabile correlazione con l'ampiezza dell'intervallo temporale da prevedere; i modelli a quattro e sei giorni non si sono tenuti in considerazione, poiché si ritiene che una quantità superiore di informazioni sull'andamento della potenza aggregata su periodi più estesi possa giovare alla rete neurale, evidenziando la ripetitività della stessa.

In Tabella 3.1.1-1 sono stati riassunti i valori del parametro *RMSE* calcolati dall'applicazione impiegata per la costruzione delle cinque reti neurali pre-programmate disponibili; se confrontati con quelli in Tabella 2.1.1-1, Tabella 2.1.1-2 e Tabella 2.1.1-3, si evince una notevole riduzione dell'errore, anche con frequenze di campionamento inferiori a quella dichiarata nella banca dati, sebbene esse non corrispondono alle prestazioni migliori.

Tabella 3.1.1-1 Valori di *RMSE* per il modello ad un mese completo dell'insieme di 11 *predictor*

MODELLO AD UN MESE					
Intervallo di campionamento	N	M	W	B	T
8 secondi	0.339	0.334	0.332	0.336	0.34
16 secondi	0.339	0.334	0.334	0.336	0.34
24 secondi	0.339	0.336	0.339	0.338	0.34

La decrescita dell'errore è diretta conseguenza di una più accurata previsione da parte della rete neurale rispetto ai casi precedenti, grazie all'incremento delle informazioni ad essa fornite durante la fase di addestramento. In particolare, l'aggiunta dei *predictor* relativi alla potenza aggregata media ha avuto un notevole impatto; a riprova di ciò, si è provveduto all'addestramento della rete con campionamenti a 8 secondi rimuovendo, prima singolarmente e poi contemporaneamente, la potenza media nell'ora (*P_mean_hours*) e quella nei minuti (*P_mean_min*) e si sono riassunti i valori di *RMSE* calcolato dal software nella Tabella 3.1.1-2.

Tabella 3.1.1-2 Valori di *RMSE* per il modello ad un mese con 10 e 9 *predictor*

MODELLO AD UN MESE (campionamento 8 secondi)					
<i>Predictor</i> rimossi	N	M	W	B	T
<i>P_mean_hours</i>	0.339	0.333	0.331	0.338	0.336
<i>P_mean_min</i>	0.763	0.752	0.718	0.749	0.728
<i>P_mean_hours</i> & <i>P_mean_min</i>	0.851	0.805	0.773	0.803	0.777

I risultati prevenienti dalla rimozione dei *predictor* correlati ai valori medi nella medesima giornata confermano il loro peso nell'errore commesso dal modello. In particolar modo, la rimozione della potenza media sul minuto nella medesima giornata ha una incidenza superiore

rispetto a quanto accade con la potenza media nell'ora: infatti, l'*RMSE* nel primo caso cresce in maniera più rilevante rispetto al secondo, prendendo come metro di paragone quanto ottenuto con il set completo di *predictor* (Tabella 3.1.1-1). È chiaro che non sempre sono disponibili le potenze medie sull'ora e sul minuto nella medesima giornata, ma esse possono ricavarsi tenendo conto della ripetitività della potenza aggregata nei giorni, come mostrato nel capitolo successivo.

Il miglioramento nei risultati è maggiormente evidente nelle rappresentazioni grafiche delle potenze; poiché la *wide neural network* con campionamenti ad 8 secondi e 12 *predictor* presenta l'*RMSE* minore, di seguito sono riportati i confronti tra la potenza campionata (in blu) e quella calcolata (in arancione) sulla giornata del 22 gennaio 2015. In tal modo si può operare un confronto diretto con quanto ottenuto nel capitolo precedente.

In particolare, è stata tracciata la potenza prevista sull'intera giornata (Figura 3.1.1-1) e due suoi dettagli: un'ora notturna (Figura 3.1.1-2) e dieci minuti della stessa ora (Figura 3.1.1-3).

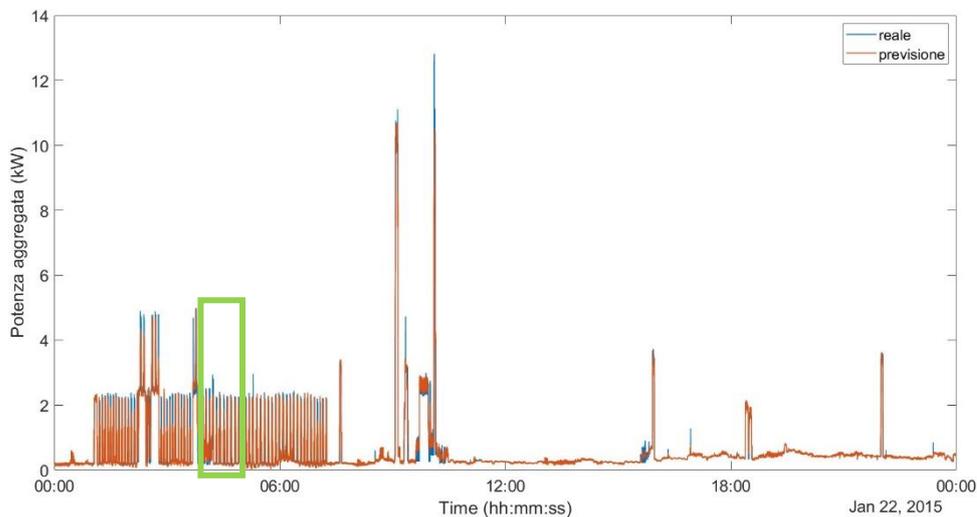


Figura 3.1.1-1 Previsione del 22 gennaio 2015 con modello ad un mese e 11 *predictor*

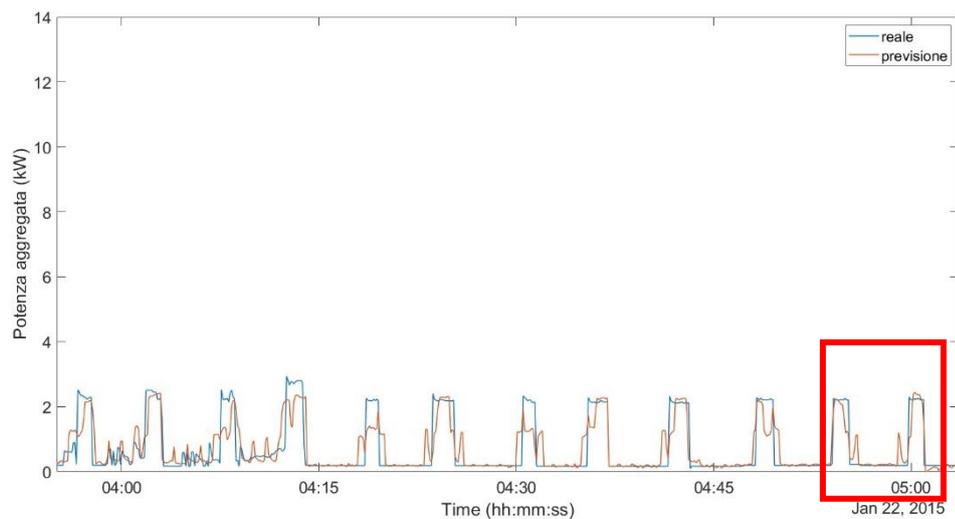


Figura 3.1.1-2 Dettaglio della porzione nel riquadro verde di Figura 3.1.1-1

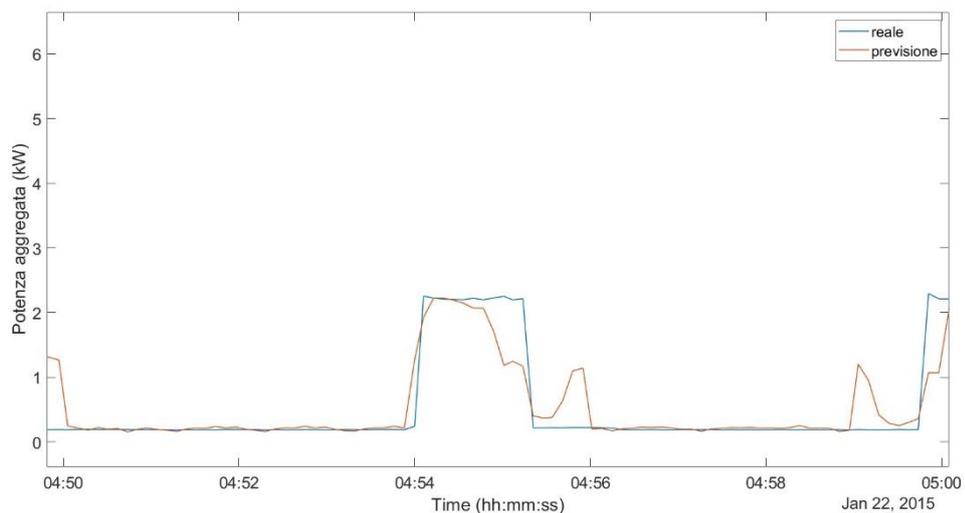


Figura 3.1.1-3 Dettaglio della porzione nel riquadro rosso di Figura 3.1.1-2

Sebbene la previsione sulla intera giornata potrebbe apparire molto vicina a quella campionata, con gli ingrandimenti gli errori diventano visibili con maggiore chiarezza. Infatti, da queste due ultime figure, la potenza prevista è sovrapposta alla reale quando questa non ha significative variazioni (tratti orizzontali inferiori).

Tuttavia, in corrispondenza dei singoli impulsi che caratterizzano le ore notturne, il modello fornisce per una parte dell'impulso un valore prossimo a quello corretto e, per la restante parte, una potenza aggregata abbastanza lontana dall'obiettivo. Tale affermazione non è però

generalizzabile, poiché sono presenti impulsi nelle ore notturne in cui la differenza tra realtà e previsione si approssima a zero. Per esempio, si osservi in Figura 3.1.1-4 quello che accade tra le ore 3:17 e le 3:19: a parte le aree nell'intorno al fronte di discesa e di salita, la previsione in corrispondenza di questo impulso è pressoché uguale alla potenza aggregata reale, sebbene nell'impulso successivo le due tornano a discostarsi. Quindi, anche l'istante di previsione è importante, sebbene la grandezza sia ripetitiva.

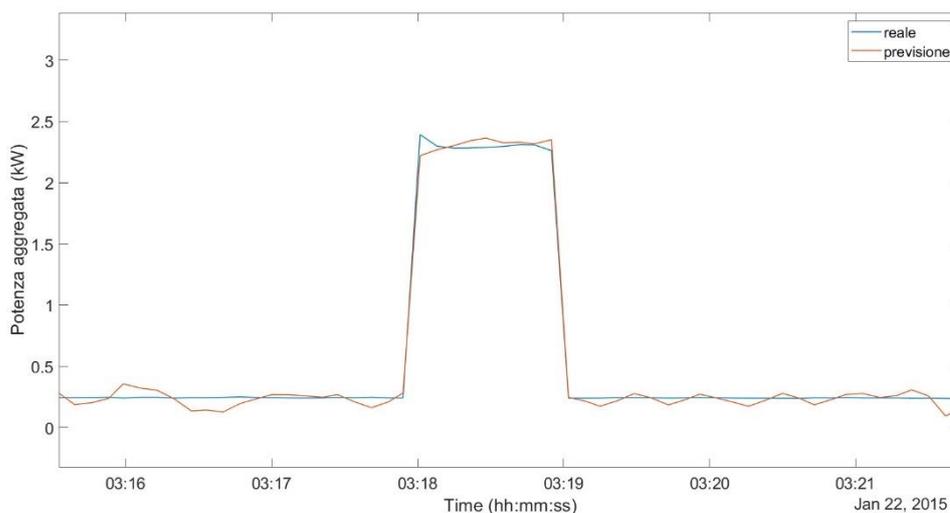


Figura 3.1.1-4 Dettaglio delle ore notturne in Figura 3.1.1-1

3.1.2 INCREMENTO DEI NEURONI

Le informazioni che una rete neurale gestisce sono correlate con il numero di neuroni presenti in ciascuno strato.

Nello strato di input, ognuno dei neuroni riceve come ingresso i valori dei *predictor* e li restituisce come output; negli strati successivi, ogni neurone presente raccoglie gli output di tutti i neuroni dello strato precedente e li elabora per restituire a sua volta un dato, pronto per lo strato successivo. Poiché l'elaborazione delle informazioni avviene all'interno degli strati nascosti, la capacità di apprendimento dell'intera rete è subordinata al comportamento e all'ammontare dei neuroni che li compongono.

Dunque, per i problemi complessi caratterizzati da un gran numero di dati e grandezze variabili in maniera casuale, l'aumento dei neuroni negli strati intermedi può condurre ad un miglioramento complessivo del risultato di calcolo. Tuttavia, se da un lato un basso numero di neuroni impedisce al modello di acquisire sufficienti informazioni per una corretta ricostruzione (*under-fitting*), dall'altro un suo aumento eccessivo non è necessariamente positivo. In quest'ultima situazione, infatti, la rete sicuramente imparerebbe in maniera più approfondita ma troppo specializzata (*over-fitting*). Dunque, in entrambi questi due casi estremi si otterrebbero delle previsioni imprecise.

Non esiste un metodo semplice ed univoco per l'identificazione a priori del numero ottimale di neuroni, differente con il variare del problema affrontato. Per tale motivazione, si è proceduto per tentativi lavorando con l'insieme completo di *predictor* disponibili per l'intero mese di gennaio 2015, così da evitare le problematiche relative alla mancanza di informazioni.

Il modello ottimale costruito sull'intero mese è la rete neurale *wide* la quale, di default, nel primo strato contiene 100 neuroni con i quali si è ottenuto un *RMSE* di 0.332 (Tabella 3.1.1-1); dall'applicazione *Regression Learner* si è proceduto a modificare l'ammontare dei neuroni sia a salire che a scendere, riportando i valori dell'*RMSE* trascritti in Tabella 3.1.2-1.

Tabella 3.1.2-1 Valori di *RMSE* per la *wide neural network* con differente numero di neuroni

wide neural network sul mese di gennaio								
Neuroni	30	40	50	60	110	120	130	140
<i>RMSE</i>	0.33269	0.33246	0.33213	0.33067	0.3299	0.32982	0.3297	0.3299

Dallo studio delle informazioni in Tabella 3.1.2-1, le variazioni del quantitativo di neuroni costituenti la rete neurale conducono ad uno spostamento minimo del valore dell'*RMSE* rispetto a quanto si era ottenuto con le impostazioni predefinite, a discapito di un notevole aumento delle tempistiche dell'addestramento della stessa. Infatti, si vuole sottolineare in questo contesto, che la variazione dei parametri e la complessità sempre superiore delle reti neurali utilizzate si correla con una lievitazione dei tempi richiesti per l'addestramento, variabili dai trenta minuti alle quattro ore.

A conferma di queste conclusioni, è stato riportato in Figura 3.1.2-1 un dettaglio della forma d'onda ottenuta con l'impiego dei modelli con in numero minimo di neuroni. Se confrontato con quanto avviene con la rete *wide* predefinita (Figura 3.1.1-2), non sono presenti differenze rilevanti. Per tale motivo si può concludere che le considerazioni esposte nel precedente paragrafo rimangono valide e che l'aumento dei neuroni in questo contesto non fornisce alcun miglioramento sostanziale.

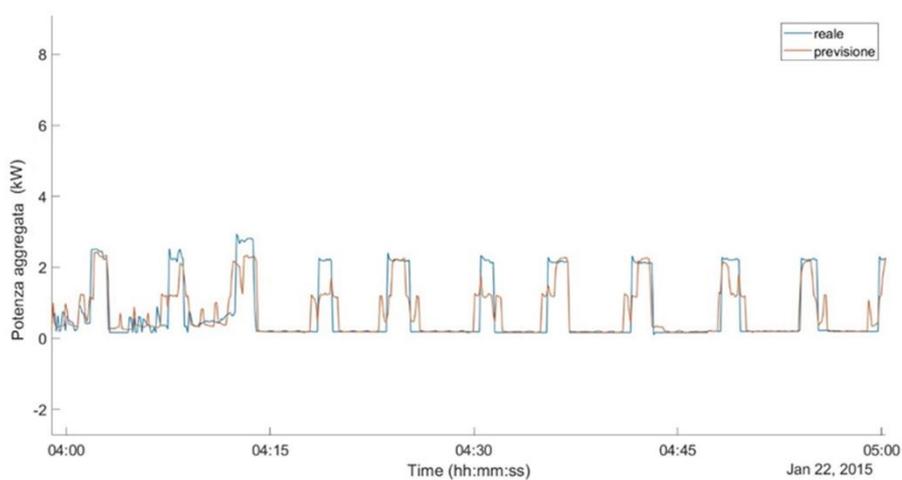


Figura 3.1.2-1 Dettaglio della previsione del 22 gennaio con rete wide 30 neuroni

3.2 CORRELAZIONE ERRORE-ISTANTE DI PREVISIONE

3.2.1 VALUTAZIONE SULLA PREVISIONE DEL 22 GENNAIO 2015

Prima di procedere oltre, si vuole aprire una parentesi per fornire degli esempi che mostrino la stretta correlazione tra la precisione del modello e l'istante di previsione. Considerando la singola simulazione, infatti, sono presenti dei momenti della giornata in cui la ricostruzione è praticamente coincidente con il campionamento, sebbene in altri istanti la stessa sia completamente scorretta.

Nei paragrafi precedenti, sono stati attenzionati i dieci minuti che intercorrono tra le 4:50 e le 5:00; per esempio, con la *wide neural network* si era arrivati a quanto riportato in Figura 3.1.1-1. Tuttavia, ingrandendo quanto ottenuto dalla stessa simulazione nei minuti tra le 3:10 e le 3:35, si può affermare che la ricostruzione riesce a prevedere la potenza realmente richiesta dall'abitazione

in molti degli impulsi che ne caratterizzano l'andamento in quelle ore notturne, come mostrato nei rettangoli neri in Figura 3.2.1-1 i quali evidenziano i momenti di maggiore accuratezza.

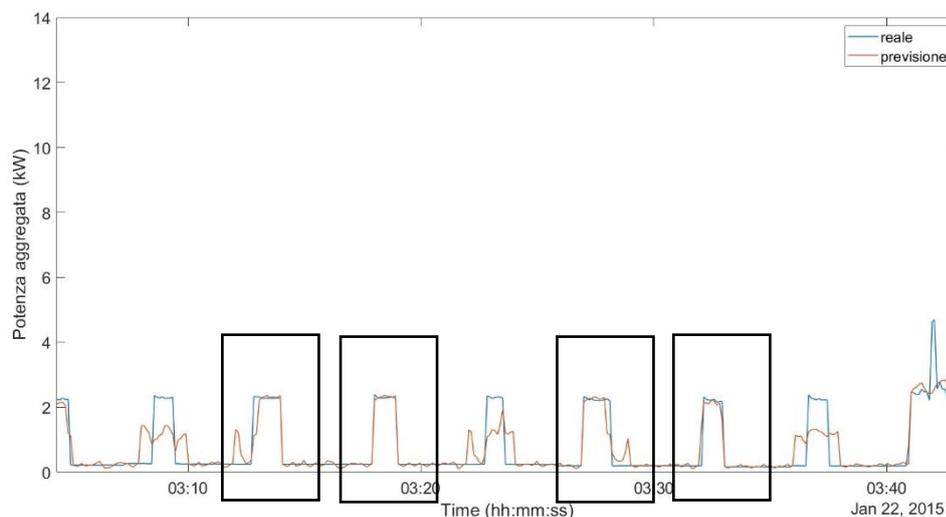


Figura 3.2.1-1 Dettaglio della previsione con rete *wide*, 50 neuroni e 11 *predictor*

3.2.2 VALUTAZIONE SULLA PREVISIONE DEL 21 MARZO 2015

Alla medesima conclusione si arriva anche conducendo le stesse analisi svolte fino a questo momento su un altro mese dell'anno. Per esempio, si è analizzato il mese di marzo 2015, nel quale le lacune sono molto più numerose rispetto a quelle di gennaio, come si osserva dalla rappresentazione grafica della potenza aggregata in questo mese (Figura 3.2.2-1). In essa compaiono due importanti lacune: la prima a cavallo tra il 15 e il 16 marzo e, la seconda, tra il giorno 22 e il 24.

L'assenza di questi dati condurrebbe a pensare ad un modello di previsione meno accurato rispetto ai casi precedenti. In parte, tale conclusione si può ritenere corretta se si fa riferimento ai soli valori di *RMSE* calcolati in fase di addestramento dall'applicazione *Regression Learner*; infatti, questi sono molto elevati, indipendentemente dal tipo di rete neurale impiegato.

Tuttavia, si è analizzato quanto accade nella ricostruzione della giornata del 21 marzo con l'impiego della rete neurale di tipo *wide*, con le sue caratteristiche predefinite (Figura 3.2.2-2).

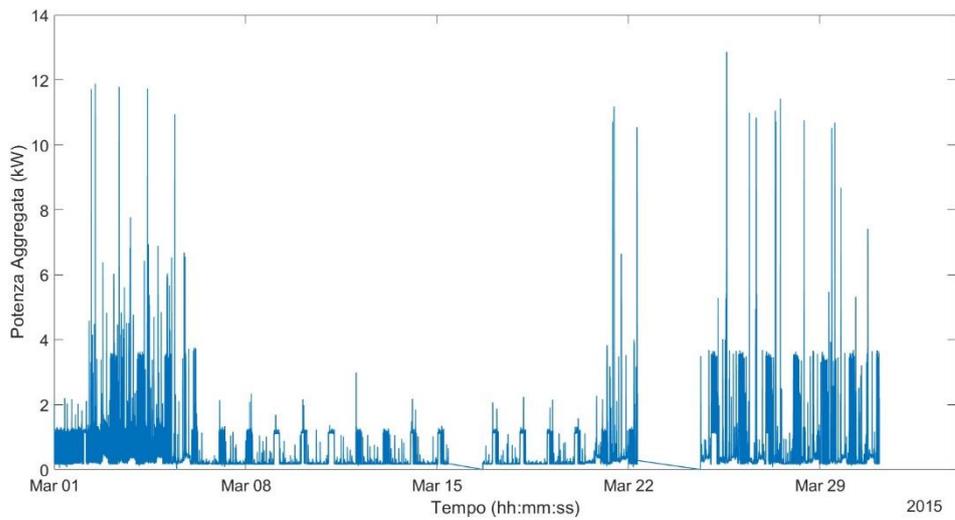


Figura 3.2.2-1 Carico aggregato reale nel mese di marzo 2015

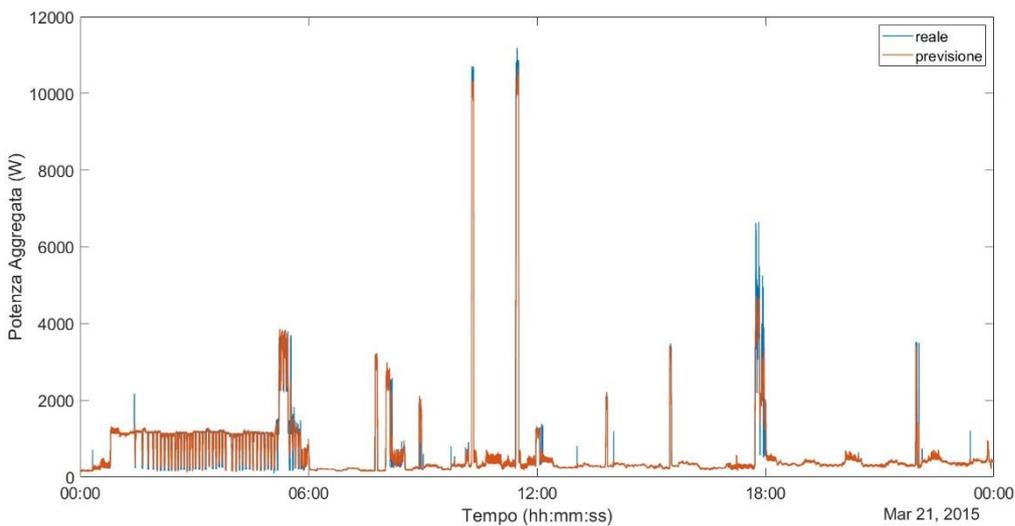


Figura 3.2.2-2 Previsione del 21 marzo con rete *wide*, pre-programmata

In Figura 3.2.2-2 è mostrata la ricostruzione dell'intera giornata, mentre in Figura 3.2.2-3 se ne mostra un ingrandimento delle ore notturne. Da queste due immagini si evince che sebbene il software abbia restituito dei valori di *RMSE* elevati, le conclusioni cui si è giunti per la previsione del mese di gennaio non sono così differenti da quelle che si possono osservare per quest'ultima simulazione. Infatti, nuovamente si presentano degli intervalli in cui la potenza prevista si discosta maggiormente da quella reale ed altri in cui le due si avvicinano.

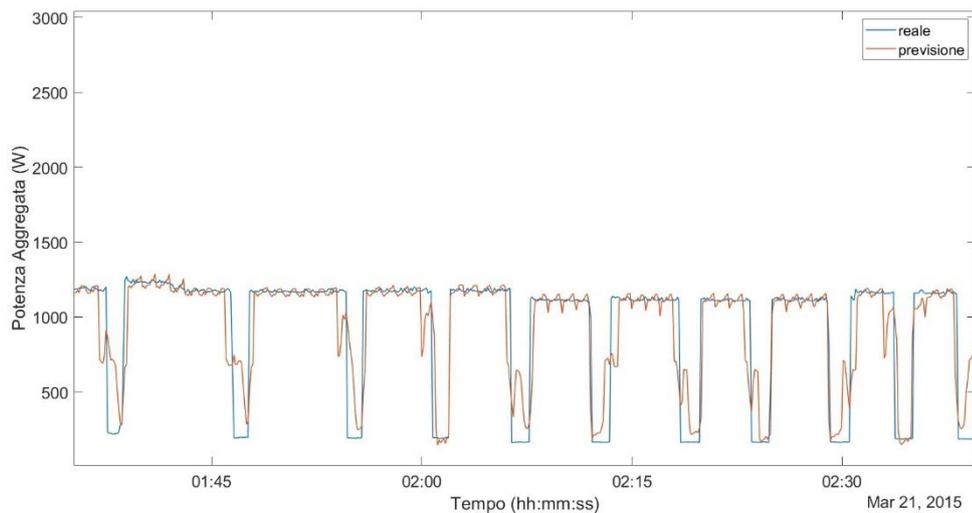


Figura 3.2.2-3 Dettaglio di Figura 3.2.2-2

Per tale ragione, si conclude che la valutazione dell'accuratezza di un modello di previsione non è un processo banale, poiché connesso ad un numero elevato di fattori, con i quali non è possibile stabilire una correlazione univoca e ben definita; tra di essi, l'intervallo temporale di previsione è un fattore di maggiore criticità poiché, nella medesima simulazione è possibile avvicinarsi quanto allontanarsi in maniera rilevante da quello che è il reale sviluppo della potenza aggregata.

3.3 SERIE TEMPORALI E NARX

Appurato il miglioramento della previsione ma la sua non completa accuratezza, studiando più approfonditamente gli andamenti in Figura 3.1.1-1 si può osservare come gli errori siano in parte frutto delle forti oscillazioni della potenza prevista al di sopra o al di sotto della potenza reale, la quale viene tuttavia ottenuta in alcuni istanti. Tali oscillazioni potrebbero essere ridotte confrontando il risultato ottenuto con i modelli trattati nei precedenti paragrafi con quelli ottenuti mediante altre simulazioni. A tale scopo, si è proceduto con l'impiego della rete neurale NARX a disposizione nel Neurale Net Time Serie App.

3.3.1 DEFINIZIONE DEI PARAMETRI DELLE RETI NARX IMPIEGATE

Come indicato dalla casa produttrice del software, vi è una correlazione tra i risultati ottenuti con il numero di neuroni e il ritardo scelto per l'*hidden layer*; quindi, le verifiche condotte sono caratterizzate per una variazione di questi due parametri, come riassunto in Tabella 3.3.1-1.

I valori di *RMSE* calcolati sono molto piccoli, con valori decrescenti all'aumentare del numero di neuroni impiegati per la rete, riduzione non altrettanto evidente con il cambiamento dei ritardi inseriti.

Tabella 3.3.1-1 Valori di *RMSE* per le reti NARX con differente numero di neuroni e ritardi

NARX									
Num. prova	1	2	3	4	5	6	7	8	9
Num. neuroni	30	40	50	30	40	50	30	40	50
Ritardo	4	4	4	6	6	6	8	8	8
<i>RMSE</i>	0.0579	0.0542	0.0549	0.0428	0.0416	0.0505	0.0339	0.0419	0.0364

Tuttavia, degli errori così ridotti non sono corrispondenti a previsioni altrettanto accurate; infatti, ci sono due aspetti che vengono messi in luce solo dalla rappresentazione grafica delle stesse.

In primo luogo, spesso gli andamenti ricostruiti hanno valori negativi, non accettabili per il problema trattato, essendo l'incognita la potenza di un carico. Questo problema viene risolto imponendo una potenza aggregata nulla negli istanti in cui la previsione è al di sotto dello zero.

In secondo luogo, le risposte ottenute sono molte disturbate, ossia, oscillano fortemente centrandosi sul valore reale incognito; questo fenomeno perdura per intervalli temporali limitati ma determina comunque un incremento sostanziale dell'errore, più marcato al crescere dei ritardi caratterizzanti la rete neurale. Questo fenomeno, sempre presente, non esclude la possibilità di ottenere ricostruzioni, soprattutto nelle ore notturne quando la grandezza reale si ripete nel tempo.

3.3.2 ISPEZIONE GRAFICA DEI RISULTATI

Quanto detto si può visualizzare bene nella rappresentazione grafica degli andamenti ottenuti con due modelli differenti; in particolare, si riportano le risposte del modello con 30 neuroni e un

ritardo pari a 4 (Figura 3.3.2-1) e del modello con medesimo numero di neuroni ma ritardo 8 (Figura 3.3.2-2), impiegati per la previsione della potenza aggregata nella giornata del 22 gennaio 2015, la stessa dei precedenti modelli. Si è scelto di riportare quanto restituito da queste due simulazioni poiché sono quelle con valori di *RMSE* della previsione maggiori rispetto agli altri e, quindi, con oscillazioni più evidenti; tuttavia, le stesse considerazioni rimangono valide anche per le altre reti NARX.

Le evoluzioni in Figura 3.3.2-1 e Figura 3.3.2-2 presentano degli evidenti gruppi di oscillazioni e valori negativi. L'insieme di questi due fenomeni conduce agli errori trascritti in Tabella 3.3.2-1, molto diversi da quelli dei modelli stessi (Tabella 3.3.1-1). Nella medesima tabella si sono riportati anche i valori di *RMSE* in seguito alla correzione dei valori negativi delle previsioni: tale semplice azione permette di ottenere un miglioramento che tuttavia non risulta sufficiente.

Tabella 3.3.2-1 Valori di *RMSE* previsioni del 22 gennaio 2015 con reti NARX costituite da 30 neuroni

NARX con 30 neuroni		
Ritardo	4	8
RMSE senza correzione	0.4368	3.2844
RMSE con correzione	0.4256	2.1570

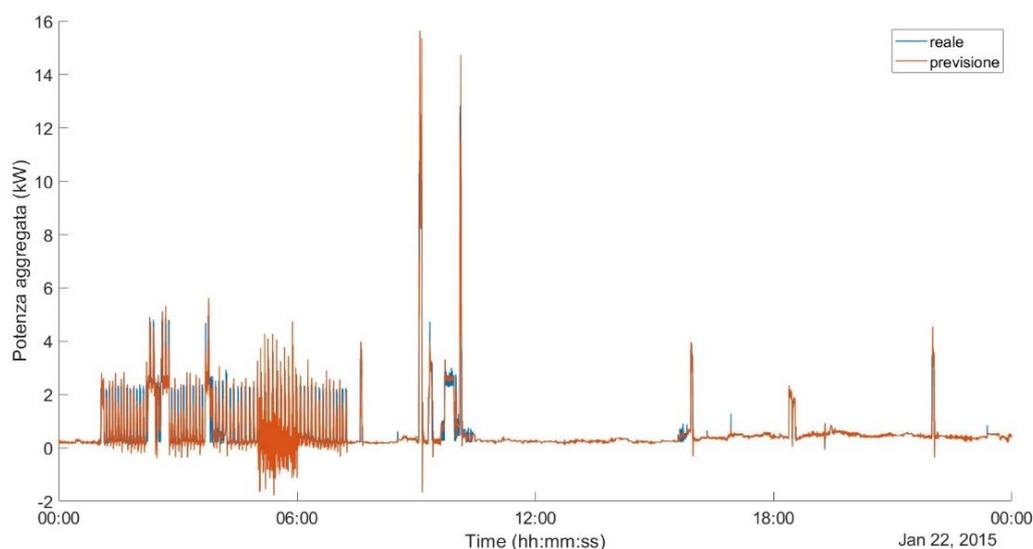


Figura 3.3.2-1 Previsione ottenuta con rete NARX costituita da 30 neuroni e ritardo 4

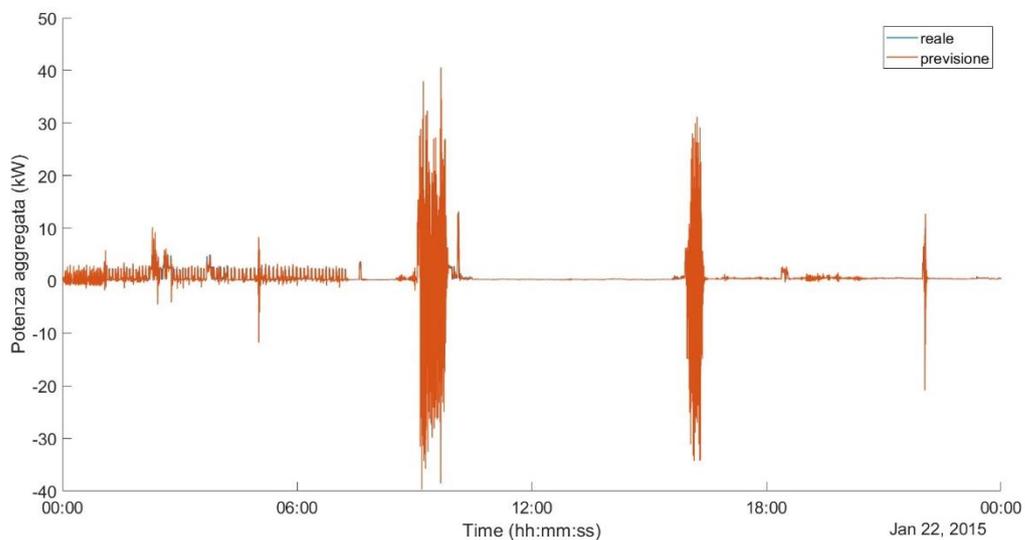


Figura 3.3.2-2 Previsione ottenuta con rete NARX costituita da 30 neuroni e ritardo 8

Nelle seguenti immagini sono riportati alcuni dettagli delle evoluzioni complessive ottenute dai due modelli; in particolare, la Figura 3.3.2-3 e la Figura 3.3.2-4 riportano l'ingrandimento delle ore in cui la previsione è particolarmente disturbata. Per la rete NARX con ritardo minore, le oscillazioni sono principalmente nelle ore notturne e di valore contenuto in un intervallo tra 5 kW e -2 kW; invece per la seconda simulazione approfondita, la potenza prevista sovrasta completamente quella reale, nascondendola dietro un gruppo di oscillazioni molto ampie, con massimi negativi che sfiorano -40 kW. Inoltre, questi disturbi si manifestano in una fascia oraria in cui la potenza reale non presenta particolari variazioni; un'anomalia simile si ha nella medesima previsione nelle ore pomeridiane ma con oscillazioni minori di quelle delle ore diurne ma superiori a quelle del modello con ritardo 4.

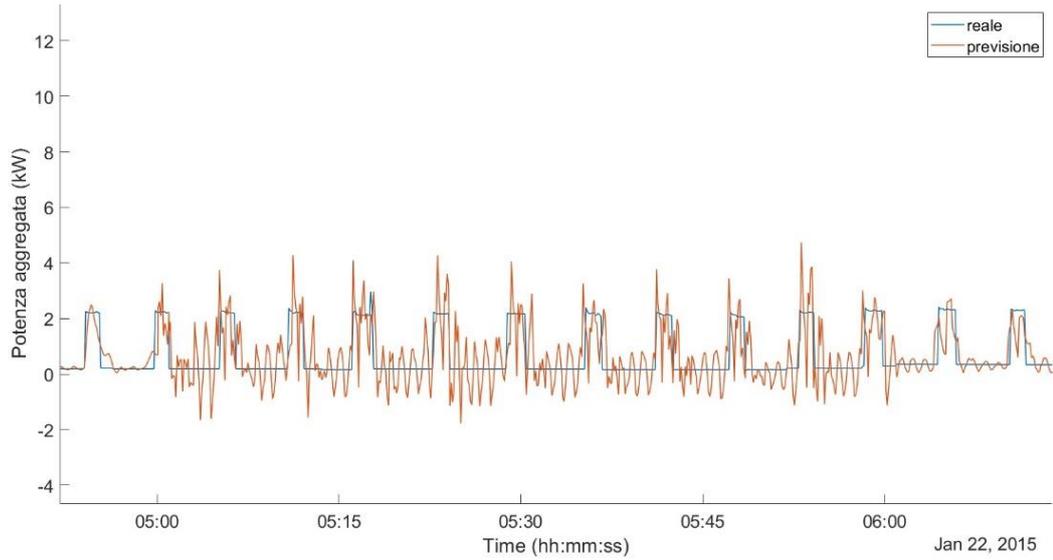


Figura 3.3.2-3 Dettaglio Figura 3.3.2-1 : disturbo nella potenza prevista con rete NARX a 30 neuroni e ritardo 4

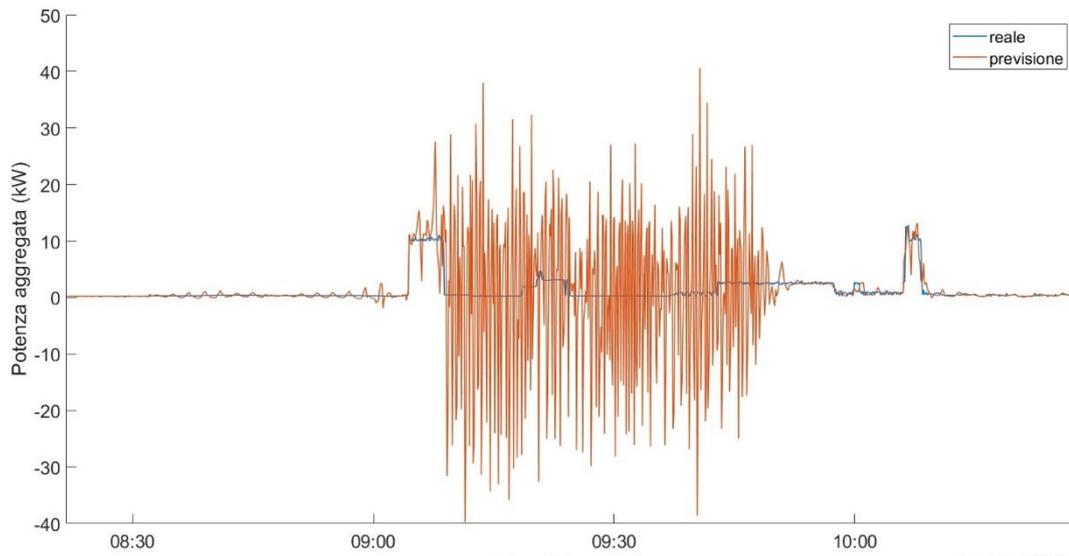


Figura 3.3.2-4 Dettaglio Figura 3.3.2-2: disturbo nella potenza prevista con rete NARX a 30 neuroni e ritardo 8

La Figura 3.3.2-5 e la Figura 3.3.2-6 riportano un secondo dettaglio delle previsioni nelle ore notturne comprese tra le 4:30 e le 6:00. Nel primo modello, la previsione è meno accurata, con errori anche di 1 kW al contrario di quanto avviene nel secondo caso, ove gli impulsi della potenza aggregata vengono seguiti con maggiore accuratezza.

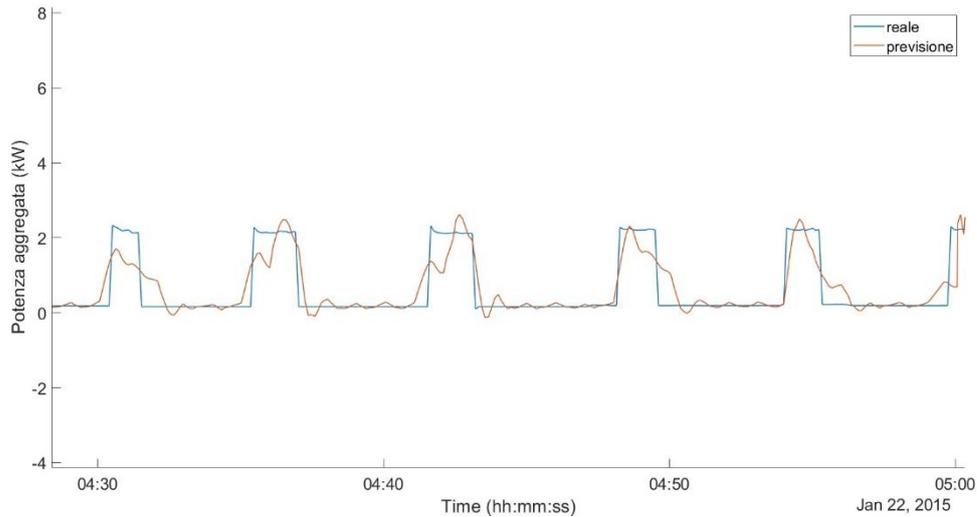


Figura 3.3.2-5 Dettaglio Figura 3.3.2-1: potenza prevista nelle ore notturne con rete NARX a 30 neuroni e ritardo 4

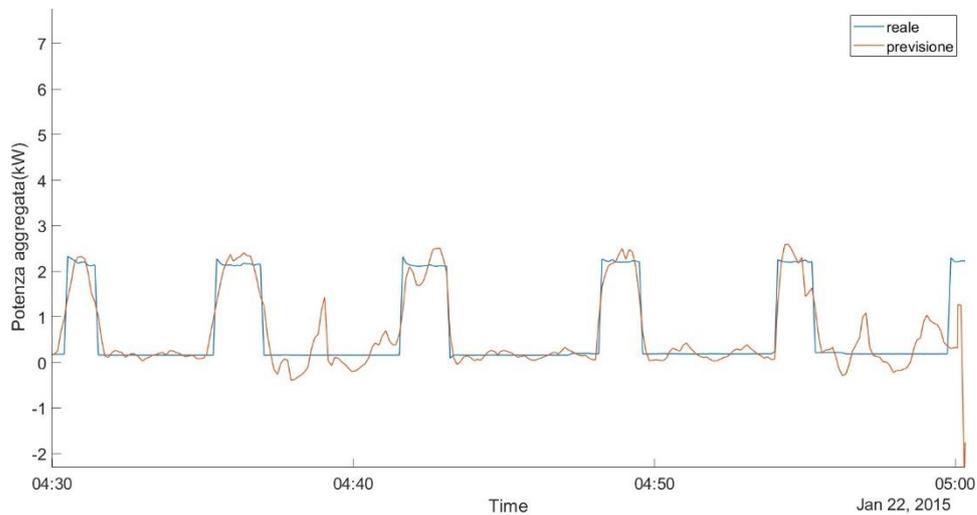


Figura 3.3.2-6 Dettaglio Figura 3.3.2-2: potenza prevista nelle ore notturne con rete NARX a 30 neuroni e ritardo 8

L'analisi di questi grafici mette in luce i punti di forza di ciascuno dei due modelli a parità di neuroni. Il cambiamento del numero di neuroni costituenti la rete neurale non conduce a particolari miglioramenti nella previsione, come si conclude paragonando quanto avviene con il modello NARX sopra analizzato, a 30 neuroni e ritardo 4, con quanto si ottiene mantenendo invariato questo secondo parametro e incrementando i neuroni a 50. La potenza aggregata calcolata da questo secondo modello è mostrata in Figura 3.3.2-7, mentre la Figura 3.3.2-8 restituisce un suo ingrandimento nelle ore notturne.

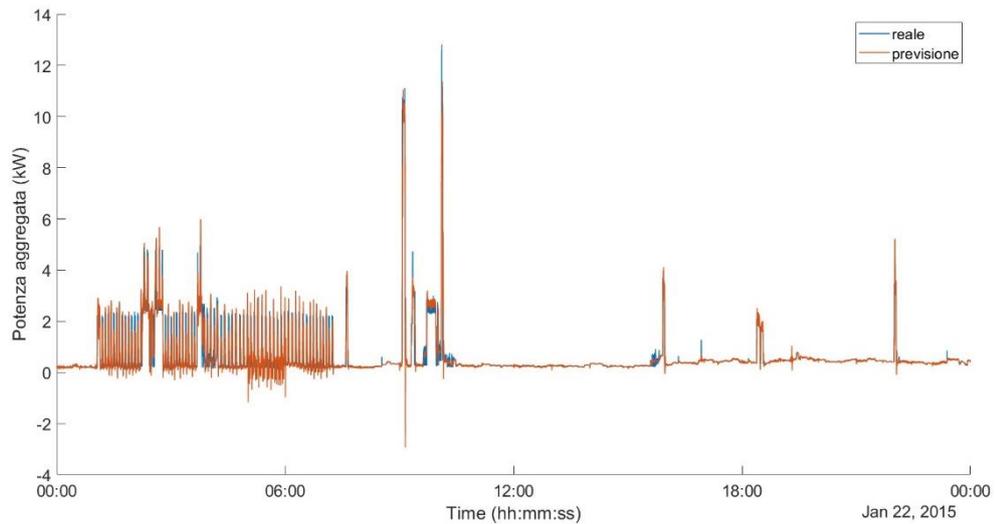


Figura 3.3.2-7 Previsione ottenuta con rete NARX costituita da 50 neuroni e ritardo 4

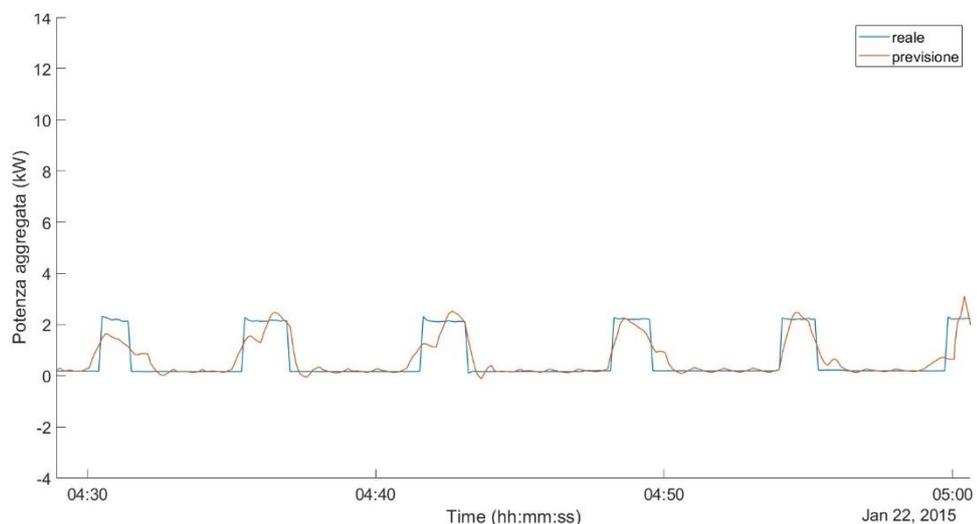


Figura 3.3.2-8 Dettaglio Figura 3.3.2-7: potenza prevista nelle ore notturne con rete NARX a 50 neuroni e ritardo 4

Le forme d'onda con la rete a 50 neuroni presentano le stesse imprecisioni della rete a 30 neuroni, con oscillazioni ridotte rispetto a quanto avviene aumentando il ritardo del modello ma con errori non trascurabili anche nelle ore notturne.

Giunti a tali conclusioni, per ridurre l'errore nelle fasi successive, si è deciso di realizzare delle reti NARX mantenendo i ritardi sotto il valore 6, essendosi manifestate in questa condizione delle oscillazioni importanti negli andamenti della potenza aggregata prevista. In particolare, si sono valutati i ritardi 1, 2, 3 e 5: per ognuno di essi si è assunto uno dei valori dei neuroni sopra visti e si sono svolte le previsioni sulla giornata del 22 gennaio 2015. In Tabella 3.3.2-2 sono riportati i valori

dell'*RMSE* di tali reti neurali e delle ricostruzioni con esse raggiunte, sia senza correzione sia nel caso di annullamento di eventuali valori negativi.

Tabella 3.3.2-2 *RMSE* per i modelli NARX aggiuntivi con ritardi inferiori a 6

NARX				
Neuroni	Ritardo	<i>RMSE</i> modello	<i>RMSE</i> previsione	<i>RMSE</i> previsione corretta
30	1	0.0675	0.3918	0.3918
30	2	0.0595	0.4425	0.3958
30	3	0.0567	0.3793	0.3712
30	5	0.0478	0.4932	0.4380
40	1	0.0672	0.3877	0.3877
40	2	0.0591	0.4459	0.4273
40	3	0.0595	0.4036	0.4011
40	5	0.0517	0.4919	0.4797
50	1	0.0631	0.3795	0.3767
50	2	0.0556	0.4118	0.4076
50	3	0.0617	0.4533	0.4489
50	5	0.0468	0.3811	0.3757

Nuovamente il valore dell'errore restituito dal software per ognuno dei modelli è molto piccolo, ma questo cresce di circa un ordine di grandezza quando essi vengono applicati alla previsione dell'evoluzione temporale del carico aggregato, come si può osservare dal grafico ad istogramma riportato in Figura 3.3.2-9; vi è infatti un salto tra le barre di colore più chiaro, relative al solo modello, e quelle più scure, determinate dallo stesso applicato alla previsione del 22 gennaio 2015, corretta dei suoi possibili valori negativi. In questa rappresentazione, si riporta anche quanto avviene con il un ritardo pari a 4, analizzato in precedenza così da avere un quadro completo per tutte le reti NARX con ritardi inferiori a 6.

Non è possibile stabilire una relazione univoca tra errore e ritardo, ma si può constatare che, per alcuni modelli, l'incremento di tale parametro conduce ad un peggioramento del risultato ottenuto. Si consideri, per esempio, la rete con 30 neuroni: l'errore quadratico medio ottenuto per il modello è massimo con ritardo pari ad 1 e minimo per ritardo pari a 5; la situazione opposta si verifica prendendo sotto esame l'errore quadratico medio della previsione. Considerazioni simili si possono

condurre anche per un numero differente di neuroni, per cui diventa complesso scegliere un modello piuttosto che un altro.

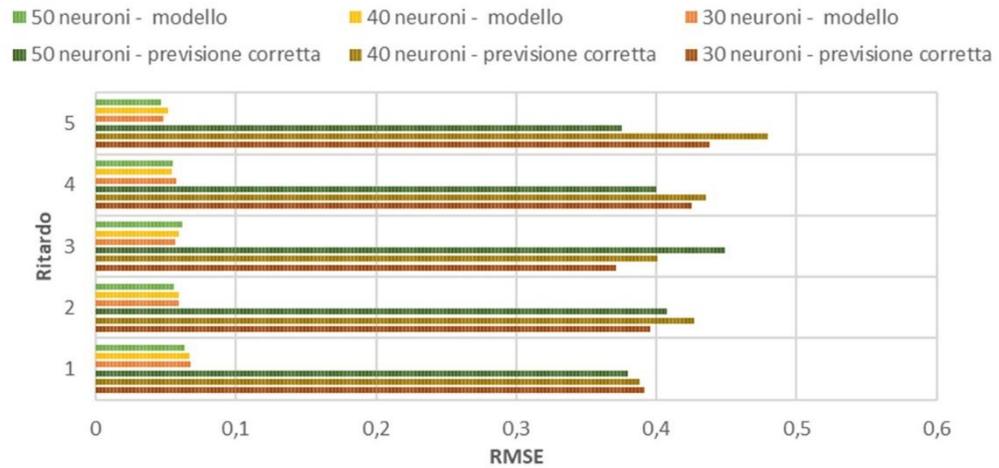


Figura 3.3.2-9 Istogramma RMSE per i modelli NARX con ritardo inferiore a 6

Tuttavia, dall'analisi degli andamenti ricostruiti da queste simulazioni, si può ancora osservare la presenza di un disturbo superiore nella potenza aggregata ricavata con reti con ritardi maggiori. A tal proposito, nelle figure seguenti si mostrano le forme d'onda della grandezza incognita (senza la correzione dei valori negativi) ottenuta con pari numero di neuroni (40) e ritardo differente, sebbene comunque al di sotto del valore 6; in particolare, in Figura 3.3.2-10 è visualizzabile la potenza aggregata campionata (in blu) e quella ricostruita (in arancione) con un modello caratterizzato da ritardo pari a 1, mentre, in Figura 3.3.2-11 sono tracciate le medesime grandezze ma con rete neurale NARX con ritardo 5.

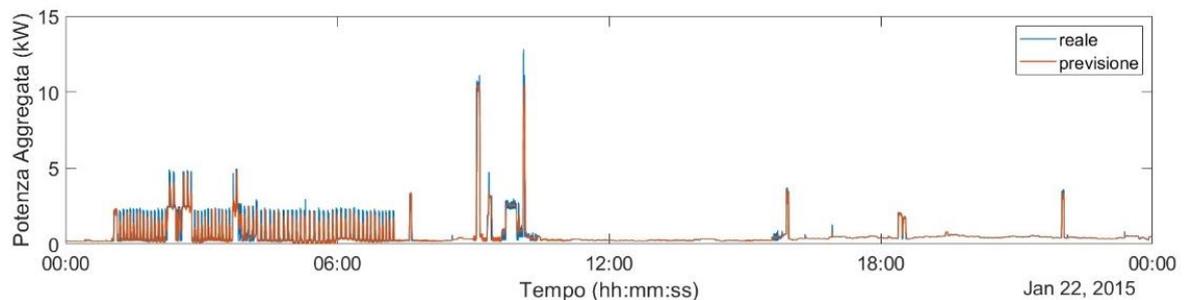


Figura 3.3.2-10 Previsione con rete NARX con 40 neuroni e ritardo 1

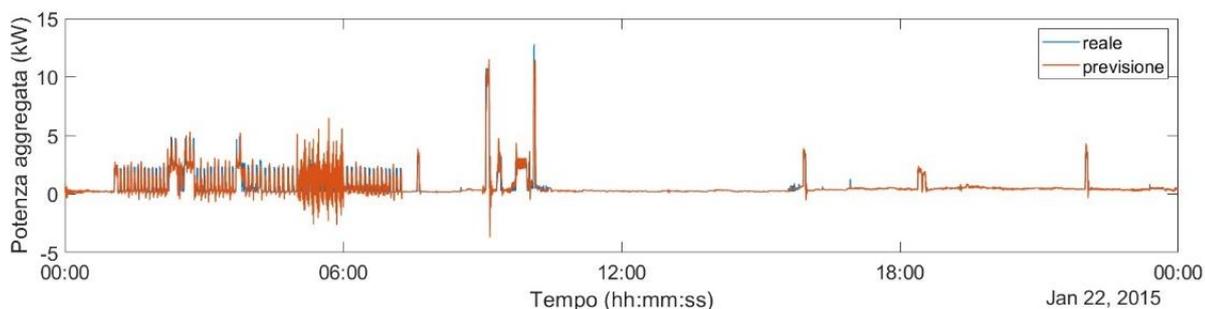


Figura 3.3.2-11 Previsione con rete NARX con 40 neuroni e ritardo 5

La differenza tra le due simulazioni è molto evidente: il disturbo presente nella seconda è molto evidente, soprattutto nelle prime ore del mattino, poco prima delle 6:00, ove le oscillazioni della potenza ricostruita sono maggiormente accentuate. In questo intervallo temporale, infatti, essa raggiunge picchi positivi di 6.5 kW e massimi negativi di -2.6 kW; tali valori sono sicuramente inferiori a quelli ottenuti con un modello NARX con ritardo 6, ma restano comunque problematici.

A valle di queste considerazioni, per il passo successivo si sono considerati i risultati delle reti NARX con neuroni da 30 a 50 e ritardi da 1 a 4.

3.4 MEDIA DEI RISULTATI

3.4.1 MODELLI DA UTILIZZARE NELLO SVOLGIMENTO DELLA MEDIA

Conteggiando, tra quelli analizzati in questo capitolo, i modelli ritenuti utilizzabili per il miglioramento dei risultati ottenuti mediante la *wide neural network*, si è giunti ad avere tredici modelli con caratteristiche differenti impiegati per la previsione del medesimo intervallo temporale, ossia la giornata del 22 gennaio 2015. Con essi, si è proceduto operando la media tra i valori di potenza aggregata prevista ricavata in ogni simulazione, istante per istante, al fine di ottenere un unico andamento. Per arrivare a questo risultato, in un primo momento si è effettuata la previsione con i singoli modelli, costruiti a partire dalla conoscenza degli undici *predictor* e con campionamenti ad otto secondi. Essi sono stati impiegati nell'ordine con cui i risultati ottenuti sono mostrati nelle seguenti figure; solo in un secondo momento, si è proceduto con la media ricostruendo l'andamento temporale di Figura 3.4.1-14.

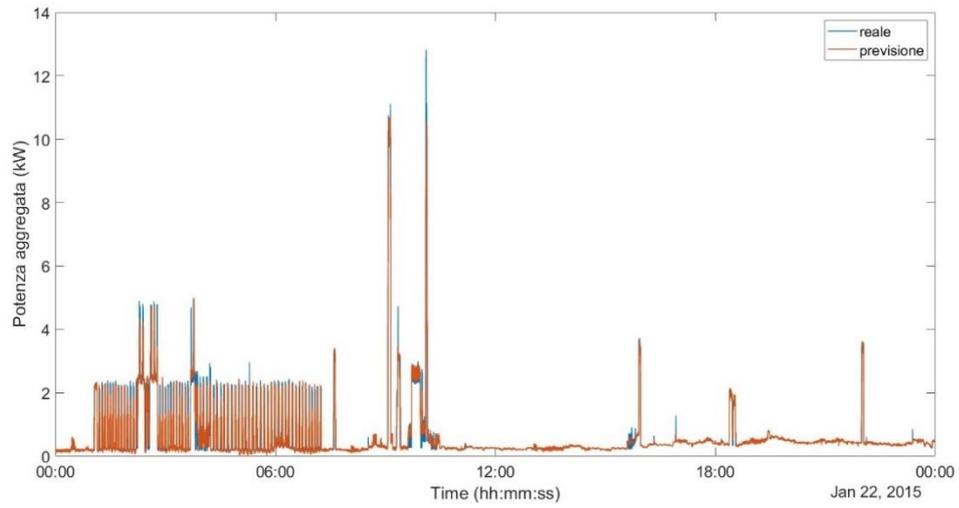


Figura 3.4.1-1 Previsione con rete neurale *wide*, 50 neuroni

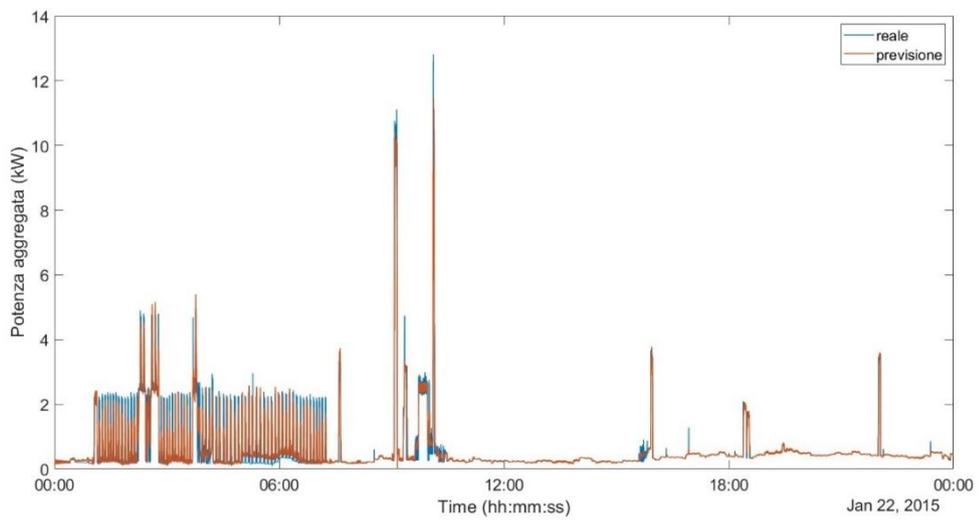


Figura 3.4.1-2 Previsione con rete neurale NARX, 30 neuroni e ritardo 1

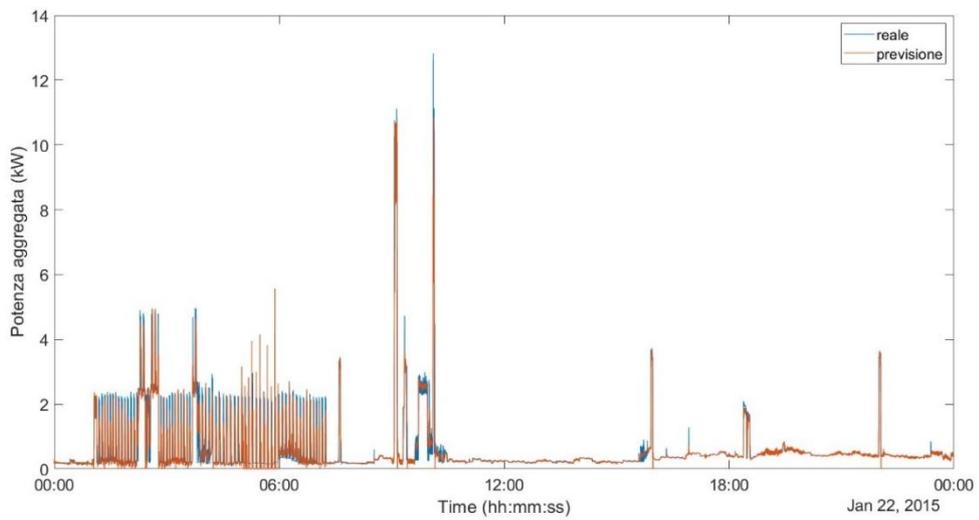


Figura 3.4.1-3 Previsione con rete neurale NARX, 30 neuroni e ritardo 2

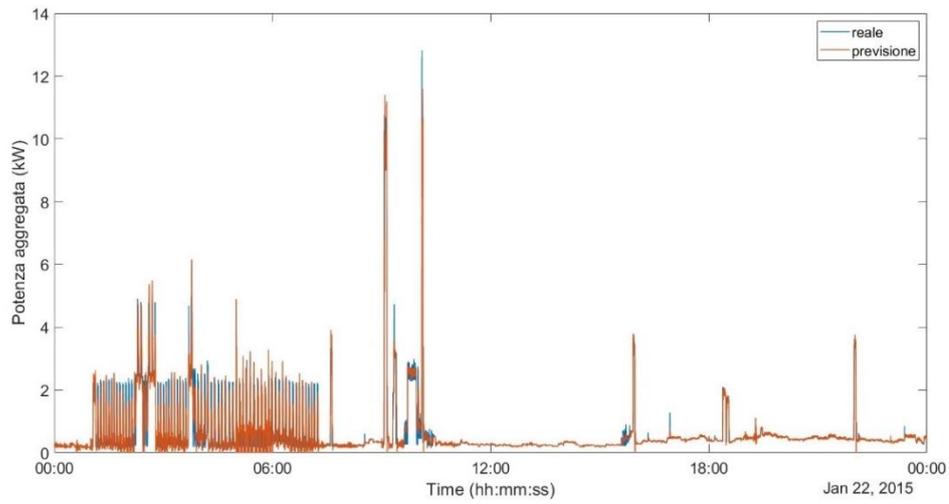


Figura 3.4.1-4 Previsione con rete neurale NARX, 30 neuroni ritardo 3

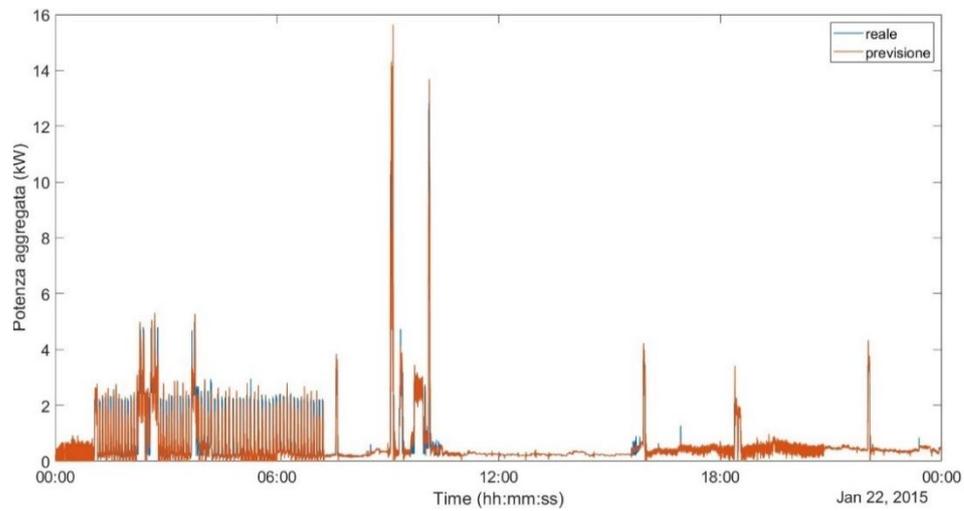


Figura 3.4.1-5 Previsione con rete neurale NARX, 30 neuroni ritardo 4

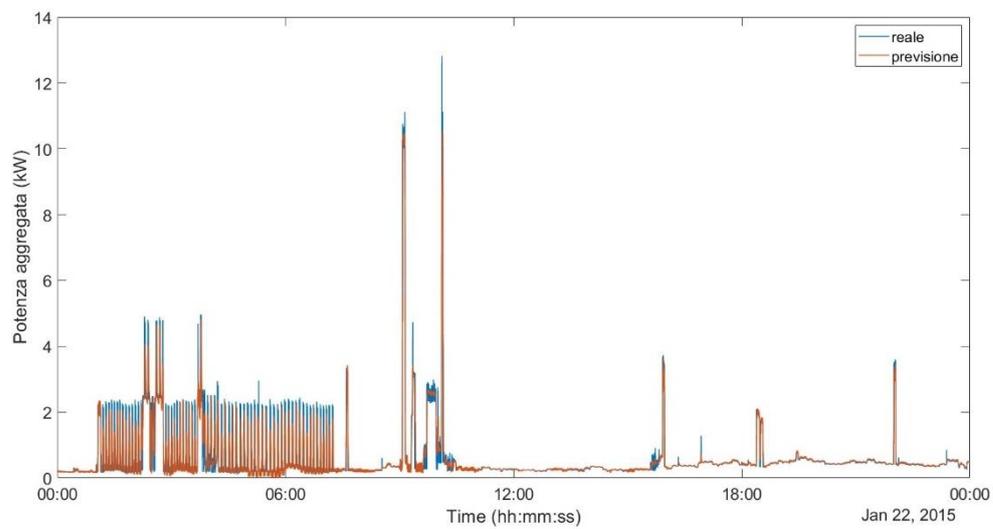


Figura 3.4.1-6 Previsione con rete neurale NARX, 40 neuroni ritardo 1

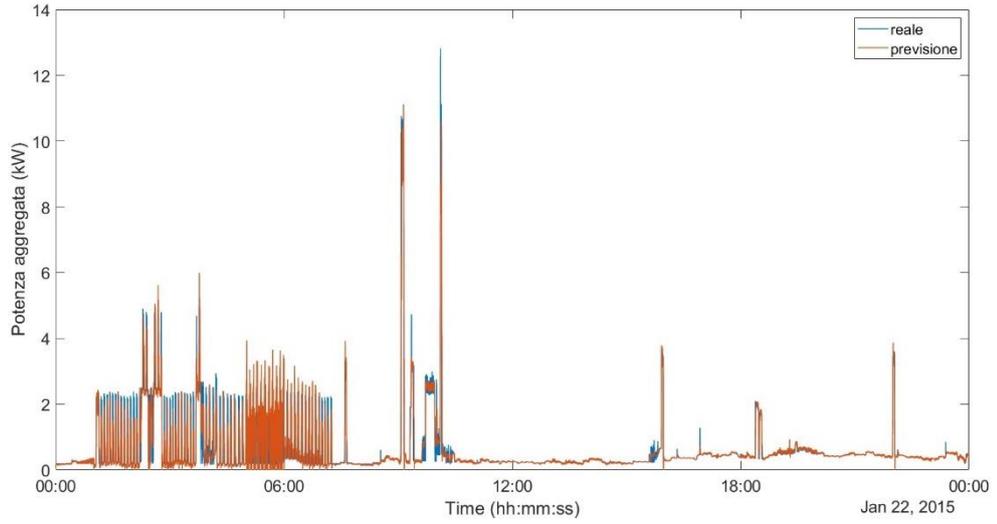


Figura 3.4.1-7 Previsione con rete neurale NARX, 40 neuroni ritardo 2

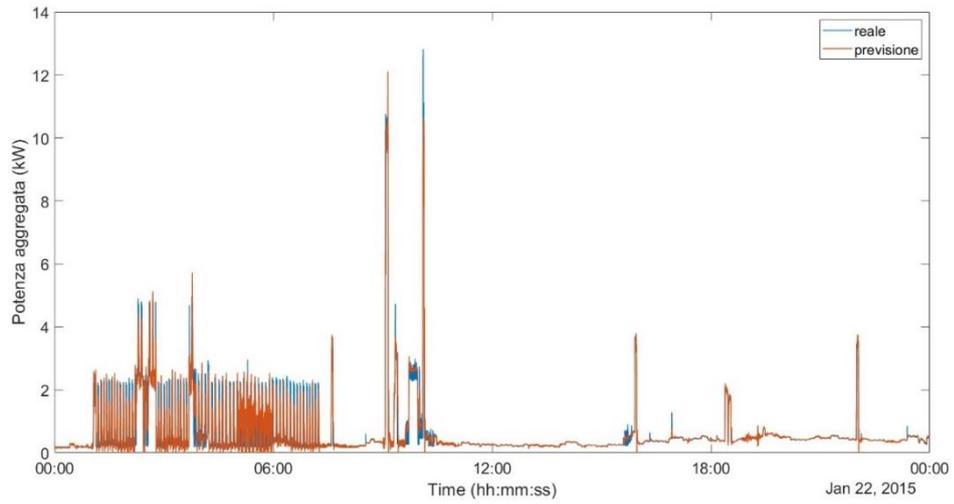


Figura 3.4.1-8 Previsione con rete neurale NARX, 40 neuroni e ritardo 3

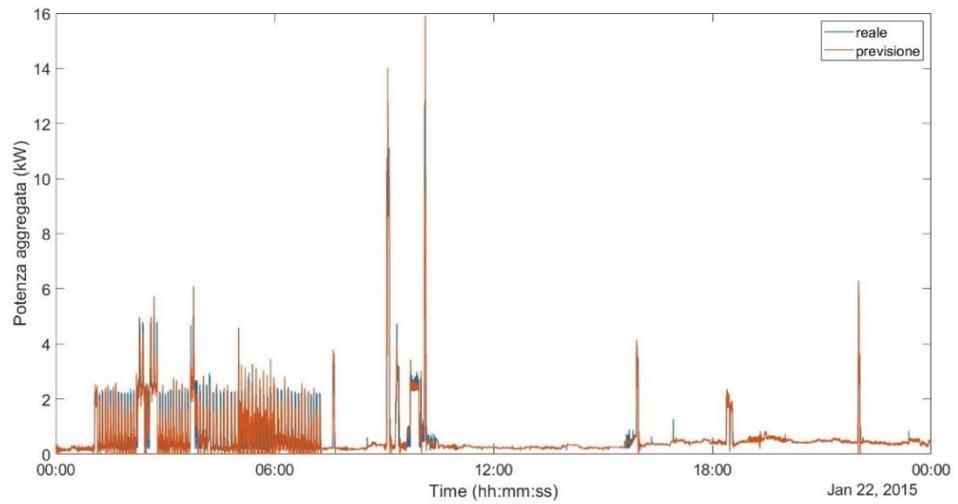


Figura 3.4.1-9 Previsione con rete neurale NARX, 40 neuroni ritardo 4

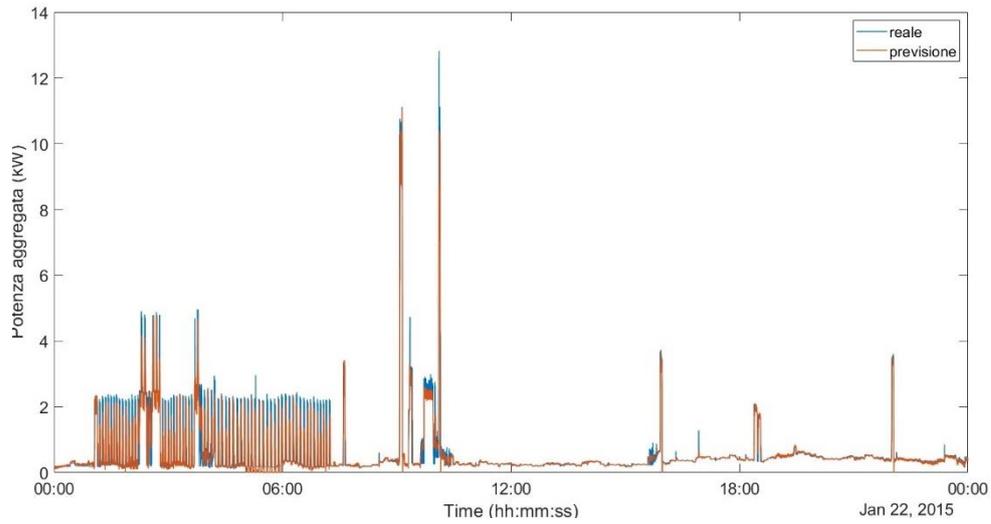


Figura 3.4.1-10 Previsione con rete neurale NARX, 50 neuroni e ritardo 1

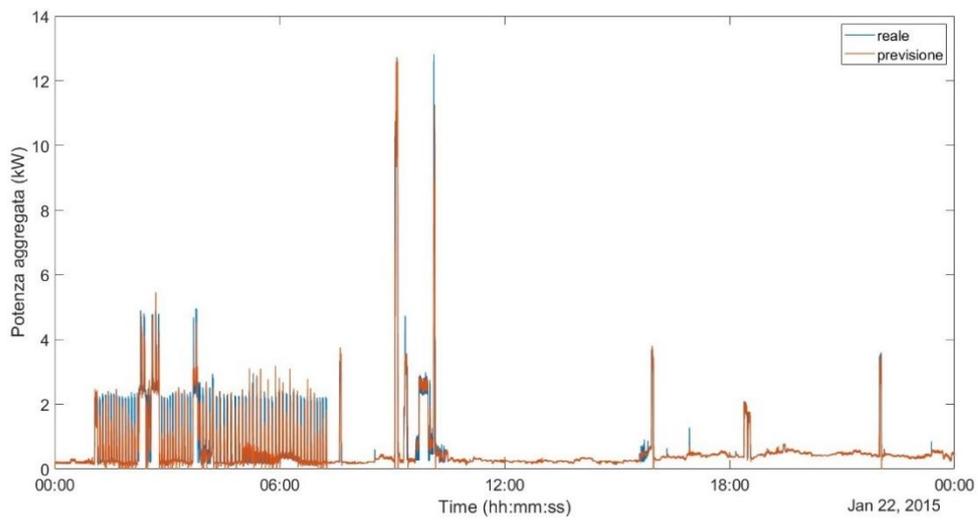


Figura 3.4.1-11 Previsione con rete neurale NARX, 50 neuroni e ritardo 2

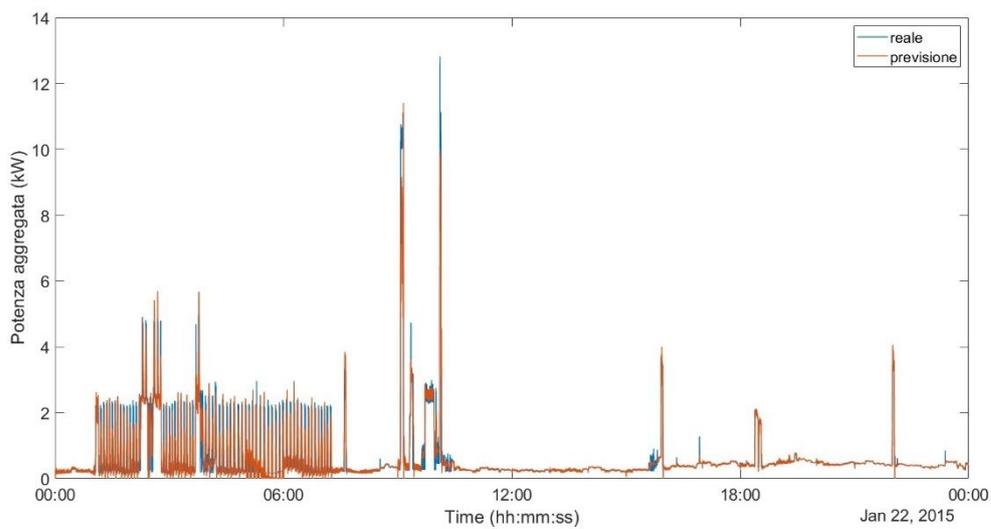


Figura 3.4.1-12 Previsione con rete neurale NARX, 50 neuroni e ritardo 3

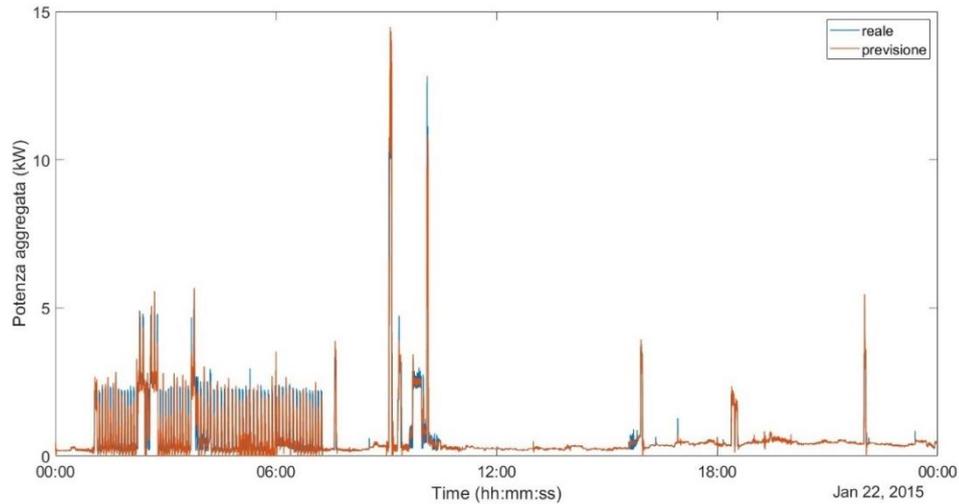


Figura 3.4.1-13 Previsione con rete neurale NARX, 50 neuroni e ritardo 4

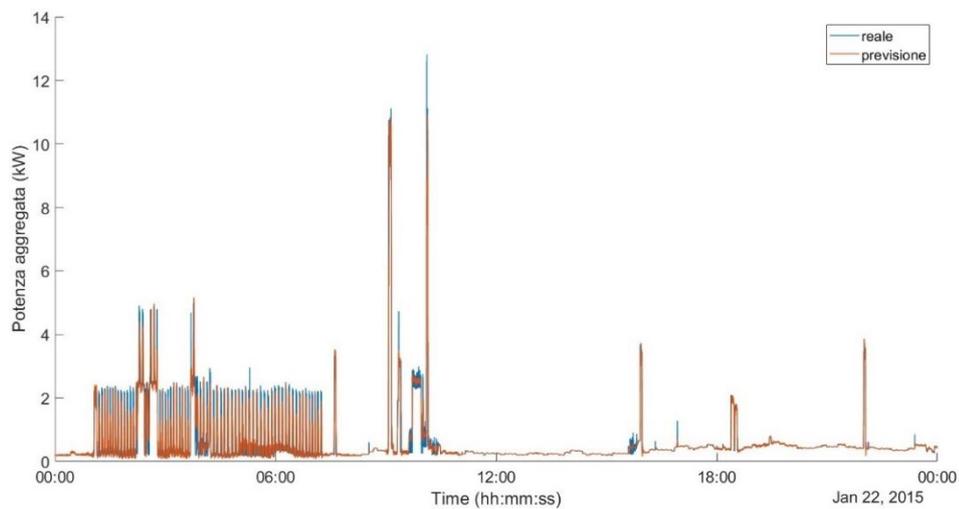


Figura 3.4.1-14 Ricostruzione dell'andamento della potenza aggregata del giorno 22 gennaio 2015 mediante media istante per istante

Confrontando quanto riportato nell'ultima rappresentazione (Figura 3.4.1-14) con la prima (Figura 3.4.1-1), relativa all'andamento che si voleva migliorare, è possibile notare senza ingrandimenti che il processo di media su tutti i modelli non ha condotto ad un miglioramento della previsione ma l'ha invece peggiorata. Questo risultato è conseguenza dell'inadeguatezza allo scopo prefissato dei modelli NARX, poiché con l'incremento del ritardo la loro accuratezza migliorava, come sopra evidenziato, ma i disturbi sul segnale ricostruito erano inaccettabili e hanno compromesso completamente il discreto risultato ottenuto con la rete neurale *wide*.

La minore precisione raggiunta con la riduzione del ritardo ha comunque condotto ad un risultato non soddisfacente: in molti istanti delle ore notturne, per i quali la rete *wide* si era mostrata efficace, nella nuova ricostruzione è visibile una netta separazione tra la linea blu (grandezza campionata) e quella arancione (grandezza ricostruita). Questo conduce a pensare che non solo la forma d'onda finale ricavata per mediatura non sia accurata, ma che il processo matematico stesso abbia condotto ad un peggioramento. A tal proposito, si confrontano per esempio, il dettaglio di Figura 3.1.1-4, relativo alla rete *wide* di partenza (di seguito nuovamente tracciato sulla sinistra per facilitare la constatazione delle differenze) e la forma d'onda finale nelle medesime ore (Figura 3.4.1-15). Il peggioramento è evidente poiché sulla destra si presenta un andamento che richiama l'impulso della potenza aggregata reale ma non lo raffigura così come nel caso a sinistra.

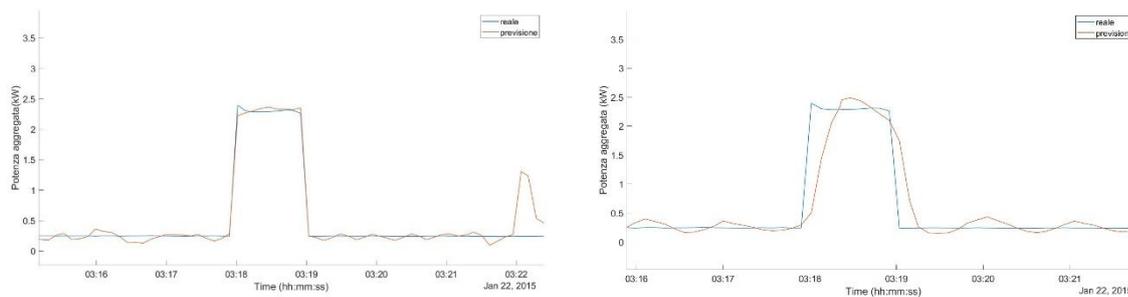


Figura 3.4.1-15 Confronto medesimi istanti della previsione con rete neurale wide e con processo di mediatura

Sebbene si potrebbe pensare ad una media pesata che dia maggiore importanza alle ricostruzioni più affidabili, istante per istante, questa non rappresenta una soluzione. Si ricordi, infatti, che gli andamenti reali sono a noi noti in via del tutto ipotetica, ossia, le linee blu nei precedenti grafici le si deve immaginare come un metro di paragone ma non come possibile strumento per la risoluzione del problema. L'obiettivo è trovare qualcosa che permetta di conoscere in maniera più o meno accurata quanto accade quando i campionamenti non sono disponibili; quindi, non avrebbe senso pesare i risultati di ciascun modello basandosi su di essi.

Altrimenti, la pesatura si potrebbe condurre sulla base del valore dell'*RMSE* che il software associa a ciascun modello, se esso fosse realmente un parametro affidabile. Rimandando alla revisione dei dati numerici ottenuti per questo parametro nella precedente trattazione e alle

considerazioni a cui si è giunti dall'osservazione delle previsioni ottenute, si conclude subito che non molti modelli con *RMSE* inferiore restituiscono andamenti peggiori di altri con tale parametro superiore. Un esempio molto chiaro lo si riscontra confrontando l'*RMSE* della rete *wide* con quello di una qualunque delle reti NARX: le seconde hanno questo parametro al di sotto dello zero di almeno due ordini di grandezza mentre la prima solo di uno. La differenza è notevole, eppure, la rete neurale *wide* ha condotto fino ad ora a risultati migliori.

Lo stesso processo di media si può condurre impiegando la rete neurale *wide* e le reti di tipo *narrow*, *medium*, *bilayered* e *trilayered*, fin ora tralasciate poiché dall'applicativo *Regression Learner* si era selezionato il modello con un valore di *RMSE* minore, come suggerito dallo stesso produttore. Tuttavia, anche le altre reti neurali sopra citate conducono a previsioni similari a quella del modello *wide*, per cui la manipolazione matematica potrebbe condurre a ottenere un unico andamento più accurato.

Come nel caso precedente, sono state eseguite le singole previsioni, nell'ordine indicato dalla numerazione con cui le stesse sono riportate nelle rappresentazioni di seguito, per poi giungere al risultato finale (Figura 3.4.1-21), eseguendo la media istante per istante dei valori numerici ottenuti da ciascun modello. In quest'ultima immagine sono stati tracciati anche dei riquadri riportanti dei dettagli dell'evoluzione complessiva, al fine di evidenziare alcuni aspetti ritenuti rilevanti. In particolare, nel riquadro viola si mostra il medesimo dettaglio della Figura 3.4.1-15: questo secondo tentativo di smussamento delle imperfezioni della previsione con la rete neurale *wide* non ne ha modificato in maniera rilevante gli aspetti positivi, mantenendo quasi inalterate le forme d'onda negli istanti in cui queste sono più ravvicinate alla realtà. Questo è conseguenza del miglior risultato ottenuto con i modelli analizzati in questo contesto rispetto a quello delle reti NARX.

Nei riquadri giallo e verde, sono mostrati due intervalli temporali in cui la previsione si avvicina alla realtà ma non riesce a seguirne tutte le variazioni essendo queste molto rapide e tra loro ravvicinate. Tuttavia, confrontando quanto accade in questi dettagli con il resto della simulazione, si può evincere che i singoli modelli e, dunque la loro composizione, sono più inclini a seguire degli

impulsi isolati, ossia, preceduti da un intervallo di durata non trascurabile nel quale la grandezza incognita si mantiene più o meno costante.

1: NARROW

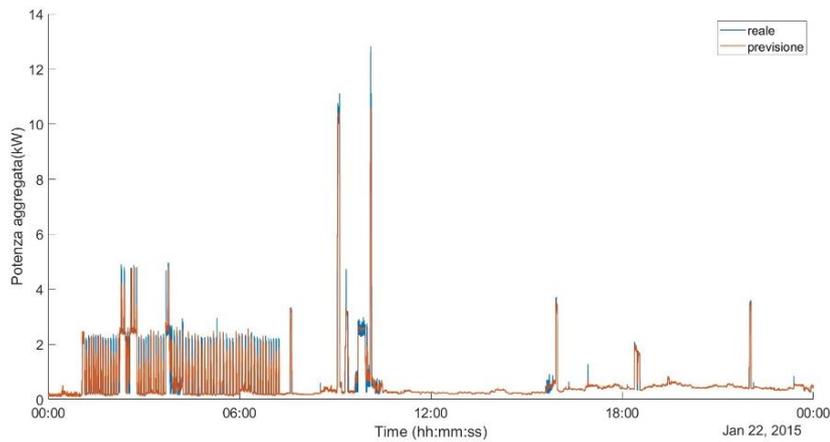


Figura 3.4.1-16 Previsione rete neurale *narrow*, campionamento 8 secondi e 11 *predictor*

2: MEDIUM

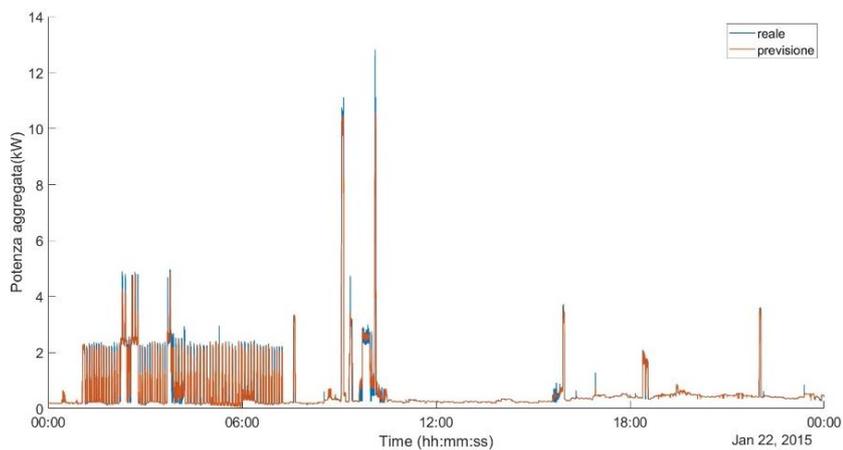


Figura 3.4.1-17 Previsione con rete neurale *medium*, campionamento 8 secondi e 11 *predictor*

3: WIDE

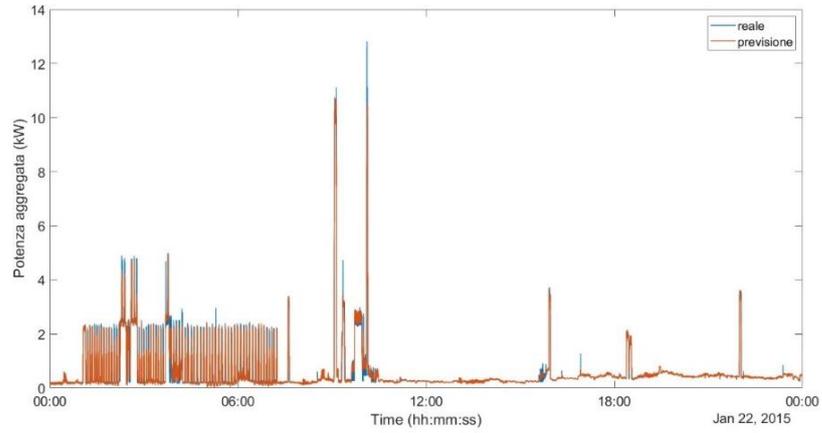


Figura 3.4.1-18 Previsione con rete neurale *wide*, campionamento 8 secondi e 11 *predictor*

4: BILAYRED

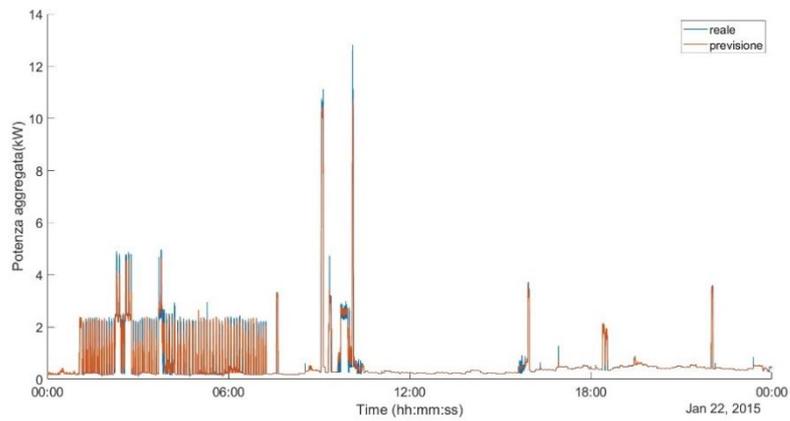


Figura 3.4.1-19 Previsione con rete neurale *bilayered*, campionamento 8 secondi e 11 *predictor*

5: TRILAYERED

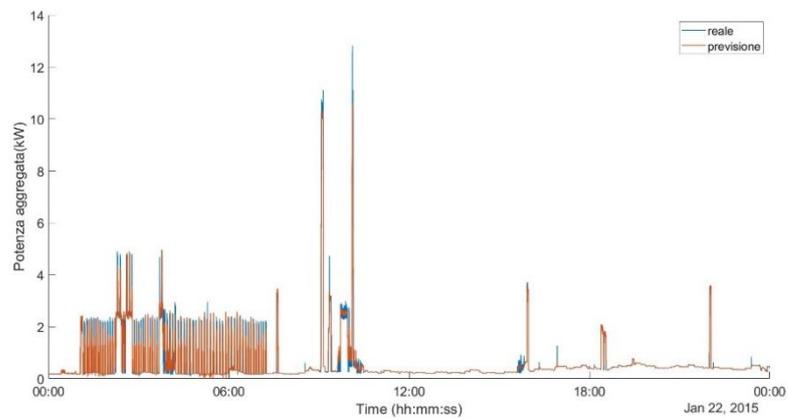


Figura 3.4.1-20 Previsione con rete neurale *trilayered*, campionamento 8 secondi e 11 *predictor*

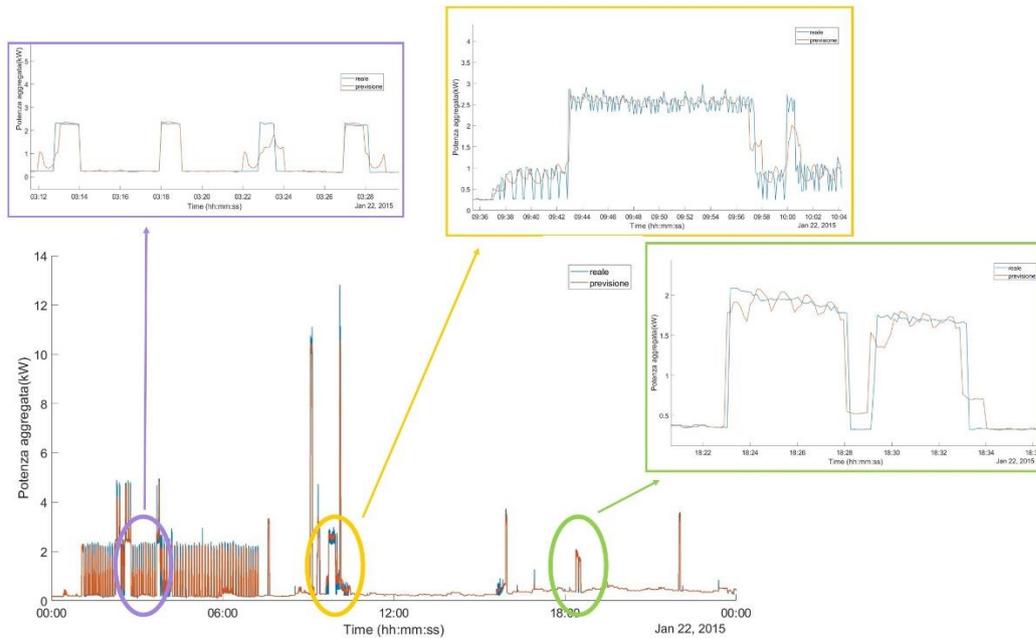


Figura 3.4.1-21 Confronto tra potenza aggregata reale e prevista con media dei risultati dei modelli precedenti, campionamento 8 secondi e 11 predictor

Inoltre, la previsione finale ottenuta evolve tra gli estremi raggiunti dalla potenza aggregata quando essa oscilla (ingrandimento nel riquadro giallo), mentre si osservano variazioni quando la grandezza incognita si mantiene costante (finestra verde). Questo conduce a pensare a un possibile miglioramento del risultato finale filtrando la risposta ricostruita, in modo da renderne meno oscillatoria l'evoluzione. Eseguendo questa operazione, si raggiunge il risultato in Figura 3.4.1-22.

In questa figura sono rappresentati in arancione l'andamento non filtrato, in blu marcato quello ottenuto dall'ultima operazione matematica ed in blu sottile la potenza aggregata reale. Dalla sola immagine rappresentante le 24 ore totali non si riesce a distinguere chiaramente questi colori poiché la linea più spessa sovrasta le altre; per tale motivo sono stati inseriti anche in questo caso dei riquadri con degli ingrandimenti di alcuni intervalli.

Nelle finestre gialla, viola e verde si riportano gli stessi dettagli visti in Figura 3.4.1-21. Nel primo caso, si vede come la linea più marcata non ha le stesse oscillazioni di quella arancione; sebbene quest'ultima tenta di seguire l'evoluzione molto variabile della grandezza campionata, l'errore commesso dal risultato dell'operazione di filtro conduce ad errori minori, poiché comunque più vicino a quanto si vuole ottenere. Le stesse conclusioni si possono trarre per la finestra in verde;

per quanto riguarda il riquadro viola, la situazione si è mantenuta inalterata rispetto a quanto si aveva prima dell'applicazione del filtro.

Rispetto alla Figura 3.4.1-21, sono stati aggiunti altri due dettagli: in arancione viene riportato il riquadro contenente un ingrandimento di un tratto in cui la potenza reale è poco variabile, con micro-oscillazioni trascurabili. In questo caso, il filtro ha condotto ad ottenere errore praticamente nullo, inferiore ad 1 W, migliorando la situazione già ottimale cui si era giunti con il processo di media. Nella finestra ciano si propone il dettaglio di un altro impulso, per il quale le considerazioni fatte per i riquadri giallo e verde sono nuovamente valide.

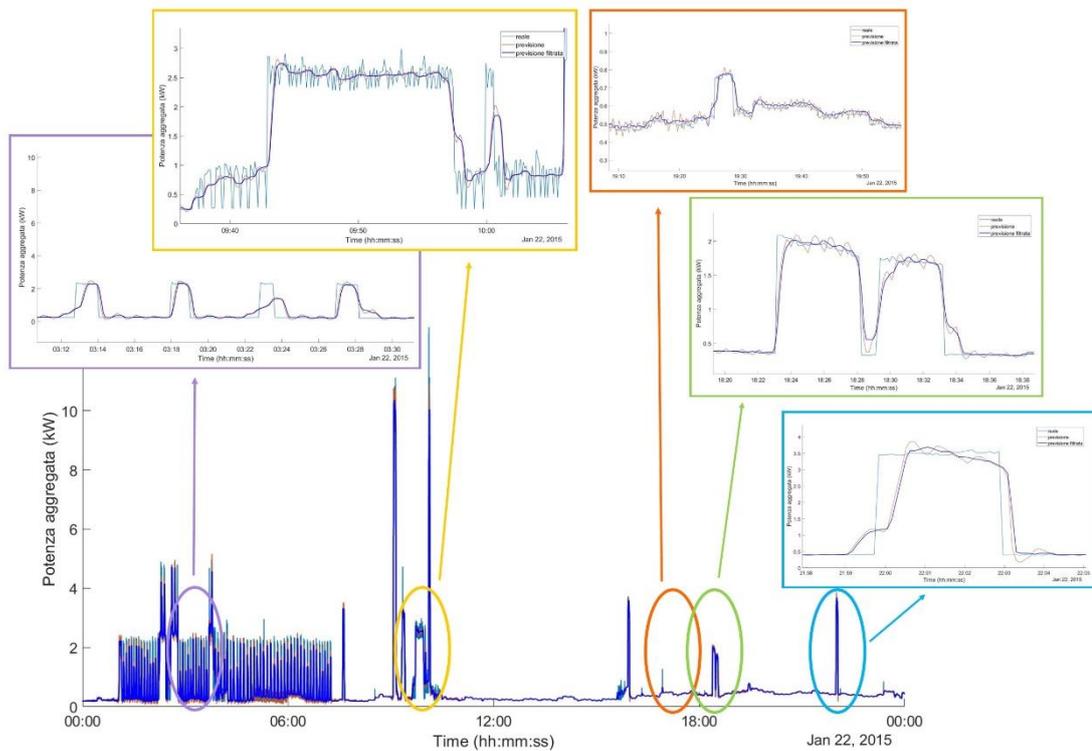


Figura 3.4.1-22 Potenza prevista ottenuta per media dei modelli precedenti filtrata

3.5 CONCLUSIONI DELLA FASE 1

Per riassumere quanto ottenuto dalle numerose simulazioni condotte, delle quali si sono riportati parzialmente i risultati raggiunti, si può affermare che l'applicativo *Regression Learner* fornito dalla *MathWorks* è un mezzo potente che si adatta solo in parte alle esigenze del problema proposto:

infatti, fornendogli le corrette informazioni, esso riesce a calcolare delle risposte non del tutto scorrette. Lo stesso non si può affermare per le reti neurali NARX poiché, nonostante la risposta da esse fornita richiami l'evoluzione della grandezza reale, la loro accuratezza è molto minore rispetto a quanto riscontrato con gli altri modelli analizzati.

Si ritiene inoltre che si potrebbero ottenere risultati di gran lunga più soddisfacenti se fosse resa disponibile l'evoluzione di un'altra grandezza, indipendente dalla potenza aggregata, nei medesimi istanti di quest'ultima. Questa sarebbe, infatti, come un *predictor* molto più affidabile degli altri, tenuto conto che, eccetto i *predictor* relativi all'identificazione dell'istante temporale del campionamento, gli altri si sono ricavati operando una media sul minuto e sull'ora. Questo implica due importanti limitazioni: in primo luogo è necessario conoscere come si evolvano queste medie in momenti in cui non si conoscono i valori istantanei; in secondo luogo, il valore medio sul minuto sarà identico per tutti gli istanti che rientrano in tale minuto, così come accade per la media sull'ora.

Dunque, sebbene i *predictor* aggiunti nel paragrafo 3.1.1 abbiano condotto a una migliore prestazione dei modelli, non si possono ritenere efficaci quanto lo sarebbe un *predictor* definito in ciascuno degli istanti in cui è stata campionata la potenza aggregata; solo in questo modo si potrebbe identificare univocamente quel momento conducendo il modello ad una minore confusione tra gli istanti che compongono il minuto o l'ora.

Sebbene si potrebbe pensare all'impiego dei campionamenti delle altre potenze relative a singoli dispositivi, si ritiene che questa non sia una soluzione applicabile. Infatti, i dispositivi sono azionati in maniera del tutto casuale nelle varie giornate e nulla garantisce che la misura ad essi correlata sia disponibile quando non lo è quella del carico completo.

Infine, considerando che la media dei risultati tra i vari modelli, siano quelli della *Regression Learner App* o le reti NARX, non ha condotto a particolari risultati se non ad un allungamento dei tempi di simulazione, si è proceduto con il solo impiego della rete *wide* con 50 neuroni, con la quale si è giunti a forme d'onda ragionevoli.

4. PERIODICITÀ E PREDICTOR MANCANTI

4.1 RICERCA DI PERIODICITÀ

4.1.1 PERIODICITÀ GIORNALIERA E SETTIMANALE

Tutti i calcoli precedenti hanno permesso di comprendere che la previsione non può essere svolta solo in funzione dei *predictor* descrittivi degli istanti temporali nei quali i campionamenti sono stati svolti, ma è richiesta qualche informazione più dettagliata, quali le potenze medie sui minuti e sulle ore. Tuttavia, è chiaro che tali grandezze medie non sono note nella giornata incognita poiché in essa si suppone la totale assenza di misure. Per questo, si propone di costruire i *predictor* P_mean_min e P_mean_hours , facendo riferimento alle giornate precedenti a quella incognita, tentando di trovare una ripetitività che aiuti almeno in parte a trovare quegli andamenti caratteristici di ogni giornata. Da questa operazione ci si aspetta una perdita di informazioni in questo passaggio, soprattutto per quanto riguarda quei picchi molto stretti ed elevati che rappresentano l'utilizzo temporaneo di qualche dispositivo da parte degli abitanti della casa considerata. Tanto l'utilizzo dei dispositivi interni all'abitazione è ripetitivo nei giorni, tanto migliore sarà la ricostruzione anche in assenza di misure.

Per chiarire il concetto, si presenta l'esempio del frigorifero e della lavatrice. Il frigorifero rappresenta un dispositivo presumibilmente sempre in funzione, per cui il suo contributo al carico aggregato sarà sempre presente, indipendentemente dalla giornata. Lo stesso non si può affermare per un elettrodomestico come la lavatrice, la quale influenza la potenza complessivamente impiegata solo quando viene attivata; la sua attivazione, tuttavia, non è una costante nel tempo. Ma, se si suppone che le abitudini degli occupanti della casa siano ripetitive, non è scorretto pensare che il dispositivo lavatrice sia impiegato in fasce orarie e giorni più o meno invariati nel tempo. Se per esempio, la famiglia occupante sfruttasse la differenziazione del costo dell'energia con le fasce orarie, in linea del tutto ipotetica la lavatrice sarebbe azionata solo nelle ore notturne. Lo stesso

ragionamento si può condurre se le persone in questione lavorano tutta la settimana e quindi sfruttano i giorni del weekend per lo svolgimento di alcune attività quali la pulizia dei panni.

Dunque, si può procedere secondo due strade. La prima opzione è quella di ipotizzare una ripetitività nel carico tra due giorni adiacenti, ossia, la routine “energetica” si ripete quotidianamente: questo implica che le stesse attività vengano svolte tutti i giorni allo stesso modo. La seconda via è quella di guardare al giorno della settimana, supponendo che la ripetitività sia sui sette giorni; in tale situazione, più realistica, si immagina che quello che accade in un determinato giorno della settimana non sia uguale al giorno precedente ma al medesimo giorno della settimana antecedente. In questa maniera, i carichi che si ripetono giorno per giorno (vedasi il frigorifero) rimangono comunque conteggiati ma si tiene in conto che altri dispositivi (vedasi lavatrice) non siano utilizzati ogni giorno ma con una periodicità maggiore.

Indipendentemente dalla strada seguita, il concetto è ricavare i due *predictor* mancanti (che sono anche quelli più influenti sui risultati della simulazione) da quanto accade nei giorni precedenti all’incognito, come schematizzato in Figura 4.1.1-1.

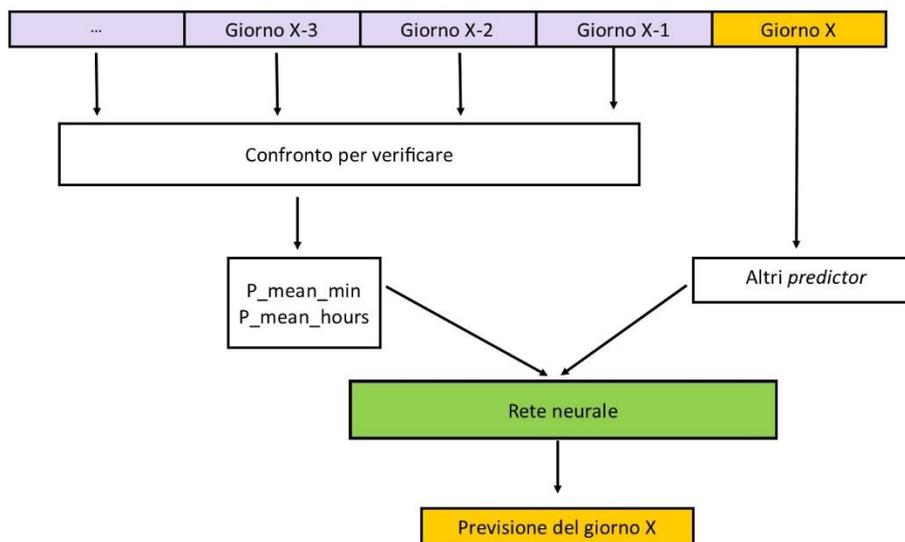


Figura 4.1.1-1 Schema di principio per identificazione *predictor* mancanti dai giorni precedenti

Il giorno X evidenziato in giallo rappresenta il giorno incognito, mentre in lilla sono indicati giorni ad esso precedenti; questi ultimi sono analizzati e studiati per valutare le periodicità e determinare quanto richiesto. Una volta a disposizione tutti i *predictor*, essi vengono utilizzati nella rete neurale per la previsione del giorno X che, nel caso in analisi, è il 22 gennaio.

4.1.2 CONFRONTO DEI GIORNI PRECEDENTI

Alla base della strategia di azione, c'è l'ipotesi di ripetitività nei giorni dei consumi interni all'abitazione, poiché, questa implica che nella giornata incognita non ci siano particolari motivazioni per presupporre un comportamento differente da quanto osservato nelle giornate precedenti. Per tale ragione, in questa sezione si rappresentano le potenze aggregate dei giorni costituenti la settimana precedente al 22 gennaio, in modo da individuarne le somiglianze negli sviluppi.

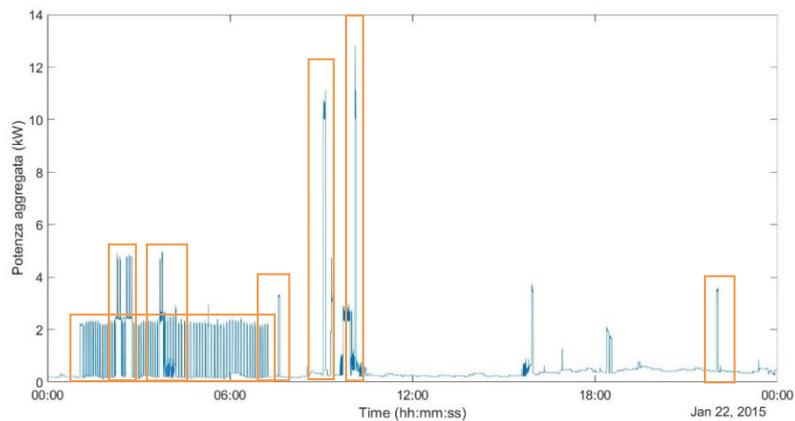


Figura 4.1.2-1 Potenza aggregata reale 22 gennaio 2015

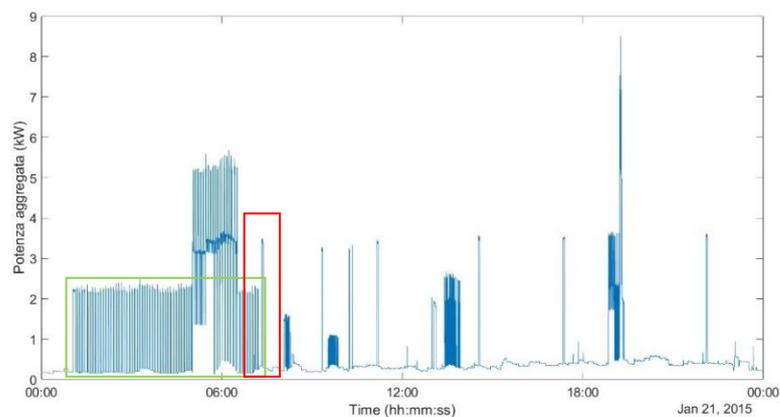


Figura 4.1.2-2 Potenza aggregata reale 21 gennaio 2015

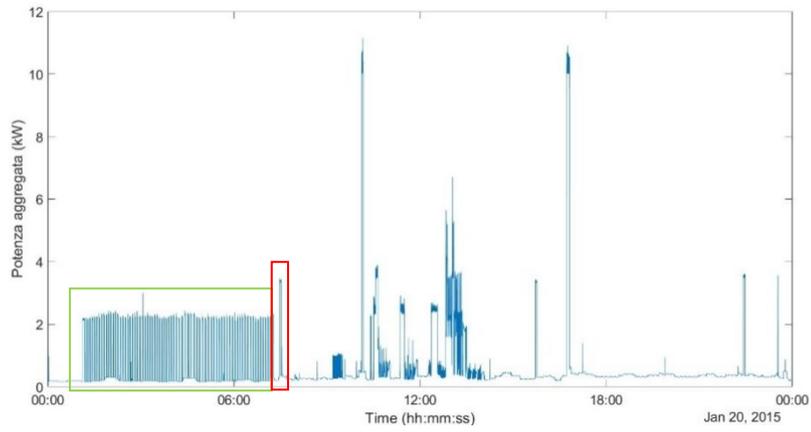


Figura 4.1.2-3 Potenza aggregata reale 20 gennaio 2015

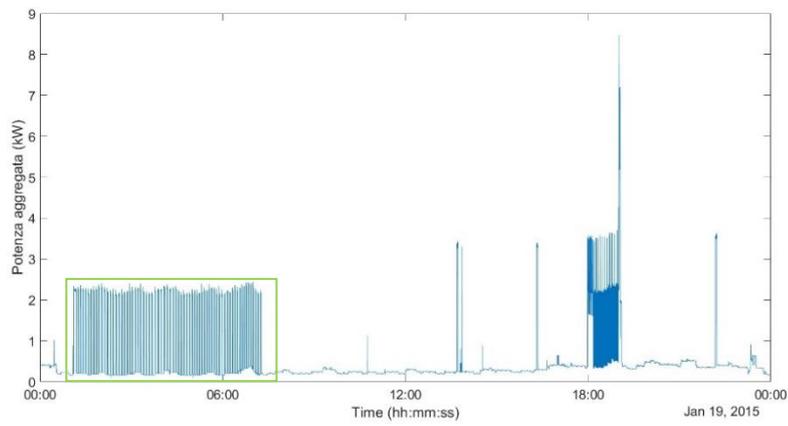


Figura 4.1.2-4 Potenza aggregata reale 19 gennaio 2015

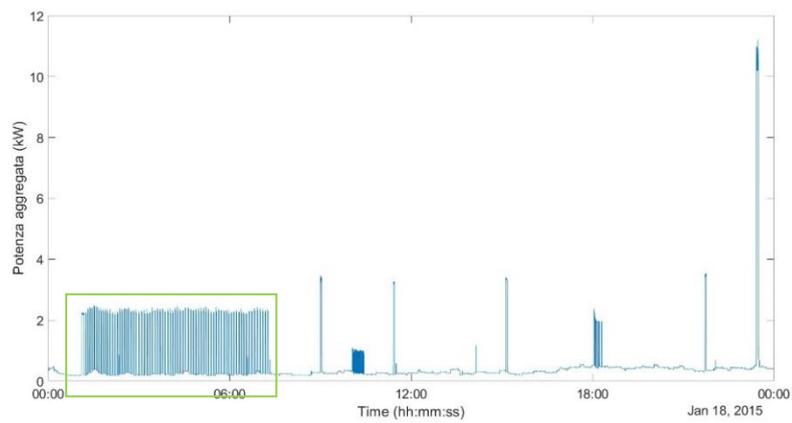


Figura 4.1.2-5 potenza aggregata reale 18 gennaio 2015

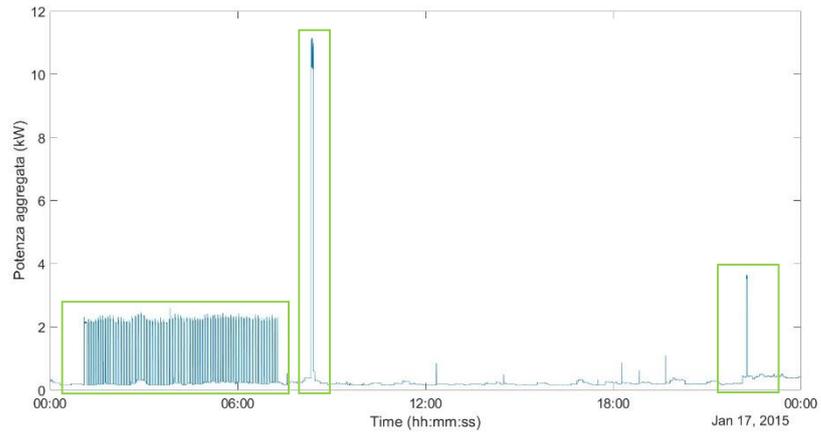


Figura 4.1.2-6 Potenza aggregata reale 17 gennaio 2015

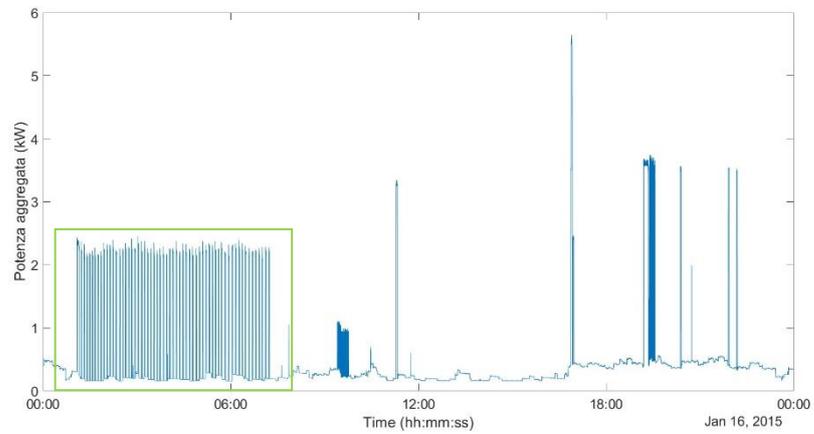


Figura 4.1.2-7 Potenza aggregata reale 16 gennaio 2015

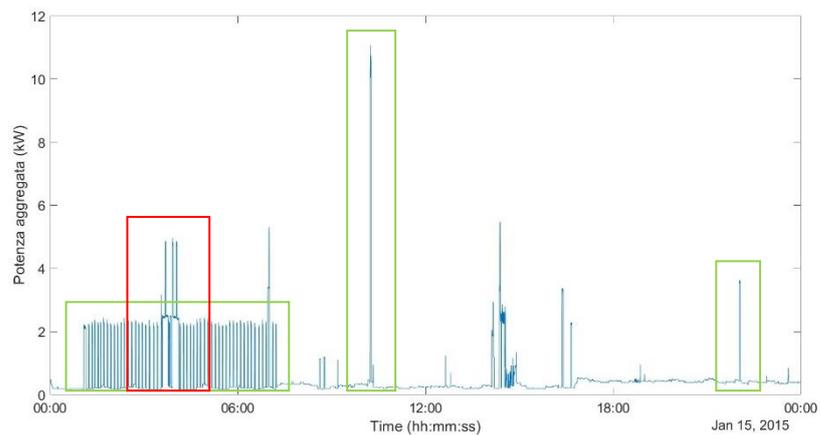


Figura 4.1.2-8 Potenza aggregata reale 15 gennaio 2015

Per rendere più evidenti le ripetizioni, nelle precedenti rappresentazioni sono stati aggiunti dei rettangoli colorati. In particolare: in arancione, sono riquadrati gli istanti nella giornata incognita nei quali l'andamento della potenza aggregata è stato riscontrato in altre giornate; in verde, sono

evidenziati gli andamenti che sono rimasti invariati nel passaggio da un giorno all'altro, con uno scarto di qualche minuto tra l'uno e l'altro; in rosso sono stati messi in luce quegli sviluppi che si possono riscontrare anche nella giornata incognita ma in istanti differenti.

Dunque, la forma d'onda della potenza aggregata nelle ore notturne si ripete sempre, con la sovrapposizione di alcune oscillazioni dovute al probabile utilizzo di qualche dispositivo che, in quelle ore, è normalmente spento; infatti, sono più numerosi i giorni in cui la successione di impulsi si ripete inalterata per tutta la notte rispetto a quelli in cui ci sono degli incrementi. Un'altra particolarità che si ripete è il piccolo cuneo nelle ore serali: anche questo è presente nella maggioranza delle giornate.

Chiaramente non si sono evidenziati i tratti con potenza poco variabile, poiché essi rimangono inferiori a 0.2 kW nelle prime ore del giorno per poi crescere a valori attorno a 0.4 kW in tutte le giornate considerate. Tuttavia, agli sviluppi sopra mostrati si vuole aggiungere anche quello dell'8 gennaio 2015, in quanto esso cade di giovedì, così come il giorno incognito, e pertanto è possibile che compaiano ulteriori somiglianze. Mantenendo la stessa convezione dei riquadri colorati sopra esposta, in Figura 4.1.2-9 è mostrato l'andamento della potenza aggregata dell'8 gennaio con l'evidenziazione delle somiglianze riscontrate.

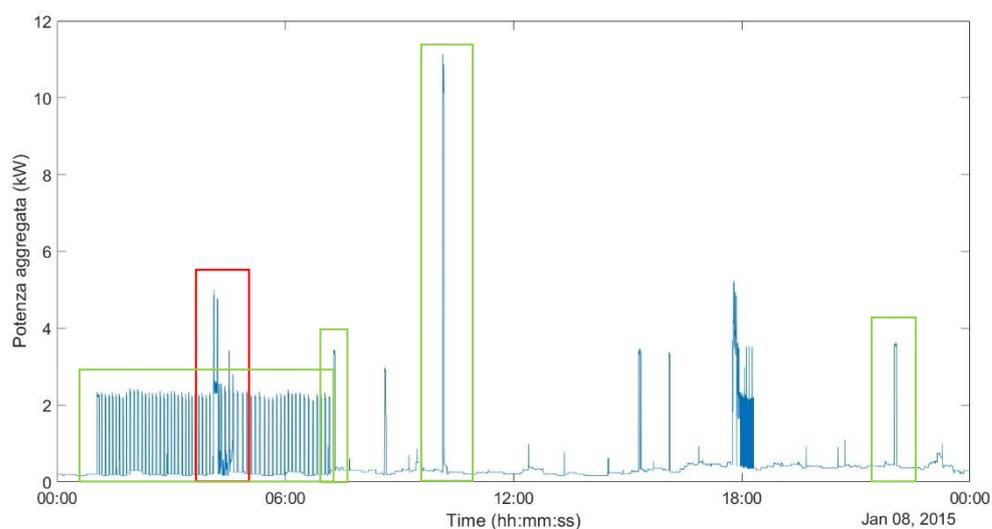


Figura 4.1.2-9 Potenza aggregata reale 8 gennaio 2015

4.1.3 ADATTAMENTO DEI DATI

Al fine di poter svolgere la previsione impiegando i valori medi ricavati dalle misurazioni di un giorno differente, indipendentemente da quale esso sia, si richiede l'adattamento dei dati disponibili poiché gli istanti di campionamento nei differenti giorni non sono coincidenti tra loro. Infatti, è anche possibile che alcuni minuti, oppure ore, presenti il 22 gennaio non ci siano nella giornata impiegata per la valutazione dei *predictor* mancanti.

Quindi si è proceduto con i seguenti passi:

- Inizialmente si identifica l'ora della i -esima riga relativa alla caratterizzazione del giorno 22 gennaio e si verifica la presenza di questa ora anche nella giornata nota;
- Se la verifica porta a un risultato positivo, si procede con lo stesso metodo per la ricerca del minuto in quell'ora, ossia, si identifica il minuto nell'ora della i -esima riga e lo si ricerca tra i minuti della medesima ora nel giorno considerato per il calcolo;
- Se anche il passo precedente si conclude con successo, allora si impone il valore della i -esima riga del vettore P_mean_min e P_mean_hours del giorno 22 gennaio pari a quelli del giorno noto nel medesimo istante.
- Se uno dei passi precedenti dovesse fallire, a causa dell'assenza dell'istante i -esimo nel giorno campionato, si procede imponendo come P_mean_min e P_mean_hours i valori assunti da tali parametri in corrispondenza degli intervalli temporali nei quali la grandezza incognita è poco variabile. L'osservazione delle forme d'onda sopra mostrate ha condotto alla scelta della prima misurazione come valore da imporre in queste situazioni; esso è poco variabile nel passaggio da un giorno al successivo e anche durante la giornata.

Questo procedimento è stato seguito per le seguenti simulazioni, sia che esse usino come riferimento il giorno precedente o il medesimo giorno della settimana antecedente.

Nei paragrafi successivi si mostra quanto ottenuto dalle due possibili vie sopra esposte.

4.1.4 METODO 1: GIORNO PRECEDENTE

Il giorno da considerarsi per l'obiettivo è il 21 gennaio 2015, nel quale la potenza aggregata ha lo sviluppo mostrato in Figura 4.1.2-2. Dal confronto con la Figura 4.1.2-1, si vede chiaramente la differenza tra la giornata incognita e la giornata nota in gran parte delle ventiquattro ore. Tuttavia, si vuole comunque provare a verificare il risultato di una previsione che usi i valori medi del 21 gennaio come *predictor*, mettendosi nella condizione in cui il 22 gennaio sia effettivamente incognito, così da valutare l'efficacia dell'approccio proposto.

Quello che si ottiene con la rete neurale *wide* applicata ai *predictor* modificati è raffigurato in arancione in Figura 4.1.4-1, insieme all'andamento (in blu) della potenza aggregata campionata nel giorno incognito. Inoltre, la Figura 4.1.4-2, la Figura 4.1.4-3 e la Figura 4.1.4-4 contengono uno *zoom* di alcuni istanti della medesima previsione, ritenuti di maggiore interesse rispetto agli altri.

Come si era già previsto, tutti i picchi successivi alle ore 7:30 sono completamente ignorati, sostituiti da altre oscillazioni, compatibili con quanto è avvenuto nel 21 gennaio ma assolutamente inesistenti nella giornata incognita. In tale situazione, in assenza di altre informazioni, si potrebbero ignorare completamente le previsioni di tali oscillazioni, poiché esse non sono neanche riscontrabili nel resto della settimana analizzata nei paragrafi precedenti. La stessa logica si può applicare alle oscillazioni ottenute nell'intorno delle ore 6:00, quando negli altri giorni della settimana la potenza aggregata in tale intervallo temporale manteneva l'andamento ripetitivo che si manifesta in tutta la notte.

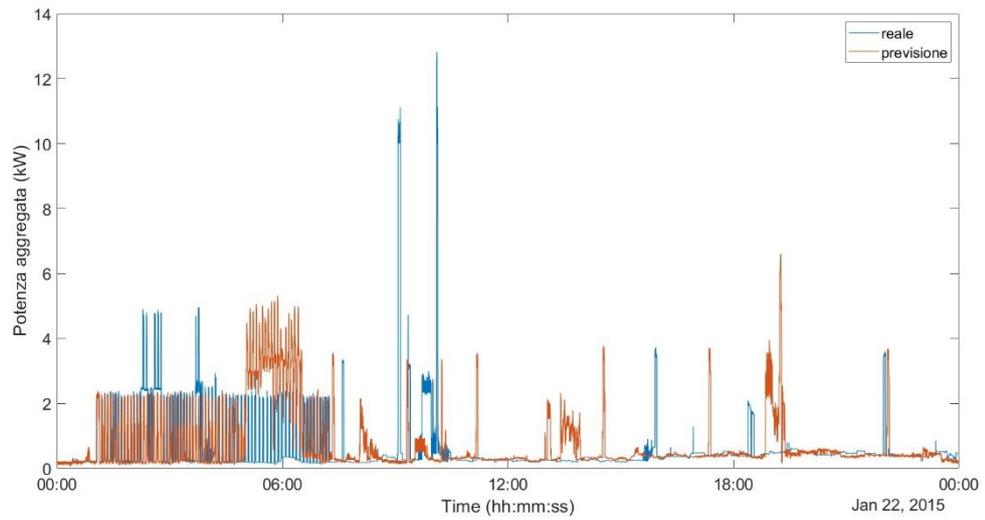


Figura 4.1.4-1 Confronto tra potenza aggregata realtà e previsione con *predictor* ottenuti dal 21 gennaio 2015

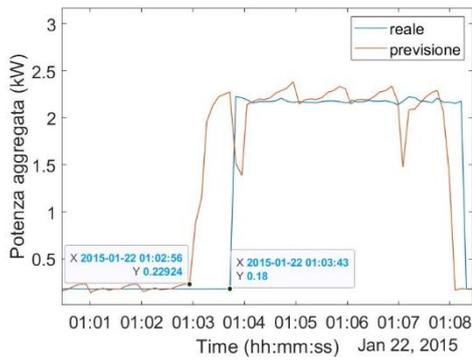


Figura 4.1.4-2 Dettaglio ore notturne della previsione in Figura 4.1.4-1

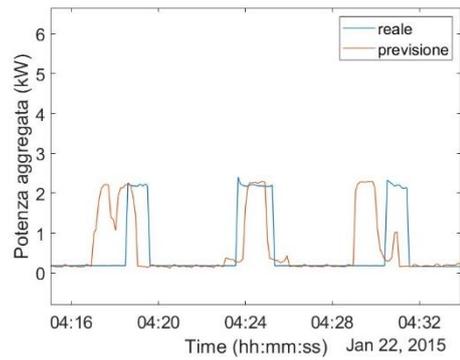


Figura 4.1.4-3 Dettaglio ore notturne della previsione in Figura 4.1.3 1

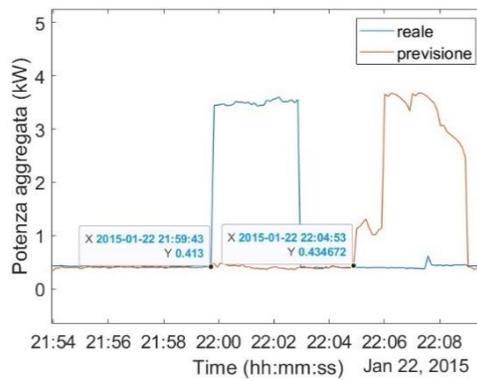


Figura 4.1.4-4 Dettaglio ore serali della previsione in Figura 4.1.3 1

Tuttavia, nella simulazione condotta, ci sono degli aspetti positivi che si possono osservare negli ingrandimenti sopra riportati.

In particolare, in Figura 4.1.4-2 si è evidenziato il primo impulso delle ore notturne: la ricostruzione manifesta tale variazione circa un minuto prima di quella reale, ma ne raggiunge il valore nei momenti successivi. Con il passare dei minuti, la previsione oscilla leggermente, esattamente come avveniva con le simulazioni che impiegavano i *predictor* corretti e, quindi, è possibile livellare questo difetto filtrando la risposta o sostituendola con una più realistica.

Procedendo oltre, in Figura 4.1.4-3 viene messo in luce la coincidenza tra l'andamento reale e quello previsto in un altro momento delle ore notturne; questo accade anche se la prima oscillazione della previsione è in anticipo rispetto a quella reale, come spiegato sopra. Dunque, l'accuratezza delle previsioni dipende ancora dall'istante analizzato, indipendentemente dai *predictor* e dal modello impiegato.

Infine, in Figura 4.1.4-4 è tracciata la potenza aggregata alle ore 22:00, ossia nell'orario in cui nella maggioranza dei giorni della settimana si verifica un salto di circa 3.5 kW. La simulazione svolta prevede una variazione di simile entità appena qualche minuto dopo, come espresso dai riquadri riportanti valore della grandezza e istante di valutazione. Questo conduce a pensare che le due oscillazioni, sebbene traslate di un breve intervallo, siano associate allo stesso evento; quindi, si potrebbe comunque prevedere questo impulso valutando in quali istanti temporali esso si manifesta nei sette giorni precedenti al 22 gennaio e considerando come plausibile l'istante medio tra quelli rilevati.

4.1.5 METODO 2: MEDESIMO GIORNO DELLA SETTIMANA PRECEDENTE

Il 22 gennaio 2015 corrisponde ad un giovedì, quindi, per effettuare questa simulazione, si è impiegato quanto si verifica il 15 gennaio 2015, medesimo giorno di quello incognito ma appartenente alla settimana precedente. Attuato l'adattamento dei dati e l'identificazione dei *predictor* esposta nel paragrafo 4.1.3, si è proceduto all'applicazione del modello *wide* a quanto ottenuto, giungendo al risultato rappresentato in Figura 4.1.5-1. Alcuni dettagli sono tracciati a seguire.

La prevalenza del colore blu nella figura relativa all'intera giornata prevista, suggeriscono che spesso la ricostruzione non è corretta e che la curva arancione si discosti da quella reale maggiormente di quanto avviene con il metodo 1. Tuttavia, è meglio attenzionare i dettagli per mettere in luce eventuali aspetti positivi, presenti nonostante una prima apparente inadeguatezza del risultato.

In Figura 4.1.6-2 sono segnalati i valori di potenza aggregata e gli istanti ad essa corrispondenti, sia per la previsione che per il reale andamento. Considerando il primo impulso previsto, alle ore 1:03:57, esso si discosta di appena 21 secondi da quello reale mentre, il secondo, alle ore 1:04:50, di poco più di un minuto. Quindi, nuovamente, questa previsione si può impiegare per l'identificazione dell'istante in cui quel carico notturno ripetuto tutti i giorni comincia il suo funzionamento, sia basandosi sulla prima oscillazione sia sulla seconda; infatti, in ogni caso il ritardo sarebbe minimo. In entrambi i casi, viene raggiunto il pianoro della potenza reale, attorno al quale si registrano piccole variazioni.

In Figura 4.1.6-3 si è voluto porre un accento particolare su quanto avviene nell'ultimo quarto d'ora delle ore 3:00. Nell'andamento reale (blu), a partire dalle 3:40:50, si manifesta un gradino più ampio con degli aumenti limitati a brevissimi intervalli di tempo; un'evoluzione simile la si può osservare anche nella forma d'onda arancione, ma traslata di un quarto d'ora rispetto alla precedente. Questo ritardo, non lo si ritiene accettabile sebbene un'analisi più superficiale, basata

sulla Figura 4.1.6-1, condurrebbe a concludere che le due rappresentazioni siano tra loro sovrapposte nell'intervallo temporale in questione.

Tuttavia, la presenza di un'evoluzione molto simile a quella reale anche nella previsione, sottolinea ulteriormente l'importanza e il valore nella considerazione della periodicità del carico, che si manifesta anche in tale occasione, sebbene in momenti tra loro differenti ma non distanti.

In Figura 4.1.6-4 vi è un ingrandimento del picco di potenza di circa 10 kW delle ore 10:00, mentre nella Figura 4.1.6-5 è riportato quanto accade alle ore 22:00. Nel primo caso, il risultato del calcolo è in ritardo di poco più di 7 minuti rispetto al campionamento ma, l'ampiezza e la durata del cuneo di potenza è invariata. Nel secondo caso, il ritardo della previsione è inferiore al minuto e nuovamente essa raggiunge il valore di circa 3.5 kW, conducendo ad un miglioramento rispetto a quanto ricavato nella precedente simulazione.

Confrontando questo metodo con il precedente, in complesso si può parlare di un miglioramento poiché, sebbene numerose siano le imprecisioni, il risultato della simulazione presenta un maggior numero di punti di incontro con la realtà. Questo conduce a pensare ad una maggiore probabilità di ripetitività settimanale, più che giornaliera.

Nuovamente, non conoscendo ipoteticamente quanto avviene il 22 gennaio, ossia, trascurando le curve blu nelle figure sottostanti, le oscillazioni della linea arancione che non compaiono in altre giornate verrebbero comunque trascurate. Quindi si terrebbero solo quelle delle ore notturne (quando ripetitive), delle ore 10:00 e delle ore 22:00, saturando il resto al minimo valore costante assunto dalla grandezza incognita.

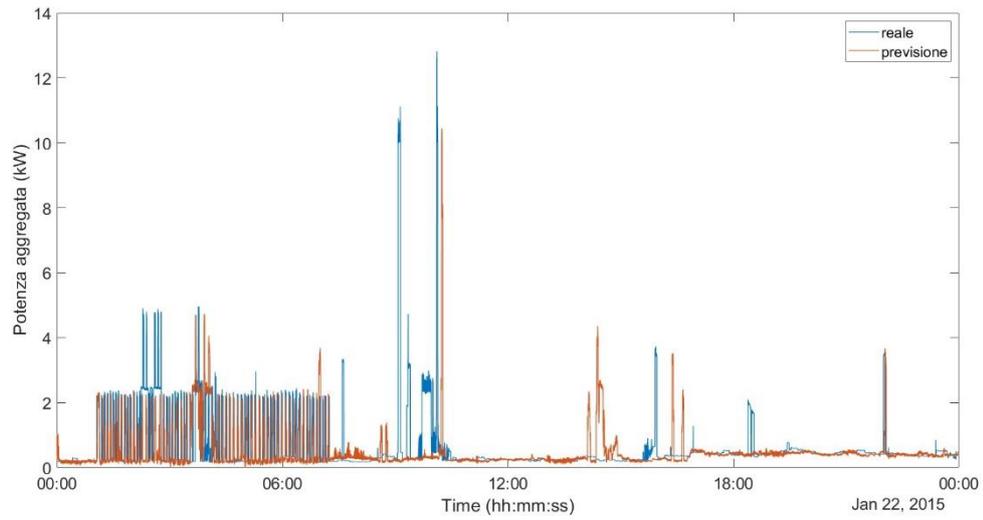


Figura 4.1.5-1 Confronto tra potenza aggregata realtà e previsione con *predictor* ottenuti dal 15 gennaio 2015

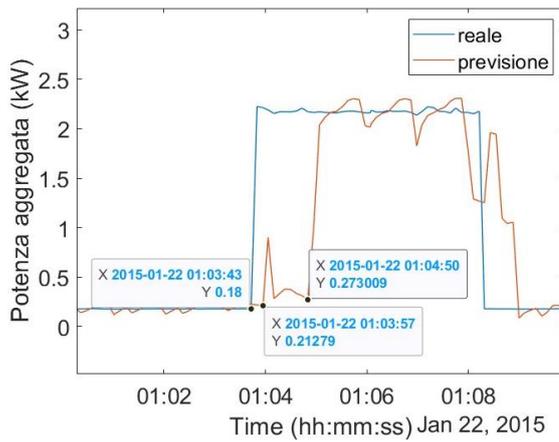


Figura 4.1.5-2 Dettaglio ore notturne della previsione in Figura 4.1.5-1

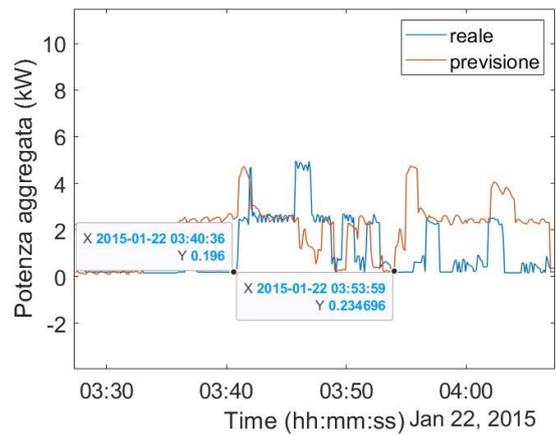


Figura 4.1.5-3 Dettaglio ore notturne della previsione in Figura 4.1.4 1

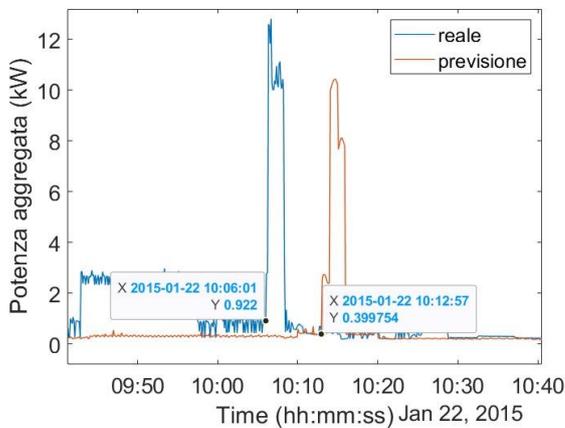


Figura 4.1.5-4 Dettaglio ore diurne della previsione in Figura 4.1.4 1

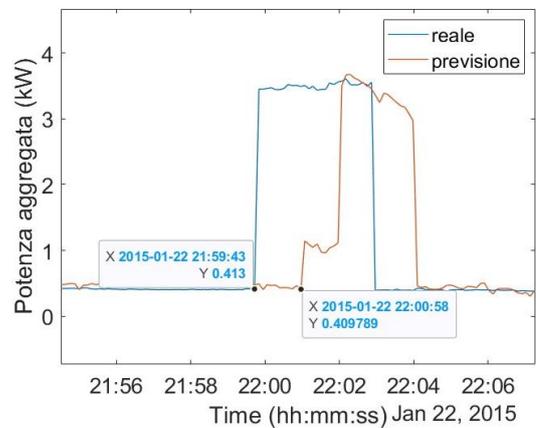


Figura 4.1.5-5 Dettaglio ore serali della previsione in Figura 4.1.4 1

4.1.6 METODO 3: MEDESIMO GIORNO DI DUE SETTIMANE PRIMA

Per confermare la migliore prestazione con l'impiego dei *predictor* ottenuti dalla stessa giornata appartenente ad una settimana differente, si è proceduto con la ripetizione della simulazione usando i dati relativi all'8 gennaio, ossia, il giovedì di due settimane prima rispetto a quella del 22 gennaio.

Il calcolo ha condotto al grafico di Figura 4.1.6-1, nel quale linea arancione è la previsione definita dal modello mentre la blu è l'andamento reale dell'incognita. Anche in questo caso, numerose sono le differenze tra la due curve che basterebbero a classificare la riproduzione inadeguata. Ma, in effetti, ci sono ottimi spunti da potersi utilizzare per migliorare quanto ottenuto; essi sono approfonditi aiutandosi con gli ingrandimenti in Figura 4.1.6-2, Figura 4.1.6-3, Figura 4.1.6-4 e Figura 4.1.6-5. Eccetto il secondo, gli istanti riportati in questi riquadri sono gli stessi che si erano analizzati nella precedente simulazione e differiscono in parte da quelli della prima. Infatti, negli ultimi due calcoli le somiglianze con la realtà sono aumentate e si manifestano negli stessi momenti della giornata. Questo conferma che l'utilizzo dei carichi da parte degli abitanti si ripete nel passaggio da una settimana all'altra con una probabilità maggiore di quanto accade tra due giorni adiacenti. Si vuole sottolineare in questa sede, che solo alcune somiglianze sono ritenute accettabili, ossia, si sono escluse quelle situazioni nelle quali la riproduzione poteva associarsi ad un'evoluzione reale ma con un anticipo o un ritardo superiore ai 10 minuti, come si è già verificato in Figura 4.1.5-3 8 (mostrata solo per evidenziare la ripetitività dei carichi) e come si evince dal cuneo prima delle ore 18:00 in Figura 4.1.6-1 (non ingrandito perché ritenuto troppo distanze da quello reale).

Per quanto riguarda la potenza aggregata nei primi minuti dopo l'una di notte (Figura 4.1.6-2), non si riscontrano particolari differenze rispetto alla Figura 4.1.6-2: la ricostruzione, questa volta in anticipo sulla realtà, si discosta da quest'ultima di neanche un minuto, quindi, si può nuovamente ritenere identificabile il momento di inizio della successione di impulsi che si verifica ogni notte.

Il secondo ingrandimento (Figura 4.1.6-3) mostra un altro momento delle ore notturne in cui il modello riesce ad ottenere una risposta pressoché coincidente con la realtà: l'istante di inizio del tratto ascendente della linea blu e di quella arancione differiscono si possono considerare sovrapposti, trascorrendo appena qualche secondo tra l'uno e l'altro. La riproduzione evolve poi mantenendosi lievemente al di sotto della grandezza campionata, raggiungendola in un solo istante, prima di decrescere con un piccolo anticipo rispetto alla realtà. Delle considerazioni simili si possono condurre in altri istanti delle ore notturne, come si può vedere anche nella Figura 4.1.6-2 di cui si è sopra discusso; tuttavia, non si sono riportate altre immagini dettagliate che non aggiungerebbero alcuna informazione a quelle già ottenute, essendo la coincidenza tra realtà e previsione del tutto casuale, proprio come avveniva anche per il modello che impiegava i valori medi della giornata incognita stessa.

Le ultime due rappresentazioni, invece, riguardano i picchi delle ore 10:00 (Figura 4.1.6-4) e delle ore 22:00 (Figura 4.1.6-5). In entrambi i casi, la riproduzione è migliorata, con una riduzione degli errori sia in termini di tempo interposto tra riproduzione e realtà sia per quanto riguarda le forme ottenute, sebbene comunque siano evidenti delle oscillazioni e dei disturbi che si possono trattare con ulteriori operazioni matematiche. La riduzione dell'errore temporale è un vantaggio importante tenuto conto che entrambi gli impulsi in questione non hanno una durata molto ampia.

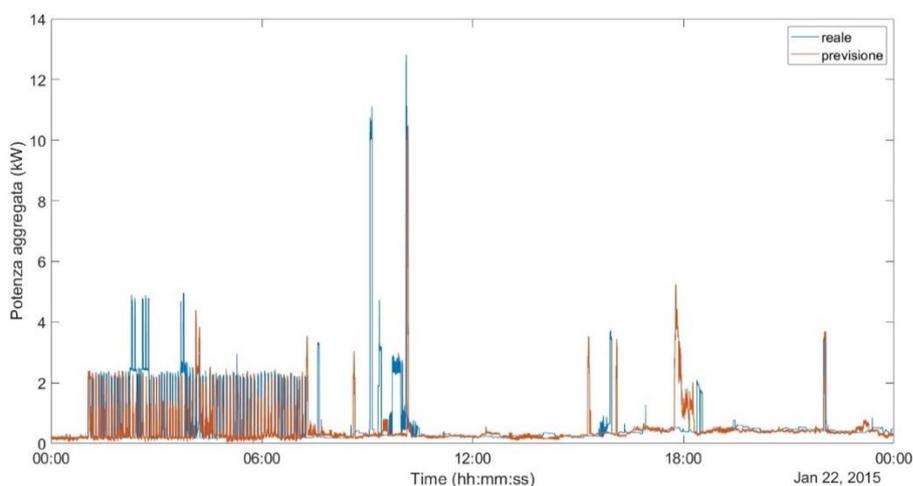


Figura 4.1.6-1 Confronto tra potenza aggregata realtà e previsione con *predictor* ottenuti dal 8 gennaio 2015

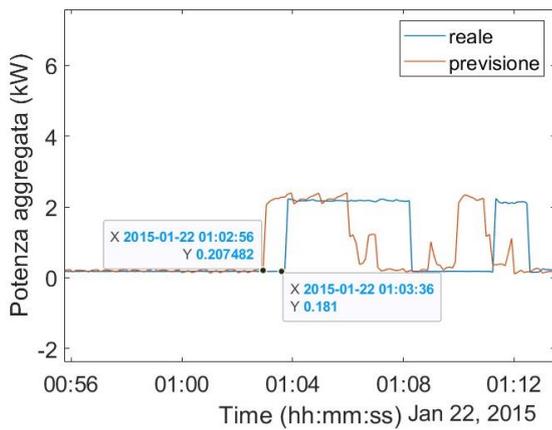


Figura 4.1.6-2 Dettaglio ore notturne della previsione in Figura 4.1.6-1

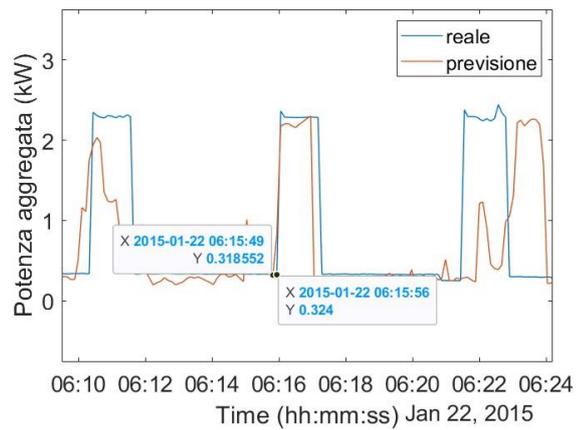


Figura 4.1.6-3 Dettaglio ore notturne della previsione in Figura 4.1.5 1

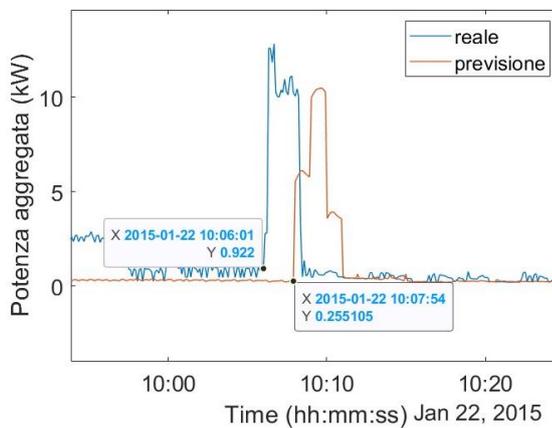


Figura 4.1.6-4 Dettaglio ore diurne della previsione in Figura 4.1.5 1

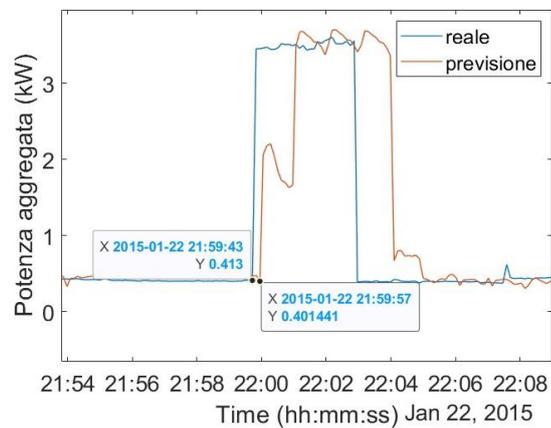


Figura 4.1.6-5 Dettaglio ore serali della previsione in Figura 4.1.5 1

4.2 CONCLUSIONI DELLA FASE 2

Le analisi condotte con l'impiego di *predictor P_mean_min* e *P_mean_hours* ricavati da giorni differenti da quello incognito hanno condotto a concludere alcuni importanti aspetti per l'ottenimento di risultati idonei.

In primo luogo, questo metodo è applicabile solo quando ci sono delle ripetitività dei carichi, situazione in cui è possibile che la rete neurale possa ricordare quanto accaduto nei giorni precedenti e riproporre un andamento più o meno vicino a quello reale. Se è vero che le simulazioni svolte evidenziano l'impossibilità di ottenere una precisione assoluta, è anche doveroso

sottolineare che i risultati ottenuti sono ottimi spunti per identificare istanti e comportamenti utilizzabili per la ricostruzione forzata del segnale, approfondita nel successivo capitolo.

È altresì deducibile con facilità che tutti questi ragionamenti condurrebbero a risultati sicuramente più accurati nel momento in cui gli occupanti dell'abitazione seguissero delle routine ben definite o si conoscessero altre informazioni, più dettagliate, su quanto avviene nella giornata incognita. Per esempio, anche se mancasse la prima condizione di ripetitività ma si conoscesse il programma di consumi che si intende seguire per il giorno successivo, la ricostruzione sarebbe più semplice. Infatti, supponendo che l'incognita sia ancora il 22 gennaio 2015, si immagini che il 21 gennaio si è programmato l'impiego della lavatrice per le ore 12:00 del giorno a seguire; anche se la misura non fosse disponibile e la ricostruzione ottenuta con i *predictor* del giorno 21 non prevedesse nessuna particolare variazione della potenza in quelle ore (come accade effettivamente e si vede in Figura 4.1.4-1), si potrebbe aggiungere manualmente questo consumo a valle di una conoscenza dell'andamento del diagramma di carico di tale dispositivo. Questo aiuterebbe molto nelle previsioni di evoluzioni future, anche perché è lecito affermare che il raggiungimento di una routine è un processo complesso ed utile allo stesso tempo per favorire una migliore gestione delle risorse disponibili.

5. MIGLIORAMENTI DEGLI ANDAMENTI OTTENUTI

5.1 TIPOLOGIA DI CARICHI RIPETITIVI

5.1.1 IDENTIFICAZIONE CARICHI RIPETITIVI

Nella condizione di ritenere valida l'ipotesi di ripetitività della potenza aggregata ad intervalli settimanali, si ignori quanto accade realmente il 22 gennaio 2015 e si isoli quanto ricavato dall'applicazione della rete neurale di tipo *wide*, con l'impiego di un set di *predictor* basati sui valori medi della grandezza incognita al minuto e dell'ora calcolati al 15 gennaio 2015. Quindi, eliminando la curva blu in Figura 5.1.1-1, rimane la sola previsione, come di seguito rappresentato.

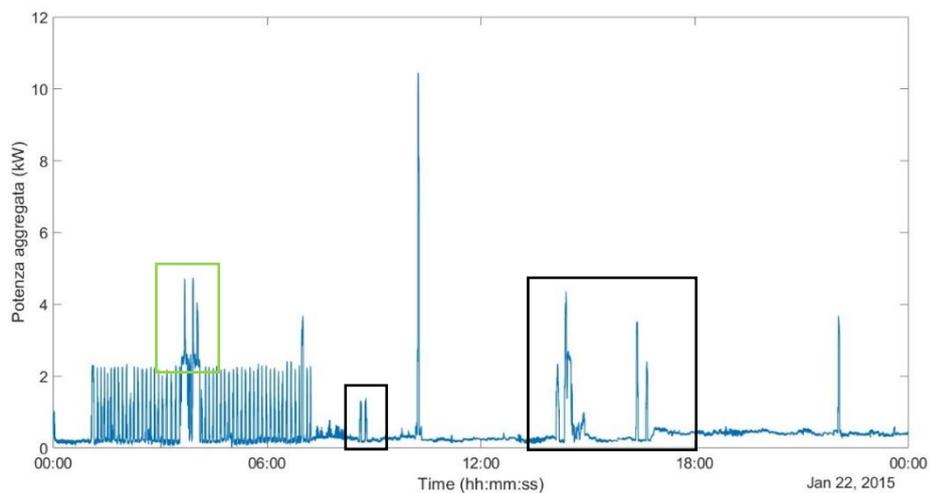


Figura 5.1.1-1 Previsione dell'andamento della potenza aggregata il 22 gennaio 2015 basandosi su quanto avviene il 15 gennaio 2015

In Figura 5.1.1-1 sono riportanti anche dei rettangoli di colore nero che evidenziano dei picchi che non si possono ritenere né corretti né errati; infatti, a valle delle considerazioni fatte nel capitolo precedente riguardanti le evoluzioni della potenza aggregata nei giorni componenti la settimana precedente al 22 gennaio, questi picchi non sono ripetitivi nei giorni e quindi su di essi non è possibile fare nessuna considerazione. Invece, nel rettangolo verde sono messi in evidenza due picchi che spesso sono riscontrabili in altri giorni ma non sempre negli stessi istanti; per questa ragione, anche su questi cunei non si può giungere a una particolare conclusione ritenendola

corretta. Per il resto, in Figura 5.1.1-1 si manifestano dei comportamenti riscontrabili in numerosi giorni del mese di gennaio. Tali comportamenti sono mostrati nelle figure successive, nelle quali viene riportato quanto accade con i campionamenti del 22 gennaio, per comodità, al fine di chiarire quali sono le evoluzioni temporali in questione. In particolare:

- COMPORAMENTO DI TIPO A (Figura 5.1.1-2): nelle ore notturne, ove la potenza aggregata evolve secondo una successione di impulsi;
- COMPORAMENTO DI TIPO B (Figura 5.1.1-3): a seguire delle ore notturne, la potenza spesso presenta un impulso superiore ai precedenti e di simile durata;
- COMPORAMENTO DI TIPO C (Figura 5.1.1-4): a mezzodì, quando si manifesta un picco di potenza di circa 10 kW o superiore;
- COMPORAMENTO DI TIPO D (Figura 5.1.1-5): nelle ultime ore della sera, dove la grandezza studiata cresce per un breve intervallo di tempo prima di tornare al valore che essa assume nel resto della giornata, quando si mantiene più o meno costante.

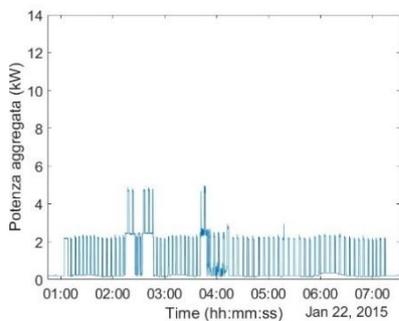


Figura 5.1.1-2 Tipologia di carico A - ore notturne

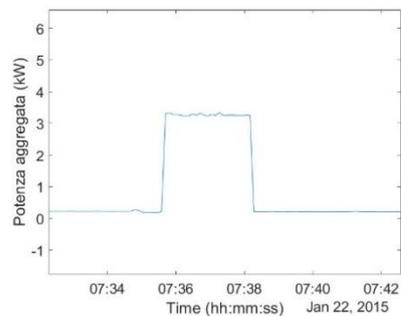


Figura 5.1.1-3 Tipologia di carico B -mattina presto

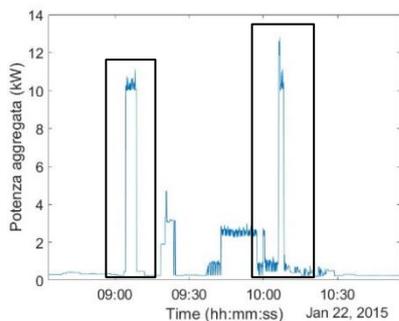


Figura 5.1.1-4 Tipologia di carico C – picchi ore diurne

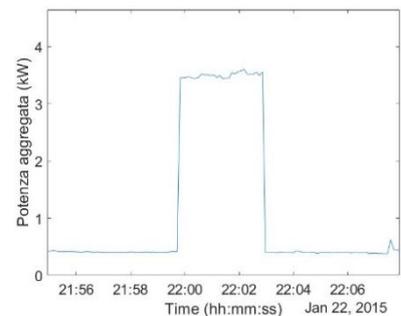


Figura 5.1.1-5 Tipologia di carico D - ore serali

Nella ricostruzione ottenuta non è visibile il comportamento di tipo B ma si possono intuire gli altri tre; quindi, in quegli intervalli, si ha l'utilizzo di alcuni dispositivi che determinano sempre la medesima variazione del carico complessivo.

L'obiettivo che ci si pone, dunque, è quello di ricostruire un diagramma verosimile della potenza richiesta da questi carichi incogniti, studiando i giorni precedenti del mese; esso non sarà ricavato da medie aritmetiche, ma da un processo di ricostruzione passo per passo. Infatti, se si operasse una media tra quanto avviene nei giorni nel mese, si rischierebbe di giungere a risultati non corretti poiché non caratterizzanti il carico reale. Se, per esempio, in 4 giorni il carico si manifesta solo in 2, operare una media condurrebbe ad avere un carico pari alla metà di quello vero. Questo è assolutamente inaccettabile. Il contributo di una sostituzione di forme d'onda si può ritenere valido solo nel momento in cui questa abbia un reale significato, attinente al fenomeno reale.

Una volta ottenuto tale andamento, lo si può sostituire a quelli incerti che il modello restituisce, aspettandosi un miglioramento della previsione, sebbene potrebbero esserci degli sfasamenti temporali.

5.1.2 STEP 1: RICERCA TIPOLOGIA CARICHI RIPETITIVI NEI GIORNI DEL MESE

Il primo passo è ricercare le tipologie di carichi sopracitati in ciascuno dei giorni del mese e verificare che essi, sebbene attivati in momenti differenti, siano sovrapponibili. Questo permetterebbe di identificare univocamente il carico e ricavarne un unico andamento da usare nelle previsioni. In Tabella 5.1.2-1 si sono riassunte le informazioni ricavate dallo studio della potenza aggregata di ciascuno giorno del mese.

Per giungere alle tipologie di carichi, si è proceduto tracciando l'andamento della grandezza incognita e ricercandovi le evoluzioni rappresentate in Figura 5.1.1-2, Figura 5.1.1-3, Figura 5.1.1-4 e Figura 5.1.1-5. A causa delle differenze temporali tra i vari momenti di attivazione del medesimo tipo di carico, nella medesima tabella, in corrispondenza della riga del giorno del mese, si è colorata la casella quando la tipologia di carico è stata rilevata e vi si è trascritto il momento di inizio.

Fatta eccezione per il carico di tipologia A, presente ogni giorno con istante di attivazione praticamente costante, le altre tipologie di carico presentano delle differenze temporali maggiori. Tuttavia, come si approfondirà in seguito, il primo tipo di carico è quello più complesso da riprodurre rispetto agli altri, che invece presentano delle maggiori regolarità che ne semplificano la rappresentazione del comportamento. Si sottolinea che nella Tabella 5.1.2-1 sono stati inseriti dei punti interrogativi nella riga del 22 gennaio poiché è la giornata incognita della quale, ipoteticamente, non è nota la potenza reale.

Tabella 5.1.2-1 Tipologia di carichi per ogni giorno del mese di gennaio 2015

GIORNO	A	B	C	D
1	01:06:56			
2	01:07:26			
3	01:07:39			
4	01:07:50			
5	01:08:31			22:09:50
6	01:08:40	07:22:35	09:46:03	21:57:11
7	01:08:41	07:27:41		21:45:12
8	01:02:58	07:16:11	10:08:22	22:00:30
9	01:03:04	07:38:25	09:51:07	22:06:34
10	01:03:28	07:47:54	09:55:48	21:44:40
11	01:04:13	08:22:22	10:49:39	21:57:40
12	01:04:20	08:12:29		22:02:12
13	01:04:31	07:22:50		22:14:53
14	01:04:39			22:01:50
15	01:04:46			22:01:44
16	01:05:15			22:10:01
17	01:05:24		08:20:44	22:14:58
18	01:06:15			21:42:50
19	01:06:29			22:10:41
20	01:07:12	07:28:16	10:06:49	22:25:09
21	01:03:10	07:19:57		22:05:44
22	?	?	?	?
23	01:04:00	7:43:24	11:54:00	21:56:32
24	01:04:04	07:48:16	10:11:11	22:02:15
25	01:04:43		11:24:58	22:09:53
26	01:05:24			
27	01:05:28	07:29:52	09:05:27	22:08:57
28	07:05:11	07:37:33		22:02:10
29	01:05:31		09:55:09	21:58:35
30	01:03:15	07:41:48	11:01:10	21:59:34
31	01:03:17	08:26:38	10:19:09	22:18:03

Si procede nei seguenti paragrafi con l'analisi di ogni famiglia di carico.

5.1.3 STEP 2: MANIPOLAZIONE DEI DATI

I valori della potenza aggregata in ogni giorno del mese sono stati isolati gli uni dagli altri sfruttando gli istanti di campionamento ad essi associati; tuttavia, non è possibile eseguire una sovrapposizione di ciascun giorno mantenendo la correlazione con il tempo. Infatti, i momenti di campionamento variano e ogni tipologia di carico viene attivata in istanti differenti delle giornate; quindi, anche imponendo l'uguaglianza dei giorni non si arriverebbe alla sovrapposizione delle curve relative ad ogni famiglia di carico. La sovrapposizione, anche parziale, è di fondamentale importanza poiché senza di esso non si riuscirebbe a constatare l'effettiva ripetitività del carico e, di conseguenza, a tracciarne un andamento caratteristico, unico per tutti i giorni.

Per risolvere tale problema, sono state isolate le potenze aggregate giorno per giorno per ciascuna famiglia di carico, immaginandole come un semplice vettore nel quale i valori si susseguono, senza tenere conto del momento in cui esse effettivamente vengono rilevate.

Così facendo, rimane comunque un secondo problema, ossia, la variabilità del numero di campioni e, quindi, le differenti lunghezze dei vettori delle potenze isolate. Questo significa che, per esempio, a parità di tipologia di carico, esso viene rappresentato da 30 punti il giorno x e 40 il giorno $x+1$. I valori qui utilizzati sono casuali poiché servono solamente ad evidenziare la difficoltà nel rappresentare la medesima situazione relativa a momenti differenti nello stesso piano.

Per ovviare a tale problema, per ogni tipologia di carico e ogni giornata si è costruito un vettore dei tempi fittizio, ossia, un vettore con valore crescente da 1 alla lunghezza del corrispondente vettore della potenza aggregata. Per questo, nei grafici delle sezioni a seguire, l'asse delle ascisse non presenta dei tempi ma dei semplici valori numerici; tale asse non viene neanche denominato poiché esente da qualunque significato fisico.

Questa manipolazione dei dati disponibili è stata condotta per ogni tipologia di carico e rappresenta la base per le analisi successive.

5.1.4 STEP 3: CARICHI DI TIPOLOGIA A

I carichi di tipologia A sono quelli che si manifestano delle ore notturne come una successione di impulsi tra un valore minimo di circa 0.2 kW, che perdura per un intervallo di circa 35 *passi*, e un pianoro di 2.23 kW, che si protrae per 15 *passi*. Con il termine *passo* si fa riferimento ai salti nel vettore tempo fittizio, costruito appositamente per questa analisi e di cui si è parlato nel paragrafo precedente.

Per giungere a tali valori, è stata necessaria l'analisi di ogni giorno e, in particolare, delle ore tra 01:00:00 e 07:30:00, intervallo temporale all'interno del quale generalmente compare questa famiglia di carichi. I valori sono stati rilevati in tutti i giorni del mese. Quindi, analizzare i dati di 30 giorni (poiché non si tiene conto del 22 gennaio, essendo questo incognito) sarebbe troppo oneroso, sia in termini di tempo che in termini quantitativi di informazioni effettivamente estraibili.

Per tale ragione, si è fatta una selezione delle giornate tenendo conto solo di quei giorni in cui è visibile solamente il carico in questione senza la sovrapposizione di altri fenomeni. Per chiarire questo concetto, vengono di seguito riportati due esempi: in Figura 5.1.4-1 è tracciato l'andamento di un carico di tipo A, relativo al 7 gennaio 2015, nel quale è visibile nel rettangolo nero la presenza dell'inserzione di qualche altro dispositivo che determina un incremento della potenza complessiva; invece, Figura 5.1.4-2 si è tracciato l'andamento di un carico della medesima famiglia, il 10 gennaio, senza la sovrapposizione di altro.

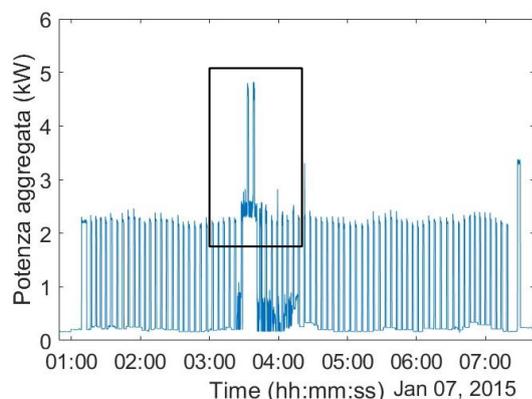


Figura 5.1.4-1 Carico di tipo A disturbato

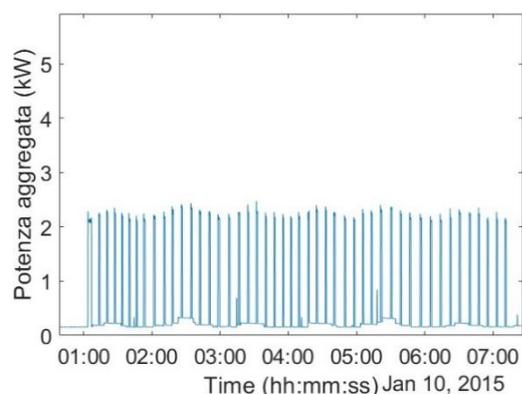


Figura 5.1.4-2 Carico di tipo A "puro"

Sebbene l'anomalia presente in Figura 5.1.4-1 sia riscontrabile in altri giorni, essa non è sempre uguale e non avviene neanche nei medesimi momenti; per tale ragione, si sono considerate solo le giornate con delle evoluzioni "pure", assimilabili a quanto mostrato in Figura 5.1.4-2. Questo riduce il numero di giorni da analizzare da 30 a 16, comunque superiore al 50% dei dati disponibili.

Sovrapponendo quanto accade in questi 16 giorni, si giunge ad una rappresentazione molto confusionaria (Figura 5.1.4-3), assimilabile ad una tavolozza di colori indistinguibili tra loro. Questo quadro poco preciso è conseguenza del fatto che tutte le curve non hanno pianori superiori e inferiori di uguale durata e dunque, negli spazi vuoti lasciati da alcune di esse si inseriscono i picchi di altre. L'unica cosa che sembra mantenersi sempre invariata è il primo impulso, ben visibile e distinto dagli altri: esso ha una durata di circa 40 passi e un'ampiezza media di 2.23 kW.

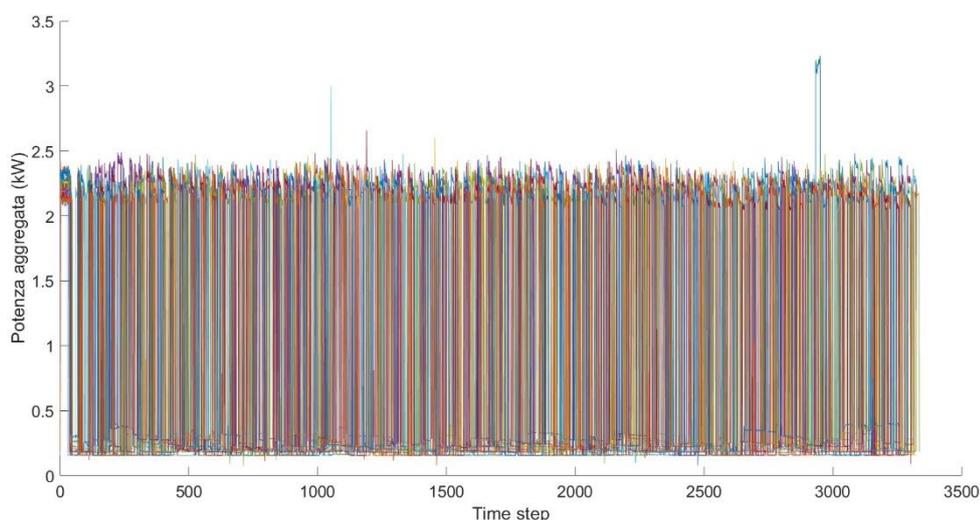


Figura 5.1.4-3 Sovrapposizione dei carichi di tipo A nei giorni in analisi

È chiaro che dedurre un'unica forma d'onda dalla sola osservazione della immagine in Figura 5.1.4-3 non è in alcun modo possibile. Quindi, si è proceduto analizzando le curve a piccoli gruppi di 4 e rilevando per ognuna di esse la durata dei pianori inferiori e di quelli superiori, in differenti punti del diagramma.

Questo ha condotto alle seguenti conclusioni:

- i pianori superiori, eccetto il primo, hanno una durata compresa tra 10 e 20 passi, con valore medio di circa 2.23 kW; dunque, si è optato per mantenere la potenza aggregata costante a tale valore per un intervallo composto da 15 passi;
- i pianori inferiori hanno durata compresa tra 30 e 40 passi e un valore medio di 0.2 kW; quindi, la potenza aggregata ricostruita rimarrà costante a questo valore per 35 passi;
- il numero di impulsi è variabile, ma comunque non inferiore a 65; quindi si è ipotizzato un numero pari a 80, poiché comunque correggibile nelle fasi successive.

Con queste informazioni si è dunque proceduto costruendo un treno di impulsi di pari ampiezza e durata che possano richiamare quanto accade nella realtà; il risultato cui si giunge è riportato in Figura 5.1.4-4.

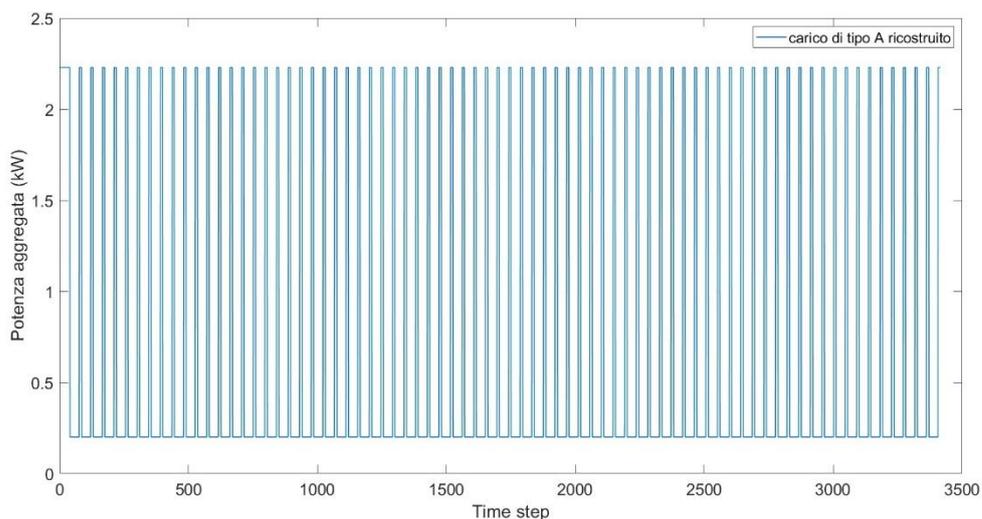


Figura 5.1.4-4 Carico di tipologia A, ricostruito

Aspettarsi che l'andamento ottenuto coincida esattamente con il carico reale non è realistico, dato che si è fatta una valutazione media ed approssimata del fenomeno; tuttavia, questo potrebbe condurre a dei miglioramenti, come si vedrà nel seguito.

Per questa ragione, la curva rappresentata in Figura 5.1.4-4 verrà sostituita in quella porzione di previsione del modello in cui il carico di tipo A è riconoscibile. Questo processo conduce a quanto mostrato in parte Figura 5.1.4-5; infatti, una raffigurazione dell'intero intervallo di durata del carico

di tipo A, non permetterebbe di visualizzare i risultati raggiunti. Come è possibile osservare, la curva gialla, rappresentante la ricostruzione del carico A, solo in alcuni momenti coincide, a volte anche per piccole porzioni, con quanto era stato preventivato dal modello, a causa della irregolarità con cui questi impulsi si verificano nei giorni. Tuttavia, una valutazione conclusiva della efficacia della correzione apportata può essere effettuata solo confrontando la ricostruzione complessiva con i campioni reali, come verrà svolto nel seguito.

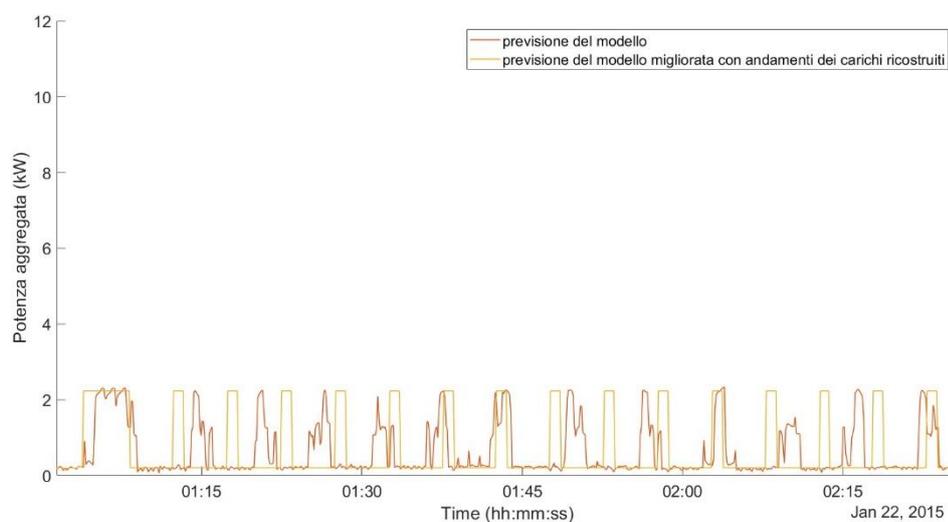


Figura 5.1.4-5 Sostituzione della porzione della previsione del modello con il carico di tipo A ricostruito

5.1.5 STEP 4: CARICHI DI TIPOLOGIA C

I carichi di tipologia C si caratterizzano per un ampio picco di breve durata; come mostrato in Figura 5.1.1-4, esistono due comportamenti simili, spesso ravvicinati tra loro: il primo si attesta attorno ai 10 kW e il secondo giunge al medesimo livello dopo un primo salto fino a 12 kW. Tenuto conto che il valore di 10 kW è da considerarsi con le dovute attenzioni (essendo questo molto elevato e di gran lunga superiore alle potenze ammesse dai più comuni contatori) e riscontrata una durata del valore superiore molto ridotta, si è ipotizzato che entrambi i cunei siano assimilabili al medesimo carico e che, nel secondo caso, la prima soglia di 12 kW sia attribuibile a qualche errore di misura e, dunque, trascurabile. Quindi, questa famiglia di carichi si caratterizza per un'ascesa

molto rapida della potenza richiesta, la quale permane costante a 10 kW tracciando un pianoro di qualche minuto per poi tornare ai valori antecedenti a questa variazione.

Si tenga conto, che i carichi di tipologia C sono i meno comuni, rilevati solo 14 su 31, come riassunto in Tabella 5.1.2-1; quindi, la loro descrizione è attuabile ma ha una affidabilità minore rispetto ai casi in cui una famiglia di andamenti si manifesta con una maggiore frequenza.

Per identificare il diagramma di carico di questa famiglia, si è scelto di considerare i campionamenti dell'11 gennaio nell'intervallo compreso tra le ore 10:08:22 e le 10:10:30. Infatti, sovrapponendo di volta in volta le curve ricavate da ciascuna giornata in corrispondenza di questi picchi molto accentuati, si è potuto osservare come molte di esse siano vicine a quella dell'11 gennaio. In realtà, alcune hanno una durata superiore, solo poche hanno durata inferiore, ma l'ipotesi più ragionevole è quella di identificare gli andamenti dei carichi di questa famiglia con quello che si presenta più volte e giace nel mezzo tra i due estremi.

Con l'intento di ottenere una migliore ricostruzione, evitando di avere delle oscillazioni indesiderate, si è utilizzata la forma d'onda dell'11 gennaio per costruire un gradino di pari durata ed ampiezza di 10 kW, per quanto sopra esposto, come rappresentato in arancione nella Figura 5.1.5-1.

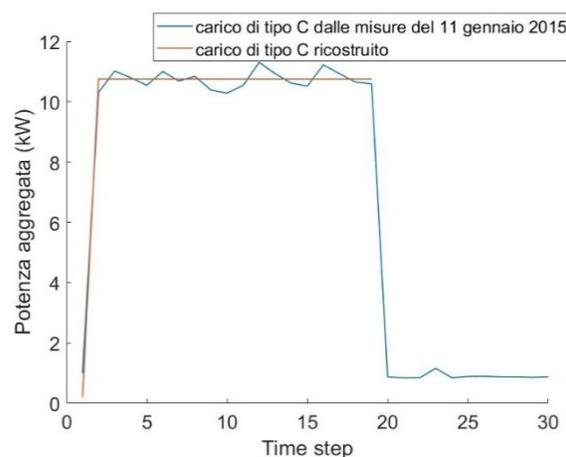


Figura 5.1.5-1 Ricostruzione andamento della potenza di un carico di tipo C

Ottenuta la descrizione della potenza richiesta di questi carichi, si è proceduto al suo inserimento nella forma d'onda calcolata con la rete *wide*: identificati i momenti in cui questo carico si

manifesta, ossia, tra le 10:13:58 e le 10:16:00, si sono sostituiti i dati in questo range temporale con quelli ottenuti dall'analisi condotta, giungendo alla curva gialla in Figura 5.1.5-2. Essa racchiude quella arancione e presenta un andamento più regolare.

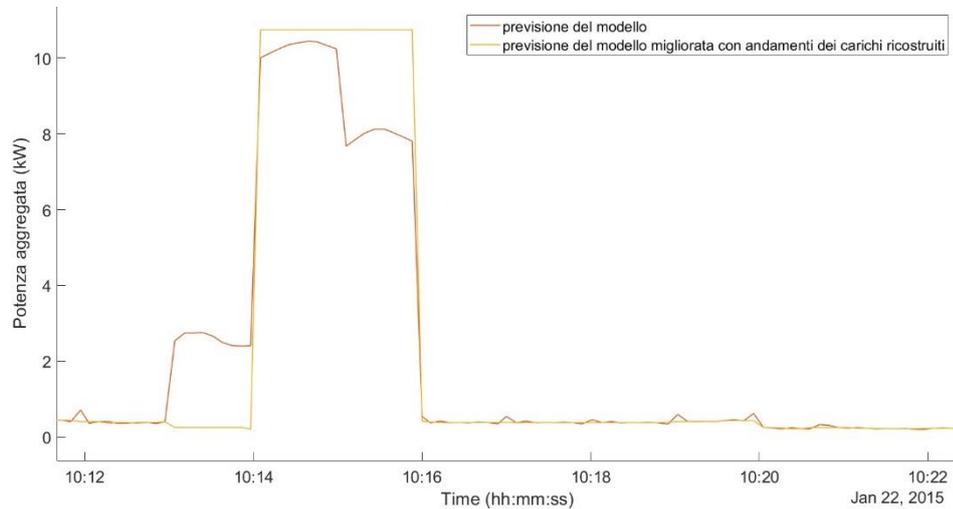


Figura 5.1.5-2 Sostituzione della porzione della previsione del modello con il carico di tipo C ricostruito

Nel seguito si analizzeranno i vantaggi ottenibili da questa sostituzione nella previsione, confrontandola con quanto campionato durante le campagne di misurazione.

5.1.6 STEP 5: CARICHI DI TIPOLOGIA D

In questa famiglia rientrano quei carichi che determinano l'incremento della potenza totalmente richiesta dall'abitazione per un breve intervallo di tempo, conducendola ad un valore di potenza tra 3.4 kW e 3.8 kW, nelle ore serali. In particolare, si è riscontrata la comparsa di tale variazione intorno alle ore 22:00 di ciascun giorno del mese, come si può dedurre dalla Tabella 5.1.2-1. Dalle informazioni nella medesima, questo comportamento non è sempre presente ma lo si ritrova in 25 giorni del mese su 31 che lo compongono; si tenga conto che il 22 gennaio non è conteggiato, poiché ritenuto incognito, e i primi giorni dell'anno probabilmente l'abitazione non era occupata. Infatti, i diagrammi di carico in quelle giornate, presentano solo le ripetitività notturne ma, durante il resto delle giornate, non si rilevano particolari variazioni di potenza che conducano a pensare all'utilizzo di qualche dispositivo.

Per procedere alla definizione di un modello di tale carico applicabile al risultato della simulazione, in Figura 5.1.6-1 è stata tracciata la sovrapposizione delle potenze aggregate ad esso attribuibili nei giorni in cui esso si è palesato. In questa rappresentazione non si è riportata la legenda poiché, eccetto nei tratti di discesa, le linee sono poste pressoché una sull'altra e, quindi, non sarebbe comunque possibile distinguerle.

Tuttavia, l'utilità di questa analisi risiede proprio nell'impossibilità di individuare le singole curve, poiché questo implica la ripetitività effettiva di questa tipologia di carico nel mese e la possibilità di elaborare un unico andamento, pulito e preferibilmente privo di disturbi indesiderati.

Nel processo di realizzazione della Figura 5.1.6-1, si è potuto osservare che l'andamento più comune è quello che si riscontra nella giornata del 6 gennaio: in numerosi altri giorni, infatti, l'ampiezza dell'impulso in analisi è approssimabile ad un valore più o meno costante di 3.6 kW e la sua durata è poco superiore ai due minuti.

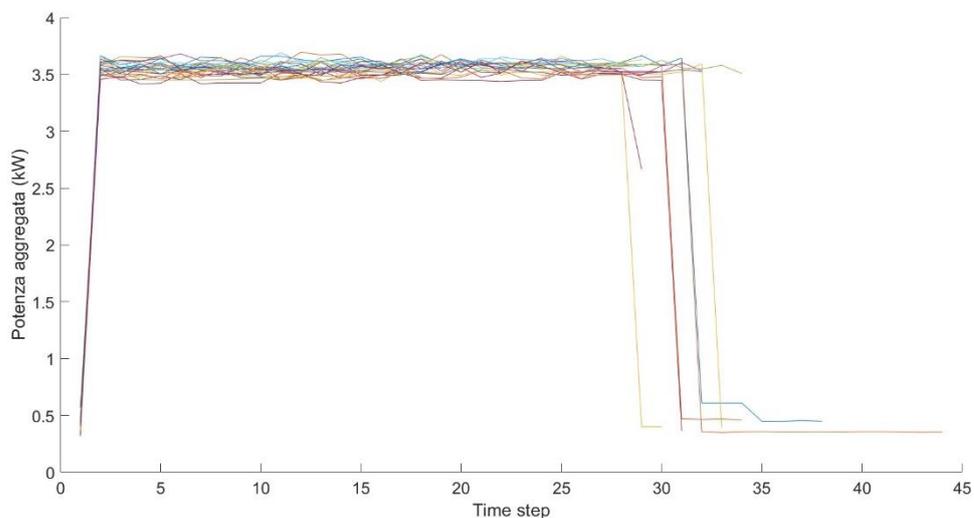


Figura 5.1.6-1 Sovrapposizione carichi di tipo D rilevati nei giorni costituenti il mese di gennaio

Pertanto, si è isolato l'andamento del 6 gennaio 2015 tra le 21:57:11 e le 22:01:01 e lo si è impiegato per ricostruire l'onda caratteristica di tali carichi, ossia, qualcosa di simile ad un'onda quadra, come in Figura 5.1.6-2.

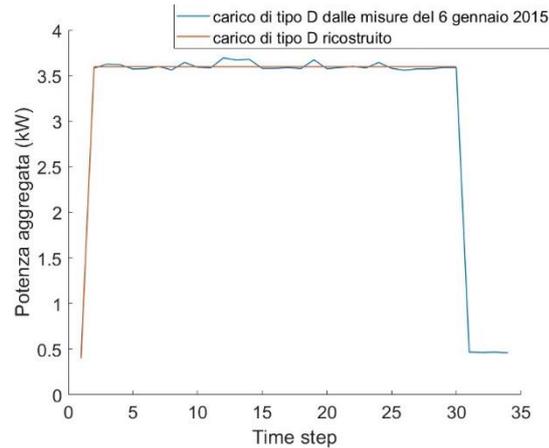


Figura 5.1.6-2 Ricostruzione del carico di tipo D

Il passo successivo è la sostituzione della curva arancione ottenuta e mostrata in Figura 5.1.6-2 nella ricostruzione effettuata con il modello *wide* (Figura 5.1.1-1). A tale scopo si isola il primo momento nel quale la previsione comincia la sua crescita attorno alle ore 22:00 e si procede con la sostituzione dei dati fintanto che la potenza rimane approssimativamente costante. Tale processo conduce al risultato riportato in Figura 5.1.6-3: questo andamento è molto simile a quello dei carichi di tipo C (Figura 5.1.5-1) ma la sua ampiezza è molto minore. In particolare, La curva arancione è ancora la previsione ottenuta dalla rete neurale, mentre in giallo è riportato quanto ottenuto dall’analisi di cui sopra.

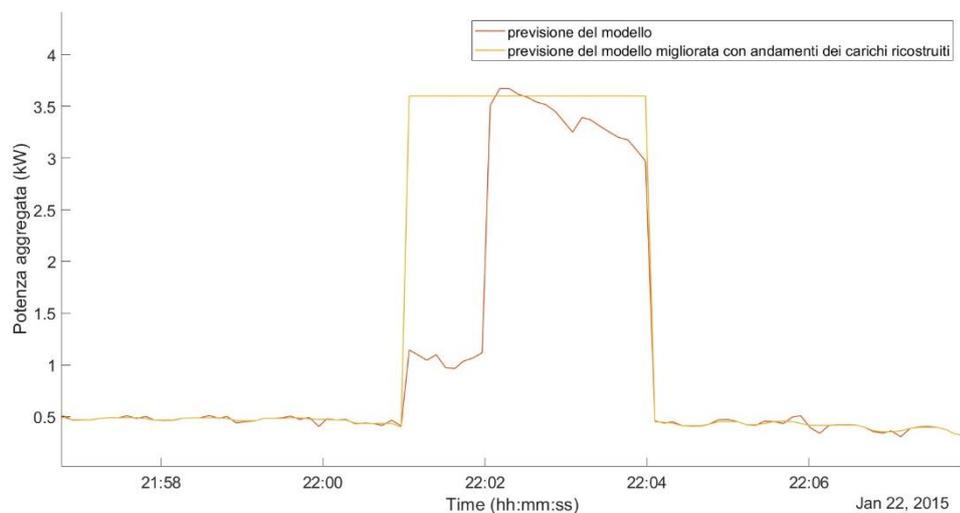


Figura 5.1.6-3 Sostituzione della porzione della previsione del modello con il carico di tipo D ricostruito

Si vedrà nel prosieguo l’impatto di questa sostituzione sul risultato finale della previsione.

5.1.7 STEP 6: CARICHI DI TIPOLOGIA B

Questa famiglia di carichi è stata lasciata per ultima poiché, nonostante essa sia riscontrabile in 16 giorni del mese (più del 50 % di quelli disponibili), comunque non vi è alcuna sua traccia nella previsione ottenuta dalla rete neurale di tipo *wide*. Per tale motivo, non si procederà ad una sua trattazione approfondita.

In questa famiglia rientrano quei carichi che determinano l'incremento della potenza totalmente richiesta dall'abitazione per un breve intervallo di tempo, conducendola ad un valore tra i 3.1 kW e 3.5 kW. Dall'analisi di ciascun giorno, si è riscontrata la comparsa di tale variazione a valle della successione di impulsi cui è soggetta la medesima grandezza nelle ore notturne; tuttavia, data la differenza di quasi 1 kW rispetto a quest'ultima, si può escludere che l'andamento in analisi sia uno strascico di quanto avvenuto nelle ore precedenti.

I passaggi per una ricostruzione del comportamento di questi carichi sono identici a quelli affrontati per le famiglie precedenti; dalla sovrapposizione degli andamenti associabili in ciascun giorno a questa famiglia di carichi, si è estratto il giorno 9 gennaio come quello adatto allo scopo perseguito, poiché mediano tra gli altri. Da esso, si è proceduto a realizzare un'onda di ampiezza media tra quelle rilevate, ossia 3.3 kW, e durata di 24 passi; il risultato è mostrato in Figura 5.1.7-1.

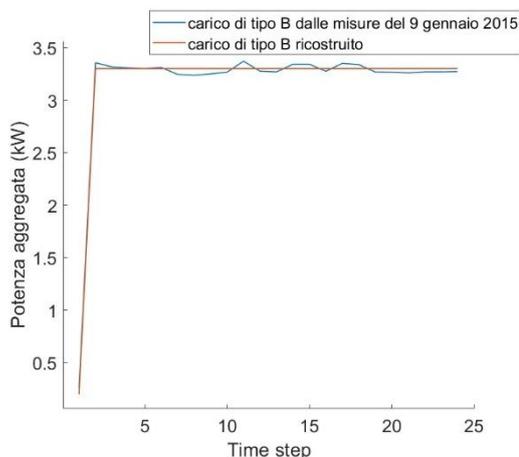


Figura 5.1.7-1 Ricostruzione del carico di tipo B

5.2 MIGLIORAMENTO DEL RISULTATO DELLA SIMULAZIONE

Ricostruite le quattro tipologie di carico ripetitivo, si è proceduto con la loro sostituzione nella previsione fornita dalla rete *wide*, ipotizzando che quest'ultima riesca a determinare gli esatti istanti in cui essi si manifestano.

A valle di questa sostituzione, si sono apportate altre due correzioni: la prima riguarda i picchi della previsione che non richiamano in alcun modo i carichi sopra analizzati; la seconda interessa le parti del diagramma di carico in cui la potenza aggregata è poco variabile. Nel primo caso, non avendo effettivamente informazioni sulla natura del carico che causa queste variazioni e non riscontrando alcuna altra sua ripetizione, non si è ritenuto opportuno mantenere queste variazioni. Quindi, negli intervalli temporali che intercorrono tra un carico noto e il successivo, ove la potenza aggregata supera il valore di 0.6 kW è stata livellata al valore costante che il modello aveva previsto prima del picco; si è scelto come limite accettabile quello di 0.6 kW poiché, nei tratti in cui la potenza aggregata varia poco, i valori che questa assume sono sempre inferiori a 0.5 kW. In particolare, nelle ore mattutine il valore si attesta nell'intorno di 0.3 kW, probabilmente associabile a un carico sempre attivo, quale il frigorifero; ad esso si sovrappone qualcos'altro nelle ore serali, portando la potenza ad attestarsi attorno a 0.45 kW.

Per tale motivo, è accettabile ipotizzare che tutto quello che è al di sopra di 0.6 kW è una variazione della quale non si può avere certezza. Questo conduce ad eliminare parte dei picchi della previsione da modello *wide*, che nel caso in analisi ove si conosce la potenza reale, effettivamente corrispondono a carichi che non sono utilizzati. Tuttavia, lo stesso livellamento si attuerebbe anche senza la conoscenza della grandezza campionata.

Si vuole sottolineare che questa è una scelta arbitraria, quindi, si potrebbe anche scegliere di mantenere queste variazioni di potenza affidandosi alla previsione del modello; infatti, non essendoci informazioni su quello che in realtà è l'evoluzione della potenza, non si può giudicare più corretta una via piuttosto che l'altra.

Per quanto riguarda i tratti con potenza poco variabile, la previsione della rete neurale è abbastanza frastagliata, ossia disturbata dalla presenza di molte oscillazioni frutto del calcolo stesso. In queste situazioni si è utilizzata la funzione *medfilt1* presente nel software *Matlab* per filtrare questo andamento e renderlo più regolare.

Questo procedimento è stato attuato in due differenti casi, al fine di sottolineare l'importanza della ripetitività dei carichi per giungere ad una corretta ricostruzione. In particolare, inizialmente si è operato prendendo come riferimento il risultato del modello wide ottenuto con i *predictor* relativi al 15 gennaio, sopra trattato. In un secondo momento, si è fatta la stessa operazione impiegando quanto ricavato con i *predictor* basati sulle misure dell'8 gennaio.

5.2.1 CASO 1: MODELLO CON PREFICTOR DEL 15 GENNAIO

Il modello wide impiegato con i *predictor* delle potenze medie al minuto e all'ora relativi alle misure del 15 gennaio ha condotto ad una forma d'onda in cui sono riscontrabili i carichi di tipo A, C e D (Figura 5.1.1-1). Quindi, ove si riscontrano tali carichi, si sono sostituiti i valori del modello con quelli ottenuti dalla ricostruzione del loro andamento tipico, giungendo a quanto mostrato in Figura 5.2.1-1. In questo riquadro si riporta in blu la potenza aggregata campionata (che dovrebbe essere incognita), in arancione quella ottenuta dal modello *wide* e in ocra quella ottenuta a valle della sostituzione dei carichi ripetitivi.

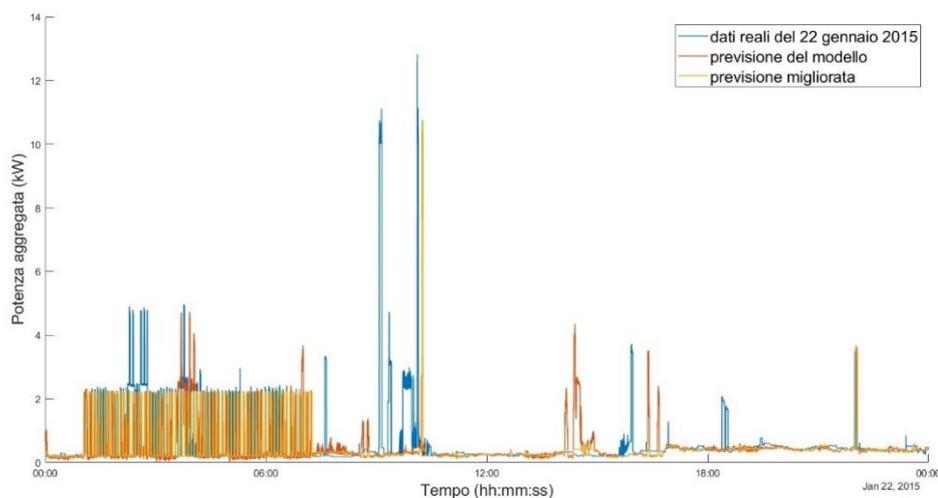


Figura 5.2.1-1 Confronto tra potenza aggregata misurata, quella prevista dal modello wide e la medesima migliorata con *predictor* del 15 gennaio 2015

Dal confronto di queste tre curve, si può immediatamente dedurre che, come previsto, alcune informazioni sono andate perse. Infatti, numerose oscillazioni della potenza reale (blu) non sono state previste dal modello wide e quindi neanche dal risultato finale. La rete neurale, invece, ha previsto la presenza di un consumo di energia tra le 12:00 e le 18:00, elemento che è stato eliminato con la rielaborazione della curva poiché non riscontrato in maniera certa ed univoca in altri giorni. Anche in corrispondenza delle ore notturne sono presenti delle variazioni che sono state ignorate: quelle reali (in blu) poiché non riconosciute dal modello wide e quelle previste (in arancione) poiché riscontrate in altri giorni ma sempre ad orari differenti. Queste ultime sono state quindi ignorate perché questo tipo di carico, sebbene presente in differenti momenti del mese, non lo è in orari regolari e dunque ci sono le stesse probabilità che la previsione sia corretta o meno. Per questo motivo, così come per le altre variazioni della giornata incognita, l'impossibilità di definire il carico come effettivamente ripetitivo ha condotto alla sua cancellazione.

La rappresentazione in Figura 5.2.1-1 non permette di visualizzare i miglioramenti che si sono ottenuti con la sostituzione dei carichi ripetitivi; per tale ragione si riportano nel seguito delle rappresentazioni ingrandite di alcune porzioni di questa immagine al fine di avere una visione più chiara delle tre evoluzioni temporali.

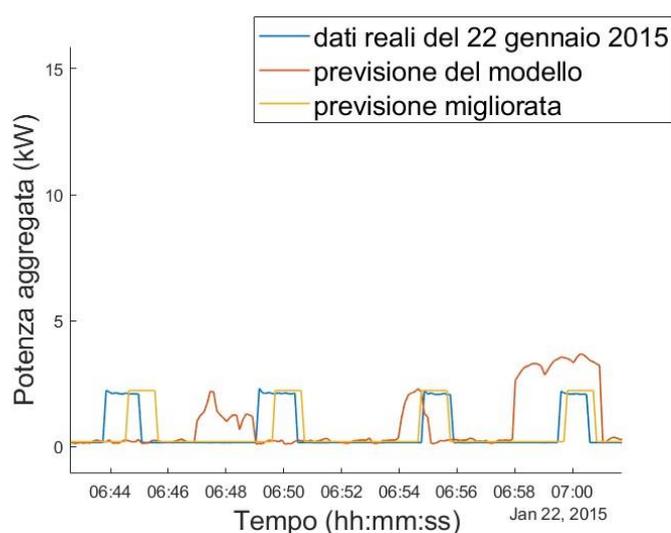


Figura 5.2.1-2 Dettaglio di Figura 5.2.1-1- ore notturne

In Figura 5.2.1-2 si osserva una porzione delle ore notturne nelle quali la potenza aggregata si caratterizza per la successione di impulsi di ampiezza più o meno costante. Si vede chiaramente che, per questa porzione, la sostituzione del carico caratteristico nella curva arancione relativa alla previsione dal modello *wide* ha condotto ad un notevole miglioramento: infatti, la curva gialla riduce l'errore della previsione avvicinandosi alla curva reale blu, talvolta sovrapponendosi ad essa.

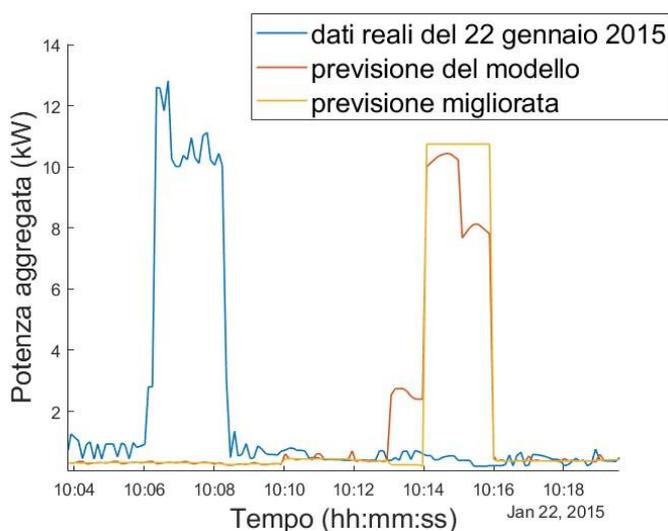


Figura 5.2.1-3 Dettaglio di Figura 5.2.1 1- ore diurne

La Figura 5.2.1-3 mostra il picco molto accentuato della potenza aggregata nelle ore diurne. In questa situazione, la previsione a valle delle correzioni rimane in ritardo rispetto alla realtà ma si evince che la sua forma è molto più simile a quella blu rispetto a quanto lo sia la curva arancione. Questo implica che la ricostruzione del carico ripetitivo e la sua sostituzione nella previsione hanno condotto ad un risultato più vicino alla realtà.

Infine, la Figura 5.2.1-4 è l'ingrandimento delle tre curve in analisi nelle ore serali e, in particolare, il picco delle ore 22:00. Anche in questo caso, la curva gialla riduce l'errore complessivamente commesso perché, essendo caratterizzato dalla medesima forma della potenza reale traslata di qualche passo temporale, compensa la mancanza della curva arancione nei primi momenti di avviamento del carico. È altresì vero che la ricostruzione finale aggiunge una piccola area negli ultimi

istanti del carico, ma tale area è comunque minore di quella iniziale, così che il bilancio rimane positivo.

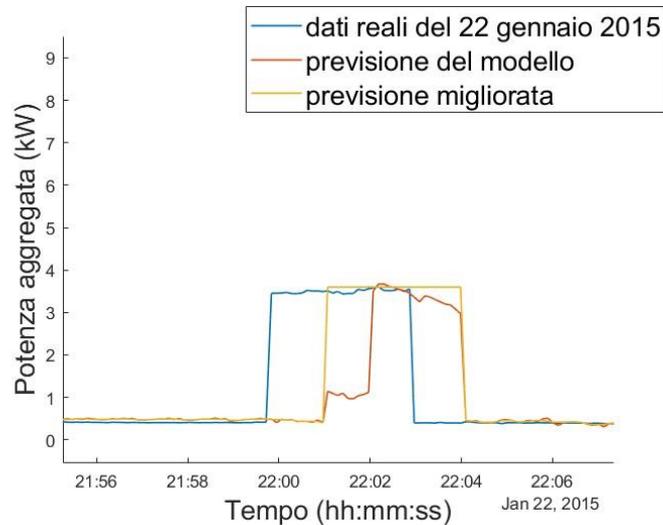


Figura 5.2.1-4 Dettaglio della Figura 5.2.1 1 - ore serali

Per attuare un confronto numerico, perché di più rapido impatto, si è calcolato l'errore totale commesso dalla previsione del modello *wide*, prima e dopo l'inserimento dei carichi caratteristici. Tale errore è stato ottenuto come somma dei valori assoluti delle differenze tra la realtà e la previsione, giungendo a $5.59 \cdot 10^3$ kW, prima delle sostituzioni, ed a $4.93 \cdot 10^3$ kW, dopo di esse. Entrambi i valori sono grandi, anche perché la potenza reale manifesta un notevole incremento in alcune aree che non viene considerato; tuttavia, essi permettono comunque di vedere un miglioramento nell'accuratezza complessiva a valle del processo di miglioramento attuato nelle fasi precedenti.

Si vuole sottolineare in questa sede, che non si può in alcun modo avere la pretesa di prevedere l'andamento della potenza se non si conosce nulla su di essa. L'obiettivo è, infatti, avvicinarsi a qualcosa che sia verosimile alla realtà sebbene identificare gli istanti esatti in cui la grandezza incognita subisce delle variazioni non è attuabile.

5.2.2 CASO 2: MODELLO CON PREDICTOR DELL'8 GENNAIO

Gli stessi passi compiuti per il modello con *predictor* del 15 gennaio vengono qui seguiti basandosi sulla previsione del modello *wide*, applicato con le informazioni prevenienti dall'8 gennaio.

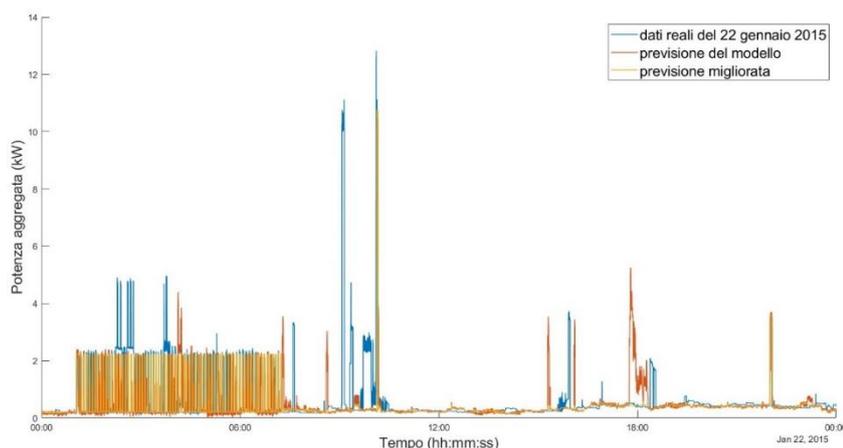


Figura 5.2.2-1 Confronto tra potenza aggregata misurata, quella prevista dal modello wide e la medesima migliorata con *predictor* dell'8 gennaio 2015

Nella Figura 5.2.2-1 sono rappresentati i due risultati ottenuti, in arancione quello dal modello *wide* e in ocrà quello a valle delle sostituzioni dei carichi ripetitivi, rispetto all'evoluzione reale della potenza (blu). Come nel caso precedente, le informazioni relative ai picchi di potenza determinati dall'utilizzo casuale dei dispositivi all'interno dell'abitazione non sono in alcun modo pervenute. Infatti, l'8 gennaio non si hanno i medesimi consumi del 22 gennaio, quindi la rete neurale impiegata per ottenere una prima sommaria previsione non sarà in grado di valutare la presenza di tali consumi. Anche in questo caso, nel passaggio dalla curva arancione a quella ocrà, non si sono solo sostituiti i tre carichi caratteristici precedentemente analizzati ma si sono anche livellati gli altri picchi presenti nella previsione, non potendo essere certi della loro correttezza.

Sebbene questa simulazione sembri non molto differente dalla precedente (Figura 5.2.1-1), in realtà essa presenta un errore totale minore, poiché già la stessa previsione dal modello *wide* è migliore. Infatti, l'errore totalmente commesso dalla rete neurale, somma degli errori istantanei, corrisponde a $5,61 \cdot 10^3$ kW. La sostituzione dei diagrammi dei carichi ripetitivi conduce ad un ulteriore miglioramento, e quindi, ad un errore minore di circa $4,74 \cdot 10^3$ kW.

Per poter visualizzare le differenze rispetto al risultato finale raggiunto con il modello *wide* con *predictor* del 15 gennaio, si ponga l'attenzione sui dettagli della Figura 5.2.2-1 di seguito raffigurati e corrispondenti a quelli riportati nella simulazione precedente: le ore notturne in Figura 5.2.2-2, il

picco nelle ore diurne in Figura 5.2.2-3 e la variazione delle ore serali nella Figura 5.2.2-4 Dettaglio di Figura 5.2.2-1- ore serali.

Nella prima rappresentazione non è possibile definire quale sia la curva più accurata tra quella arancione e quella ocra, poiché in alcuni momenti l'una si sovrappone alla linea blu e, in altri, è l'altra a confondersi con quella reale. Si può solo affermare che la curva ocra ha un'evoluzione praticamente identica con quella reale e, se quest'ultima avesse degli impulsi equi-distanziati, si giungerebbe ad una sovrapposizione quasi totale.

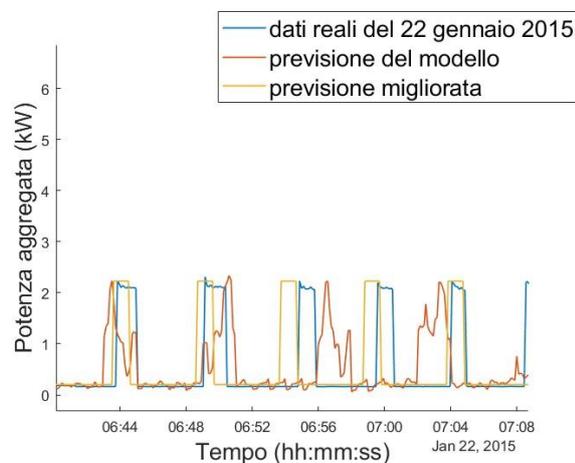


Figura 5.2.2-2 Dettaglio di Figura 5.2.2-1- ore notturne

In Figura 5.2.2-3 è rappresentato il picco subito dopo le ore 10:00. In questo caso, il risultato finale cui si giunge è migliore rispetto al precedente (Figura 5.2.1-3) poiché il modello *wide* con i *predictor* relativi all'8 gennaio fornisce una previsione che in parte si sovrappone all'andamento reale. Questo implica che il momento di inizio considerato per la sostituzione del diagramma di carico di tipo C rientra nell'intervallo di durata del picco reale e quindi, in parte lo si riesce a prevedere. In aggiunta, la sostituzione nella previsione dal modello *wide* ha permesso di recuperare una parte del carico reale, nel momento di inizio, dove la curva arancione è minore rispetto a quella blu di circa 4 kW, e di eliminare una porzione in eccesso nei momenti finali.

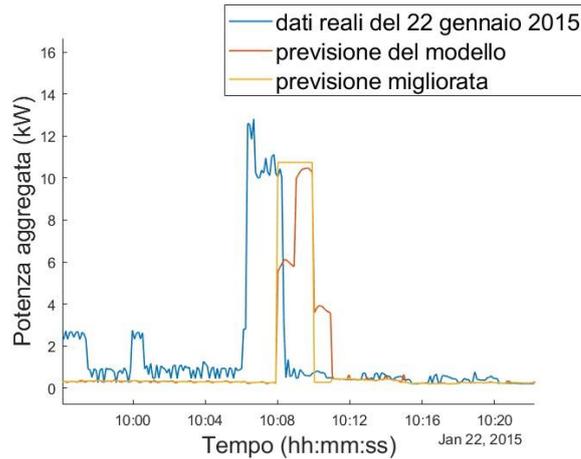


Figura 5.2.2-3 Dettaglio della Figura 5.2.2-1 - ore diurne

Infine, in Figura 5.2.2-4, la sostituzione del carico caratteristico di tipo D nel risultato del modello apporta un contributo assolutamente positivo; infatti, la curva ocra è molto vicina a quella blu e ha permesso di eliminare porzioni in eccesso, aggiungendole quando la curva arancione era in difetto sulla curva blu.

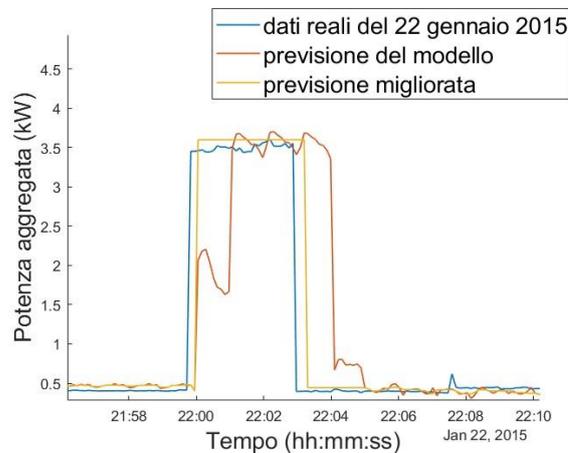


Figura 5.2.2-4 Dettaglio di Figura 5.2.2-1- ore serali

In conclusione, il risultato finale che si è ottenuto si può ritenere più accurato di quello precedente. Questo è determinato da una maggiore somiglianza tra l'andamento della potenza aggregata durante la giornata dell'8 gennaio e quello nel giorno incognito.

Per tale ragione, si può concludere nuovamente, che la ripetitività dei carichi è la chiave per poter giungere a previsioni più accurate, poiché si avrebbero delle informazioni precise sui momenti di impiego dei carichi e sull'effettivo utilizzo dei dispositivi.

5.3 RIDUZIONE DEGLI INTERVALLI DI ANALISI

In questo contesto, si vuole attenzionare un aspetto importante. Tutte le simulazioni svolte si basano sulla volontà di prevedere un intervallo temporale piuttosto ampio, pari alle 24 ore; considerando che la banca dati impiegata dichiara la realizzazione di campionamenti ogni 8 secondi, questo implica l'obiettivo di volere ricostruire cosa succede in 10800 momenti della giornata, nei quali la potenza richiesta dall'abitazione evolve in maniera casuale.

Tuttavia, se si restringesse il campo di analisi, si potrebbe giungere a risultati migliori, perché, presumibilmente, si potrebbero conoscere gli istanti esatti in cui i dispositivi corrispondenti ai carichi caratteristici sono attivati. In questa situazione si potrebbe prevedere comunque quello che accadrebbe, anche senza l'ausilio della rete neurale ma solamente impiegando i diagrammi costruiti per i carichi ripetitivi.

Nel seguito si riporta un esempio nel quale si ipotizzano mancanze di dati nella giornata del 22 gennaio 2015 inferiori alle 24 ore e si ricostruiscono tali andamenti senza l'ausilio delle reti neurali.

5.3.1 MANCANZA DI 30 MINUTI NELLE ORE SERALI

Si ipotizzi l'assenza delle misure della potenza aggregata tra le ore 22:00 e le 22:30 mentre risulta noto quanto avviene a monte e a valle di tale intervallo. Per visualizzare questa situazione, si è annullata la grandezza in analisi in questo range temporale, ottenendo quando riportato in Figura 5.3.1-1.

In tale immagine viene rappresentata solo la porzione di interesse della potenza aggregata, evidenziando in rosso il tratto nel quale essa viene volontariamente annullata per riconoscere il range nel quale ipoteticamente non è nota. Tuttavia, poco prima del tratto rosso, è presente un picco della grandezza incognita del quale si è identificato il valore con l'impiego della funzione *Data Tips* delle figure di *Matlab*. Dunque, sebbene non si sappia cosa avvenga nei 30 minuti successivi alle 22:00, si osserva un picco di 3.45 kW poco prima di questo orario. Tenuto conto dell'analisi dei carichi ripetitivi fatta precedentemente e, in particolare, del carico di tipo D, tale variazione può

essere associata a questa categoria di utenza, sia per la sua ampiezza che per l'orario in cui essa si presenta.

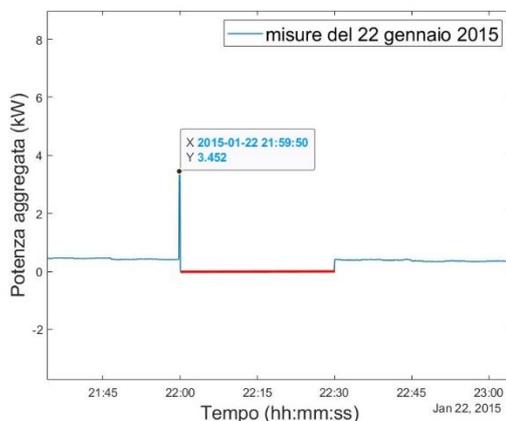


Figura 5.3.1-1 Misure reali del 22 gennaio con annullamento della potenza aggregata nell'intervallo 22:00-22:30

Per tali ragioni, si può ricostruire quello che avviene nei minuti mancanti impiegando il diagramma di carico caratteristico ottenuto in Figura 5.1.7-1, facendo attenzione al momento di inizio del carico. Infatti, in Figura 5.3.1-1 si constata la presenza di alcuni punti del carico ripetitivo nelle misure reali e, di conseguenza, per giungere ad una rappresentazione con la durata corretta, il suo diagramma di carico deve essere inserito sottraendo un numero di punti iniziali pari a quelli già presenti. Con tale accortezza, il risultato finale cui si giunge è quello in Figura 5.3.1-2.

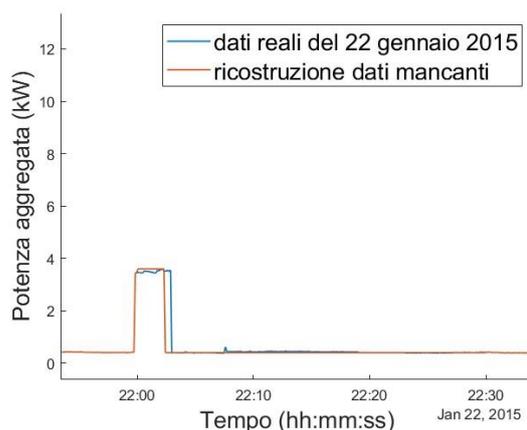


Figura 5.3.1-2 Confronto tra potenza reale e ricostruita con l'impiego del solo carico caratteristico di tipo D

Dalla rappresentazione grafica ottenuta, si evince che la ricostruzione finale è quasi coincidente con le misure reali, eccetto per un leggero sfasamento temporale tra i due tratti di discesa.

La ricostruzione si può considerare perfetta nel tratto successivo all'impulso, poiché la linea blu è distinguibile da quella arancione solo a causa della presenza di lievi oscillazioni nella sua evoluzione. Quindi, si è osservato come l'analisi delle ripetitività dei carichi nei giorni può condurre a previsioni corrette, anche senza l'ausilio delle reti neurali.

È altresì vero che se l'intervallo di 30 minuti ipotizzato per la mancanza di misure avesse inizio poco prima del picco delle 22:00, per esempio alle 21:50 come evidenziato dalla linea rossa in Figura 5.3.1-3, il solo diagramma di carico caratteristico non sarebbe sufficiente ad ottenere una previsione corretta, perché non potremmo neanche essere sicuri della presenza di tale picco. Questo aspetto diventa ancora più rilevante tanto meno è presente un carico caratteristico negli altri giorni del mese. Infatti, la variazione delle ore serali è presente 25 giorni su 30 (Tabella 5.1.2-1) quindi, è corretto ipotizzarne la presenza anche se le mancanze fossero tra le 21:50 e le 22:20. Lo stesso non si può affermare per altri carichi caratteristici, come per esempio quello di tipo B; esso, infatti, è stato riscontrato in appena 16 giorni su 30 (Tabella 5.1.2-1) e quindi la sua presenza è possibilmente vera con una probabilità di circa il 50%.

In virtù di questo, se le misure mancassero nel range tra le 7:20 e le 7:50 (in rosso in Figura 5.3.1-4) non si potrebbe affermare la presenza del carico di tipo B con la stessa certezza con cui lo si farebbe per il carico D.

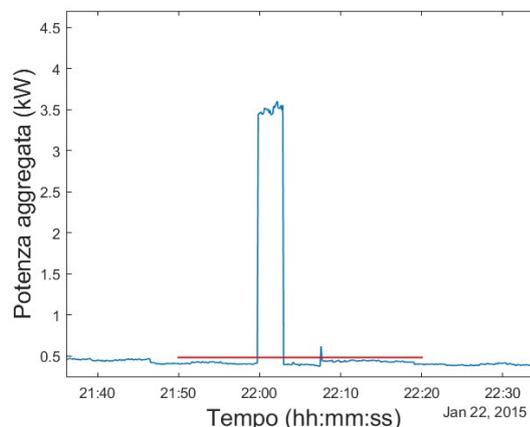


Figura 5.3.1-3 Intervallo con l'ipotetica mancanza di misure 21:50-22:20 il giorno 22 gennaio 2015

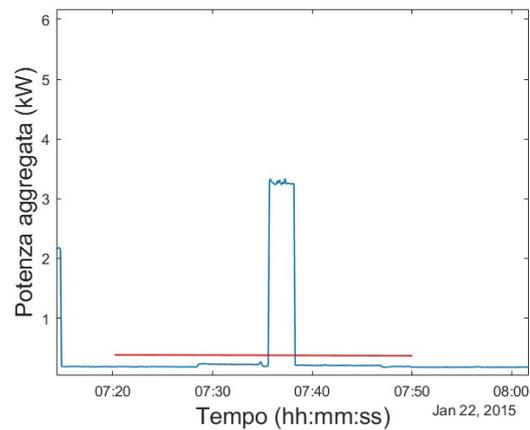


Figura 5.3.1-4 Intervallo con l'ipotetica mancanza di misure 7:20-7:50 il giorno 22 gennaio 2015

Lo stesso principio lo si potrebbe seguire per qualunque altro intervallo temporale per cui si conclude che la ripetitività del carico è un elemento di importanza prevalente al fine di attuare una buona previsione. In virtù di questo, l'esperto di dominio può tentare una previsione anche senza l'ausilio di modelli di previsioni quali le reti neurali, le quali diventano tuttavia utili nel momento in cui le mancanze di misure si manifestano in corrispondenza di istanti nei quali non è possibile affermare nulla con certezza, a causa di insufficienti informazioni fornite dai dati a disposizione.

6. SIMULAZIONE FINALE

A conclusione del percorso compiuto, si vuole condurre un'ultima simulazione, ipotizzando che la giornata incognita sia sabato 17 gennaio 2015. Trattandosi di un sabato, per l'identificazione dei *predictor* relativi alla potenza media sul minuto e sull'ora si è utilizzata la potenza aggregata campionata il 10 marzo, il sabato della settimana precedente.

6.1 SCALETTA RIASSUNTIVA

Si riassumono di seguito i passaggi seguiti al fine di fornire una eventuale scaletta da impiegare per ulteriori simulazioni:

- PASSO 1: PREPARAZIONE DEI DATI

A partire dai momenti di campionamento, si determinano i *predictor* temporali mentre, utilizzando le misure della potenza totalmente impiegata nell'abitazione, si ricavano i *predictor* delle potenze medie nell'ora e nel minuto di tutti i giorni.

- PASSO 2: DIVISIONE DEI DATI PER L'ADDESTRAMENTO E PER LA PREVISIONE

La matrice ottenuta al passo precedente viene divisa in due sottomatrici: la prima utilizzata per l'addestramento della rete, composta sia dalla potenza campionata che dai *predictor* nei medesimi istanti, e la seconda per la previsione, le cui colonne contengono i valori dei soli *predictor*. Si ricordi che in questa seconda matrice si deve provvedere alla sostituzione dei *predictor* medi relativi alla giornata incognita con quella del medesimo giorno della settimana precedente.

- PASSO 3: ADDESTRAMENTO DELLA RETE NEURALE WIDE A 50 NEURONI

La prima matrice ottenuta al passo precedente viene inserita nell'applicazione *Regression Learner* e si procede con il processo di addestramento della rete *wide* con 50 neuroni, essendo questa risultata essere quella più prestante; alla fine di tale processo,

si esporta il modello nel *workspace* di Matlab così da poterlo impiegare semplicemente nelle fasi successive.

- **PASSO 4: PREVISIONE DELLA GIORNATA INCOGNITA**

Il modello importato viene utilizzato con i dati della seconda matrice ottenuta al passo 2 per prevedere la potenza nella giornata incognita.

- **PASSO 5: MIGLIORAMENTO DELLA PREVISIONE CON LA SOSTITUZIONE DEI CARICHI RIPETITIVI**

Nella previsione ottenuta si identificano l'eventuale presenza dei carichi ripetitivi studiati nel capitolo precedente e si sostituiscono alla stessa per migliorarne l'accuratezza.

6.2 RISULTATI DELLA PREVISIONE DEL 17 GENNAIO 2015

6.2.1 VISUALIZZAZIONE GRAFICA DEI RISULTATI

Per simulare quello che sarebbe un reale processo di previsione, inizialmente si analizza la Figura 6.2.1-1 rappresentante quanto ottenuto dal modello *wide*. In questa ricostruzione sono evidenziati in rettangoli neri gli andamenti che fanno pensare alla presenza di uno dei carichi per i quali è stato ricostruito l'andamento caratteristico. In particolare, con riferimento alla nomenclatura impiegata nel capitolo 5, nel riquadro 1 si osserva il carico di tipologia A, nel 2 la tipologia B (assente nella precedente simulazione), nel 3 il carico C e nell'ultimo il tipo D. Sono dunque presenti tutti e quattro i carichi analizzati. A questi si aggiungono delle oscillazioni nelle ore pomeridiane che non sono classificabili con qualche andamento caratteristico. Infine, si constata un notevole disturbo della potenza aggregata prima della sequenza di impulsi delle ore notturne.

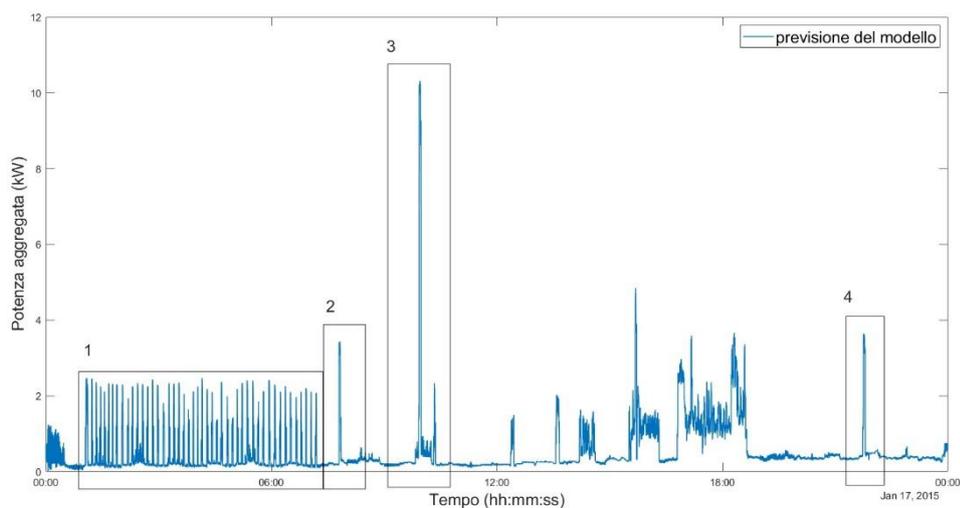


Figura 6.2.1-1 Previsione del 17 gennaio 2015 con modello *wide* a 50 neuroni e predictor relativi al 10 gennaio 2015

Constatati i margini per un miglioramento, si è proceduto con la sostituzione degli andamenti nei riquadri neri con quelli ricostruiti per i carichi caratteristici e con l'eliminazione di tutte le variazioni incerte della potenza. Si è inoltre filtrata l'evoluzione così ottenuta per ridurre le oscillazioni della grandezza incognita, sicuramente attribuibili alle incertezze del calcolo svolto dalla rete neurale. Il risultato a cui si giunge è riportato in Figura 6.2.1-2, ove si osserva in blu l'andamento calcolato dal modello *wide* e in arancione quanto ricavato a valle delle correzioni.

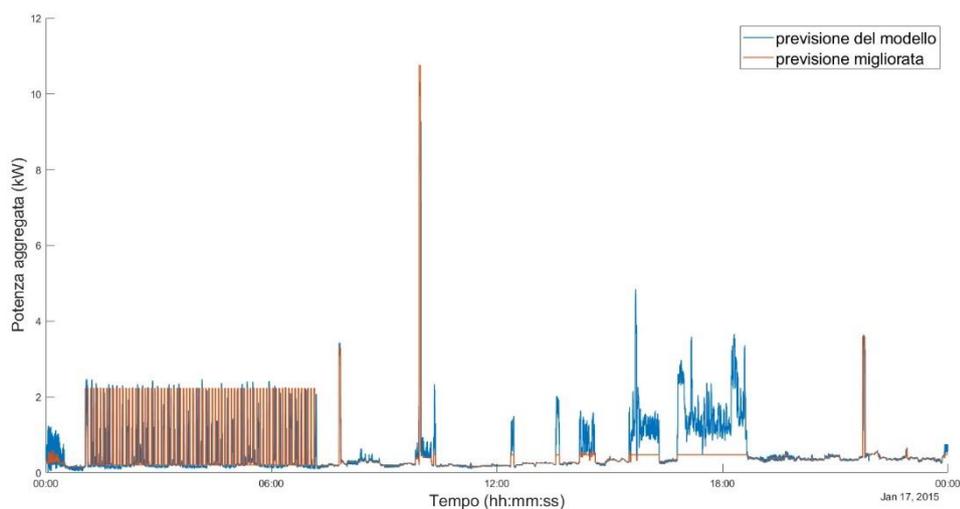


Figura 6.2.1-2 Confronto tra risultato del modello *wide* prima e dopo i miglioramenti apportati

A questo punto, rimane da confrontare la potenza campionata con la previsione finale per poter effettivamente valutare la bontà di quest'ultima; con tale obiettivo, sono state rappresentate le due curve precedenti nel medesimo piano della potenza reale, come mostrato in Figura 6.2.1-3.

Dall'osservazione delle tre curve si giunge subito alla conclusione che le correzioni apportate conducono ad un miglioramento della previsione, come viene confermato dal calcolo dell'errore complessivo. Tale parametro, infatti, si attesta a $5.99 \cdot 10^3$ kW senza correzioni, per decrescere a $4.42 \cdot 10^3$ kW a valle delle stesse. Questa variazione è imputabile all'eliminazione delle variazioni di potenza nelle ore pomeridiane, che dalla Figura 5.3.1-2 si deducono essere inesistenti.

Tuttavia, l'errore viene sostenuto anche dall'inserimento dei carichi caratteristici. Infatti, ad esclusione del carico di tipo A, gli altri sono tutti previsti in istanti scorretti, quindi il contributo all'errore sarà raddoppiato: una volta per l'assenza della previsione, quando il carico è effettivamente utilizzato, e una volta quando esso viene previsto ma non è realmente presente. Questa problematica deriva dal fatto di ipotizzare la potenza aggregata ripetitiva nei giorni.

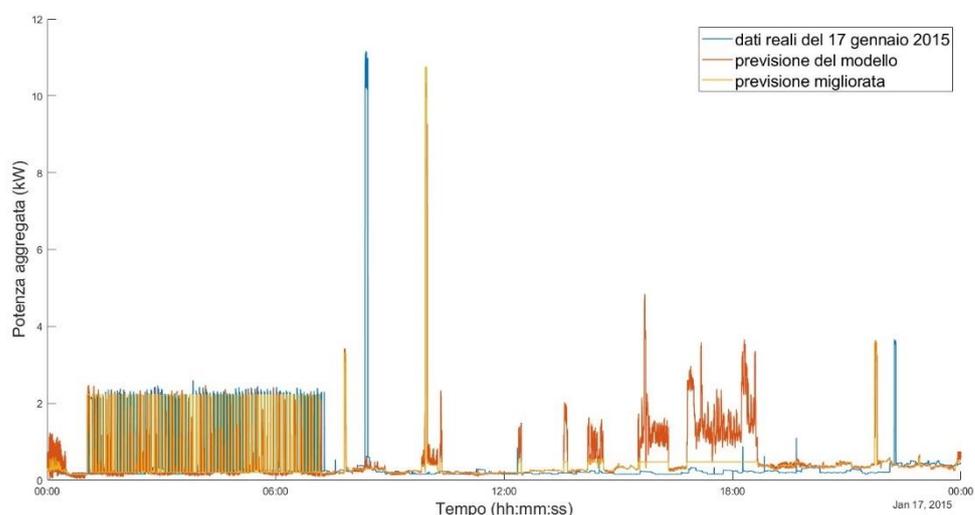


Figura 6.2.1-3 Confronto tra potenza campionata e risultati dei modelli di previsione

Quindi si conclude che anche l'identificazione del carico caratteristico e la sua sostituzione nella previsione da rete neurale hanno un limite, correlato a quanto quest'ultima calcola: il carico infatti potrebbe presentarsi in momenti differenti della giornata o non esserci affatto. Poiché la

sostituzione dei valori viene fatta in virtù del calcolo della rete neurale che, a sua volta, lavora su *predictor* relativi a giornate differenti da quella incognita, la conclusione è nuovamente l'impossibilità di una previsione corretta se non sono fornite delle informazioni aggiuntive e se i carichi non sono effettivamente ripetitivi.

Tuttavia, nella giornata del 17 gennaio non si constata l'impiego di dispositivi particolari durante la giornata; infatti, l'andamento della potenza aggregata misurata (in blu Figura 6.2.1-3) ha molti tratti regolari intervallati da variazioni associabili ai carichi caratteristici analizzati. Per tale motivo, tenuto conto che le variazioni nella previsione da rete neurale che non sono associate a uno di tali carichi vengono livellate, la ricostruzione finale risulta essere più vicina alla realtà rispetto a quanto si era ottenuto per le previsioni del 22 gennaio, essendo la potenza di questa giornata molto più variabile.

Pertanto, una maggiore regolarità del diagramma di carico aiuta a raggiungere un errore minore nella previsione ottenuta.

7. CONCLUSIONI

L'analisi condotta con questa tesi ha avuto inizio dall'intento di ricostruire i diagrammi di carico di un'abitazione impiegando i *toolbox* di *Matlab*, note le misurazioni della potenza totale in momenti differenti da quello incognito. In particolare, si sono impiegate le reti neurali *pre-programmate* fornite dal software, le quali vengono addestrate sulla base dei valori noti della grandezza incognita e dei *predictor* e poi utilizzate per la previsione su un intervallo predefinito.

Tuttavia, durante lo studio, sono venuti alla luce i limiti che questi sistemi di calcolo manifestano a causa della loro stessa natura; essi, infatti, trattano i dati che ricevono in ingresso come numeri, senza tenere conto dell'effettivo fenomeno fisico che tali valori rappresentano. Per tale ragione, meno dettagliate sono le informazioni che ricevono, maggiore è l'errore commesso nelle ricostruzioni. Per questo, ai *predictor* 'temporali', basati sugli istanti di campionamento, si sono aggiunti i valori medi della potenza aggregata, sul minuto e sull'ora.

Nonostante questo, i limiti di cui sopra non vengono meno e incrementano la loro incidenza nel momento in cui la grandezza in analisi perde la sua ripetitività nel tempo, poiché le reti neurali imparano da quanto già è successo in passato: se un evento non si è verificato in altri momenti, il modello non avrà la possibilità di prevederlo perché non contemplato tra i fenomeni che si possono verificare.

Per tali motivazioni, si è reso necessario l'intervento 'umano' derivante da una maggiore conoscenza nel dominio trattato; questo ha permesso di evidenziare la presenza di eventi che si ripetono e che possono essere associati a precisi carichi. Questi, sebbene in istanti differenti ma non temporalmente distanti, causano le medesime variazioni della potenza totalmente richiesta dall'abitazione.

Identificati tali carichi, dei quali tuttavia non si conosce la natura, si è proceduto alla loro caratterizzazione e alla costruzione di un diagramma di carico che ne descrivesse la richiesta di

potenza nel tempo. Noto tale diagramma, è stato possibile apportare un miglioramento delle previsioni da reti neurali procedendo con un inserimento nella previsione stessa, fermo restando che tutto il processo si è basato sull'assunzione che il modello possa prevedere l'istante in cui il carico si manifesta. Questa ipotesi poggia le sue basi sulla verifica della ripetitività della potenza nei giorni: sebbene tale grandezza presenti variazioni non prevedibili, l'analisi dell'evoluzione della potenza giorno per giorno mette in luce un consumo che si ripete nel tempo e, dunque, è lecito pensare ad una ripetizione dell'andamento del carico anche negli istanti incogniti.

La gran parte del lavoro si è svolta sul tentativo di miglioramento delle previsioni condotte su 24 ore, ma il metodo di sostituzione del diagramma di un carico ripetitivo può essere impiegato per ricostruire i dati mancanti anche su intervalli più piccoli. Se infatti, con i dati noti, si riconosce la presenza di un determinato carico in un certo momento, si può semplicemente impiegare l'evoluzione della potenza che lo caratterizza come previsione degli istanti successivi.

In conclusione, i modelli di calcolo forniti dal software *Matlab* sono sicuramente molto avanzati e possono raggiungere dei gradi di accuratezza molto elevati, a patto che si abbia un certo numero di informazioni disponibili nel giorno incognito: maggiore è tale numero, migliore sarà il risultato ottenuto. Inoltre, le previsioni presentano errori minori se la grandezza analizzata si ripete nel tempo perché questo, non solo permette alle reti neurali di utilizzare quanto imparato nella fase di addestramento dai dati precedenti, ma fornisce la possibilità all'esperto di dominio di dare un contributo significativo. Esso, infatti, l'esperto è in grado di poter condurre delle previsioni in autonomia rispetto al software. Tali previsioni diventano maggiormente accurate se la grandezza analizzata ha delle caratteristiche che si ripetono nel tempo.

Si vuole sottolineare come l'idea di ricostruire, con la precisione del secondo, cosa accade in un intervallo temporale in cui i dati non sono disponibili, è una pretesa irrealizzabile perché qualunque evento in esso può avere luogo e rimanere fuori dalla possibilità di previsione sia per il calcolatore sia per l'essere umano.

BIBLIOGRAFIA

[1] “A voltage and current measurement dataset for plug load appliance identification in households”, Roberto Medico, Leen De Baets, Jingkun Gao, Suman Giri, Emre Kara, Tom Dhaene, Chris Develder, Mario Bergés & Dirk Deschrijver, SCIENTIFIC DATA, 2020.

[2] “*GREEND: An Energy Consumption Dataset of Households in Italy and Austria*”, Andrea Monacchi, Dominik Egater, Wilfried Elmenreich, Salvatore D’alessandro, Andrea M. Tonello, IEEE SMARTGRIDCOMM, 2014, Venice, Italy, 3-6 November 2014.

[3] “Data Descriptor: An electrical load measurements dataset of United Kingdom households from a two-year longitudinal study”, David Murray, Lina Stankovic, Vladimir Stankovic, SCIENTIFIC DATA, 35 January 2015.

SITOGRAFIA

[4] *REFIT: Electrical Load Measurements (Cleaned)*, David Murray, Lina Stankovic & Vladimir Stankovic, University Strathclyde Glasgow, <https://pureportal.strath.ac.uk/en/datasets/refit-electrical-load-measurements-cleaned>

[5] *Machine Learning in MATLAB*, MathWorks, https://it.mathworks.com/help/stats/machine-learning-in-matlab.html#responsive_offcanvas

[6] *Regression Learner App*, MathWorks, <https://it.mathworks.com/help/stats/regression-learner-app.html>

[7] *Electricity Load and Price Forecasting Webinar Case Study*, MathWorks, <https://it.mathworks.com/matlabcentral/fileexchange/28684-electricity-load-and-price-forecasting-webinar-case-study>

[8] *Design Time Series NARX Feedback Neural Networks*, MathWorks, <https://it.mathworks.com/help/deeplearning/ug/design-time-series-narx-feedback-neural-networks.html>

RINGRAZIAMENTI

Scrivo questi ringraziamenti non perché lo ritengo necessario bensì perché ci sono persone senza le quali non sarei giunta a questo traguardo e alle quali sento di dovere dedicare almeno queste poche righe.

In primo luogo, ringrazio il **mio relatore** e il **mio correlatore** per avermi offerto l'opportunità di ampliare le mie abilità e conoscenze, guidandomi con utili consigli e suggerimenti.

A seguire, non avrò mai abbastanza parole per ringraziare adeguatamente ogni membro della mia numerosa e meravigliosa famiglia, dal più piccolo al più grande, dal più vicino al più lontano; a loro voglio dedicare l'intero lavoro svolto. Alle mie rocce, **papà e mamma**, per l'amore di cui mi riempite, per la fiducia che riponete in me e per le vostre parole; alla mia unica e insostituibile **zia**, per la tua presenza costante e forte, per la tua attenzione al mio umore, per il sostegno in ogni momento; ai **miei fratelli** per i sorrisi e gli aiuti che mi avete donato, per le nostre interminabili conversazioni che mi hanno portato il buon umore e mi hanno tenuto compagnia in questa fredda città; e ai miei tre **ometti**, così piccoli ma così importanti, ai loro luminosi sorrisi, perché ogni giorno mi ricordate che ci sono tantissimi motivi per non mollare la presa e cercare di rendere questo un posto migliore.

Ringrazio **Elena**, l'Amica che tutti vorrebbero al loro fianco, sostegno costante, di giorno e di notte, nonostante gli anni, nonostante le distanze: per la tua pazienza nell'ascoltare i miei sproloqui, per la tua capacità di calmare le mie ansie, nonché per le follie vissute spalla a spalla.

Non resta che ringraziare te, **Davide**; ti ho messo alla fine non certo per importanza bensì per dedicarti l'attenzione che meriti. Grazie perché rendi la mia vita una successione di giorni meravigliosi, regalandomi ad ogni respiro un ricordo "arancione" da aggiungere agli altri. Ti ringrazio infinitamente perché hai saputo guardarMI e hai creduto in me quando anche io non l'ho fatto. Grazie per essere sempre il mio posto sicuro.

Per ultimo, ringrazio **me stessa** perché, come le "tigri", ho imparato ad atterrare sempre in piedi, sempre fiera di quello che sono diventata.