



Ricerca di Sistema elettrico

Metodi per l'analisi prestazionale di reti di edifici e test su dati sperimentali e simulatore

S. P. Corgnati, A. Capozzoli, I. Khan, F. Lauro, D. Raimondo, N. Talà

METODI PER L'ANALISI PRESTAZIONALE DI RETI DI EDIFICI E TEST SU DATI SPERIMENTALI E SIMULATORE

S. P. Corgnati, A. Capozzoli, I. Khan, F. Lauro, D. Raimondo, N. Talà (Politecnico di Torino, Dipartimento di Energia)

Settembre 2013

Report Ricerca di Sistema Elettrico

Accordo di Programma Ministero dello Sviluppo Economico – ENEA

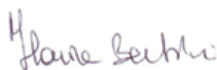
Piano Annuale di Realizzazione 2012

Area: Razionalizzazione e risparmio nell'uso dell'energia elettrica

Progetto: Risparmio di energia elettrica nei settori: civile, industria e servizi

Obiettivo: Gestione ottimale delle reti di edifici

Responsabile del Progetto: Ilaria Bertini, ENEA



Il presente documento descrive le attività di ricerca svolte all'interno dell'Accordo di collaborazione "Monitoraggio e diagnostica remota di reti di edifici"

Responsabile scientifico ENEA: Stefano Pizzuti



Responsabile scientifico del Politecnico di Torino: S.P. Corgnati

Indice

SINTESI DEL RAPPORTO TECNICO	4
DESCRIZIONE DELLE ATTIVITÀ	6
Attività 1	6
<i>Sviluppo e studio di un piano di monitoraggio energetico e ambientale e definizione di un kit di misura ottimale per edifici in contesto urbano con differente destinazione d'uso</i>	6
Caso Studio 1: Scuola elementare.....	9
Monitoraggio energetico	12
Monitoraggio ambientale.....	14
Caso Studio 2: Edificio per uffici.....	16
Monitoraggio energetico	19
Monitoraggio ambientale.....	22
Attività 2	23
<i>Verifica funzionale in fase operativa dei sistemi di monitoraggio in campo operanti su “smartdistrict” e analisi della qualità dei dati raccolti</i>	23
Analisi qualità dei dati n° 1.....	25
Analisi qualità dei dati n° 2.....	31
Attività 3	38
<i>Elaborazione e analisi dati energetici e microclimatici del monitoraggio in campo ai fini diagnostici di “primo livello”: profili di consumo e qualità climatica indoor, tendenze e benchmarks</i>	38
Attività 4	44
<i>Sviluppo di modelli di diagnostica avanzata (“secondo livello”) attraverso i dati del monitoraggio in campo</i>	44
MODELLAZIONE DEI DATI DELL’EDIFICIO F40 E DEL CLUSTER DI EDIFICI PER LA DIAGNOSTICA AVANZATA	44
VERIFICA E APPLICAZIONE DI METODI DI DATA MINING PER ANALISI DI FAULT DETECTION	45
<i>Risultati dell’analisi di fault detection per l’edificio F40 con tecniche di data mining</i>	54
<i>Risultati dell’analisi diagnostica avanzata con tecniche di data mining per il cluster o distretto di edifici</i>	61
VERIFICA E APPLICAZIONE DI TECNICHE DI “NEURAL ENSEMBLING” PER ANALISI DI FAULT DETECTION	89
<i>Risultati dell’ analisi di fault detection per l’edificio F40 con tecniche di neural ensembling</i>	90
<i>Risultati dell’ analisi di fault detection per il cluster o distretto di edifici con tecniche di neural ensembling</i>	94
INDIVIDUAZIONE DI OUTLIER COMUNI INDIVIDUATI ATTRAVERSO METODI DI DATA MINING E BEM NEURALI.....	96
ANALISI DIAGNOSTICA RELATIVA ALL’UTILIZZO ILLUMINAZIONE ARTIFICIALE BASATA SULLA DISPONIBILITÀ DI LUCE NATURALE	98
ANALISI DI OTTIMIZZAZIONE.....	101
CONCLUSIONI	104
RIFERIMENTI BIBLIOGRAFICI.....	106
CURRICULUM VITAE DEGLI AUTORI DEL RAPPORTO TECNICO.....	107
STEFANO PAOLO CORGNATI	107
ALFONSO CAPOZZOLI	107
IMRAN KHAN.....	107
FIORELLA LAURO	107
DANIELA RAIMONDO	108
NOVELLA TALÀ	108
APPENDICE.....	109

Sintesi del rapporto tecnico

La presente relazione illustra le attività e le risultanze frutto degli approfondimenti sviluppati nell'ambito della collaborazione tra ENEA e Politecnico di Torino - Dipartimento Energia sui temi del monitoraggio, della diagnostica remota e dell'ottimizzazione per la gestione energetica degli edifici.

Infatti, attraverso l'accordo di Programma Ministero dello Sviluppo Economico – ENEA sulla Ricerca di Sistema Elettrico, Piano Annuale di Realizzazione 2012, Progetto C1: "Risparmio di energia elettrica nei settori: civile, industria e servizi", si è messo in atto un programma di ricerca di interesse congiunto dal titolo "Sviluppo e studio di un piano di monitoraggio energetico e ambientale e definizione di un kit di misura ottimale per edifici in contesto urbano con differente destinazione d'uso".

Obiettivo della ricerca è quindi quello di mettere a punto e approfondire le tematiche riguardanti il monitoraggio energetico ambientale di insiemi di edifici dal quale ottenere dati per elaborazioni base ed avanzate. Già di per sé l'edificio è un organismo energeticamente complesso costituito da una struttura portante, da un involucro che ne confina l'ambiente interno, da dotazioni impiantistiche finalizzate alla climatizzazione e all'espletamento delle funzioni per le quali l'edificio è costruito e da vettori energetici che lo alimentano. Perciò è corretto parlare di "sistema edificio", intendendolo come organismo complesso e integrato la cui prestazione energetica è legata alle prestazioni delle singole parti costituenti il sistema e dalle reciproche interazioni delle parti che costituiscono l'organismo.

Nel percorso di ricerca sviluppato, il "sistema edificio" ed il monitoraggio delle sue prestazioni energetiche e ambientali è stato oggetto degli approfondimenti effettuati negli anni passati. La ricerca attualmente condotta è da considerarsi la naturale prosecuzione del percorso in atto, già parzialmente cominciata lo scorso anno, muovendosi dal monitoraggio del singolo edificio a quello di un gruppo di edifici, che generalizzando possono divenire un distretto di edifici a diverse destinazioni d'uso. Questo passaggio complessifica significativamente i termini del problema ed apre numerose vie di approfondimento.

Sono state in particolare definite quattro linee di investigazione alla base del programma annuale di ricerca:

- Sviluppo e studio di un piano di monitoraggio energetico e ambientale e definizione di un kit di misura ottimale per edifici in contesto urbano con differente destinazione d'uso (Attività 1)
- Verifica funzionale in fase operativa dei sistemi di monitoraggio in campo operanti su "smartdistrict" e analisi della qualità dei dati raccolti (Attività 2)
- Elaborazione e analisi dati energetici e microclimatici del monitoraggio in campo ai fini diagnostici di "primo livello": profili di consumo e qualità climatica indoor, tendenze e benchmarks (Attività 3)
- Sviluppo di modelli di diagnostica avanzata ("secondo livello") attraverso i dati del monitoraggio in campo (Attività 4).

Ciascuna delle linee di investigazione sopra introdotte è illustrata nel seguito della presente relazione, evidenziandone le finalità, le metodologie di analisi impiegate e i risultati ottenuti.

Una considerazione generale e comune emersa nello sviluppo della attività è stata l'assoluta necessità di mettere a disposizione delle analisi dati di "qualità". Per ottenere ciò è necessario in primis definire un chiaro piano di monitoraggio in grado di raccogliere con continuità i dati essenziali per l'analisi, senza inutili ridondanze che accumulino grandi quantità di dati non utili. A ciò si aggiunge l'affidabilità dell'apparato di misura di restituire dati corretti nel giusto intervallo di precisione: quindi un sistema robusto.

L'attività svolta in questo anno di verifica operativa della qualità dei dati raccolti sullo "smartdistrict" oggetto di studio (Attività 2) è stata essenziale per mettere in atto un processo di miglioramento continuo nella acquisizione, raccolta e fruizione dei dati monitorati. L'attività ha portato ad un crescente e costante miglioramento della qualità dei dati raccolti, e conseguentemente degli studi effettuati. Questa attività è infatti fondamentale per un corretto svolgimento della attività di elaborazione e analisi, sia di base che avanzata, fino allo sviluppo di nuove tecniche diagnostiche real-time.

Questo anno di ricerca ha consentito di aggiungere conoscenze e innovazione intorno al tema del monitoraggio continuo e real-time di reti di edifici, aprendo la visione a nuove tematiche di approfondimento per ricerche future in continuità con le presenti. Tra queste, l'individuazione di kit minimi di strumenti di monitoraggio da mettere in campo per azioni di monitoraggio energetico-ambientale su scala urbana, metodi di analisi real-time sull'affidabilità delle misure condotte, sviluppo di tecniche di benchmarking dinamico e implementazioni di processi di diagnostica avanzata con logiche multi-decisionali.

Descrizione delle attività

Attività 1

Sviluppo e studio di un piano di monitoraggio energetico e ambientale e definizione di un kit di misura ottimale per edifici in contesto urbano con differente destinazione d'uso

L'attività consiste nella definizione delle specifiche di un kit di monitoraggio ottimale dal punto vista tecnico ed economico per lo sviluppo di un piano di monitoraggio di edifici situati in contesto urbano.

La ricerca ha previsto l'identificazione di una metodologia di monitoraggio ambientale ed energetico applicabile in edifici di diversa destinazione d'uso e che, oltre all'attività di misura ed elaborazione dati, permettesse di agire, attraverso specifiche attuazioni, sul controllo degli usi energetici.

La redazione e la conseguente messa in opera di un piano di monitoraggio energetico ed ambientale prevede varie fasi:

1. Innanzitutto deve essere effettuata un'analisi conoscitiva dell'edificio oggetto di studio, allo scopo di evidenziare lo stato di fatto del sistema edificio-impianti, la gestione energetica dell'edificio e i consumi che ne derivano, l'occupazione degli ambienti e le condizioni ambientali di esercizio. Questa prima fase può essere definita fase di audit energetico ed ambientale.
2. A seconda dei risultati ottenuti dalla prima analisi, si può procedere alla fase di definizione di un piano di monitoraggio. Tramite la fase di audit, infatti, si possono evidenziare i percorsi dell'energia elettrica, termica e frigorifera richiesta dall'edificio ai fini dello svolgimento delle attività previste e del mantenimento di determinati livelli di comfort termico, visivo e di qualità dell'aria prefissati.
3. Unitamente al piano di monitoraggio deve essere progettata anche la strategia di attuazione: a seconda del sistema di controllo installato (manuale, semiautomatico, automatico con Building Management System), si deve infatti prevedere in che modo si potrà agire ad esempio sull'accensione o spegnimento degli interruttori di energia elettrica, oppure sulle attivazioni o disattivazioni dell'impianto di condizionamento/climatizzazione.
4. La fase successiva prevede l'installazione del sistema di monitoraggio e attuazione, nonché la sua configurazione su piattaforma di raccolta ed elaborazione dati accessibile da remoto.
5. L'ultima operazione prevede la verifica e validazione del sistema, che si concretizza in un periodo di tempo in cui il sistema va testato allo scopo di ottimizzarlo affinché possa rispondere ai requisiti previsti dal piano di monitoraggio. (Tale fase è meglio approfondita di seguito, all'interno dell'Attività 2).

Il monitoraggio energetico ed ambientale può essere più o meno approfondito a seconda di quelli che sono le finalità per cui questo viene progettato. A seconda del livello di dettaglio che si vuole raggiungere, infatti, il numero dei punti di misura può variare notevolmente e, di conseguenza, anche il costo della messa in opera del sistema si comporta di conseguenza.

Per meglio illustrare il concetto appena introdotto, in Figura 1 e Figura 2 sono rispettivamente illustrati i diversi livelli di approfondimento di analisi in caso di monitoraggio energetico e di monitoraggio

ambientale. Come evidente, l'approfondimento varia in entrambi i casi passando dalla scala di intero edificio alla scala di singolo ambiente.

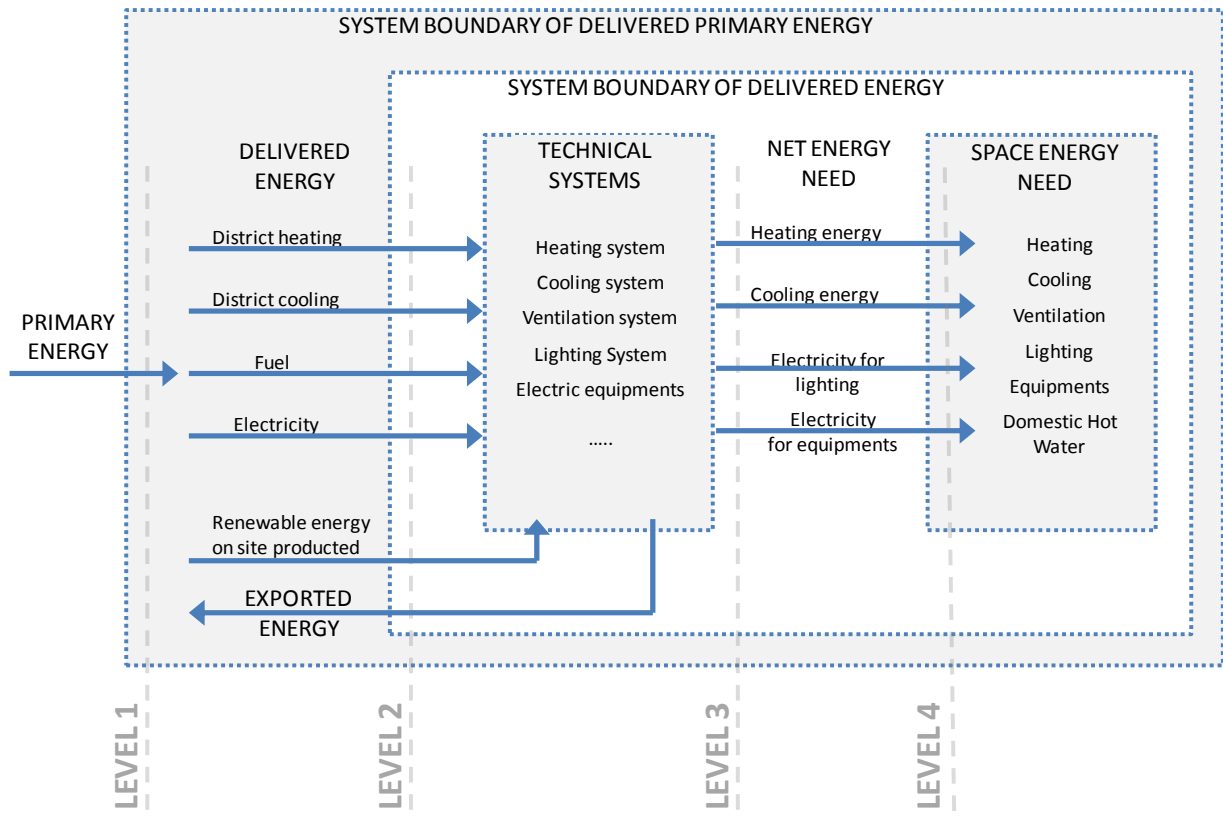


Figura 1: Livelli di dettaglio in caso di monitoraggio energetico.

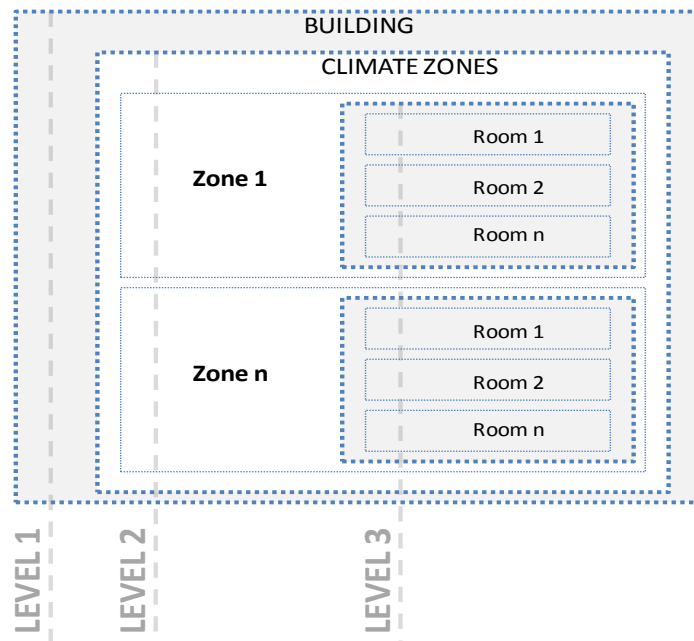


Figura 2: Livelli di dettaglio in caso di monitoraggio ambientale

Tra gli obiettivi dell'Attività 1, come indicato all'inizio del paragrafo, vi è quello di realizzare un sistema di monitoraggio economico e adattabile a diverse tipologie di edifici. Sicuramente questi fattori escludono la possibilità di spingere il livello di dettaglio oltre certi limiti, sia per l'impiego di strumentazione necessaria, sia perché maggiore è il livello di dettaglio, minore sarà la possibilità di adattare il sistema di monitoraggio progettato a più edifici.

Al fine di rispondere alle richieste della ricerca, l'attività si è sviluppata attraverso uno studio sperimentale condotto attraverso l'utilizzo di casi studio: è stata infatti condotta un'attenta analisi prendendo in considerazione un campione rappresentativo di edifici pubblici, all'interno del quale ne sono stati identificati due con destinazione d'uso diversa e ritenuti particolarmente adatti per conseguire gli obiettivi prefissati.

Il primo è una scuola elementare, mentre il secondo è un edificio con destinazione d'uso ufficio. Dato lo scopo dell'attività, i due edifici scelti rappresentano un riferimento per altri edifici della stessa destinazione d'uso. Come viene di seguito dimostrato, i casi studio presentano distribuzioni funzionali che comunemente trovano spazio in altri edifici della stessa categoria. Nel caso studio della scuola vi sono infatti prevalentemente aule scolastiche, poi uffici, laboratori, spazi distributivi, palestra ecc., mentre nel secondo caso studio trovano collocazione uffici di diverse dimensioni, spazi di distribuzione e servizi. Come verrà mostrato ci possono essere suddivisioni diverse degli impianti termici ed elettrici, e, proprio per questo, si è scelto di illustrare un caso in cui la suddivisione avviene per aree funzionali e uno in cui avviene per usi energetici.

Scorrendo i casi studio, risulterà evidente che in entrambi i casi si può dimostrare la ripetibilità dell'applicazione di una metodologia di monitoraggio prestabilita, ovviamente apportando, ove necessario, piccole variazioni affinché il sistema possa essere adeguato al caso specifico. Sia nella scuola che nell'ufficio si è infatti predisposto un monitoraggio a livello di intero edificio, un monitoraggio a valle dei sistemi energetici ed un monitoraggio/attuazione a livello di singole utenze o gruppi di utenze. Entrambi gli edifici sono stati oggetto di audit energetico, a cui è seguita una fase di elaborazione dati che ha consentito di restituire le informazioni reperite attraverso strumenti grafici. Tali strumenti sono mirati ad esprimere schematicamente il sistema edificio-impianti, innanzitutto evidenziando le aree funzionali con le relative superfici, volumetrie ed orari di occupazione, passando poi all'analisi dell'impiantistica per verificare la possibilità di isolare porzioni di edificio tra loro energeticamente differenti.

L'attività ha portato alla definizione di un kit minimo di monitoraggio, di facile installazione, configurazione e di basso impatto, in grado di identificare un equilibrio tra la quantità/qualità della strumentazione installata in campo e i relativi costi, individuando una soluzione capace di operare processi diagnostici efficaci e realisticamente esportabile sul mercato in contesti urbani e applicabile in edifici con diversa destinazione d'uso.

Caso Studio 1: Scuola elementare



Figura 3: Inquadramento dell'edificio.

Il primo caso studio analizzato è una scuola elementare situata all'interno del centro urbano de L'Aquila. L'edificio presenta una geometria ad L e si compone di 2 piani fuori terra ed uno seminterrato. La struttura portante è in calcestruzzo armato, mentre i tamponamenti sono realizzati in laterizio. L'edificio è stato oggetto di risanamento nel 2009. In tale occasione sono stati riprogettati i sistemi elettrici, l'impianto antincendio e l'impianto di illuminazione. Sono stati sostituiti i quadri di derivazione dell'energia elettrica ed è stato inserito un sistema di telecontrollo.

Di seguito si riportano alcune immagini dell'edificio in questione (Figura 4).



Figura 4: Scuola elementare, vista Ovest ed Est.

L'edificio presenta più aree funzionali, sulla base delle quali si ripartiscono le linee dell'energia elettrica. Essendo un edificio scolastico, gli ambienti che occupano la maggiore volumetria sono le aule, seguite dalle

aree distributive e dagli uffici amministrativi. Vi sono poi aule speciali, come laboratori di informatica, biblioteca, ecc., i servizi igienici ed una palestra. (Figura 5)



Figura 5: Principali aree funzionali all'interno dell'edificio.

Essendo un edificio scolastico, gli orari di fruizione dello stesso corrispondono con quelli dello svolgimento dell'attività didattica, il che vuol dire, in questo caso specifico, solo durante le ore del mattino. Solo la palestra viene utilizzata in orari differenti, in quanto si presta anche per attività extrascolastiche.

Il riscaldamento è garantito grazie ad un impianto composto da un solo generatore di calore a servizio di tutto l'edificio, con radiatori come terminali. Non è presente un impianto di raffrescamento, né un impianto per il trattamento dell'aria primaria.

Annessa all'edificio, in un locale del piano interrato, vi è l'abitazione del custode della scuola. In tali ambienti l'energia elettrica è controllata da un interruttore posto all'interno del quadro elettrico generale, ma non viene contabilizzata separatamente rispetto alla totalità dell'edificio. L'energia termica fornita a questi ambienti deriva dall'unico generatore di calore presente nella scuola e non vi è una diramazione della rete di distribuzione dedicata: questo vuol dire che il generatore, in periodo invernale, resta in funzione a pieno regime anche oltre gli orari di fruizione della scuola, compresi gli orari notturni.

L'impianto elettrico è suddiviso principalmente per aree funzionali. Un quadro elettrico generale si trova al piano terra e da questo si dipartono più linee, alcune delle quali raggiungono sottoquadri di zona (al piano terra, al piano primo e in centrale termica).

Di seguito (Figura 6 e Figura 7) si riportano alcune immagini dei quadri elettrici situati in centrale termica, al piano terra e al primo piano, e del generatore di calore in centrale termica.



Figura 6. Quadro elettrico centrale termica, del piano terra e del primo piano



Figura 7. Centrale termica: Generatore di calore e pompe di circolazione

Monitoraggio energetico

Grazie all'operazione di audit energetico è stato possibile risalire alla distribuzione dei flussi energetici, schematizzandola come rappresentato nello schema di Figura 8.

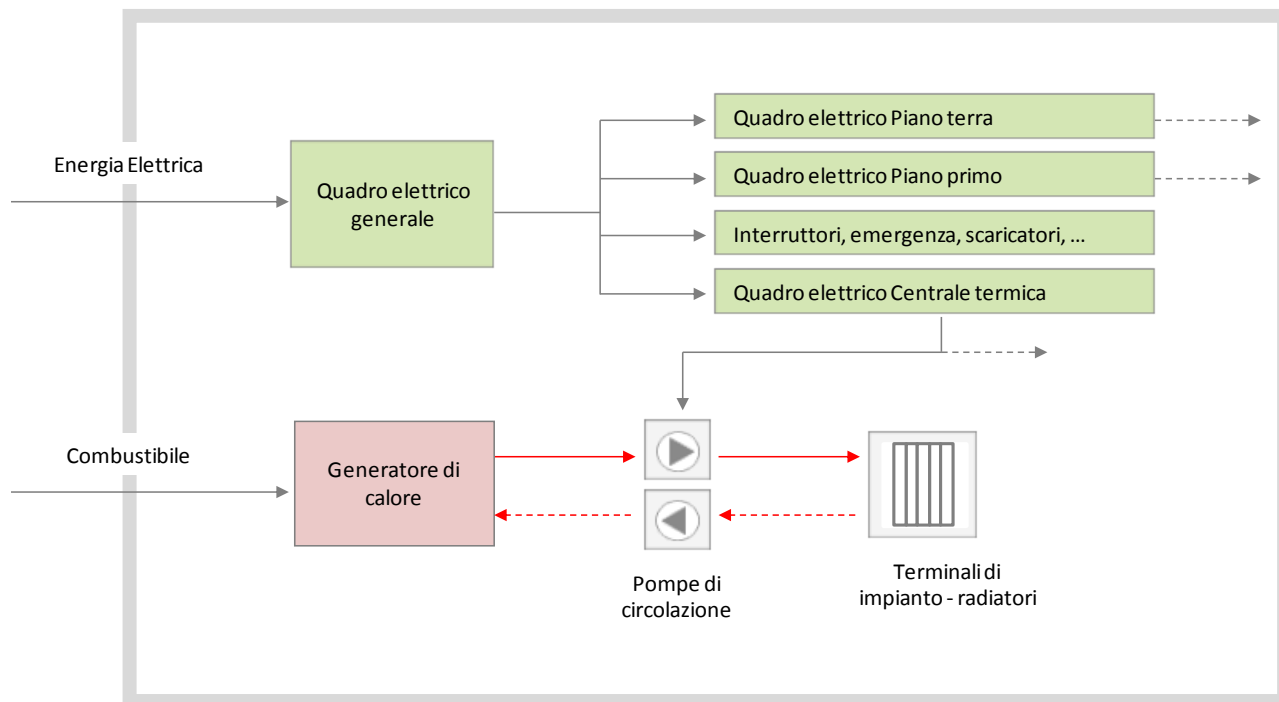


Figura 8: Schema rappresentativo dei vettori energetici

Questo tipo di rappresentazione è un passaggio fondamentale per la definizione del kit minimo: intercettando i vettori energetici con strumenti di misura e attuazione, è infatti possibile individuare in modo veloce ed efficace quelli che saranno gli strumenti da installare.

L'immagine in Figura 9 illustra un dettagliato schema unifilare del percorso che l'energia elettrica fa all'interno dell'edificio. Sulla figura sono anche indicati i punti di misura dell'energia e i punti di attuazione previsti dal piano di monitoraggio.

In questo caso i "pallini" rossi indicano i punti in cui si intende effettuare solo la misura dell'energia, mentre quelli verdi indicano i punti in cui oltre alla misura dell'energia si vuole effettuare anche l'attuazione degli interruttori ai quadri.

Gli strumenti di misura da impiegare in questo caso sono:

- Uno strumento che consenta di rilevare in analogico le letture del contatore posto in prossimità del quadro elettrico generale, in modo da poter monitorare l'energia elettrica totale assorbita dall'edificio.
- 3 powermeters con uscita impulsiva e con inserzione indiretta a livello di quadro elettrico di Piano Terra, Piano primo e Centrale Termica, e altri 4 powermeters con le stesse caratteristiche dei precedenti da inserire come indicato negli interruttori delle derivazioni del piano terra.

La scelta di determinati punti di attuazione rispetto ad altri dipende in molti casi dalla specifica costruzione dell'impianto elettrico. In questo caso, ad esempio, metodologicamente, le logiche di attuazione

vorrebbero disattivare gli interruttori generali di ogni sottoquadro quando l'area funzionale da essi servita non viene utilizzata. Ciò vuol dire che, fuori dall'orario scolastico, la condizione ottimale sarebbe quella di disattivare la corrente elettrica in tutti gli ambienti, tranne la palestra. Il problema che si è riscontrato in questo caso specifico è che il sistema di emergenza e il sistema antincendio non avevano una linea dedicata e separata rispetto al resto dell'impianto. Per tale motivo, inserire un'attuazione sull'interruttore del piano terra sarebbe stato inutile in quanto disattivare tale interruttore significherebbe disattivare anche l'impianto antincendio. Le attuazioni sono state quindi inserite su tutti gli altri interruttori del quadro del piano terra, in modo da poter disattivare l'energia elettrica in aule e uffici al di fuori dell'orario scolastico, in palestra durante le ore notturne, e in modo da non poter interferire in alcun modo sull'impianto antincendio.

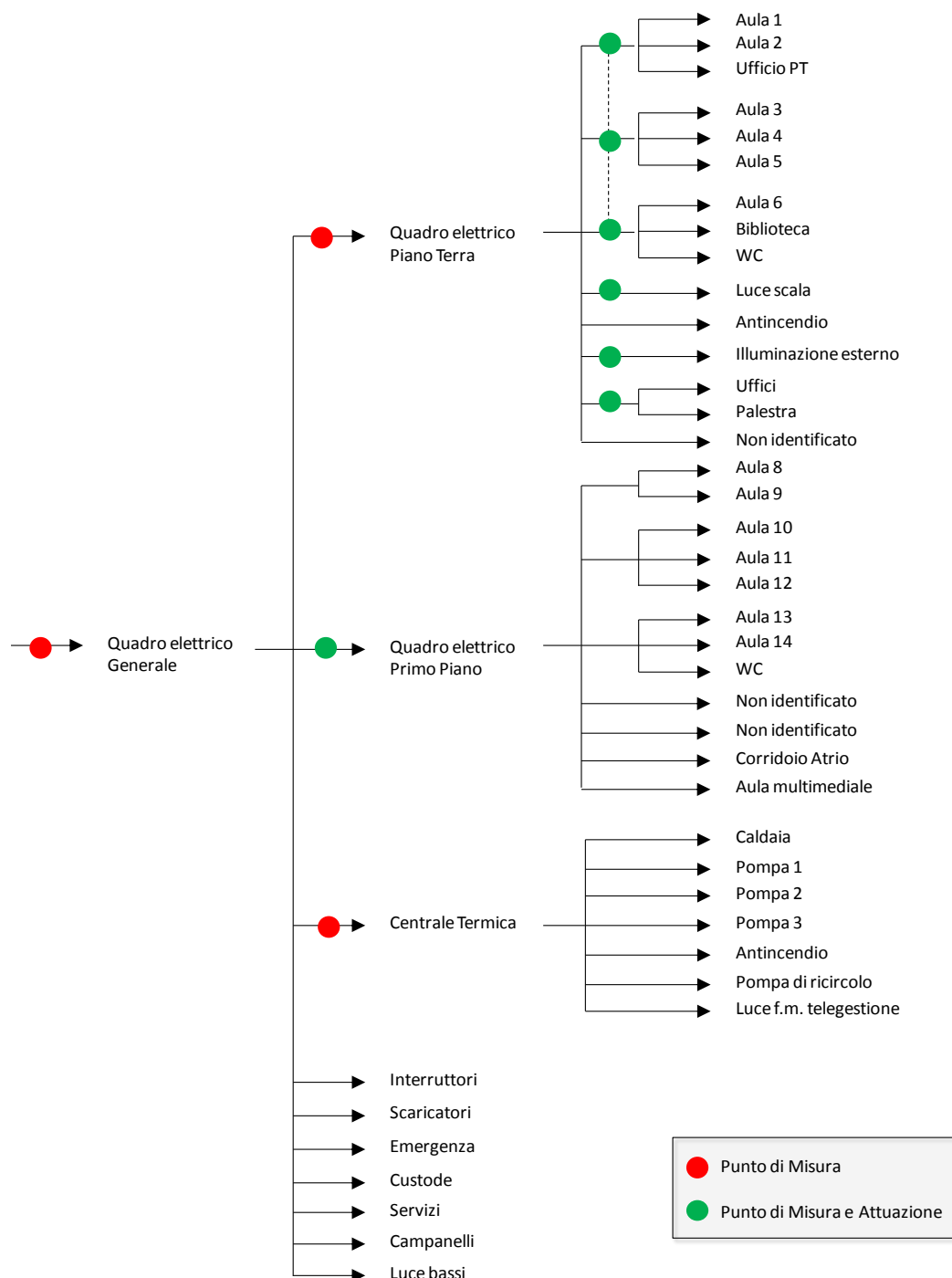


Figura 9: Percorso e distribuzione dell'energia elettrica all'interno della scuola elementare

Situazione diversa si presenta invece al primo piano: l'impianto antincendio in questo caso è alimentato dal quadro del piano terra e, dunque, l'interruttore generale del sottoquadro del primo piano può essere disattivato quando nelle aule non c'è nessuno. In centrale termica invece vale la regola del piano terra. Proprio per tale ragione in centrale termica si è deciso di monitorare solo l'energia elettrica totale a livello di quadro, senza andare ad inserire attuazioni di alcun tipo.

Tale approccio evidenzia come, in alcuni casi, la metodologia proposta possa subire delle piccole variazioni dettate dalle esigenze di strutturazione dell'impianto. In taluni casi è comunque più conveniente, sia a livello tecnico che economico, apportare delle modifiche all'impianto (e ri-certificarlo), piuttosto che eccedere con l'uso di strumenti di monitoraggio.

A livello di energia termica il monitoraggio si presenta in questo caso molto semplice. Come indicato in Figura 10 i punti di misura in questo caso sono 2: il primo ha lo scopo di misurare la quantità di combustibile in ingresso al generatore di calore, mentre il secondo di misurare l'energia termica in uscita dal generatore. Per questo nel primo caso, se il contatore del gas metano è già dotato di comparatore elettronico, basterà richiedere l'uscita impulsiva e registrare gli impulsi, mentre nel secondo caso si deve provvedere con l'installazione di un contabilizzatore di calore da installare sui tubi di mandata e ritorno del fluido caldo in uscita/entrata del generatore di calore.

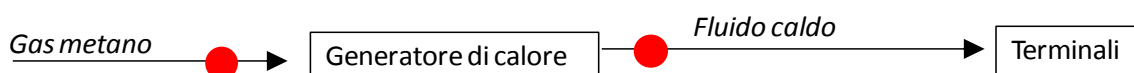


Figura 10: Percorso e distribuzione dell'energia termica all'interno della scuola elementare

In questo caso le attuazioni vengono realizzate attraverso l'inserimento di elettrovalvole e non si prevede nessun intervento di variazione dell'impianto esistente.

Il problema introdotto relativamente alla presenza dell'abitazione del custode della scuola può essere risolto in due modi, entrambi in fase di valutazione: il primo è il trasferimento del custode in un altro edificio (operazione in programma per l'anno seguente), mentre il secondo è quello di provvedere i locali adibiti ad abitazione di un sistema di riscaldamento autonomo con gestione separata. Dato che queste soluzioni sono in corso di valutazione, il piano di monitoraggio energetico della scuola potrebbe ancora subire piccole variazioni nel periodo precedente la fase di installazione.

Monitoraggio ambientale

Come definito per il monitoraggio energetico, anche il monitoraggio ambientale avviene per aree funzionali. Per monitoraggio ambientale, in questo specifico caso, si fa riferimento alla sola analisi di temperatura ed umidità relativa in punti diversi della struttura, al fine di verificare che il comfort termo igrometrico sia mantenuto all'interno degli ambienti secondo quanto prescritto dalla normativa internazionale in materia di comfort EN ISO 15251:2008.

In Figura 11 sono illustrati i punti in cui da piano di monitoraggio si prevede di inserire i sensori di temperatura e umidità relativa. Come indicato si intendono monitorare tre aree funzionali diverse, ossia gli uffici, la palestra e le aule. Dato che le aule sono però numerose, situate su piani diversi, e con orientamenti diversi, si prevede di monitorare aule significative al fine di avere una mappatura completa del comfort termo igrometrico negli ambienti.



Figura 11: Monitoraggio ambientale

Essendo gli ambienti scolastici, come si è introdotto precedentemente, ambienti occupati durante orari ben precisi, si prevedono attuazioni regolate a seconda degli orari. Non è dunque necessario inserire in questa tipologia di edifici dei sensori di presenza al fine dell'attuazione. I sensori di presenza, se si vogliono inserire, possono essere utili al solo fine dell'attivazione/spegnimento degli apparecchi illuminanti in zone distributive e servizi igienici.

Tutti gli strumenti previsti dal piano di monitoraggio trasmettono i dati via wifi ad un gateway che li invierà ad una piattaforma di acquisizione, elaborazione e controllo da remoto.

Caso Studio 2: Edificio per uffici



Figura 12: Inquadramento del cluster di edifici.

L'edificio con destinazione d'uso ufficio preso in esame fa parte di un complesso di 8 edifici ed è situato all'interno del Centro Ricerche ENEA Casaccia. Come visibile in Figura 12 gli edifici sono dislocati in due blocchi che si attestano su un viale. Un primo blocco costituito da tre edifici contigui è orientato secondo l'asse NO-SE, mentre il secondo blocco costituito da 5 edifici a pettine ha come orientamento principale l'asse NE-SO. L'edificio che è stato preso in analisi fa parte di quest'ultimo blocco. L'edificio ha forma rettangolare, si compone di un unico piano fuori terra, ed è realizzato in calcestruzzo armato. Le pareti esterne hanno uno spessore di 30 cm, ed al loro interno ospitano una pannellatura di materiale isolante di circa 5 cm. La facciata esterna è in calcestruzzo a vista, mentre all'interno il cls è rivestito in cartongesso e polistirolo di circa 3 cm. Gli infissi sono scorrevoli e con intelaiatura in alluminio. Il controllo della radiazione solare negli uffici avviene attraverso tende veneziane esterne, mentre nei bagni sono installate delle persiane fisse in alluminio sul lato esterno.

Di seguito si riportano alcune immagini dell'edificio in questione (Figura 13)



Figura 13: Edificio per uffici: facciata longitudinale e vista dell'ingresso.

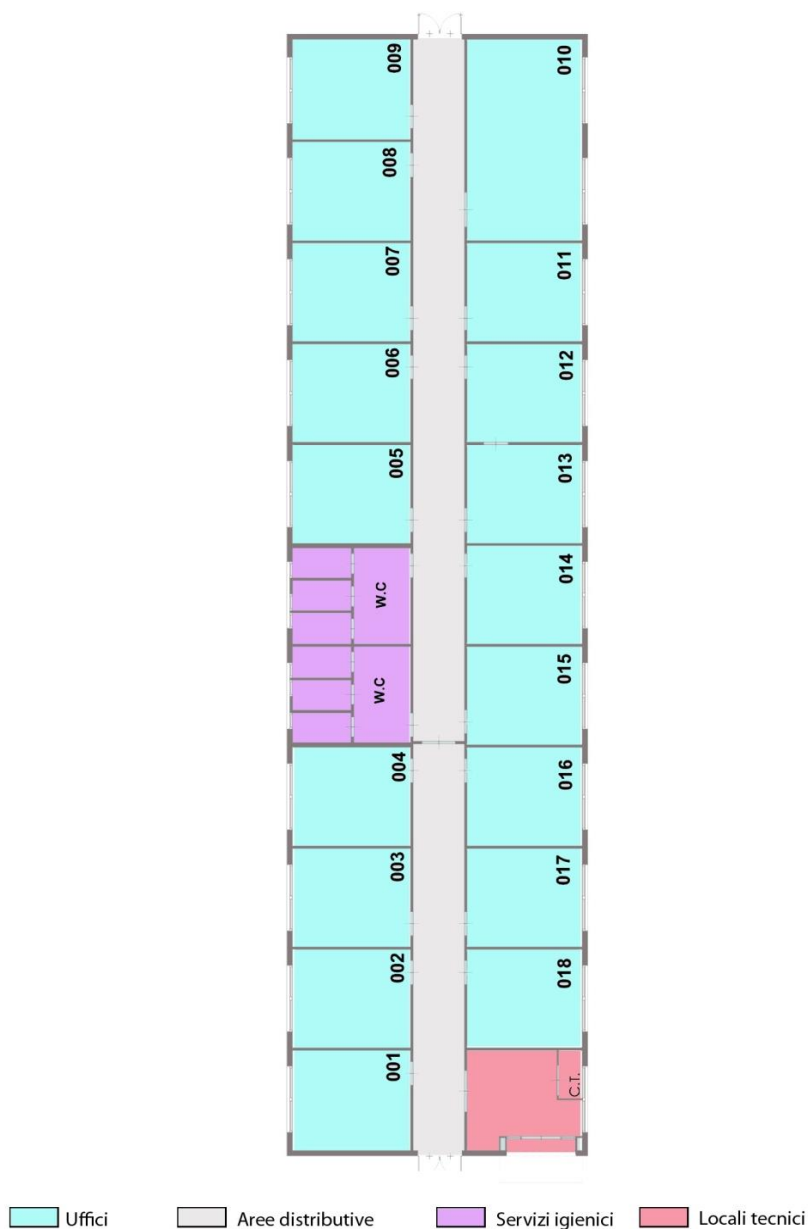


Figura 14: Principali funzioni all'interno dell'edificio.

Come evidenziato in Figura 14 l'edificio ha la sola funzione di ufficio. Le altre aree funzionali evidenziate nell'immagine sono le aree distributive e i servizi igienici. All'ingresso vi è un locale tecnico dove trova collocazione la diramazione dell'impianto di riscaldamento /raffrescamento, l'unità di trattamento aria e il quadro elettrico di edificio.

Dal punto di vista impiantistico tutti gli edifici facenti parte del cluster sono serviti da un'unica centrale termica, dislocata presso un locale tecnico in un fabbricato dedicato, e da una cabina elettrica (che serve questi 8 edifici più un altro). In centrale termica uno scambiatore di calore ed un gruppo frigo producono il fluido caldo/freddo che alimenta un circuito idronico sotterraneo a servizio di tutto il cluster. Ogni edificio è poi provvisto di una sottocentrale termica, in cui il fluido caldo/freddo viene prelevato dal circuito principale ed inserito nei due circuiti dell'impianto di riscaldamento dell'edificio

(destro e sinistro) attraverso un'elettropompa circolare con portata di 200l/1-3 min. collocata sulla tubazione di ritorno del fluido caldo/freddo. I terminali di stanza sono fancoils a 2 tubi. Ogni singolo ufficio è riscaldato attraverso fancoils e l'aria primaria è garantita attraverso un impianto di ventilazione con diffusore in ambiente. Anche nei corridoi il riscaldamento/raffrescamento è garantito da fancoils. L'unità di trattamento aria della sottocentrale immette negli ambienti una portata di aria esterna totale pari a 1750 m³/h. Non vi è recupero di calore, né estrazione forzata di aria dagli ambienti.

All'interno dell'edificio vi è un solo quadro elettrico, che comprende un contatore generale dell'energia elettrica totale fornita all'edificio, più una suddivisione delle linee generale luce, forza motrice e generale condizionamento.

Per maggiori informazioni relativamente al sistema impiantistico del cluster di edifici si rimanda al documento *"Sviluppo di una metodologia per la diagnostica remota di reti di edifici integrata nella rete della pubblica illuminazione: applicazione su smartvillage C.R. Casaccia"* ENEA (2012).



Figura 15: Quadro elettrico e sottocentrale termica dell'edificio.

Monitoraggio energetico

Anche in questo caso l'intero sistema impiantistico è stato schematizzato al fine di evidenziare i percorsi e le trasformazioni dell'energia. Per meglio spiegare questi percorsi, in Figura 16 si riporta non solo la schematizzazione del sistema impiantistico della sottocentrale termica di edificio, ma anche quello della centrale termica generale in cui avviene la produzione del fluido caldo/freddo.

È da sottolineare che nel sistema di sottocentrale il piano di monitoraggio ha previsto l'inserimento di una nuova elettrovalvola a 3 vie a valle dello spillamento del circuito principale, in modo da poter agire con le attuazioni direttamente su questa consentendo di bloccare il passaggio del fluido durante gli orari prefissati.

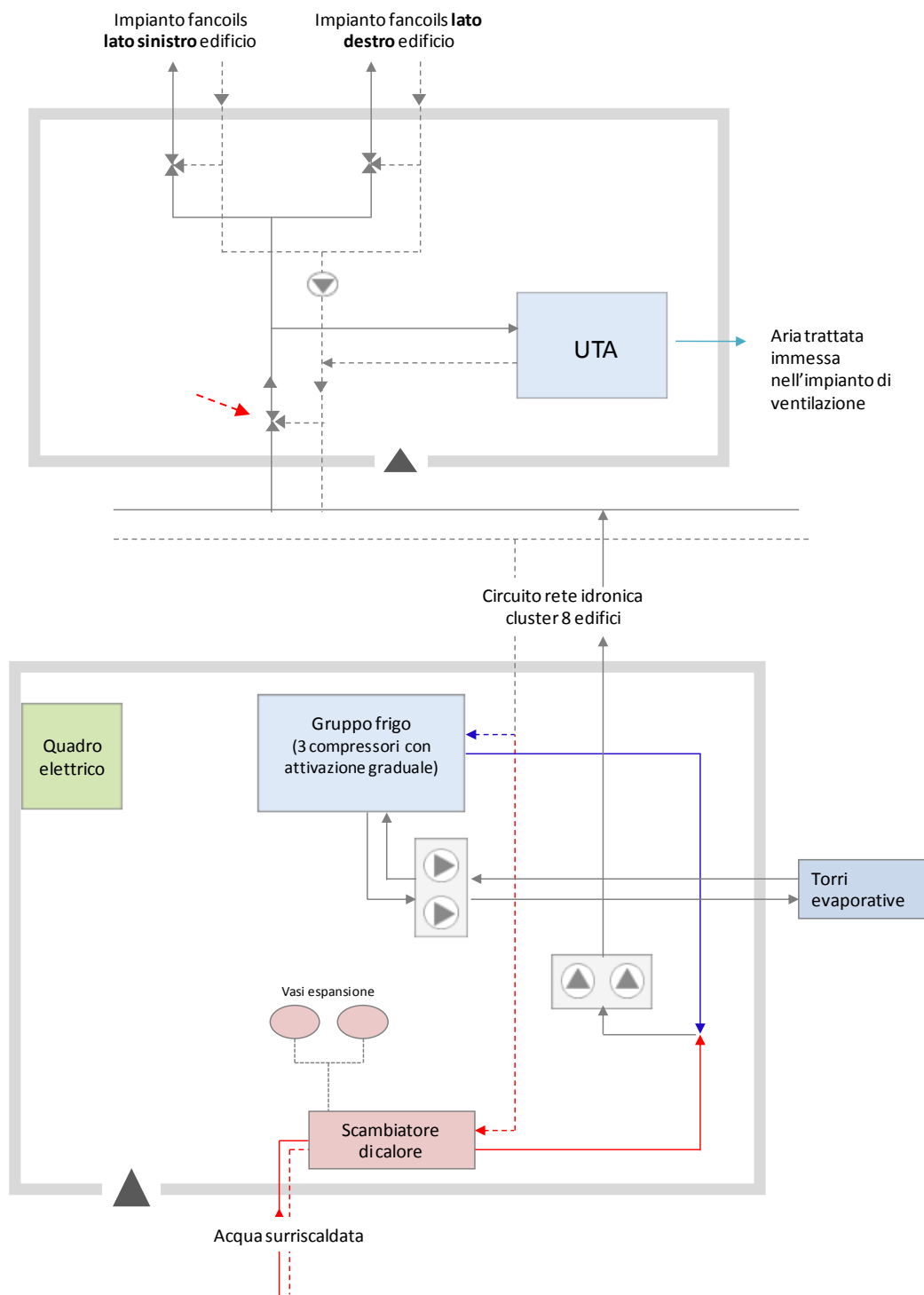


Figura 16: Schema dei vettori energetici dalla centrale termica all'edificio oggetto di studio

Come anticipato precedentemente, a differenza del caso studio relativo alla scuola elementare, essendoci in questo caso una sola area funzionale specifica, gli interruttori al quadro sono divisi a seconda degli usi elettrici: per l'appunto illuminazione, forza motrice e condizionamento. Le attuazioni in questo caso presenteranno una logica molto diversa da quella illustrata nel caso precedente.

Il piano di monitoraggio prevede infatti, in questo caso, di installare strumenti di misura e attuazione direttamente sulle linee relative all'illuminazione e condizionamento, e ricavando l'informazione relativa alla forza motrice per differenza con l'energia elettrica totale fornita all'edificio, (Figura 17).

Le attuazioni sono poste solo sulle linee luce e condizionamento. Come accadeva nel caso studio precedente, infatti, la linea di emergenza non deve mai essere disattivata e in questo caso tale linea si trova sotto la voce "forza motrice".

L'attuazione degli interruttori della luce dipenderà dalla presenza o meno di persone all'interno dell'edificio: questa sarà monitorata principalmente attraverso la registrazione delle timbrature dei badges personali dei dipendenti ai tornelli d'ingresso del complesso di ENEA, ma il piano prevede anche l'installazione di un sensore di presenza all'ingresso principale dell'edificio. Per quanto riguarda l'attuazione della linea di condizionamento, invece, il controllo sarà principalmente orario per l'attivazione, mentre dipenderà dall'occupazione degli ambienti per lo spegnimento: in tal modo si garantisce l'accensione dell'impianto per il pre-riscaldamento/raffrescamento degli ambienti al mattino, permettendo all'aria di raggiungere la temperatura di comfort all'arrivo dei dipendenti.

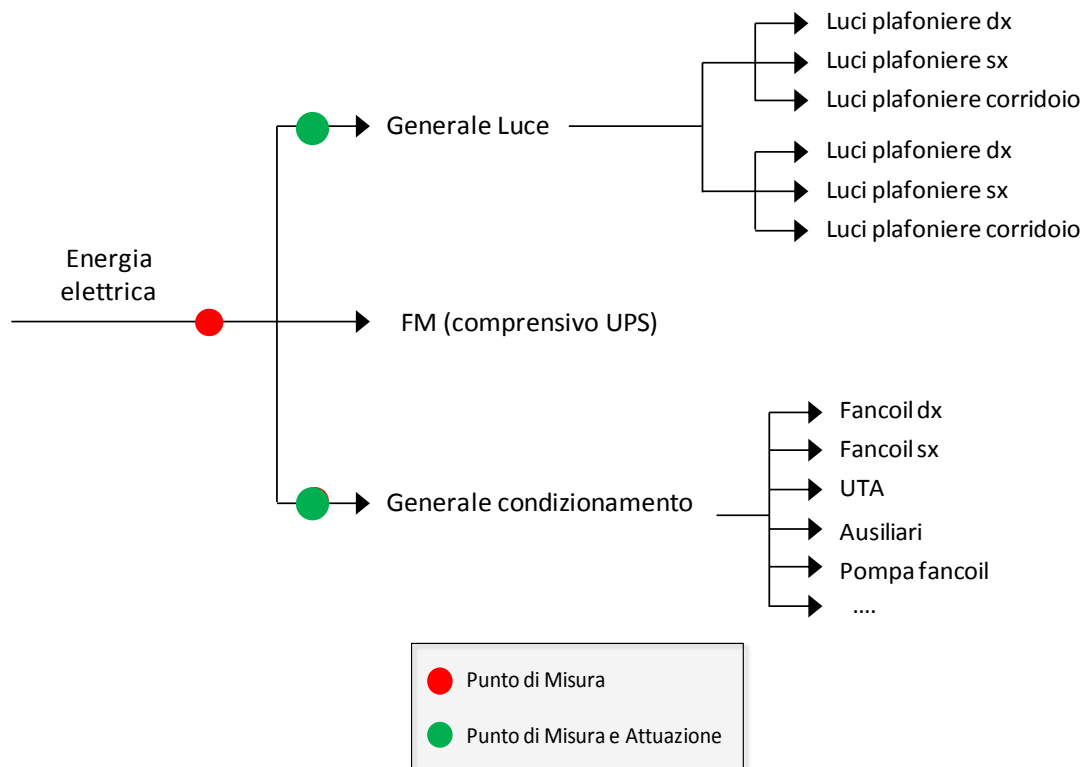


Figura 17: Schema unifilare energia elettrica dell'edificio

Il piano di monitoraggio dell'energia termica prevede l'installazione di un contabilizzatore di calore da installarsi sulle condotte di mandata e di ritorno del circuito idronico dell'edificio, a valle dello spillamento dal circuito del cluster e a monte della suddivisione dell'impianto di edificio in due diramazioni(Figura 18). Il contabilizzatore di calore prevede l'inserimento di un contatore volumetrico di portata del fluido, più due sonde di temperatura (PT100) da inserire all'interno delle condotte di mandata e di ritorno, per valutare la differenza di temperatura dei due fluidi e quindi il calore ceduto/guadagnato all'ambiente.

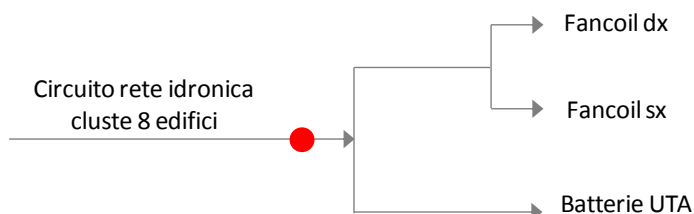


Figura 18: Schema unifilare in sottocentrale di edificio

Monitoraggio ambientale

Il monitoraggio ambientale prevede il posizionamento di sonde per la misurazione di temperatura dell'aria e umidità relativa in almeno 2 punti rappresentativi dell'edificio. Tali punti rappresentativi sono stati evidenziati in 2 uffici di uguale superficie ma con esposizioni diverse (Figura 19).

La presenza delle persone all'interno degli edifici, come già evidenziato, sarà monitorata principalmente attraverso la registrazione delle timbrature dei badges personali dei dipendenti ai tornelli di ingresso del complesso di ENEA. Il sistema prevede però l'installazione di un sensore di presenza all'ingresso principale di ogni edificio, al fine soprattutto dell'attivazione della luce esterna.

Come nel caso precedente, anche qui la strategia di comunicazione, acquisizione dati e controllo è interamente wireless e gestita da remoto.

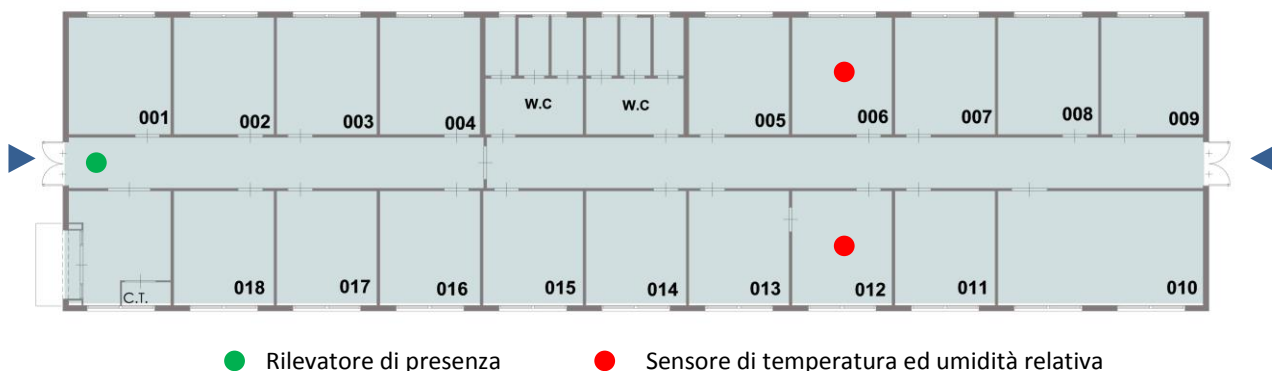


Figura 19: Posizionamento sensori di temperatura e rilevatori di presenza all'interno dell'edificio

Attività 2

Verifica funzionale in fase operativa dei sistemi di monitoraggio in campo operanti su “smartdistrict” e analisi della qualità dei dati raccolti.

L’attività consiste nella verifica del corretto funzionamento della sensoristica del sistema di monitoraggio energetico e ambientale indoor implementato sul distretto di edifici presso il centro ENEA della Casaccia.

Il piano di monitoraggio energetico ed ambientale sviluppato per il complesso di edifici, nonché i criteri che hanno portato alla sua definizione, sono stati ampiamente illustrati all’interno del documento “Sviluppo di una metodologia per la diagnostica remota di reti di edifici integrata nella rete della pubblica illuminazione: applicazione su smartvillage C.R. Casaccia” ENEA (2012).

Questa seconda sezione del documento si concentra sulla verifica della corrispondenza tra tale piano di monitoraggio e l’installazione/configurazione della sensoristica inserita all’interno/esterno degli otto edifici componenti il cluster (e connessa centrale termica del distretto).

Dall’analisi effettuata si sono evidenziate differenze rispetto al piano di monitoraggio solo per quanto riguarda la misura dei parametri ambientali: mentre nella versione originale per ogni edificio erano previsti due sensori di temperatura ed umidità relativa da disporsi in locali rappresentativi con diverse esposizioni, la nuova configurazione presenta un solo sensore di temperatura situato al centro del corridoio di ciascun edificio, non risultando quindi rappresentativo della qualità termo igrometrica indoor. Anche il rilevatore di presenza previsto dal piano non è stato installato: la presenza degli occupanti all’interno dell’edificio è determinata attraverso il passaggio dei badge personali al tornello di ingresso.

In Figura 20 si illustra l’attuale configurazione dei punti di monitoraggio rispetto al piano di monitoraggio.

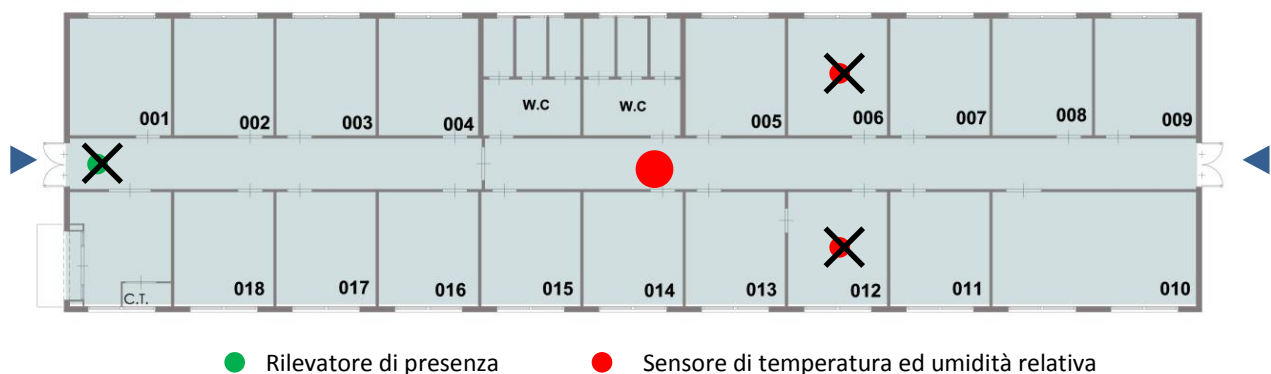


Figura 20: Variazione dei punti di monitoraggio rispetto al piano di installazione

Oltre alla disposizione fisica dei sensori l’attività ha anche previsto la verifica del loro corretto funzionamento. Per tale ragione ogni sensore è stato analizzato separatamente: è stata verificata l’accuratezza della grandezza misurata, il periodo di campionamento, la frequenza temporale di acquisizione, il numero totale di record registrati durante il periodo di monitoraggio, la quantità di dati mancanti e l’attendibilità dei dati acquisiti. I dati collezionati dalla piattaforma di acquisizione sono stati estratti ed elaborati fuori linea. Per ogni edificio si sono analizzati i dati di energia termica, energia elettrica e i parametri ambientali, per verificarne il loro possibile impiego nelle successive fasi di elaborazione ai fini del benchmarking e della diagnostica avanzata (Attività 3 e 4).

Il processo di verifica della qualità dei dati è un'operazione che si ripete periodicamente al fine di ottimizzare il sistema di monitoraggio installato, e che si struttura come illustrato nello schema di Figura 21.

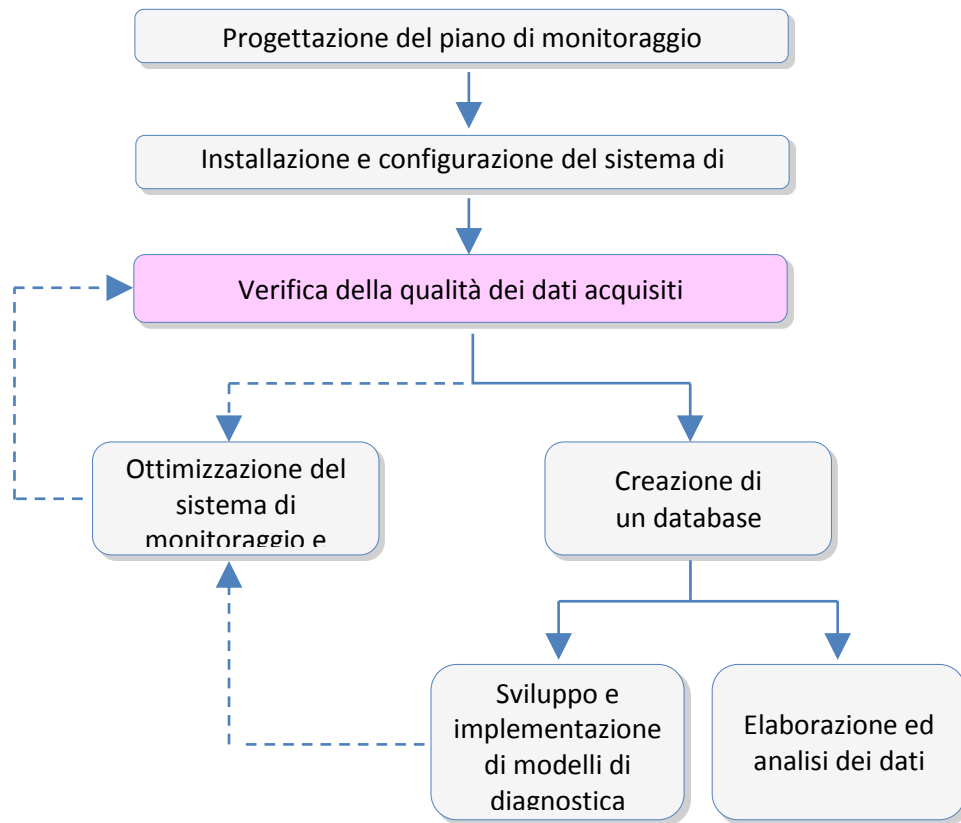


Figura 21: Processo ottimizzazione del sistema di monitoraggio ed elaborazione dati

In Figura 21 è illustrato l'intero processo di creazione ed ottimizzazione del sistema di monitoraggio ed elaborazione dati. Come indicato, successivamente alle prime fasi di redazione di un piano di monitoraggio energetico ed ambientale e la sua effettiva messa in opera e configurazione, inizia un processo di analisi dati destinato all'ottimizzazione del sistema. Una volta che il sistema è stato configurato e l'acquisizione dei dati di monitoraggio è iniziata, si passa ad una fase di analisi dei dati sia dal punto di vista qualitativo che quantitativo. Occorre, infatti, verificare che i parametri descritti all'interno del piano di monitoraggio siano rispettati (ad esempio la frequenza di acquisizione e l'assenza di dati mancanti), dopodiché è necessario analizzare l'effettiva coerenza ed attendibilità dei dati di monitoraggio rispetto ai risultati attesi. Nel caso in cui quest'analisi evidenzi delle anomalie, queste devono essere corrette e, solo successivamente, si potrà procedere alle fasi di elaborazione successive. Nel caso in cui il sistema funzioni correttamente, i dati di monitoraggio andranno a costituire un database che costituirà la base delle informazioni utili alle fasi di elaborazione ed analisi dati e di sviluppo di modelli di diagnostica. Inoltre, attraverso l'implementazione nel sistema dei modelli di diagnostica, è possibile evidenziare eventuali anomalie del sistema di monitoraggio: anche in questo caso i warnings, outputs della fase di diagnostica, diventeranno inputs per l'ottimizzazione del sistema. È molto importante sottolineare che questo processo di verifica della qualità dei dati dovrà essere ripetuto svariate volte, fintanto che l'intero sistema di monitoraggio non risponderà a tutti i requisiti di partenza. Sarà oggetto di approfondimento futuro l'identificazione del tempo standard di ottimizzazione del sistema sulla base di diverse tipologie di monitoraggio e di diverse tipologie di edificio o distretto analizzato.

Analisi qualità dei dati n° 1.

La prima fase di analisi dei dati è consistita nello studio dei profili temporali delle grandezze misurate tra il 24-10-2012 e il 01-07-2013. Per ogni singolo edificio sono stati analizzati puntualmente i profili di tutte le grandezze e si sono evidenziati gli errori presenti nel sistema che necessitavano particolare attenzione e che richiedevano correzioni. Ogni edificio ha presentato anomalie ed errori diversi, che sono riassunti successivamente in Tabella 1. Per illustrare la metodologia applicata, a titolo di esempio sono riportati di seguito i vari passaggi dell'analisi effettuata per l'edificio F66. In Figura 22 sono riportati i profili di energia termica (rosso) ed energia frigorifera (blu) restituiti dal sistema di monitoraggio. I profili evidenziano anomalie nel sistema: picchi di potenza molto elevati si alternano a valori molto bassi e quasi costanti.

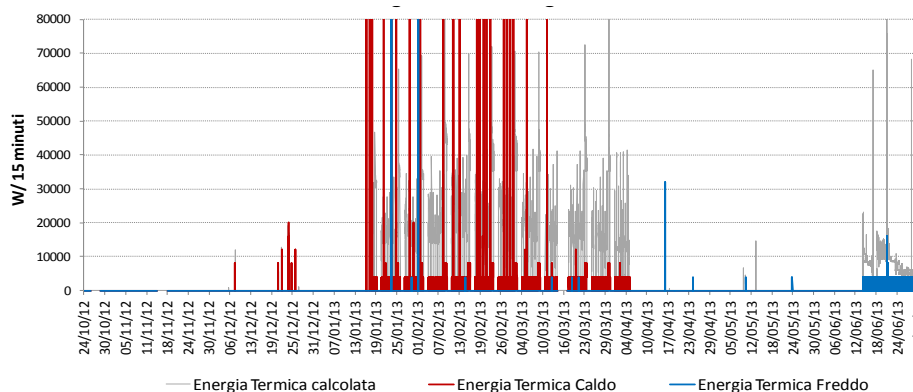


Figura 22: Energia termica ed energia frigorifera (24/10/2012 – 01/07/2013)

Confrontando tali profili con i valori di energia termica derivati applicando l'equazione

$$Q/A = m \cdot c_p \cdot \Delta T \quad (1)$$

dove: m = portata di fluido

c_p = calore specifico dell'acqua

ΔT = differenza tra temperatura di mandata e di ritorno all'interno del circuito idronico

si è ottenuto il profilo indicato in Figura 22 con il colore grigio. Come si può notare dal grafico, la differenza tra i profili di energia termica misurata e derivata è evidente. Questo fattore evidenzia un'anomalia nel sistema, che dunque necessita correzione. Inoltre, analizzando i profili di (m) e di (ΔT) riportati in figura Figura 23, si nota come non si tratti di un errore di misura dei singoli parametri, quindi non imputabile ai sensori, ma sia probabilmente attribuibile alle opzioni di impostazione dell'apparato elettronico del contabilizzatore di calore.

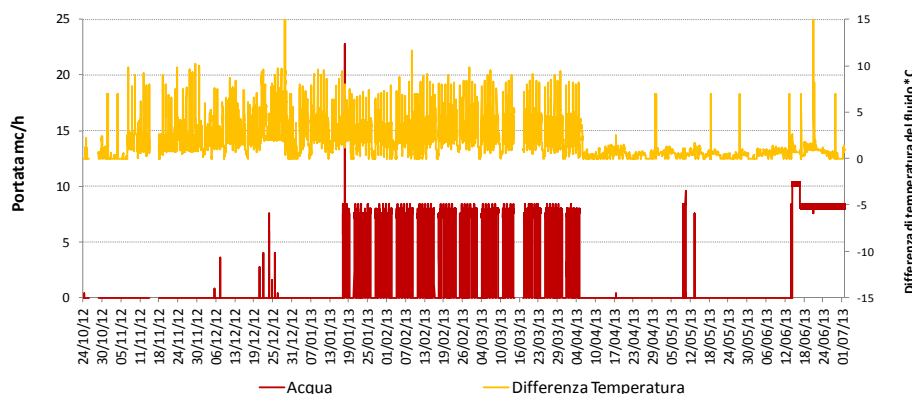


Figura 23: Portata d'acqua nel circuito idronico a servizio dell'edificio e differenza di temperatura tra l'acqua di ritorno dall'impianto del singolo edificio e l'acqua di mandata in uscita dalla CT (24/10/2012 – 01/07/2013)

Per quanto riguarda il monitoraggio dell'energia elettrica a livello di singolo edificio sono stati rilevati la Potenza attiva e la Potenza reattiva relativamente all'intero edificio, la potenza attiva dell'impianto di illuminazione e dell'impianto di condizionamento. La Potenza attiva relativa alla Forza Motrice è stata derivati per differenza

$$Pa_{FM} = [Pa_{totale} - (Pa_{illuminazione} + Pa_{condizionamento})] \tag{2}$$

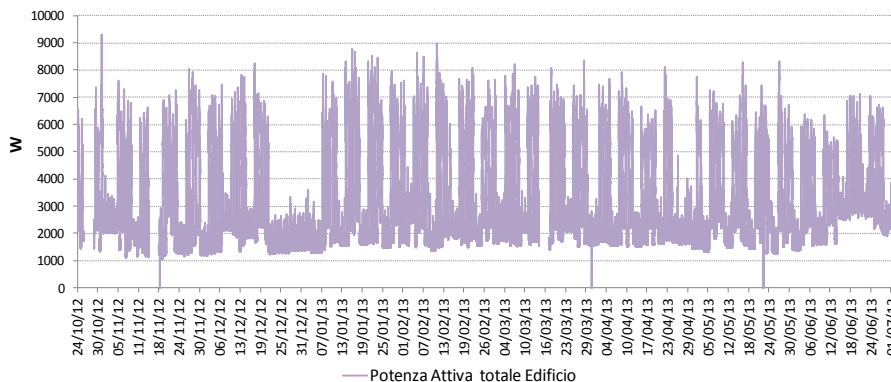


Figura 24: Potenza attiva totale dell'edificio (24/10/2012 – 01/07/2013)

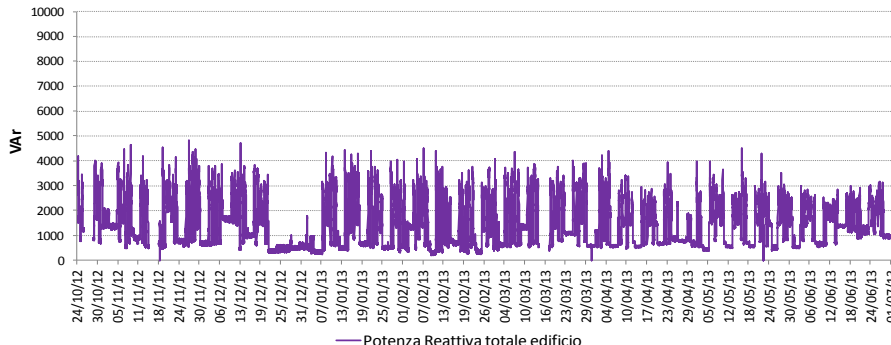


Figura 25: Potenza reattiva totale dell'edificio (24/10/2012 – 01/07/2013)

I profili di Potenza attiva e reattiva riportati rispettivamente in Figura 24 e Figura 25 si presentano coerenti riguardo alle aspettative e, inoltre, i dati mancanti sono inferiori al 4%. In questo caso non si sono riscontrate anomalie nel sistema.

La stessa considerazione può essere fatta analizzando il profilo della Potenza attiva dell'impianto di illuminazione (Figura 26), ma non per l'impianto di condizionamento. Come illustrato infatti in Figura 27, il profilo mostra potenze molto basse e non attendibili vista la tipologia di impianto di condizionamento installato e dati gli orari di funzionamento dello stesso. In tal caso l'anomalia può essere attribuita ad un errato fattore di scala applicato nel processo di conversione del dato misurato. Tale errore influenza di conseguenza anche la definizione del profilo della Potenza attiva di Forza Motrice, essendo questo valore derivato, come precedentemente descritto, dalla potenza attiva totale, di illuminazione e di condizionamento.

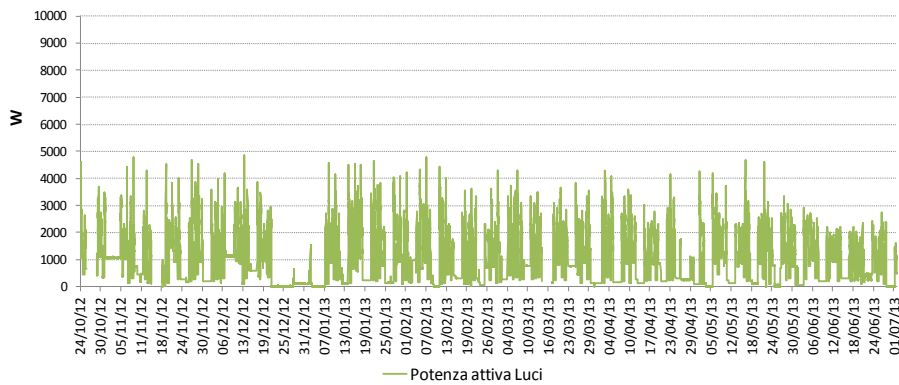


Figura 26: Potenza attiva impianto di illuminazione (24/10/2012 – 01/07/2013)

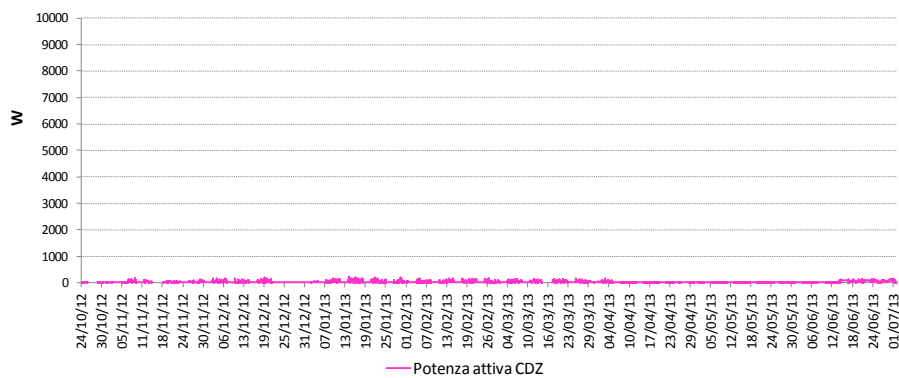


Figura 27: Potenza attiva impianto di condizionamento (24/10/2012 – 01/07/2013)

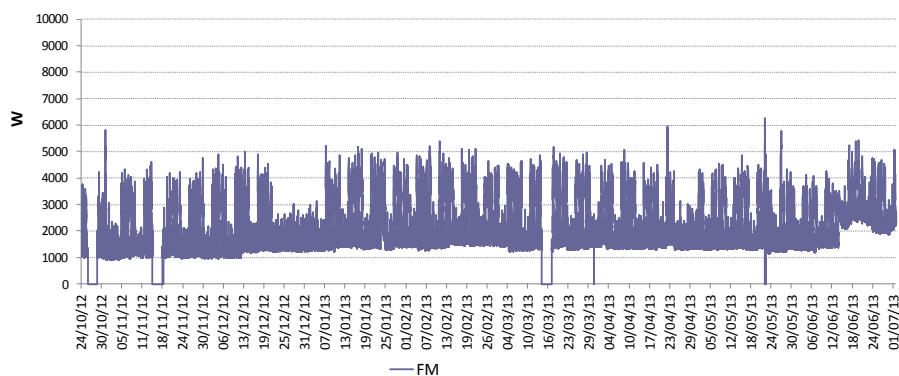


Figura 28: Potenza attiva Forza Motrice (24/10/2012 – 01/07/2013)

Il profilo di temperatura dell'ambiente interno (Figura 29) non presenta particolari anomalie. Anche in questo caso i dati mancanti sono inferiori al 4% e il grafico presenta valori coerenti con quanto previsto dalle normative in materia di comfort in ambienti ad uso ufficio.

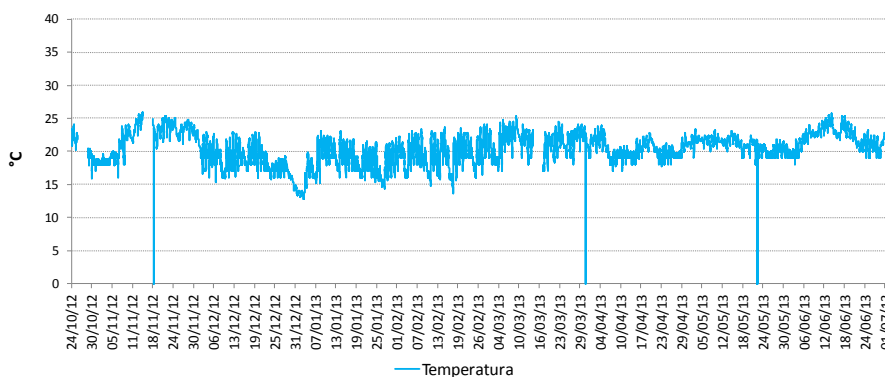


Figura 29: Temperatura degli ambienti all’interno dell’edificio (24/10/2012 – 01/07/2013)

Oltre all’analisi dei profili si è anche valutata la qualità dei dati in termini di numero dei records disponibili attraverso il calcolo dei “missing data”, ossia dei dati non acquisiti dal sistema. Affinchè i campioni collezionati siano considerati accettabili per l’elaborazione, i dati mancanti non devono essere inferiori al 5%.

Di seguito (Tabella 1 e Tabella 2) si riporta il sommario dell’analisi eseguita durante la prima fase di analisi. All’interno di tale sommario si illustrano, per ogni edificio, per la centrale termica e per i parametri climatici, il numero totale di records registrati nel periodo compreso tra il 24-10-2012 e il 01-07-2013. Il sommario evidenzia inoltre le anomalie che si sono riscontrate nel sistema attraverso l’analisi dei profili.

Tabella 1: Qualità acquisizione dati durante il primo intervallo di analisi - Cluster

Denominazione edificio	Periodo monitoraggio		Intervallo temporale di acquisizione	Numero totale record registrati	Missing Data %	Anomalie
	Inizio	Fine				
F66	24/10/20 12 00:00	01/07/2013 23:45	15 minuti	209025 su 216864 totali	3,6 %	<ul style="list-style-type: none"> - Anomalie riscontrate nella misura dell’energia termica e frigorifera. Fattore imputabile ad errore nella misura della portata volumetrica di fluido. - Anomalia riscontrata nella misura della potenza attiva per condizionamento. Possibile fattore di scala.
F67	24/10/20 12 00:00	01/07/2013 23:45	15 minuti	252445 su 265056 totali	4,8 %	<ul style="list-style-type: none"> - Anomalie riscontrate nella misura dell’energia termica e frigorifera. - Anomalia relativa alla potenza attiva illuminazione. Impianto al piano terra sembra funzionante anche di notte e in periodi di chiusura del sito. - Anomalia riscontrata nella misura della potenza attiva per condizionamento. Possibile fattore di scala.
F68	24/10/20 12 15:45	20/06/2013 23:45	15 minuti	198594 su 216864 totali	8,4 %	<ul style="list-style-type: none"> - Anomalie riscontrate nella misura dell’energia termica e frigorifera. - Inizio e fine periodo di analisi non

						<p>corrispondente con il resto del cluster.</p> <ul style="list-style-type: none"> - Anomalia riscontrata nella misura della temperatura interna. La temperatura supera i 30 °C durante il periodo di riscaldamento.
F69	25/10/20 12 00:00	01/07/2013 23:45	15 minuti	208057 su 216864 totali	4,1 %	<ul style="list-style-type: none"> - Anomalie riscontrate nella misura dell'energia termica e frigorifera. - Anomalia riscontrata nella misura della potenza attiva per condizionamento. Possibile fattore di scala.
F70	24/10/20 12 00:00	01/07/2013 23:45	15 minuti	208458 su 216864 totali	3,9 %	<ul style="list-style-type: none"> - Misure di energia termica e frigorifera assenti fino a metà febbraio a causa di un errore di misura della portata di fluido. - Anomalie riscontrate nella misura dell'energia termica e frigorifera.
F71	24/10/20 12 00:00	01/07/2013 23:45	15 minuti	209743 su 216864 totali	3,6 %	<ul style="list-style-type: none"> - Anomalie riscontrate nella misura dell'energia termica e frigorifera. - Misure attendibili di potenza attiva illuminazione e condizionamento dal 06/12/2012. - Anomalia relativa alla potenza attiva illuminazione. Impianto al sembra funzionante anche di notte e in periodi di chiusura del sito.
F72	24/10/20 12 00:00	01/07/2013 23:45	15 minuti	208964 su 216864 totali	3,9 %	<ul style="list-style-type: none"> - Anomalie riscontrate nella misura dell'energia termica e frigorifera. Fattore imputabile ad errore nella misura della portata volumetrica di fluido. - Anomalia relativa alla potenza attiva illuminazione. Impianto al sembra funzionante anche di notte e in periodi di chiusura del sito. - Misure attendibili di potenza attiva illuminazione e condizionamento dal 06/12/2012.
F73	24/10/20 12 00:00	01/07/2013 23:45	15 minuti	209034 su 216864 totali	3,6%	<ul style="list-style-type: none"> - Anomalie riscontrate nella misura dell'energia termica e frigorifera. - Anomalia relativa alla potenza attiva illuminazione. Impianto al sembra funzionante anche di notte e in periodi di chiusura del sito. - Misure attendibili di potenza attiva illuminazione e condizionamento dal 06/12/2012.
Centrale termica	24/10/20 12 00:00	01/07/2013 23:45	15 minuti		3,3%	<ul style="list-style-type: none"> - Fino al 18/02/2013 i valori indicati come energia frigorifera si riferiscono ad energia termica.

Tabella 2: Qualità acquisizione dati durante il primo intervallo di analisi – Clima esterno

Parametro monitorato	Periodo monitoraggio		Intervallo temporale di acquisizione	Numero totale di records registrati	Missing Data %	Anomalie
	Inizio	Fine				
Temperatura dell'aria esterna	24/10/2012 00:00	01/07/2013 23:45	15 minuti	18718 su 24096 totali	22%	Presenza di periodi di discontinuità.
Umidità relativa aria esterna	24/10/2012 00:00	01/07/2013 23:45	15 minuti	18718 su 24096 totali	22%	
Velocità del vento	24/10/2012 15:45	20/06/2013 23:45	15 minuti	18718 su 24096 totali	22%	
Direzione del vento	25/10/2012 00:00	01/07/2013 23:45	15 minuti	18718 su 24096 totali	22%	
Pressione	25/10/2012 00:00	01/07/2013 23:45	15 minuti	18718 su 24096 totali	22%	
mm di pioggia	25/10/2012 00:00	01/07/2013 23:45	15 minuti	18718 su 24096 totali	22%	
Radiazione solare	24/10/2012 00:00	01/07/2013 23:45	15 minuti	18718 su 24096 totali	22%	

In Figura 30 sono rappresentati i record registrati e i missing data di tutti i parametri misurati in rapporto al tempo. Dal grafico emerge che, al di là dei dati meteo, che presentano discontinuità durante tutto il periodo di misura, tutti gli altri dati presentano brevi interruzioni in periodi pressoché identici, il che porta a pensare che vi sia stato un problema di comunicazione generale del sistema di acquisizione dati. Si evidenzia inoltre il fatto che i dati mancanti sono quasi sempre al di sotto della valore massimo del 5% e che, se non fosse per altre anomalie rilevate dall'analisi dei profili, il risultato ottenuto potrebbe definirsi soddisfacente ai fini delle elaborazioni successive.

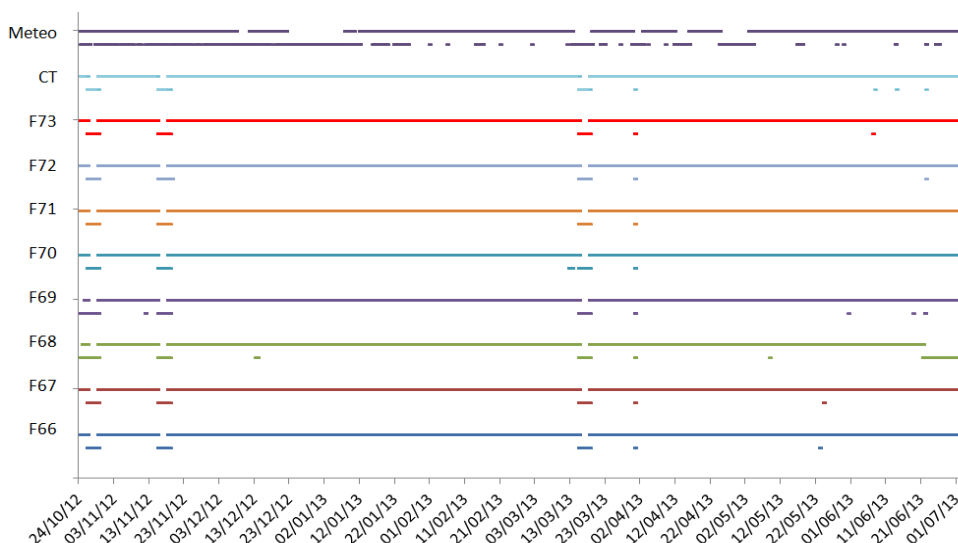


Figura 30: Evidenza dei missing data totali presenti nei dati di monitoraggio degli otto edifici componenti il cluster e della centrale termica.

Analisi qualità dei dati n° 2.

Come descritto nello schema di Figura 21, l'analisi della qualità dei dati deve essere ripetuta periodicamente allo scopo di ottimizzare il sistema di monitoraggio e acquisizione dati installato, riducendo conseguentemente gli errori di misura che si presentano nella fase di start up del sistema stesso.

Il periodo temporale di acquisizione dati trattato nella prima fase di analisi comprendeva circa 9 mesi di misura. Tal intervallo è servito sia per valutare eventuali errori ciclici del sistema nel tempo, sia per avere una visione completa del sistema di monitoraggio installato e del suo funzionamento in periodo invernale e in periodo estivo.

Gli errori e le anomalie emerse dalla prima analisi sono stati oggetto di studio e correzione, ed i risultati di tale operazione sono visibili nella seconda analisi di seguito descritta. Questa volta l'intervallo di tempo selezionato è stato inferiore, ma sufficiente per verificare l'effettiva messa in opera delle operazioni di ottimizzazione del sistema.

La Tabella 3, a differenza di quanto era descritto in Tabella 1, illustra come la percentuale di dati mancanti nel secondo periodo di analisi sia pari a zero, il che evidenzia come il sistema di acquisizione sia stato migliorato sulla base delle segnalazioni evidenziate durante la prima analisi. Lo stesso risultato positivo è illustrato in Figura 31. Ciò non toglie che le anomalie non siano ancora state del tutto eliminate e che in alcuni casi, ne sono emerse di nuove.

Tabella 3: Qualità acquisizione dati durante il secondo intervallo di analisi

Denominazione edificio	Periodo monitoraggio		Intervallo temporale di acquisizione	Numero totale record registrati	Missing Data %	Anomalie
	Inizio	Fine				
F66	29/07/2013 01:00	23/07/2013 23:45	15 minuti	2492 su 2492 totali	0,0 %	<ul style="list-style-type: none"> - Anomalia relativa alla misura della portata volumetrica di fluido (portata presente anche di notte intorno a 8 mc/h). - Anomalia relativa alla potenza attiva illuminazione. Impianto al sembra funzionante anche di notte e in periodi di chiusura del sito. - Anomalia relativa alla misura della potenza attiva per condizionamento: l'impianto sembra sempre in funzione anche durante gli orari notturni.
F67	29/07/2013 01:00	23/07/2013 23:45	15 minuti	2492 su 2492 totali	0,0 %	<ul style="list-style-type: none"> - Anomalia relativa alla misura della portata volumetrica di fluido (portata presente anche di notte intorno a 11 mc/h). - Anomalia relativa alla potenza attiva illuminazione. Impianto al sembra funzionante anche di notte e in periodi di chiusura del sito. - Anomalia relativa alla misura della

						potenza attiva per condizionamento: l'impianto sembra sempre in funzione anche durante gli orari notturni.
F68	29/07/20 13 01:00	23/07/2013 23:45	15 minuti	2492 su 2492 totali	0,0 %	<ul style="list-style-type: none"> - Anomalia relativa alla misura della portata volumetrica di fluido (portata presente anche di notte intorno a 9 mc/h). - Anomalia relativa alla potenza attiva illuminazione. Impianto al sembra funzionante anche di notte e in periodi di chiusura del sito. - Anomalia relativa alla misura della potenza attiva per condizionamento: l'impianto sembra sempre in funzione anche durante gli orari notturni.
F69	29/07/20 13 01:00	23/07/2013 23:45	15 minuti	2492 su 2492 totali	0,0 %	<ul style="list-style-type: none"> - Anomalia relativa alla misura della portata volumetrica di fluido (portata presente anche di notte intorno a 7 mc/h). - Anomalia relativa alla misura della potenza attiva per condizionamento: l'impianto sembra sempre in funzione anche durante gli orari notturni.
F70	29/07/20 13 01:00	23/07/2013 23:45	15 minuti	2492 su 2492 totali	0,0 %	<ul style="list-style-type: none"> - Anomalia relativa all'energia termica e frigorifera. Risulta presente energia termica in periodo estivo. Dato trascurabile perché determinato dalle differenze di temperatura che di notte si invertono rispetto al giorno: l'errore dipende dal fatto che sembra che la portata di fluido sia costante a 6 mc/h anche di notte. - Anomalia relativa alla misura della potenza attiva per condizionamento: l'impianto sembra sempre in funzione anche durante gli orari notturni.
F71	29/07/20 13 01:00	23/07/2013 23:45	15 minuti	2492 su 2492 totali	0,0 %	<ul style="list-style-type: none"> - Anomalia relativa alla misura della portata volumetrica di fluido. - Anomalia relativa alla misura della potenza attiva per condizionamento: l'impianto sembra sempre in funzione anche durante gli orari notturni.
F72	29/07/20 13 01:00	23/07/2013 23:45	15 minuti	2492 su 2492 totali	0,0 %	<ul style="list-style-type: none"> - Anomalia relativa alla misura della portata volumetrica di fluido. - Anomalia relativa alla misura della potenza attiva per condizionamento: l'impianto sembra sempre in funzione anche

						durante gli orari notturni.
F73	29/07/20 13 01:00	23/07/2013 23:45	15 minuti	2492 su 2492 totali	0,0 %	<ul style="list-style-type: none"> - Anomalia relativa alla potenza attiva illuminazione. Impianto al sembra funzionante anche di notte e in periodi di chiusura del sito. - Anomalia relativa alla misura della potenza attiva per condizionamento: l'impianto sembra sempre in funzione anche durante gli orari notturni.
Centrale termica	29/07/20 13 01:00	23/07/2013 23:45	15 minuti	2492 su 2492 totali	0,0 %	---

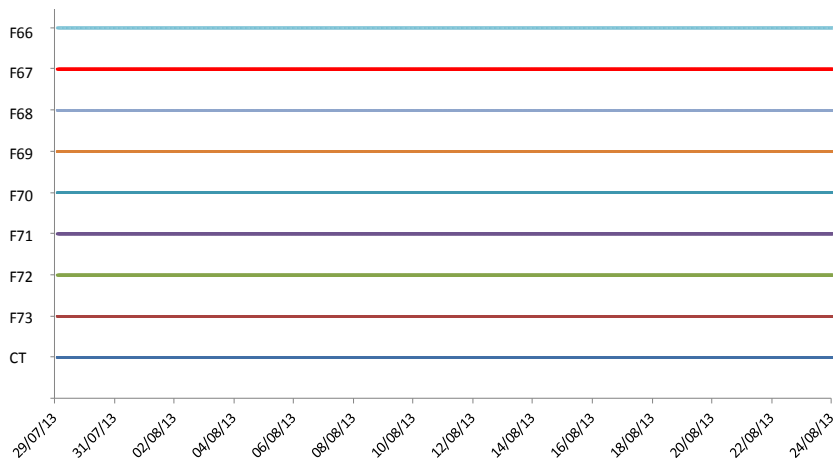


Figura 31: Evidenza dei missing data totali presenti nei dati di monitoraggio degli otto edifici componenti il cluster e della centrale termica.

A dimostrazione delle operazioni di ottimizzazione del sistema di monitoraggio, si riportano di seguito i profili di energia e potenza relativi alla seconda fase di analisi e su cui si erano riscontrate anomalie durante la prima fase.

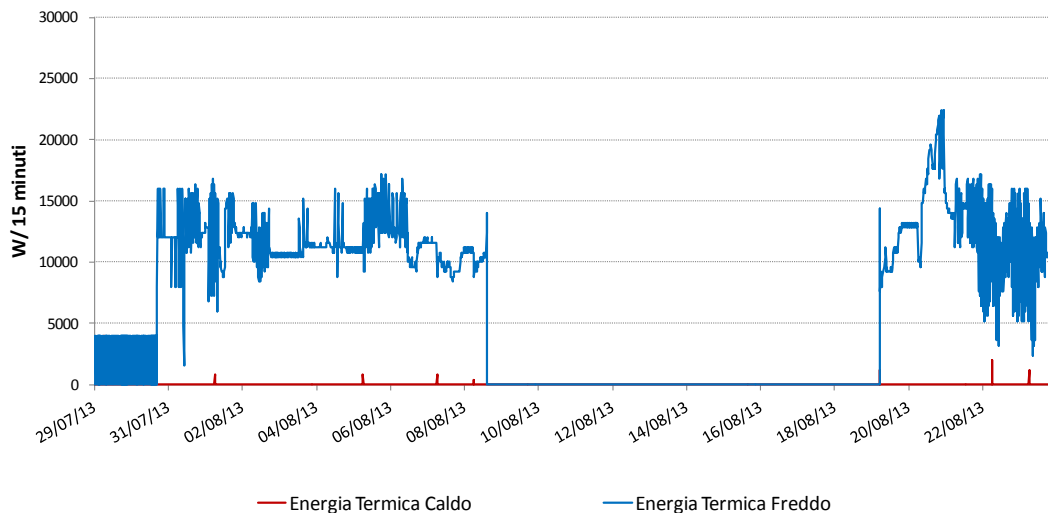


Figura 32: Energia termica ed energia frigorifera (29/10/2012 – 22/08/2013)

La Figura 32 dimostra come le anomalie di funzionamento riscontrate nella fase precedente siano state corrette, e da inizio agosto 2013 i dati si presentino completi. Restano però delle anomalie su cui è opportuno indagare: com'è dimostrato dal grafico in

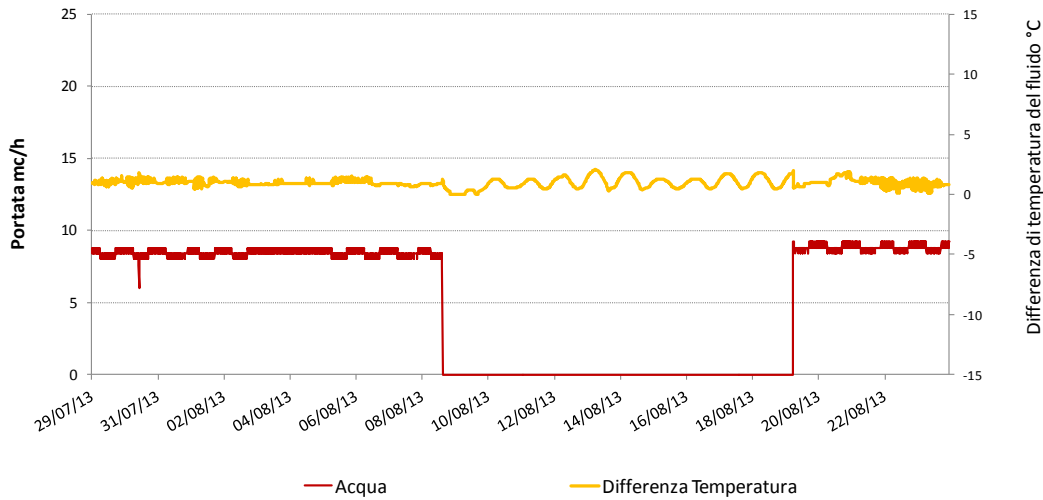
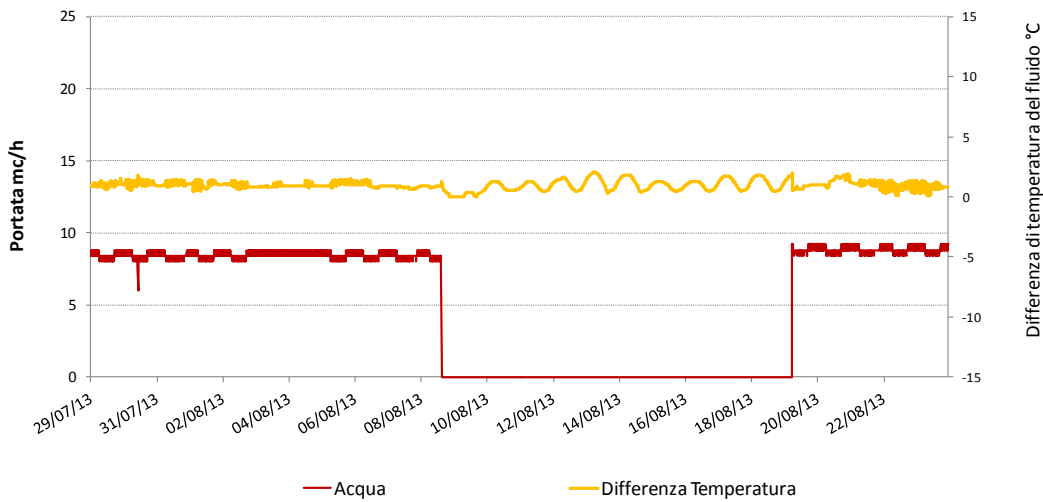


Figura 33, la

portata volumetrica di fluido sembra avere un profilo costante, variando tra la notte e il giorno di circa 1 m³/h, ed essendo sempre presente tranne che nel periodo di inattività dell'edificio dal 08/08 al 18/08. In tal modo, anche l'energia frigorifera misurata risulta sempre positiva e con valori superiori ai 40 kWh anche in orari notturni. Per tale motivo occorre ancora effettuare delle indagini per verificare che tale risultato sia da imputare ai sistemi di misura o ai sistemi di controllo dell'impianto in oggetto.



Figura

33: Portata d'acqua presente nel circuito idronico a servizio dell'edificio e differenza di temperatura tra l'acqua di ritorno dall'impianto del singolo edificio e l'acqua di mandata in uscita dalla Centrale Termica (29/10/2012 – 22/08/2013)

Non sempre dopo le correzioni effettuate al sistema di monitoraggio successivamente alla prima fase di analisi si riescono a risolvere tutte le anomalie, e soprattutto, non è detto che dalla seconda fase non ne emergano di nuove. Il profilo della potenza attiva dell'impianto di illuminazione dell'edificio F66 rappresentato in Figura 26, ad esempio, sembrava evidenziare una situazione di corretto funzionamento del sistema durante la prima analisi. Andando ora a verificare ciò che accade allo stesso impianto di illuminazione durante il secondo periodo di analisi (Figura 34), si nota che vi sono alcune irregolarità: l'impianto sembra avere una potenza attiva costante superiore ai 200 W anche durante i periodi notturni e nei giorni di chiusura degli uffici. In un caso come questo bisogna ancora indagare se tali anomalie

dipendono da una cattiva gestione dell'impianto di illuminazione o se si tratta di errori dovuti al sistema di acquisizione dati.

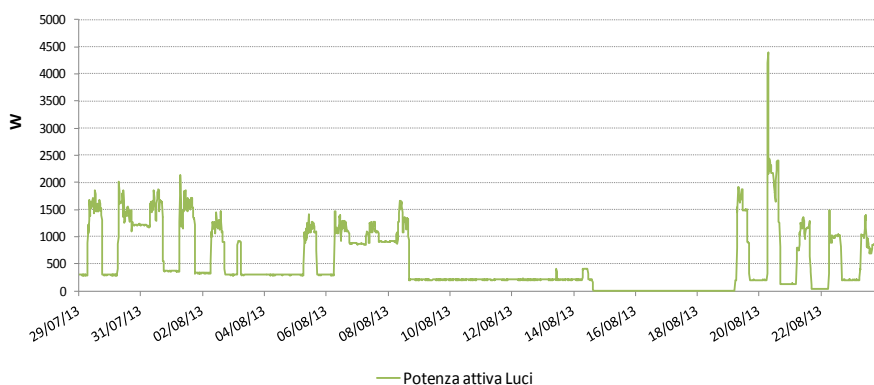


Figura 34: Potenza attiva impianto di illuminazione (29/07/2013 – 22/08/2013)

Per quanto riguarda dati di potenza attiva e reattiva di tutto l'edificio (Figura 35 e Figura 36) i profili evidenziano che il sistema ha continuato a funzionare bene. Data la leggibilità, nel profilo di potenza attiva totale, di alcuni andamenti imputabili alla potenza dell'impianto di illuminazione, questa analisi porta a pensare che le anomalie evidenziate in Figura 34 siano da attribuire ad uno scorretto utilizzo dell'impianto di illuminazione o, come seconda ipotesi, ad una scorretta suddivisione delle linee elettriche a livello di quadro.

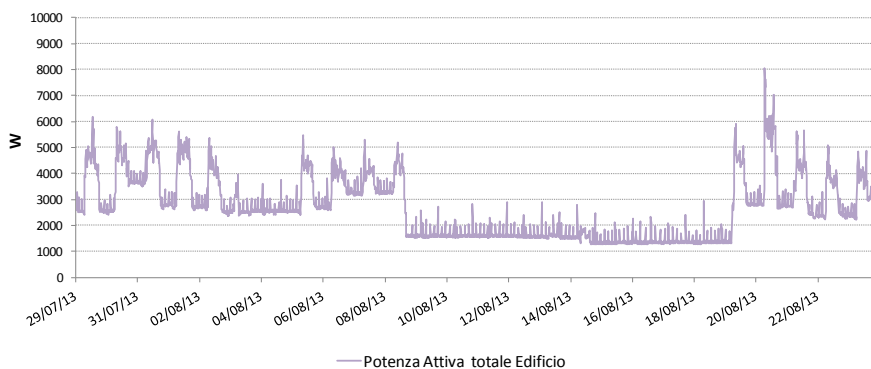


Figura 35: Potenza attiva totale dell'edificio (29/07/2013 – 22/08/2013)

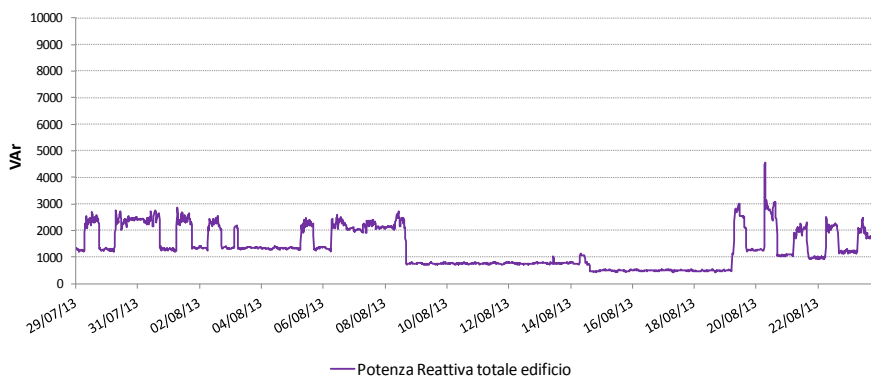


Figura 36: Potenza reattiva totale dell'edificio (29/07/2013 – 01/07/2013)

Per meglio evidenziare quanto appena detto si riporta l'esempio dell'edificio F67. In questo caso l'impianto di illuminazione del piano terra, come evidente dalla Figura 37, sembrerebbe funzionante anche durante gli orari notturni e durante i periodi di chiusura dell'edificio.

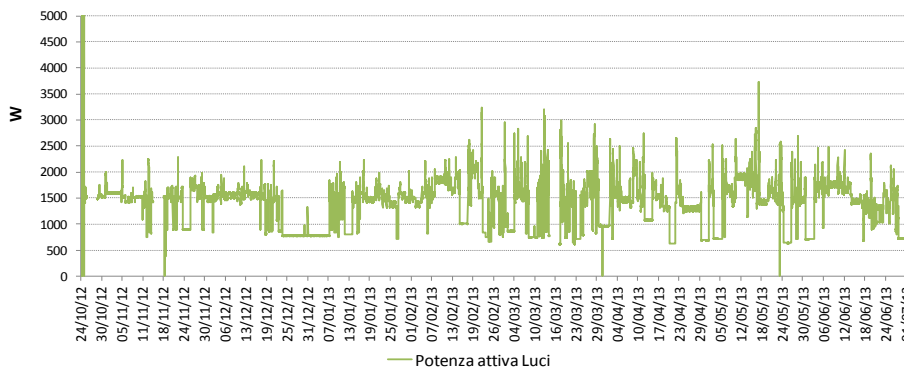


Figura 37: Potenza attiva impianto di illuminazione (29/10/2012 – 01/07/2013)

Facendo un focus su un periodo di tre settimane, e confrontando entrambe le fasi di analisi (Figura 38 e Figura 39), si evidenzia ancor più quanto sopra descritto e si denota come anche dopo una prima fase di revisione del sistema, permangano anomalie. L'impianto sembra infatti essere sempre in funzione a potenza pressoché costante durante la settimana (esempio dal 29/07 al 03/08), e durante orari di chiusura degli uffici (esempio dal 12/08 al 15/08).

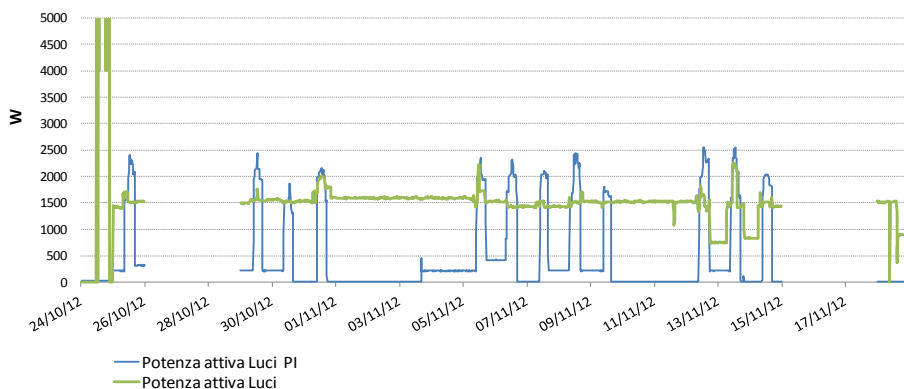


Figura 38: Potenza attiva impianto di illuminazione (24/10/2012 – 19/11/2012)

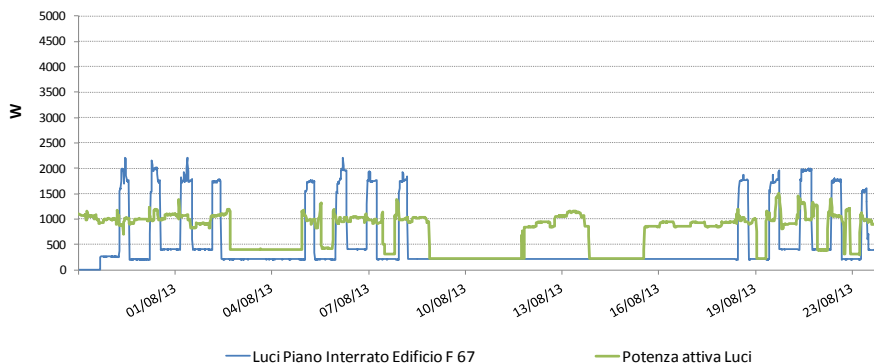


Figura 39: Potenza attiva impianto di illuminazione (29/07/2013 – 24/08/2013)

Anche in questo caso, come per l'edificio F66, dovranno essere fatte valutazioni basandosi anche su altri dati di monitoraggio, oltre a quelli della sola Potenza attiva dell'impianto di illuminazione. Successivamente verranno fatte verifiche sulla gestione degli impianti e se non basterà si passerà alla verifica della

ripartizione dell'energia nei quadri elettrici. Solo in seguito a queste operazioni si potrà poi procedere con un'ulteriore fase di analisi della qualità dei dati.

Continuando l'analisi dell'edificio F66, e analizzando questa volta la potenza attiva dell'impianto di condizionamento, si evidenzia che le anomalie emerse dalla prima fase di analisi sono state corrette ed ora i dati raccolti sono coerenti con le aspettative. Resta però da comprendere il motivo per cui, durante la settimana, l'impianto sia in funzione anche durante gli orari notturni.

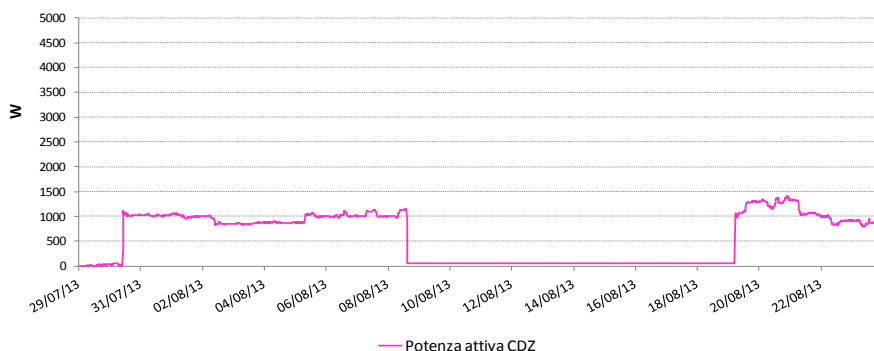


Figura 40: Potenza attiva impianto di condizionamento (29/07/2013 – 22/08/2013)

Come per la prima fase, anche in questo caso la potenza assorbita dalla Forza Motrice è stata derivata ed è quindi dipendente da altri parametri monitorati. Considerando attendibili i valori misurati di potenza totale, di illuminazione e di condizionamento, sembrerebbe che la forza motrice richieda una potenza costante superiore ai 1500 W, anche durante i periodi di inattività dell'edificio. Queste assunzioni, insieme ad altre emerse durante l'analisi di tutti gli edifici, costituiranno la base di partenza per far partire il secondo processo di ottimizzazione del sistema.

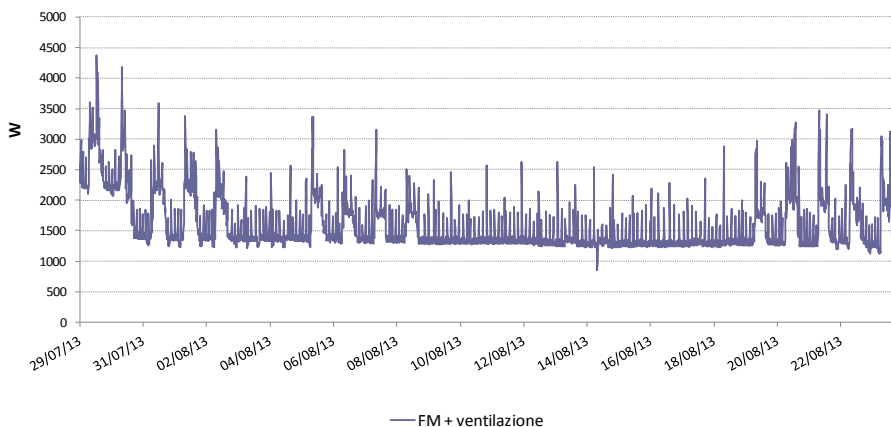


Figura 41: Potenza attiva Forza Motrice (29/07/2013 – 22/08/2013)

Malgrado la presenza di alcune anomalie e dati missing data, nel complesso le verifiche effettuate hanno confermato l'attendibilità dei dati al fine della loro successiva elaborazione (Attività 3 e Attività 4).

Attività 3

Elaborazione e analisi dati energetici e microclimatici del monitoraggio in campo ai fini diagnostici di “primo livello”: profili di consumo e qualità climatica indoor, tendenze e benchmarks

L’elaborazione e l’analisi dei dati raccolti per mezzo del sistema di monitoraggio installato sul cluster di 8 edifici oggetto d’esame si è sviluppata principalmente secondo due direttrici: la prima è l’analisi dei profili di consumo, la seconda la valutazione e il confronto di indicatori di prestazione energetica attraverso lo sviluppo di un modello di regressione multipla a due variabili indipendenti atto a identificare un “piano medio” di prestazione.

Propedeutico allo svolgimento delle attività relative all’analisi dei profili e allo sviluppo, applicazione e validazione di metodologie di benchmarking è stata l’identificazione di possibili ed eventuali tendenze o eventi ricorrenti nelle dinamiche dei consumi attraverso il confronto trasversale tra edifici degli andamenti dei consumi e dei consumi specifici di energia elettrica e termica mensili e giornalieri [1].

I **dati a disposizione** si riferiscono al gruppo di 8 edifici facenti parte del complesso di ENEA Casaccia. Gli 8 edifici risultano sottoposti al monitoraggio dei consumi di **energia elettrica totale, per l’illuminazione e per il condizionamento degli ambienti**. Viene altresì misurata **l’energia elettrica totale entrante in cabina elettrica** al servizio degli 8 edifici monitorati e di un ulteriore edificio escluso dal piano di monitoraggio. Tra i dati di natura elettrica raccolti si annoverano anche i **consumi totali della centrale termica** ospitante al suo interno il gruppo frigo e lo scambiatore di calore per la produzione del fluido freddo/caldo necessario all’alimentazione del circuito idronico del cluster. La presenza in centrale termica di un sottocontatore permette di scorporare dai **consumi elettrici** della centrale quelli attribuiti al **gruppo frigo**.

I **dati di natura termica** si riferiscono all’energia termica/frigorifera in uscita dalla centrale termica e consegnata al cluster di edifici oltre che l’energia termica/frigorifera in ingresso a ciascuno di essi.

L’esame dei consumi elettrici degli edifici appartenenti al cluster permette di evidenziare che in termini assoluti l’edificio più energivoro è l’edificio F67 che è anche il più grande e l’unico ad essere composto da un piano fuori terra ed un piano interrato. L’edificio F69 presenta invece i consumi più bassi rispetto al gruppo considerato (vedi **Figura 42**).

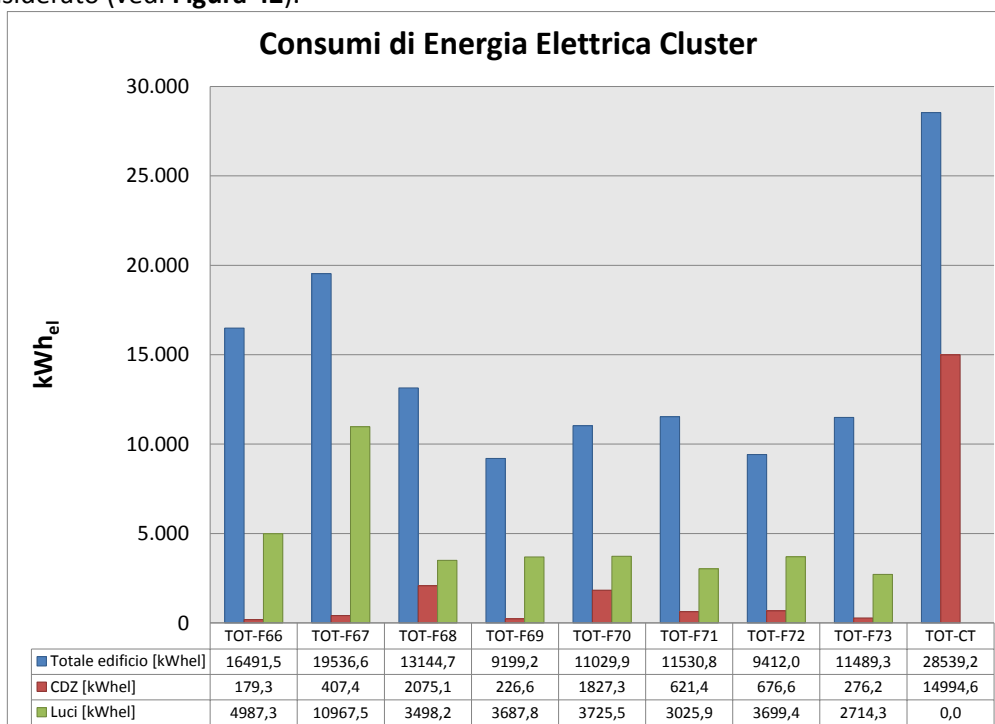


Figura 42: Consumi di Energia Elettrica del cluster nel periodo novembre 2012-giugno 2013

Il confronto degli edifici in termini di consumi specifici rivela che è l'edificio F66 ha presentare consumi più alti con valori di circa il 20% superiore alla media. L'andamento mensile dei consumi e dei consumi specifici di energia elettrica del cluster conferma il risultato appena esposto (vedi Figura 43 e Figura 44).

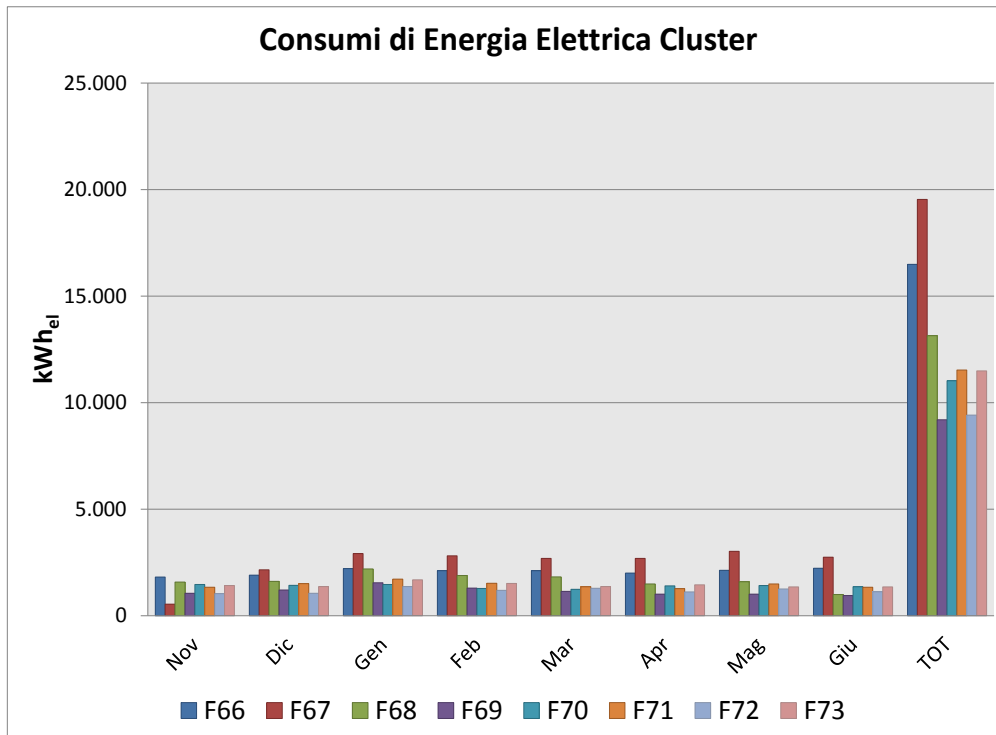


Figura 43: Andamento mensile dei consumi di Energia Elettrica del cluster

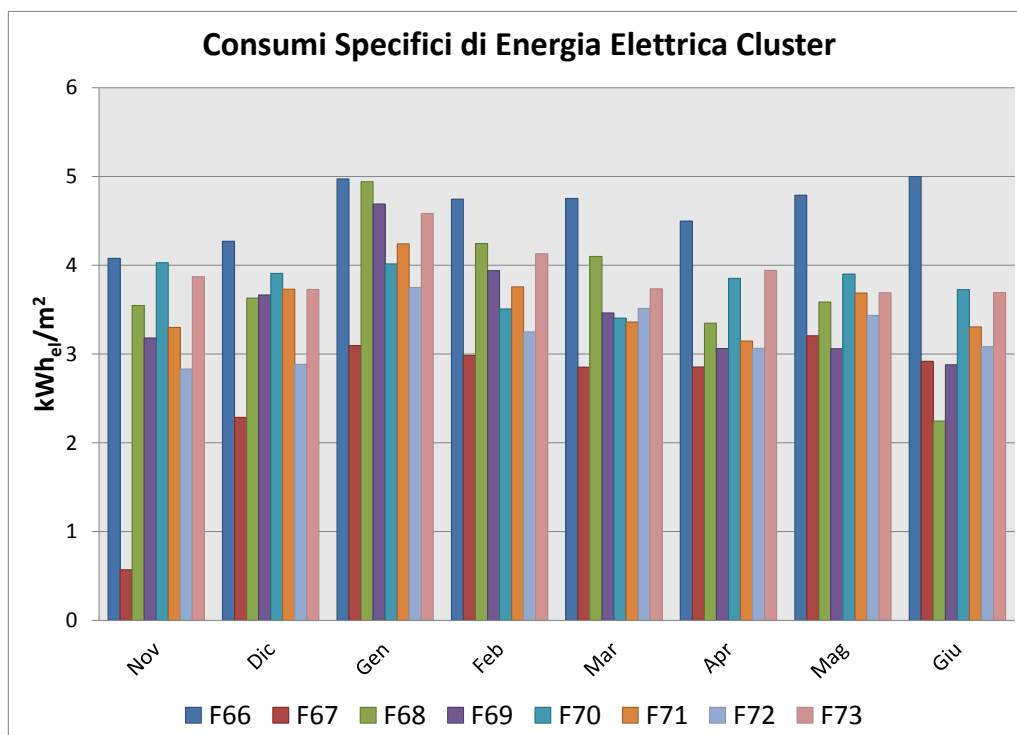


Figura 44: Andamento mensile dei consumi specifici di Energia Elettrica del cluster

Le analisi degli andamenti mensili dei consumi di Energia Elettrica suddivisi per usi finali (condizionamento e illuminazione) condotte per ciascuno degli edifici appartenenti al cluster appaiono in buon accordo per

quanto attiene all'illuminazione mentre suggeriscono ulteriori approfondimenti in relazione al condizionamento come da grafici riportati in Appendice 1 in Tabella 45.

Con riferimenti agli andamenti dei consumi di Energia Termica le analisi evidenziano e confermano valori ampiamente variabili da edificio a edificio la cui causa è attribuibile alla scarsa attendibilità dei dati misurati (vedi Appendice 1 in Tabella 46)

L'esame del consumo specifico medio giornaliero calcolato mese per mese per tutto il periodo in esame e per ciascuno degli edifici consente di determinare l'intervallo entro cui tali consumi variano. I valori calcolati nei mesi in cui i consumi totali registrano valori più elevati (gennaio e febbraio 2013) risultano essere compresi tra 0,169 kWh/m²/giorno-mese e 0,1 kWh/m²/giorno-mese. Il valore massimo si presenta a febbraio 2013 per l'edificio F66 mentre il valore minimo si ha per l'edificio F67 a gennaio 2013. I risultati descritti sono riportati in Appendice 1 Tabella 47.

Successivamente lo studio degli andamenti giornalieri dei consumi suddivisi per usi finali (condizionamento e illuminazione) nei mesi di gennaio e giugno 2013 per ciascun edificio ha permesso di verificare la correttezza delle modalità di consumo settimanale con consumi maggiori e pressoché regolari nel corso dei giorni feriali e in quelli festivi (Appendice 1 Tabella 48).

Conclusa l'analisi preliminare delle tendenze dei consumi, l'attività si è concentrata sullo studio di approcci semplificati per l'identificazione dei cambiamenti di comportamento degli edifici attraverso la lettura dei profili di consumo energetico. Rispetto agli obiettivi del progetto non è stato possibile condurre analisi della qualità climatica indoor perché, diversamente da quanto previsto nel piano di monitoraggio ambientale proposto, secondo cui sarebbero state posizionate 2 sonde per la misurazione della temperatura e umidità relativa in due uffici di uguale superficie e diversa esposizione, nel sistema di monitoraggio installato si è deciso di installare un solo sensore di temperatura nei corridoi di ciascun edificio come anche riportato nella relazione dell'attività relativa alla verifica funzionale in fase operativa dei sistemi di monitoraggio in campo operanti su "smartdistrict" e analisi della qualità dei dati raccolti.

Nel tentativo di caratterizzare la "forma del profilo", sono stati definiti e calcolati sui profili orari medi su base stagionale (inverno e primavera 2013) un insieme di parametri statistici che consentono una schematizzazione della forma in tratti associati ai momenti della giornata [2].

Tipicamente il profilo elettrico medio orario per la stagione estiva di un edificio a destinazione d'uso ufficio presenta una forma schematizzabile composta da una successione di linee. Semplificando al massimo la descrizione della forma, è possibile distinguere un primo tratto della curva che corrisponde agli assorbimenti elettrici misurati durante la notte quando l'edificio è in uno stato che può definirsi "spento". Durante le prime ore del mattino quando si attivano i sistemi di climatizzazione e riprendono le attività all'interno dell'edificio, i consumi cominciano a crescere (secondo tratto), anche per effetto dell'aumento della temperatura esterna, fino a stabilizzarsi durante le ore centrali della giornata (terzo tratto). Nel tardo pomeriggio, dal momento in cui l'edificio comincia a svuotarsi, gli assorbimenti diminuiscono fino a riportarsi ai valori di consumo dell'edificio inattivo (quarto tratto).

In alcuni casi è anche possibile distinguere sulla curva del profilo, nel corso delle prime ore del mattino, un picco in corrispondenza del passaggio di setpoint di temperatura tra notte e giorno. Quando al mattino la temperatura all'interno dell'edificio è minore di quella richiesta dal setpoint, gli impianti di climatizzazione si attivano improvvisamente facendo registrare un picco la cui durata dipende dalla temperatura a cui si trova l'edificio quando il fenomeno si manifesta.

Per rendere oggettivo l'esame visivo dei profili, sono stati definiti sulla curva una serie di parametri statistici che consentono di stabilire "i tempi di risposta" dell'edificio alle condizioni esterne imposte dagli utilizzatori e dal clima. La Tabella 4 riassume tali parametri rappresentati in Figura 45.

Tabella 4: Parametri caratterizzanti la forma del profilo di carico

NearPeakLoad	Picco di carico pari al 97,5esimo percentile del carico giornaliero
Near Base Load	Base di carico, pari al 2,5esimo percentile del carico giornaliero
NBL+0,1*(NPL-NBL)	Inizio del periodo cosiddetto di "carico alto"
NPL-0,1*(NPL-NBL)	Fine del periodo cosiddetto di "carico alto"
(NPL+NBL)/2	50esimo percentile del carico giornaliero

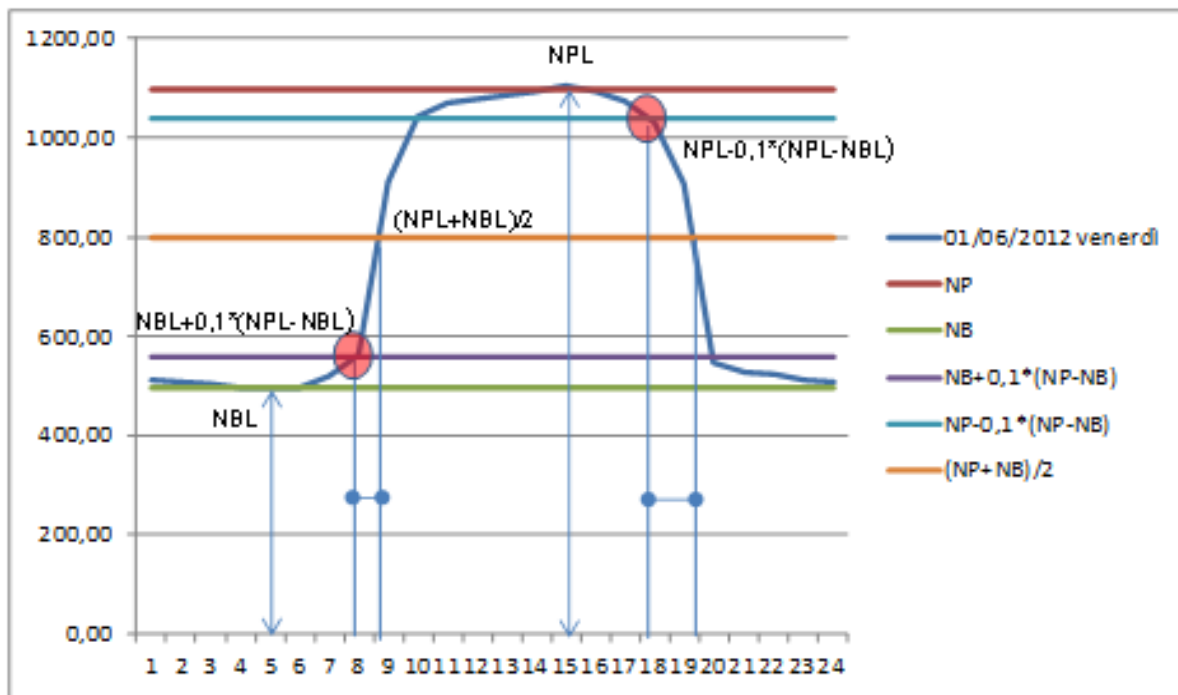


Figura 45: Rappresentazione dei parametri caratterizzanti il profilo di carico

In Appendice 1 Tabella 49 e Tabella 52 si riportano i *Box-Plots* dei profili orari medi per l’inverno e la primavera 2013 rispettivamente per ciascuno degli edifici appartenenti al cluster. Inoltre per ognuno dei profili sono calcolati i parametri statistici, definiti sopra, per stagione.

Con riferimento alla stagione invernale, il picco di carico presenta un coefficiente di variazione compreso tra 0,37 e 0,57 mentre la base di carico ha valori compresi tra 0,12 e 0,36. Tali intervalli permettono di valutare la dispersione dei parametri caratterizzanti i profili e indicano una sufficiente omogeneità dei profili.

Per la stagione estiva il coefficiente di variazione è compreso all’interno degli intervalli 0,38 e 0,65 per il picco di carico e 0,16 e 0,41 per la base di carico.

Le Tabelle contengono inoltre i *Carpet Plots* che facilitano la visualizzazione di lunghe serie temporali di una singola variabile (nel presente studio la variabile è rappresentata dal consumo totale di energia elettrica) per mezzo di mappe di colori che rivelano ‘patterns’ tipici che si ripetono con una qualche ciclicità (ora della giornata, giorno della settimana, settimana del mese e mese dell’anno),

Infine la Tabella 50 e Tabella 53 in Appendice 1 riportano i risultati delle elaborazioni condotte per la stagione invernale su ciascuno degli 8 edifici del cluster e rappresentano i profili orari di carico ottenuti calcolando ora per ora la media dei valori corrispondenti ai giorni feriali, ai giorni di sabato ed alle domeniche comprese nell’intero periodo di analisi. Analogo studio è stato eseguito per la stagione primaverile e ciò che emerge dalle elaborazioni è che la base di carico varia da 2 Wh/m² a 4 Wh/m² per la stagione invernale e da 3 Wh/m² a 4 Wh/m² per quella primaverile mentre il picco di carico presenta valori tra 9 Wh/m² e 14 Wh/m² nel periodo invernale e tra 7,5 Wh/m² e 9 Wh/m² in quello primaverile. Quest’ultimo risultato conferma quanto emerso nelle precedenti analisi dei consumi mensili di energia elettrica secondo cui i valori maggiori si registrano proprio nei mesi di gennaio e febbraio.

I profili di carico ottenuti calcolando il consumo medio di energia elettrica per ogni giorno e per ogni mese ad ogni ora sono riportati in Tabella 51 per la stagione invernale e in Tabella 54 per la stagione primaverile avvalorano le precedenti considerazioni di consumi più alti a febbraio e maggio.

L’attività relativa alla valutazione e al confronto di indicatori di prestazione energetica è avvenuta attraverso lo sviluppo di un modello di regressione multipla a due variabili indipendenti atto a identificare un “piano medio” di prestazione [3].

L'analisi di benchmarking è stata sviluppata considerando i dati elettrici e costruendo un set di dati contenente l'energia elettrica assorbita in relazione alla superficie per ciascun edificio come variabile indipendente e le superfici ed il numero di presenze rilevate nel periodo di analisi considerato come variabili dipendenti.

L'equazione del modello assume la forma:

$$\min_{\beta_0, \beta_1, \beta_2, \varepsilon_i} \left\{ \sum_{i=1}^8 \varepsilon_i^2 \mid E_i = \beta_0 + \beta_1(m^2)_i + \beta_2 \left(\frac{\text{presenze}}{\text{anno}} \right)_i + \varepsilon_i \quad \forall i = 1 \div 8 \right\}$$

dove i coefficienti $\beta_0, \beta_1, \beta_2$ rappresentano l'intercetta ed i coefficienti di regressione; 8 è il numero di osservazioni campionarie, cioè il numero di edifici che compongono il campione oggetto di interesse; $E_i, (m^2)_i, \left(\frac{\text{presenze}}{\text{anno}} \right)_i$ si riferiscono alla i -esima osservazione e ε_i è la componente d'errore della i -esima osservazione.

La soluzione del sistema di equazioni attraverso il metodo dei minimi quadrati consente di calcolare le incognite $\beta_0, \beta_1, \beta_2$ che definiscono il piano di benchmark.

La Tabella 5 riporta i dati utilizzati per lo sviluppo del modello di regressione lineare multipla e l'errore di approssimazione.

Tabella 5: Dati di input del modello di regressione lineare multipla

	kWh _(nov-giu 2013) /m ²	m ²	presenze _(nov-giu)	kWh ^{stimato} _(nov-giu 2013) /m ²	residuo
F73	31,38	366,15	6649,22	29,86	1,52
F72	25,81	364,60	9885,00	30,31	-4,49
F71	28,53	404,16	12588,00	30,13	-1,60
F70	30,35	363,44	9023,75	30,21	0,14
F69	27,95	329,17	7193,25	30,43	-2,48
F68	29,65	443,40	11765,00	29,49	0,15
F67	20,77	940,50	5067,50	21,94	-1,16
F66	37,11	444,42	9477,25	29,18	7,93

Per gli edifici F72, F71, F69 ed F67 i valori misurati risultano minori di quelli previsti (residuo negativo) ciò denota un comportamento energetico più efficiente rispetto quello atteso.

Nel prosieguo dell'attività lo studio della distribuzione dei residui ottenuti dal modello di regressione potrà consentire la costruzione di una tabella di benchmark.

I risultati dell'analisi di regressione sono sintetizzati nella Tabella 6 e Tabella 7.

Tabella 6: Statistica della regressione

R multiplo	0,61
R al quadrato	0,38
R al quadrato corretto	0,13
Errore standard	4,37
Osservazioni	8,00

La Tabella 6 riporta il coefficiente di correlazione r tra la variabile dipendente e le variabili indipendenti, il coefficiente di determinazione R^2 che presenta un valore basso probabilmente a causa di una elevata variabilità dei dati, la stima della deviazione standard degli errori casuali ed il numero di osservazioni.

La Tabella 7 presenta le stime dei coefficienti $\beta_0, \beta_1, \beta_2$, le stime delle rispettive deviazioni standard, i p-value ed i rispettivi intervalli di fiducia al 95%.

La Figura 46 permette di visualizzare il posizionamento degli edifici rispetto al piano di regressione identificato.

Tabella 7: Coefficienti e valori di significatività

	<i>Coefficienti</i>	<i>Errore standard</i>	<i>Valore di significatività</i>	<i>Inferiore 95%</i>	<i>Superiore 95%</i>
Intercetta	33,91	9,73	0,02	8,90	58,91
m²	-0,01	0,01	0,22	-0,04	0,01
presenze(nov-giu 2013)	0,00	0,00	0,87	0,00	0,00

Rappresentazione tridimensionale del piano di regressione

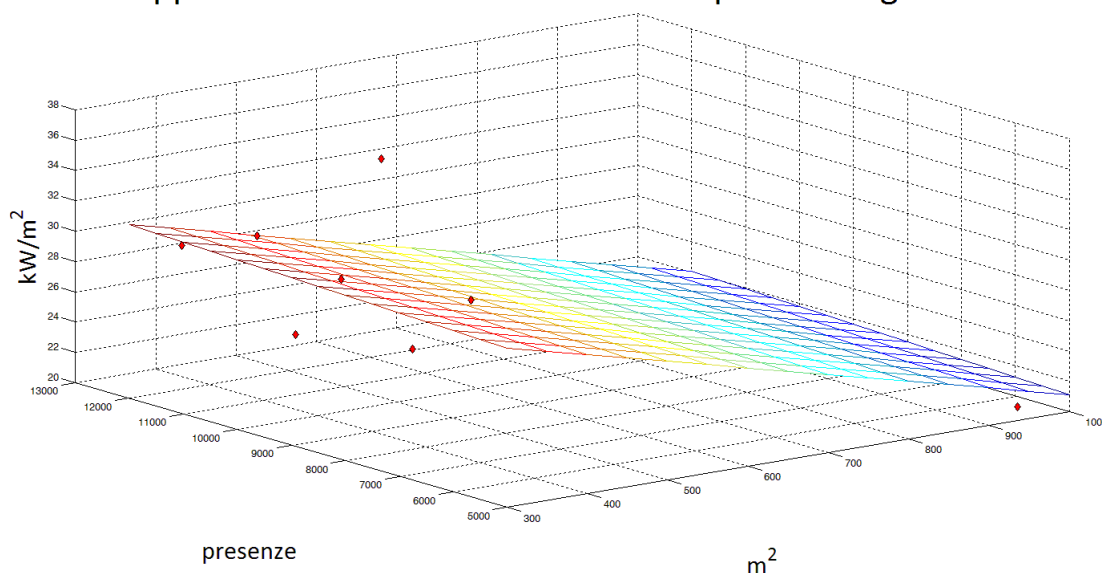


Figura 46: Rappresentazione tridimensionale del piano di regressione

Attività 4

Sviluppo di modelli di diagnostica avanzata (“secondo livello”) attraverso i dati del monitoraggio in campo

La ricerca relativa a questa attività ha lo scopo di verificare ed applicare metodi robusti per l’individuazione automatica di valori anomali di energia e potenza relativi a diversi usi finali per edifici per uffici. L’applicazione di questi metodi è finalizzata a condurre analisi diagnostiche avanzate per il risparmio energetico relativamente all’edificio F40 e al distretto di edifici (di seguito chiamato anche cluster di edifici) presso il centro ENEA Casaccia. Per questa attività sono stati condotti studi sull’utilizzo sia di modelli inversi puramente empirici che di metodi statistici, seguendo un approccio del tipo top down per il rilevamento automatico di anomalie sulla base di dati monitoraggio relativi all’intero edificio (whole building approach). In particolare, la ricerca condotta ha previsto l’applicazione e la verifica delle potenzialità di impiego di “reti neurali (RNA) e diverse tecniche di ensembling di reti neurali (BEM), così come di metodi di clustering, classificazione e di pattern recognition per la caratterizzazione e la stima degli usi finali di energia negli edifici, e per lo sviluppo di tecniche innovative di FDD (Fault Detection and Diagnostic).

L’attività condotta ha previsto anche lo studio e l’implementazione di tecniche diagnostiche per il rilevamento automatico di anomalie connesse all’uso della luce artificiale sulla base di modelli stimativi del livello di illuminamento naturale interno di ogni ufficio nell’edificio F40.

L’analisi di diagnostica condotta ha previsto, inoltre, l’individuazione di opportune funzioni obiettivo per condurre analisi di ottimizzazione energetica nell’edificio F40 attraverso l’applicazione di algoritmi genetici (in collaborazione con l’Università Politecnica della Marche e l’Università Roma 3) relativamente alle variabili temperatura interna di set-point e temperatura di mandata dei fluidi termovettori caldo/freddo dell’impianto di climatizzazione ambientale.

Gli studi condotti hanno portato anche allo sviluppo di regole diagnostiche concepite secondo la suddivisione logica delle anomalie individuate, degli eventi connessi all’anomalia, e delle cause effettive. A supporto della complessa logica decisionale legata all’individuazione delle cause connesse alle anomalie di consumo o di qualità ambientale, riscontrabili a partire dai dati di monitoraggio elaborati secondo grandezze e indicatori aggregati, sono state individuate, infatti, due regole diagnostiche connesse all’utilizzo dell’illuminazione artificiale attraverso metodi di data fusion. Per un approfondimento sullo sviluppo di tali regole diagnostiche si consulti il Report RdS di ENEA.

Di seguito sono descritte le attività sviluppate per le diverse indagini effettuate evidenziando metodologie adottate e finalità perseguite.

Modellazione dei dati dell’edificio F40 e del cluster di edifici per la diagnostica avanzata

Nello studio effettuato sono stati applicati sia metodi basati sull’utilizzo di modelli inversi completamente empirici per la modellazione del comportamento energetico dell’edificio, sia metodi statistici di data mining per l’analisi delle serie di dati di monitoraggio relativi alla richiesta di energia elettrica per diversi usi finali. Pertanto si è seguito un approccio di tipo *top down* finalizzato a condurre analisi di rilevamento di anomalie sulla base di dati di monitoraggio relativi all’intero edificio F40 o a una sua porzione (singolo piano) così come a tutti gli edifici facenti parte del cluster descritto nella prima parte di questa relazione.

Di seguito è presentato lo sviluppo di modelli basati su tecniche di data mining e di metodi basati sull’utilizzo di reti neurali attraverso cui è stato possibile individuare in maniera automatica valori anomali di consumo degli edifici oggetto di indagine e di attivare conseguentemente il processo diagnostico. L’obiettivo finale è l’individuazione della tipologia di modelli in grado di rappresentare con l’errore più basso possibile il comportamento energetico degli edifici oggetto di indagine a partire da forzanti quali dati climatici ed occupazione dell’edificio. Tutte le elaborazioni sono state eseguite utilizzando dati relativi al

funzionamento dell'edificio in assenza di qualsiasi dispositivo e strategia di controllo e attuazione automatici.

Verifica e applicazione di metodi di data mining per analisi di fault detection

La ricerca condotta ha previsto la verifica delle potenzialità e dei limiti di tre differenti tecniche di data mining per condurre analisi di fault detection su valori di energia e potenza relativi all'uso finale illuminazione. Per il cluster di edifici si è analizzato anche il vettore potenza totale di energia elettrica. In particolare sono state sperimentate le potenzialità di un algoritmo di classificazione (CART), e di due tecniche di clustering (K-Means e DBSCAN). L'obiettivo primario è stato quello di individuare metodologie robuste e affidabili finalizzate alla partizione dei data set analizzati in classi omogenee prima di condurre analisi di individuazione delle anomalie. L'ipotesi alla base delle analisi effettuate è che un evento anomalo può essere efficacemente rilevato all'interno di una serie di dati, soltanto se questi sono stati generati sotto le stesse condizioni al contorno. La ricerca effettuata ha permesso di individuare le potenzialità connesse alla metodologia attraverso cui la partizione del dato viene effettuata. Le tecniche di classificazione e di clustering adottate, infatti, prevedono una diversa logica di suddivisione del data set.

Una struttura ad albero, output della procedura di classificazione, a prima vista sembra presentare numerosi elementi di contatto con un dendrogramma di una cluster analysis gerarchica. In realtà, invece, le due tecniche presentano differenze sostanziali sia negli obiettivi che nelle modalità di analisi. La finalità dell'analisi di clustering è quella di accorpate le unità statistiche in gruppi o classi che sono ignote all'analista. Tale raggruppamento avviene attraverso la ricerca di gruppi in cui le osservazioni siano omogenee rispetto alle p variabili X osservate. Diversamente, in un albero di classificazione, i gruppi a cui le unità appartengono sono già noti a priori in una variabile indicata con Y , e attraverso la metodologia si individuano le relazioni tra le p variabili esplicative X che spiegano il perché un'osservazione appartenga ad una classe della Y piuttosto che ad un'altra.

Tale struttura di relazioni, una volta definita, consente anche di predire, per nuove unità statistiche, la classe di appartenenza quando questa non sia nota a priori. Questi due differenti approcci sono noti in letteratura come:

- **classificazione supervisionata**

come nel caso degli alberi di classificazione in cui la ricerca della partizione è guidata (supervisionata) dalla conoscenza a priori della Y ;

- **classificazione non supervisionata**

come nel caso della cluster analysis in cui la ricerca della partizione è effettuata unicamente sulla base della somiglianza delle unità rispetto alle caratteristiche osservate (le variabili X).

Con i metodi CART e K-Means il rilevamento dei valori anomali è stato condotto per ogni classe o cluster individuato, attraverso il metodo GESD (*generalized extreme studentized deviate many outlier procedure*). Quest'ultimo consente di rilevare uno o più valori anomali partendo da un set di dati che seguono approssimativamente una distribuzione normale. A partire dall'individuazione dei valori estremi rispetto al valor medio del vettore oggetto di indagine, i valori anomali vengono individuati dal confronto tra ogni deviazione massima e un valore critico valutato sulla base della distribuzione t-student.

Il metodo DBSCAN, invece, ha consentito di individuare per via diretta i valori anomali attraverso lo stesso processo di clusterizzazione. La metodologia, infatti, prevede l'individuazione di un particolare cluster (cluster 0) nel quale tutti gli outlier e in generale il *rumore* vengono isolati e identificati.

Di seguito sono brevemente descritti i principi teorici dei metodi di classificazione e di clustering utilizzati così come dei metodi selezionati per l'individuazione in ogni classe o cluster dei valori anomali.

Classificazione dei dati con metodo CART

Riguardo al metodo statistico, sulla base degli studi condotti in letteratura, è emersa la necessità di provvedere a una classificazione dei dati in funzione degli attributi rappresentati dalle variabili indipendenti misurate (nei casi oggetto del presente lavoro sono rappresentati dall'energia richiesta o dalla potenza per

ogni step temporale per diversi usi finali), al fine di suddividere il campione di dati in classi omogenee prima di condurre l'analisi di individuazione delle anomalie. *Un evento anomalo può essere efficacemente rilevato soltanto all'interno di una serie di dati, soltanto se questi sono stati generati sotto le stesse condizioni al contorno.*

Tra i metodi di classificazione è stato selezionato il metodo CART (classification and regression tree), poiché esso è risultato flessibile sia rispetto alla quantità di dati da trattare che rispetto alla loro tipologia.

Si tratta di una procedura molto flessibile che consente di maneggiare facilmente variabili indipendenti continue, categoriche o congiuntamente categoriche e continue. L'albero finale include solo le variabili indipendenti che risultano essere predittive della variabile dipendente; le altre variabili indipendenti non predittive non hanno effetto sul risultato finale e sotto questo aspetto il metodo CART si differenzia dalle altre procedure statistiche tradizionali come la regressione.

Inoltre tale tecnica di classificazione, come sarà più chiaro avanti, è basata sulla costruzione di alberi decisionali dai quali è possibile estrapolare le regole (rules based approach) che consentono di classificare i dati che in tempo reale arrivano dal sistema di monitoraggio. Pertanto il metodo proposto è finalizzato all'individuazione di alberi decisionali dai quali è possibile individuare regole basate sui valori che assumono le variabili indipendenti per la classificazione dei dati e per la successiva ed efficace individuazione delle anomalie. Esso quindi risulta particolarmente adatto per condurre analisi di *fault detection in tempo reale*.

Gli **alberi di classificazione** (o di *segmentazione*) rappresentano pertanto una metodologia che ha l'obiettivo di ottenere una segmentazione gerarchica di un insieme di unità statistiche mediante l'individuazione di "regole" che sfruttano la relazione esistente tra una classe di appartenenza e le variabili rilevate per ciascuna unità. Essi vengono utilizzati per individuare l'appartenenza di unità statistiche alle classi di una variabile dipendente conoscendo i valori o le modalità di una o più variabili esplicative. La regola individuata viene successivamente impiegata per classificare nuove unità statistiche di cui si ignora la categoria di appartenenza (*albero decisionale*). Il metodo CART risulta particolarmente flessibile poiché può essere utilizzato sia con variabili numeriche che categoriali e adatto nell'individuazione degli outliers. L'algoritmo attraverso cui viene eseguito lo split dei dati consente di isolare i valori anomali in alcuni nodi terminali. Inoltre un'importante proprietà pratica del metodo CART è che la struttura della sua classificazione è invariante rispetto alle trasformazioni monotone delle variabili indipendenti, per cui può essere sostituito il valore di ogni variabile con il suo logaritmo o il suo quadrato senza che cambi la struttura dell'albero.

Definizione di una struttura ad albero

Per albero si intende un modello grafico costituito da un insieme finito di elementi, detti nodi, che si dipartono da un nodo iniziale denominato nodo radice.

La struttura di un classification tree include i nodi non terminali (*parent nodes*), i quali hanno due discendenti diretti (*child nodes*), ed i nodi terminali che non subiscono ulteriori bipartizioni (*terminal nodes*). Il primo nodo non terminale o parent node contiene tutte le osservazioni che vengono immediatamente suddivise in due gruppi o in due child nodes. Ogni child node diventa un parent node e viene ancora suddiviso in due child nodes. Il processo continua fino a che l'albero non conclude la sua crescita. I nodi terminali o terminal nodes sono i nodi finali dell'albero decisionale e contengono insiemi di osservazioni che vanno a formare classi omogenee al loro interno ma diverse tra loro. Il metodo è binario e ricorsivo: binario poiché ogni parent node si divide in due nodi e ricorsivo poiché i nodi (non terminali) nati dallo split del nodo genitore possono diventare, a loro volta, nodi genitore e suddivisi in due nodi figli successivi.

L'idea di base dalla segmentazione binaria è quella suddividere ricorsivamente un insieme di unità in gruppi sempre più fini, cioè di numerosità inferiore, e sempre più omogenei internamente (rispetto alla distribuzione della variabile risposta). Per definizione i nodi terminali rappresenteranno un grado di omogeneità interna maggiore rispetto al gruppo di partenza. Il ruolo di generatore delle possibili partizioni, o split, viene assunto dai predittori, i quali caratterizzano il passaggio delle unità statistiche della variabile Y da un nodo ai suoi discendenti.

La segmentazione può essere d'ausilio per soddisfare due obiettivi:

1. **esplorativo**, finalizzato principalmente all'individuazione e descrizione delle relazioni esistenti tra le diverse variabili esplicative e l'appartenenza ad una classe piuttosto che ad un'altra della variabile di risposta;
2. **decisionale**, finalizzato alla costruzione di un albero che consenta di classificare nuove unità statistiche per le quali non è nota la classe di appartenenza ma unicamente le modalità assunte per le variabili esplicative.

Le fasi della procedura

Ogni procedura di segmentazione è caratterizzata da un certo numero di fasi che guidano la costruzione dell'albero:

- Creazione dell'*insieme degli split*, cioè dell'insieme dei potenziali tagli binari (ottenuti attraverso le variabili esplicative) che consentono di dividere le unità contenute in un nodo padre in due insiemi che formano i nodi figli;
- Il *criterio di partizione*, passaggio fondamentale consistente in un algoritmo di partizione che genera, a partire dal nodo radice, gruppi sempre più omogenei internamente ed eterogenei dall'esterno;
- La *regola di arresto* della procedura, essenziale per il controllo della dimensione dell'albero finale;
- L'*assegnazione della risposta*, che si esplica con l'assegnazione di una classe alle unità presenti in un nodo terminale;
- La *potatura dell'albero*, che consente di individuare, a partire dall'albero finale, un sottoalbero ottimale che possa essere utilizzato poi a fini decisionali.

E' necessario quindi adottare un criterio di bontà dello split ossia di bontà della tecnica con i cui i nodi vengono suddivisi da parent nodes a child nodes (split criterion); stabilire una regola di arresto di crescita dell'albero (stopping rule); e, una volta che un nodo viene riconosciuto come terminale, fissare una regola che permetta di stabilire come classificare i casi che in esso sono contenuti. Il processo inizia con una domanda binaria che stabilisce se l'osservazione si collocherà nel nodo di destra o di sinistra. Se l'osservazione soddisfa il criterio sarà collocata a destra nel caso contrario a sinistra. Il processo continuerà fino a che tutte le osservazioni saranno divise in child nodes e non verranno raggiunti nodi terminali che saranno tra di loro più differenti possibile ma che andranno a costituire classi molto omogenee al loro interno. Il criterio con i cui i nodi vengono suddivisi da parent nodes a child nodes sarà tanto buono quando sarà alto il grado di omogeneità in ogni nodo terminale ovvero quanto minimo sarà il grado di impurità di ciascun nodo. La regola di arresto o stopping rule è opzionale ma utile ad evitare problemi di overfit. Un'altra strategia consiste nel lasciare che l'albero cresca liberamente e, successivamente, applicare una procedura definita "pruning", che prevede la "potatura" dell'albero eliminando i rami meno significativi. Lo step finale consiste nell'assegnare ciascun nodo terminale ad una classe. Il risultato finale della procedura, consiste, quindi, nell'individuazione delle classi costituite ognuna dagli elementi della variabile Y con carattere di omogeneità rispetto alle variabili esplicative.

Costruzione dell'insieme di split

Primo passo della segmentazione binaria consiste nella individuazione di tutte le potenziali domande dicotomiche (binarie) generate dalle variabili X originarie. Le possibili suddivisioni dipendono dalla natura quantitativa o qualitativa dei predittori.

La fase di ripartizione consiste quindi nell'individuazione, tra l'insieme di tutti i possibili split generati nella fase precedente, del taglio binario "migliore" rispetto ad un dato criterio statistico.

Nella metodologia CART, il criterio statistico che guida la scelta del migliore split si basa sul concetto di decremento di impurità.

Essendo l'obiettivo della segmentazione quello di ottenere nodi sempre più puri, *si individua tra tutti i possibili tagli, lo split ottimo come quello che massimizza la riduzione di impurità che si ottiene tagliando un nodo padre in due nodi figli.*

Sia t_p il nodo padre e t_l e t_r rispettivamente i nodi figlio di sinistra e destra del nodo padre t_p . Si consideri un campione con la matrice delle variabili X con M pari a n numero di variabili ed N osservazioni. Sia il vettore Y costituito da N osservazioni con un valore totale di K classi. L'albero di classificazione è costruito in

accordo con le regole di split. Ogni valore della variabile deve essere diviso in due parti con la massima omogeneità:

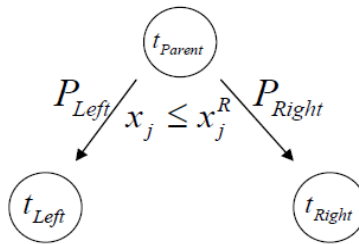


Figura 47: Algoritmo per lo split del metodo CART.

con x_j variabile j e x_j^R il valore relativo del criterio migliore di split per le variabili x_j . La massima omogeneità dei nodi figlio è definita attraverso la funzione di impurità $i(t)$. Dato che il grado di impurità del nodo padre t_p è costante per ognuno dei possibili split $x_j \leq x_j^R$, $j = 1, \dots, M$ la massima omogeneità dei nodi figlio sinistro e destro sarà equivalente alla massimizzazione del decremento della funzione di impurità.

Il decremento di impurità ad un nodo t ottenuto con uno split s , si definisce come:

$$\Delta i_Y(t,s) = \{i_Y(t) - [i_Y(t_L)p(t_L) + i_Y(t_R)p(t_R)]\}$$

dove $i_Y(t)$ è il grado di impurità nel nodo padre t ;

$i_Y(t_L)$ è il grado di impurità nel nodo figlio di sinistra e $p(t_L)$ è la proporzione di unità contenute nel nodo di sinistra;

$i_Y(t_R)$ è il grado di impurità nel nodo figlio di destra e $p(t_R)$ è la proporzione di unità contenute nel nodo di destra.

Il migliore split s^* è quello che massimizza il decremento di impurità:

$$s^* \rightarrow \max \Delta i_Y(t,s)$$

L'algoritmo CART quindi, ad ogni nodo genera l'insieme S di tutte le possibili partizioni binarie (split), calcola il decremento di impurità e determina la miglior partizione cui è associato il massimo decremento di impurità.

In sostanza, l'algoritmo CART si compone dei seguenti passi:

- *Passo 1. si genera l'insieme S di tutte le possibili partizioni binarie ottenute dal set di predittori X ;*
- *Passo 2. per ogni split s dell'insieme S si calcola il decremento di impurità;*
- *Passo 3. si determina la miglior partizione a cui è associato il massimo decremento di impurità.*

L'algoritmo è applicato ad ogni nodo fino a che la costruzione dell'albero non si arresta. Il costo computazionale di questa metodologia è molto elevato. Infatti basti pensare al caso in cui i predittori impiegati sono in numero sostanzioso ed inoltre parte di essi sono in scala numerica o nominale. In questo caso il numero di split S che deve essere generato ad ogni nodo è considerevole soprattutto se si pensa che ogni volta per ognuno di esso, l'algoritmo deve calcolare il decremento di impurità per poi selezionare la migliore partizione.

Nel presente lavoro per applicare la metodologia CART è stato adottato quale misura d'impurità, per gli alberi di classificazione, l'indice **H di eterogeneità di Gini**.

Esso si definisce in generale come:

$$H = 1 - \sum_{j=1}^J f_j^2$$

dove f_j rappresenta la frequenza relativa di osservazioni la cui modalità della variabile è pari a j .
Nell'ambito della segmentazione binaria, l'**impurità in un nodo** sarà quindi pari a:

$$i_Y(t) = 1 - \sum_{j=1}^J p^2(t|Y=j)$$

dove $i_Y(t)$ è la misura di impurità in un generico nodo t e $p(t|Y=j)$ è la proporzione di unità nel nodo t che appartengono alla j -esima classe della variabile di risposta Y .

Regole di arresto

Le regole d'arresto della procedura rappresentano l'insieme di criteri che determinano quando un nodo debba essere dichiarato terminale e quindi non più ulteriormente divisibile in ulteriori nodi figli.

Esse consistono nelle seguenti condizioni: "**Un nodo t diventa terminale se**":

- La numerosità dello stesso è inferiore ad una certa soglia prefissata;*
Si fissa una soglia minima per il numero di osservazioni contenute in un nodo padre o eventualmente nei nodi figli generati da questo. La regola serve ad ottenere alberi i cui nodi non siano espressione di singole o pochissime unità fornendo così percorsi poco informativi;
- Il grado di impurità del nodo t è al di sotto di una certa soglia prefissata;*
Se il nodo ha un grado di purezza elevato allora sue ulteriori partizioni non produrranno alcun miglioramento nell'accuratezza della struttura ma unicamente uno svantaggio misurato dalla crescita della complessità dell'albero.

Si rende pertanto necessaria una procedura di induzione dell'albero che da una parte lo semplifichi e dall'altra consideri l'accuratezza di previsione per nuove unità.

Tale procedura prende il nome di **pruning** (*potatura*).

Il pruning si prefigge l'obiettivo di individuare le branche meno rilevanti o addirittura dannose per il processo decisionale e di rimuoverle.

Il pruning

Il pruning della metodologia CART è un metodo che genera una sequenza ottimale nidificata di sottoalberi potati tra i quali ne viene selezionato uno finale quale regola di decisione per nuove unità.

Molto sinteticamente, la procedura di pruning opera nel modo seguente:

Si definisce una misura che tenga conto del trade-off tra il costo (*l'aumento della impurità dell'albero*) e il beneficio (*semplificazione della struttura*) legato alla potatura di un albero. Tale misura α_t , detta **parametro di costo complessità** è calcolata ad ogni nodo interno dell'albero:

$$\alpha_t = \frac{R(t) - R(T_t)}{|\tilde{T}| - 1} = \frac{\text{aumento impurità}}{\text{riduzione della complessità}}$$

Dove al numeratore si misura la crescita di mal classificati come differenza tra il tasso al nodo t (che se potato diventerebbe terminale) e quello del ramo T_t che diparte da t . Mentre al denominatore si misura la riduzione della complessità della struttura in termini di riduzione del numero di nodi terminali.

Ad ogni passo verrà potato il ramo il cui nodo di partenza presenta l'alfa minimo dell'intera struttura ad albero.

In questo modo, ripetendo iterativamente la procedura, si otterrà una sequenza di alberi, via via più piccoli, tutti potenzialmente candidati ad essere scelti come albero decisionale.

Simulando in ognuno di essi un nuovo campione di dati (il **campione test**) si potrà scegliere l'albero migliore come quello che minimizza il tasso di errata classificazione del campione test.

I principali step della metodologia statistica adottata per l'analisi di "fault detection" sono di seguito riassunti:

- Analisi di sensibilità sui dati di monitoraggio al fine di individuare le variabili indipendenti di maggiore importanza rispetto alla variazione delle variabili dipendenti (energia o potenza);
- Classificazione dei dati attraverso il metodo CART e utilizzo del criterio di GINI per effettuare lo splitting dei dati;
- Applicazione di un metodo statistico robusto per la determinazione dei valori anomali per ogni classe individuata;
- Individuazione di indicatori statistici adimensionali sintetici per valutare il grado di importanza o di severità di ogni valore anomalo individuato in ogni classe (outlier individuati).

Clustering dei dati con i metodi K-Means e DBSCAN

Gli algoritmi selezionati per l'analisi di clustering appartengono a due diverse categorie: il metodo K-Means è un classico metodo di partizione mentre il metodo DBSCAN è un metodo basato sulla densità delle unità statistiche o oggetti nel campo di interesse. Questi metodi richiedono la definizione di una opportuna metrica per computare la distanza tra gli oggetti nel data set; in questo studio si è adottata la distanza Euclidea valutata sulla base dei dati opportunamente normalizzati. I metodi di partizione prevedono la suddivisione del data set costituito da n oggetti in k partizioni o cluster dove $k < n$. Il criterio generale adottato attraverso cui è operata la partizione dei dati prevede che gli oggetti del data set siano assegnati allo stesso cluster quando questi sono vicini, e a cluster differenti quando invece risultano lontani sulla base di una particolare metrica che consenta di valutare la distanza.

In generale i metodi di partizione risultano efficaci solo nel caso di cluster a forma sferica e se i cluster sono bene separati, inoltre risultano in generale sensibili alla presenza di outlier.

I metodi del tipo density based sono concepiti per gestire anche cluster di forma non sferica e in generale risultano meno sensibili alla presenza di valori anomali. L'obiettivo di questa tipologia di metodi è identificare porzioni dello spazio di dati caratterizzati da una elevata densità di oggetti. La densità è definita come il numero di oggetti che sono in una particolare area di uno spazio n -dimensionale. La strategia generale di tale metodo è l'esplorazione dello spazio di dati aumentando il numero di cluster fino a quando il numero di oggetti nel relativo cluster supera una data soglia.

Metodo di partizione (K-Means)

Il metodo K-Means richiede in ingresso la conoscenza del parametro k che rappresenta il numero di partizioni in cui il data set verrà suddiviso. Esso rappresenta ogni cluster con il valore medio degli oggetti che esso aggrega e viene chiamato centroide. L'algoritmo alla base del metodo è basato su una procedura iterativa, preceduta da una fase di set up, dove k oggetti del dataset sono scelti in modo random come valori centroidi iniziali. Ogni iterazione è caratterizzata da due fasi; nella prima fase ogni oggetto è assegnato al cluster il cui centroide è il più vicino dall'oggetto. Nella seconda fase i centroidi sono ricollocati attraverso il calcolo del valore medio degli oggetti all'interno di ogni cluster. Il processo iterativo continua finché i k centroidi non cambiano più. Il metodo K-Means risulta efficace per i cluster di forma sferica mentre per cluster di forma differente esso risulta efficace solo se gli oggetti sono ben separati.

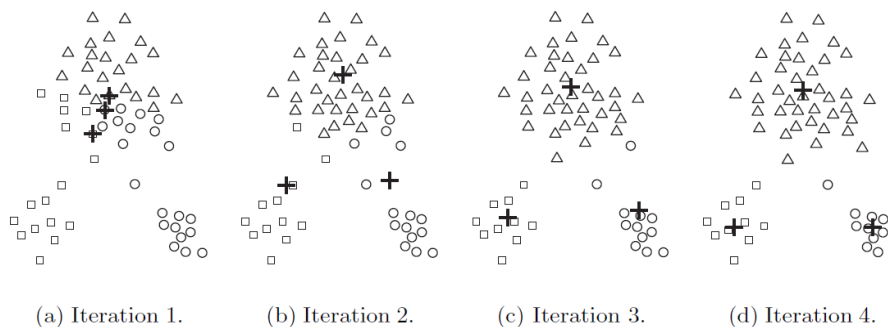


Figura 48: Algoritmo iterativo per la formazione dei cluster nel metodo K-Means.

Density Based (DBSCAN)

Il metodo DBSCAN richiede invece in ingresso il valore di due parametri, il raggio r e il minimo numero di punti $minPts$ definiti per individuare un valore di densità di riferimento nel campo spaziale dei dati.

Un'area ad alta densità nello spazio dei dati rappresenta una sfera n -dimensionale con un raggio pari ad r che contiene almeno un numero di oggetti pari a $minPts$.

In altre parole, se esiste un numero di oggetti superiore a $minPts$ la cui distanza dall'oggetto considerato come riferimento è minore di r , allora questo oggetto e tutti quelli vicini danno luogo a un nuovo cluster. Dato che l'efficacia di questo metodo è fortemente influenzata dal valore assunto da questi due parametri, è stata condotta un'opportuna analisi di sensibilità conducendo numerose simulazioni per individuarne correttamente i valori.

Il metodo DBSCAN risulta efficace nell'individuazione di cluster di qualsiasi forma e soprattutto presenta la caratteristica di segregare i valori anomali in un cluster attraverso la ricerca di aree a bassa densità di oggetti nello spazio di dati.

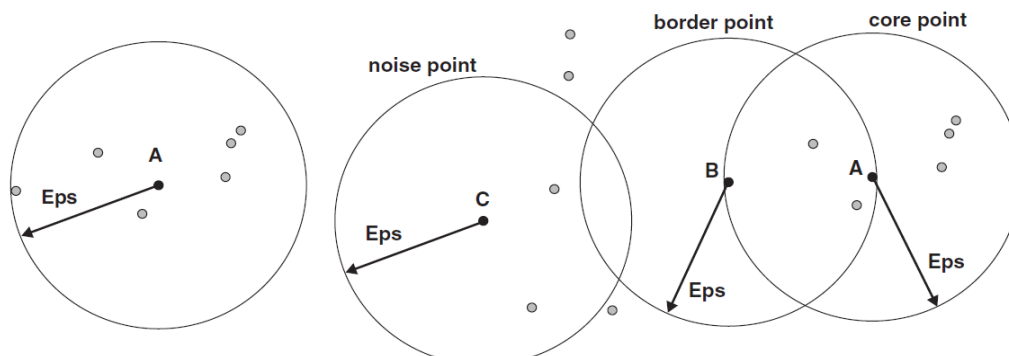


Figura 49: Algoritmo per la formazione dei cluster nel metodo DBSCAN.

Metodi per il rilevamento e l'identificazione di valori anomali di energia e potenza

Un valore anomalo è definito come un'osservazione che appare inconsistente rispetto alla maggioranza delle osservazioni di un data set di dati. I valori anomali in un data set possono avere differenti cause come errori umani, errori di monitoraggio, variazione nella prestazione dei sistemi o fault nei sistemi. In questo studio per la determinazione dei valori anomali nel data set sono stati implementati due diversi metodi i cui modelli si basano su diversi principi metodologici.

Di seguito sono descritti le basi teoriche dei metodi utilizzati.

Metodo del peak detection

Il primo metodo implementato, "peak detection method" [4], consente di identificare ed analizzare i picchi presenti in una serie temporale, esigenza essenziale, in tutte quelle applicazioni (quali il monitoraggio dei

consumi energetici) che richiedono l'analisi di grandi quantità di dati in tempo reale. L'analisi condotta ha richiesto pertanto la formalizzazione della nozione di picco al fine di mettere a punto un algoritmo capace di rilevarli automaticamente in una data serie temporale.

Sulla base di un esteso studio di letteratura l'attenzione si è focalizzata su algoritmi per la rilevazione dei picchi nelle serie di dati basati sull'utilizzo di una funzione picco S , che associa un punteggio (score) ad ogni elemento della serie: un elemento viene considerato un picco quando lo score ad esso associato è positivo e supera una certa soglia θ definita dall'utente o opportunamente calcolata.

Sia $T = x_1, x_2, \dots, x_N$ una serie temporale univariata campionata in maniera uniforme, contenente N valori. Sia x_i un i -esimo punto assegnato in T , con relativo score $S(i, x_i, T)$ associato dalla funzione S : se tale score supera la soglia θ ($S(i, x_i, T) \geq \theta$), allora x_i è un picco. Non tutti i picchi individuati tuttavia sono "veri" picchi: un picco locale è un vero picco se ha un valore ragionevolmente grande anche nel contesto globale. Negli step che seguono, in cui si articola l'algoritmo implementato per la peak detection, si specifica come eliminare alcuni dei picchi locali rilevati utilizzando il contesto globale (l'intera serie temporale).

- 1) Si inizializza l'insieme dei picchi O come insieme vuoto;
- 2) Si inizializza il valore di k (tipicamente tra 1 e 5) e di un ulteriore costante h (tipicamente tra 1 e 3);
- 3) Si calcola il valore della funzione picco per ognuno degli N punti in T ;
- 4) Si calcola la media m' e la deviazione standard s' di tutti i valori positivi della funzione picco S ;
- 5) Si eliminano i picchi locali che sono "piccoli" nel contesto globale. In particolare, si inseriscono nell'insieme O solo i punti x_i della serie temporale che soddisfano la seguente condizione:

$$S(k, i, x_i, T) - m' > h * s'$$

- 6) Si eliminano quei picchi che sono troppo "vicini" tra di loro (ossia nella stessa finestra di dimensione k). Per ogni coppia di picchi adiacenti x_i e x_j in O , se $|j-i| \leq k$ si rimuove dall'insieme O il valore più piccolo tra x_i e x_j .

Metodo GESD

Il secondo metodo utilizzato, GESD (generalized extreme studentized deviate many outlier procedure), consente di rilevare uno o più valori anomali partendo da un set di dati che seguono approssimativamente una distribuzione normale. A partire dall'individuazione dei valori estremi rispetto al valor medio del vettore oggetto di indagine, i valori anomali vengono individuati dal confronto tra ogni deviazione massima e un valore critico valutato sulla base della distribuzione t student.

La procedura ESD consente di identificare gli elementi di un set di dati che rappresentano valori anomali. Il modello prevede che l'utente debba specificare la probabilità α di individuare dei falsi outlier e il numero massimo n_u di potenziali outliers.

Nel presente lavoro, seguendo le indicazioni di Carey et al. [21], il numero n_u è stato individuato trovando il più grande intero che soddisfa la seguente disuguaglianza: $n_u \leq 0,5 (n - 1)$.

Di seguito è descritta per step la procedura alla base del metodo per l'individuazione dei valori anomali di un set di dati X :

- 1) Viene fissato inizialmente un numero di outlier n_{out} pari a zero.
- 2) Viene calcolata la media degli elementi nel set X :

$$\bar{x} = \frac{\sum_{j=1}^n x_j}{n}$$

dove x_j è un elemento del vettore X ed n è uguale al numero di elementi in X

- 3) Viene calcolata la deviazione standard degli elementi nel vettore X . Questa è pari a:

$$s = \sqrt{\frac{\sum_{j=1}^n (x_j - \bar{x})^2}{n-1}}$$

Se la deviazione standard è uguale a zero allora tutti gli elementi nel vettore X hanno lo stesso valore per cui non ci sono ulteriori valori anomali.

- 4) Viene individuato il valore estremo $x_{e,i}$ nel vettore X. L'elemento è identificato come l'osservazione del vettore X che presenta la massima distanza dal valor medio.

$$x_{e,i} = \max |x_j - \bar{x}|$$

- 5) Viene calcolata la massima deviazione R_i che rappresenta la misura normalizzata della distanza dell' i-esimo valore estremo dal valor medio del campione:

$$R_i = \frac{|x_{e,i} - \bar{x}|}{s}$$

- 6) Viene calcolato l' i-esimo valore critico λ_i attraverso la seguente equazione

$$\lambda_i = \frac{(n-i)t_{n-i-1,p}}{\sqrt{(n-i+1)(n-i-1)t_{n-i-1,p}^2}}$$

dove $t_{n-i-1,p}$ rappresenta la distribuzione t student con (n-i-1) gradi di libertà e con probabilità p determinata attraverso la seguente relazione:

$$p = \frac{\alpha}{2(n-i+1)}$$

- 7) Se il valore R_i risulta maggiore del valore λ_i per il corrispondente valore $x_{e,i}$, il valore di n_{out} (numero degli outliers) viene posto pari a i.
- 8) Viene rimosso il valore estremo $x_{e,i}$ dal vettore X per cui quest'ultimo presenta un numero di valori pari a n-i. Se i è uguale ad n_u allora la procedura si ferma, diversamente ricomincia un nuovo ciclo su i considerando il nuovo vettore X costituito da un numero di valori pari n-i dove i rappresenta il numero di valori estremi che rispettano la disuguaglianza di cui al punto 8.
- 9) Il modello restituisce alla fine il vettore costituito dai valori estremi $x_{e,i}$ considerati anomali. Si osserva che il modello non considera valori anomali tutti i valori estremi trovati durante la procedura ma soltanto i primi valori $i=n_u$.

Indicatori adimensionali per la diagnostica degli edifici

A seguito dell'individuazione dei valori anomali sono stati introdotti alcuni indici sintetici che consentono di quantificare la severità della anomalie riscontrate, e quindi di fornire informazioni circa la priorità di intervento da associare ad ognuna di esse. Tali indici si basano sulla determinazione del grado e della

direzione della variazione di ogni outlier da una distribuzione normale. A tal fine sono stati valutati gli indici *z score*, *s function* e *modified z score*.

Quest'ultimo, è basato sulla distanza e la direzione di ogni valore anomalo individuato rispetto al valore medio di osservazioni normali (osservazioni che non contengono valori anomali).

Il metodo prevede la determinazione dell'indice *modified zscore* sulla base della stima del valore medio e della deviazione standard per il set di dati X osservati.

Sulla base della conoscenza del vettore costituito dagli n_{out} valori anomali individuati con i metodi sopra descritti, si individua il set di valori del vettore X, $X_{non-out}$, costituito dai soli valori non anomali $n-n_{out}$.

Si calcola la media e la deviazione standard del vettore $X_{non-out}$:

$$\bar{X}_{rob} = \frac{\sum_{j=1}^{n-n_{out}} X_j}{n-n_{out}}, \quad S_{rob} = \sqrt{\frac{\sum_{j=1}^{n-n_{out}} (X_j - \bar{X}_{rob})^2}{n-n_{out}-1}}$$

Il valore di *modified zscore* viene calcolato per ogni valore anomalo individuato nel seguente modo:

$$Z_m = \frac{X_{out} - \bar{X}_{rob}}{S_{rob}}$$

Il valore di *zscore* viene invece calcolato per ogni valore anomalo individuato nel seguente modo:

$$Z = \frac{X_{out} - \bar{X}}{S}$$

Allo stesso modo è stato assegnato un punteggio anche ai valori del vettore che non sono risultati outlier calcolando anche per essi l'indice *modified zscore* e *zscore*:

$$Z_m = \frac{X_{non-out} - \bar{X}_{rob}}{S_{rob}}$$

$$Z = \frac{X_{non-out} - \bar{X}}{S}$$

Per il calcolo della *s function*, che cattura la "spigolosità" del punto x_i in un contesto locale, esistono differenti approcci. L'approccio utilizzato nei test effettuati, per un assegnato punto x_i in T, calcola la media tra la massima distanza di x_i dai suoi k vicini a sinistra e i suoi k vicini a destra (dove k è un parametro intero positivo stabilito dall'utente, che deve essere tanto più piccolo quanto più i picchi da individuare sono di breve durata). I valori di $s(i, x_i, T)$ indicano l'"importanza" dell'altezza del picco all'istante di tempo i -esimo:

$$s(k, i, x_i, T) = \frac{\max\{x_i - x_{i-1}, x_i - x_{i-2}, \dots, x_i - x_{i-k}\} + \max\{x_i - x_{i+1}, x_i - x_{i+2}, \dots, x_i - x_{i+k}\}}{2}$$

Tale funzione calcola per ogni punto della serie il suo valore utilizzando la finestra locale di dimensione $2k$ attorno a quel punto. Tutti i punti con un valore positivo della funzione picco sono picchi candidati.

Risultati dell'analisi di fault detection per l'edificio F40 con tecniche di data mining

La prima analisi di fault detection è stata eseguita sull'edificio F40 situato nel centro ENEA alla Casaccia. L'edificio è costituito da tre piani ed è stato strumentato con un sistema di monitoraggio finalizzato alla misura dell'energia elettrica e termica e delle condizioni ambientali interne ed esterne. Inoltre ogni ufficio è caratterizzato dalla presenza di un sensore di presenza. Nel presente studio gli esperimenti sono stati

condotti sui dati di energia elettrica per illuminazione riferiti al primo piano. In questo piano sono presenti 13 uffici di differente dimensione con una superficie variabile tra 13 ai 36 m² e due sale CED ognuna con una superficie di circa 20 m²

A partire dai dati di monitoraggio sono stati condotti numerosi test con l'applicazione dei metodi sopra citati alle variabili energia, potenza massima, elaborate con frequenza temporale oraria, considerando come attributi la data, il giorno della settimana, l'ora (solo per il metodo CART), la radiazione solare, il numero di stanze attive (una stanza viene considerata attiva quando è presente almeno un occupante all'interno) e il numero di occupanti corrispondenti. In particolare l'energia e la potenza massima per l'illuminazione sono stati analizzati per i mesi di Dicembre 2013 e Gennaio 2013.

Al fine di verificare l'accuratezza e l'efficacia dei metodi proposti sono stati creati due fault artificiali accendendo il 24 e 25 Gennaio 2013 alla fine dell'orario di lavoro tra le 17:30 e le 18:00 con una bassa occupazione tutte le luci del primo piano creando in questo modo un picco di richiesta di energia elettrica richiesta.

A seguito dell'individuazione dei valori anomali per ogni classe o cluster individuata sono stati calcolati gli indici sintetici che consentono di quantificare la severità della anomalie riscontrate, e quindi di fornire informazioni circa la priorità di intervento da associare ad ognuna di esse. Tali indici, come si è descritto sopra, si basano sulla determinazione del grado e della direzione della variazione di ogni outlier da una distribuzione normale. A tal fine sono stati valutati gli indici z score, e modified z score.

L'analisi condotta per l'edificio F40 ha pertanto lo scopo di verificare le potenzialità e i limiti dei metodi di data mining applicati per il rilevamento di anomalie nei dati di energia e potenza e di operare un confronto sulla base della capacità di rilevamento dei fault artificiali creati.

Analisi e risultati ottenuti attraverso la classificazione CART

Per la classificazione dei dati relativi all'energia e la potenza massima richiesta nell'edificio F40 per l'uso finale illuminazione sono stati costruiti opportuni alberi di classificazione. Dall'osservazione dei dati si deduce che il massimo numero di persone e di stanza attive sono rispettivamente pari a 21 e 14. Inoltre i valori di irradianza solare variano da 0 a 476.45 W/m². Il numero degli occupanti è più alto nell'intervallo temporale 0900-1600 e cioè durante le ore lavorative. L'analisi di sensibilità condotta sui dati ha dimostrato che sia la variabile energia che potenza elettrica presentano una relazione e dipendenza con le altre variabili misurate come la presenza degli occupanti, la radiazione solare, il giorno della settimana e le stanze attive, per cui è evidente che valori estremi nel data set non possono definirsi per via diretta valori anomali, poiché ci sono altre variabili che possono influenzare l'entità del consumo energetico. Risulta essenziale quindi classificare il data set secondo gli attributi che influenzano il consumo di energia elettrica prima di individuare la presenza di valori anomali. Pertanto, sia con la variabile energia che potenza elettrica sono stati costruiti differenti alberi decisionali considerando come attributi il giorno della settimana, l'ora del giorno, il numero di stanze attive, il numero di occupanti, e i valori di irradianza solare. A seguito dell'applicazione dell'algoritmo CART e considerando un opportuno criterio di pruning dell'albero, il data set è stato suddiviso in 9 classi per la variabile energia e 10 classi per la variabile potenza. Le classi individuate per l'energia e la potenza sono state analizzate separatamente e nelle tabelle seguenti viene presentato un quadro riassuntivo circa le loro caratteristiche. Nelle tabelle i valori di irradianza solare minori di 150 W/m² sono considerati bassi.

Tabella 8: Classi individuate con il metodo CART considerando la variabile energia e descrizione delle principali caratteristiche

Class	Time	People Presence	Active Room	Solar Radiations	Day
3	Mostly early morning	Zero or one	Zero or one	Mostly zero	Thu-Fri
6	1800-2100 hrs	Mostly < 7	Mostly < 7	zero	Weekdays
8	Weekends-all day Weekdays evenings	Zero or one	Zero or one	Zero except weekends	Weekends and Mon-Wed
9	0700, 0800 hrs	80 % ≤10	≤10	Lower values	Weekdays
11	0600-0700 hrs	zero	zero	Mostly zero	Weekends
12	0600-0700 hrs	zero	zero	zero	Mon-Wed
13	Different timing	Mostly ≥ 10	Mostly ≥ 10	Lower values	Weekdays
15	1200-1600 hrs	≥ 10	≥ 10	Higher values	Weekdays
16	0900-1100 hrs and 1700hrs	≥ 10	≥ 10	Higher values	Weekdays

Tabella 9: Classi individuate con il metodo CART considerando la variabile potenza e descrizione delle principali caratteristiche

Class	Time	People Presence	Active Room	Solar Radiations	Day
6	1800-2100 hrs	80% ≤ 5	80% ≤ 5	zero	Weekdays
7	0600-0800 hrs	Zero	Zero	Zero or lower values	Thu-Fri
9	Evening and early morning	Mostly zero	Mostly zero	Mostly zero	Thu-Fri
10	Weekends-all day, Weekdays early morning	Zero	Zero	Zero except weekends	Weekends and Mon-Wed
11	0700, 0800 hrs	Mostly ≤ 10	Mostly ≤ 10	Lower values	Weekdays
13	0600-0800 hrs	Zero	Zero	Mostly zero	Weekends
14	0600-0800 hrs	Zero	Zero	Mostly zero	Mon-Wed
15	Different timing	Almost 60 % ≥ 10	Almost 60 % ≥ 10	Lower values	Weekdays
17	0900-1700 hrs	Almost 70 % ≥ 10	Almost 60 % ≥ 10	Medium range	Weekdays
18	0900-1500 hrs	Mostly ≥ 10	Mostly ≥ 10	Higher values	Weekdays

Le analisi eseguite anche attraverso metodi grafici di rappresentazione dei dati contenuti in ogni classe hanno dimostrato che le classi 6, 9, 11, 13 e 15 relativamente alla variabile energia sono pure e cioè non presentano valori anomali. Allo stesso modo per la variabile potenza le classi 6, 7, 10, 11, 13, 14 e 18 possono considerarsi pure.

Nella figura seguente invece sono presentati alcuni scatter plots relativi alla classe 3 per la variabile energia e alla classe 17 per la variabile potenza. Nei grafici, i valori di energia e potenza per le classi selezionate sono rappresentati in funzione del numero di occupanti e dell'ora del giorno. E' evidente attraverso questa efficace rappresentazione come alcuni valori sia di energia che di potenza elettrica sembrano avere carattere anomalo rispetto agli altri dati contenuti nella stessa classe in relazione alle variabili indipendenti selezionate.

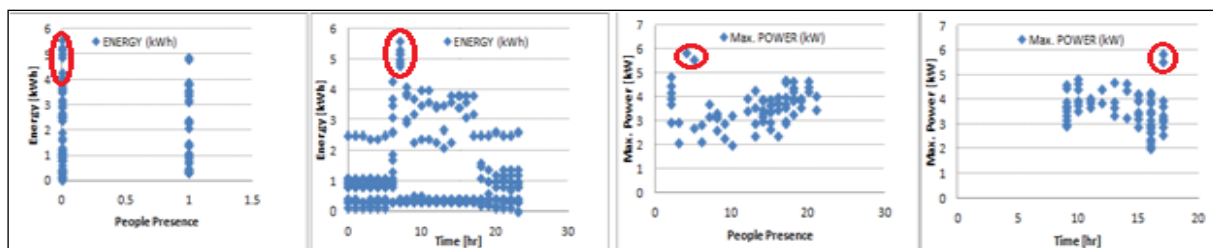


Figura 50: Scatter plots relativi alla variabile energia e potenza per due classi di dati (i valori anomali sono cerchiati)

Per ogni classe individuata sia per l'energia che per la potenza elettrica è stato applicato il metodo GESD. L'entità della variazione da un valore normale è stato determinato attraverso l'indicatore Mzscore

valutando il valore medio e la deviazione standard dei dati per ogni classe sulla base dei valori anomali individuati dal metodo GESD. I valori anomali individuati per l'energia e la potenza dimostrano che la costruzione delle classi attraverso il metodo CART è stata correttamente effettuata. Per esempio nella classe 3 per la variabile energia si è riscontrato che il consumo è maggiore di zero quando non vi sono occupanti nell'edificio. I più significativi valori anomali sono stati individuati al mattino presto tra le 6:00 e le 7:00 dove il valore dell'energia elettrica presenta valori molto simili a quelli che assume durante l'orario di lavoro con una presenza di almeno 15 occupanti. Allo stesso modo nella classe 8 relativa alla variabile energia i principali valori anomali sono stati individuati alle 8:00 del mattino quando il numero di occupanti è molto basso o pari a zero. Tuttavia, considerando il valore medio di energia elettrica oraria come variabile dipendente, i due fault artificiali creati non sono stati individuati attraverso il metodo GESD, anche se entrambi i valori di energia in corrispondenza dei picchi artificiali creati sono stati inclusi con il metodo CART nella stessa classe (classe 16). I risultati ottenuti, considerando invece la variabile potenza massima e l'applicazione dell'algoritmo GESD, sono risultati particolarmente positivi poiché la metodologia seguita ha consentito di individuare i due fault artificiali creati, oltre che una serie di valori anomali considerabili come fault reali. Si è compreso, pertanto, che i due fault creati possono essere identificati se viene considerata la variabile potenza elettrica massima oraria piuttosto che il valore medio dell'energia elettrica per l'illuminazione. Nella figure seguenti è rappresentato l'andamento nel tempo della potenza per l'illuminazione relativa al primo piano dell'edificio F40 con l'evidenza dei valori anomali individuati e del relativo valore di Mzscore in una classe selezionata. E' evidente come attraverso il metodo CART i due fault artificiali siano correttamente individuati nella classe 17. Inoltre è chiaro dalle figure presentate come i valori anomali individuati con efficacia quando sono inseriti nella classe, risultano invece difficilmente identificabili dall'osservazione del data set non classificato. I valori anomali all'interno della classe di appartenenza sono molto più evidenti come dei picchi e possono essere localizzati con più facilità. E' evidente inoltre come il valori di Mzscore risultino amplificati in corrispondenza dei valori anomali, rispondendo alle proprietà richieste di indice diagnostico utile all'identificazione della severità dell'outlier individuato.

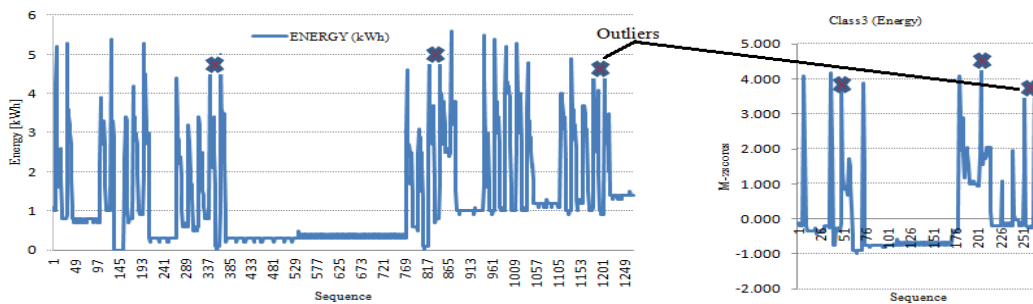


Figura 51: Andamento della potenza elettrica attiva per illuminazione nel tempo con dati non classificati e variazione dell'indice Mzscore in un classe di dati (classificazione con metodo CART) con evidenza dei principali outlier individuati con GESD

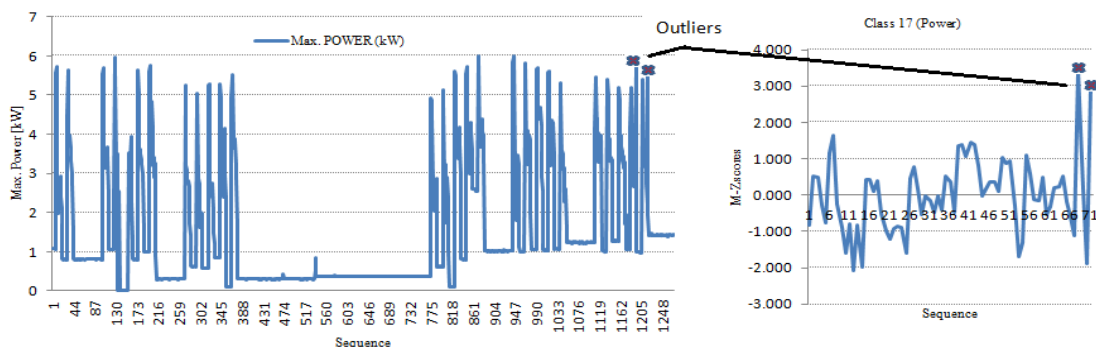


Figura 52: Andamento della potenza elettrica attiva per illuminazione nel tempo con dati non classificati e variazione dell'indice Mzscore in un classe di dati (classificazione con metodo CART) con evidenza dei principali outlier individuati con GESD

Il metodo CART (classification and regression tree) è risultato particolarmente flessibile sia rispetto alla quantità di dati da trattare che rispetto alla loro tipologia. Attraverso la costruzione di alberi decisionali è stato possibile estrapolare le regole (rules based approach) per la classificazione dei dati che in tempo reale arrivano dal sistema di monitoraggio. Pertanto il metodo proposto è stato possibile individuare regole basate sui valori che assumono le variabili indipendenti per la classificazione dei dati e per la successiva individuazione delle anomalie.

Analisi e risultati ottenuti attraverso l'applicazione di metodi di clustering

Gli algoritmi di clustering risultano particolarmente efficaci nell'identificazione di particolari correlazioni e similarità tra i dati monitorati (energia, presenza occupanti, stanza attive, irradianza solare) al fine di rilevare in modo automatico e con maggiore accuratezza i valori identificabili come anomalie.

Stante il fatto che le variabili legate al tempo (ora, giorno, data) non possono essere utilizzate in questi algoritmi (K-Means e DBSCAN), il data set, sia per il valore medio di energia, che di potenza massima, è stato preventivamente suddiviso in valori relativi al periodo diurno e notturno tenendo conto dei giorni feriali e dei fine settimana. Pertanto la ricerca condotta è stata finalizzata anche a comprendere la potenzialità di applicazione di algoritmi di clustering per il fault detection in assenza di informazioni relative ai pattern temporali. Prima di applicare l'analisi di clustering i dati misurati sono stati normalizzati attraverso il metodo z-score.

Applicazione del metodo K-Means

Particolare cura è stata posta nella modellazione degli algoritmi di clustering selezionati. Il metodo K-Means richiede come parametri di input il numero di partizioni k in cui il data set può essere suddiviso. Il numero di cluster k è stato individuato attraverso un algoritmo gerarchico applicando il metodo Ward e conducendo una opportuna analisi di sensibilità. La distanza tra gli oggetti del data set è stata misurata attraverso l'approccio Euclideo considerando i dati opportunamente normalizzati. Per l'individuazione di cluster sono stati analizzati separatamente i valori normalizzati di energia e potenza considerando inoltre le variabili di occupazione, stanza attive e irradianza solare.

Al fine di individuare i valori anomali è stato applicato il metodo GESD a ogni cluster costruito sia per la variabile energia che per la variabile potenza. Sulla base dei risultati ottenuti con l'analisi di clustering, nelle tabelle seguenti sono presentate le principali caratteristiche dei dati per ogni cluster individuato, considerando solo le ore diurne relativamente alla variabile energia e potenza massima.

Tabella 10: Cluster individuati con il metodo K-Means considerando la variabile energia e descrizione delle principali caratteristiche (ore diurne)

Clu No	People Presence	Active Room	Solar Radiations	Standard deviation of Energy [kWh]
1	Almost 60 % ≥ 10	Almost 60 % ≥ 10	Higher values	0.6204
2	Almost 80 % ≤ 5	Almost 80 % ≤ 5	Lower values	0.9579
3	Always ≥ 7	Always ≥ 7	Different range	0.9112
4	Almost 85% ≤ 3	Almost 85% ≤ 3	Mostly zero	0.4388

Tabella 11: Cluster individuati con il metodo K-Means considerando la variabile potenza massima e descrizione delle principali caratteristiche (ore diurne)

Clu No	People Presence	Active Room	Solar Radiations	Standard deviation of Power [kW]
1	Almost 60 % ≤ 5	Almost 60 % ≤ 5	Higher values	0.5945
2	Almost 95 % ≤ 5	Almost 95 % ≤ 5	Lower values	1.8911
3	Always ≥ 10	Always ≥ 10	Higher values	0.6362
4	Almost 60% ≥ 15	Almost 65% ≥ 10	Lower values	1.1504

I risultati dell'analisi effettuata dimostrano che molti dei cluster individuati attraverso il metodo K-Means non sono puri e i valori anomali risultano pertanto disseminati tra questi. Relativamente alla variabile energia i fault artificiali sono stati inseriti nello stesso cluster (cluster 2) mentre relativamente alla variabile potenza questi sono presenti nei cluster 2 e 4. In particolare il cluster 4 risulta non puro sia per la variabile energia che per la variabile potenza massima. Nel cluster 4 per la variabile energia sono stati individuati

alcuni falsi positivi soprattutto alla sera e al mattino presto, e all'opposto per la variabile potenza massima. Il motivo del rilevamento di alcuni falsi positivi risiede probabilmente nel fatto che alcune variabili in corrispondenza di questi falsi positivi, presentano valori molto simili a quelli dei fault reali. L'applicazione del metodo GESD e successivamente del metodo boxplot ai cluster individuati non ha condotto all'identificazione dei fault artificiali presenti nel cluster 2 per la variabile energia. Invece è stato correttamente identificato il fault presente nel cluster 4 considerando la variabile potenza massima. Nella figura seguente è rappresentato l'andamento nel tempo della potenza massima e la variazione dell'indice Mzscore relativamente al cluster 4 con l'evidenza dei valori anomali identificati (compreso il fault artificiale) attraverso il metodo GESD. Sulla base dei risultati ottenuti si osserva che l'applicazione del metodo K-Means per la ripartizione dei dati preventiva all'analisi di fault detection non si è rivelata una metodologia robusta rispetto all'identificazione dei fault artificiali creati.

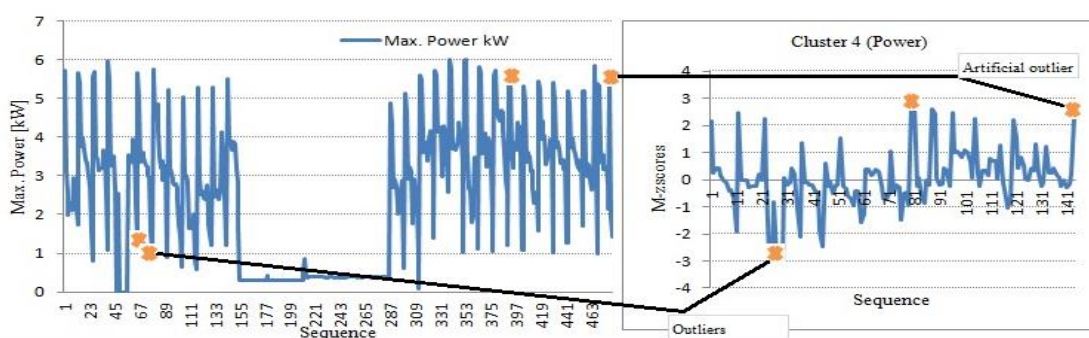


Figura 53: Andamento della potenza elettrica attiva per illuminazione nel tempo con dati non classificati e variazione dell'indice Mzscore in un cluster di dati (classificazione con metodo K-Means) con evidenza dei principali outlier individuati con GESD

Applicazione del metodo DBSCAN

Come già osservato in precedenza il metodo DBSCAN richiede in ingresso la conoscenza dei parametri r e $minPts$, e in generale la sua efficacia è fortemente dipendente dalla scelta di questi parametri. Al contrario del metodo K-Means questo metodo permette di segregare tutti gli outlier in un particolare cluster che nel seguito sarà identificato come *cluster 0*. Al fine di individuare in maniera corretta i suddetti parametri è stata condotta una opportuna analisi di sensibilità attraverso una serie di simulazioni condotte per l'intero data set (considerando sia la variabile energia che potenza massima) in cui sono stati valutati diversi valori dei parametri analizzando i risultati in termini di valori anomali individuati.

Nella tabella seguente sono presentati alcuni risultati dell'analisi di sensibilità condotta relativamente alla variabile energia e considerando il solo periodo diurno. In particolare si è osservato che tenendo costante uno dei due parametri e facendo variare l'altro i risultati variano fortemente. Se il valore di $minPts$ viene tenuto costante e il valore di r decresce allora il numero di outlier individuati tende a crescere così come il numero totale di cluster individuati. I risultati ottenuti da quest'analisi sono stati attentamente analizzati e sulla base delle similarità riscontrate all'interno dei cluster sono stati individuati i valori ottimali per i due parametri di ingresso al metodo.

Tabella 12: Analisi di sensibilità condotta sui valori dei parametri r e $minPts$ in ingresso all'algoritmo DBSCAN

S. No	MinPts	Radius 'r'	No. of Clusters	No. of Outliers
1	5	1	2	1
2	5	0.9	2	1
3	5	0.8	2	4
4	5	0.7	3	10
5	5	0.6	4	23
6	5	0.5	5	52
7	7	0.7	3	16
8	9	0.7	3	29

I risultati ottenuti dimostrano che attraverso il metodo DBSCAN è possibile identificare cluster caratterizzati dalla stessa densità di oggetti e con spiccate similarità.

Nella tabella seguente (tabella 13) sono presentati i dati di energia e dei relativi attributi inclusi nel cluster 0 considerando il periodo diurno e come parametri di ingresso quelli relativi al caso numero 4 della tabella precedente (tabella 12). I valori segregati possono sicuramente considerarsi valori anomali. Per esempio con un basso di numero di occupanti e di stanza attive (casi 1, 5, 6, 10) l'energia elettrica richiesta per l'illuminazione risulta alta e all'opposto per i casi 2, 3, 4 e 9. L'algoritmo risulta particolarmente efficace per l'individuazione di fault reali ma non ha consentito l'individuazione dei due fault artificiali. Il motivo risiede probabilmente nella natura dei fault artificiali creati, che risultano fortemente legati alla variabile tempo che non è stato possibile considerare come attributo per la creazione dei cluster.

Tabella 13: Outlier presenti nel cluster 0 per il periodo diurno individuati con il metodo DSCAN per la variabile energia elettrica (edificio F68)

S. No	Energy [kWh]	People Presence	Active Room	Solar Radiation [W/m2]
1	4.3	2	2	343.6
2	1.4	13	10	190.3
3	1.8	19	14	218.83
4	2.1	15	12	445
5	2.9	6	5	476.45
6	3.3	8	6	429.32
7	3.0	13	12	413.75
8	3.1	15	13	408.08
9	1.9	10	7	330.12
10	2.4	7	6	209.6

Nella tabella seguente (tabella 14) sono presentati alcuni valori anomali comuni individuati dai tre metodi di data mining analizzati. In particolare sono stati inclusi tutti i valori anomali di potenza massima con i relativi attributi identificati da almeno due dei tre metodi utilizzati e tra questi è presente anche uno dei fault artificiali. E' evidente dall'analisi della tabella che i valori anomali individuati rappresentano fault reali del sistema per cui si è spesso riscontrato che ad alti valori di potenza corrispondono valori molto bassi del tasso di occupazione e viceversa. In generale, pertanto, è possibile osservare che i metodi proposti sono efficaci nell'individuazione di anomalie e quindi di essere adatti per l'applicazione di una diagnosi avanzata.

Tabella 14: Outlier comuni individuati da almeno due dei tre metodi analizzati per la potenza elettrica attiva per illuminazione

S.No	Day	Date	Time	Max. Power [kW]	People Presence	Active Room	Solar Radiation [W/m2]
1	Fri	07/12/2012	8	5.42	7	6	38.5
2	Tue	11/12/2012	8	5.48	10	9	0.00
3	Fri	14/12/2012	11	4.84	2	2	343.6
4	Mon	14/01/2013	7	5.98	4	4	1.83
5	Mon	14/01/2013	14	1.95	19	14	218.83
6	Tue	15/01/2013	7	5.82	5	5	3.8
7	Fri	25/01/2013	17	5.55	5	4	1.75

I metodi proposti e implementati si sono rivelati adatti ai fini diagnostici anche se con differenti potenzialità limiti per l'edificio F40. In particolare il metodo CART accoppiato all'algoritmo GESD è risultato particolarmente robusto e accurato nella ricerca dei valori anomali di energia e potenza. Con il metodo K-Means alcuni dei cluster individuati sono risultati impuri e i valori anomali spesso sono disseminati tra più cluster. Il metodo DBSACN si è rivelato particolarmente adatto a raggruppare dati in cluster caratterizzati dalla stessa densità e con valori simili; inoltre esso consente di raggruppare tutti i valori anomali in un singolo cluster. Questo metodo è risultato adatto all'individuazione di valori anomali oltre a quelli artificialmente creati che si sono riscontrati soprattutto nelle prime ore del mattino e durante la notte.

In generale i metodi di clustering analizzati sono risultati poco affidabili nel rilevare fault fortemente legati alla variabile tempo. Quest'ultima è una importante osservazione che va considerata con attenzione in funzione della natura dei fault che possono riscontrarsi in un edificio.

Risultati dell'analisi diagnostica avanzata con tecniche di data mining per il cluster o distretto di edifici

In questa sezione sono presentati i risultati dell'analisi diagnostica avanzata condotta per il cluster di edifici per il quale è presente una dettagliata descrizione nella prima parte della seguente relazione. Per ogni edificio appartenente al cluster sono stati applicati i metodi di classificazione e di clustering sopra descritti.

Sulla base di un'analisi di sensibilità preliminare, è stato verificato che la potenza elettrica per ogni step temporale risulta essere la grandezza elaborata in fase di preprocessing più efficace per il rilevamento delle anomalie. Pertanto, per tutte le simulazioni condotte nel presente lavoro, è stato considerato il valore della potenza elettrica attiva come variabile dipendente.

In particolare le variabili analizzate per ogni edificio sono la potenza elettrica attiva per l'illuminazione e la potenza elettrica attiva totale. La frequenza temporale considerata è di 15 minuti. Con lo stesso time stamp sono stati analizzati per la costruzione delle classi o dei cluster anche le variabili data, giorno della settimana, ora del giorno (solo per il metodo CART), radiazione solare, numero di occupanti corrispondenti e temperatura media interna. Il data set a disposizione per tutti gli edifici del distretto di edifici è stato preventivamente oggetto di analisi dettagliata, per cui sulla base delle analisi condotte nelle attività 1, 2 e 3 del seguente documento si è provveduto a effettuare un *cleaning* dei dati e alla ricostruzione dei missing data quando questo è risultato possibile e affidabile.

In particolare i dati di potenza elettrica attiva per illuminazione e totale sono stati analizzati separatamente per il periodo invernale Gennaio – Febbraio 2013 e per il periodo primaverile/estivo Maggio- Giugno 2013.

La scelta dei periodi temporali da analizzare ha permesso quindi anche di valutare le tipologie di fault riscontrabili in periodi in cui le condizioni climatiche relative all'irradianza solare sono differenti. I periodi temporali selezionati, infatti, presentando disponibilità di luce naturale differenti consentono di verificare potenziali fault associati al comportamento degli occupanti nella gestione dell'illuminazione artificiale.

Riguardo al metodo CART per ogni simulazione effettuata è stato costruito un albero di classificazione e regressione relativamente a ogni edificio, per le due variabili considerate (potenza illuminazione e potenza totale) e per i due periodi temporali analizzati. Per ogni costruzione è stata condotta un'analisi di sensibilità sulle variabili indipendenti al fine di verificare la loro influenza sulla variazione della variabile dipendente (potenza elettrica). Tutte le simulazioni sono state condotte operando una *cross validation* sul grado di purezza di ogni nodo terminale ed inoltre è stato effettuato un opportuno *pruning* dell'albero.

Si è riscontrato che per tutti gli edifici, sia considerando la variabile potenza elettrica per illuminazione che la potenza elettrica totale, le variabili maggiormente influenzanti sono rappresentate dalla presenza degli occupanti e dalla radiazione solare globale incidente. Durante la stagione estiva anche la variabile temperatura esterna è risultata particolarmente importante per il processo di classificazione. Inoltre, il numero di classi individuate per ogni caso esaminato sono comprese tra 5 e 10.

Riguardo all'analisi di clustering effettuata (con i metodi K-Means e DBSCAN) al fine di superare i limiti degli algoritmi che non consentono l'utilizzo di variabili legate a pattern temporali, i data set sono stati preventivamente suddivisi considerando il periodo lavorativo (dalle 7:30 alle 17:30), il periodo non lavorativo e i fine settimana. L'approccio adottato per lo *splitting* dei dati deriva anche dall'esperienza maturata sull'edificio F40 per il quale la suddivisione del data set in ore diurne, ore notturne e fine settimana non si è rivelata particolarmente efficace per la natura dei fault presenti nella tipologia di edifici oggetto di indagine. Infatti molti falsi positivi individuati per l'edificio F40 con i metodi di clustering sono legati alla variabile tempo per cui la suddivisione in ore diurne e notturne non ha permesso di valutare con efficacia i valori anomali presenti nelle prime ore del mattino e alla fine dell'orario di lavoro.

L'analisi di fault detection è stata effettuata applicando il metodo GESD a ogni classe individuata con il metodo CART e a ogni cluster trovato con il metodo K-Means. Con il metodo DBSCAN, invece, i valori anomali sono stati individuati analizzando direttamente il cluster 0.

Nelle figure e tabelle seguenti sono presentati i risultati ottenuti dall'analisi di fault detection condotta per ogni edificio separatamente, per le due variabili considerate, e per i due periodi temporali analizzati.

Riguardo all'analisi CART nelle figure seguenti è riportata la rappresentazione grafica dell'indice diagnostico Mzscore per una particolare classe selezionata contenente i principali fault identificati, e la variazione della variabile dipendente per un determinato periodo temporale con dati non classificati.

L'andamento temporale per la potenza dei dati non classificati con l'evidenza degli outlier individuati attraverso l'analisi effettuata classificando preventivamente i dati consente di evidenziare come alcune anomalie non possono essere rilevate senza un processo di classificazione. È evidente, infatti, come nella serie di dati le anomalie non risultano evidenti né rilevabili se i dati non sono preventivamente suddivisi in classi omogenee.

Per ogni valore anomalo individuato con il metodo GESD è riportata l'ora del giorno, il giorno della settimana e la relativa data. È stato verificato che per tutte le simulazioni effettuate l'indice Mzscore presenta valori estremi in valore assoluto in corrispondenza delle anomalie rilevate dal metodo GESD, per cui esso risulta un ottimo indicatore diagnostico per l'analisi di fault detection.

Inoltre per ogni edificio esaminato (relativamente alla potenza elettrica attiva per l'illuminazione) e per la classe selezionata con il metodo CART, nei grafici seguenti sono evidenziati anche i valori anomali individuati che risultano comuni ai metodi di clustering applicati. Nei grafici i valori anomali comuni evidenziati con colore differente dal giallo sono inclusi nella classe selezionata ma non appartengono al periodo temporale rappresentato con dati non classificati.

Questa tipologia di analisi risulta particolarmente importante poiché ha consentito di verificare la correttezza nella classificazione dei dati sulla base dei valori anomali identificati nella classe selezionata e individuati anche attraverso il processo di clustering operato con i metodi K-Means e DBSCAN.

In generale i valori anomali sono stati identificati in tre differenti periodi della giornata. Il primo periodo in cui sono stati riscontrati valori anomali è relativo al mattino presto tra le 6:30 e le 7:30. Questa tipologia di fault è risultata spesso comune nella sua identificazione ai tre metodi di data mining analizzati.

Al mattino presto infatti sia relativamente ai dati di potenza elettrica per l'illuminazione che totale sono presenti dei picchi in corrispondenza di una presenza di occupanti molto bassa.

Il secondo periodo temporale in cui sono stati identificati molti valori anomali è compreso tra le 12:30 e le 13:30 e cioè durante la pausa pranzo. Il terzo periodo è relativo alla fine dell'orario di lavoro tra le 17:00 e le 17:30 dove si è osservato che a una diminuzione del numero di occupanti dell'edificio non corrisponde una diminuzione della potenza elettrica assorbita per l'illuminazione, e della potenza elettrica totale la quale tiene conto anche delle FEM e del condizionamento ambientale.

Nelle pagine seguenti sono riportate anche delle tabelle in cui sono esaminati in dettaglio i valori anomali identificati con i metodi K-Means a DBSCAN e sono evidenziati quelli comuni sia per la potenza elettrica attiva per illuminazione che per la potenza elettrica attiva totale considerando differenti periodi della giornata.

Si osserva che per gli edifici F69 e F72 (per il periodo Gennaio-Febbraio) e per gli edifici F68, F71, F72 e F73 (per il periodo Maggio – Giugno) i valori anomali comuni ai due metodi di clustering sono stati individuati o solo per il vettore potenza elettrica per illuminazione, o solo per il vettore potenza elettrica totale. Ulteriore analisi condotta e riportata nelle tabelle seguenti riguarda l'individuazione dei valori anomali riscontrati per il vettore potenza elettrica totale e comuni ai tre metodi di data mining analizzati. Dall'analisi delle tabelle si osserva che alcuni valori anomali comuni nel periodo Gennaio-Febbraio sono stati individuati durante la notte. Gli altri fault individuati e comuni ai tre metodi per la potenza elettrica totale sono stati trovati al mattino presto, a metà mattinata e durante la pausa pranzo.

Infine dall'analisi di fault detection effettuata sono stati individuati anche fault che probabilmente sono imputabili a una misura non corretta operata attraverso i sistemi di monitoraggio. Anche da questo punto di vista la metodologia introdotta è risultata robusta rispetto all'individuazione di anomalie relative alla strumentazione adottata per rilievo dei dati.

F66 – Jan-Feb

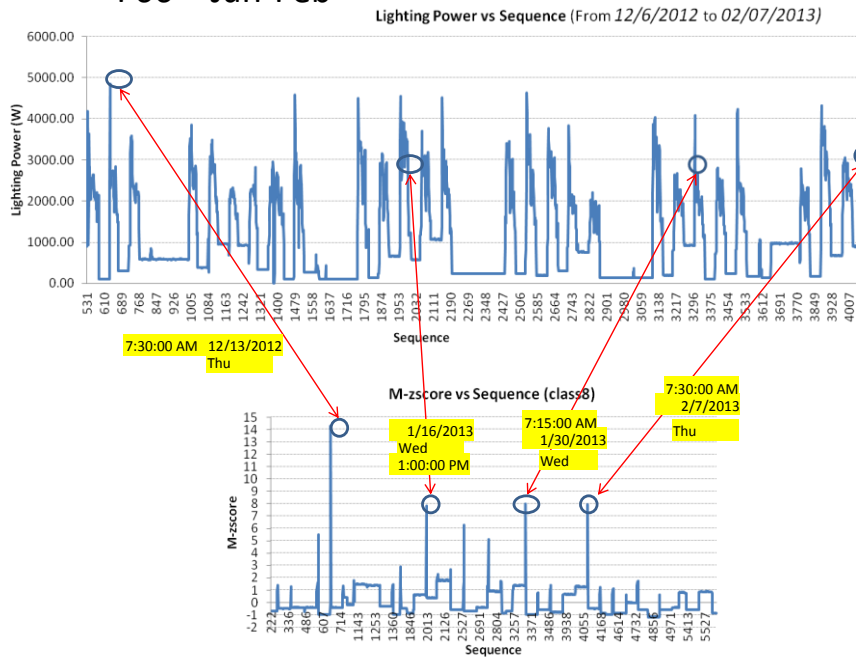


Figura 54: Andamento della potenza elettrica attiva per illuminazione nel tempo con dati non classificati e variazione dell'indice Mzscore in un classe di dati (classificazione con metodo CART) con evidenza dei principali outlier individuati con ESD (edificio F66)

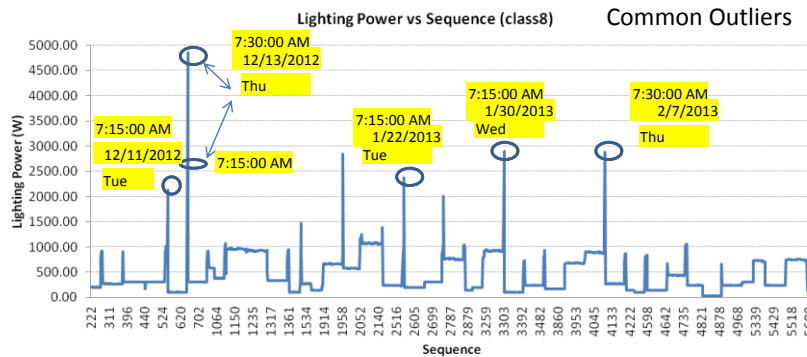


Figura 55: Andamento della potenza elettrica attiva per illuminazione per una classe di dati (classificazione con metodo CART) con evidenza dei principali outlier comuni individuati dai tre metodi analizzati (CART – K-Means – DBSCAN) (edificio F66)

F68-Jan-Feb

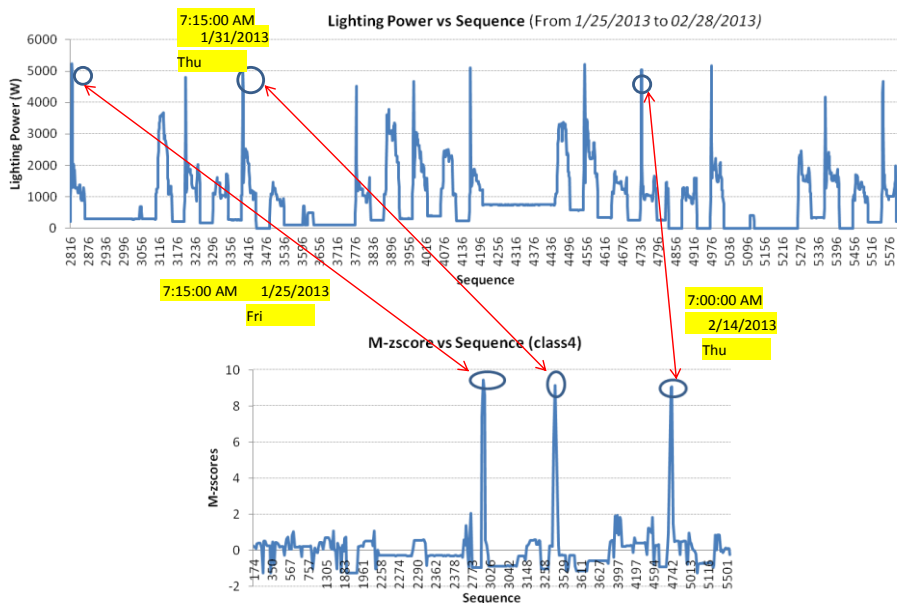


Figura 56: Andamento della potenza elettrica attiva per illuminazione nel tempo con dati non classificati e variazione dell'indice Mzscore in un classe di dati (classificazione con metodo CART) con evidenza dei principali outlier individuati con ESD (edificio F68)

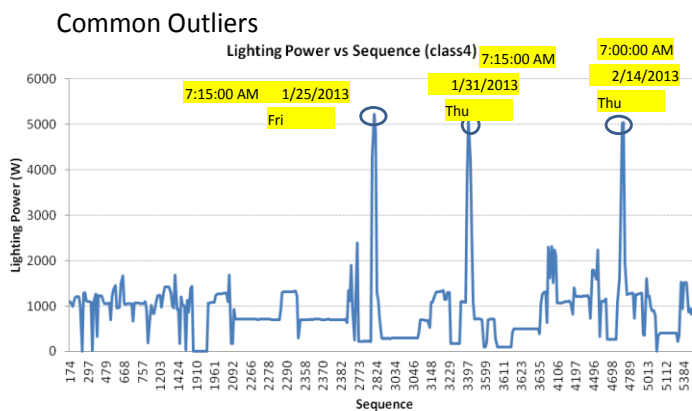


Figura 57: Andamento della potenza elettrica attiva per illuminazione per una classe di dati (classificazione con metodo CART) con evidenza dei principali outlier comuni individuati dai tre metodi analizzati (CART – K-Means – DBSCAN) (edificio F68)

F69-Jan-Feb

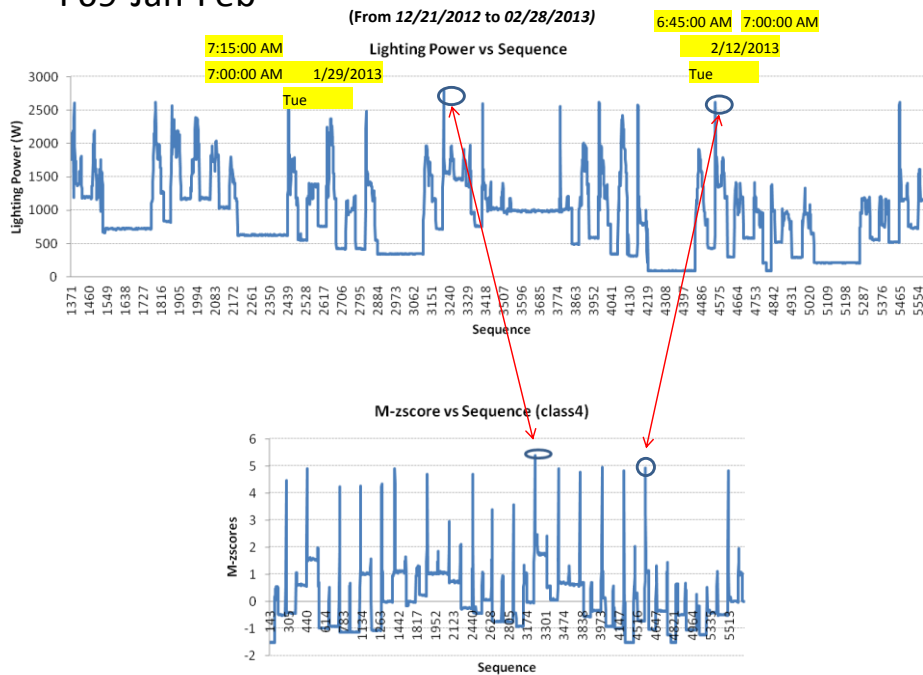


Figura 58: Andamento della potenza elettrica attiva per illuminazione nel tempo con dati non classificati e variazione dell'indice Mzscore in un classe di dati (classificazione con metodo CART) con evidenza dei principali outlier individuati con ESD (edificio F68)

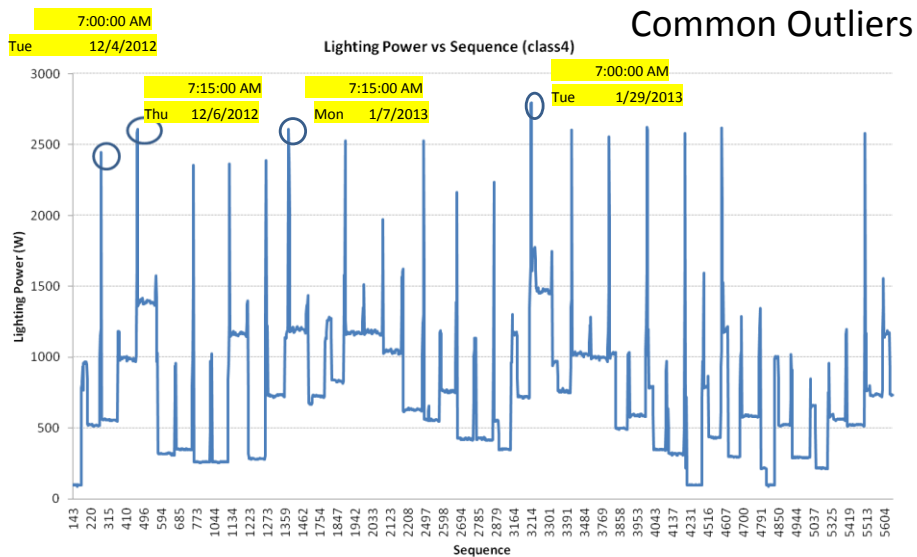


Figura 59: Andamento della potenza elettrica attiva per illuminazione per una classe di dati (classificazione con metodo CART) con evidenza dei principali outlier comuni individuati dai tre metodi analizzati (CART – K-Means – DBSCAN) (edificio F68)

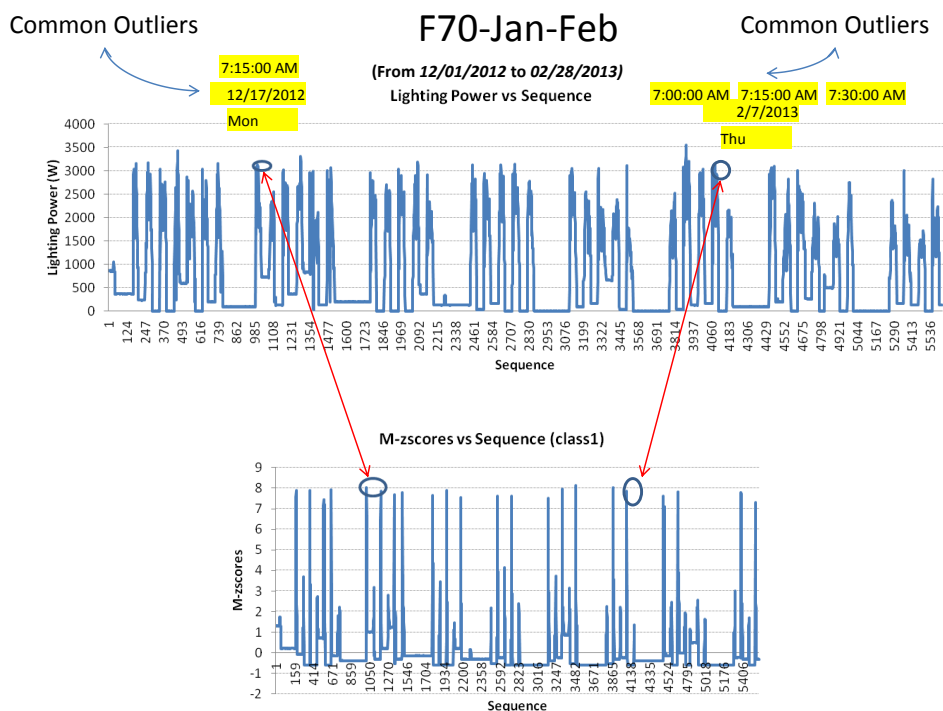


Figura 60 : Andamento della potenza elettrica attiva per illuminazione nel tempo con dati non classificati e variazione dell'indice Mzscore in un classe di dati (classificazione con metodo CART) con evidenza dei principali outlier comuni individuati dai tre metodi analizzati (CART(ESD) – K-Means (ESD) – DBSCAN) (edificio F70)

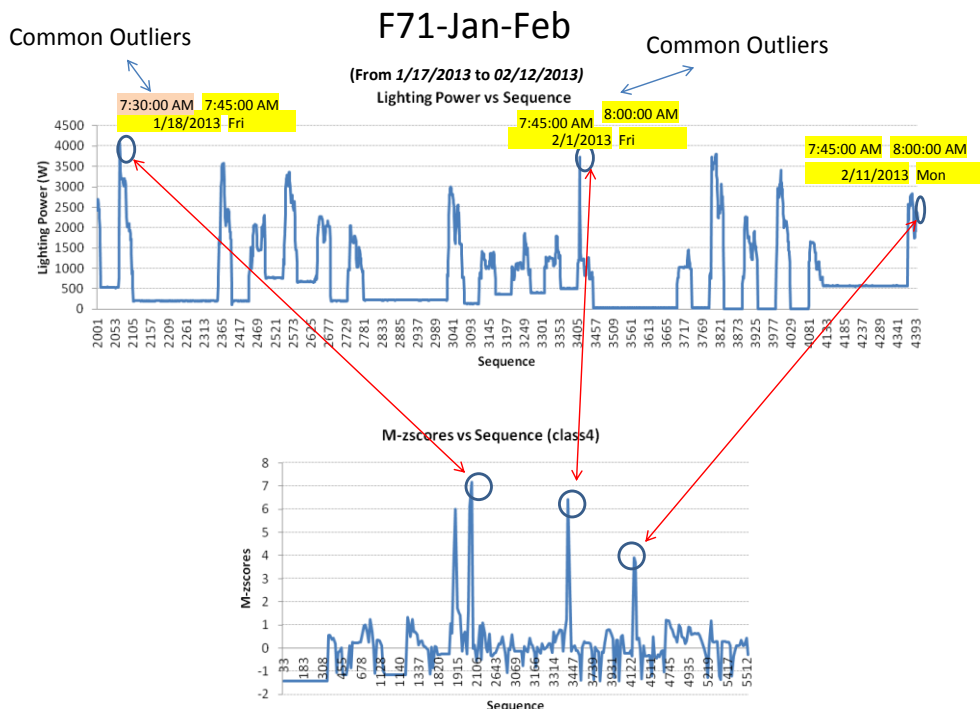


Figura 61: Andamento della potenza elettrica attiva per illuminazione nel tempo con dati non classificati e variazione dell'indice Mzscore in un classe di dati (classificazione con metodo CART) con evidenza dei principali outlier comuni individuati dai tre metodi analizzati (CART(ESD) – K-Means (ESD) – DBSCAN) (edificio F71)

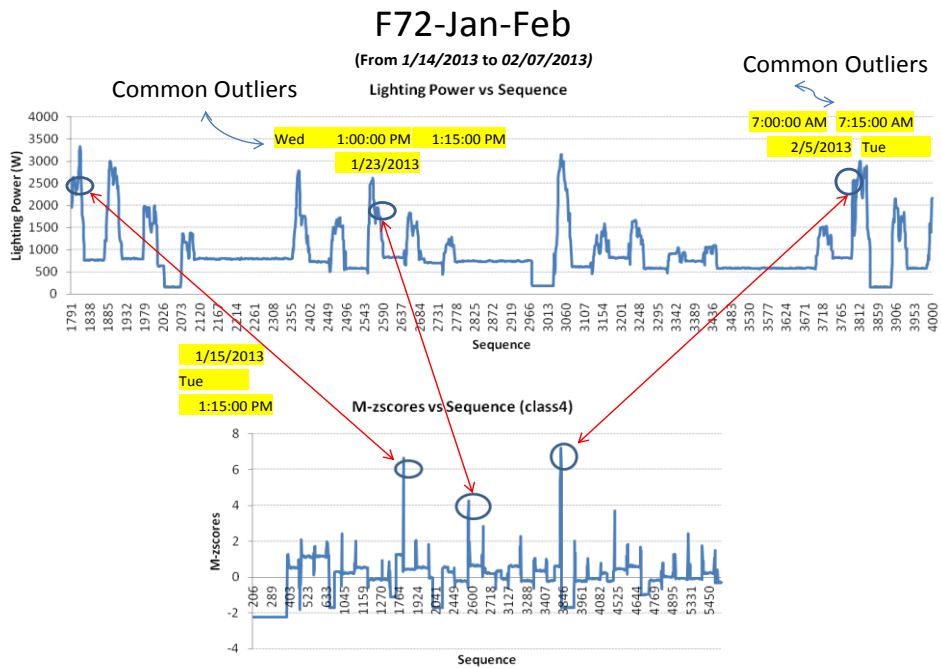


Figura 62: Andamento della potenza elettrica attiva per illuminazione nel tempo con dati non classificati e variazione dell'indice Mzscore in un classe di dati (classificazione con metodo CART) con evidenza dei principali outlier comuni individuati dai tre metodi analizzati (CART(ESD) – K-Means (ESD) – DBSCAN) (edificio F72)

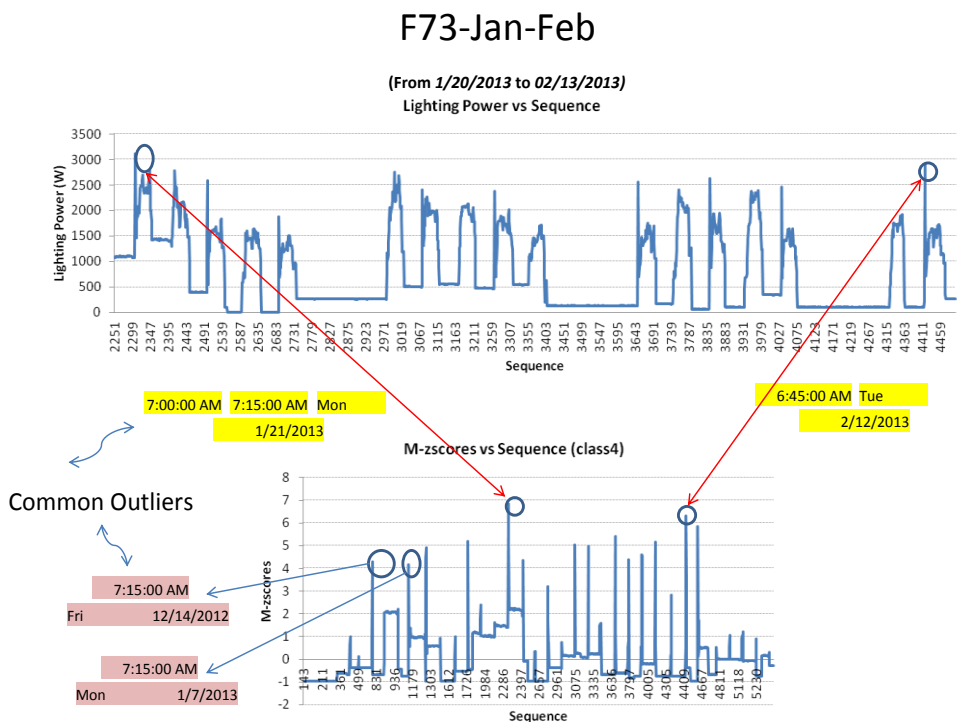


Figura 63: Andamento della potenza elettrica attiva per illuminazione nel tempo con dati non classificati e variazione dell'indice Mzscore in un classe di dati (classificazione con metodo CART) con evidenza dei principali outlier comuni individuati dai tre metodi analizzati (CART(ESD) – K-Means (ESD) – DBSCAN) (edificio F73)

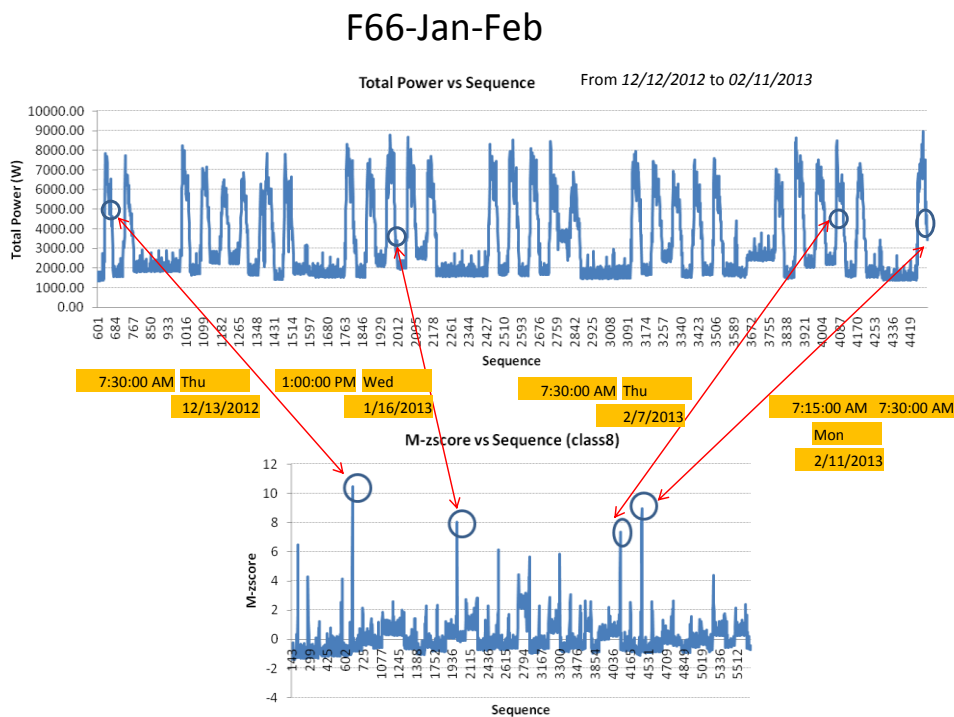


Figura 64: Andamento della potenza elettrica attiva totale nel tempo con dati non classificati e variazione dell'indice Mzscore in un classe di dati (classificazione con metodo CART) con evidenza dei principali outlier individuati con ESD (edificio F66)

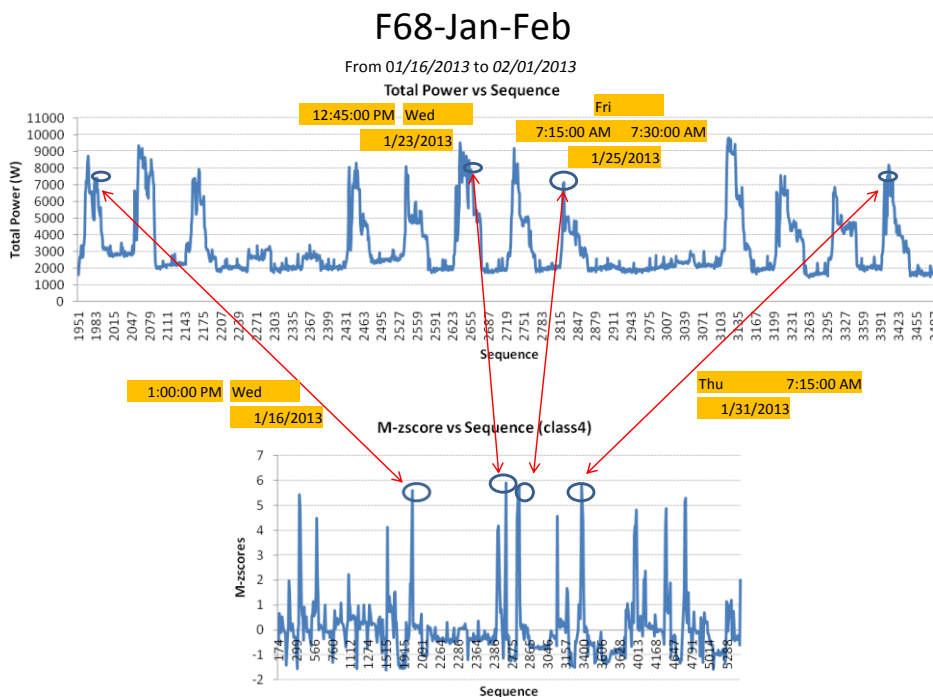


Figura 65: Andamento della potenza elettrica attiva totale nel tempo con dati non classificati e variazione dell'indice Mzscore in un classe di dati (classificazione con metodo CART) con evidenza dei principali outlier individuati con ESD (edificio F68)

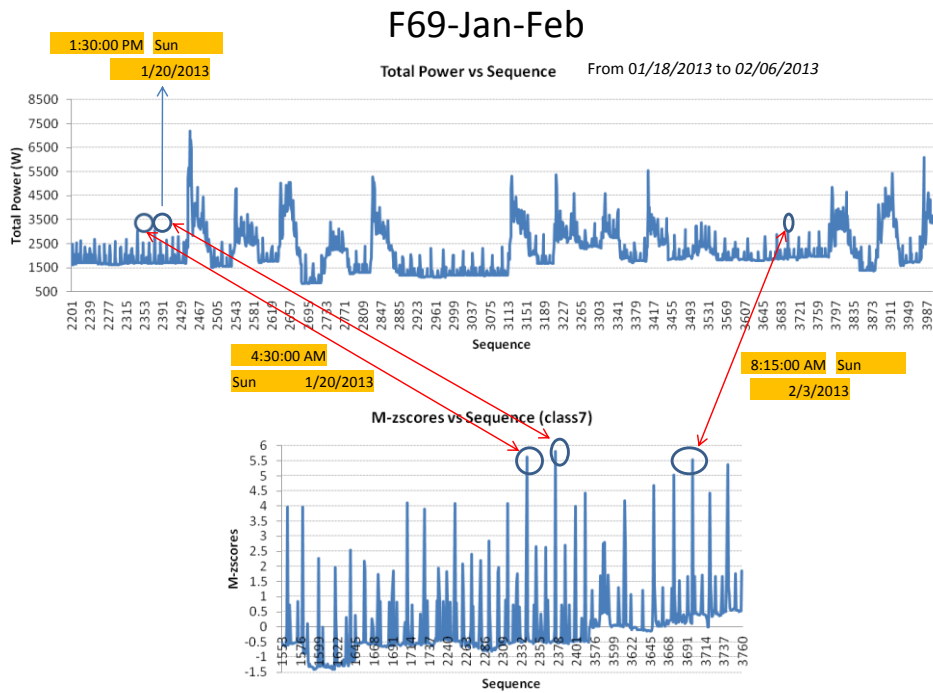


Figura 66: Andamento della potenza elettrica attiva totale nel tempo con dati non classificati e variazione dell'indice Mzscore in un classe di dati (classificazione con metodo CART) con evidenza dei principali outlier individuati con ESD (edificio F69)

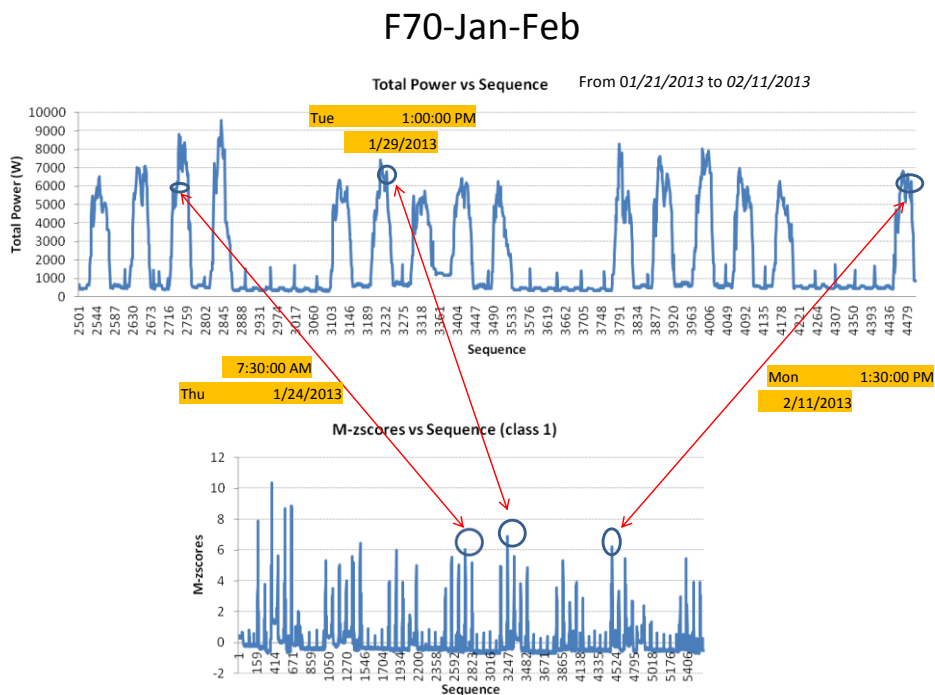


Figura 67: Andamento della potenza elettrica attiva totale nel tempo con dati non classificati e variazione dell'indice Mzscore in un classe di dati (classificazione con metodo CART) con evidenza dei principali outlier individuati con ESD (edificio F70)

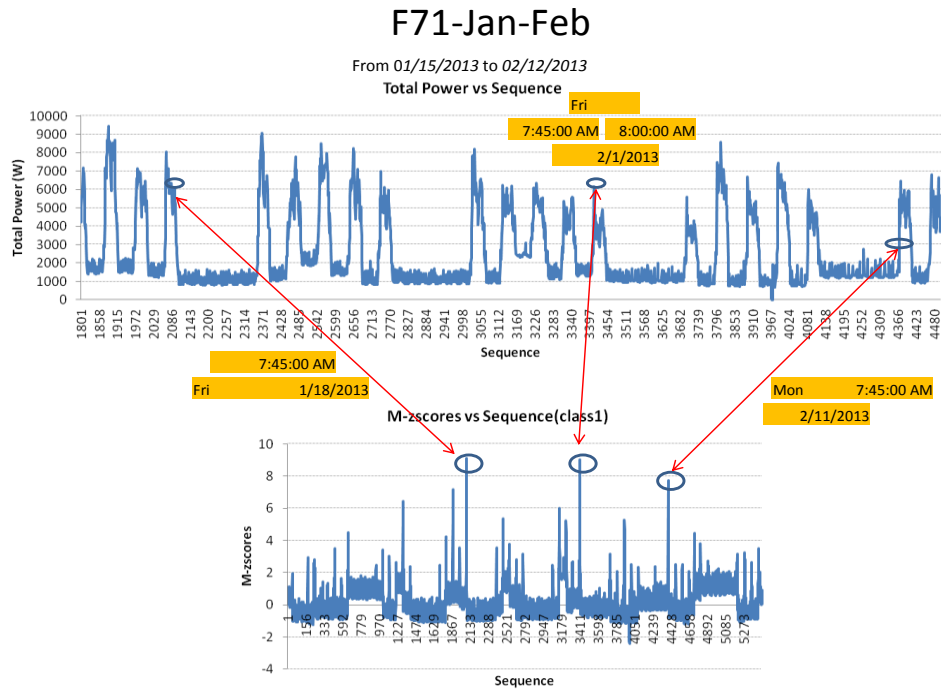


Figura 68: Andamento della potenza elettrica attiva totale nel tempo con dati non classificati e variazione dell'indice Mzscore in un classe di dati (classificazione con metodo CART) con evidenza dei principali outlier individuati con ESD (edificio F71)

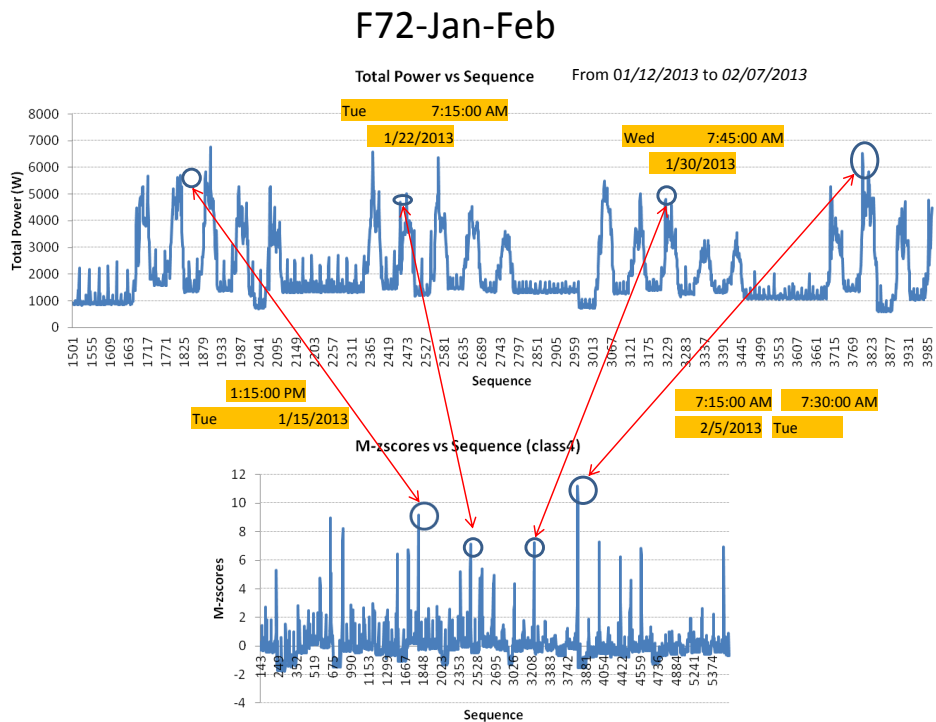


Figura 69: Andamento della potenza elettrica attiva totale nel tempo con dati non classificati e variazione dell'indice Mzscore in un classe di dati (classificazione con metodo CART) con evidenza dei principali outlier individuati con ESD (edificio F72)

F73-Jan-Feb

From 01/21/2013 to 02/13/2013

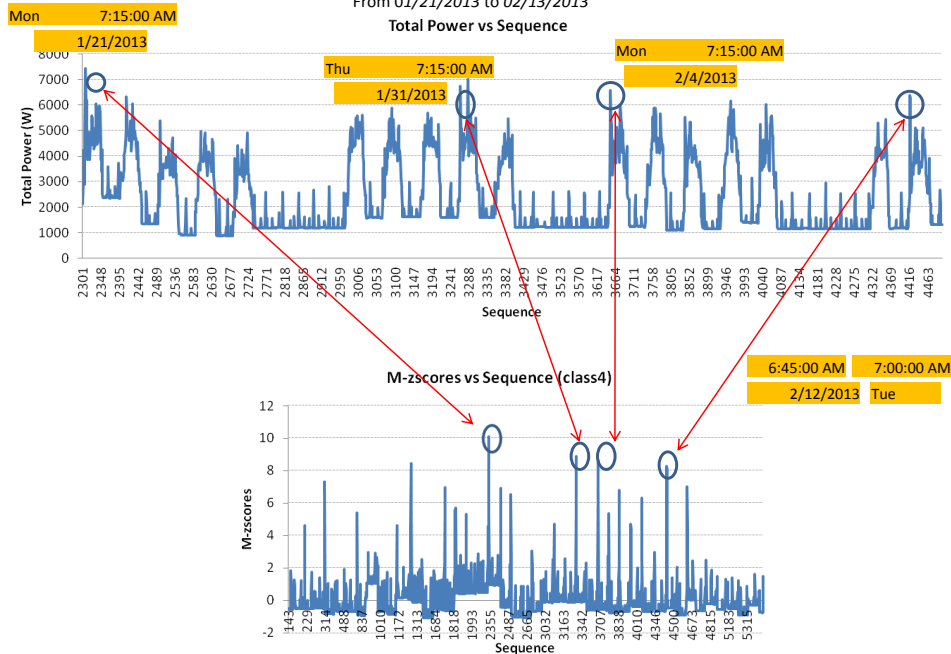


Figura 70: Andamento della potenza elettrica attiva totale nel tempo con dati non classificati e variazione dell'indice Mzscore in un classe di dati (classificazione con metodo CART) con evidenza dei principali outlier individuati con ESD (edificio F73)

Tabella 15: Outlier individuati con il metodo DSCAN e comuni con il metodo K-Means (GESD) per la potenza elettrica attiva per illuminazione e potenza elettrica attiva totale (edificio F66)

S.No	DATE	Id DAY	HOUR	Electric active Power for lighting F66	Total electric active power F66	Indoor temperature F66	People presence F66	Outdoor temperature	Global solar radiation	
1	05/12/2012	4	7.30.00	2430.00	6311.39	16.15	1	6.2	141	
2	06/12/2012	5	6.45.00	300.00	2535.39	18.69	0	4.6	339	DBSCAN-Lighting Power
3	06/12/2012	5	7.00.00	175.00	1958.19	18.83	0	5.2	381	DBSCAN-Total Power
4	06/12/2012	5	7.15.00	135.00	2001.40	18.9	1	5.6	292	Common-DBSCAN&K-Means
5	06/12/2012	5	7.30.00	450.00	2616.60	18.95	1	5.8	360	
6	07/12/2012	6	7.30.00	2615.00	5211.00	16.35	1	3.7	107	
7	13/12/2012	5	7.15.00	2515.00	4729.20	17.69	0	4.9	311	
8	13/12/2012	5	7.30.00	4855.20	7655.79	16.47	0	5.2	369	
9	14/12/2012	6	7.30.00	920.00	3181.10	19.51	1	9.7	512	
10	17/12/2012	2	7.15.00	585.00	2807.89	18.83	1	12.7	67	
11	17/12/2012	2	7.30.00	970.00	3406.19	19.02	1	13.1	466	
12	20/12/2012	5	7.30.00	845.00	3322.19	16.99	1	8	248	
13	08/01/2013	3	7.00.00	1470.00	3450.00	16	0	10	185	
14	08/01/2013	3	7.15.00	2640.00	4722.70	15.76	1	9.9	100	
15	08/01/2013	3	7.30.00	3065.00	5405.79	15.69	1	9.8	111	
16	16/01/2013	4	7.30.00	3510.00	6110.60	16.6	2	3.7	0	
17	24/01/2013	5	7.30.00	2655.00	6253.29	16.02	2	5.5	0	
18	30/01/2013	4	7.30.00	3590.00	6100.20	16.26	2	5.2	7	
19	07/02/2013	5	7.30.00	2885.00	5980.00	18.93	0	3.7	0	
20	11/02/2013	2	7.30.00	4445.20	6852.39	15.09	0	3.2	0	
21	13/02/2013	4	7.30.00	3185.00	7012.70	16.72	3	3.8	11	

Tabella 16: Outlier individuati con il metodo DSCAN e comuni con il metodo K-Means (GESD) per la potenza elettrica attiva per illuminazione e potenza elettrica attiva totale (edificio F68)

S.No	DATE	Id DAY	HOUR	Electric active Power for lighting F68	Total electric active power F68	Indoor temperature F68	People presence F68	Outdoor temperature	Global solar radiation	
1	04/12/2012	3	6.45.00	3840	5462.7	22.41	0	5.6	107	
2	04/12/2012	3	7.00.00	4685.2	6237.5	22.67	0	5.7	88	DBSCAN-Lighting Power
3	06/12/2012	5	6.30.00	1425	3440.1	19.39	0	4.1	241	DBSCAN-Total Power
4	06/12/2012	5	6.45.00	3695	6013	20.33	0	4.6	339	Common-DBSCAN&K-Means
5	06/12/2012	5	7.00.00	5025.2	7362.2	20.52	0	5.2	381	
6	06/12/2012	5	7.15.00	3010	5282.7	20.78	0	5.6	292	
7	13/12/2012	5	6.45.00	3695	6013	20.33	0	4.4	227	
8	13/12/2012	5	7.00.00	5025.2	7362.2	20.52	0	4.7	186	
9	13/12/2012	5	7.15.00	3010	5282.7	20.78	0	4.9	311	
10	14/12/2012	6	7.00.00	4015.19	5713.7	21	0	8.8	116	
11	14/12/2012	6	7.15.00	2730	4517.7	21.37	0	9.2	302	
12	14/12/2012	6	7.30.00	1405	3637.1	21.51	0	9.7	512	
13	17/12/2012	2	7.00.00	95	1764.59	21.31	0	11.3	246	
14	17/12/2012	2	7.30.00	215	1978.3	21.91	0	13.1	466	
15	18/12/2012	3	7.30.00	835	2347.3	24.41	1	9.6	343	
16	07/01/2013	2	7.00.00	3750	5476.79	18.92	0	10.8	336	
17	07/01/2013	2	7.15.00	5250	6762.2	19.46	0	11.2	346	
18	07/01/2013	2	7.30.00	4995.2	7264.39	19.28	0	11.7	358	
19	08/01/2013	3	7.00.00	1065	2837.69	20.26	0	10	185	
20	23/01/2013	4	17.45.00	2640	5092.5	27.28	5	7.2	0	
21	31/01/2013	5	7.00.00	3690	5590.2	22.99	1	8.7	0	
22	06/02/2013	4	7.30.00	2295	4137.1	21.66	1	9.8	0	
23	13/02/2013	4	17.45.00	1445	3499.69	31.05	5	7.2	0	
24	28/02/2013	5	18.45.00	215	1704.3	30.77	4	8.9	0	

Tabella 17: Outlier individuati solo con il metodo DSCAN e comuni con il metodo K-Means (GESD) per la potenza elettrica attiva per illuminazione e potenza elettrica attiva totale (edificio F69)

S.No	DATE	Id DAY	HOUR	Electric active Power for lighting F69	Indoor temperature F69	People presence F69	Outdoor temperature	Global solar radiation	
1	06/12/2012	5	7.45.00	1810	22.62	1	6.1	394	
2	17/12/2012	2	7.45.00	1575	20.59	2	14.1	137	DBSCAN-Lighting Power
3	17/12/2012	2	8.00.00	1585	20.79	2	13.8	214	
4	17/12/2012	2	8.15.00	1580	21.1	4	14.1	420	
5	07/01/2013	2	7.45.00	2010	18.4	0	12.2	369	
6	07/01/2013	2	12.45.00	1390	21.62	4	14.6	19	
7	07/01/2013	2	13.00.00	1395	21.27	2	13.5	7	
8	14/01/2013	2	12.30.00	1990	23.85	4	11.9	345	
9	14/01/2013	2	13.00.00	1820	23.87	1	11.9	448	
10	14/01/2013	2	13.15.00	1810	23.87	2	11.9	448	
11	29/01/2013	3	17.30.00	1740	23.46	0	9.4	0	
12	04/02/2013	2	9.30.00	990	20.54	9	6.1	28	
13	04/02/2013	2	9.45.00	1010	20.57	9	6.4	30	
14	04/02/2013	2	10.00.00	1010	20.5	9	6.9	39	
15	04/02/2013	2	13.15.00	1015	20.68	1	12.1	471	
16	07/02/2013	5	13.45.00	1940	23.63	8	6.3	499	
17	14/02/2013	5	17.30.00	210	25.36	0	9.6	9	
18	28/02/2013	5	12.45.00	1165	25.92	2	13.4	626	
19	28/02/2013	5	13.00.00	1165	25.92	1	13.3	628	
20	28/02/2013	5	13.15.00	1170	25.79	2	13.6	603	
21	28/02/2013	5	13.30.00	1165	25.62	1	13.4	571	

Tabella 18: Outlier individuati con il metodo DSCAN e comuni con il metodo K-Means (GESD) per la potenza elettrica attiva per illuminazione e potenza elettrica attiva totale (edificio F70)

S.No	DATE	Id DAY	HOUR	Electric active Power for lighting F70	Total electric active power F70	Indoor temperature F70	People presence F70	Outdoor temperature	Global solar radiation	
1	03/12/2012	2	7.30.00	1815	4678.5	16.9	1	7.1	0	
2	05/12/2012	4	18.00.00	795	3195.19	22.06	3	1.9	0	DBSCAN-Lighting Power
3	06/12/2012	5	6.30.00	0	2961.89	18.29	0	4.1	241	DBSCAN-Total Power
4	06/12/2012	5	7.00.00	0	3367.39	18.5	0	5.2	381	Common-DBSCAN&K-Means
5	06/12/2012	5	7.30.00	1030	5007	19.17	1	5.8	360	
6	11/12/2012	3	6.45.00	1600	3199.6	15.8	0	7	330	
7	11/12/2012	3	7.00.00	2790	8029.79	16.36	0	7.2	343	
8	11/12/2012	3	7.15.00	1760	5702.2	16.86	0	7.7	364	
9	11/12/2012	3	7.30.00	1195	4021.69	17.06	0	7.9	373	
10	12/12/2012	4	17.45.00	945	2713.69	23.1	4	3.2	0	
11	13/12/2012	5	7.00.00	2045	6369.79	17.33	0	4.7	186	
12	13/12/2012	5	7.15.00	3045	8175.1	18.13	1	4.9	311	
13	13/12/2012	5	7.30.00	2875	8079.2	18.51	1	5.2	369	
14	14/12/2012	6	7.30.00	860	2639.8	18.12	0	9.7	512	
15	17/12/2012	2	7.00.00	2360	3927.19	16.68	0	11.3	246	
16	17/12/2012	2	7.15.00	3080	4784.2	17.28	0	12.7	67	
17	17/12/2012	2	7.30.00	2545	5271.5	17.48	1	13.1	466	
18	19/12/2012	4	7.00.00	2535	4377.79	18.07	0	9.1	329	
19	19/12/2012	4	7.15.00	2960	4893.5	18.68	0	9.6	350	
20	19/12/2012	4	7.30.00	3025	5057.5	19.07	1	9.7	357	
21	20/12/2012	5	7.30.00	1175	4235.89	19.37	1	8	248	
22	20/12/2012	5	17.45.00	1880	4456	23.11	5	7.1	0	
23	20/12/2012	5	18.00.00	1385	2250.19	22.82	4	7.1	0	
24	08/01/2013	3	6.45.00	1405	2848.19	15.95	0	9.8	221	
25	08/01/2013	3	7.00.00	2830	4907.39	16.56	0	10	185	
26	08/01/2013	3	7.15.00	3005	4977.39	17.25	0	9.9	100	
27	08/01/2013	3	7.30.00	2585	6218.39	17.81	0	9.8	111	
28	05/02/2013	3	17.45.00	1140	2676.39	22.86	4	10.2	0	
29	06/02/2013	4	17.45.00	1330	2836	22.42	5	8.7	0	
30	07/02/2013	5	7.00.00	2740	4374.5	17.95	1	3.8	0	
31	07/02/2013	5	7.15.00	3025	4622.5	18.43	1	3.8	0	
32	07/02/2013	5	7.30.00	2400	3980.1	18.67	1	3.7	0	
33	12/02/2013	3	17.45.00	1710	3524.8	23.55	4	9.6	0	
34	12/02/2013	3	18.00.00	1450	2304.6	22.85	4	9.1	0	
35	14/02/2013	5	17.45.00	1730	3576.5	22.62	5	9.1	0	

Tabella 19: Outlier individuati solo con il metodo DSCAN e comuni con il metodo K-Means (GESD) per la potenza elettrica attiva per illuminazione e potenza elettrica attiva totale (edificio F71)

S.No	DATE	Id DAY	HOUR	Electric active Power for lighting F71	Total electric active power F71	Indoor temperature F71	People presence F71	Outdoor temperature	Global solar radiation	
1	03/12/2012	2	7.30.00	1.18	2259.8	19.71	3	7.1	0	DBSCAN-Lighting Power
2	03/12/2012	2	17.45.00	1.63	3015.1	22.87	4	2.6	0	DBSCAN-Total Power
3	03/12/2012	2	18.15.00	0.55	1430.19	22.62	2	3.1	0	Common-DBSCAN&K-Means
4	05/12/2012	4	7.30.00	0.85	1863.4	17.67	1	6.2	141	
5	05/12/2012	4	17.45.00	1.26	2302.5	22.81	3	2.1	0	
6	06/12/2012	5	6.45.00	0.29	1317.3	19.47	0	4.6	339	
7	06/12/2012	5	7.15.00	0.43	2101	19.71	1	5.6	292	
8	06/12/2012	5	7.30.00	0.84	1839.4	19.9	2	5.8	360	
9	07/12/2012	6	7.30.00	790	2381.39	19.22	2	3.7	107	
10	11/12/2012	3	7.30.00	870	3424.89	18.58	2	7.9	373	
11	12/12/2012	4	17.45.00	160	1156.5	22.37	2	3.2	0	
12	13/12/2012	5	7.15.00	160	1409.59	18.24	1	4.9	311	
13	13/12/2012	5	7.30.00	165	1456.59	18.21	1	5.2	369	
14	14/12/2012	6	7.30.00	1020	3765.19	19.6	2	9.7	512	
15	17/12/2012	2	7.30.00	1260	2833.1	20.34	2	13.1	466	
16	18/12/2012	3	7.30.00	1015	2196.1	20.3	2	9.6	343	
17	19/12/2012	4	7.15.00	1075	3156.39	19.48	1	9.6	350	
18	19/12/2012	4	7.30.00	1075	2444.5	16.59	2	9.7	357	
19	20/12/2012	5	7.15.00	1660	2929.69	20.34	1	7.8	272	
20	20/12/2012	5	7.30.00	1490	2729	20	1	8	248	
21	07/01/2013	2	6.30.00	200	1145.59	16.34	0	9.9	288	
22	07/01/2013	2	6.45.00	205	1361.5	16.39	0	10.3	313	
23	07/01/2013	2	7.00.00	500	1865	16.53	0	10.8	336	
24	07/01/2013	2	7.15.00	1280	2375.19	16.61	2	11.2	346	
25	07/01/2013	2	7.30.00	1290	2462.3	16.79	2	11.7	358	
26	08/01/2013	3	7.15.00	1040	2774.69	18.11	1	9.9	100	
27	08/01/2013	3	7.30.00	1055	2678.8	17.05	1	9.8	111	
28	14/01/2013	2	17.45.00	2065	3939.1	22.53	4	9.1	0	
29	14/01/2013	2	18.00.00	2025	4209.79	22.14	3	8.9	0	
30	14/01/2013	2	18.15.00	1315	2625.3	22.16	3	8.9	0	
31	18/01/2013	6	7.30.00	3595	5069	19.3	1	4.5	0	
32	22/01/2013	3	17.45.00	1580	3190.8	21.75	4	8.3	0	
33	24/01/2013	5	17.45.00	335	2087.69	23.69	3	8.6	0	
34	29/01/2013	3	7.30.00	805	3526.6	16.98	2	5.8	5	
35	30/01/2013	4	17.45.00	1415	3251.1	22.04	4	10.6	0	
36	30/01/2013	4	18.00.00	1065	2691.69	22.18	3	10.1	0	
37	21/02/2013	5	7.30.00	1465	2822.89	15.88	1	3.4	21	

Tabella 20: Outlier individuati con il metodo DSCAN e comuni con il metodo K-Means (GESD) per la potenza elettrica attiva per illuminazione e potenza elettrica attiva totale (edificio F72)

S.No	DATE	Id DAY	HOUR	Total electric active power F72	Indoor temperature F72	People presence F72	Outdoor temperature	Global solar radiation	
1	01/12/2012	7	4.15.00	2614.5	23.21	0	7.2	0	DBSCAN-Lighting Power
2	01/12/2012	7	18.45.00	2256.6	25.21	0	9.3	0	DBSCAN-Total Power
3	01/12/2012	7	19.00.00	2574.1	25.09	0	9.2	0	Common-DBSCAN&K-Means
4	02/12/2012	1	1.30.00	2574.5	23.57	0	8.6	0	
5	02/12/2012	1	7.45.00	2517	22.75	0	8.8	0	
6	02/12/2012	1	20.00.00	2588.5	21.9	0	6.2	0	
7	16/12/2012	1	8.30.00	1098.8	19.97	0	14.3	288	
8	12/01/2013	7	13.00.00	1407.4	23.36	0	11.8	371	
9	12/01/2013	7	13.45.00	1627.09	23.58	0	11.6	121	
10	12/01/2013	7	14.00.00	2222.39	23.62	0	11.6	121	
11	13/01/2013	1	10.45.00	2316.1	20.34	0	9.2	134	
12	19/01/2013	7	2.30.00	2529.89	20.09	0	0.4	0	
13	19/01/2013	7	2.45.00	2403.5	20.06	0	0.2	0	
14	19/01/2013	7	8.15.00	2606.89	19.83	0	2.6	12	
15	19/01/2013	7	8.30.00	2005.09	19.91	0	2.6	9	
16	19/01/2013	7	14.45.00	2699.1	21.31	0	4.5	51	
17	20/01/2013	1	3.15.00	2628.19	20.04	0	9.2	0	
18	26/01/2013	7	13.00.00	1738.09	21.74	1	7.9	443	
19	27/01/2013	1	7.45.00	1734.4	19.34	1	1.5	11	
20	27/01/2013	1	13.30.00	1679.8	19.47	1	10.1	454	
21	27/01/2013	1	16.30.00	1743.8	19.37	1	9.2	40	
22	27/01/2013	1	16.45.00	1347.8	19.3	1	8.5	26	
23	02/02/2013	7	1.00.00	2105.39	22.8	0	11.2	0	
24	02/02/2013	7	11.15.00	1533.9	23.2	0	13.1	123	
25	02/02/2013	7	12.15.00	1081.8	23.39	0	14.4	411	
26	02/02/2013	7	15.00.00	2031	23.9	0	12.5	12	
27	03/02/2013	1	11.00.00	1501.59	21.11	0	9.2	415	
28	03/02/2013	1	13.15.00	1205.9	21.31	0	12.1	661	
29	03/02/2013	1	14.30.00	1073.8	21.09	0	11.2	373	
30	03/02/2013	1	15.15.00	1132.9	21.04	0	10.4	316	
31	03/02/2013	1	16.30.00	2014.19	20.75	0	9.5	21	
32	09/02/2013	7	10.30.00	1540.09	21.9	0	5.4	443	
33	09/02/2013	7	14.30.00	1453.4	23.55	0	7.2	338	
34	09/02/2013	7	18.30.00	1637	23.86	0	4.4	0	
35	09/02/2013	7	19.30.00	2107.5	23.41	0	3.1	0	
36	10/02/2013	1	7.15.00	1975.19	20.05	0	0.1	0	
37	10/02/2013	1	12.15.00	1520.9	19.81	0	6.2	538	
38	10/02/2013	1	15.00.00	1571.3	20.06	0	7.9	346	
39	10/02/2013	1	18.15.00	2011.09	19.43	0	4.2	0	
40	23/02/2013	7	10.15.00	2025.69	22.76	1	5.4	35	
41	24/02/2013	1	10.00.00	1645.59	21.06	0	6.6	165	
42	24/02/2013	1	11.30.00	1296.09	20.84	0	8.1	687	

Tabella 21: Outlier individuati con il metodo DSCAN e comuni con il metodo K-Means (GESD) per la potenza elettrica attiva per illuminazione e potenza elettrica attiva totale (edificio F73)

S.No	DATE	Id DAY	HOUR	Electric active Power for lighting F73	Total electric active power F73	Indoor temperature F73	People presence F73	Outdoor temperature	Global solar radiation	
1	06/12/2012	5	7.00.00	2.82	5838	20.89	0	5.2	381	DBSCAN-Lighting Power
2	06/12/2012	5	7.15.00	1.22	5427	21.07	0	5.6	292	DBSCAN-Total Power
3	07/12/2012	6	7.30.00	595	2907.89	20.88	1	3.7	107	Common-DBSCAN&K-Means
4	11/12/2012	3	7.15.00	880	2660.19	20.55	1	7.7	364	
5	11/12/2012	3	7.30.00	695	3525.5	20.55	1	7.9	373	
6	14/12/2012	6	7.15.00	2115	4722.5	20.57	0	9.2	302	
7	14/12/2012	6	7.30.00	1085	4114.1	20.57	1	9.7	512	
8	17/12/2012	2	7.00.00	130	2073.89	20.3	0	11.3	246	
9	17/12/2012	2	7.30.00	385	1879.5	20.26	0	13.1	466	
10	18/12/2012	3	7.00.00	3105	6389.2	21.72	1	8.3	311	
11	18/12/2012	3	7.15.00	1905	6200.39	21.86	1	9.2	334	
12	18/12/2012	3	7.30.00	900	5207.2	21.42	1	9.6	343	
13	19/12/2012	4	7.30.00	950	2841.8	20.31	1	9.7	357	
14	20/12/2012	5	6.45.00	2070	4264.5	21.38	0	7.1	223	
15	07/01/2013	2	6.30.00	5	1720.4	18.45	0	9.9	288	
16	07/01/2013	2	6.45.00	5	1705.09	18.75	0	10.3	313	
17	07/01/2013	2	7.00.00	2190	4285	18.98	0	10.8	336	
18	07/01/2013	2	7.15.00	2360	6475.29	19.38	0	11.2	346	
19	07/01/2013	2	7.30.00	670	4925.5	19.61	0	11.7	358	
20	08/01/2013	3	7.00.00	860	2531.19	20.91	0	10	185	
21	08/01/2013	3	7.15.00	895	2512.8	21.1	1	9.9	100	
22	08/01/2013	3	7.30.00	900	2532.39	21.31	1	9.8	111	
23	09/01/2013	4	17.45.00	1575	2680.1	24.54	9	10.4	0	
24	17/01/2013	5	17.45.00	2045	4186	24.09	5	2.9	0	
25	21/01/2013	2	7.15.00	3120	7425.39	20.6	0	8.4	0	
26	05/02/2013	3	19.00.00	110	1328.8	23.51	3	9.3	0	
27	08/02/2013	6	7.30.00	870	4066	20.95	2	4.4	7	
28	27/02/2013	4	17.45.00	1345	3173.6	24.49	7	10.1	9	

Tabella 22: Outlier comuni individuati dai tre metodi analizzati (CART(ESD) – K-Means (ESD) – DBSCAN) per la potenza elettrica attiva totale (edificio F66)

S.No	DATE	Id DAY	HOUR	Total electric active power F66	Indoor temperature F66	People presence F66	Outdoor temperature	Global solar radiation
1	12/13/2012	5	7:15:00 AM	4729.20	17.69	0	4.9	311
2	12/13/2012	5	7:30:00 AM	7655.79	16.47	0	5.2	369
3	1/25/2013	6	7:00:00 AM	5064.00	16.32	0	5.4	0
4	1/30/2013	4	7:15:00 AM	5175.50	16.8	0	5.5	0
5	2/7/2013	5	7:30:00 AM	5980.00	18.93	0	3.7	0
6	2/11/2013	2	7:30:00 AM	6852.39	15.09	0	3.2	0

Tabella 23: Outlier comuni individuati dai tre metodi analizzati (CART(ESD) – K-Means (ESD) – DBSCAN) per la potenza elettrica attiva totale (edificio F68)

S.No	DATE	Id DAY	HOUR	Total electric active power F68	Indoor temperature F68	People presence F68	Outdoor temperature	Global solar radiation
1	12/4/2012	3	6:45:00 AM	5462.7	22.41	0	5.6	107
2	12/4/2012	3	7:00:00 AM	6237.5	22.67	0	5.7	88
3	1/21/2013	2	7:15:00 AM	8055.5	21.2	0	8.4	0

Tabella 24: Outlier comuni individuati dai tre metodi analizzati (CART(ESD) – K-Means (ESD) – DBSCAN) per la potenza elettrica attiva totale (edificio F69)

S.No	DATE	Id DAY	HOUR	Total electric active power F69	Indoor temperature F69	People presence F69	Outdoor temperature	Global solar radiation
1	12/13/2012	5	10:15:00 AM	6639.1	21.79	6	6	46
2	1/21/2013	2	7:45:00 AM	7197.39	19.93	1	8.3	5
3	1/21/2013	2	8:00:00 AM	6614.89	20.11	1	8.2	5
4	1/21/2013	2	8:15:00 AM	6817.89	20.3	2	8.4	16
5	1/21/2013	2	8:30:00 AM	6414.39	20.55	3	8.3	26
6	1/26/2013	7	1:45:00 PM	2327.8	19.96	0	8.2	401
7	1/27/2013	1	11:30:00 AM	2111.89	18.77	0	7	441
8	2/10/2013	1	1:45:00 PM	1997.8	19.6	0	7.1	471

Tabella 25: Outlier comuni individuati dai tre metodi analizzati (CART(ESD) – K-Means (ESD) – DBSCAN) per la potenza elettrica attiva totale (edificio F72)

S.No	DATE	Id DAY	HOUR	Total electric active power F72	Indoor temperature F72	People presence F72	Outdoor temperature	Global solar radiation
1	12/2/2012	1	7:45:00 AM	2517	22.75	0	8.8	0
2	12/2/2012	1	8:00:00 PM	2588.5	21.9	0	6.2	0
3	12/14/2012	6	6:30:00 AM	5514.5	18.94	0	8.5	91
4	12/14/2012	6	6:45:00 AM	4492.1	18.3	0	8.6	103
5	12/17/2012	2	7:30:00 AM	4815.6	19.67	1	13.1	466
6	1/14/2013	2	7:30:00 AM	4498.2	19.4	0	9	0
7	1/19/2013	7	2:30:00 AM	2529.89	20.09	0	0.4	0
8	1/19/2013	7	2:45:00 AM	2403.5	20.06	0	0.2	0
9	1/19/2013	7	8:15:00 AM	2606.89	19.83	0	2.6	12
10	1/19/2013	7	2:45:00 PM	2699.1	21.31	0	4.5	51
11	1/20/2013	1	3:15:00 AM	2628.19	20.04	0	9.2	0
12	1/22/2013	3	7:15:00 AM	4694.2	18.76	0	4	0
13	2/5/2013	3	7:15:00 AM	6533.6	19.19	0	5.9	0
14	2/5/2013	3	7:30:00 AM	6010.5	19.03	0	5.9	0
15	2/7/2013	5	7:30:00 AM	4765.6	20.93	1	3.7	0
16	2/13/2013	4	7:30:00 AM	4558	21.64	1	3.8	11

Tabella 26: Outlier comuni individuati dai tre metodi analizzati (CART(ESD) – K-Means (ESD) – DBSCAN) per la potenza elettrica attiva totale (edificio F73)

S.No	DATE	Id DAY	HOUR	Total electric active power F73	Indoor temperature F73	People presence F73	Outdoor temperature	Global solar radiation
1	1/7/2013	2	7:15:00 AM	6475.29	19.38	0	11.2	346
2	1/19/2013	7	10:00:00 PM	3719.1	21.85	0	8.1	0
3	1/20/2013	1	1:45:00 PM	3883.69	20.55	0	13.4	51
4	1/21/2013	2	7:15:00 AM	7425.39	20.6	0	8.4	0
5	1/31/2013	5	7:15:00 AM	6732	21.91	0	8.6	0

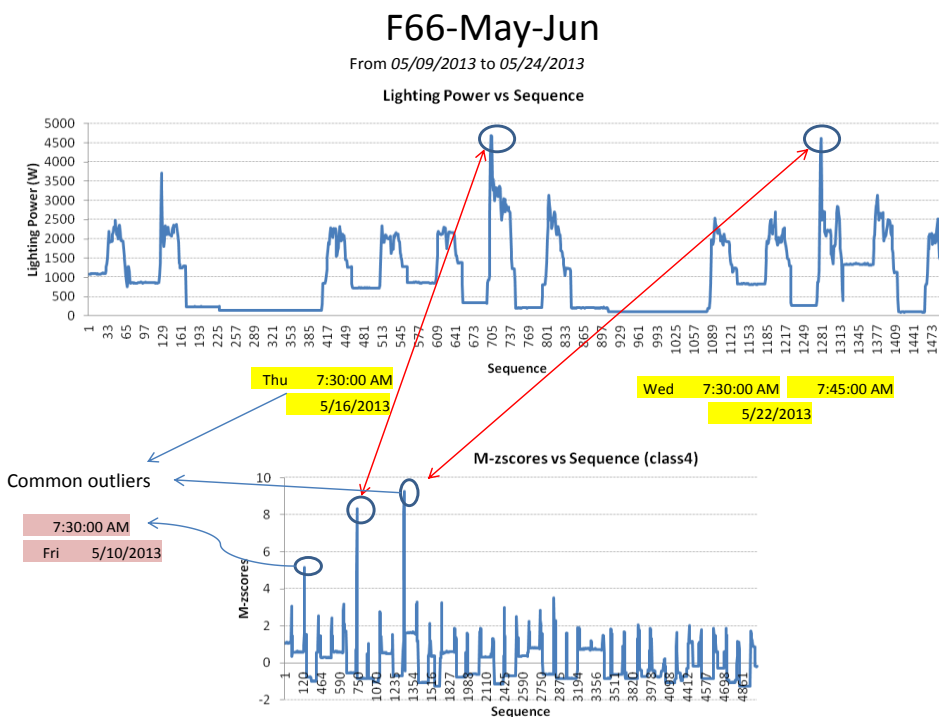


Figura 71: Andamento della potenza elettrica attiva per illuminazione nel tempo con dati non classificati e variazione dell'indice Mzscore in un classe di dati (classificazione con metodo CART) con evidenza dei principali outlier comuni individuati dai tre metodi analizzati (CART(ESD) – K-Means (ESD) – DBSCAN) (edificio F66)

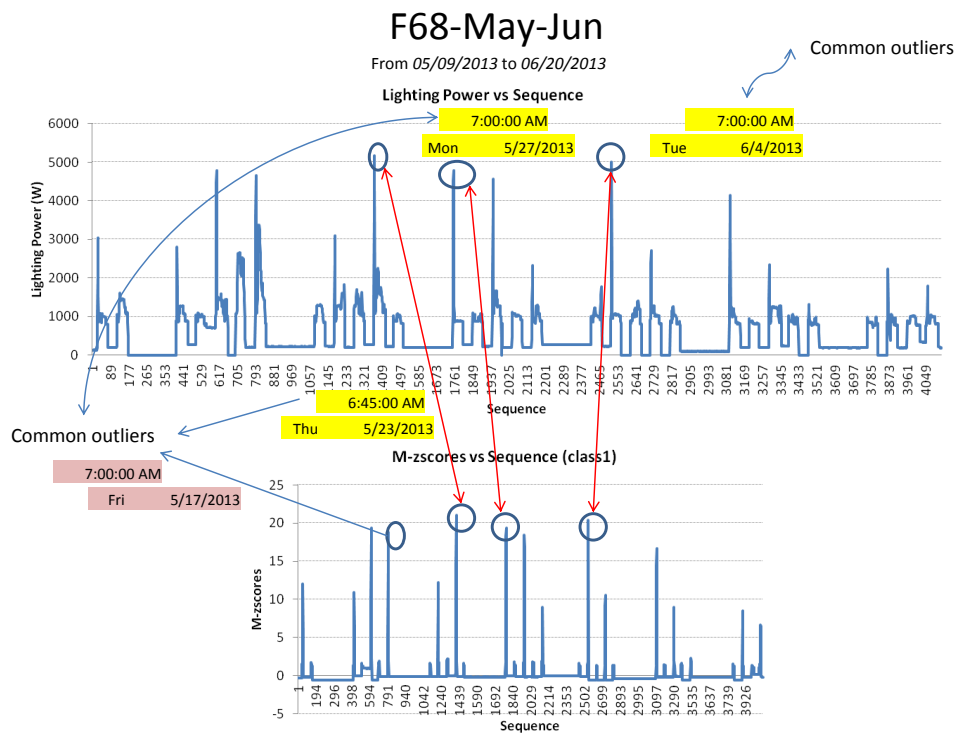


Figura 72: Andamento della potenza elettrica attiva per illuminazione nel tempo con dati non classificati e variazione dell'indice Mzscore in un classe di dati (classificazione con metodo CART) con evidenza dei principali outlier comuni individuati dai tre metodi analizzati (CART(ESD) – K-Means (ESD) – DBSCAN) (edificio F68)

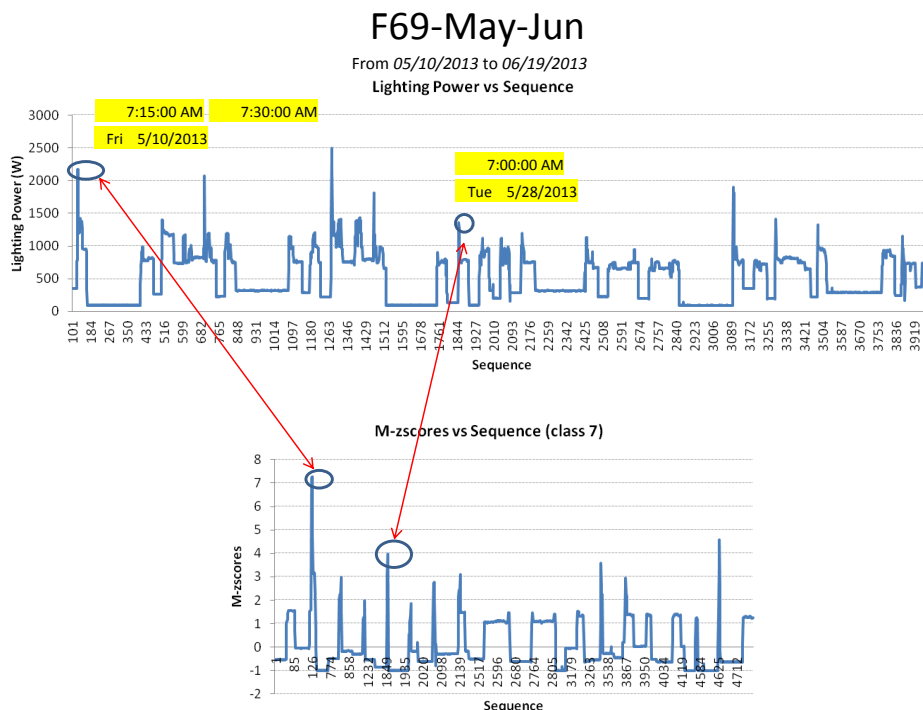


Figura 73: Andamento della potenza elettrica attiva per illuminazione nel tempo con dati non classificati e variazione dell'indice Mzscore in un classe di dati (classificazione con metodo CART) con evidenza dei principali outlier individuati con metodo (CART(ESD) – K-Means (ESD) – DBSCAN) (edificio F69)

Common Outliers

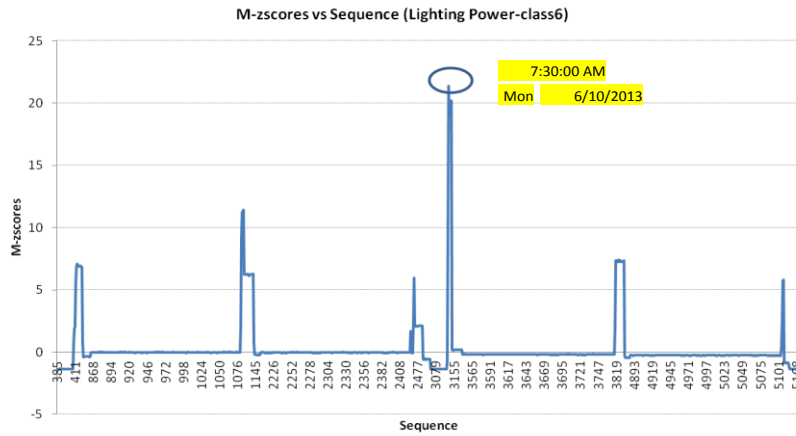


Figura 74 : Variazione dell'indice Mzscore in un classe di dati (classificazione con metodo CART) con evidenza dei principali outlier comuni individuati dai tre metodi analizzati (CART(ESD) – K-Means (ESD) – DBSCAN) (edificio F69)

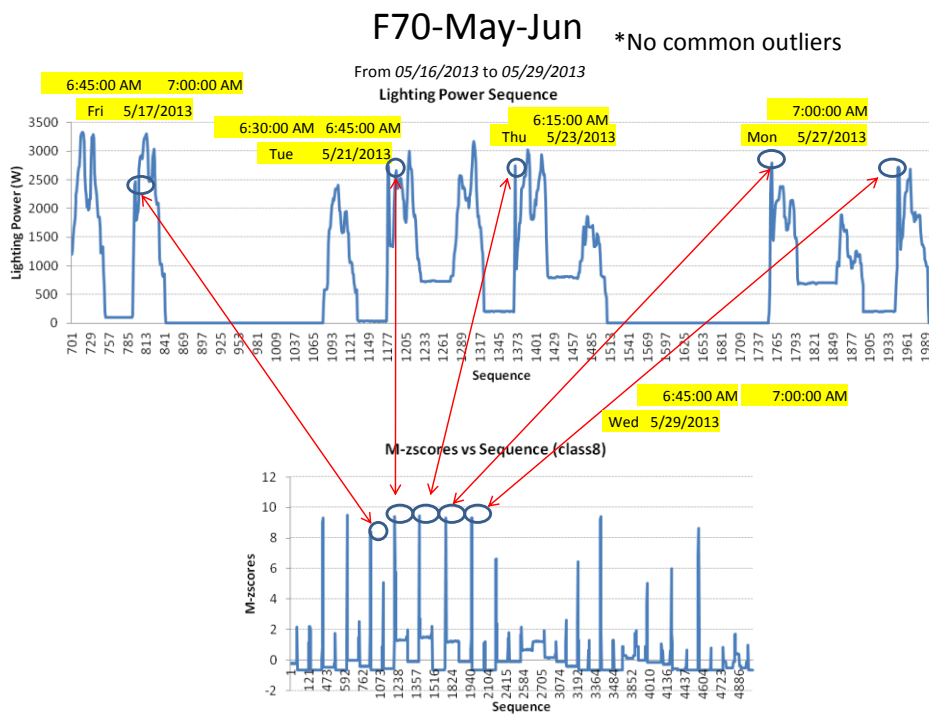


Figura 75: Andamento della potenza elettrica attiva per illuminazione nel tempo con dati non classificati e variazione dell'indice Mzscore in un classe di dati (classificazione con metodo CART) con evidenza dei principali outlier individuati con ESD (edificio F70)

F71-May-Jun

From 05/09/2013 to 07/01/2013

Lighting Power vs Sequence

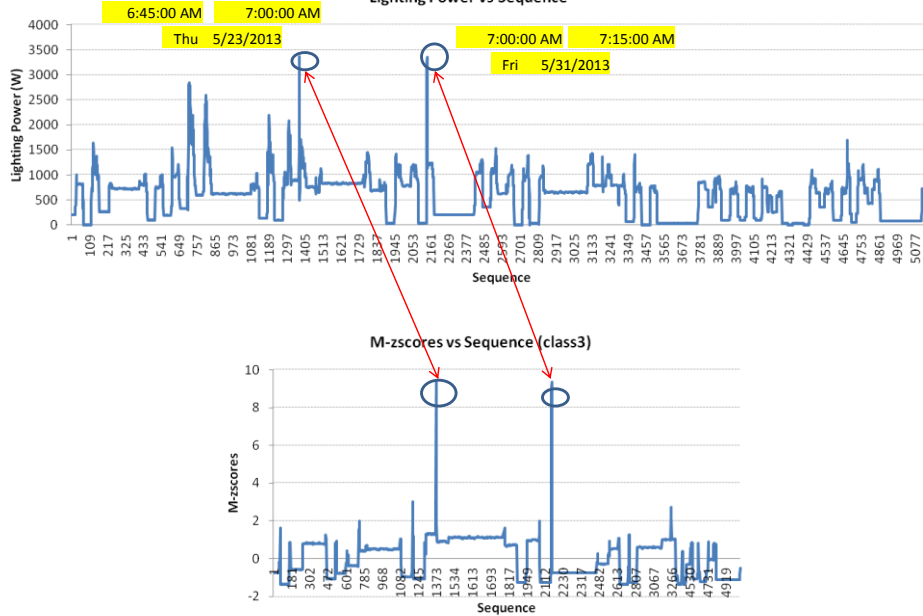


Figura 76: Andamento della potenza elettrica attiva per illuminazione nel tempo con dati non classificati e variazione dell'indice Mzscore in un classe di dati (classificazione con metodo CART) con evidenza dei principali outlier individuati con metodo (CART(ESD) – K-Means (ESD) – DBSCAN) (edificio F71)

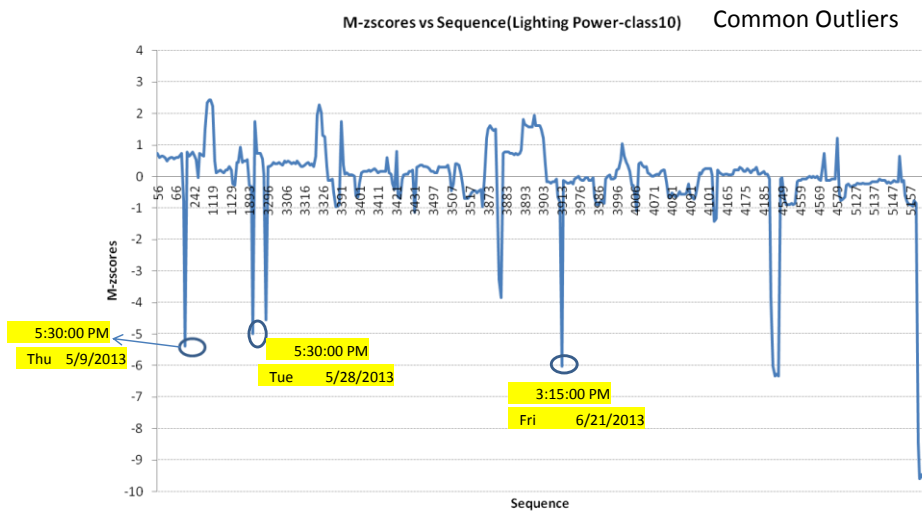


Figura 77 : Variazione dell'indice Mzscore in un classe di dati (classificazione con metodo CART) con evidenza dei principali outlier comuni individuati dai tre metodi analizzati (CART(ESD) – K-Means (ESD) – DBSCAN) (edificio F71)

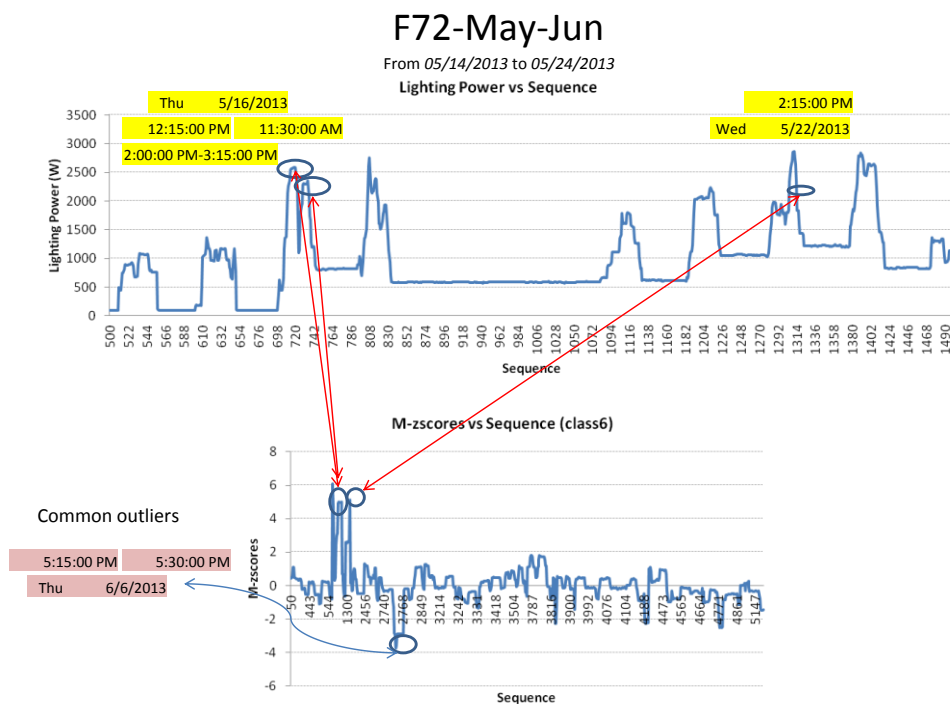


Figura 78: Andamento della potenza elettrica attiva per illuminazione nel tempo con dati non classificati e variazione dell'indice Mzscore in un classe di dati (classificazione con metodo CART) con evidenza dei principali outlier comuni individuati dai tre metodi analizzati (CART(ESD) – K-Means (ESD) – DBSCAN) (edificio F72)

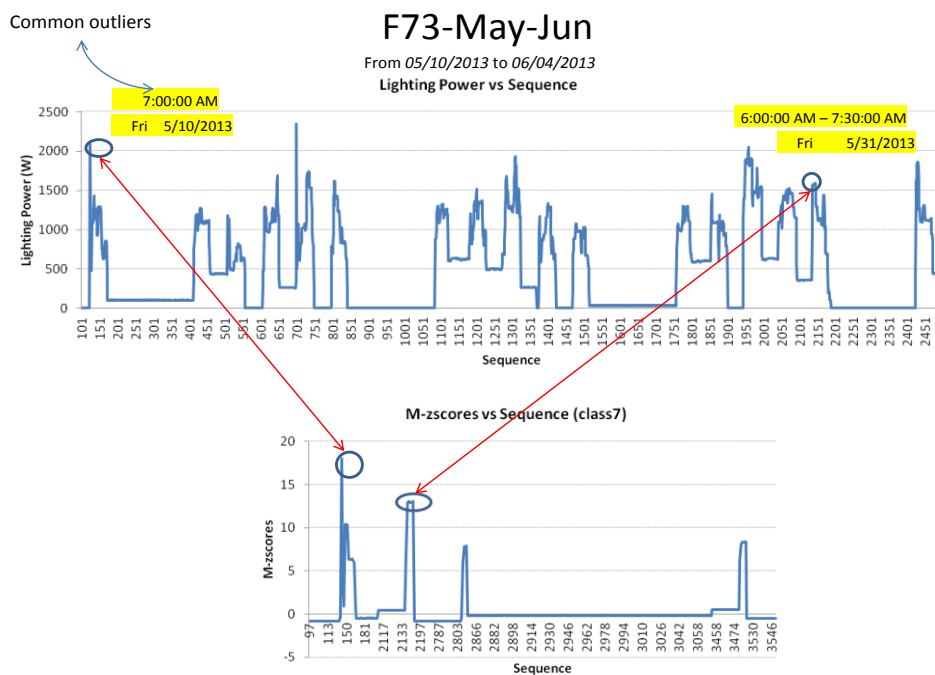


Figura 79: Andamento della potenza elettrica attiva per illuminazione nel tempo con dati non classificati e variazione dell'indice Mzscore in un classe di dati (classificazione con metodo CART) con evidenza dei principali outlier comuni individuati dai tre metodi analizzati (CART(ESD) – K-Means (ESD) – DBSCAN) (edificio F72)

F66-May-Jun

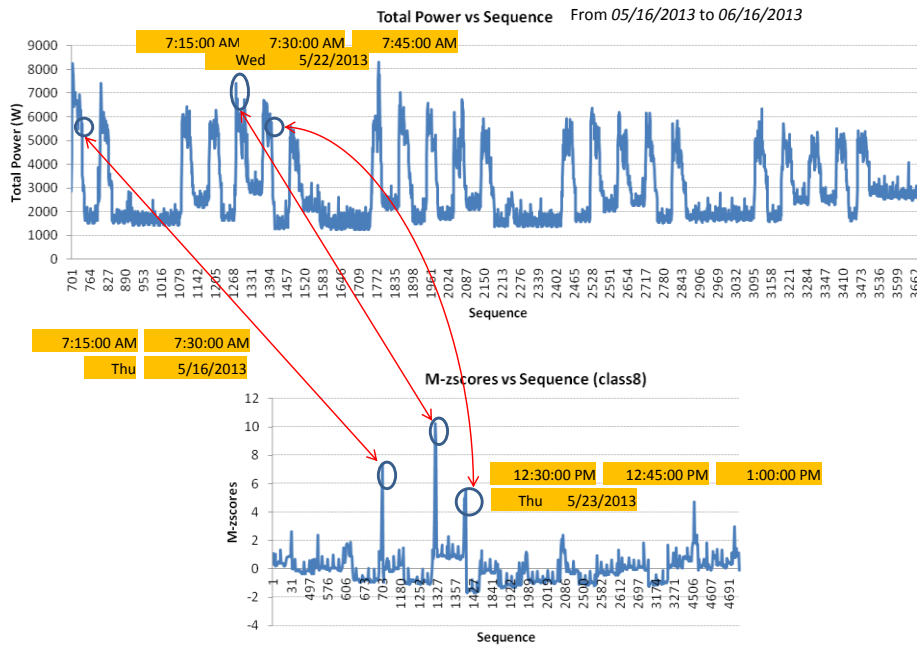


Figura 80: Andamento della potenza elettrica attiva totale nel tempo con dati non classificati e variazione dell'indice Mzscore in un classe di dati (classificazione con metodo CART) con evidenza dei principali outlier individuati con ESD (edificio F66)

F68-May-Jun

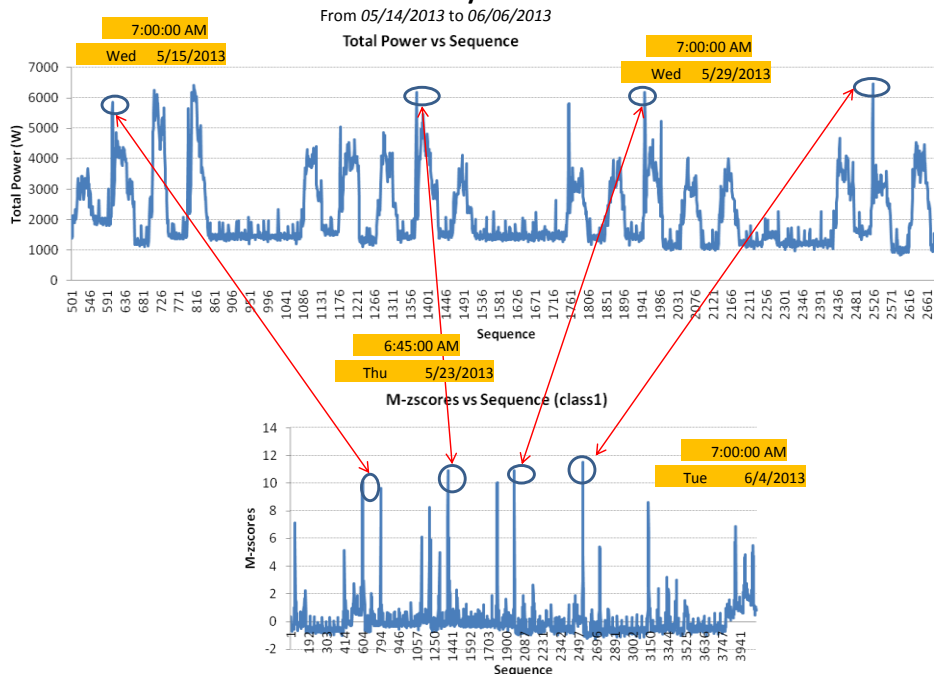


Figura 81: Andamento della potenza elettrica attiva totale nel tempo con dati non classificati e variazione dell'indice Mzscore in un classe di dati (classificazione con metodo CART) con evidenza dei principali outlier individuati con ESD (edificio F68)

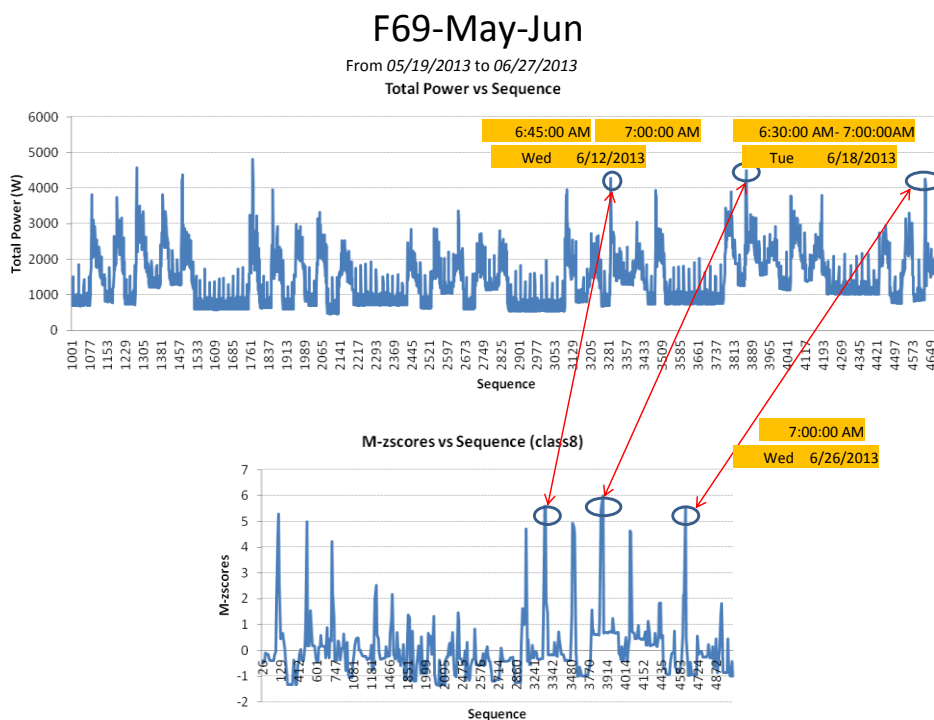


Figura 82: Andamento della potenza elettrica attiva totale nel tempo con dati non classificati e variazione dell'indice Mzscore in un classe di dati (classificazione con metodo CART) con evidenza dei principali outlier individuati con ESD (edificio F73)

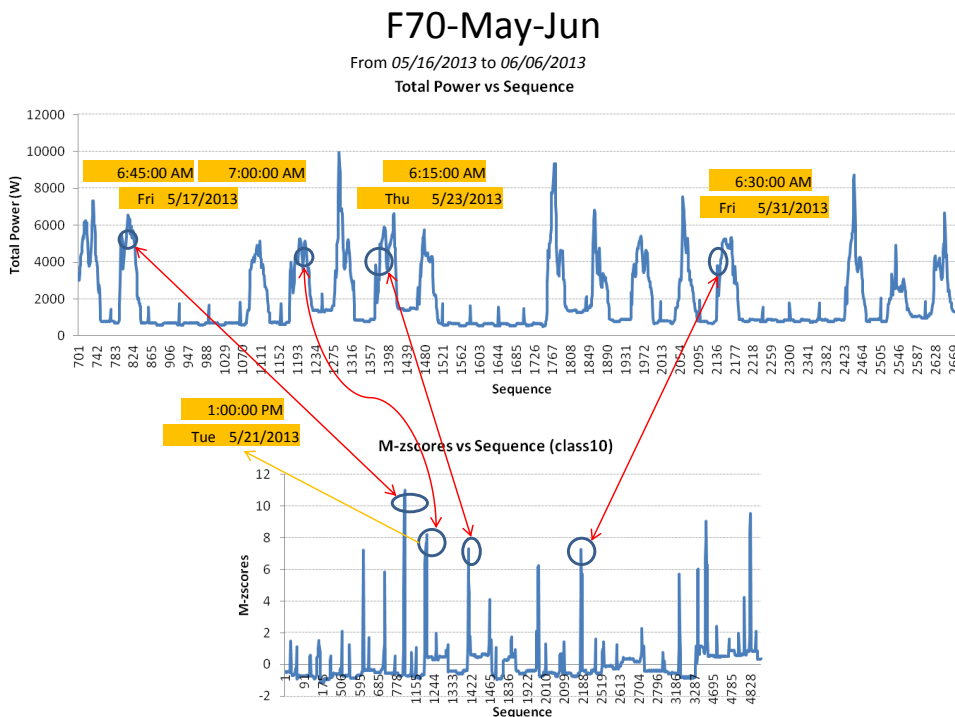


Figura 83: Andamento della potenza elettrica attiva totale nel tempo con dati non classificati e variazione dell'indice Mzscore in un classe di dati (classificazione con metodo CART) con evidenza dei principali outlier individuati con ESD (edificio F70)

F71-May-Jun

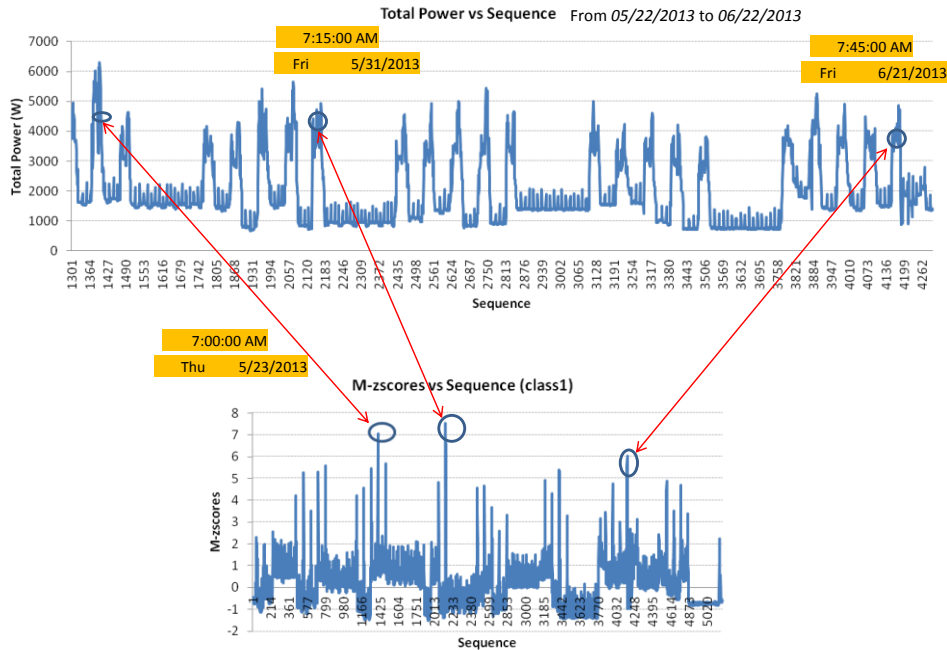


Figura 84: Andamento della potenza elettrica attiva totale nel tempo con dati non classificati e variazione dell'indice Mzscore in un classe di dati (classificazione con metodo CART) con evidenza dei principali outlier individuati con ESD (edificio F71)

F72-May-Jun

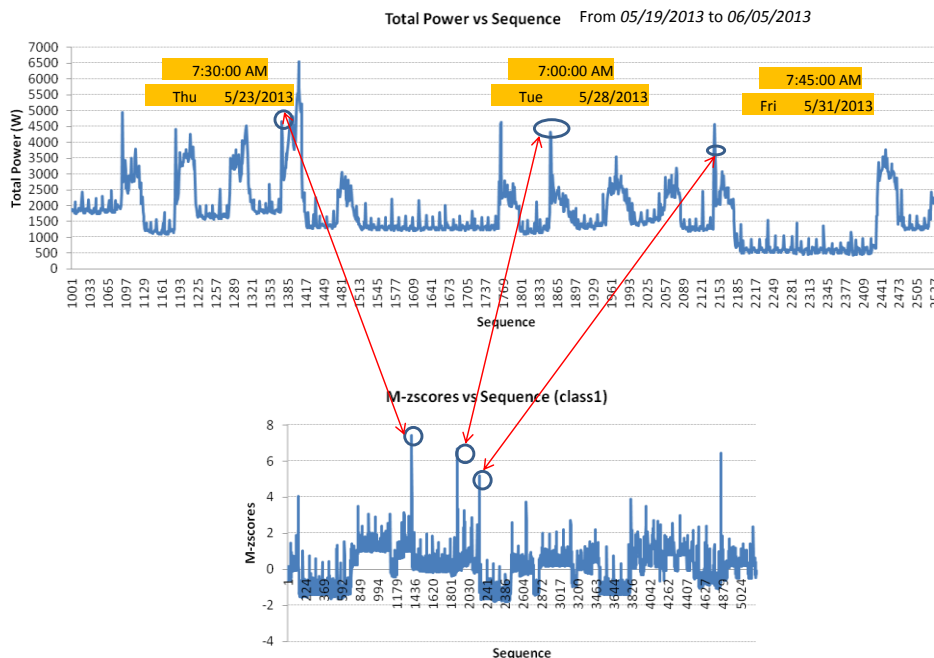


Figura 85: Andamento della potenza elettrica attiva totale nel tempo con dati non classificati e variazione dell'indice Mzscore in un classe di dati (classificazione con metodo CART) con evidenza dei principali outlier individuati con ESD (edificio F72)

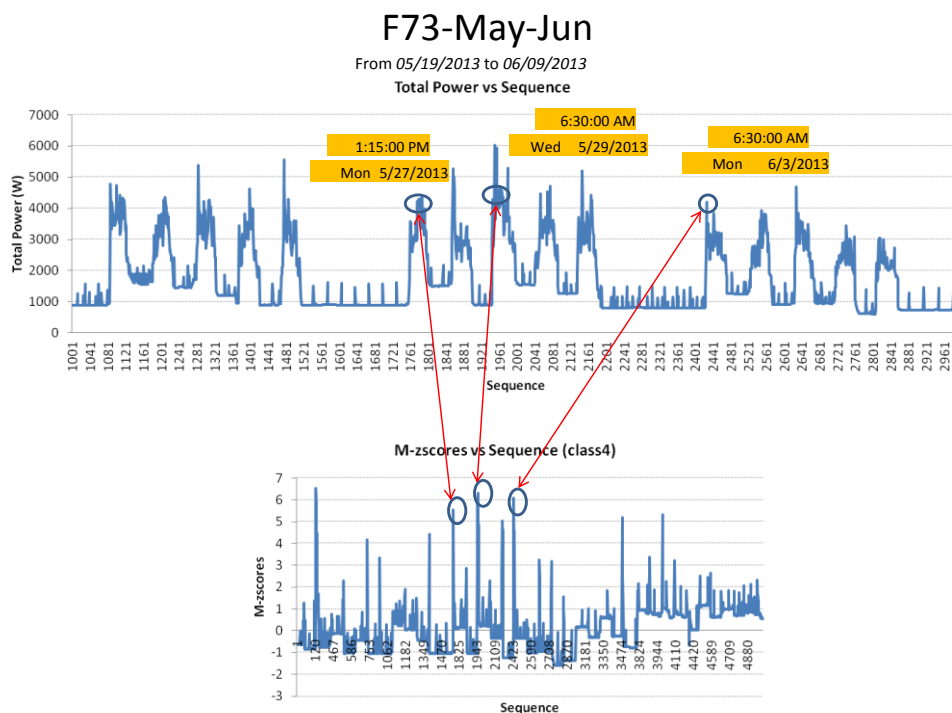


Figura 86: Andamento della potenza elettrica attiva totale nel tempo con dati non classificati e variazione dell'indice Mzscore in un classe di dati (classificazione con metodo CART) con evidenza dei principali outlier individuati con ESD (edificio F73)

Tabella 27: Outlier individuati solo con il metodo DSCAN e comuni con il metodo K-Means (GESD) per la potenza elettrica attiva per illuminazione e potenza elettrica attiva totale (edificio F66)

S.No	DATE	Id DAY	HOUR	Electric active Power for lighting F66	Total electric active power F66	Indoor temperature F66	People presence F66	Outdoor temperature	Global solar radiation	
1	10/05/2013	6	7.30.00	2825	4583.2	20.79	1	15.4	40	DBSCAN-Lighting Power
2	10/05/2013	6	17.45.00	1295	3293.39	22.44	0	18.6	176	DBSCAN-Total Power
3	10/05/2013	6	18.00.00	1295	2963.6	22.57	0	18.8	111	Common-DBSCAN&K-Means
4	16/05/2013	5	7.30.00	4200	5950.29	20.92	1	13.4	39	
5	16/05/2013	5	17.45.00	1230	3113.39	22.34	4	22.8	123	
6	20/05/2013	2	18.15.00	1250	2710.8	22.16	1	18.2	204	
7	22/05/2013	4	7.15.00	2530	4392.89	20.07	1	13.8	40	
8	22/05/2013	4	7.30.00	4165.2	7422.79	19.34	2	13.9	65	
9	27/05/2013	2	7.15.00	720	2111.6	19.18	1	9.9	83	
10	29/05/2013	4	7.15.00	845	2227.6	20.21	0	14	160	
11	29/05/2013	4	17.45.00	305	2348.8	21.65	0	17.4	221	
12	30/05/2013	5	7.30.00	1365	3196.8	19.45	2	11.1	37	
13	31/05/2013	6	7.30.00	1020	2645.6	20.08	0	10.4	127	
14	03/06/2013	2	7.30.00	2490	4283	19.05	3	14.4	72	
15	05/06/2013	4	7.30.00	1855	3813.69	20.68	4	14.7	76	
16	05/06/2013	4	17.45.00	1545	3137	22.16	3	21.7	146	
17	10/06/2013	2	17.45.00	1735	3678.6	24.09	5	21.8	47	
18	10/06/2013	2	18.00.00	1515	3251.89	24.09	5	21.8	46	
19	12/06/2013	4	17.45.00	1660	3595.19	25	4	28.4	51	
20	12/06/2013	4	18.00.00	1545	3267.69	24.99	3	28.2	49	
21	12/06/2013	4	18.15.00	1225	2790.39	25.03	3	28.1	47	
22	18/06/2013	3	17.45.00	1590	4448.7	25.03	3	30.7	63	
23	18/06/2013	3	18.00.00	1600	4452.7	24.89	3	30.3	63	
24	21/06/2013	6	17.45.00	860	3762.39	23.54	1	26.9	98	
25	25/06/2013	3	17.45.00	1215	3667.8	23.13	3	27.1	63	
26	25/06/2013	3	18.00.00	900	3158.69	23.04	3	26.3	107	
27	25/06/2013	3	18.15.00	220	2500.6	23.12	1	25.5	148	

Tabella 28: Outlier individuati solo con il metodo DSCAN e comuni con il metodo K-Means (GESD) per la potenza elettrica attiva per illuminazione e potenza elettrica attiva totale (edificio F68)

S.No	DATE	Id DAY	HOUR	Electric active Power for lighting F68	Total electric active power F68	Indoor temperature F68	People presence F68	Outdoor temperature	Global solar radiation	
1	17/05/2013	6	7.00.00	3235	4342.39	21.91	0	15.3	63	DBSCAN-Total Power
2	23/05/2013	5	6.45.00	5165.2	6191.29	21.64	0	13.1	23	Common-DBSCAN&K-Means
3	27/05/2013	2	6.45.00	4530	5791.39	20.07	0	9.4	28	
4	27/05/2013	2	7.00.00	4785.2	5826.2	20.12	0	9.5	35	
5	27/05/2013	2	7.15.00	3840	4860	20.1	0	9.9	83	
6	29/05/2013	4	7.15.00	4160	5200.6	21.28	0	14	160	
7	04/06/2013	3	7.00.00	5015.2	6462.2	21.85	0	14.2	26	
8	10/06/2013	2	6.45.00	3640	4515.6	24.09	0	16.3	105	
9	10/06/2013	2	7.00.00	4140	5202.79	24.09	0	16.4	90	
10	10/06/2013	2	7.15.00	2910	4194.89	24.07	1	16.5	102	

Tabella 29: Outlier individuati solo con il metodo DSCAN e comuni con il metodo K-Means (GESD) per la potenza elettrica attiva per illuminazione e potenza elettrica attiva totale (edificio F69)

S. No	DATE	Id DAY	HOUR	Total electric active power F69	Indoor temperature F69	People presence F69	Outdoor temperature	Global solar radiation	
1	25/05/2013	7	17.45.00	1376	21	0	14.4	146	Common-DBSCAN&K-Means
2	26/05/2013	1	3.30.00	1470.19	20.6	0	7.9	0	
3	26/05/2013	1	8.45.00	1303.59	20.44	0	12.1	248	
4	26/05/2013	1	13.30.00	1543.9	20.58	0	15.8	274	
5	26/05/2013	1	18.45.00	1736.5	20.86	0	16.4	40	
6	02/06/2013	1	10.15.00	1350.69	21.85	0	18.6	357	
7	09/06/2013	1	10.45.00	1973.3	24.28	0	22.6	441	
8	22/06/2013	7	17.15.00	2157	24.96	0	27	264	
9	23/06/2013	1	11.45.00	2176	24.08	0	26.9	889	
10	23/06/2013	1	18.30.00	2151.69	24.76	0	25.8	44	
11	30/06/2013	1	8.45.00	2003.4	23.92	0	19.8	485	

Tabella 30: Outlier individuati solo con il metodo DSCAN e comuni con il metodo K-Means (GESD) per la potenza elettrica attiva per illuminazione e potenza elettrica attiva totale (edificio F70)

S.No	DATE	Id DAY	HOUR	Electric active Power for lighting F70	Total electric active power F70	Indoor temperature F70	People presence F70	Outdoor temperature	Global solar radiation	
1	09/05/2013	5	17.45.00	270	1048.5	22.45	2	22.2	162	DBSCAN-Lighting Power
2	10/05/2013	6	17.45.00	1050	1880.4	21.68	0	18.6	176	DBSCAN-Total Power
3	15/05/2013	4	17.45.00	1580	2582.19	22.32	3	17.3	25	
4	15/05/2013	4	18.00.00	1260	2055	22.33	3	17.2	18	
5	16/05/2013	5	17.45.00	1700	3657.89	22.32	6	22.8	123	
6	16/05/2013	5	18.00.00	1220	3498.69	22.18	6	22.8	111	
7	16/05/2013	5	18.15.00	1275	2542.89	22.19	4	23	100	
8	16/05/2013	5	18.30.00	1335	2320	22.22	4	23.1	93	
9	17/05/2013	6	7.15.00	1790	5143.7	19.42	1	15.6	139	
10	20/05/2013	2	17.45.00	630	1606.59	21.7	3	18.2	186	
11	20/05/2013	2	18.00.00	540	1317.09	21.65	2	18	162	
12	20/05/2013	2	18.15.00	535	1265.4	21.61	2	18.2	204	
13	23/05/2013	5	17.45.00	2130	3116.5	21.2	6	19.5	97	
14	23/05/2013	5	18.00.00	1490	2199.5	21.16	4	19	51	
15	29/05/2013	4	7.15.00	2690	3543.6	20.05	1	14	160	
16	29/05/2013	4	17.45.00	995	2102.19	21.06	3	17.4	221	
17	31/05/2013	6	17.45.00	1300	2431.1	20.64	3	16.6	153	
18	04/06/2013	3	18.30.00	520	1201.09	22.3	2	20.8	158	
19	06/06/2013	5	17.45.00	1395	2683.3	23.28	6	24.3	63	
20	06/06/2013	5	18.00.00	1215	2108.89	23.3	5	24.1	62	
21	19/06/2013	4	7.15.00	1590	3667.39	25.16	1	23	35	
22	24/06/2013	2	7.30.00	930	2850.6	23.71	2	18.9	118	
23	25/06/2013	3	17.45.00	675	2480.19	24.28	5	27.1	63	
24	25/06/2013	3	18.00.00	695	2213.39	24.15	4	26.3	107	
25	27/06/2013	5	18.30.00	555	2072.69	23.12	2	23.8	137	
26	28/06/2013	6	18.00.00	405	1641.8	22.44	0	18.4	186	

Tabella 31: Outlier individuati solo con il metodo DSCAN e comuni con il metodo K-Means (GESD) per la potenza elettrica attiva per illuminazione e potenza elettrica attiva totale (edificio F71)

S.No	DATE	Id DAY	HOUR	Electric active Power for lighting F71	Indoor temperature F71	People presence F71	Outdoor temperature	Global solar radiation	
1	09/05/2013	5	17.30.00	330	23.72	2	22.4	429	DBSCAN-Lighting Power
2	16/05/2013	5	8.15.00	2260	21.21	8	14.9	394	
3	16/05/2013	5	8.30.00	2540	21.34	8	15.9	446	Common-DBSCAN&K-Means
4	16/05/2013	5	11.45.00	2250	22.04	12	21.4	564	
5	16/05/2013	5	13.30.00	1760	22.13	8	23.7	385	
6	17/05/2013	6	8.45.00	2000	21.2	8	16.5	267	
7	17/05/2013	6	9.00.00	2155	21.31	10	16.7	552	
8	17/05/2013	6	9.30.00	2300	21.68	13	17.8	272	
9	17/05/2013	6	10.30.00	2485	21.74	13	19.2	181	
10	17/05/2013	6	12.00.00	2075	21.85	9	20.2	976	
11	21/05/2013	3	9.15.00	2180	21.58	12	14.9	165	
12	21/05/2013	3	9.30.00	2195	21.62	12	15.1	178	
13	28/05/2013	3	17.30.00	360	22.93	3	21.8	197	
14	31/05/2013	6	7.45.00	1155	18.95	3	10.9	111	
15	31/05/2013	6	17.30.00	205	22.3	1	16.6	172	
16	06/06/2013	5	16.30.00	190	24.6	6	25	624	
17	06/06/2013	5	16.45.00	135	24.13	4	24.8	596	
18	07/06/2013	6	7.45.00	410	21.83	2	18.9	46	
19	11/06/2013	3	16.15.00	1195	25.54	7	24.9	81	
20	14/06/2013	6	17.00.00	75	27.54	1	28.1	519	
21	14/06/2013	6	17.15.00	35	27.39	0	28.9	227	
22	14/06/2013	6	17.30.00	35	27.31	0	28.7	70	
23	21/06/2013	6	15.15.00	280	25.64	6	29.3	747	
24	26/06/2013	4	15.00.00	1550	24.32	14	23.4	95	
25	26/06/2013	4	15.15.00	1700	24.07	14	21.4	125	
26	28/06/2013	6	16.45.00	175	24.05	1	22.8	107	
27	28/06/2013	6	17.00.00	80	23.94	0	22.9	153	
28	28/06/2013	6	17.15.00	80	23.88	0	23.5	410	
29	28/06/2013	6	17.30.00	85	23.85	0	23.4	204	

Tabella 32: Outlier individuati solo con il metodo DSCAN e comuni con il metodo K-Means (GESD) per la potenza elettrica attiva per illuminazione e potenza elettrica attiva totale (edificio F72)

S.No	DATE	Id DAY	HOUR	Total electric active power F72	Indoor temperature F72	People presence F72	Outdoor temperature	Global solar radiation	
1	10/05/2013	6	16.45.00	3192	22.87	1	19.4	225	DBSCAN-Total Power
2	10/05/2013	6	17.00.00	3126.8	23.1	0	19.3	197	Common-DBSCAN&K-Means
3	16/05/2013	5	11.45.00	5472.5	22.57	11	21.4	564	
4	20/05/2013	2	8.30.00	4948.2	22.72	5	15.4	454	
5	23/05/2013	5	7.45.00	4162	21.88	0	13.2	95	
6	23/05/2013	5	8.00.00	2991.1	21.85	1	13.3	151	
7	23/05/2013	5	15.00.00	6559.5	22.73	10	16.8	206	
8	27/05/2013	2	9.00.00	4648.5	21.09	6	14.8	515	
9	28/05/2013	3	7.45.00	2869.39	20.11	1	14.3	151	
10	31/05/2013	6	7.45.00	3650.6	19.42	1	10.9	111	
11	31/05/2013	6	8.00.00	4568.79	19.65	3	11.4	144	
12	31/05/2013	6	8.15.00	3477.8	20.03	5	12.1	404	
13	06/06/2013	5	17.15.00	3872.6	23.31	4	24.6	287	
14	10/06/2013	2	12.30.00	3567.6	25.36	7	21.8	1044	
15	10/06/2013	2	13.00.00	2578.19	25.22	0	22.2	1060	
16	10/06/2013	2	13.15.00	2435.8	25.16	0	21.9	1007	
17	12/06/2013	4	8.45.00	3597.8	24.6	5	21.1	510	
18	14/06/2013	6	10.45.00	3367.19	26.77	9	26.2	111	
19	14/06/2013	6	17.00.00	1386.69	27.63	1	28.1	519	
20	14/06/2013	6	17.15.00	1299	27.67	1	28.9	227	
21	14/06/2013	6	17.30.00	1237.3	27.62	1	28.7	70	
22	18/06/2013	3	10.45.00	2981.69	25.89	8	30.4	93	
23	19/06/2013	4	10.45.00	2512	23.57	6	30.7	77	
24	20/06/2013	5	9.30.00	4673.29	23.86	9	27.6	613	
25	28/06/2013	6	7.45.00	4215.39	21.85	1	17	74	
26	01/07/2013	2	10.45.00	3003.1	25.51	10	25.7	88	

Tabella 33: Outlier individuati solo con il metodo DSCAN e comuni con il metodo K-Means (GESD) per la potenza elettrica attiva per illuminazione e potenza elettrica attiva totale (edificio F73)

S.No	DATE	Id DAY	HOUR	Electric active Power for lighting F73	Indoor temperature F73	People presence F73	Outdoor temperature	Global solar radiation	
1	09/05/2013	5	17.45.00	920	25.34	4	22.2	162	DBSCAN-Lighting Power
2	10/05/2013	6	7.00.00	2110	23.86	0	15.2	28	
3	10/05/2013	6	17.45.00	855	25.04	0	18.6	176	Common-DBSCAN&K-Means
4	13/05/2013	2	17.45.00	1120	25.34	5	21.9	88	
5	15/05/2013	4	17.45.00	1375	25.05	5	17.3	25	
6	15/05/2013	4	18.00.00	1180	25.16	4	17.2	18	
7	16/05/2013	5	6.45.00	2345	23.88	0	12.2	30	
8	16/05/2013	5	17.45.00	975	24.83	4	22.8	123	
9	21/05/2013	3	7.30.00	820	24.35	2	13.1	118	
10	21/05/2013	3	17.45.00	855	24.62	3	17.4	74	
11	22/05/2013	4	17.45.00	1390	24.81	4	18.6	104	
12	22/05/2013	4	18.00.00	1285	24.81	4	18.4	91	
13	27/05/2013	2	17.45.00	655	23.56	4	17.7	179	
14	28/05/2013	3	7.30.00	1255	22.06	2	13.7	146	
15	28/05/2013	3	17.45.00	955	24.37	4	21.3	148	
16	28/05/2013	3	18.00.00	795	24.43	4	20.6	109	
17	29/05/2013	4	7.15.00	1760	21.31	0	14	160	
18	29/05/2013	4	7.30.00	1885	21.33	1	14.1	76	
19	29/05/2013	4	17.45.00	1535	24.51	3	17.4	221	
20	31/05/2013	6	7.30.00	1575	21.42	0	10.4	127	
21	03/06/2013	2	7.30.00	1855	22.13	1	14.4	72	
22	12/06/2013	4	17.45.00	880	28.04	6	28.4	51	
23	19/06/2013	4	17.45.00	790	26.05	6	32.8	54	
24	19/06/2013	4	18.00.00	785	26.07	4	32.6	53	
25	19/06/2013	4	18.15.00	805	26.2	3	32.4	51	
26	19/06/2013	4	18.30.00	815	26.31	3	31.9	49	
27	20/06/2013	5	17.45.00	600	25.86	4	30.1	67	
28	20/06/2013	5	18.00.00	605	25.93	4	29.9	67	
29	24/06/2013	2	17.45.00	685	24.33	3	27.1	63	
30	27/06/2013	5	19.00.00	100	23.3	2	23.8	63	
31	28/06/2013	6	18.00.00	105	23.35	0	18.4	186	

Tabella 34: Outlier comuni individuati dai tre metodi analizzati (CART(ESD) – K-Means (ESD) – DBSCAN) per la potenza elettrica attiva totale (edificio F66)

S.No	DATE	Id DAY	HOUR	Total electric active power F66	Indoor temperature F66	People presence F66	Outdoor temperature	Global solar radiation
1	5/16/2013	5	7:30:00 AM	5950.29	20.92	1	13.4	39
2	5/16/2013	5	8:00:00 AM	8265.79	20.63	8	14.3	54
3	5/22/2013	4	7:30:00 AM	7422.79	19.34	2	13.9	65
4	6/22/2013	7	9:45:00 AM	4537.39	20.62	2	24.5	668
5	6/23/2013	1	9:00:00 PM	4387.89	21.32	0	21.4	0

Tabella 35: Outlier comuni individuati dai tre metodi analizzati (CART(ESD) – K-Means (ESD) – DBSCAN) per la potenza elettrica attiva totale (edificio F68)

S.No	DATE	Id DAY	HOUR	Total electric active power F68	Indoor temperature F68	People presence F68	Outdoor temperature	Global solar radiation
1	5/23/2013	5	6:45:00 AM	6191.29	21.64	0	13.1	23
2	5/27/2013	2	6:30:00 AM	3800.69	20.06	0	9.2	35
3	5/27/2013	2	6:45:00 AM	5791.39	20.07	0	9.4	28
4	5/27/2013	2	7:00:00 AM	5826.2	20.12	0	9.5	35
5	5/29/2013	4	7:15:00 AM	5200.6	21.28	0	14	160
6	6/4/2013	3	7:00:00 AM	6462.2	21.85	0	14.2	26
7	6/10/2013	2	7:00:00 AM	5202.79	24.09	0	16.4	90

Tabella 36: Outlier comuni individuati dai tre metodi analizzati (CART(ESD) – K-Means (ESD) – DBSCAN) per la potenza elettrica attiva totale (edificio F69)

S.No	DATE	Id DAY	HOUR	Total electric active power F69	Indoor temperature F69	People presence F69	Outdoor temperature	Global solar radiation
1	5/26/2013	1	6:45:00 PM	1736.5	20.86	0	16.4	40
2	5/27/2013	2	10:30:00 AM	4811.89	20.61	7	17	149
3	6/9/2013	1	10:45:00 AM	1973.3	24.28	0	22.6	441
4	6/22/2013	7	5:15:00 PM	2157	24.96	0	27	264
5	6/23/2013	1	11:45:00 AM	2176	24.08	0	26.9	889
6	6/23/2013	1	6:30:00 PM	2151.69	24.76	0	25.8	44
7	6/30/2013	1	8:45:00 AM	2003.4	23.92	0	19.8	485

Tabella 37: Outlier comuni individuati dai tre metodi analizzati (CART(ESD) – K-Means (ESD) – DBSCAN) per la potenza elettrica attiva totale (edificio F70)

S.No	DATE	Id DAY	HOUR	Total electric active power F70	Indoor temperature F70	People presence F70	Outdoor temperature	Global solar radiation
1	5/22/2013	4	10:15:00 AM	8814.2	20.66	8	16.8	638
2	5/27/2013	2	11:00:00 AM	9312.7	20.05	8	17.6	853
3	5/27/2013	2	11:15:00 AM	9328.59	20.11	8	18	918
4	6/3/2013	2	11:15:00 AM	8749.5	20.38	8	18.9	1042

Tabella 38: Outlier comuni individuati dai tre metodi analizzati (CART(ESD) – K-Means (ESD) – DBSCAN) per la potenza elettrica attiva totale (edificio F71)

S.No	DATE	Id DAY	HOUR	Total electric active power F71	Indoor temperature F71	People presence F71	Outdoor temperature	Global solar radiation
1	5/14/2013	3	7:30:00 AM	3169	18.87	2	12.4	26
2	5/23/2013	5	6:45:00 AM	4066.6	21.79	0	13.1	23
3	5/23/2013	5	7:00:00 AM	4238.7	21.15	0	13.2	32
4	5/31/2013	6	6:45:00 AM	3205.3	20.42	0	9	26
5	5/31/2013	6	7:00:00 AM	4292	20.59	0	9.3	51
6	5/31/2013	6	7:15:00 AM	4433.1	20.23	0	9.8	74
7	5/31/2013	6	7:30:00 AM	4124.5	19.5	2	10.4	127
8	6/11/2013	3	7:15:00 AM	3361.19	23.35	1	16.1	33
9	6/13/2013	5	7:15:00 AM	3561.69	23.89	0	19.4	42
10	6/21/2013	6	7:30:00 AM	3633.19	24.45	1	20.4	60
11	6/21/2013	6	3:15:00 PM	918.4	25.64	6	29.3	747
12	6/27/2013	5	7:30:00 AM	3271.8	20.01	2	17.1	35

Tabella 39: Outlier comuni individuati dai tre metodi analizzati (CART(ESD) – K-Means (ESD) – DBSCAN) per la potenza elettrica attiva totale (edificio F72)

S.No	DATE	Id DAY	HOUR	Total electric active power F72	Indoor temperature F72	People presence F72	Outdoor temperature	Global solar radiation
1	5/16/2013	5	11:45:00 AM	5472.5	22.57	11	21.4	564
2	5/23/2013	5	7:30:00 AM	4664.29	21.9	0	13.2	83
3	5/23/2013	5	7:45:00 AM	4162	21.88	0	13.2	95
4	5/23/2013	5	3:00:00 PM	6559.5	22.73	10	16.8	206
5	5/28/2013	3	7:00:00 AM	4323.5	20.17	0	13.2	53
6	5/28/2013	3	7:15:00 AM	4135.7	20.05	0	13.4	130
7	6/28/2013	6	7:45:00 AM	4215.39	21.85	1	17	74

Tabella 40: Outlier comuni individuati dai tre metodi analizzati (CART(ESD) – K-Means (ESD) – DBSCAN) per la potenza elettrica attiva totale (edificio F73)

S.No	DATE	Id DAY	HOUR	Total electric active power F73	Indoor temperature F73	People presence F73	Outdoor temperature	Global solar radiation
1	5/24/2013	6	9:00:00 AM	5557.29	23.32	7	16.6	533
2	5/29/2013	4	7:30:00 AM	6035.79	21.33	1	14.1	76
3	5/29/2013	4	8:15:00 AM	5681.29	21.49	4	15.6	152
4	5/29/2013	4	8:30:00 AM	5930.89	21.84	5	16	114
5	6/28/2013	6	7:30:00 AM	5257.29	21.08	1	16.5	70
6	6/29/2013	7	12:45:00 AM	2520.89	22.38	0	15.8	0
7	6/29/2013	7	8:00:00 AM	2538.8	21.99	0	16.1	352
8	6/30/2013	1	3:45:00 PM	2729.69	24.79	0	27	715

Verifica e applicazione di tecniche di “neural ensembling” per analisi di fault detection

La ricerca relativa a questa attività ha previsto l’applicazione di tecniche di “neural ensembling” e di modelli statistici per l’individuazione automatica di valori anomali di consumo degli edifici e la conseguente attivazione del processo diagnostico. In particolare l’analisi diagnostica avanzata è finalizzata al risparmio energetico nell’ambito dell’edificio F40 e del distretto di edifici del Centro Ricerche ENEA Casaccia.

Sono state adoperate due tecniche di ensembling di reti neurali artificiali (RNA) ai fini della fault detection dell’uso finale illuminazione sulla base di dati di monitoraggio relativi all’edificio F40 o a una sua porzione (singolo piano) e al distretto di edifici. Un ensembling consiste in un insieme di modelli (in questo caso RNA) che lavorano sullo stesso problema in modo da ottenere risultati più robusti rispetto a quelli dei singoli modelli.

Per poter fornire ai modelli neurali una base di conoscenza adeguata è stato utilizzato un set di dati opportunamente elaborati provenienti dall’attività di monitoraggio svolta sugli edifici di interesse. Tali dataset contengono informazioni concernenti:

- consumo di potenza di luce di edificio/piano su base oraria/quindici minuti;
- numero di presenze di edificio/piano su base oraria/quindici minuti;
- solo per l’edificio F40, numero di stanze attive di piano su base oraria (una stanza è stata considerata “attiva” se almeno una persona è presente al suo interno);
- radiazione solare globale su base oraria/quindici minuti;
- ora e giorno della settimana.

Basic Ensemble Method di RNA per analisi di fault detection

Di seguito si presenta lo sviluppo di modelli basati su reti neurali, finalizzato all’individuazione della tipologia di modelli in grado di rappresentare con l’errore più basso possibile il comportamento energetico degli edifici. Le RNA sono utilizzate in questo caso come stimatori per la modellazione del consumo di illuminazione degli edifici sulla base delle forzanti climatiche e di occupazione.

I modelli neurali sono costruiti utilizzando solo una parte dei dati disponibili e vengono applicati alla parte rimanente dei dataset non utilizzata per la calibrazione o training. In particolare, il training delle RNA è eseguito su dataset privi di fault in modo da condurre in seguito un’analisi di fault detection nei periodi di testing basata sul valore dei residui tra valore di consumo stimato dall’ensembling di RNA e valore di consumo reale.

La tipologia di Rete Neurale utilizzata è feed-forward ad uno strato nascosto, con un numero di input e output pari al numero delle variabili considerate di volta in volta in ingresso e in uscita e 15 neuroni interni. Le simulazioni sono eseguite in ambiente MATLAB (R2012b) sfruttando le function messe a disposizione dal Neural Networks Toolbox. Le funzioni di attivazione per il livello nascosto e per il livello di uscita sono quelle di default, tangente iperbolica (tansig) e lineare (purelin) rispettivamente.

Per ogni prova effettuata sono stati eseguiti 10 run (training e testing) con le RNA appena descritte. Gli output delle 10 RNA sono stati mediati per ottenere l'output complessivo del modello secondo uno schema Basic Ensemble Method (BEM) che effettua una media aritmetica degli output delle singole RNA. Quindi si ha che:

$$Y = \frac{1}{10} \sum_{i=1}^{10} y_i$$

dove Y è l'output dell'Ensemble, mentre y_i sono gli output delle RNA.

I criteri adoperati per valutare le prestazioni dei modelli sono l'errore medio assoluto (*MAE*, mean absolute error) e l'errore massimo assoluto (*MAX*, maximum absolute error). In formule:

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i|$$

$$MAX = \max \left\{ |y_i - \hat{y}_i| \right\}_{i=1}^N$$

dove:

- y_i : consumo elettrico reale all'ora i ;
- \hat{y}_i : consumo elettrico stimato all'ora i ;
- $|y_i - \hat{y}_i|$: valore assoluto della differenza tra consumo elettrico reale e consumo elettrico stimato all'ora i (errore di stima assoluto);
- N : numero dei campioni orari di consumo.

Si sottolinea che la dinamica dell'edificio (il consumo stesso) non viene considerata tra i dati utilizzati in input: ciò permette di valutare l'influenza degli aspetti connessi all'utenza e ai fattori climatici sull'andamento dei consumi elettrici, rispondendo alle finalità di diagnostica avanzata per cui viene adoperato questo tipo di modellazione.

Al termine della fase di modellazione, si procede al calcolo dei residui tra valore di consumo stimato dal BEM e valore di consumo reale. Un fault coincide con la presenza di un picco nella serie temporale dei residui, la cui severità è quantificata da indici calcolati a partire dalla procedura GESD e dalla procedura Peak Detection Method, già incontrate nella precedente sezione.

Majority Voting Ensemble Method di RNA per analisi di fault detection

Nella seconda fase dell'analisi di fault detection, le RNA sono state utilizzate non più come stimatori dei valori di consumo ma come classificatori delle condizioni operative del sistema edificio. Al fine di sviluppare un ensembling di RNA per la classificazione delle situazioni di esercizio, il training delle RNA è stato realizzato utilizzando dati rappresentativi di condizioni operative dell'edificio sia normali che anomale. Una volta allenate le reti neurali, il processo di fault detection and diagnosis diventa quindi un task di pattern recognition.

Gli input delle RNA sono pertanto un insieme di caratteristiche che definisce l'output della rete, ossia lo stato operativo dell'edificio, come "normale" o "anomalo". Oltre al numero delle presenze, al numero delle stanze attive, alle informazioni su ora e tipo di giorno, ai dati climatici, per il classificatore neurale anche i valori di consumo illuminazione sono dati in input. Le caratteristiche e il numero di test delle RNA di cui fare l'ensembling sono uguali al caso precedente. Per codificare i dati in output, è stato adottato il seguente pattern numerico: valore numerico 0 per le situazioni normali, valore numerico 1 per le situazioni anomale. Gli output dei modelli neurali sono poi combinati secondo lo schema del Majority Voting Ensemble Method (MVEM): questo schema associa l'uscita complessiva del sistema (0 o 1) alla classe sulla quale converge almeno la metà dei singoli modelli.

[Risultati dell'analisi di fault detection per l'edificio F40 con tecniche di neural ensembling](#)

Per l'analisi di fault detection dell'edificio F40 si fa riferimento a un dataset orario di circa due mesi (Dicembre 2012 – Gennaio 2013). La modellazione del consumo illuminazione effettuata con l'ensembling neurale riguarda l'energia attiva e la potenza attiva massima del primo piano dell'edificio, sebbene successivamente per l'analisi di fault detection si faccia riferimento solo alla potenza attiva massima perché più rappresentativa delle variazioni di consumo. La Tabella seguente illustra la qualità della stima ottenuta, valutata secondo il MAE e il MAX descritti in precedenza.

Tabella 41: Risultati sperimentali di stima dei consumi (training e testing)

	TRAINING	RNA	BEM	TESTING	RNA	BEM
ACTIVE ENERGY	MAE (kWh)	0.33 (± 0.01)	0.31	MAE (kWh)	0.66 (± 0.04)	0.63
	MAX (kWh)	1.41	1.06	MAX (kWh)	4.26	3.75
MAXIMUM ACTIVE POWER	MAE (kW)	0.35 (± 0.02)	0.33	MAE (kW)	0.81 (± 0.05)	0.78
	MAX (kW)	1.76	1.40	MAX (kW)	4.78	4.51

Si può osservare come i risultati ottenuti con il BEM neurale siano sempre migliori di quelli ottenuti con le singole reti. Come mostrato nelle seguenti figure, nel periodo di training il BEM neurale stima con elevata accuratezza la potenza attiva massima; nel periodo di testing la potenza stimata segue piuttosto bene quella reale, ad eccezione di alcuni evidenti valori anomali.

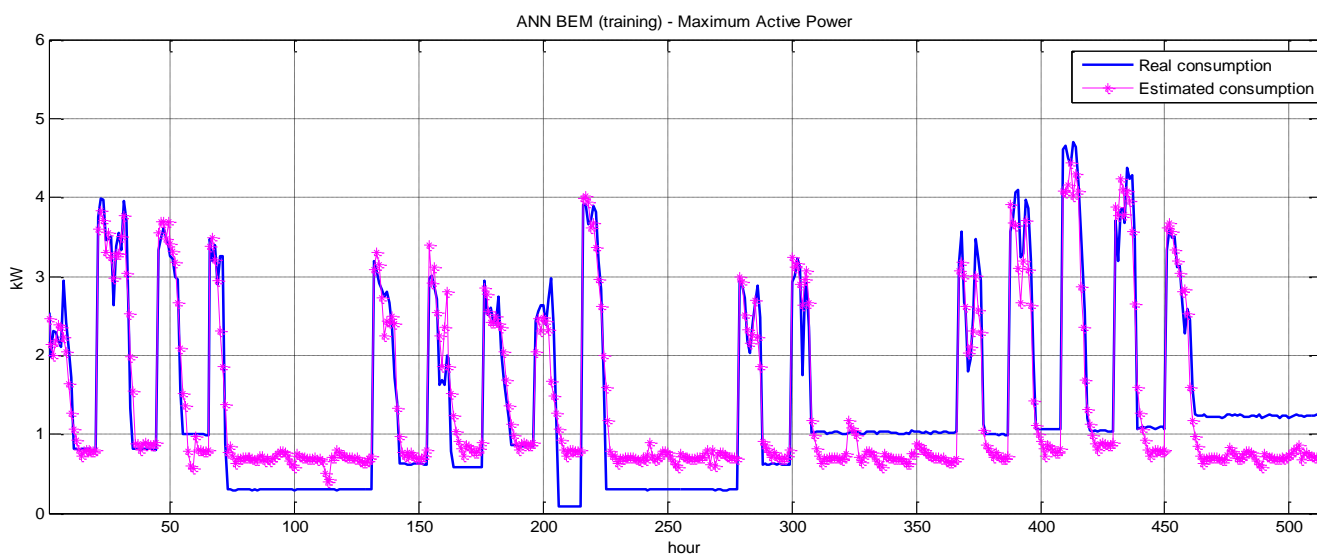


Figura 87: RNA BEM (training) – Maximum active power

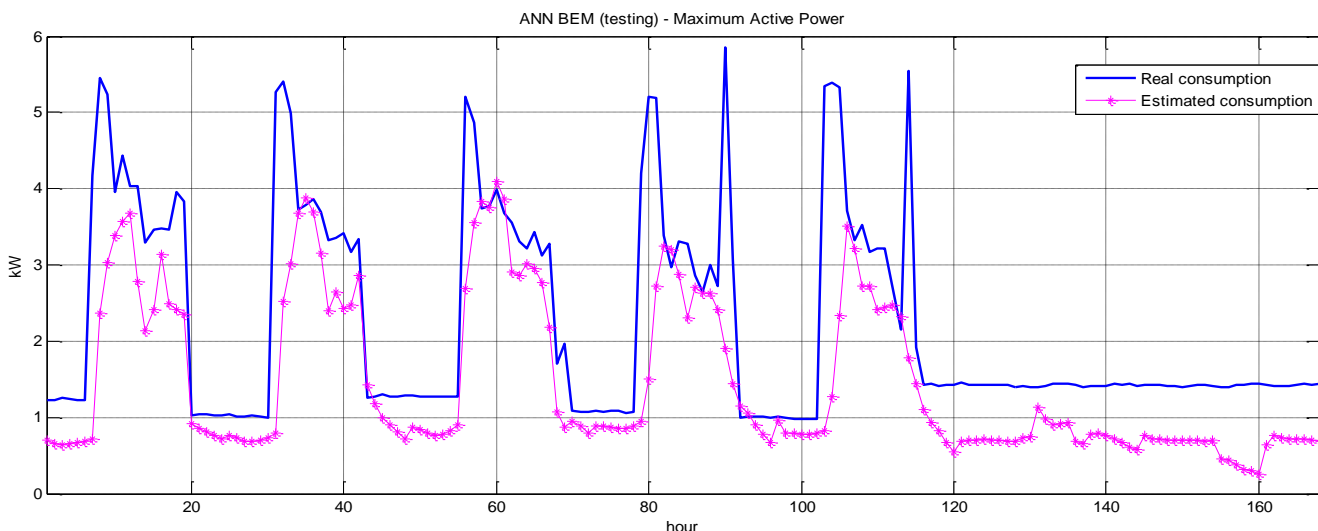


Figura 88: RNA BEM (testing) – Maximum active power

Per l’analisi di fault detection è stata analizzata la differenza tra i valori stimati e i valori reali (residui) di potenza attiva massima. In particolare, il Peak Detection Method è stato applicato alla serie dei residui nel periodo di testing. Al fine di testare la validità delle tecniche di fault detection proposte, nella settimana di testing sono stati generati due fault: in due giorni diversi, al termine dell’orario di lavoro (tra le 17:30 e le 18:00) e quindi in condizioni di bassa presenza, tutte le luci degli uffici del primo piano sono state accese. Come illustrato nella seguente figura, il Peak Detection Method applicato alla serie dei residui è in grado di individuare i due fault artificiali appena descritti e un insieme di altri fault corrispondenti al consumo verificatosi nelle prime ore mattutine quando all’interno dell’edificio è presente unicamente il personale di pulizia.

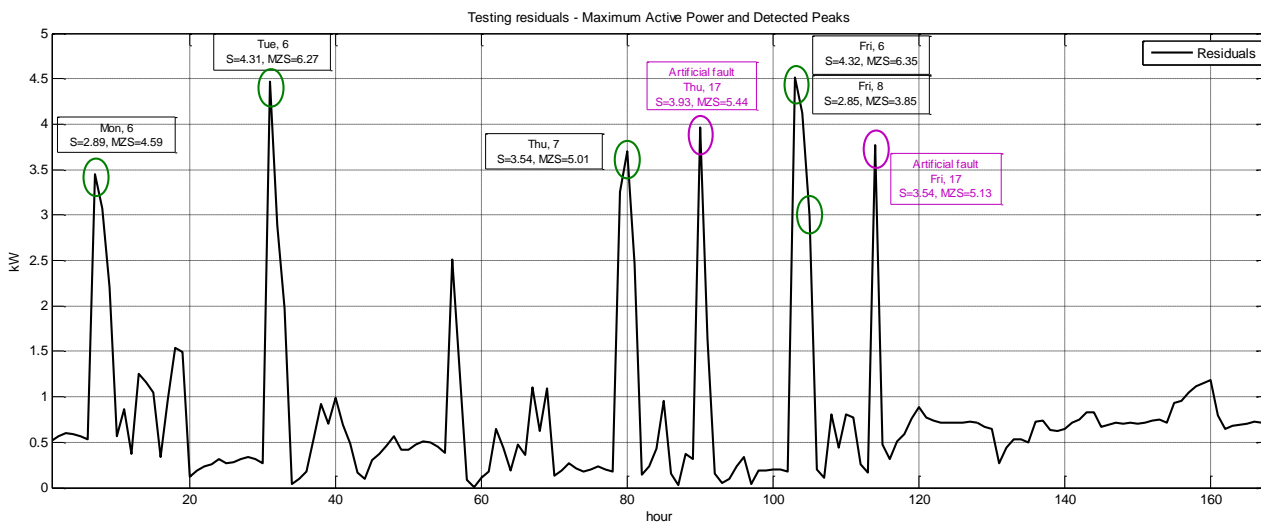


Figura 89: Picchi individuati sui residui di testing

I risultati confermano che l’analisi dei residui generati attraverso il BEM neurale rappresenta una tecnica efficace per la fault detection dei valori anomali di consumo di illuminazione dell’edificio.

Il Peak Detection Method è stato poi applicato alla serie dai dati di consumo di potenza attiva massima. Nella figura seguente sono illustrati gli outlier individuati nel periodo di testing con i relativi valori degli indici S Function e MZ Score. Si osserva che vengono nuovamente individuati i due fault artificiali e un insieme di fault che avvengono nelle prime ore mattutine: in tutti questi casi gli indici assumono correttamente valori più elevati. Tuttavia, poiché i dati di consumo sono correlati a variabili quali le

presenze, le stanze attive, la radiazione solare e il tipo di giorno, quando si effettua un'analisi di fault detection univariata, che non tiene cioè conto delle variabili indipendenti come in questo caso, aumenta il rischio di individuare tra gli outlier diversi falsi positivi.

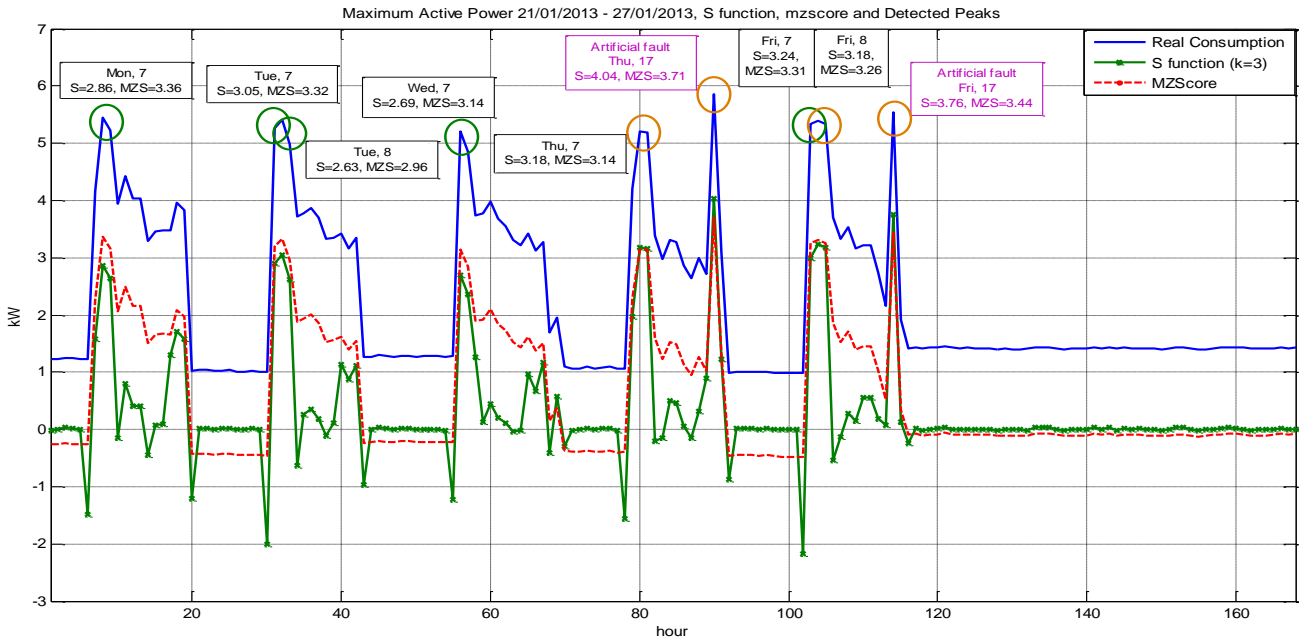


Figura 90: Maximum active power (periodo di testing), valori di S function, mzscore e picchi individuati (i picchi comuni sono evidenziati in arancione)

Un ulteriore step dell'analisi di fault detection relativa al primo piano dell'edificio F40 ha previsto lo sviluppo di un ensembling neurale per la classificazione del consumo di potenza attiva massima di illuminazione in "normale" o "anomalo". La tecnica di ensembling utilizzata è il Majority Voting e il classificatore neurale è stato costruito in base a come descritto nel precedente paragrafo. La figura seguente mostra come il classificatore neural ensembling messo a punto sia in grado di individuare i due fault artificiali così come altri valori anomali di consumo. Gli outlier comuni ai metodi analizzati sono evidenziati in verde.

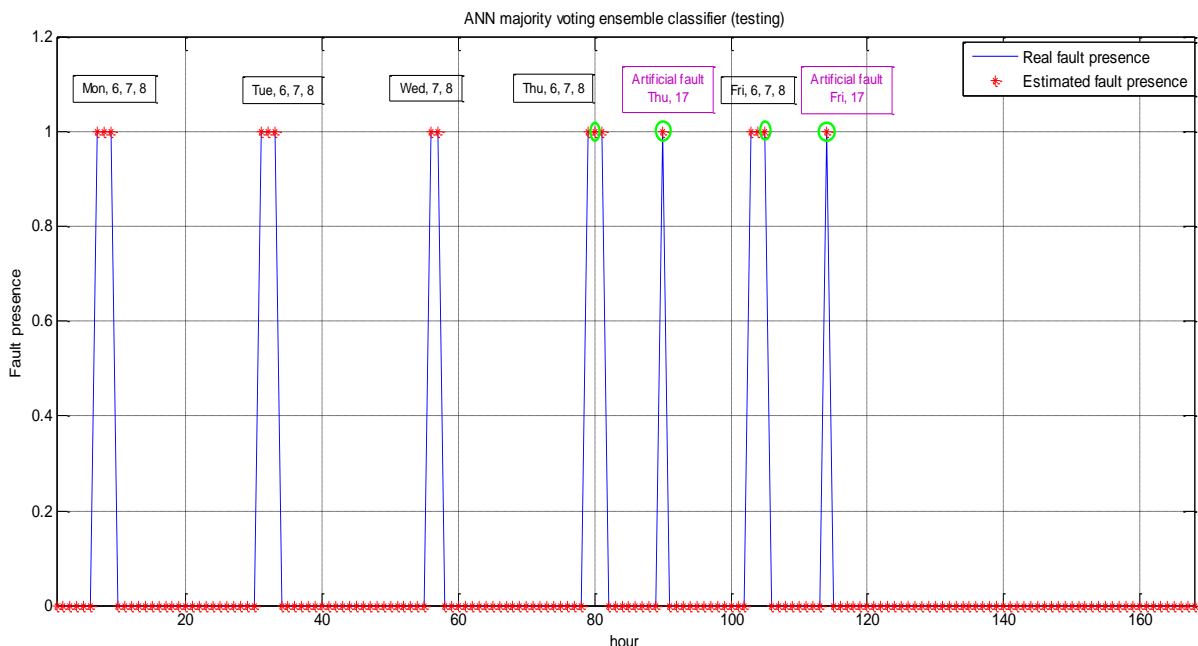


Figura 91: Classificatore RNA majority voting ensembling

Nella tabella seguente viene riportato il valore relativo percentuale dell'errore di classificazione: si osserva che anche in questo caso l'ensembling neurale offre prestazioni migliori rispetto alle singole reti.

Tabella 42: Errore percentuale di classificazione (training e testing)

TRAINING	PERCENTAGE ERROR (%)	TESTING	PERCENTAGE ERROR (%)
BEST ANN	0.00	BEST ANN	0.00
WORST ANN	0.35	WORST ANN	0.60
CLASSIFIER	0.00	CLASSIFIER	0.00

I risultati sperimentali dimostrano anche in questo caso l'elevata efficacia del classificatore neurale Majority Voting nell'individuazione delle anomalie di consumo illuminazione verificatesi nella settimana di testing.

Risultati dell' analisi di fault detection per il cluster o distretto di edifici con tecniche di neural ensembling

Così come già illustrato nell'analisi di fault detection sul distretto di edifici basata su tecniche di data mining, anche nel caso della fault detection basata su tecniche di "neural ensembling" la maggior parte degli outlier individuati nel consumo elettrico illuminazione delle 8 palazzine si verifica nelle prime ore del mattino. Di nuovo, ciò è da ricondurre all'accensione delle luci di edificio al mattino presto effettuata dal personale di pulizia che non risulta peraltro tra i dati di presenza.

L'analisi di fault detection è stata portata avanti applicando i metodi Peak Detection e GESD alle seguenti serie dati relative a ogni edificio:

- Residui di potenza attiva illuminazione calcolati tramite l'applicazione del BEM neurale;
- Residui di potenza attiva totale - potenza attiva illuminazione calcolati tramite l'applicazione del BEM neurale;
- Valori di consumo di potenza attiva illuminazione;
- Valori di consumo di potenza attiva totale - potenza attiva illuminazione.

Gli outlier individuati assumono diversi significati a seconda del criterio applicato. Gli outlier relativi alla serie dei residui sono, infatti, valori anomali di consumo sulla base del numero di occupanti, delle condizioni meteorologiche, dell'ora e del tipo di giorno; quelli individuati a partire dai soli dati di consumo, invece, tengono conto solo di brusche variazioni che si registrano nei valori di consumo stesso. Il confronto dei risultati così ottenuti e l'utilizzo degli indicatori adimensionali discussi nelle precedenti sezioni hanno permesso di quantificare la severità e la significatività degli outlier individuati.

Per l'analisi di fault detection del distretto di edifici si fa riferimento a dati misurati ogni quindici minuti per un periodo di circa due mesi (Maggio 2013 – Giugno 2013). Di seguito si presenta qualche esempio significativo di outlier rilevati per diversi edifici.

Edificio F66.

Nella figura seguente si illustra l'individuazione di un picco di potenza di consumo illuminazione (4695.2 W) avvenuto il 16 Maggio alle 07.45 del mattino.

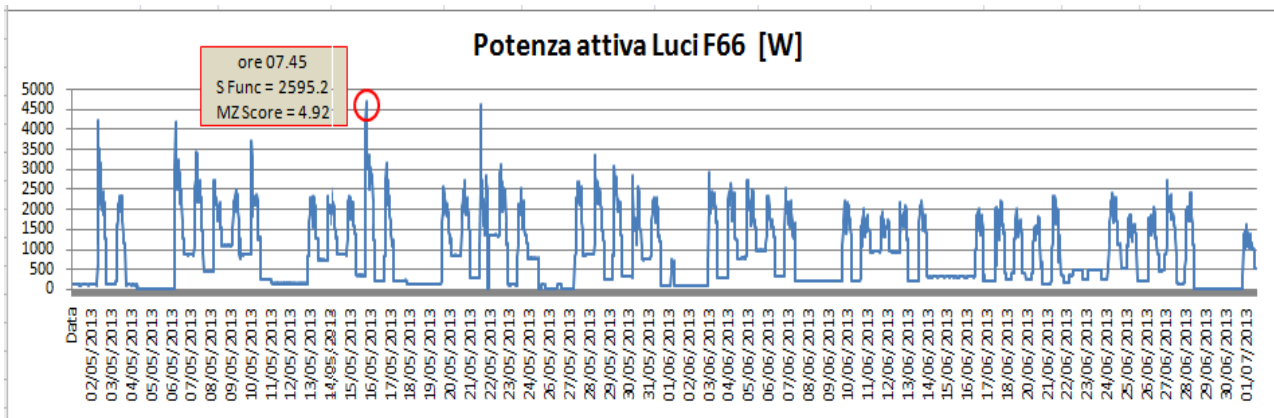


Figura 92: Edificio F66: esempio di outlier individuato su consumo di potenza attiva illuminazione

A seguire si illustra invece l'individuazione di un picco nei residui di potenza di illuminazione avvenuto il 26 Giugno alle 07.30 del mattino, corrispondente a un valore consumo troppo basso (200 W) rispetto all'ora e alle presenze nell'edificio.

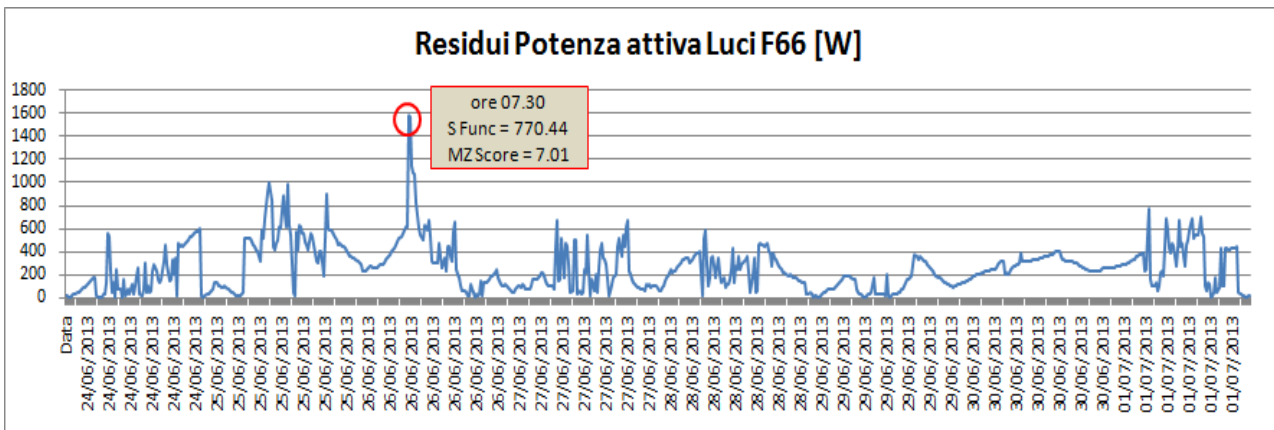


Figura 93: Edificio F66: esempio di outlier individuato sui residui di potenza attiva illuminazione

Edificio F68.

Nella figura seguente si illustra l'individuazione di un picco di potenza di consumo illuminazione (5170 W) avvenuto il 7 Maggio alle 07.00 del mattino.

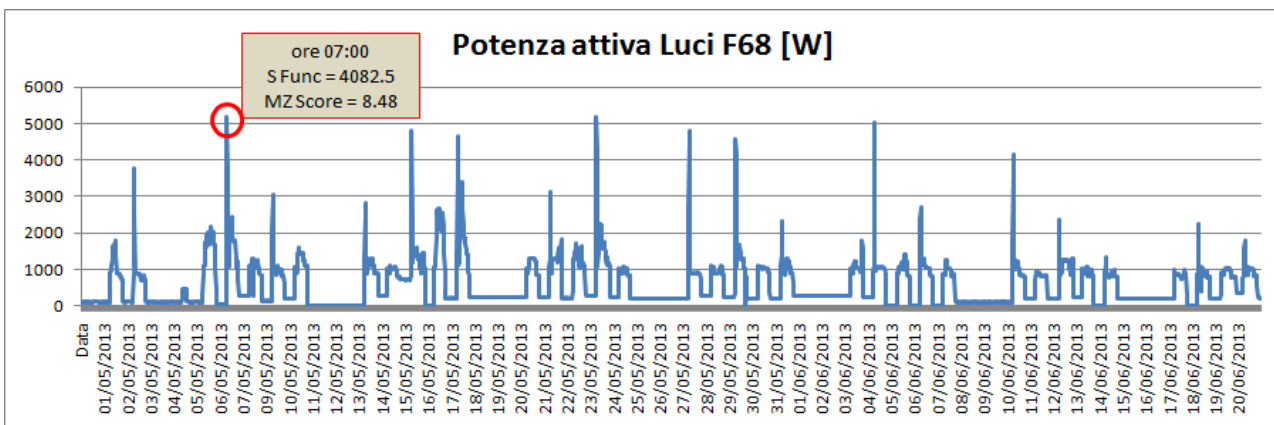


Figura 94: Edificio F68: esempio di outlier individuato su consumo di potenza attiva illuminazione

A seguire si illustra invece l'individuazione di un picco nei residui di potenza di illuminazione avvenuto il 12 Giugno alle 08:00 di sera, corrispondente a un valore consumo molto alto (1280 W) rispetto all'ora ed a una sola presenza nell'edificio.

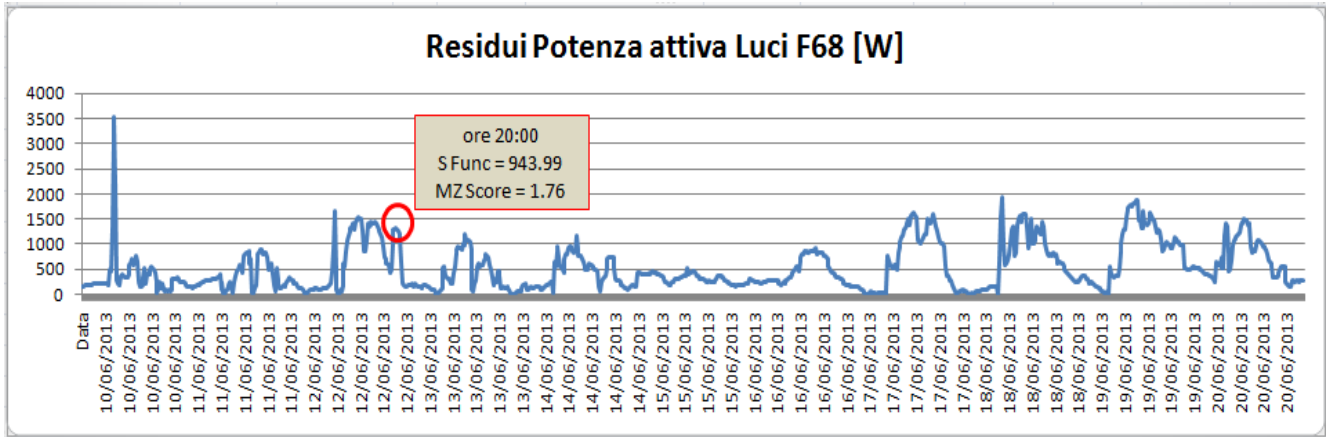


Figura 95: Edificio F68: esempio di outlier individuato sui residui di potenza attiva illuminazione

Individuazione di outlier comuni individuati attraverso metodi di data mining e BEM neurali

Nella presente sezione sono riassunti, attraverso opportune rappresentazioni grafiche, i principali fault identificati attraverso i modelli di data mining e attraverso l'applicazione di metodi di outlier detection ai residui ottenuti dal BEM neurale per alcuni edifici appartenenti al cluster nel periodo Maggio-Giugno. Gli outlier di consumo individuati da più metodi corrispondono con maggiore probabilità a valori di consumo realmente anomali, permettendo quindi quanto più possibile di non considerare dei falsi positivi.

In particolare nella figura seguente sono evidenziati per l'edificio F66 gli outlier comuni individuati attraverso l'applicazione del metodo GESD a un cluster o una classe individuata attraverso il metodo CART, con il metodo DBSCAN (cluster 0) e con il metodo del peak detection all'intero set di dati.

Per gli altri edifici, invece, partendo da una classe di dati creata attraverso il metodo CART, sono evidenziati i valori anomali comuni identificati attraverso l'applicazione del metodo GESD alla classe di dati, a un cluster di dati (K-Means), ai residui ottenuti attraverso il BEM neurale e attraverso il cluster 0 dell'algoritmo DBSCAN.

F66(May-Jun)common outliers in CART and Peak Detection Model (PDM) (class 4)

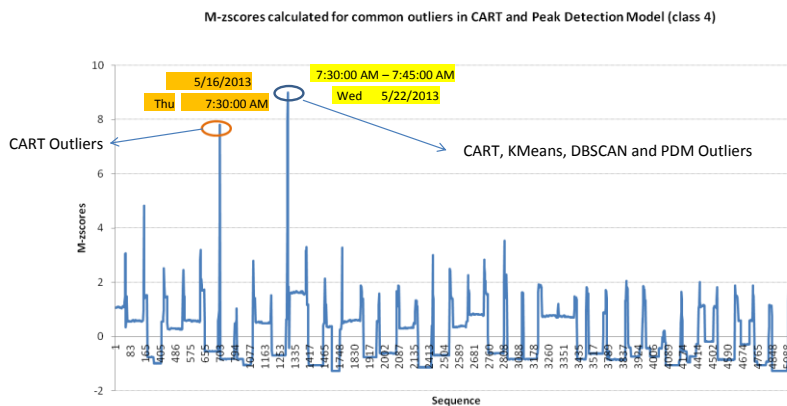


Figura 96: Variazione dell'indice Mzscore in un classe di dati (classificazione con metodo CART) con evidenza dei principali outliers comuni individuati dal metodo peak detection e CART(ESD) (edificio F66)

F68(May-Jun)Common Outliers CART & Residuals NN (class1)

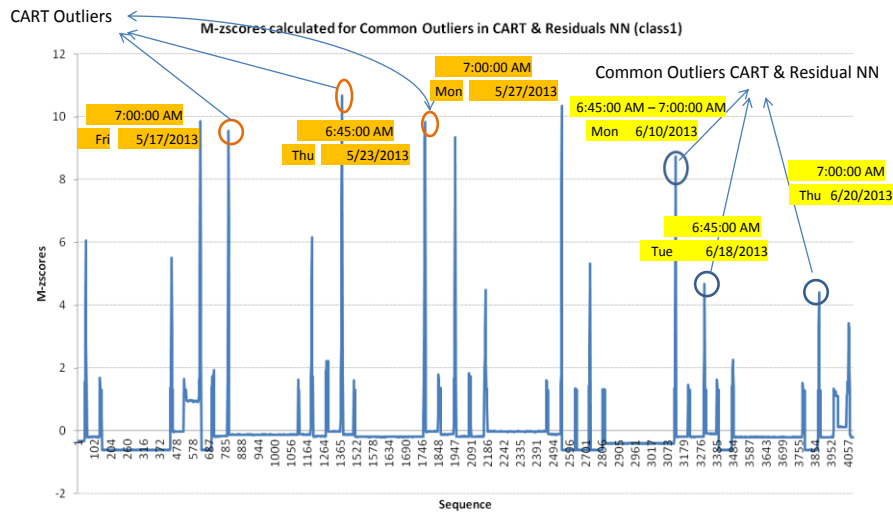


Figura 97: Variazione dell'indice Mzscore in un classe di dati (classificazione con metodo CART) con evidenza dei principali outliers comuni individuati dal metodo NN residuals (GESD) e CART(ESD) (edificio F68)

F69(May-Jun)Common Outliers in CART & Residuals NN (class7)

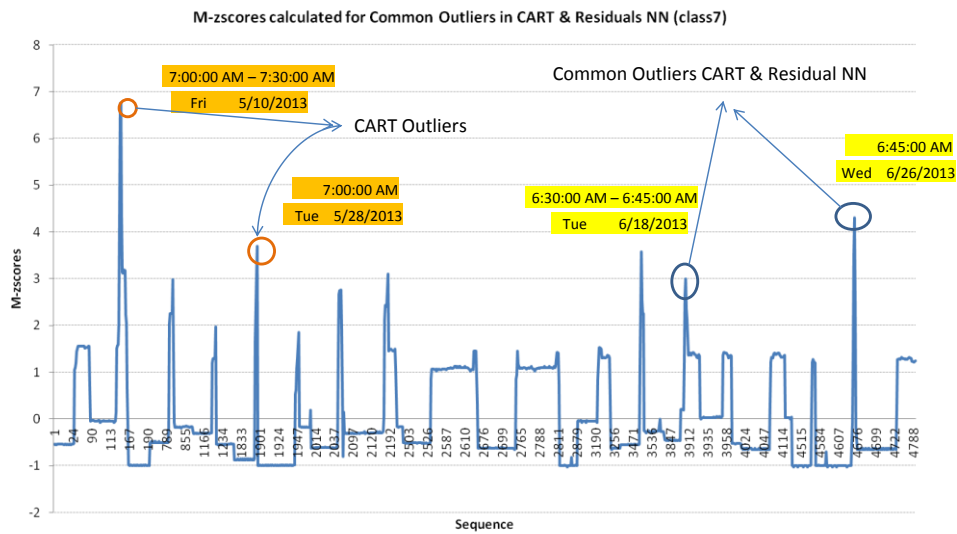


Figura 98: Variazione dell'indice Mzscore in un classe di dati (classificazione con metodo CART) con evidenza dei principali outliers comuni individuati dal metodo NN residuals (GESD) e CART(ESD) (edificio F69)

F70(May-Jun)Common Outliers in CART & Residuals NN (class8)

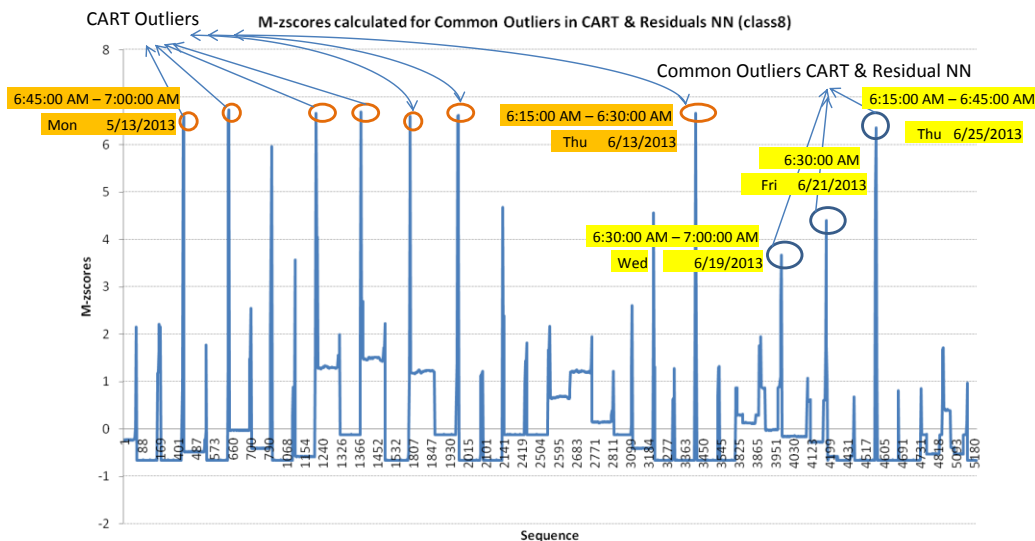


Figura 99: Variazione dell'indice Mzscore in un classe di dati (classificazione con metodo CART) con evidenza dei principali outlier comuni individuati dal metodo NN residuals (GESD) e CART(ESD) (edificio F70)

F73(May-Jun)Common Outliers in CART & Residuals NN (class7)

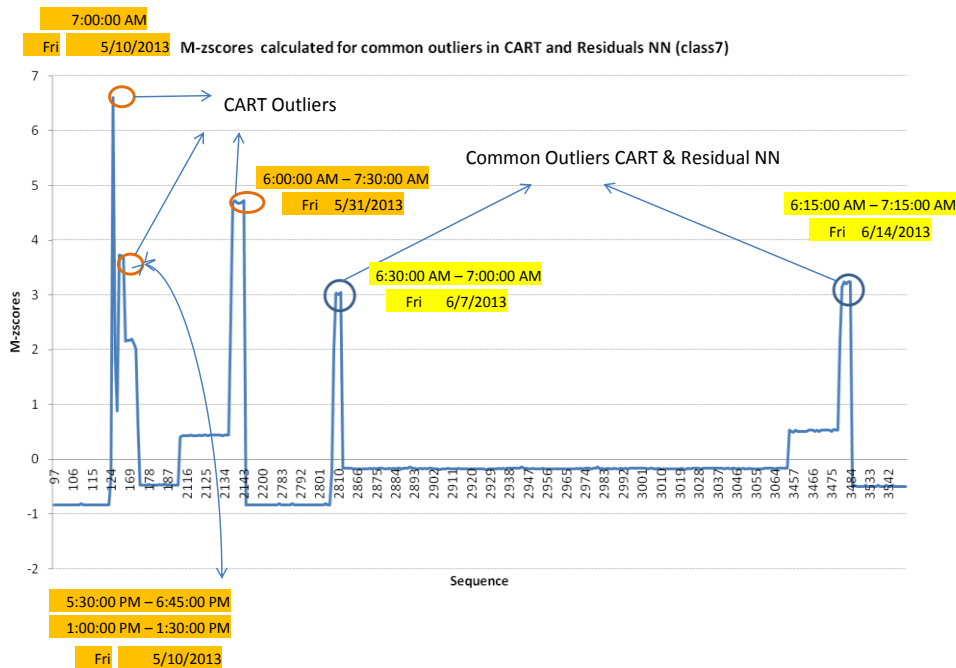


Figura 100: Variazione dell'indice Mzscore in un classe di dati (classificazione con metodo CART) con evidenza dei principali outlier comuni individuati dal metodo NN residuals (GESD) e CART(ESD) (edificio F73)

Analisi diagnostica relativa all'utilizzo illuminazione artificiale basata sulla disponibilità di luce naturale

Nell'ambito del risparmio energetico legato all'illuminazione uno degli aspetti che è opportuno considerare con attenzione e che presenta generalmente un forte impatto sul bilancio energetico di un edificio è lo sfruttamento della luce naturale. Sulla base di questa considerazione l'attività di ricerca condotta ha previsto l'analisi e la determinazione numerica per alcuni uffici dell'edificio F40 della disponibilità di luce naturale, al fine di individuare attraverso robuste indagini diagnostiche opportunità legate al suo corretto

sfruttamento. Il fine dell'analisi diagnostica è quindi la razionalizzazione della richiesta energetica legata all'illuminazione artificiale attraverso l'analisi della disponibilità di luce naturale (daylight harvesting).

Lo studio ha comportato dapprima un'approfondita analisi delle caratteristiche influenzanti le prestazioni illuminotecniche dell'edificio oggetto di analisi quali l'esposizione, la disposizione, le caratteristiche dei punti luce, i fattori geometrici, la presenza di ostacoli, le caratteristiche dei rivestimenti dei pavimenti, muri o soffitti, l'estensione e le proprietà delle superfici vetrate. Al fine di condurre un'analisi diagnostica sull'opportunità di sfruttamento della luce naturale, è stato costruito un modello numerico con il codice Daysim. Attraverso il codice è stata simulato in regime dinamico il livello di illuminamento interno garantito attraverso il daylighting considerando un'opportuna griglia di punti per due uffici tipologici caratterizzati da una diversa esposizione. Al fine di ottenere un modello calibrato sono state condotte alcune misure in campo per la determinazione di alcune importanti variabili in ingresso al modello (per esempio coefficienti di riflessione e trasmissione delle superfici vetrate).

Sulla base dei risultati delle simulazioni condotte, lo studio si è focalizzato sulla costruzione di un modello inverso del tipo black box che permettesse di stimare il valore di illuminamento medio interno utile note le condizioni climatiche esterne e le caratteristiche dell'ambiente oggetto di studio.

L'edificio di riferimento è anche in questo caso l'edificio F40. La tecnica utilizzata per la modellazione dell'illuminamento è stata il Basic Ensemble Method di reti neurali. In particolare, nell'esperimento sono stati utilizzati i seguenti dati: come input dei modelli neurali sono stati considerati i valori di radiazione solare globale, altezza solare (misurati attraverso la centralina meteo installata nel Centro) e ora del giorno corrispondente; come output del modello è stata considerata la percentuale di punti nella griglia considerata con valori di illuminamento maggiori di 500 lux per due diverse stanze dell'edificio. I dati sono stati simulati per un intero anno ma lo studio è stato suddiviso in 3 fasi temporali: Gennaio – Aprile, Maggio – Agosto, Settembre – Dicembre. Ogni dataset è stato suddiviso in due parti: la prima, contenente circa il doppio dei dati della seconda, è stata utilizzata per la calibrazione (training) dei modelli, la seconda invece per valutarne le performance (testing). Per ogni periodo di studio sono state realizzate 10 prove con le reti neurali e i calcoli relativi all'ensemble sono stati svolti a partire dalla combinazione (media aritmetica) degli output di queste 10 reti.

I criteri adoperati per valutare le prestazioni dei modelli sono l'errore medio assoluto (MAE), l'errore massimo assoluto (MAX), entrambi descritti nei precedenti paragrafi, e la radice dell'errore quadratico medio (Root Mean Square Error, RMSE). In formule:

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{N}}$$

Di seguito sono illustrati i risultati di stima ottenuti in relazione al periodo Maggio – Agosto.

Tabella 43: Risultati sperimentali (training e testing)

TRAINING	ANN	BEM	TESTING	ANN	BEM
MAE [% punti]	5.65 (±0.24)	5.46	MAE [% punti]	6.94 (±0.28)	5.77
MAX [% punti]	57.54	47.48	MAX [% punti]	81.46	44.29
RMSE [% punti]	6.96	6.33	RMSE [% punti]	8.68	7.45

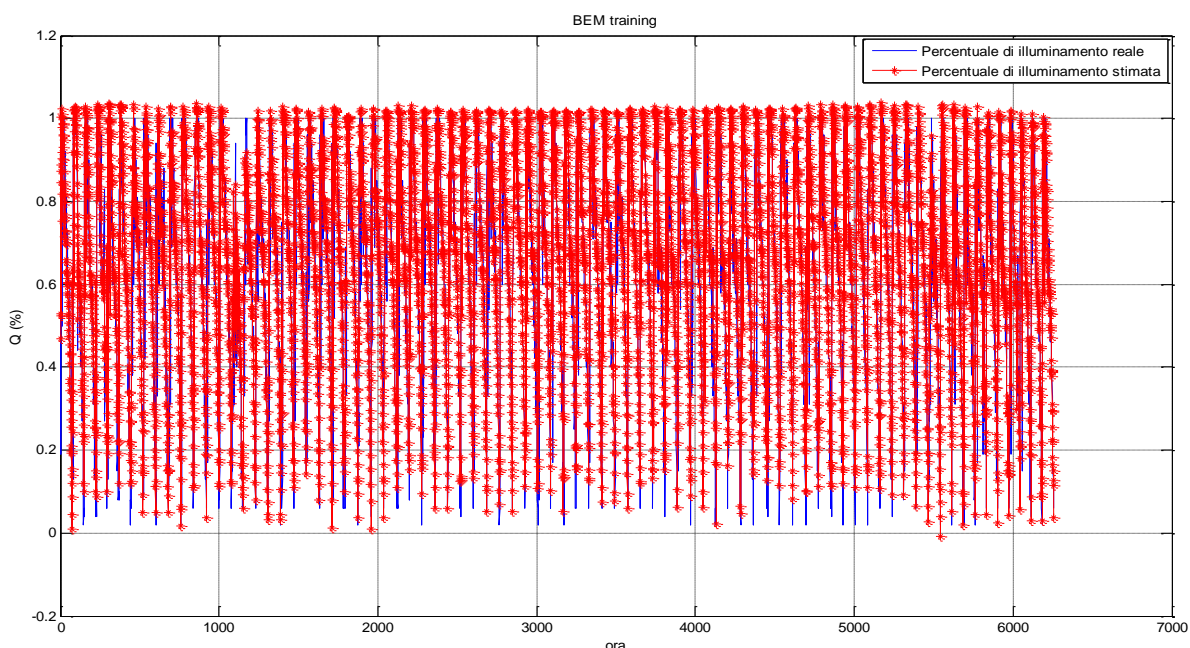


Figura 101: RNA BEM (training) – Percentuale di punti con valori di illuminamento maggiori di 500 lux

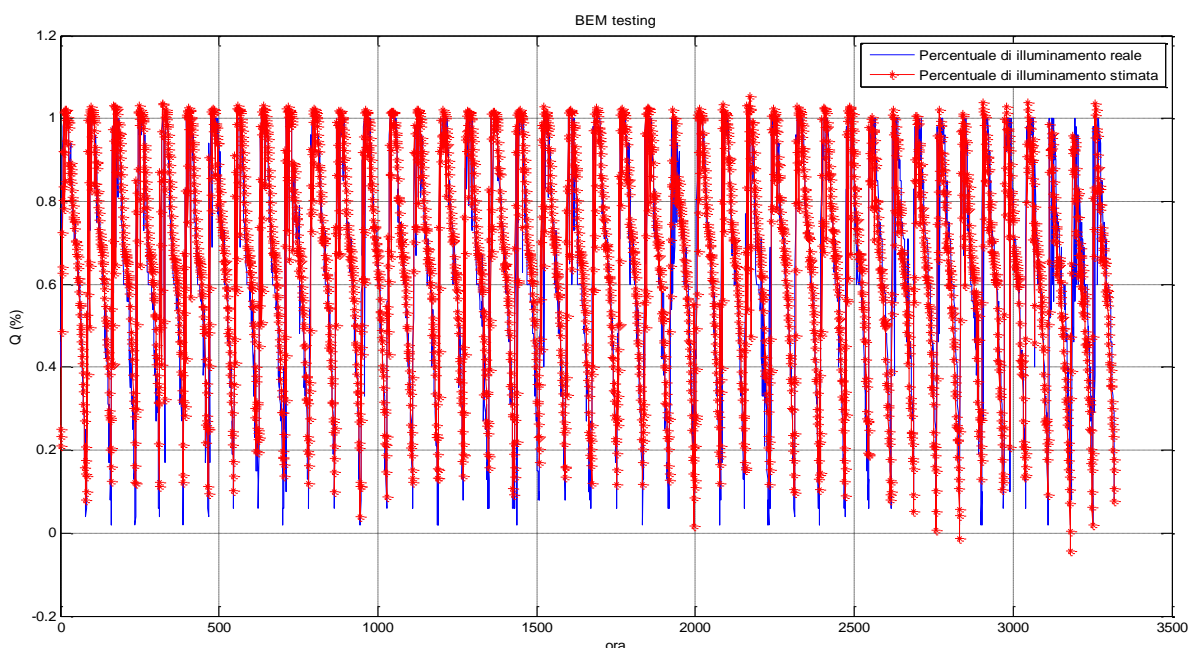


Figura 102: ANN BEM (testing) – Percentuale di punti con valori di illuminamento maggiori di 500 lux

Le simulazioni hanno dimostrato una precisione della stima del modello neurale pari a circa il 7%. L’obiettivo dell’analisi è stato quello svolgere un’analisi FDD mirata ad individuare gli utilizzi “anomali” di luce artificiale in condizioni di illuminamento interno naturale sufficiente. Il valore di illuminamento medio interno è stato considerato sufficiente quando in almeno il 75% dei punti della griglia per ogni stanza l’illuminamento supera il valore di 500 lux. Per individuare il valore di soglia è stato considerato il posizionamento dei piani di lavoro rispetto alla griglia di calcolo considerata. Attraverso il modello stimativo del livello medio di illuminamento interno sviluppato, è risultata possibile l’implementazione di un’ulteriore indagine diagnostica avanzata. L’algoritmo di fault detection in questo caso si basa sul confronto del valore di illuminamento medio interno disponibile stimato con il valore di soglia individuato e la corrispondente analisi dell’energia richiesta per l’illuminazione artificiale rispetto a

quella teoricamente attesa, considerando inoltre la presenza di occupanti in ogni stanza. Attraverso tale analisi è quindi possibile individuare anomalie rispetto all'uso delle luci artificiali in ogni stanza che risultano accese in presenza di occupanti e in corrispondenza di livelli medi di illuminamento naturale sufficiente. Sulla base di questo algoritmo di fault detection è in corso di sviluppo un'ulteriore regola diagnostica attraverso la quale è possibile risalire alla causa "Accensione luci in presenza di illuminamento naturale sufficiente" attraverso il passaggio logico di pre-processing e situation assessment. Per i dettagli si rimanda al contributo ENEA.

Analisi di ottimizzazione

Come completamento all'attività di gestione del comportamento energetico dell'edificio F40, accanto all'applicazione di metodi di FDD sono state poste le basi per condurre un'analisi di ottimizzazione multi obiettivo con il fine di ottimizzare la richiesta energetica per la climatizzazione ambientale e le condizioni di comfort interno per l'edificio F40. Le variabili considerate per l'ottimizzazione sono relative alla gestione dei termostati di stanza e al settaggio dei parametri di impianto. In particolare le variabili di progetto considerate per l'ottimizzazione sono le temperature di setpoint interne dei termostati di stanza e la temperatura del fluido vettore di mandata proveniente dalla centrale termica e frigorifera. L'obiettivo è stato quello di individuare il fronte con le soluzioni di ottimo considerando il valore delle suddette temperature in grado di garantire il miglior compromesso tra la potenza termica da erogare/sottarre (e, quindi, i costi economici da sostenere) ed il comfort degli occupanti. Per condurre l'analisi di ottimizzazione è stato sviluppato un simulatore della prestazione energetica dell'edificio F40 in ambiente MATLAB/SIMULINK dall'Università Politecnica delle Marche (Figura 2), per i cui dettagli implementativi si rimanda al corrispondente contributo (RdS/2013/120).

Il simulatore considera un modello di edificio a 15 zone: in particolare, per ognuno dei 3 piani dell'edificio sono state considerate 4 zone (ottenute raggruppando più uffici) ed 1 zona per il corridoio. Ad esclusione dei corridoi, ogni zona prevede un certo numero di fancoil per un totale di 147 fancoil in tutto l'edificio. Per poter simulare il comportamento dell'edificio in modalità invernale e in modalità estiva, il simulatore adopera dei modelli di centrale termo frigorifera e di fancoil in grado di supportare i due diversi funzionamenti (riscaldamento e raffrescamento) attraverso la semplice variazione di un set di parametri e con riferimento agli opportuni dati climatici. Il simulatore prevede la modellazione del controllo di temperatura di zona tramite termostati: ogni termostato comanda la velocità di ventilazione dei fancoil di zona in modo da garantire una temperatura interna di zona pari a quella di set point considerando una banda di tolleranza di ± 0.5 °C. Le velocità del ventilatore di ogni fan coil nel modello variano tra 0 (convezione naturale, potenza termica immessa in ambiente pari al 15% della potenza massima) e 3 (corrispondente a 180 m³/h).

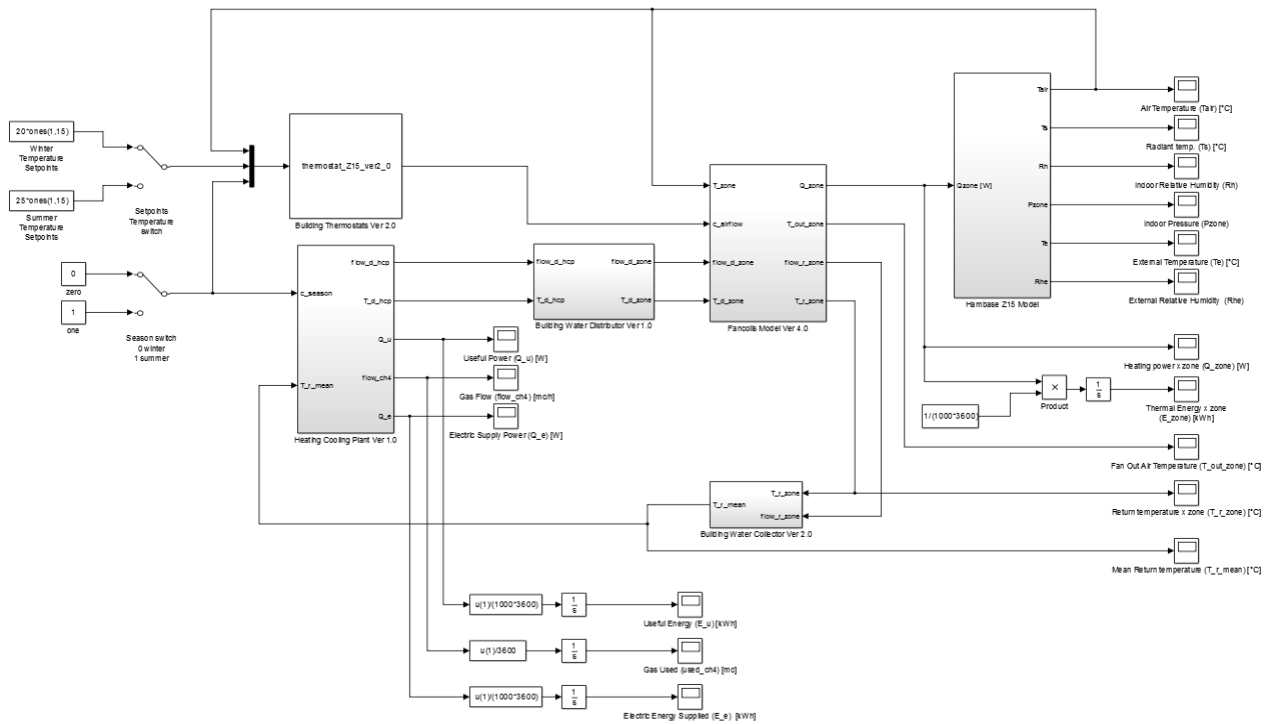


Figura 103: Simulatore edificio F40

Al fine di Individuare le temperature ottime di setpoint dei termostati di zona e del fluido vettore di mandata della centrale termo frigorifera considerando più funzioni obiettivo si è scelto di utilizzare tecniche basate su algoritmi evolutivi, ed in particolare algoritmi genetici. L’algoritmo genetico per ottimizzare il problema in questione è risultato agevole nell’implementazione a meno della complessità del calcolo della funzione obiettivo. Tale funzione infatti, essendo dipendente dai valori in uscita del simulatore, richiede un run del simulatore ad ogni accesso. Tale costo computazionale annullerebbe completamente i vantaggi degli algoritmi genetici che in genere richiedono una complessità costante della fitness. Tale problema è stato affrontato con un approccio di approssimazione della funzione di fitness. Le tecniche di fitness approximation (FA) consentono di diminuire il numero di richieste di calcolo della funzione di fitness utilizzando valori stimati oltre ai valori ottenuti dalla funzione originaria.

La funzione obiettivo considerata per tenere conto del comfort interno è basata sull’utilizzo dell’indice Predicted Mean Vote (PMV) introdotto nel modello di Fanger. Questo indice è universalmente riconosciuto anche da standard internazionali come un ottimo indicatore del livello di comfort interno e della sensazione percepita dell’ambiente termico da parte degli occupanti.

Un valore di PMV pari a zero corrisponde alla più bassa percentuale di insoddisfatti dell’ambiente termico in una popolazione statisticamente significativa. Un valore assoluto dell’indice PMV pari a 0.5 è generalmente considerato il limite accettabile per la zona di comfort. Al fine di ottimizzare il comfort interno sono stati considerati due diversi parametri per la funzione di fitness:

- Il valore medio assoluto dell’indice PMV per il periodo temporale selezionato e per le zone termiche occupate;
- Il numero di ore Ndis in cui il valore assoluto dell’indice PMV è superiore a 0.5. Esso rappresenta in altre parole il tempo in ore cumulato in cui si registrano condizioni di discomfort nel periodo considerato;

In particolare la funzione obiettivo relativa al comfort interno è stata costruita considerando il valore del parametro Ndis come termine di penalty al valore medio assoluto del PMV per considerare un vincolo sulle soluzioni di ottimo.

La funzione di fitness relativa alla prestazione energetica invece tiene conto della somma dell’energia primaria per l’intero edificio richiesta per il riscaldamento (nella stagione invernale) o per il raffrescamento

(nella stagione estiva) e dell'energia elettrica convertita in primaria richiesta dai ventilatori dei fan coil. Anche per questa funzione è stato considerato un vincolo attraverso il termine di penalty rappresentato dal parametro Ndis.

Il fine dell'analisi condotta è quello di individuare i valori della temperatura di set point interna e della temperatura di mandata dei fluidi termivettori giornalieri corrispondenti alle soluzioni individuate attraverso il fronte di Pareto che ottimizzano le funzioni di fitness individuate.

Conclusioni

Attività 1:

- L'attività ha affrontato il tema della definizione di una metodologia di analisi atta all'individuazione di un kit minimo di monitoraggio energetico ed ambientale. Tale kit minimo deve soddisfare due richieste fondamentali: la misura dei parametri energetici ed ambientali e la possibilità di controllare da remoto gli usi dell'energia attraverso delle attuazioni.
- Si è dimostrato, attraverso casi studio, come la definizione di un sistema di monitoraggio non possa prescindere da alcune operazioni preliminari, quali l'audit energetico ambientale e l'analisi delle informazioni raccolte. Tali operazioni sono infatti indispensabili al fine di adattare il kit minimo di monitoraggio al caso specifico.
- I casi studio presentati hanno due destinazioni d'uso diverse e sono rappresentativi della categoria a cui appartengono: eventuali variazioni ai sistema di monitoraggio presentati, dipenderanno dall'architettura dei sistemi edificio-impianti dei singoli casi specifici.

Attività 2:

- Il processo di verifica della qualità dei dati è un'operazione che deve essere ripetuta periodicamente e fintanto che l'intero sistema di monitoraggio non sarà ottimizzato e non risponderà a tutti i requisiti di partenza previsti dal piano.
- Non sempre dopo le correzioni effettuate al sistema di monitoraggio successivamente alla prima fase di analisi si riescono a risolvere tutte le anomalie, e soprattutto, non è detto che dalla seconda fase non ne emergano di nuove.
- Sarà oggetto di approfondimento futuro l'identificazione del tempo standard di ottimizzazione del sistema sulla base di diverse tipologie di monitoraggio e di diverse tipologie di edificio o distretto analizzato.
- L'analisi di qualità dei dati si basa su metodi diversi, come l'analisi dei dati mancanti e l'analisi dei profili, che devono essere svolti in parallelo al fine di poter fornire una corretta valutazione dell'attendibilità e completezza del database che si va a costruire.

Attività 3:

Le analisi condotte suggeriscono delle tendenze, sia relativamente alle forme dei profili di consumo che ai corrispondenti parametri statistici, che potranno essere ulteriormente approfondite estendendo i metodi esposti nel presente lavoro ad un periodo temporale più ampio con l'obiettivo di evidenziare differenze nei profili orari medi e negli andamenti dei consumi soprattutto in riferimento ai dati di natura termica parzialmente esaminati in questo studio. Per quanto attiene all'esame delle forme dei profili di carico, la valutazione del rapporto incrementale tra variazione di carico e corrispondente differenza calcolata sull'asse dei tempi consentirà la clusterizzazione dei profili con conseguente possibile ottimizzazione delle modalità di gestione della domanda attiva anche in relazione ai costi.

Attività 4:

La ricerca effettuata nell'ambito dell'Attività 4 ha condotto allo sviluppo di modelli robusti per la diagnosta avanzata degli edifici a supporto della complessa logica decisionale legata all'individuazione delle cause connesse alle anomalie di consumo o di qualità ambientale identificabili a partire dai dati di monitoraggio elaborati secondo grandezze e indicatori aggregati. L'attività di ricerca ha previsto l'applicazione di metodologie fondate sull'utilizzo sia di modelli inversi empirici sia di metodi statistici, seguendo un approccio del tipo top down per l'edificio F40 e per il distretto di edifici in ENEA Casaccia.

In particolare la ricerca è stata finalizzata alla verifica delle potenzialità di impiego di reti neurali (RNA) ed ensembling di reti neurali (BEM), così come di metodi di clustering e pattern recognition per la caratterizzazione e stima degli usi finali di energia negli edifici e per lo sviluppo di tecniche innovative di FDD (Fault Detection and Diagnosis).

I metodi proposti e implementati si sono rivelati adatti ai fini diagnostici anche se con differenti potenzialità e limiti. In particolare, sono state utilizzate due tecniche di ensembling di reti neurali artificiali (RNA) e nel caso di stima dei consumi sono stati applicati modelli statistici ai corrispondenti residui ai fini della fault detection dell'uso finale illuminazione dell'edificio F40 e del distretto di edifici. Gli ensembling neurali si sono sempre dimostrati più robusti rispetto ai singoli modelli neurali. L'analisi dei residui ha inoltre permesso di individuare outlier di consumo in relazione alle condizioni al contorno degli edifici (presenze, condizioni meteo, ora e tipo di giorno), e quindi outlier che molto più probabilmente rispecchiano reali valori anomali di consumo rispetto agli outlier individuati dai metodi statistici applicati ai soli dati di consumo. Il metodo CART accoppiato all'algoritmo GESD è risultato particolarmente robusto e accurato nella ricerca dei valori anomali di energia e potenza per illuminazione artificiale e di potenza elettrica totale. Alcuni dei cluster individuati con il metodo K-Means sono risultati impuri e i valori anomali spesso sono risultati disseminati e di difficile identificazione. Il metodo DBSACN si è rivelato particolarmente adatto a raggruppare dati in cluster caratterizzati dalla stessa densità e con valori simili; inoltre il metodo è risultato efficace nel raggruppare tutti i valori anomali in un singolo cluster. In generale si è riscontrato per i metodi di clustering analizzati che un opportuno e preventivo splitting dei dati permette di superare il limite intrinseco che caratterizza gli algoritmi alla loro base circa l'impossibilità di considerare variabili temporali per la formazione dei cluster. Dall'analisi diagnostica condotta sul distretto di edifici si è riscontrato come la suddivisione del data set in valori appartenenti al periodo giornaliero di lavoro, al periodo non lavorativo e ai fine settimana abbia consentito di individuare anche attraverso il metodo K-Means accoppiato all'algoritmo GESD e attraverso il metodo DBSCAN importanti anomalie che sono risultate comuni sia a quelle individuate attraverso il metodo CART che attraverso l'ensembling di reti neurali.

L'analisi di fault detection fin qui portata avanti prevede tra gli sviluppi futuri:

- l'individuazione di valori anomali di consumo relativamente agli usi finali termici di riscaldamento/climatizzazione degli edifici;
- lo sviluppo di un'analisi di fault detection a livello di intera rete per il distretto di edifici, valutando quindi le anomalie di consumo totale del distretto e come esse siano legate alle anomalie di consumo che si verificano nei singoli edifici.

L'attività condotta ha portato anche all'individuazione di tecniche diagnostiche per il rilevamento automatico di anomalie connesse all'uso della luce artificiale sulla base di modelli stimativi del livello di illuminamento naturale interno di ogni ufficio nell'edificio F40 basati su ensembling di reti neurali. I modelli sviluppati hanno posto le basi per l'implementazione di una tecnica diagnostica attraverso cui è possibile individuare anomalie rispetto all'uso delle luci artificiali in ogni stanza che risultano accese in presenza di occupanti e in corrispondenza di livelli medi di illuminamento naturale stimati sufficienti.

L'analisi di diagnostica condotta ha previsto, inoltre, l'individuazione di opportune funzioni obiettivo relative alla qualità dell'ambiente interno e alla richiesta energetica per condurre analisi numeriche di ottimizzazione energetica nell'edificio F40 attraverso l'applicazione di algoritmi genetici (in collaborazione con l'Università Politecnica della Marche e l'Università Roma 3) relativamente alle variabili temperatura interna di set-point e temperatura di mandata dei fluidi termovettori caldo/freddo dell'impianto di climatizzazione ambientale.

Riferimenti bibliografici

1. J.Haberl and M.Abbas. "Development of graphical indices for viewing building energy data: Part I-II". Journal of Solar Energy Engineering, vol120, pp.156-167,1998.
2. P.Price, "Methods for Analyzing Electric Load Shape and its Variability". Ernest Orlando Lawrence Berkeley National Laboratory-Environmental Energy Technologies Division. May 2012.
3. W.Chung, "Review of building energy-use performance benchmarking methodologies". Applied Energy, volume 88, Issue 5, May 2011, Pages 1470-1479.
4. G. K. Palshikar, "Simple algorithms for peak detection in time-series", Tata Research Development and Design Centre (TRDDC), India.

Curriculum vitae degli Autori del Rapporto Tecnico

Stefano Paolo Corgnati

Vercellese di Livorno Ferraris, laureato con lode in ingegneria meccanica e dottore di ricerca in energetica, è professore associato di “Fisica tecnica ambientale” presso il Dipartimento Energia del Politecnico di Torino. Svolge la sua attività di ricerca nel gruppo TEBE (www.polito.it/tebe) nel settore dell’uso razionale dell’energia e del risparmio energetico negli edifici. E’ coordinatore operativo del Centro di Ricerca “Indoor Environment & Energy Management Competence Centre” di Telecom presso il Politecnico di Torino. E’ autore di oltre duecento pubblicazioni a carattere scientifico, didattico e divulgativo. Per le sue ricerche, nel 2009 è vincitore del premio REHVA “Young Scientist Award”. Nel 2011 è nominato “Rehva Fellow” e nello stesso anno è eletto Vicepresidente di Rehva (Federazione Europea delle Associazioni del Condizionamento dell’Aria) e co-chair del Educational Committee. Nei gruppi di ricerca della Agenzia Internazionale per l’Energia (IEA), è stato leader del Subtask-C nel progetto ECBCS-Annex 53 “Total Energy Use in Buildings”. Attualmente è Co-Leader del SubTask_A nel progetto ECBCS - Annex 59 “High Temperature Cooling & Low Temperature Heating in Buildings”. E’ inoltre Chair delle Rehva-TaskForce “Indoor Climatic Control Assessment” e “Reference buildings for energy performance and cost-optimal analyses”. Nel maggio 2013 è stato eletto Sindaco di Livorno Ferraris, nel cui Comune dal 2004 al 2010 ha ricoperto la carica di Assessore all’Ambiente e all’Energia.

Alfonso Capozzoli

(1979), ingegnere meccanico e dottore di ricerca in ingegneria dei sistemi meccanici, è attualmente ricercatore a tempo determinato presso il Dipartimento Energia del Politecnico di Torino. Dal 2007 al 2012 è stato professore a contratto in fisica tecnica, fisica tecnica ambientale, impianti termotecnici, presso l’Università degli Studi di Napoli Federico II ed il Politecnico di Torino. Ha svolto attività di ricerca in qualità di ricercatore a contratto post doc presso il Politecnico di Torino dal 2008 al 2012. E’ autore di circa 60 pubblicazioni scientifiche, tecniche e didattiche sugli impianti di climatizzazione, sulla simulazione energetica dinamica della prestazione degli edifici, sulla fisica dell’edificio, sul controllo dell’umidità e sulla modellazione inversa. E’ membro del consiglio direttivo della sezione italiana di IBPSA (International Building Performance Simulation Association) e partecipa in qualità di esperto a gruppi di lavoro in ambito CTI (Comitato Termotecnico Italiano) sul tema delle prestazioni energetiche degli edifici. E’ membro del comitato scientifico del master di II livello della I Facoltà di Architettura del Politecnico di Torino in edilizia sostenibile ed efficienza energetica. E’ coinvolto in diversi progetti nazionali e internazionali sulla prestazione energetica degli edifici.

Imran Khan

(1980), ha ottenuto il suo master in “Building Services Engineering” presso la University of Leicestershire, UK e il suo BSc in ingegneria meccanica presso la N.W.F.P University of Engineering and Technology in Pakistan dove ha svolto anche il ruolo di lecturer. E’ attualmente dottorando in Energetica presso il Dipartimento di Energia DENERG del Politecnico di Torino.

Fiorella Lauro

Ha conseguito la Laurea Specialistica in Ingegneria dell’Automazione presso l’Università della Calabria. Nel lavoro di tesi, svolto presso l’unità UTTEI del Centro Ricerche ENEA di Roma, si è occupata dello sviluppo di un nuovo approccio per la modellazione dei consumi energetici degli edifici al fine di realizzare un innovativo sistema di gestione remota per la diagnostica delle anomalie e l’ottimizzazione del comportamento energetico delle reti di edifici. E’ autore di tre pubblicazioni scientifiche sull’ensembling di

reti neurali e su modelli ibridi applicati ai consumi energetici degli edifici. Attualmente è ricercatore a contratto presso il Politecnico di Torino e la sua attività di ricerca riguarda la modellazione dei consumi energetici degli edifici attraverso approcci metodologici inversi e l'individuazione di efficaci metodologie diagnostiche.

Daniela Raimondo

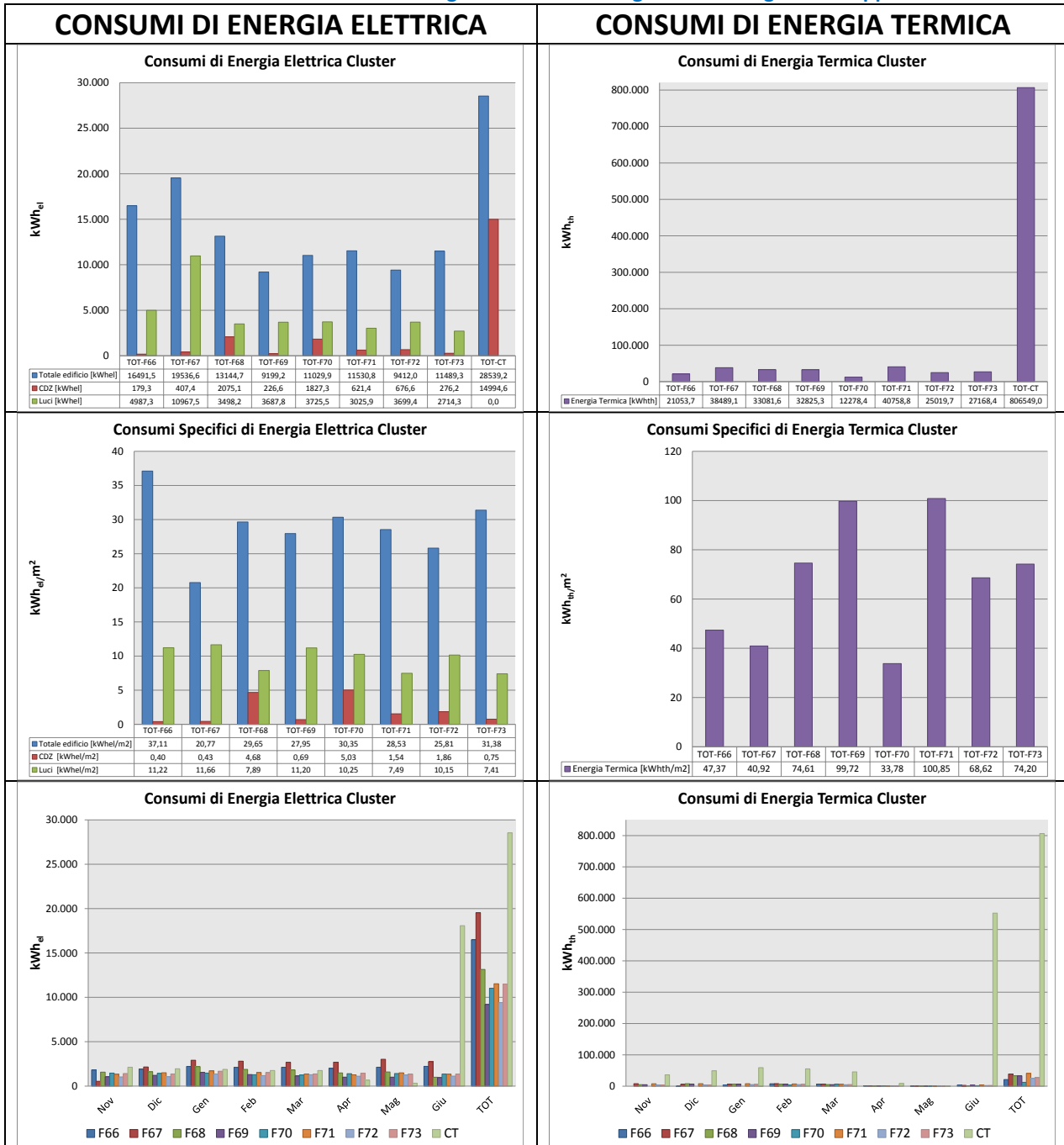
Daniela Raimondo, laureata in Scienze dell'Architettura e dottore di ricerca in Innovazione per l'ambiente costruito (Fisica tecnica ambientale), svolge attualmente il ruolo di ricercatore a contratto all'interno del gruppo Tebe, presso il Dipartimento Energia del Politecnico di Torino. Durante il percorso di formazione ha approfondito i temi del comfort ambientale legato ai consumi energetici degli edifici, focalizzando parte della ricerca su sistemi radianti innovativi a bassa domanda energetica. Ha svolto parte del dottorato di ricerca presso l'International Center of Indoor Environment and Energy del DTU di Copenhagen, dove ha svolto attività di monitoraggio energetico ed ambientale che sono ora, insieme al tema dell'audit energetico ed ambientale, alla base della sua ricerca presso l'IEEM (Indoor Environment and Energy Management Competence Center) di Telecom. Da 5 anni copre il ruolo di assistenza al corso di progettazione sostenibile presso il corso di laurea magistrale di Architettura al Politecnico di Torino. È attualmente coinvolta in svariati progetti di ricerca ed è autrice di oltre 15 pubblicazioni scientifiche nazionali e internazionali.

Novella Talà

Ha conseguito la Laurea in Ingegneria Civile presso il Politecnico di Torino nel 2003. Nel periodo ottobre 2003-marzo 2008 ha lavorato presso imprese di costruzioni come assistente direttore tecnico. Attualmente è ricercatore a contratto presso il Dipartimento Energia del Politecnico di Torino. La sua attività di ricerca è focalizzata principalmente nell'area dei metodi di analisi statistica finalizzati allo studio dell'utilizzo dell'energia negli edifici e alla previsione dei consumi energetici degli edifici. Ing. Talà è coautore di oltre 10 pubblicazioni in atti di convegni internazionali e rapporti interni per industrie. Ing. Talà collabora con ENEA e TELECOM in progetti nazionali e con diverse istituzioni accademiche e industriali nell'ambito di progetti europei ANNEX e BECA.

APPENDICE

Tabella 44: Confronto tra i consumi di energia elettrica ed energia termica degli edifici appartenenti al cluster



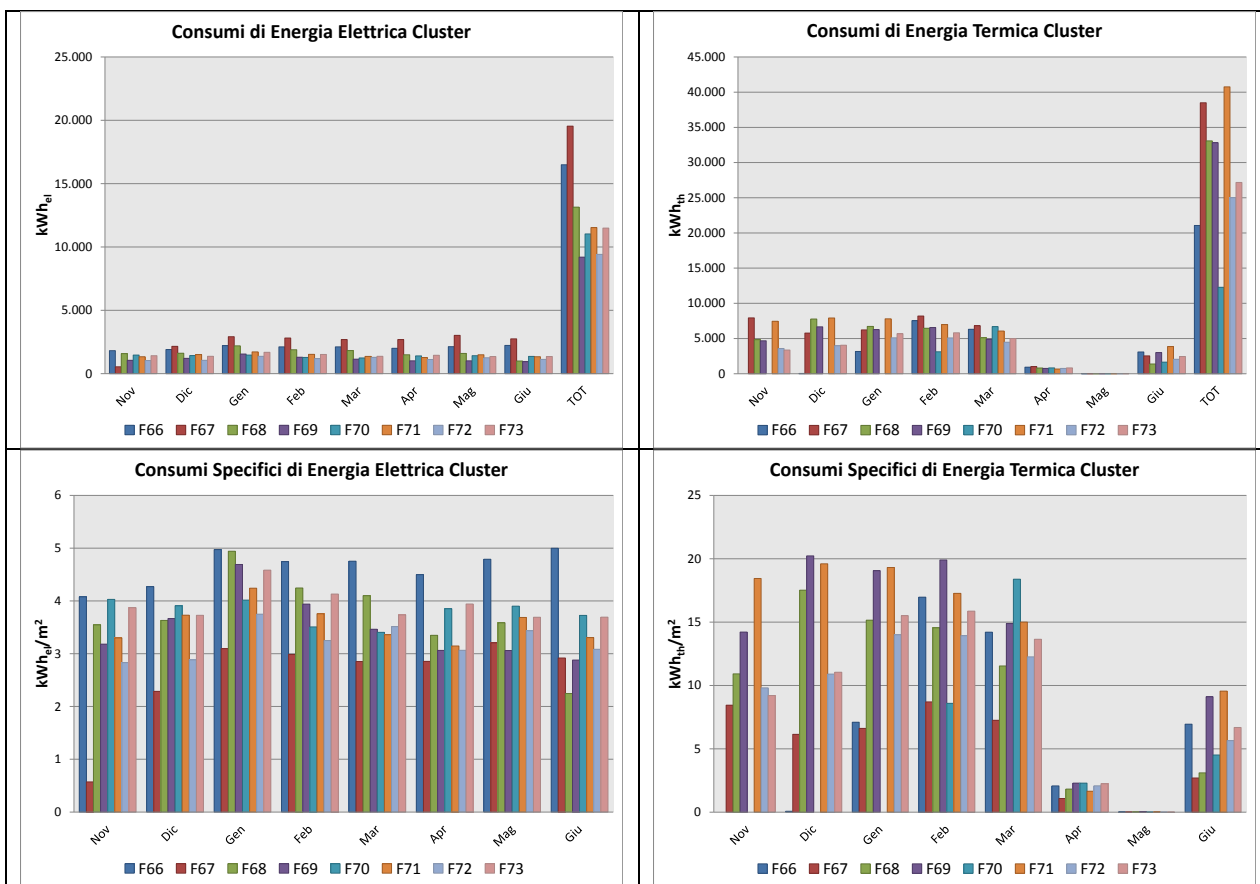
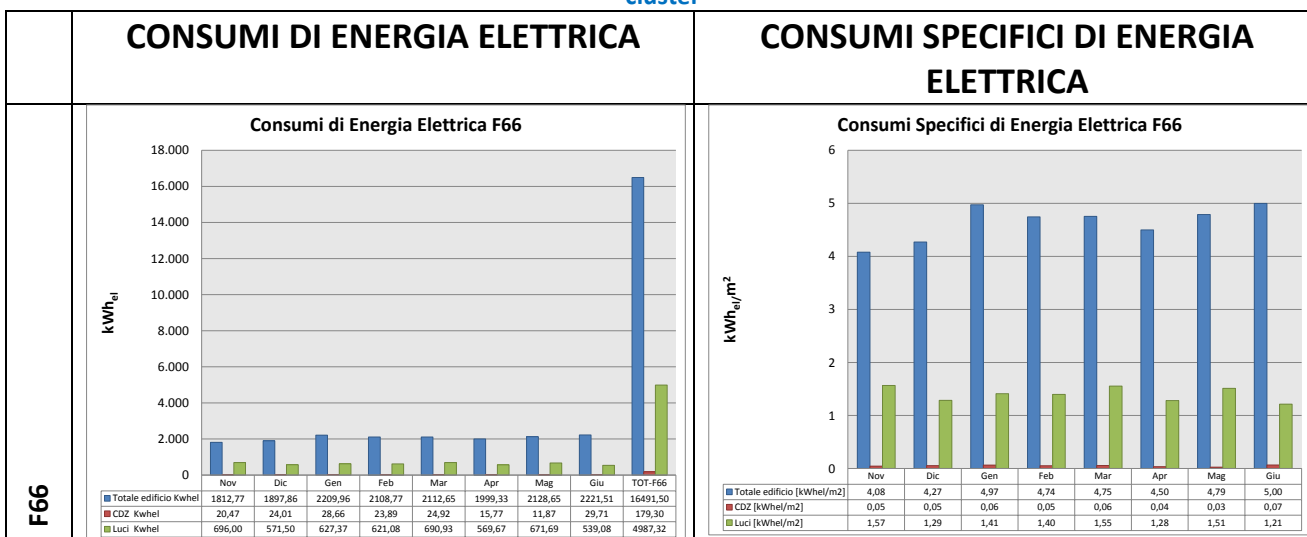
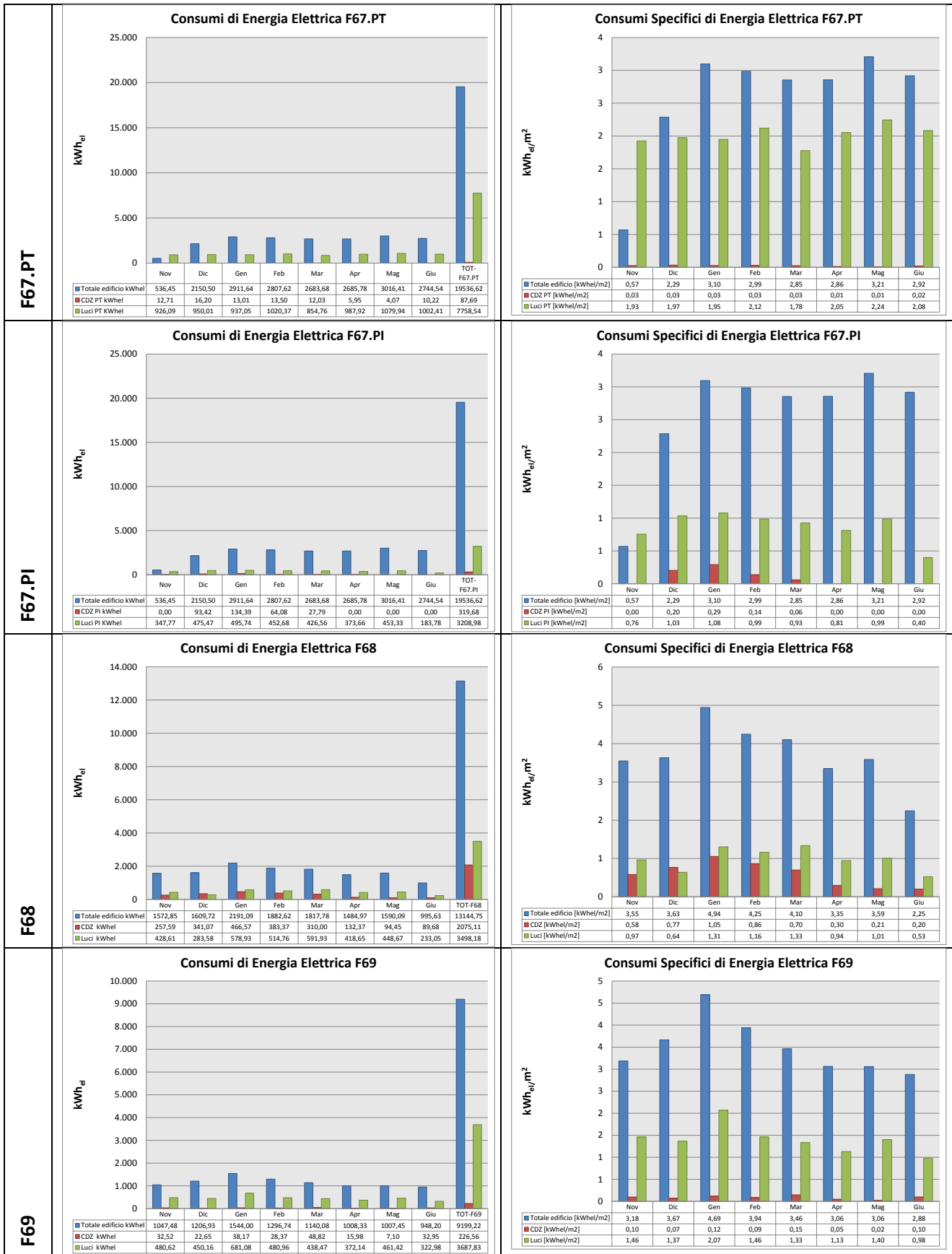
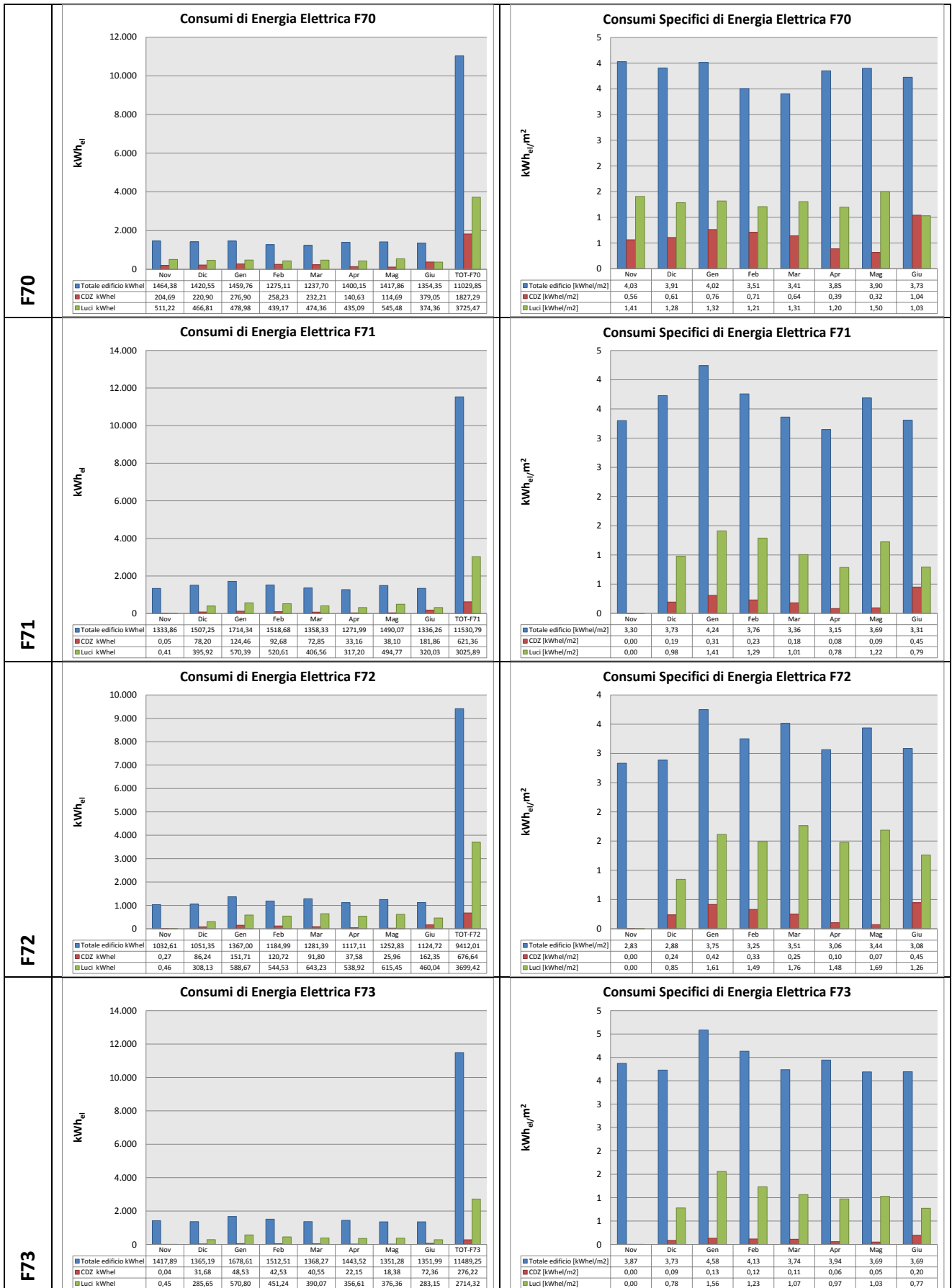


Tabella 45: Confronto tra consumi assoluti e consumi specifici di energia elettrica degli edifici appartenenti al cluster







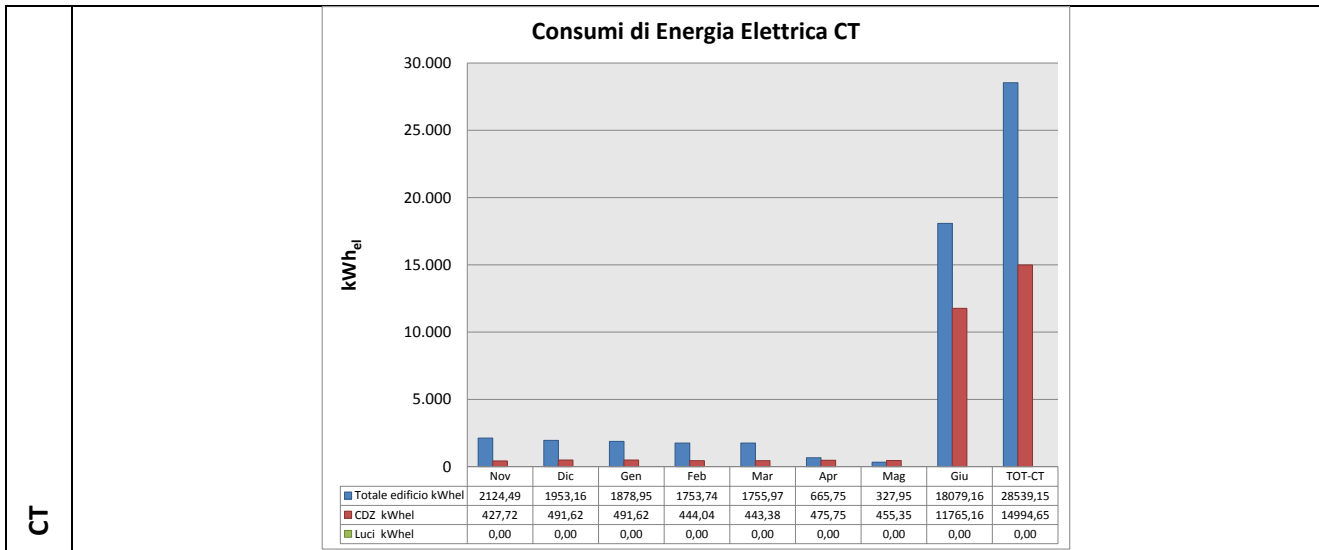
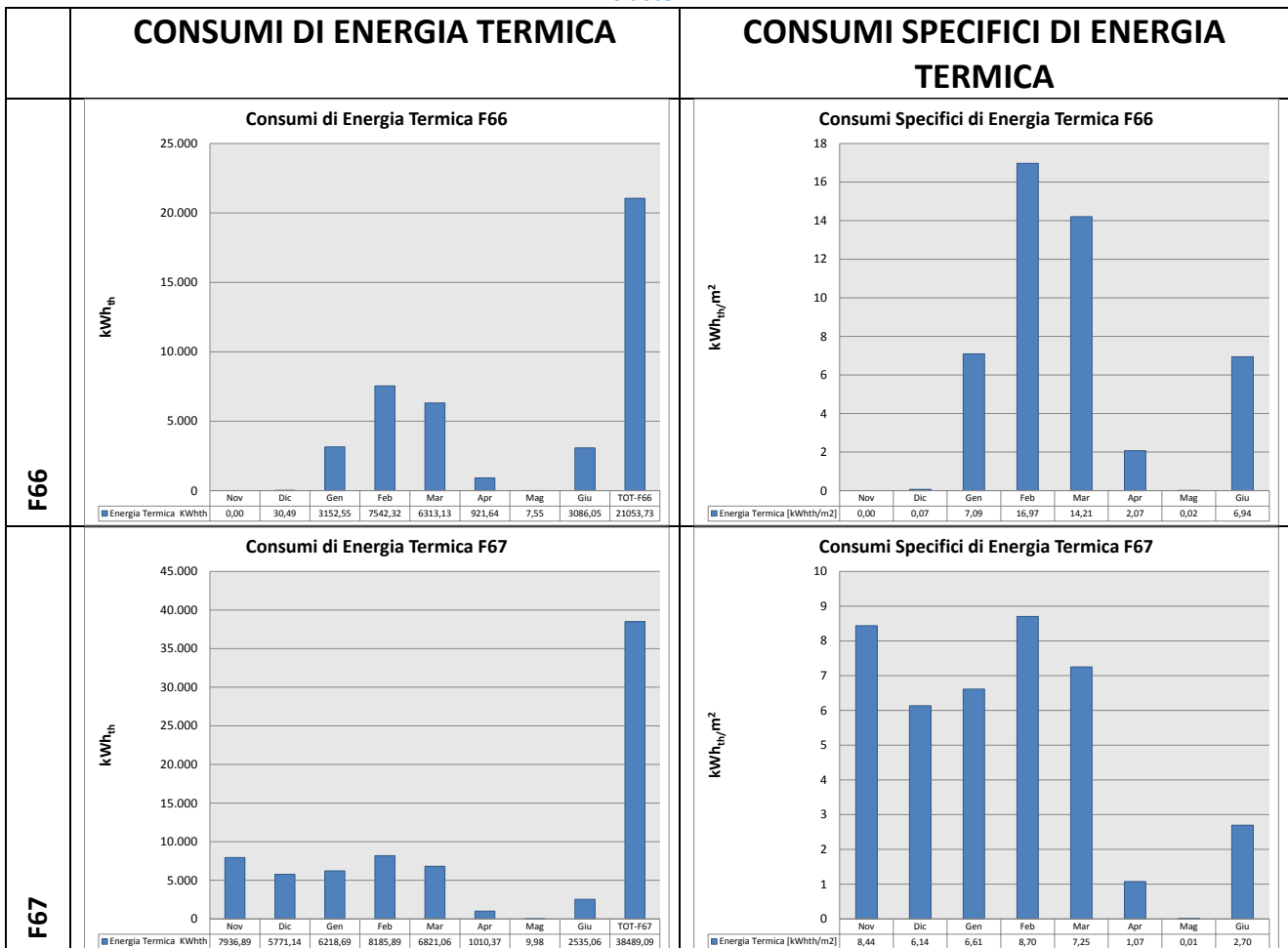
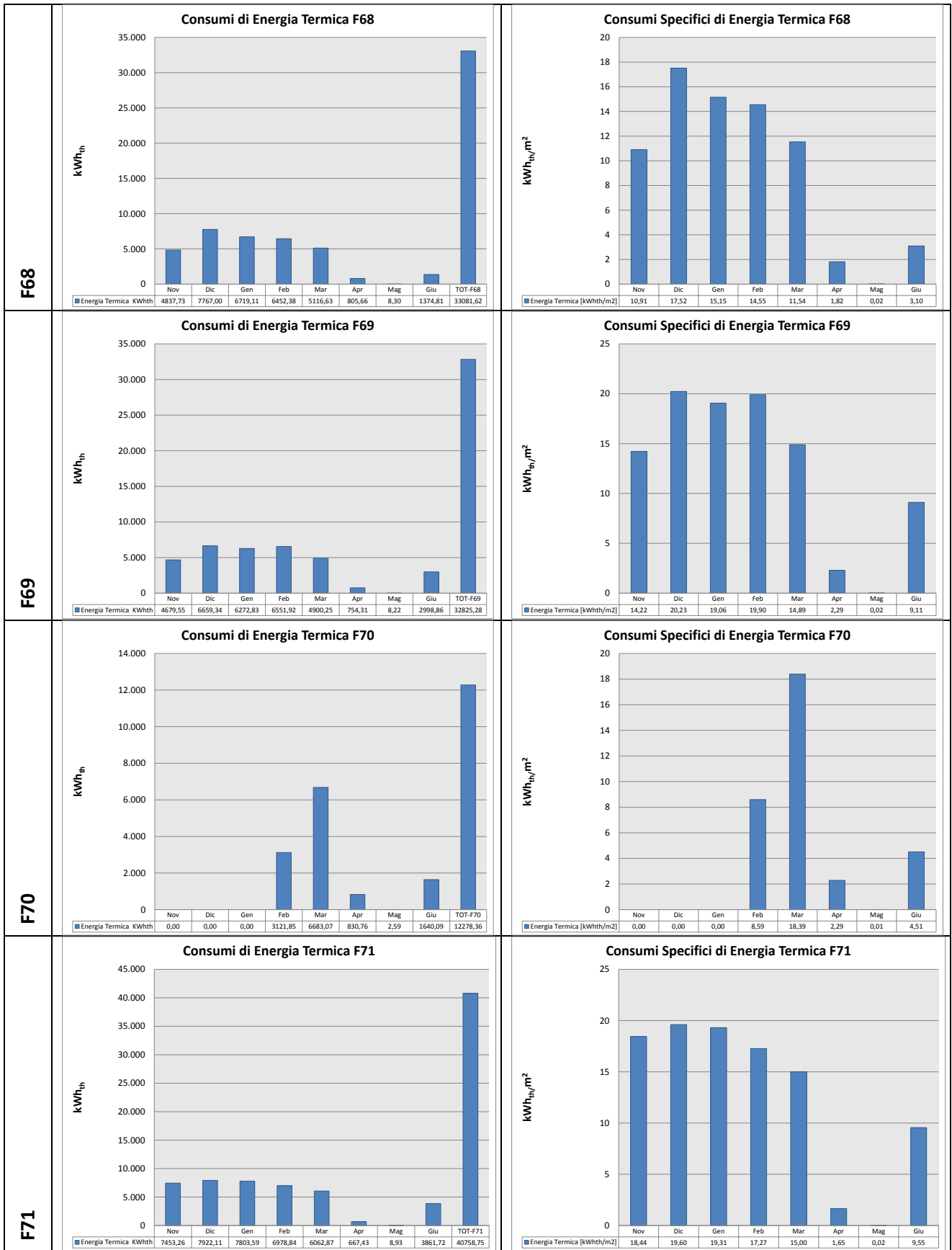


Tabella 46: Confronto tra consumi assoluti e consumi specifici di energia termica degli edifici appartenenti al cluster





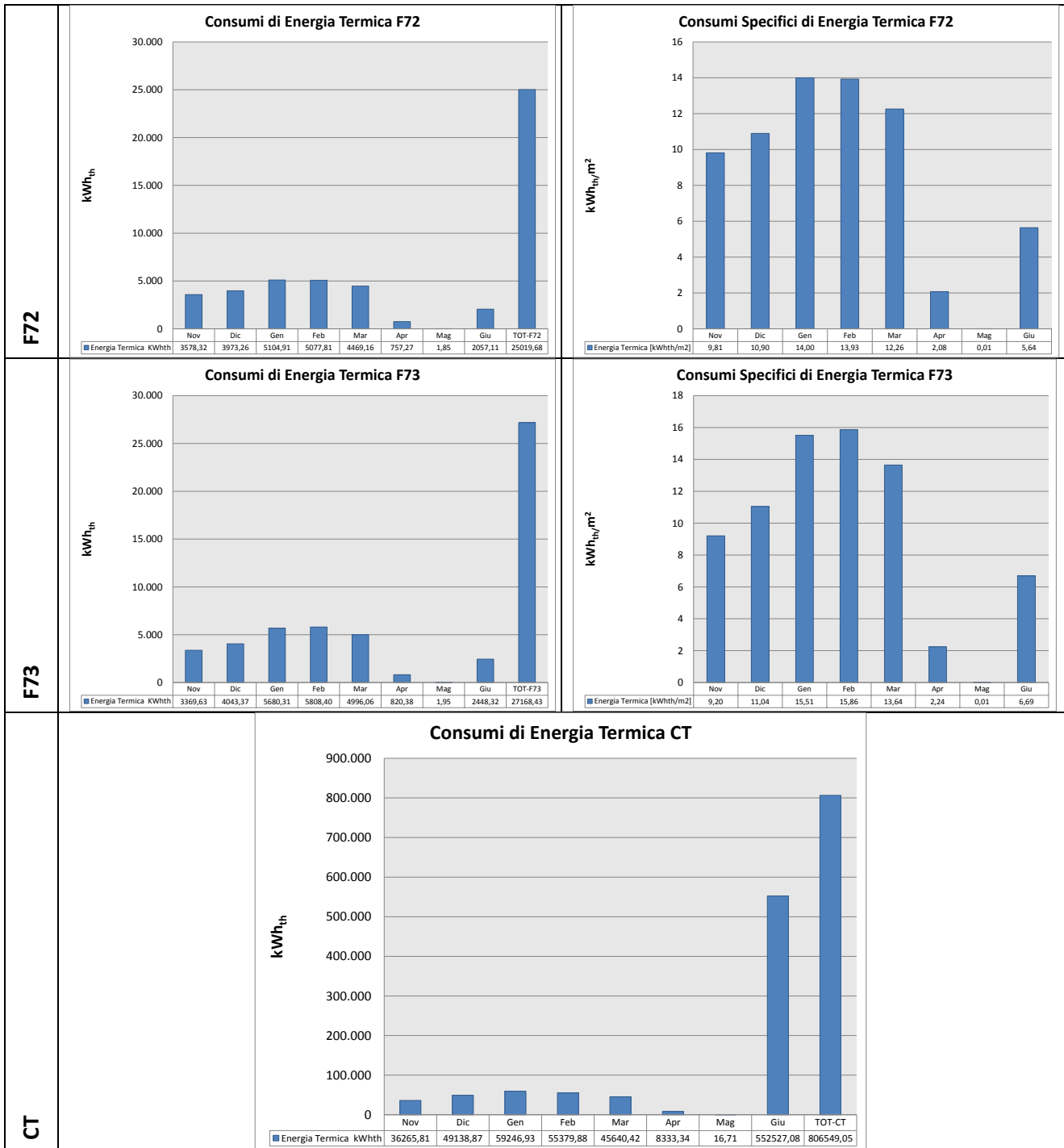
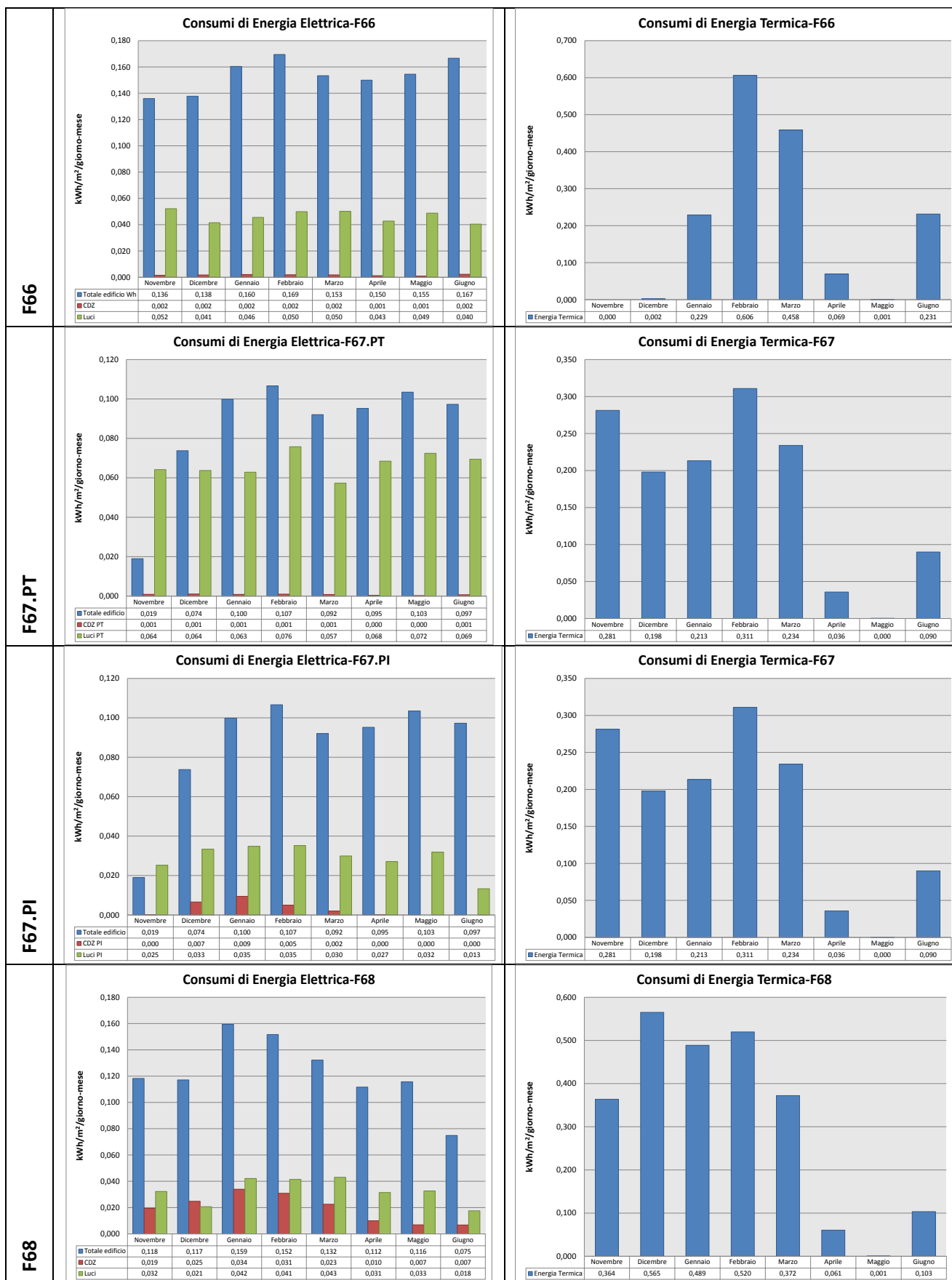
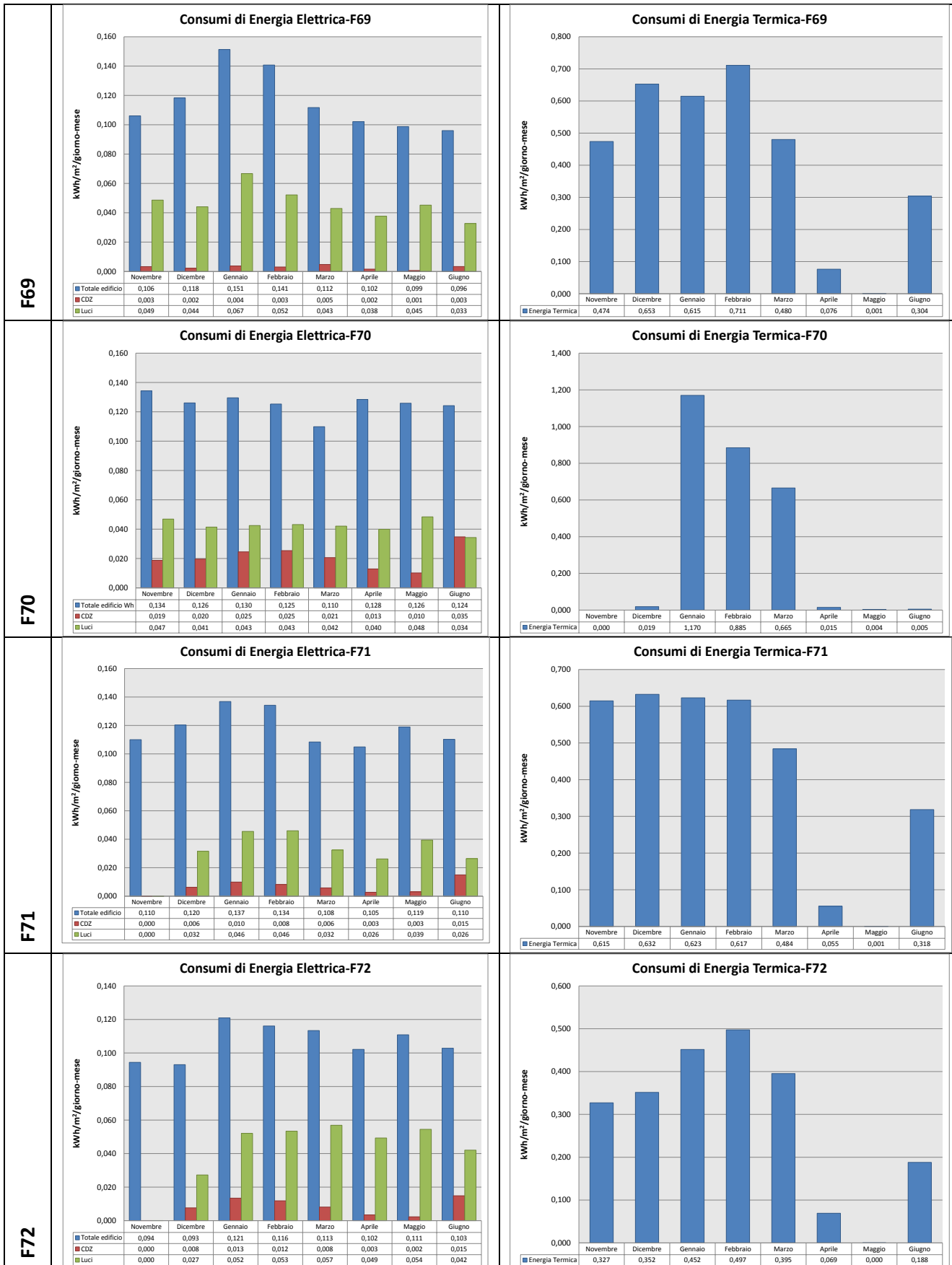


Tabella 47: Andamento mensile dei consumi specifici di energia elettrica e di energia termica per ciascuno degli edifici appartenenti al cluster

CONSUMI DI ENERGIA ELETTRICA	CONSUMI DI ENERGIA TERMICA
-------------------------------------	-----------------------------------





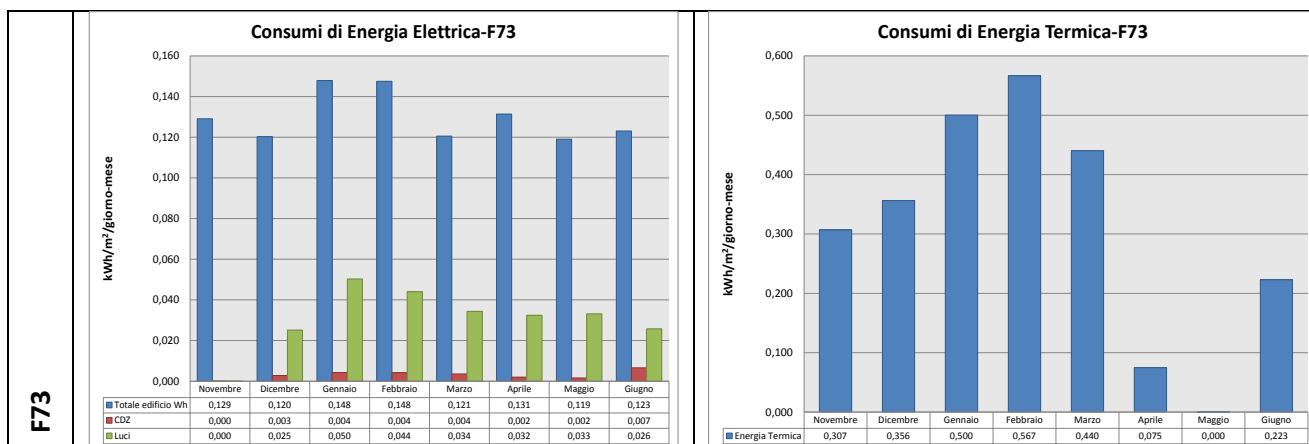
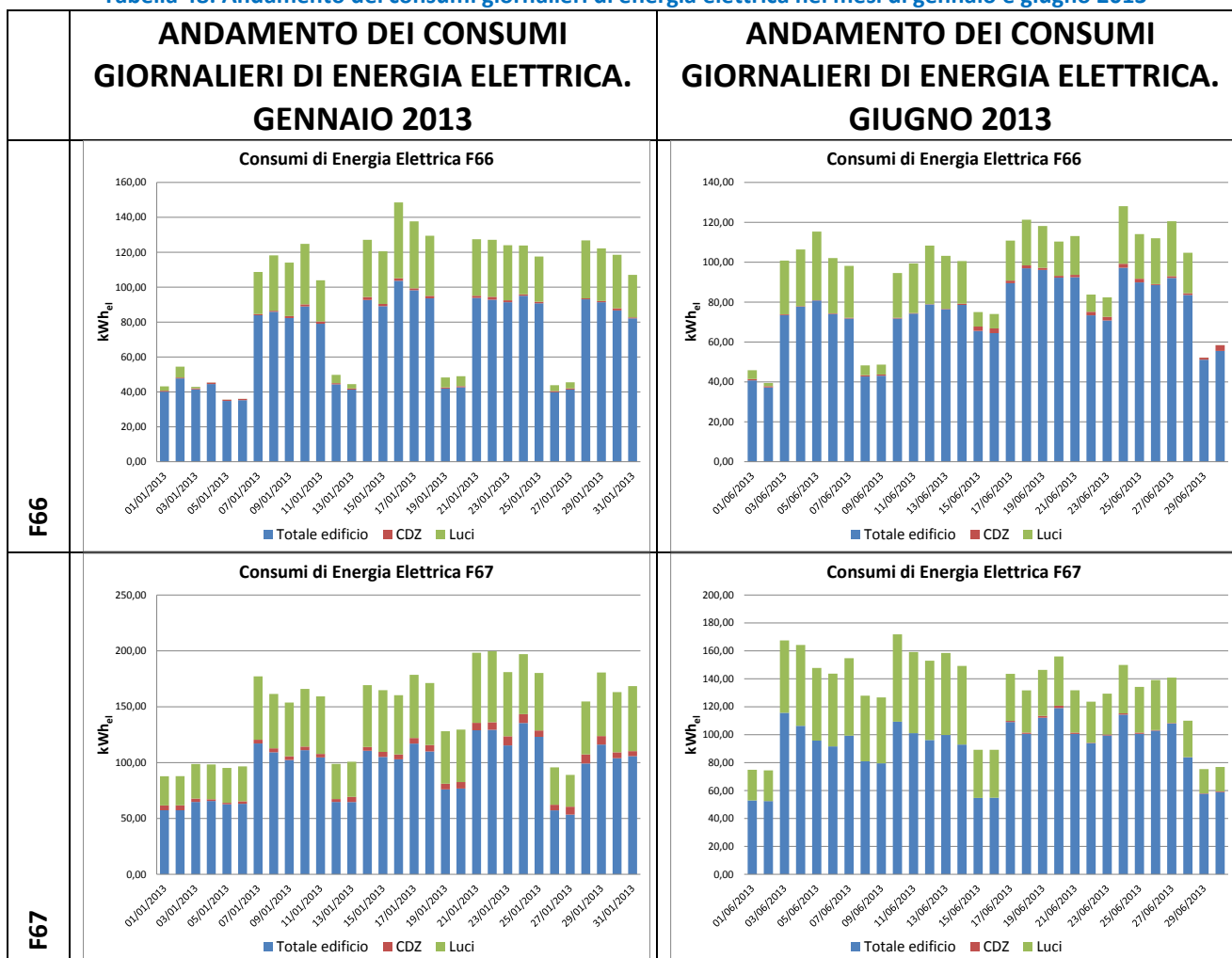
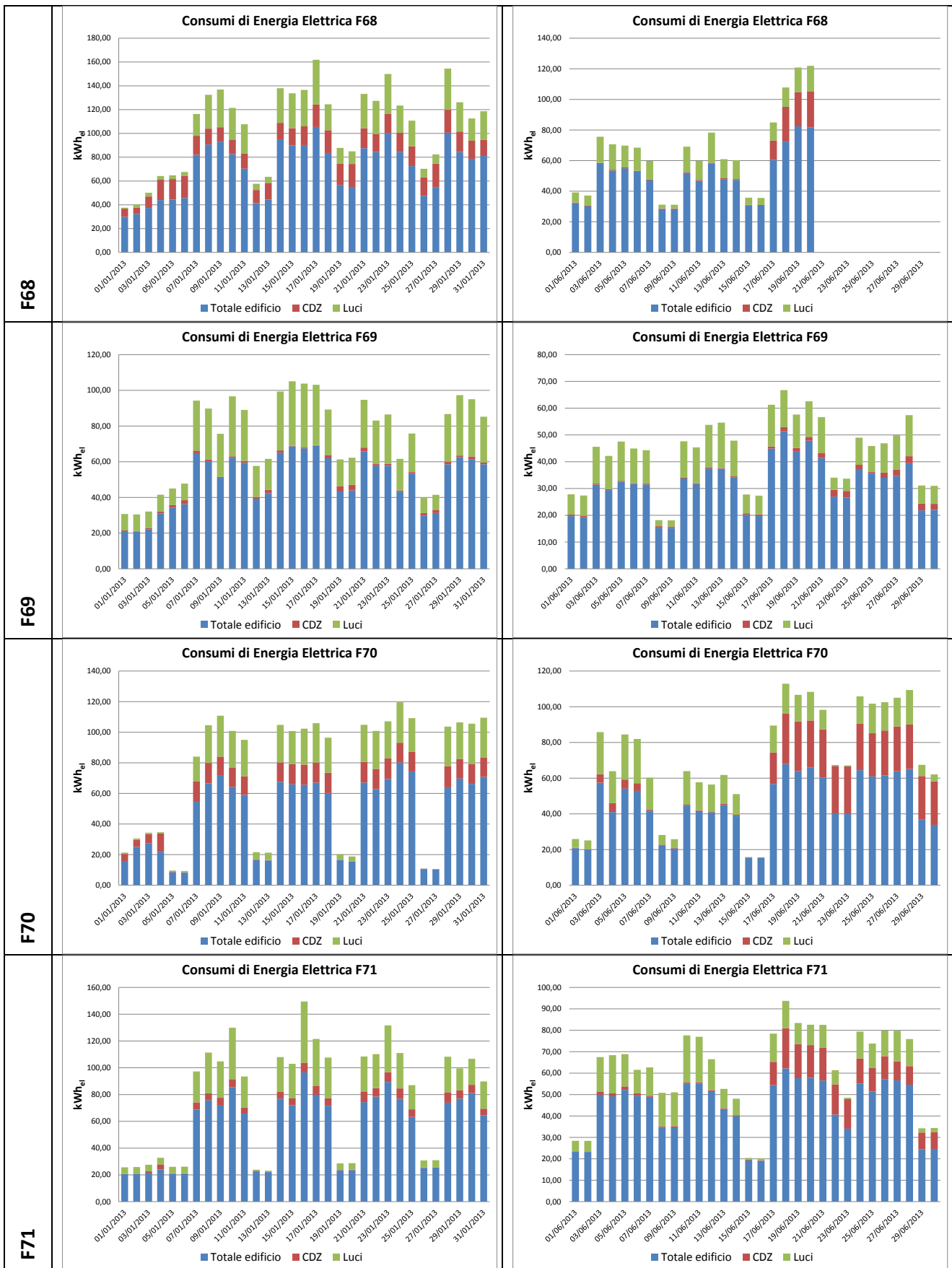


Tabella 48: Andamento dei consumi giornalieri di energia elettrica nei mesi di gennaio e giugno 2013





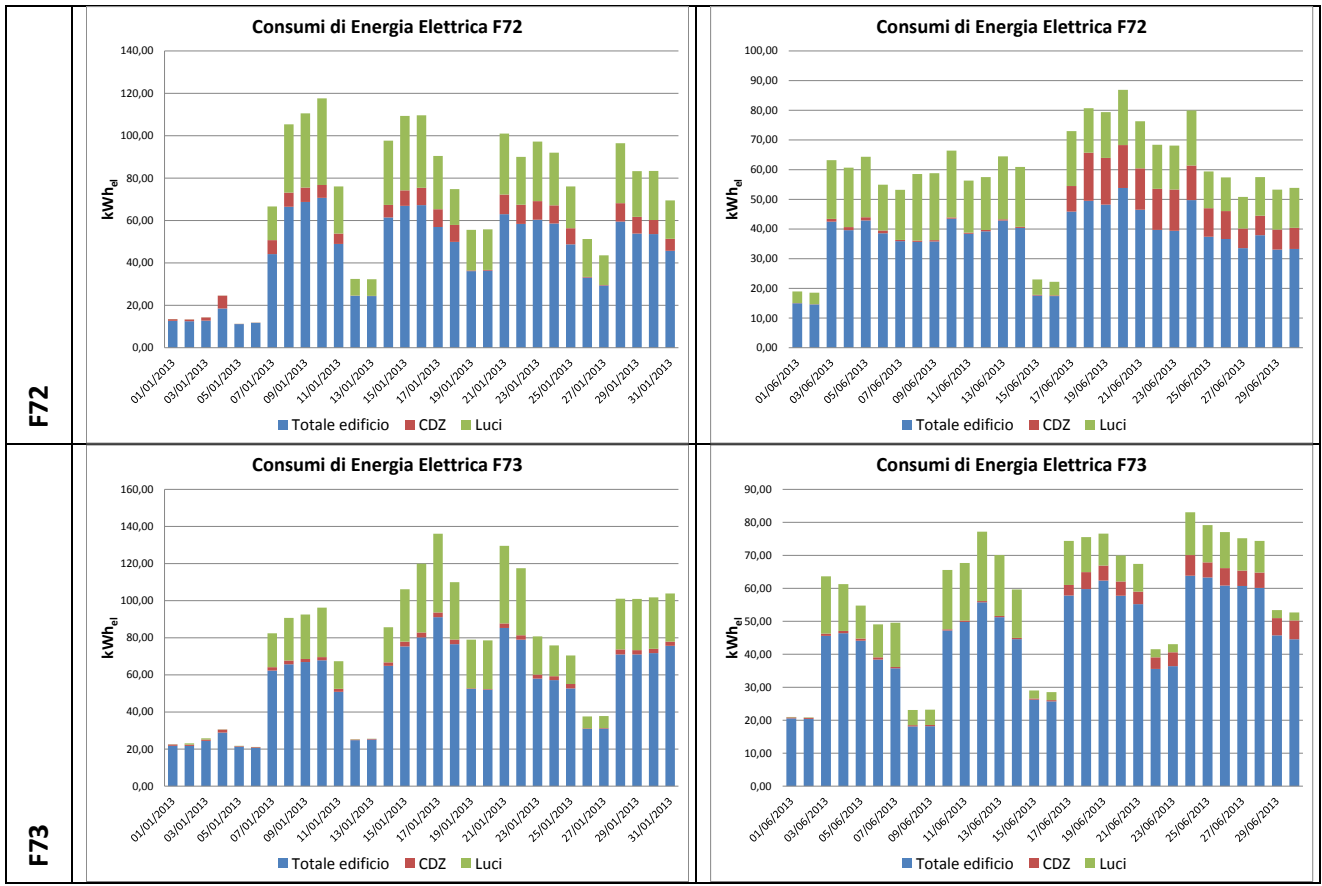
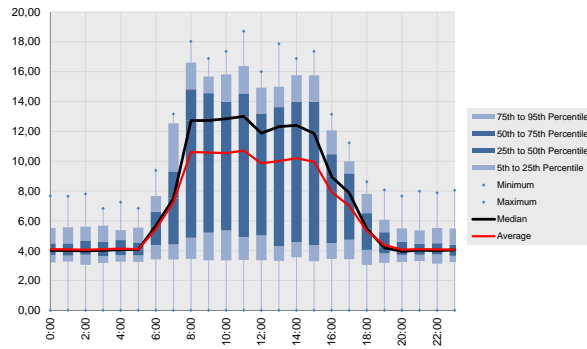


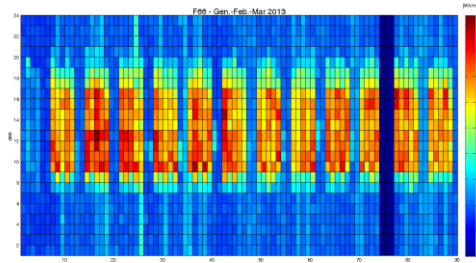
Tabella 49: Box-Plots e Carpet Plots dei consumi di energia elettrica per ciascuno degli edifici appartenenti al cluster nella stagione invernale

	F66
	F67

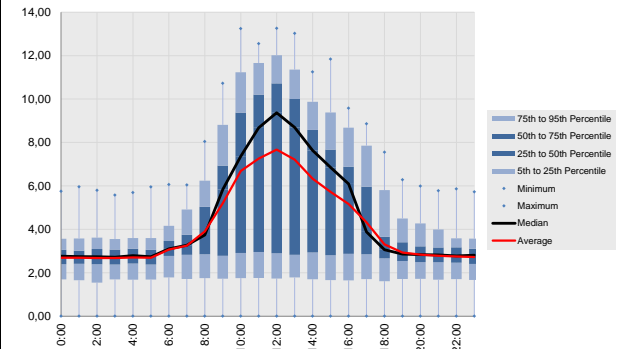
Gennaio-Febbraio-Marzo 2013



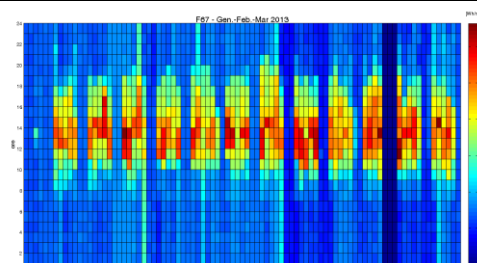
gennaio-febbraio-marzo 2013	Min	Max	Deviazione standard	Media	Coeff. di Var. COV	Mediana
Near Peak Load	3,65	17,93	4,71	11,86	0,40	14,32
Near Base Load	2,05	5,49	0,46	3,73	0,12	3,65
NBL+0,1*(NPL-NBL)	2,30	5,96	0,67	4,54	0,15	4,74
NPL-0,1*(NPL-NBL)	3,59	16,60	4,24	11,04	0,38	13,26
(NPL+NBL)/2	3,29	11,28	2,40	7,79	0,21	8,97



F68

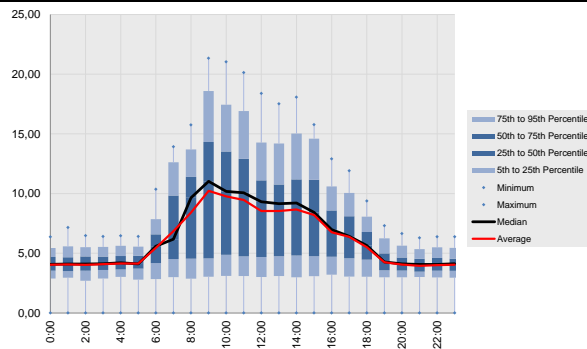


gennaio-febbraio-marzo 2013	Min	Max	Deviazione standard	Media	Coeff. di Var. COV	Mediana
Near Peak Load	1,93	12,84	3,61	8,07	0,45	9,98
Near Base Load	1,36	3,55	0,48	2,46	0,20	2,52
NBL+0,1*(NPL-NBL)	1,49	4,21	0,62	3,02	0,20	3,08
NPL-0,1*(NPL-NBL)	1,89	11,73	3,26	7,50	0,43	9,16
(NPL+NBL)/2	1,75	7,89	1,87	5,26	0,35	6,12

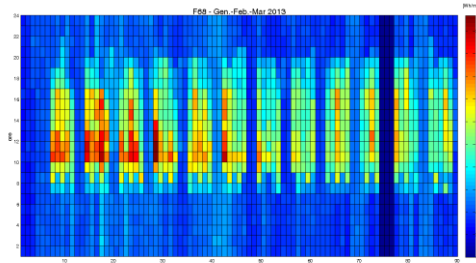


F69

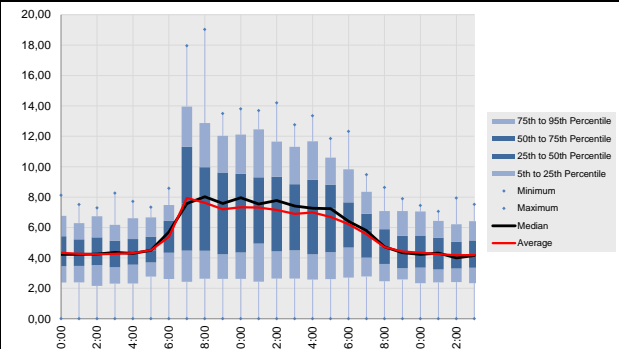
Gennaio-Febbraio-Marzo 2013



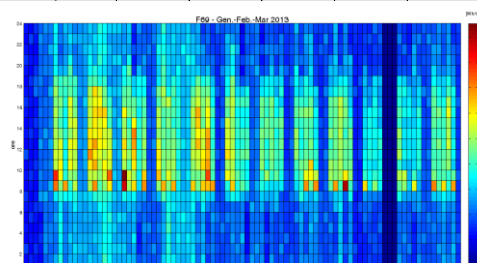
gennaio-febbraio-marzo 2013	Min	Max	Deviazione standard	Media	Coeff. di Var. COV	Mediana
Near Peak Load	3,19	21,16	4,80	10,80	0,44	11,88
Near Base Load	1,65	5,49	0,65	3,74	0,17	3,70
NBL+0,1*(NPL-NBL)	1,82	6,61	0,83	4,45	0,19	4,40
NPL-0,1*(NPL-NBL)	3,13	19,47	4,34	10,10	0,43	11,02
(NPL+NBL)/2	2,50	12,68	2,49	7,27	0,34	7,53



F70



gennaio-febbraio-marzo 2013	Min	Max	Deviazione standard	Media	Coeff. di Var. COV	Mediana
Near Peak Load	3,30	18,14	3,36	8,95	0,38	9,45
Near Base Load	1,71	6,47	1,02	3,67	0,28	3,42
NBL+0,1*(NPL-NBL)	1,94	7,04	1,08	4,19	0,26	3,99
NPL-0,1*(NPL-NBL)	3,20	16,79	3,06	8,42	0,36	8,84
(NPL+NBL)/2	2,80	11,42	1,92	6,31	0,30	6,27



F71

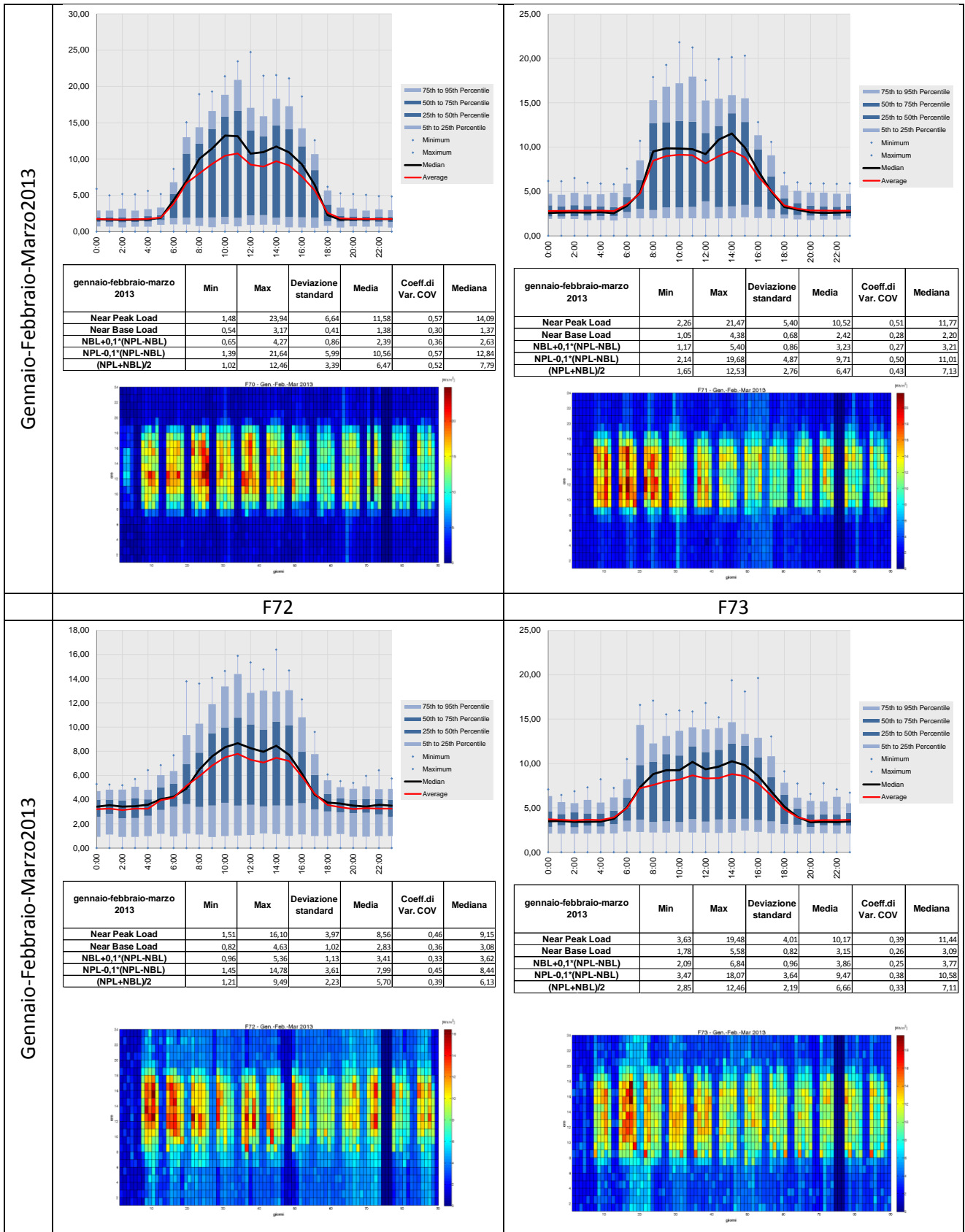
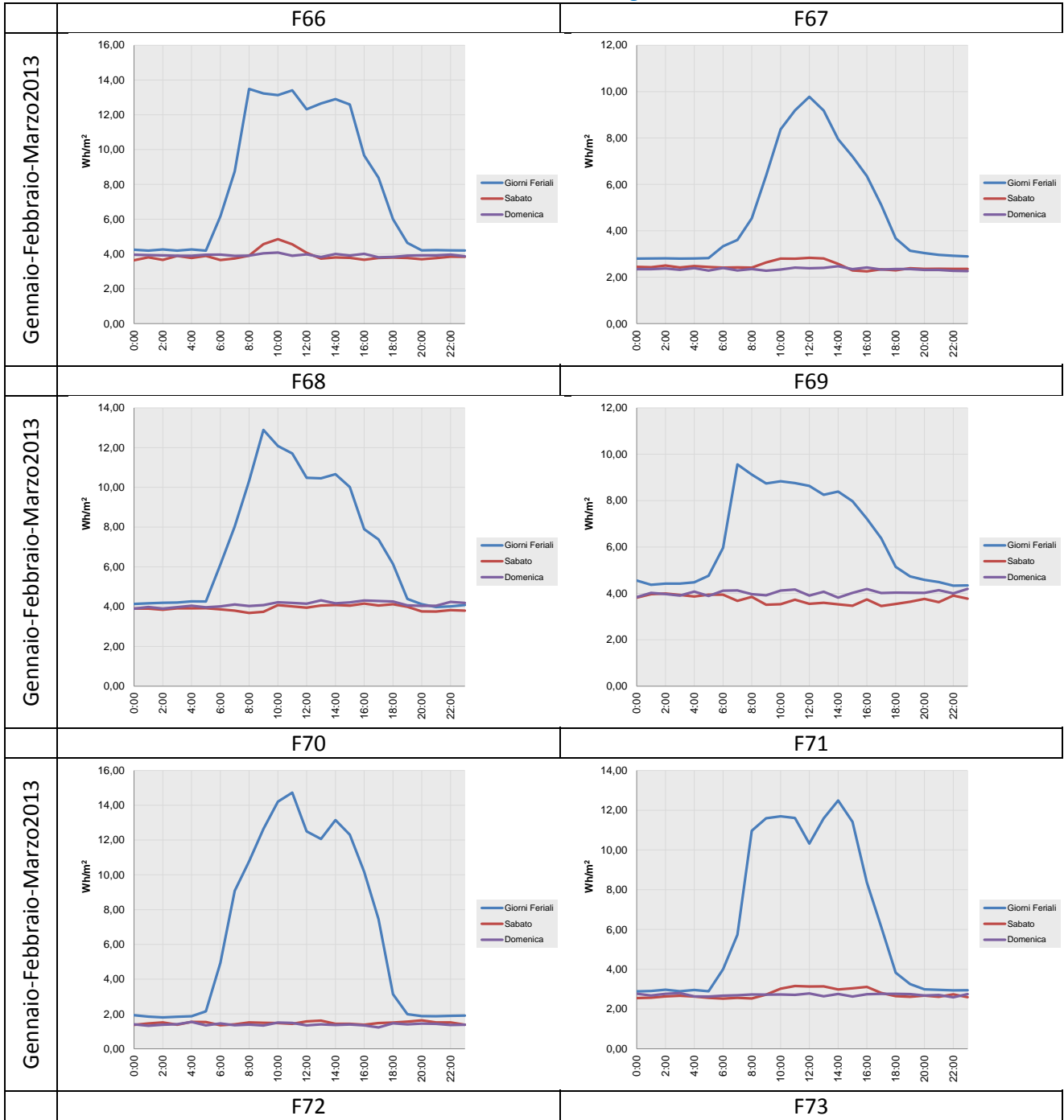


Tabella 50: Profili orari di carico medio stagionale di energia elettrica corrispondenti ai giorni feriali, ai giorni di sabato ed alle domeniche nella stagione invernale



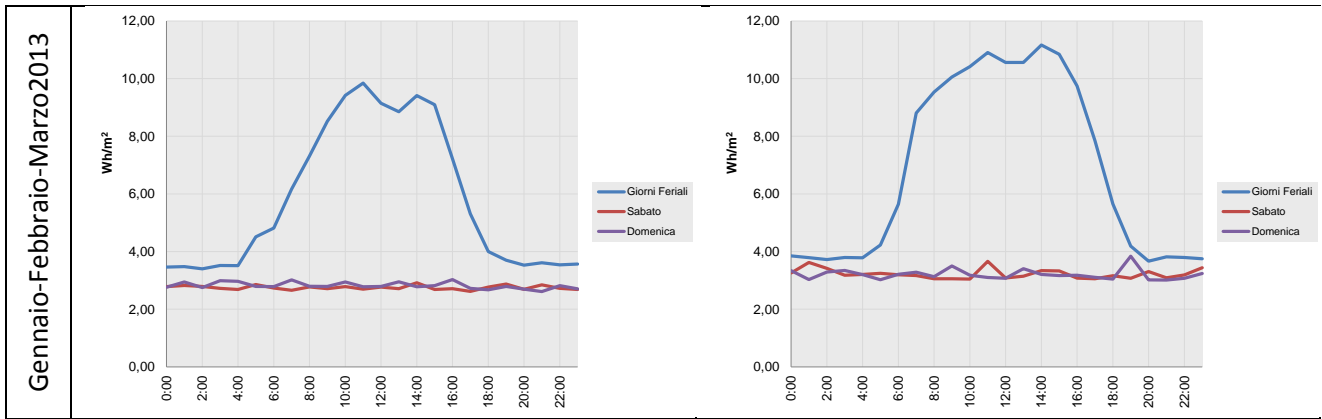
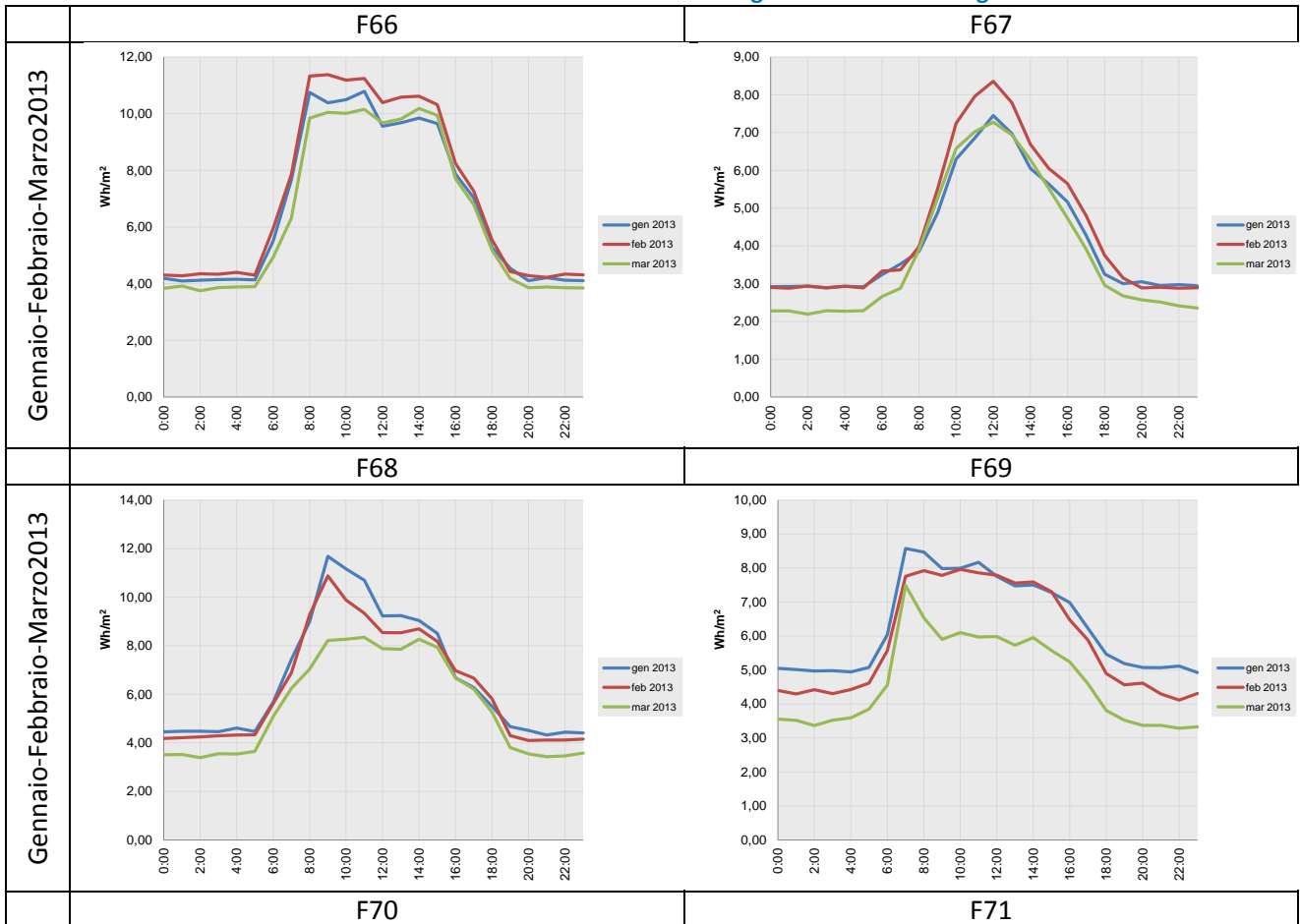


Tabella 51: Profili orari di carico medio mensile di energia elettrica nella stagione invernale



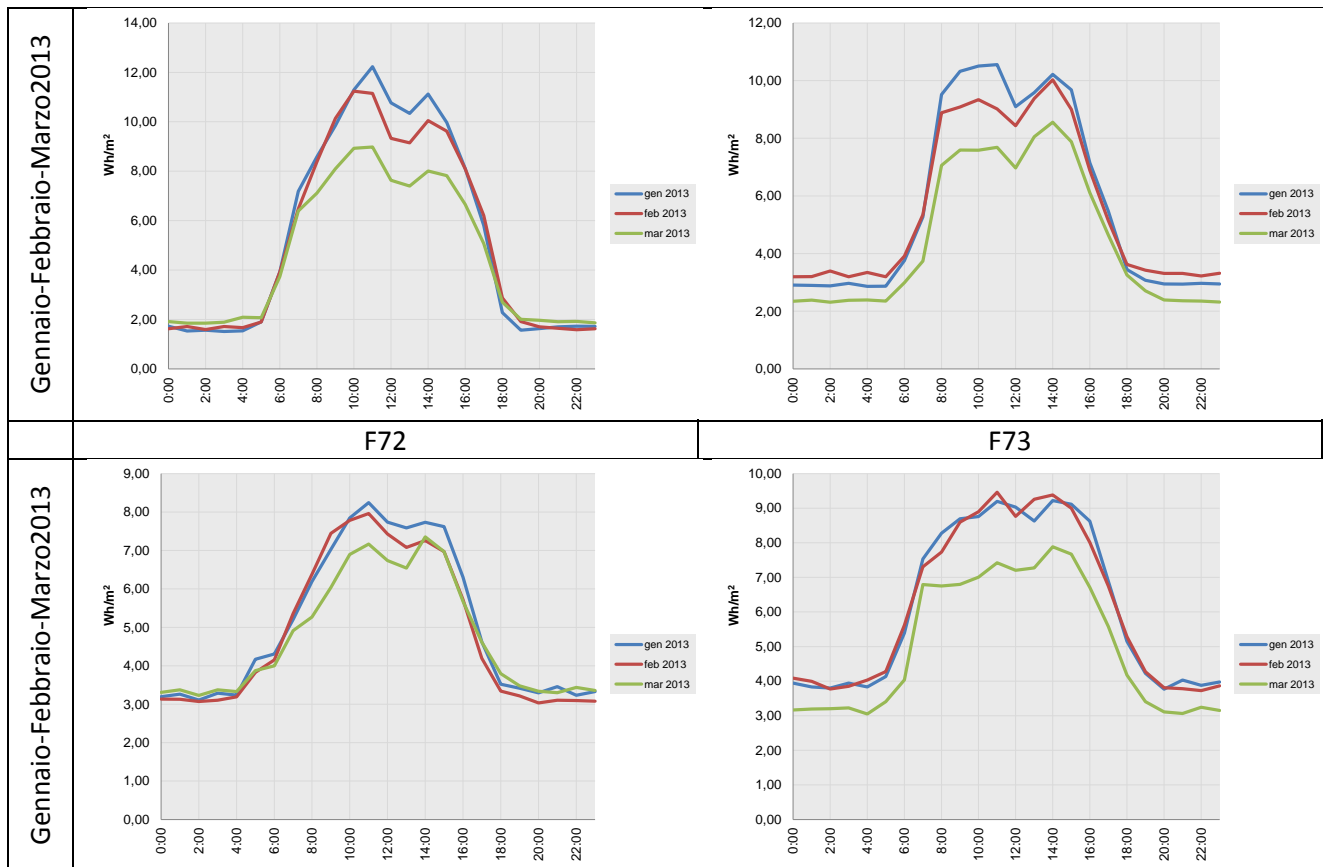
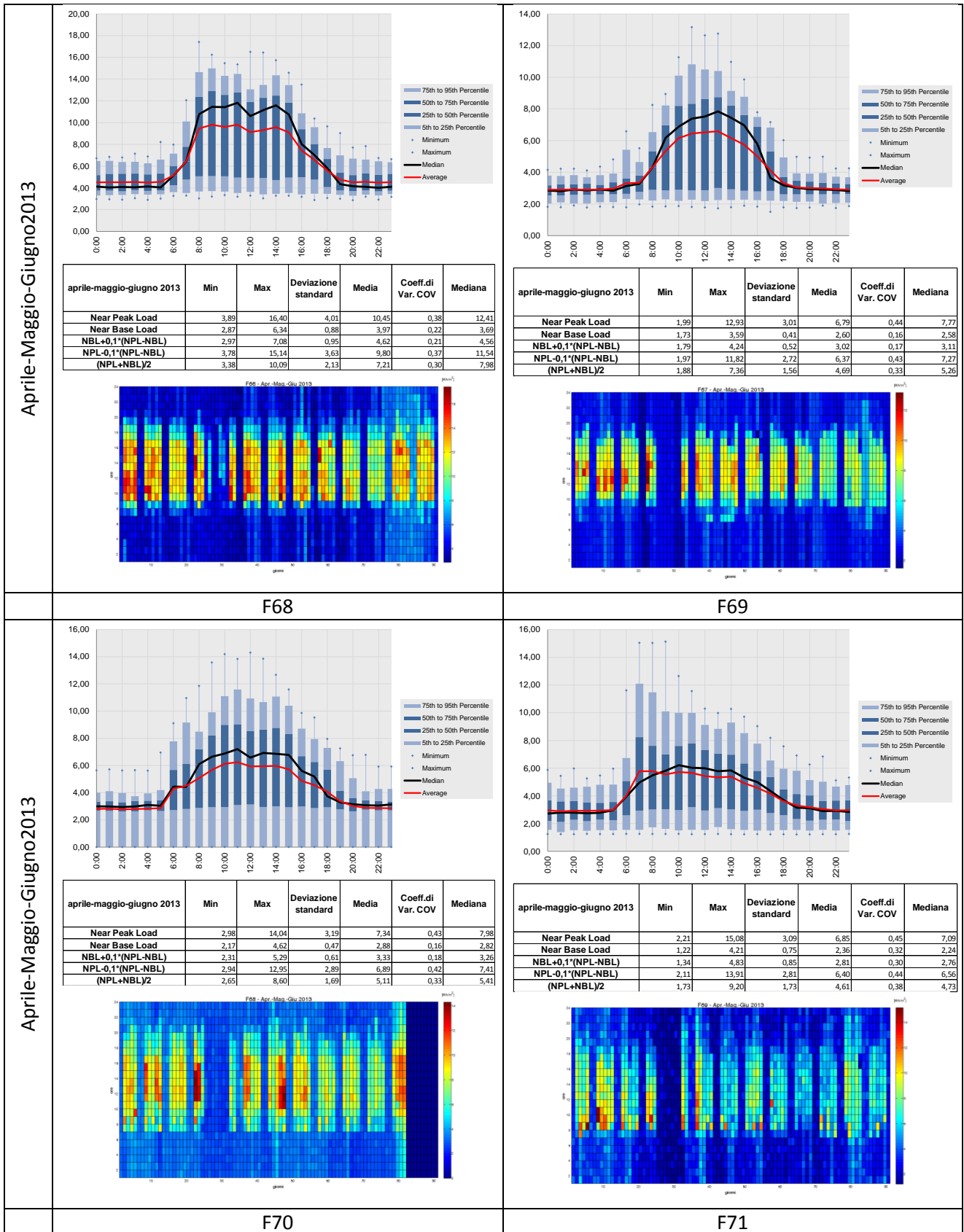
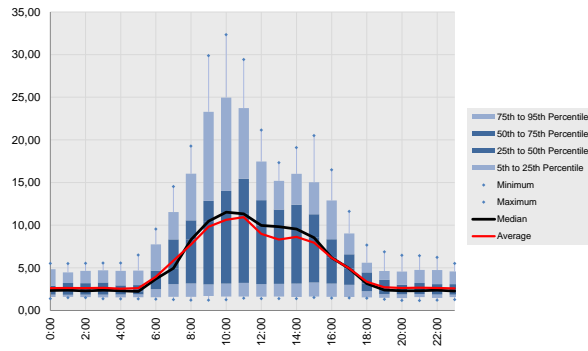


Tabella 52: Box-Plots e Carpet Plots dei consumi di energia elettrica per ciascuno degli edifici appartenenti al cluster nella stagione primaverile

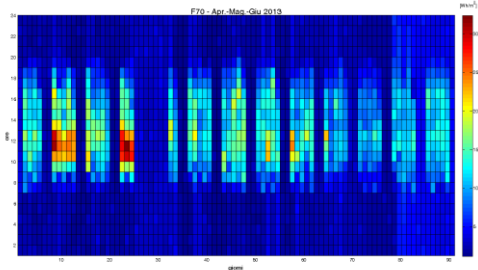
F66	F67
-----	-----



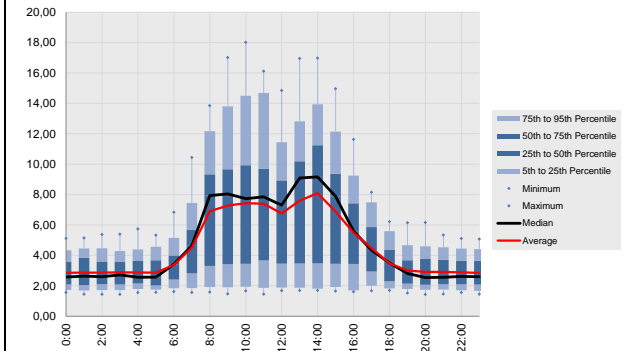
Aprile-Maggio-Giugno2013



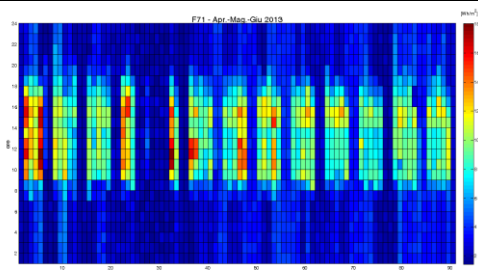
april-maggio-giugno 2013	Min	Max	Deviazione standard	Media	Coeff.di Var. COV	Mediana
Near Peak Load	1,91	30,92	7,47	11,42	0,65	12,28
Near Base Load	1,16	5,22	0,90	2,21	0,41	1,95
NBL+0,1*(NPL-NBL)	1,27	6,02	1,08	3,13	0,35	2,97
NPL-0,1*(NPL-NBL)	1,84	28,00	6,72	10,50	0,64	11,22
(NPL+NBL)/2	1,56	16,31	3,75	6,82	0,55	7,22



F70

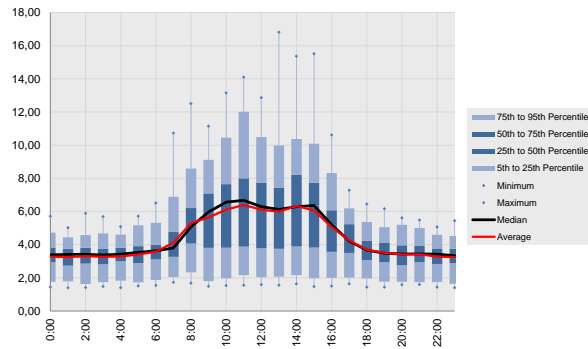


april-maggio-giugno 2013	Min	Max	Deviazione standard	Media	Coeff.di Var. COV	Mediana
Near Peak Load	1,98	17,44	4,31	8,43	0,51	9,67
Near Base Load	1,43	4,50	0,75	2,38	0,31	2,10
NBL+0,1*(NPL-NBL)	1,51	5,14	0,84	2,98	0,28	3,04
NPL-0,1*(NPL-NBL)	1,95	15,86	3,89	7,83	0,50	9,00
(NPL+NBL)/2	1,82	9,55	2,23	5,41	0,41	6,08

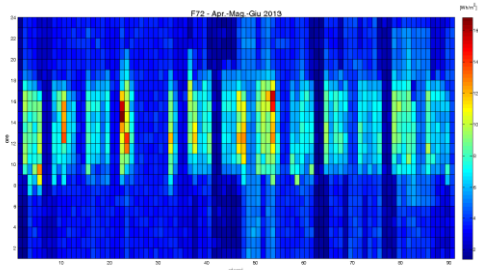


F71

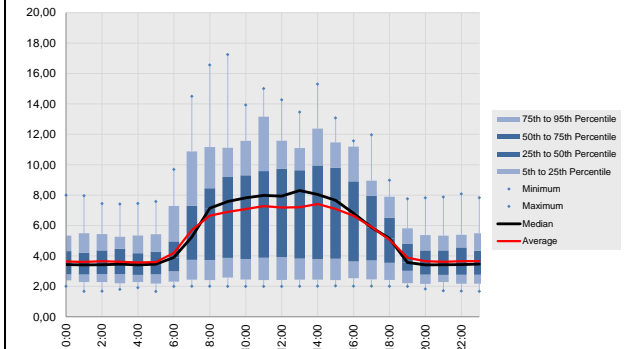
Aprile-Maggio-Giugno2013



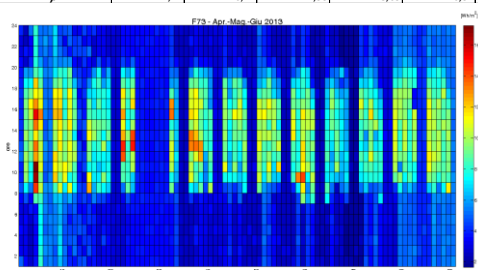
april-maggio-giugno 2013	Min	Max	Deviazione standard	Media	Coeff.di Var. COV	Mediana
Near Peak Load	2,19	15,98	2,94	6,90	0,43	7,06
Near Base Load	1,39	4,91	0,80	2,84	0,28	3,11
NBL+0,1*(NPL-NBL)	1,50	5,20	0,80	3,25	0,25	3,34
NPL-0,1*(NPL-NBL)	2,11	14,69	2,66	6,50	0,41	6,59
(NPL+NBL)/2	1,81	9,52	1,56	4,87	0,32	4,82



F72



april-maggio-giugno 2013	Min	Max	Deviazione standard	Media	Coeff.di Var. COV	Mediana
Near Peak Load	2,48	16,85	3,59	8,14	0,44	9,31
Near Base Load	1,66	5,00	0,89	3,17	0,28	2,89
NBL+0,1*(NPL-NBL)	2,04	5,63	0,94	3,66	0,26	3,44
NPL-0,1*(NPL-NBL)	2,43	15,52	3,25	7,64	0,43	8,70
(NPL+NBL)/2	2,24	10,21	1,93	5,65	0,34	6,02



F73

Tabella 53: Profili orari di carico medio stagionale di energia elettrica corrispondenti ai giorni feriali, ai giorni di sabato ed alle domeniche nella stagione primaverile

Aprile-Maggio-Giugno2013	<p style="text-align: center;">F66</p>	<p style="text-align: center;">F67</p>
	Aprile-Maggio-Giugno2013	<p style="text-align: center;">F68</p>
Aprile-Maggio-Giugno 2013		<p style="text-align: center;">F70</p>
	Aprile-Maggio-Giugno2013	<p style="text-align: center;">F72</p>

Tabella 54: Profili orari di carico medio mensile di energia elettrica nella stagione primaverile

