



**Politecnico  
di Torino**

**POLITECNICO DI TORINO**

**Corso di Laurea Magistrale  
in Ingegneria Energetica e Nucleare**

**Sistemi di Controllo Avanzato per l'Ottimizzazione Energetica  
ed Economica di Impianti Civili e Industriali**

**Relatori**

Prof. Marco Badami

**Correlatore**

Ing. Andrea Franzoso

**Candidato**

Dalila Calò





## **ABSTRACT**

*Gli edifici, nelle loro diverse destinazioni d'uso, rappresentano una quota significativa dei consumi energetici globali e delle emissioni di gas serra, rendendo necessario lo sviluppo di strategie per migliorare l'efficienza senza compromettere il comfort interno. La presente tesi analizza e confronta diversi metodi di controllo per la gestione termica ed energetica di edifici adibiti ad usi differenti fra loro, con l'obiettivo di fornire un quadro globale delle prestazioni dei controllori in contesti eterogenei.*

*È stato sviluppato un modello dinamico dell'edificio fisicamente coerente, capace di rappresentare l'inerzia termica dell'involucro e di integrarsi con il bilancio energetico elettrico che include la produzione fotovoltaica, scambi con la rete e gestione dello stato di carica della batteria. Su questa base sono stati sviluppati due differenti approcci di controllori avanzati.*

*L'analisi è condotta su casi studio con caratteristiche differenti: un edificio residenziale, un edificio industriale e un edificio intermedio di tipo scolastico, introdotto per valutare in modo graduale la capacità di adattamento dei controlli al variare di dimensioni, profili di carico e dinamiche operative.*

*Il primo è il Model Predictive Control (MPC), formulato come un problema di ottimizzazione vincolato su un orizzonte finito, in grado di considerare contemporaneamente dinamica termica, vincoli di comfort e costi energetici. Il secondo è un approccio di Deep Reinforcement Learning (DRL) che consente di apprendere una politica di controllo attraverso l'interazione ripetuta con l'ambiente, senza la necessità di un modello esplicito del sistema. Il DRL si distingue per la capacità di apprendere direttamente dal comportamento del sistema, adattandosi anche a dinamiche non lineari e difficili da modellare, come sono quelle degli edifici, perfezionando nel tempo le proprie decisioni sulla base dei risultati ottenuti. Per la sua implementazione è stato utilizzato un algoritmo Actor-Critic di tipo Deep Deterministic Policy Gradient (DDPG) in grado di apprendere politiche di controllo mediante reti neurali profonde.*

*Un ulteriore aspetto del lavoro riguarda l'applicazione del Transfer Learning. Un agente inizialmente addestrato sul caso residenziale è stato trasferito verso gli altri edifici, analizzando sia l'applicazione diretta della politica appresa, sia un successivo riaddestramento a partire dai parametri acquisiti. Questo ha permesso di valutare l'efficacia del trasferimento e la riduzione dei tempi di addestramento in scenari diversi da quello di partenza.*

*I risultati evidenziano che l'MPC garantisce prestazione stabili e prevedibili nel rispetto dei vincoli e nella minimizzazione dei costi energetici, con riduzioni dei costi fino al 10-15% rispetto agli approcci basati su apprendimento e ai controlli tradizionali nei diversi casi studio analizzati. Il DRL mostra, invece, maggiore flessibilità adattativa, soprattutto in scenari dinamici, ottenendo prestazioni complessivamente comparabili ma con una gestione meno ottimizzata delle risorse energetiche, in particolare per quanto riguarda l'utilizzo del sistema di accumulo. L'applicazione del Transfer Learning consente di accelerare la convergenza nei domini simili a quello di partenza, mantenendo prestazioni prossime a quelle di un agente addestrato da zero, mentre in presenza di differenze più marcate tra i domini, come nel caso studio industriale, le prestazioni risultano più sensibili alla necessità di una fase di riaddestramento.*

*In conclusione, il lavoro fornisce un confronto strutturato tra controllo predittivo e adattativo nel settore edilizio, evidenziando potenzialità, limiti e prospettive di scalabilità in applicazioni reali.*

## SOMMARIO

1. Introduzione.....	14
2. Metodologia.....	17
2.1. Modello del sistema energetico .....	17
2.1.1. Bilancio energetico e ipotesi di base .....	17
2.1.2. Modello a tre nodi .....	22
2.1.3. Modello del sistema di climatizzazione.....	24
2.1.4. Modello dell'impianto fotovoltaico e sistema di accumulo.....	29
2.2. Sistemi di controllo del modello energetico .....	31
2.2.1. Model Predictive Control (MPC) .....	31
2.2.2. Deep Reinforcement Learning.....	40
2.2.3. Transfer Learning .....	48
2.2.4. Controllo tradizionale PID + Rule-Based.....	50
3. Casi studio e analisi dei risultati .....	51
3.1. Caso studio residenziale .....	52
3.1.1. Controllo con MPC.....	54
3.1.2. Controllo con Deep Reinforcement Learning .....	59
3.1.3. Applicazione del Transfer Learning .....	63
3.1.4. Controllo tradizione PID + RB.....	68
3.1.5. Confronto delle strategie di controllo .....	70
3.2. Caso studio scolastico.....	73
3.2.2. Controllo con MPC.....	75
3.2.3. Controllo con Deep Reinforcement Learning .....	78
3.2.4. Applicazione del Transfer Learning .....	81
3.2.5. Controllo tradizione PID + RB.....	87
3.2.6. Confronto delle strategie di controllo.....	90

3.3.	Caso studio industriale .....	92
3.3.1.	Caso industriale 1 .....	95
3.3.2.	Caso Industriale 2 .....	112
4.	Discussione.....	129
4.1.	Prestazioni delle strategie di controllo.....	129
4.2.	Ruolo del sistema di accumulo e dell'autoconsumo .....	131
4.3.	Capacità di generalizzazione e Transfer Learning.....	132
5.	Conclusioni e sviluppi futuri .....	135
6.	Riferimenti.....	138

## ***APPENDICE FIGURE***

Figura 1:Sistema 3R3C del modello a 3 nodi .....	23
Figura 2: Descrizione funzionamento MPC .....	31
Figura 3: Orizzonte temporale MPC [16] .....	32
Figura 4: L'interazione agente-ambiente in un RL [20].....	40
Figura 5:Integrazione ambiente-agente nel nostro modello in Simulink.....	47
Figura 6: Temperatura esterna del periodo considerato .....	51
Figura 7:Andamento temperatura interna ed esterna con controllore MPC - caso studio residenziale .....	55
Figura 8:Andamento potenza elettrica consumata dalla pompa di calore con controllo MPC - caso studio residenziale .....	55
Figura 9: Andamento stato di carica della batteria con controllo MPC - caso studio residenziale .....	56
Figura 10:Andamento degli stati di carica e scarica della batteria con controllo MPC - caso studio residenziale .....	56
Figura 11:Andamento degli import ed export dalla rete con controllo MPC - caso studio residenziale .....	57
Figura 12: Bilancio della potenza elettrica con controllo MPC - caso studio residenziale .....	57
Figura 13:Bilancio orario della potenza elettrica con controllo MPC - caso studio residenziale .....	57
Figura 14:Andamento temperatura interna ed esterna con controllore DRL - caso studio residenziale .....	60
Figura 15:Bilancio della potenza elettrica con controllo DRL - caso studio residenziale .....	61
Figura 16:Bilancio orario della potenza elettrica con controllo DRL - caso studio residenziale .....	61
Figura 17:Andamento stato di carica della batteria con controllo DRL - caso studio residenziale .....	62

Figura 18:Andamento temperatura interna ed esterna con TL zero-shot - caso studio residenziale Livigno .....	63
Figura 19:Andamento temperatura interna con TL zero-shot - caso studio residenziale Livigno .....	64
Figura 20:Bilancio della potenza elettrica con TL zero-shot - caso studio residenziale Livigno .....	64
Figura 21:Bilancio orario della potenza elettrica con TL zero-shot - caso studio residenziale Livigno .....	65
Figura 22:Andamento temperatura interna ed esterna con TL fine-tuning - caso studio residenziale Livigno .....	66
Figura 23:Andamento temperatura interna ed esterna con TL fine-tuning - caso studio residenziale Livigno .....	66
Figura 24:Bilancio della potenza elettrica con TL zero-shot - caso studio residenziale Livigno .....	67
Figura 25:Bilancio orario della potenza elettrica con TL fine-tuning - caso studio residenziale Livigno .....	67
Figura 26:Andamento temperatura interna ed esterna con controllore PID + RB - caso studio residenziale .....	68
Figura 27:Andamento stato di carica della batteria con controllo PID + RB - caso studio residenziale .....	69
Figura 28:Bilancio della potenza elettrica con controllo PID + RB - caso studio residenziale .....	69
Figura 29:Bilancio orario della potenza elettrica con controllo PID + RB - caso studio residenziale .....	70
Figura 30:Andamento temperatura interna ed esterna con controllore MPC - caso studio scolastico.....	75
Figura 31:Andamento potenza elettrica consumata dalla pompa di calore con controllo MPC- caso studio scolastico.....	76
Figura 32: Andamento stato di carica della batteria con controllo MPC - caso studio scolastico .....	76

Figura 33:Andamento degli import ed export dalla rete con controllo MPC - caso studio scolastico.....	77
Figura 34:Bilancio della potenza elettrica con controllo MPC - caso studio scolastico .....	77
Figura 35:Bilancio orario della potenza elettrica con controllo MPC - caso studio scolastico .....	77
Figura 36:Andamento temperatura interna ed esterna con controllore DRL - caso studio scolastico.....	79
Figura 37:Bilancio della potenza elettrica con controllo DRL - caso studio scolastico .....	79
Figura 38:Bilancio orario della potenza elettrica con controllo DRL - caso studio scolastico .....	80
Figura 39:Andamento stato di carica della batteria con controllo DRL - caso studio scolastico .....	80
Figura 40:Andamento temperatura interna ed esterna con TL zero-shot - caso studio scolastico .....	82
Figura 41:Andamento stato di carica della batteria con TL zero-shot - caso studio scolastico .....	83
Figura 42:Bilancio della potenza elettrica con TL zero-shot - caso studio scolastico	83
Figura 43:Bilancio orario della potenza elettrica con TL zero-shot - caso studio scolastico .....	84
Figura 44:Andamento temperatura interna ed esterna con TL fine-tuning - caso studio scolastico .....	85
Figura 45:Andamento stato di carica della batteria con TL fine-tuning - caso studio scolastico .....	85
Figura 46:Bilancio della potenza elettrica con TL fine-tuning - caso studio scolastico .....	86
Figura 47:Bilancio orario della potenza elettrica con TL fine-tuning - caso studio scolastico .....	86

Figura 48:Andamento temperatura interna ed esterna con controllore PID + RB - caso studio scolastico.....	87
Figura 49:Andamento stato di carica della batteria con controllo PID + RB - caso studio scolastico.....	88
Figura 50:Bilancio della potenza elettrica con controllo PID + RB - caso studio scolastico .....	88
Figura 51:Bilancio orario della potenza elettrica con controllo PID + RB - caso studio scolastico .....	89
Figura 52:Andamento temperatura interna, esterna e di mandata UTA con controllore MPC - caso studio industriale1 .....	95
Figura 53:Andamento potenza elettrica consumata dalla pompa di calore con controllo MPC- caso studio industriale1 .....	96
Figura 54:Andamento stato di carica della batteria con controllo MPC - caso studio industriale1 .....	97
Figura 55:Andamento degli import ed export dalla rete con controllo MPC - caso studio industriale1 .....	97
Figura 56:Bilancio della potenza elettrica con controllo MPC - caso studio industriale1 .....	98
Figura 57:Bilancio orario della potenza elettrica con controllo MPC - caso studio industriale1 .....	98
Figura 58:Andamento temperatura interna, esterna e di mandata UTA con controllore DRL - caso studio industriale 1 .....	99
Figura 59:Andamento stato di carica della batteria con controllo DRL - caso studio industriale 1 .....	100
Figura 60:Bilancio della potenza elettrica con controllo DRL - caso studio industriale 1 .....	100
Figura 61:Bilancio orario della potenza elettrica con controllo DRL - caso studio industriale 1 .....	101
Figura 62:Andamento temperatura interna, esterna e di mandata UTA con TL zero-shot - caso studio industriale 1 .....	102

Figura 63: Bilancio della potenza elettrica con TL zero-shot - caso studio industriale 1 .....	103
Figura 64: Bilancio orario della potenza elettrica con TL zero-shot - caso studio industriale 1 .....	103
Figura 65: Andamento stato di carica della batteria con TL zero-shot - caso studio industriale 1 .....	104
Figura 66: Andamento temperatura interna, esterna e di mandata UTA con TL fine-tuning - caso studio industriale1 .....	105
Figura 67: Andamento stato di carica della batteria con TL fine-tuning - caso studio industriale 1 .....	105
Figura 68: Bilancio della potenza elettrica con TL fine-tuning - caso studio industriale 1 .....	106
Figura 69: Bilancio orario della potenza elettrica con TL fine-tuning - caso studio industriale 1 .....	106
Figura 70: Andamento temperatura interna ed esterna con controllore PID +RB - caso studio industriale 1 .....	107
Figura 71: Andamento stato di carica della batteria con controllo PID + RB - caso studio industriale 1 .....	108
Figura 72: Bilancio della potenza elettrica con controllo PID + RB - caso studio industriale 1 .....	108
Figura 73: Bilancio orario della potenza elettrica con controllo PID + RB - caso studio industriale 1 .....	109
Figura 74: Andamento temperatura interna, esterna e di mandata UTA con controllore MPC - caso studio industriale 2 .....	113
Figura 75: Andamento potenza elettrica consumata dalla pompa di calore con controllo MPC - caso studio industriale 2 .....	113
Figura 76: Andamento stato di carica della batteria con controllo MPC - caso studio industriale 2 .....	114
Figura 77: Andamento degli import ed export dalla rete con controllo MPC - caso studio industriale 2 .....	114

Figura 78: Bilancio della potenza elettrica con controllo MPC - caso studio industriale 2 .....	115
Figura 79: Bilancio orario della potenza elettrica con controllo MPC - caso studio industriale 2 .....	115
Figura 80: Andamento temperatura interna, esterna e di mandata UTA con controllore DRL - caso studio industriale 2 .....	116
Figura 81: Andamento stato di carica della batteria con controllo DRL - caso studio industriale 2 .....	117
Figura 82: Bilancio della potenza elettrica con controllo DRL - caso studio industriale 2 .....	117
Figura 83: Bilancio orario della potenza elettrica con controllo DRL - caso studio industriale 2 .....	118
Figura 84: Andamento temperatura interna, esterna e di mandata UTA con TL zero-shot - caso studio industriale 2 .....	120
Figura 85: Bilancio della potenza elettrica con TL zero-shot - caso studio industriale 2 .....	120
Figura 86: Bilancio orario della potenza elettrica con TL zero-shot - caso studio industriale 2 .....	121
Figura 87: Andamento stato di carica della batteria con TL zero-shot - caso studio industriale 2 .....	121
Figura 88: Andamento temperatura interna, esterna e di mandata UTA con TL fine-tuning - caso studio industriale 2 .....	122
Figura 89: Andamento stato di carica della batteria con TL fine-tuning - caso studio industriale 2 .....	123
Figura 90: Bilancio della potenza elettrica con TL fine-tuning - caso studio industriale 2 .....	123
Figura 91: Bilancio orario della potenza elettrica con TL fine-tuning - caso studio industriale 2 .....	124
Figura 92: Andamento temperatura interna, esterna e di mandata UTA con controllore PID +RB - caso studio industriale 2 .....	125

Figura 93:Andamento stato di carica della batteria con controllo PID +RB - caso studio industriale 2 .....	125
Figura 94:Bilancio della potenza elettrica con controllo PID +RB - caso studio industriale 2 .....	126
Figura 95:Bilancio orario della potenza elettrica con controllo PID +RB- caso studio industriale2 .....	126

## ***APPENDICE TABELLE***

Tabella 1:Parametri del modello energetico - caso residenziale.....	54
Tabella 2: Confronto fra le diverse strategie di controllo - caso studio residenziale.	71
Tabella 3:Confronto fra le diverse strategie di controllo - caso studio residenziale Livigno .....	72
Tabella 4: Parametri a confronto del modello energetico - caso residenziale e caso scolastico .....	75
Tabella 5:Confronto fra le diverse strategie di controllo - caso studio scolastico .....	90
Tabella 6:Parametri a confronto del modello energetico - caso residenziale e caso industriale .....	94
Tabella 7:Confronto fra le diverse strategie di controllo - caso studio industriale	1 110
Tabella 8:Confronto fra le diverse strategie di controllo - caso studio industriale	2 127

## ***LISTA DELLE ABBREVIAZIONI***

<b>Abbreviazione</b>	<b>Significato</b>
----------------------	--------------------

---

<b>MPC</b>	Model Predictive Control
<b>RL</b>	Reinforcement Learning
<b>DRL</b>	Deep Reinforcement Learning
<b>TL</b>	Transfer Learning
<b>UTA</b>	Unità di Trattamento Aria
<b>HVAC</b>	Heating, Ventilation and Air Conditioning
<b>SOC</b>	State of Charge
<b>PID</b>	Proportional Integral Derivative
<b>RB</b>	Rule-Based
<b>PV</b>	Photovoltaic

## ***1. Introduzione***

Il riscaldamento globale e il cambiamento climatico rappresentano oggi una delle sfide più rilevanti per l'umanità e per il pianeta. Negli ultimi anni, la comunità internazionale ha riconosciuto l'urgenza di affrontare questi fenomeni con politiche e accordi condivisi. In questo contesto, l'Accordo di Parigi del 2015 stabilisce l'obiettivo di contenere l'aumento della temperatura media globale ben al di sotto dei 2 °C rispetto ai livelli preindustriali, puntando idealmente a non superare 1,5 °C[1].

Gli edifici rivestono un ruolo centrale in questo scenario, poiché incidono in modo significativo sui consumi energetici e sulle emissioni di gas serra a livello globale. Secondo l'International Energy Agency, l'edilizia è responsabile di circa il 30 % del consumo finale di energia mondiale, contribuendo notevolmente alle emissioni di CO<sub>2</sub> [2] mentre le Nazioni Unite stimano circa il 40 % del consumo energetico totale globale e circa il 33 % delle emissioni complessive di anidride carbonica [3].

Le tendenze attuali indicano inoltre che, in assenza di misure efficaci, il fabbisogno energetico degli edifici continuerà ad aumentare entro il 2030. Questo scenario, unito agli effetti già evidenti del riscaldamento globale, sottolinea la necessità di intervenire in modo strutturale sul modo in cui l'energia viene utilizzata nel settore edilizio.

In questo contesto, l'ottimizzazione energetica degli edifici non è soltanto un obiettivo desiderabile, ma una necessità reale. Le politiche europee, in particolare la Energy Performance of Buildings Directive (EPBD), promuovono interventi per migliorare l'efficienza dell'involucro edilizio e dei sistemi impiantistici. Queste politiche incentivano l'uso di tecnologie ad alte prestazioni e l'integrazione di fonti energetiche rinnovabili[4].

Parallelamente, l'ottimizzazione dell'energia degli edifici tramite algoritmi sta assumendo un ruolo sempre più importante. I moderni sistemi energetici degli edifici sono infatti caratterizzati da una crescente complessità e da un comportamento fortemente dinamico, dovuto all'interazione tra impianti di generazione, distribuzione ed emissione dell'energia, fonti rinnovabili, sistemi di accumulo e profili di utilizzo variabili. In questo scenario, l'impiego di metodologie algoritmiche avanzate permette una gestione più efficiente e adattiva dell'energia, migliorando le prestazioni complessive degli edifici e contribuendo al raggiungimento degli obiettivi di sostenibilità energetica [5][6].

La letteratura presenta varie strategie di controllo per gestire l'energia negli edifici. I controllori tradizionali, come i regolatori PID, sono ampiamente usati per la loro semplicità e affidabilità, ma hanno limiti significativi nella gestione di sistemi complessi e multivariati [7].

Per superare queste limitazioni, negli anni sono stati sviluppati controllori avanzati, tra cui i controllori predittivi o adattativi. Per i controllori predittivi, il Model Predictive Control (MPC) si è dimostrato particolarmente efficace nel ridurre i consumi energetici mantenendo il comfort interno. L'MPC utilizza un modello del sistema per prevedere il comportamento futuro e determinare le azioni di controllo migliori, rispettando i vincoli operativi e di comfort. Studi come quello di Ma et al.[6], hanno dimostrato che l'MPC può ridurre notevolmente i consumi energetici rispetto ai controlli normali, mantenendo buoni livelli di comfort anche in caso di disturbi e sfruttando l'inerzia termica dell'edificio. Oldewurtel et al. [8] hanno proposto un metodo per valutare la flessibilità energetica degli edifici. Utilizzando un controllo basato su Model Predictive Control e segnali di prezzo dell'energia, gli autori analizzano la capacità di un edificio di spostare nel tempo il proprio consumo energetico, ad esempio spostando parte dei consumi verso le ore in cui il costo dell'energia è minore, mantenendo al contempo le condizioni di comfort interno

Più recentemente, approcci basati su controllori adattativi come il Reinforcement Learning (RL) hanno attirato sempre più attenzione. Questi metodi possono apprendere politiche di controllo ottimali semplicemente interagendo con l'ambiente, senza bisogno di un modello del sistema. In questi approcci, un agente di controllo aggiorna gradualmente la propria strategia attraverso un processo di 'trial and error', selezionando azioni di controllo e valutandone gli effetti sul sistema attraverso un segnale di ricompensa (reward), adattandosi così alle dinamiche operative dell'edificio.

Nei metodi Deep Reinforcement Learning (DRL), questo processo di apprendimento è supportato da reti neurali profonde che permettono di affrontare problemi con dinamiche complesse e non lineari. Per questo, i controlli basati sull'approccio di apprendimento sono di tipo adattativo, in quanto l'agente apprende in autonomia in risposta ai cambiamenti delle condizioni operative, migliorando progressivamente le prestazioni del controllo in funzione della ricompensa ottenuta.

Wei et al.[9]hanno dimostrato che tecniche di Deep Reinforcement Learning possono essere applicate con successo per il controllo di impianti HVAC, in cui mantenendo i livelli di comfort si ottengono riduzioni dei consumi energetici. Infine, un'analisi sistematica degli approcci di Reinforcement Learning per la gestione energetica degli edifici è presentata da Vázquez-Canteli e Nagy [10], nel quale gli autori evidenziano sia il potenziale di tali metodi sia le principali difficoltà, tra queste i tempi di addestramento molto lunghi e la difficoltà di riutilizzare politiche di controllo apprese in edifici diversi.

In questo contesto, il Transfer Learning si presenta come una strategia promettente per migliorare la scalabilità e l'efficienza dei controllori basati su apprendimento. L'idea è riutilizzare, in tutto o in parte, le conoscenze acquisite da un agente già addestrato, applicandole al controllo di altri edifici diversi che possono avere geometrie, destinazioni d'uso o sistemi energetici differenti.

Studi recenti, come quello di Zhang et al. [11], hanno dimostrato che trasferire modelli o politiche di controllo può ridurre notevolmente i tempi di addestramento e migliorare le prestazioni favorendo l'applicazione dei metodi di Deep Reinforcement Learning in contesti reali.

Sulla base di queste considerazioni, il presente lavoro analizza e confronta diverse strategie di controllo avanzate per la gestione energetica degli edifici, con particolare attenzione al confronto tra approcci predittivi e basati su apprendimento. L'analisi viene condotta su tre edifici con caratteristiche e destinazioni d'uso differenti, valutando anche il potenziale del transfer learning nel trasferimento delle politiche di controllo tra domini diversi, al fine di individuare la soluzione più efficace dal punto di vista energetico ed economico.

Nel Capitolo 2 si esaminerà il bilancio energetico dei tre edifici oggetto di studio e si svilupperanno i modelli utilizzati per l'implementazione delle strategie di controllo. Questo capitolo fornirà le basi teoriche e modellistiche necessarie alla descrizione dei casi studio, che verranno approfonditi nel Capitolo 3.

I Capitoli 4 e 5 saranno dedicati all'analisi e alla discussione dei risultati ottenuti. In questa fase, si potranno confrontare le prestazioni dei controllori adottati, evidenziando i principali vantaggi e limiti in termini di efficienza energetica ed economica, comportamento dinamico e capacità di adattamento.

## ***2. Metodologia***

### ***2.1. Modello del sistema energetico***

#### ***2.1.1. Bilancio energetico e ipotesi di base***

Per garantire il comfort termoigrometrico agli occupanti, occorre determinare la quantità di calore necessaria a mantenere una temperatura interna confortevole negli edifici durante tutto l'anno, in base alle stagioni.

Questo calore deve compensare le dispersioni termiche attraverso l'involucro edilizio rappresentate da:

- **Trasmissione:** perdita di calore attraverso pareti, tetti, pavimenti e finestre a causa della differenza di temperatura tra interno ed esterno.
- **Ventilazione e infiltrazioni:** scambi d'aria tra interno ed esterno dovuti a ricambi controllati (impianti di ventilazione) o incontrollati (spifferi, fessure, porte aperte).

Oltre a queste dispersioni non dobbiamo trascurare l'inerzia termica nei materiali edilizi che permette di accumulare calore al loro interno e i guadagni termici gratuiti come gli apporti interni (persone, elettrodomestici) e l'irraggiamento solare diretto o sulle superfici opache.

Andiamo a determinare il bilancio energetico dell'edificio considerando il primo principio della termodinamica. Il principio stabilisce che l'energia non si crea né si distrugge, ma si trasforma. Il bilancio energetico è quindi una rappresentazione di questa legge applicata a un sistema complesso come un edificio in cui l'energia entrante deve essere uguale a quella uscente.

Iniziamo determinando la zona interna come un volume di controllo con temperatura uniforme  $T_i(t)$  e capacità volumetrica  $C_v$  ( $J/m^3K$ ).

$$C_v \frac{dT_i}{dt} = \sum q_{in} - \sum q_{out} \quad (1)$$

Dove gli apporti uscenti  $q_{out}$  sono rappresentati dalle perdite per trasmissione e ventilazione, dovute allo scambio termico con l'esterno e  $q_{in}$ , invece, i guadagni interni, solari e dell'impianto di climatizzazione.

Tutti i contributi sono espressi come potenza volumetrica [W/m<sup>3</sup>].

### ***Perdite per trasmissione e ventilazione***

Le dispersioni attraverso l'involucro sono descritte dall'equazione:

Dove  $\dot{Q}_T$  è il flusso scambiato per trasmissione [W],  $U$  è la trasmittanza degli elementi

$$\dot{Q}_T = U \cdot S \cdot (T_e(t) - T_i(t)) \quad (2)$$

[W/m<sup>2</sup>] e  $S$  la superficie disperdente [m<sup>2</sup>].

Questo contributo deve essere considerato sia per gli elementi opachi sia per quelli trasparenti, poiché hanno proprietà termiche differenti.

La trasmittanza rappresenta la quantità di energia che attraversa un metro quadrato di superficie per ogni grado di differenza di temperatura tra interno ed esterno ed è definita come l'inverso della resistenza termica  $R$ . Quest'ultima è determinata a partire dalle proprietà dei materiali costruttivi, generalmente riportata nella scheda tecnica del produttore.

$$U = \frac{1}{\frac{1}{h_e} + \sum_{i=1}^n \frac{s_i}{\lambda_i} + \sum_{j=1}^n R_j + \frac{1}{h_i}} \quad (3)$$

Dove:

- $h_i$  e  $h_e$  sono i coefficienti di scambio termico liminari rispettivamente interno ed esterno. Per i valori si è tenuto conto della norma UNI EN 6946
- $R_j$  è la resistenza termica del j-esimo elemento non omogeneo
- $\lambda_i$  è la conducibilità termica della i-esimo elemento omogeneo
- $s_i$  è lo spessore dello i-esimo elemento omogeneo

Per la trasmittanza delle finestre si considera:

$$U_w = \frac{(U_g \cdot A_g + U_f \cdot A_f + \Psi_g * l_g)}{(A_g + A_f)} \quad (4)$$

Dove:

- $U_g$  è la trasmittanza termica del vetro [W/m<sup>2</sup>].
- $A_g$  è l'area del vetro [m<sup>2</sup>].
- $U_f$  è la trasmittanza termica del telaio [W/m<sup>2</sup>].
- $A_f$  è l'area del telaio [m<sup>2</sup>].
- $l_g$  è il perimetro totale della vetrata [W/m].
- $\Psi_g$  è la trasmittanza termica lineica del ponte termico dovuta al distanziatore [m].

Le dispersioni per ventilazione ( $\dot{Q}_{ve}$ ) sono espresse come:

$$\dot{Q}_{ve} = H_{ve} (T_e(t) - T_i(t)) \quad (5)$$

Dove  $H_{ve}$  è il coefficiente di perdita per ventilazione [W/K].

Il coefficiente di perdite per ventilazione  $H_{ve}$  è determinato a partire dalla portata di aria di rinnovo che dipende dalla tipologia di edificio e dalla normativa tecnica di riferimento adottata. Per edifici ad uso residenziale e commerciale, i valori di riferimento sono generalmente assunti dalle norme sulla qualità dell'aria interna, quali la UNI EN 16798-1. Nel caso di edifici uso industriale, la determinazione delle portate di ventilazione è invece basata sulla valutazione dei rischi specifici presenti (ad esempio carichi inquinanti, polveri, vapori o calore), facendo riferimento alla normativa sulla sicurezza nei luoghi di lavoro e ai criteri di buona tecnica.

Una volta determinata la portata di ventilazione in funzione del caso specifico, il coefficiente di perdite per ventilazione è calcolato mediante la relazione:

$$H_{ve} = \frac{Q_{ve}}{1000} \cdot \rho \cdot c_p \quad (6)$$

Dove  $\rho$  e  $c_p$  sono rispettivamente la densità e il calore specifico dell'aria.

Le modalità di determinazione delle portate di ventilazione adottate nei diversi casi studio considerati in questo lavoro sono descritte in dettaglio nei successivi sottocapitoli.

### ***Guadagni interni***

I guadagni termici interni comprendono i contributi dovuti alla presenza di occupanti, apparecchiature, impianti e processi interni. Tali apporti possono includere componenti sia di calore sensibile sia di calore latente e risultano fortemente dipendenti dalla tipologia di edificio e dalle attività svolte al suo interno.

Poiché le modalità di valutazione non sono univoche e variano tra edificio ad uso domestico, commerciale e industriale, nel presente capitolo essi vengono definiti in modo specifico nei successivi capitoli dedicati ai casi studio in funzione della destinazione d'uso.

Il contributo complessivo dei guadagni interni viene considerato nel bilancio energetico dell'edificio come termine additivo.

### ***Guadagni solari***

I guadagni solari sono suddivisi in due contributi principali: il guadagno solare diretto attraverso le superfici trasparenti e il guadagno solare assorbito dalle superfici opache dell'involucro edilizio.

Il primo è il guadagno solare diretto attraverso le superfici vetrate:

$$\dot{Q}_{sol,dir,t} = g \cdot A_v \cdot G_d(t) \quad (7)$$

Dove  $g$  il fattore solare del vetro,  $A_v$  è l'area vetrata,  $G_d(t)$  è l'irradianza solare diretta sul piano della finestra.

Il secondo è il guadagno solare assorbito dalle superfici opache:

$$\dot{Q}_{sol,op,t} = \alpha \cdot A_{op} \cdot Irr_{op}(t) \quad (8)$$

Dove  $\alpha$  è l'assorbimento della parete opaca,  $A_{op}$  è l'area opaca ( $m^2$ ) e  $Irr_{op}(t)$  è l'irradianza solare incidente sulla superficie opaca, data dalla somma della componente diretta e riflessa [ $W/m^2$ ].

### **Forma finale del bilancio energetico**

Tutti questi contributi permettono di riformulare l'equazione **Errore. L'origine riferimento non è stata trovata.** passando da un bilancio energetico volumetrico ad uno globale dell'edificio, ottenendo:

$$C \frac{dT_i}{dt} = \sum_i^N U_i \cdot S_i + H_{ve} (T_e(t) - T_i(t)) + \dot{Q}_{int,tot} + \dot{Q}_{sol,dir,t} + Q_{sol,op,t} + \dot{Q}_{HC} \quad (9)$$

Sostituiamo:

$$C \frac{dT_i}{dt} = \sum_i^N U_i \cdot A_{disp,i} + H_{ve} (T_e(t) - T_i(t)) + \dot{Q}_{int,tot} + g \cdot A_v \cdot G_d(t) + \alpha \cdot A_{op} \cdot Irr_{op}(t) + \dot{Q}_{HC} \quad (10)$$

Dove:

- $\dot{Q}_{HC}$  è il contributo dell'impianto di climatizzazione per mantenere il comfort interno [ $W/m^3$ ].
- $A_v$  è l'area degli elementi trasparenti;
- $A_{op}$  è l'area degli involucri opachi
- $A$  è l'area di tutte le superfici opache disperdenti verso l'esterno

La forma finale della relazione rappresenta quindi il bilancio energetico dinamico dell'edificio, ottenuto dall'applicazione della prima legge della termodinamica.

### **2.1.2. Modello a tre nodi**

La modellazione termica degli edifici può essere affrontata con diversi livelli di complessità, che variano da modelli a nodo singolo, nei quali l'intero edificio è rappresentato mediante una sola capacità termica equivalente, fino a modelli multinodo che consentono una descrizione più dettagliata dell'inerzia dell'edificio.

I modelli a uno o due nodi permettono una rappresentazione più semplificata della dinamica termica, ma risultano limitati nella descrizione dei fenomeni transitori. In particolare, concentrando l'intera massa termica in un unico nodo equivalente, tali modelli tendono a sovrastimare o sottostimare i tempi di risposta del sistema.

Per garantire un compromesso tra accuratezza fisica e complessità computazionale, è stato adottato un modello a tre nodi, in cui vengono distinti in aria interna, superficie interna della parete e superficie esterna della parete.

L'edificio è descritto da una rete 3R3C in cui:

- tra aria interna e superficie interna è presente una resistenza di convezione dell'aria interna ( $R_i$ )
- tra superficie interna e superficie esterna è presente la resistenza di conduzione della parete ( $R_w$ );
- tra superficie esterna e ambiente esterno è presente una resistenza di convezione esterna ( $R_e$ ).

La capacità termica della parete viene ripartita sui due nodi superficiali (avremo  $C_{wi}$ ,  $C_{we}$ ): per ciascun nodo si considera metà della capacità termica totale della stratigrafia, in modo da rappresentare in maniera più accurata l'inerzia termica e la propagazione del calore attraverso l'involucro.

Il sistema RC è riportato in Figura 1: Sistema 3R3C del modello a 3 nodi

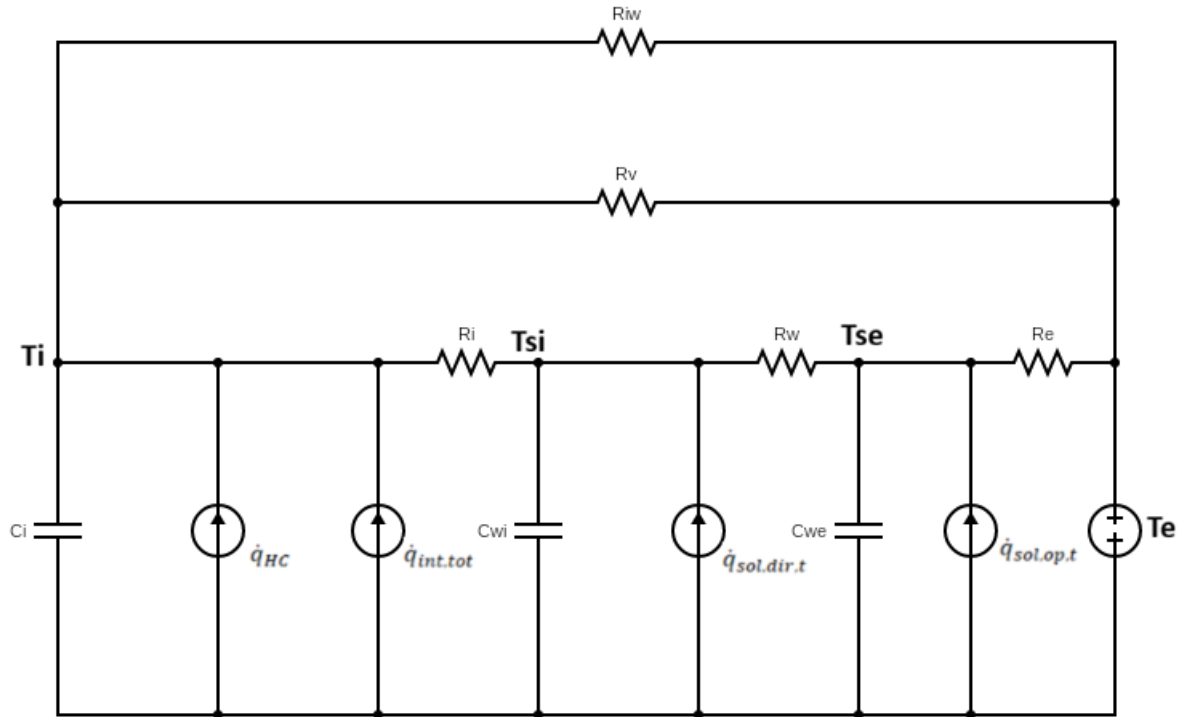


Figura 1: Sistema 3R3C del modello a 3 nodi

Le equazioni che descrivono il modello saranno:

$$\frac{dT_i}{dt} = -\frac{1}{R_i * C_i} (T_i - T_{sup,i}) - \frac{H_{ve}}{C_i} (T_i - T_e) + \frac{H_{win}}{C_i} (T_i - T_e) + \frac{q_{int}}{C_i} + \frac{q_{HC}}{C_i} \quad (11)$$

$$\frac{dT_{sup,i}}{dt} = -\frac{1}{R_w * C_{w,i}} (T_{sup,i} - T_{sup,e}) - \frac{1}{R_i * C_{w,i}} (T_i - T_{sup,i}) + \frac{g * Area_v}{C_{w,i}} Gd(t) \quad (12)$$

$$\frac{dT_{sup,e}}{dt} = +\frac{1}{R_w * C_{w,e}} (T_{sup,i} - T_{sup,e}) - \frac{1}{R_e * C_{w,e}} (T_{sup,e} - T_e) + \frac{alpha * Area_{op}}{C_w} * Irr_{op}(t) \quad (13)$$

Le matrici  $A_d$ ,  $B_d$  e  $E_d$  saranno:

$$A_d = \begin{bmatrix} -\left(\frac{1}{R_i} + H_{ve} + H_{win}\right) & \frac{1}{R_i \cdot C_i} & 0 \\ \frac{1}{R_i \cdot C_{w,i}} & -\left(\frac{1}{R_i} + \frac{1}{R_w}\right) & \frac{1}{R_w \cdot C_{w,i}} \\ 0 & \frac{1}{R_w \cdot C_{w,e}} & -\left(\frac{1}{R_e} + \frac{1}{R_w}\right) \end{bmatrix}$$

$$B_d = \begin{bmatrix} 1 \\ C_i \\ 0 \\ 0 \end{bmatrix}$$

$$E_d = \begin{bmatrix} \frac{H_{ve} + H_{win}}{C_i} & 0 & 0 & \frac{1}{C_i} \\ 0 & 0 & \frac{g \cdot Area_v}{C_{w,i}} & 0 \\ \frac{1}{R_e \cdot C_{w,e}} & \frac{alpha \cdot Area_{op}}{C_{w,e}} & 0 & 0 \end{bmatrix}$$

Rispetto ai modelli più semplici, il modello a tre nodi consente una descrizione più realistica della dinamica termica dell'edificio, mantenendo al tempo stesso una struttura lineare compatibile con l'implementazione dei sistemi di controllo. Per questo il modello a tre nodi è adottato come unico modello di riferimento per le simulazioni e per l'integrazione con le strategie di controllo.

### 2.1.3. Modello del sistema di climatizzazione

#### *Calcolo del carico termico in regime invernale*

Il dimensionamento del sistema di riscaldamento è stato effettuato tenendo conto delle condizioni invernali di progetto. L'obiettivo è determinare la potenza termica necessaria a garantire il mantenimento della temperatura interna di comfort anche nelle condizioni climatiche più sfavorevoli.

Per il calcolo del carico termico invernale sono state adottate alcune ipotesi semplificative finalizzate a stimare in modo conservativo il fabbisogno di riscaldamento. In particolare, l'analisi è stata condotta assumendo un regime termico stazionario, ipotizzando una temperatura costanti nel tempo e trascurando gli effetti dinamici legati ai transitori termici.

La temperatura esterna di progetto  $\theta_e^*$  è stata assunta pari a  $-8^\circ\text{C}$ , che rappresenta le condizioni climatiche più rigide per la località considerata, Torino, secondo la norma UNI EN 12831. Questo assicura un corretto dimensionamento dell'impianto anche nei periodi più sfavorevoli. Inoltre, non sono stati considerati i guadagni termici gratuiti, come l'irraggiamento solare e gli apporti interni dovuti alla presenza di persone e apparecchiature elettriche. Questa evita il rischio di sottodimensionare l'impianto e di garantisce il mantenimento delle condizioni di comfort anche in assenza di contributi favorevoli.

Il carico termico invernale è stato quindi determinato considerando esclusivamente:

- le perdite per trasmissione attraverso l'involucro edilizio, valutate sulla base delle trasmittanze termiche delle superfici disperdenti;
- le perdite per ventilazione, calcolate in funzione dei ricambi d'aria previsti.

L'equazione di bilancio per la determinazione del carico termico sensibile di progetto in regime invernale si riduce alla seguente espressione:

$$\tilde{\Phi}_p^* = -\Phi_c - \Phi_a \quad (14)$$

con  $\Phi_c$  flusso termico disperso attraverso i componenti d'involucro per trasmissione e  $\Phi_a$  flusso termico connesso alla ventilazione.

Il segno negativo indica che entrambi i contributi rappresentano perdite di energia dal volume riscaldato. Il valore  $\tilde{\Phi}_p^*$  corrisponde pertanto alla potenza termica che il sistema di riscaldamento deve fornire per mantenere le condizioni interne di progetto in regime stazionario.

La norma UNI 12831 fornisce un metodo dettagliato per il calcolo delle perdite di calore per trasmissione, introducendo coefficienti specifici che tengono conto della

complessità degli scambi termici tra l'ambiente riscaldato e le zone adiacenti. Tali coefficienti sono indicati come coefficienti di perdita per trasmissione  $H_T$ .

Il flusso termico disperso per:

$$\phi_c = (H_{T,e} + H_{T,nr} + H_{T,i} + H_{T,t}) \cdot (\tilde{\theta}_i - \theta_e^*) \quad (15)$$

dove:

- $\tilde{\theta}_i$  è la temperatura interna media (°C).
- $\theta_e^*$  è la temperatura esterna di progetto, che rappresenta il valore medio più basso atteso (°C).

Le perdite di calore dovute alla ventilazione, che contribuiscono in modo significativo al carico termico invernale, sono state valutate mediante il coefficiente di perdita per ventilazione  $H_{ve}$ , secondo la relazione:

$$\phi_a = H_{ve} * (\tilde{\theta}_i - \theta_e^*) \quad (16)$$

Il valore di  $\tilde{\Phi}_p^*$  rappresenta la potenza termica utile richiesta dall'edificio nelle condizioni più sfavorevoli. Tale valore non tiene conto delle perdite associate al sistema impiantistico di emissione, distribuzione e regolazione del calore, che vengono considerate separatamente nella fase successiva di dimensionamento del sistema di generazione, in accordo con la metodologia prevista dalla norma tecnica.

### ***Scelta del sistema di generazione e considerazione delle perdite***

Il valore di carico termico invernale ottenuto rappresenta la potenza termica utile necessaria a compensare le perdite dell'edificio nelle condizioni di progetto. Sulla base di tale valore, nel presente lavoro si è scelto di adottare una pompa di calore acqua-acqua, alimentata da un pozzo geotermico, come sistema di generazione del calore.

Una pompa di calore è un sistema di generazione che consente di trasferire calore da una sorgente esterna (aria, acqua, terreno) all'interno dell'edificio mediante un ciclo termodinamico a circuito chiuso che utilizza solitamente un fluido refrigerante ed è

alimentato elettricamente. Questa tecnologia è molto usata negli edifici in quanto consente di sfruttare fonti rinnovabili, riducendo l'uso di combustibili fossili.

Per il funzionamento invernale, la pompa sottrae calore alla sorgente fredda e lo cede al fluido dell'impianto di riscaldamento, che può essere costituito da aria o acqua, in base al sistema di riscaldamento adottato.

La scelta di pompa di calore di tipo acqua-acqua è motivata dalla maggiore stabilità della sorgente termica rispetto ai sistemi aria-acqua, in quanto la temperatura dell'acqua di falda risulta meno soggetta a variazioni stagionali. Questa caratteristica consente un funzionamento più regolare del generatore e prestazioni più affidabili durante l'intero periodo di riscaldamento.

Ai fini delle simulazioni e del dimensionamento energetico, alla pompa di calore è stato associato un coefficiente di prestazione (COP) pari a 2. Tale valore non deriva da un dato di targa nominale, ma è stato assunto sulla base dell'analisi di un caso reale osservato durante un'attività di tirocinio, relativo a un impianto acqua-acqua effettivamente installato e operante in condizioni comparabili a quelle oggetto di studio. Il valore adottato è da considerarsi cautelativo e rappresentativo del funzionamento reale del sistema.

Si precisa che il coefficiente di prestazione (COP) non influisce sulla determinazione della potenza termica di progetto, ma viene utilizzato esclusivamente per la valutazione della potenza elettrica assorbita e dei consumi energetici del sistema di generazione.

Il carico termico utile calcolato non coincide direttamente con la potenza che il generatore deve effettivamente fornire, in quanto occorre tener conto delle perdite associate al sistema impiantistico di emissione, distribuzione e regolazione del calore. Secondo la UNI/TS 11300-2, tali perdite possono essere ricondotte a un rendimento globale dell'impianto, espresso come prodotto dei rendimenti parziali:

$$\eta_{\text{impianto}} = \eta_e \cdot \eta_d \cdot \eta_r \quad (17)$$

Nel presente lavoro, la pompa di calore costituisce il sistema di generazione di calore a servizio dell'Unità di Trattamento Aria (UTA), che provvede sia al riscaldamento dell'aria immessa negli ambienti sia alla ventilazione degli stessi. Nei sistemi di

riscaldamento ad aria, le perdite di emissione risultano generalmente più elevate rispetto ai sistemi radianti, a causa di fenomeni di stratificazione termica, differenze tra temperatura dell'aria e temperatura operante e dispersioni associate alle distribuzioni dell'aria nei locali.

Studi sperimentali presenti in letteratura evidenziano come le perdite di emissione dipendono dalla tipologia di sistema di riscaldamento. In particolare, per sistemi a radiatori le perdite di emissione risultino dell'ordine del 2,3 %, comprese tra 5,5% e 9,9% per sistemi radianti a pavimento, e possono raggiungere valori prossimi ai 9,9% per sistemi di riscaldamento ad aria [12]. In accordo con tali indicazioni, in questo lavoro il rendimento di emissione del sistema UTA è stato assunto pari a:

$$\eta_e = 0,90 \quad (18)$$

In assenza di informazione di dettaglio sull'impianto di distribuzione e sul sistema di regolazione, sono stati valori rappresentativi di un impianto moderno correttamente progettato, in accordo con i valori indicativi forniti dalla UNI/TS 11300-2:

$$\eta_d = 0,94 \quad \eta_r = 0,95 \quad (19)$$

Il rendimento globale dell'impianto risulta pertanto pari a:

$$\eta_{impianto} = 0,90 \cdot 0,94 \cdot 0,95 \approx 0,8 \quad (20)$$

La potenza termica di progetto richiesta al sistema di generazione è quindi calcolata come:

$$\tilde{\Phi}_{P,gen}^* = \frac{\tilde{\Phi}_P^*}{\eta_{impianto}} \quad (21)$$

Il valore così ottenuto rappresenta la potenza termica effettivamente richiesta al generatore e costituisce il riferimento per il dimensionamento della pompa di calore, la

quale alimenta la batteria di riscaldamento dell'Unità di Trattamento Aria. Questa configurazione garantisce il mantenimento delle condizioni di comfort interno anche in presenza delle principali perdite impiantistiche associate ai sistemi di climatizzazione ad aria.

#### ***2.1.4. Modello dell'impianto fotovoltaico e sistema di accumulo***

Il calcolo del dimensionamento dell'impianto fotovoltaico e del sistema di accumulo elettrico ha come obiettivo il miglioramento dell'integrazione tra produzione e consumo di energia all'interno dell'edificio. Questo riduce il prelievo dalla rete elettrica e aumenta l'autoconsumo dell'energia prodotta localmente. Questo aspetto è particolarmente importante nel contesto di questo lavoro, che include la presenza di una pompa di calore elettrica, il cui funzionamento incide in modo significativo sui consumi complessivi.

Per definire criteri di dimensionamento realistici, sono stati presi in considerazione studi presenti in letteratura che analizzano il comportamento dei sistemi fotovoltaici residenziali con batterie di accumulo, come il lavoro di J. Weniger et al. [13]. In particolare, l'articolo esaminato dimostra come il rapporto tra potenza fotovoltaica installata, capacità di accumulo e profilo dei consumi giochi un ruolo cruciale nel raggiungimento di elevati livelli di autoconsumo e autosufficienza energetica [13].

Dall'analisi risulta che aumentare la potenza fotovoltaica e della capacità della batteria porta inizialmente a un miglioramento significativo delle prestazioni del sistema. Tuttavia, oltre certe soglie, i benefici aggiuntivi diventano progressivamente più limitati. In altre parole, un sovradimensionamento eccessivo dell'impianto o dell'accumulo non garantisce un uso più efficiente dell'energia prodotta, poiché la batteria tende a rimanere parzialmente inutilizzata durante i cicli giornalieri [13].

In particolare, lo studio sottolinea come, per utenze residenziali, una potenza fotovoltaica proporzionata ai consumi annui consente già di ottenere buoni livelli di autoconsumo. L'aggiunta di una batteria di accumulo moderata permette di incrementare ulteriormente l'energia fotovoltaica utilizzata direttamente nell'edificio, spostando parte dei consumi verso le ore serali e notturne. Al contrario, avere una capacità di accumulo eccessive comporta vantaggi marginali, a fronte di una maggiore complessità del sistema [13].

Alla luce di queste considerazioni, il dimensionamento dell'impianto fotovoltaico e del sistema di accumulo è stato effettuato seguendo un approccio equilibrato. Questo si basa sulla relazione tra consumi elettrici annuali dell'edificio, produzione fotovoltaica e capacità della batteria. L'obiettivo non è quello di massimizzare indiscriminatamente la produzione o l'accumulo, ma di individuare una configurazione che consente un utilizzo efficace dell'energia rinnovabile prodotta, compatibile con il profilo di carico dell'edificio e con le strategie di gestione energetica adottate [13].

Indicazioni simili emergono anche in studi non residenziali, quali edifici pubblici o strutture del settore terziario, come mostrato nel lavoro di U. Acar e O. Kaska[14]. In questi casi, la dimensione dei sistemi fotovoltaici e di accumulo non è orientata dalla massima autosufficienza energetica, ma a un incremento tecnico-economico tra incremento dell'autoconsumo e sostenibilità dell'investimento. Lo studio mostra che l'aumento della capacità di accumulo porta a miglioramenti nei tassi di autoconsumo e autosufficienza, ma con rendimenti decrescenti man mano che cresce la dimensione della batteria. Allo stesso tempo, le analisi economiche indicano una diminuzione progressiva del Valore Attuale Netto (VAN) con l'aumento della capacità di accumulo anche con valori negativi dovuti anche ad un investimento iniziale notevole.

Alla luce di queste indicazioni, nel presente lavoro i valori di potenza fotovoltaica installata (kWp) e di capacità di sistema di accumulo (kWh) vengono determinati a partire dai consumi elettrici annui dell'edificio, adottando i rapporti suggeriti in letteratura come riferimento per una configurazione realistica ed energeticamente efficace. Pur mentendo un criterio di proporzionalità tra i due sistemi, i valori adottati non sono univoci, ma sono stati definiti in funzione della destinazione d'uso dell'edificio e delle caratteristiche del relativo profilo di carico.

Questi valori sono la base per le successive analisi di simulazione e per l'integrazione con il sistema di generazione termica e di controllo dell'edificio.

## 2.2. Sistemi di controllo del modello energetico

### 2.2.1. Model Predictive Control (MPC)

Il Model Predictive Control (MPC), è una strategia di controllo avanzata che utilizza un modello matematico del sistema per prevederne il comportamento futuro su un determinato orizzonte temporale. A partire dallo stato attuale e da una serie ipotizzata di ingressi di controllo, l'MPC calcola l'evoluzione futura delle variabili di interesse e determina l'azione di controllo ottimale risolvendo, a ogni istante di campionamento, un problema, o più, di ottimizzazione.

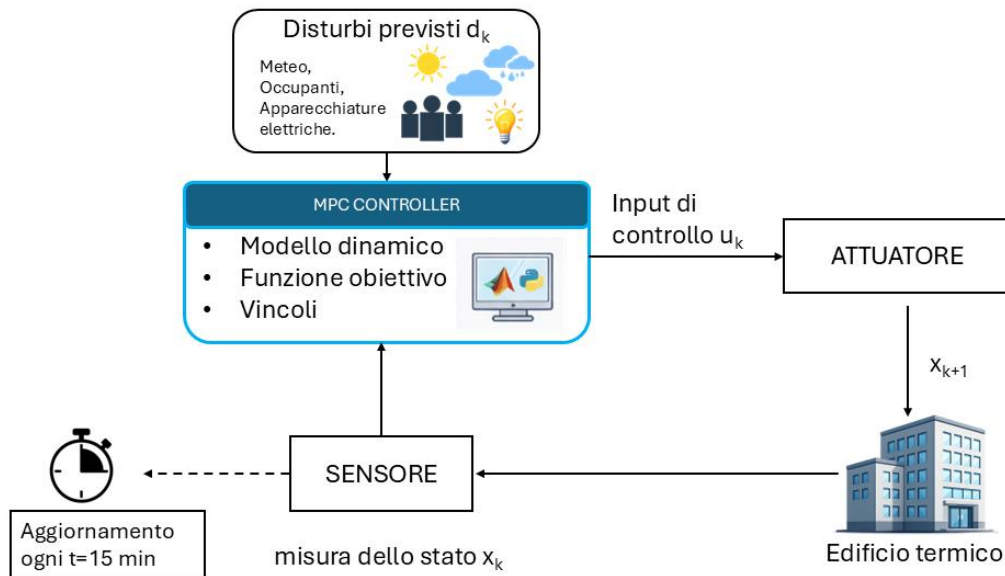


Figura 2: Descrizione funzionamento MPC

L'approccio MPC si basa su tre elementi fondamentali:

1. un modello predittivo del sistema;
2. una funzione di costo che esprime gli obiettivi di controllo, come il tracciamento di un riferimento o la minimizzazione dei consumi energetici;
3. un algoritmo di ottimizzazione in grado di determinare la sequenza ottimale delle azioni di controllo rispettando i vincoli operativi.

L'ottimizzazione viene eseguita su un orizzonte temporale finito, chiamato orizzonte di predizione. Sebbene venga calcolata una sequenza di comandi futuri, solo il primo

valore viene effettivamente applicato al sistema. Al passo di campionamento successivo, il problema di ottimizzazione viene riformulato utilizzando le nuove misure disponibili.

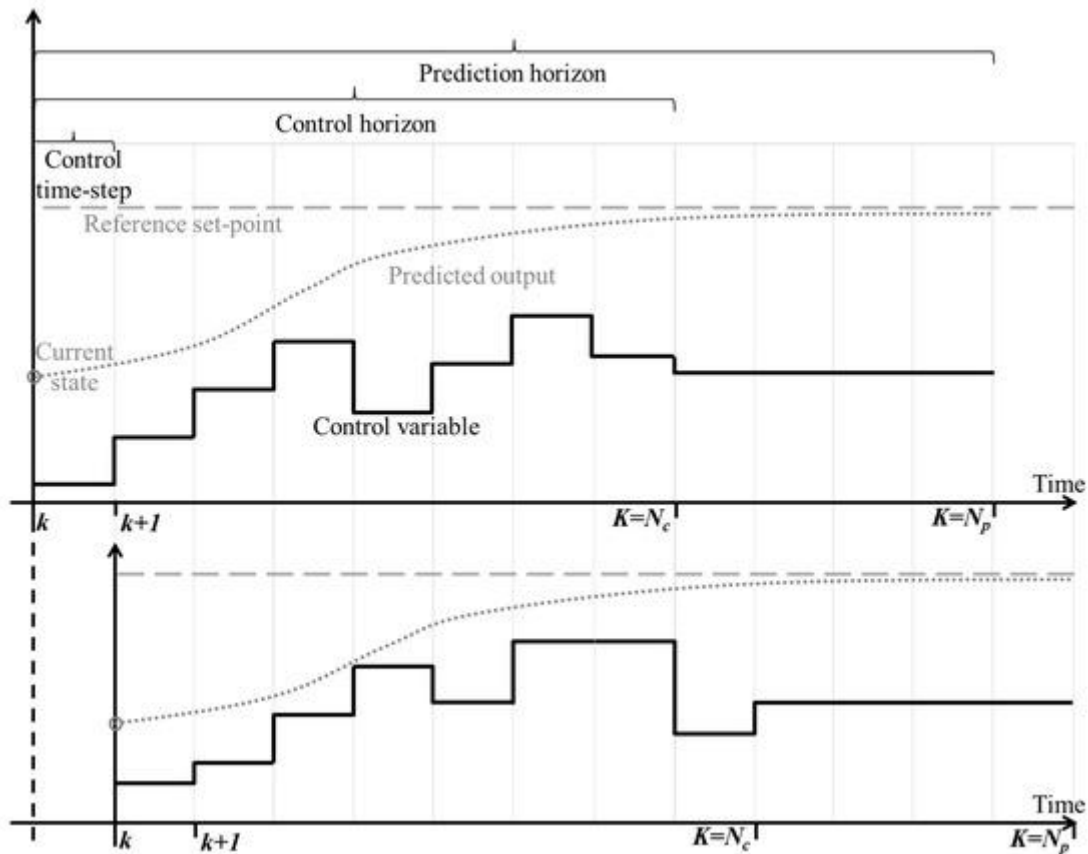


Figura 3: Orizzonte temporale MPC [16]

Questo meccanismo è noto come orizzonte mobile (receding horizon) e permette al controllore di compensare disturbi, incertezze di modello e variazioni di condizioni operative.

Una caratteristica distintiva del controllo predittivo è la possibilità di includere esplicitamente vincoli sugli ingressi di controllo, sugli stati e sulle uscite del sistema. Tali vincoli rendono l'MPC particolarmente adatto al controllo di sistemi complessi multi-ingresso e multi-uscita (MIMO), nei quali le interazioni tra le variabili sono rilevanti e le limitazioni fisiche non possono essere trascurate.

A seconda del tipo di modello utilizzato, si hanno diverse formulazioni di controllo predittivo. Nel caso di modelli lineari e funzioni di costo quadratiche, il problema di ottimizzazione assume la forma di un programma quadratico (QP), mentre per sistemi

non lineari si ricorre a formulazioni di tipo non lineare (NMPC). Per sistemi inizialmente non lineari, è comune ricorrere a tecniche di linearizzazione del modello attorno a un punto di funzionamento, così da applicare un controllo predittivo lineare con minore carico computazionale. In questo lavoro si adatta proprio questo approccio; i passaggi per la linearizzazione del modello sono descritti nel capitolo successivo. In entrambi i casi, il controllo predittivo offre buone prestazioni in termini di stabilità, robustezza e capacità di gestione delle incertezze, a fronte di un maggiore onere computazionale rispetto ai controllori tradizionali, come il PID.

Grazie a queste caratteristiche, l'MPC trova applicazione in numerosi ambiti dell'ingegneria, tra cui il controllo di processo industriale, i sistemi energetici e la gestione termica degli edifici, dove è richiesta una regolazione efficiente nel rispetto di vincoli di comfort e di consumo energetico [17][18][19].

### ***Implementazione del MPC***

In questo sottocapitolo viene descritta l'implementazione del controllo predittivo adottato nel presente lavoro. Partendo dalla formulazione matematica del problema di controllo, si introducono le equazioni di dinamica del sistema, la discretizzazione del modello e la definizione della funzione obiettivo e dei vincoli. Infine, viene illustrata l'implementazione numerica del controllore MPC.

### ***Discretizzazione del modello dinamico***

La dinamica termica dell'edificio è descritta mediante un modello a stati in forma continua, che consente di derivare il bilancio energetico dei nodi considerati. Tale modello può essere espresso come:

$$\dot{x}(t) = A_c x(t) + B_c u(t) + E_c d(t) \quad (22)$$

Dove:

- $x(t)$  rappresenta il vettore degli stati del sistema
- $u(t)$  l'ingresso di controllo
- $d(t)$  il vettore dei disturbi esterni

Considerando la nostra equazione precedente (10) discretizziamo fino ad ottenere la stessa struttura dell'equazione superiore (22).

Iniziamo dividendo tutto per il termine costante che è presente a sinistra in modo da lasciare solo la derivata nel tempo:

$$\frac{dT_i}{dt} = \frac{\Sigma U \cdot A_{dips} + H_{ve}}{VC_v} (T_e(t) - T_i(t)) + \frac{\dot{q}_{int,tot}}{C_v} + \frac{g \cdot A_v \cdot G_d(t)}{VC_v} + \frac{\alpha \cdot A_{op} \cdot Irr_{op}(t)}{VC_v} + \frac{\dot{q}_{HC}(t)}{C_v} \quad (23)$$

Disponiamo i termini dell'equazione secondo il termine che si deriva a sinistra; quindi, avremo la precedenza su  $T_i$  che sarà lo stato da controllare:

$$\frac{dT_i}{dt} = -\frac{\Sigma U \cdot A_{dips} + H_{ve}}{VC_v} T_i(t) + \frac{\dot{q}_{HC}(t)}{C_v} + \frac{\Sigma U \cdot A_{dips} + H_{ve}}{VC_v} T_e(t) + \frac{\dot{q}_{int,tot}}{C_v} + \frac{g \cdot A_v \cdot G_d(t)}{VC_v} + \frac{\alpha \cdot A_{op} \cdot Irr_{op}(t)}{VC_v} \quad (24)$$

Avendo quest'ultima equazione (24) possiamo determinare:

- Stato

Nel caso ad un nodo:

$$x(t) = T_i(t) \quad (25)$$

- Input di controllo:

Il controllo è la potenza termica di riscaldamento/raffrescamento:

$$u(t) = \dot{q}_{HC}(t) \quad (26)$$

- Disturbi:

Nel vettore dei disturbi sono incluse tutte le grandezze non direttamente controllabili, quali le condizioni climatiche esterne e i guadagni solari:

$$d(t) = \begin{bmatrix} T_e(t) \\ G_d(t) \\ Irr_{op}(t) \\ \dot{q}_{int,tot} \end{bmatrix}$$

È quindi possibile ricavare le matrici corrispondenti del modello continuo, che variano in funzione del numero di nodi considerati. Queste le andremo ad osservare nel prossimo sottocapitolo.

Poiché il controllo predittivo opera in tempo discreto, il modello continuo viene discretizzato assumendo un intervallo di campionamento  $\Delta t$  e ipotizzando che ingressi e disturbi restino costanti durante ciascun intervallo di campionamento. Sotto tali ipotesi, il modello discreto assume la forma:

$$x_{k+1} = x_k * A_d + B_d * u_k + E_d * d_k \quad (27)$$

Dove:

- $x_k$  e  $x_{k+1}$  sono le temperature allo stato presente e successivo
- $u_k$  rappresenta l'output della potenza di riscaldamento/raffrescamento
- $d_k$  sono i disturbi esterni (e.g. temperatura esterna, irradiazioni etc....)

Le matrici del modello discreto sono ottenute tramite discretizzazione esatta del sistema continuo e sono definite come:

$$A_d = e^{A_c \Delta t} \quad (28)$$

$$B_d = B_c \int_0^{\Delta t} e^{A_c \tau} d\tau \quad (29)$$

$$E_d = E_c \int_0^{\Delta t} e^{A_c \tau} d\tau \quad (30)$$

Nel caso in cui la matrice  $A_c$  sia invertibile, gli integrali possono essere calcolati sotto forma chiusa come:

$$\int_0^{\Delta t} e^{A_c \tau} d\tau = A_c^{-1}(e^{A_c \Delta t} - I) \quad (31)$$

Dove  $I$  è la matrice identità, questa è una matrice quadrata che:

- ha 1 sulla diagonale principale
- ha 0 in tutte le altre posizioni
- ha la proprietà di non modificare una matrice o un vettore quando viene moltiplicata.

Dall'equazione (31) si ottengono:

$$B_d = B_c^{-1}(B_c - I) B_c \quad (32)$$

$$E_d = A_c^{-1}(A_c - I) E_c \quad (33)$$

Questa formulazione consente di ottenere una rappresentazione discreta accurata della dinamica del sistema, preservando le caratteristiche del modello continuo.

Il modello discreto consente di ottenere il vincolo dinamico del problema di controllo predittivo ed è utilizzato all'interno del controllore MPC per la previsione dell'evoluzione futura degli stati lungo l'orizzonte di predizione. In particolare, le matrici  $A_d$ ,  $B_d$ ,  $E_d$  vengono calcolate numericamente e successivamente impiegate nella formulazione del problema di ottimizzazione descritto nelle sezioni successive.

### ***Variabili, vincoli e parametri del problema MPC***

Il problema di controllo predittivo descritto nelle sezioni precedenti è stato implementato in ambiente Python mediante la formulazione esplicita di un problema di ottimizzazione vincolata, risolto tramite il solver commerciale Gurobi, scelto per l'elevata

efficienza computazionale e l'affidabilità nella risoluzione di problemi convessi di grandi dimensioni.

L'implementazione si articola in due funzioni principali:

1. la costruzione del modello di ottimizzazione
2. la risoluzione del problema MPC ogni istante di campionamento

La prima funzione per il caso a tre nodi ha il compito di costruire il problema di controllo predittivo a partire dal modello dinamico discretizzato dell'edificio e dai parametri operativi del sistema energetico.

### ***Variabili decisionali***

Nel modello MPC vengono introdotte le seguenti variabili decisionali:

- il vettore di stato  $x_k \in \mathbb{R}^3$ , che rappresenta la temperatura dei tre nodi termici del modello dell'edificio, definito per  $k=0, H$ ;
- l'ingresso di controllo  $u_k \in \mathbb{R}$ , che rappresenta la potenza termica che deve essere fornita, definito per  $k=0, \dots, H-1$ ;
- la variabile di slack  $s_k \in \mathbb{R}^3$ , introdotta per consentire violazioni controllate dei vincoli di comfort termico;
- la potenza di scambio con la rete elettrica,  $p_{imp,k}$  e  $p_{exp,k}$ , rispettivamente potenza prelevata e immessa in rete;
- la potenza di ricarica della batteria  $p_{ch,k}$  e  $p_{dis,k}$ ;
- lo stato di ricarica della batteria  $SOC_k$ , espresso in kWh e definito per  $k=0, \dots, H$ .

La presenza di tali variabili consente di integrare, all'interno dello stesso problema di ottimizzazione, la dinamica termica dell'edificio e la gestione energetica del sistema elettrico.

### ***Vincoli di comfort termico***

Per garantire condizioni di comfort negli ambienti interni, lo stato del sistema è vincolato entro limiti inferiori e superiori:

$$x_{min} \leq x_k \leq x_{max} \quad (34)$$

dove  $x_{min}$  e  $x_{max}$  rappresentano rispettivamente i limiti inferiori e superiori di comfort termico.

Il valore  $x_{ref}$  indica la temperatura interna di riferimento desiderata.

Tali vincoli sono ammorbiditi mediante l'introduzione di variabili di  $s_k$ , che permettono violazioni limitate dei limiti di temperatura in caso di condizioni operative particolarmente gravose. Le violazioni vengono successivamente penalizzate nella funzione obiettivo.

### ***Integrazione del sistema elettrico***

Il modello MPC include un bilancio di potenza elettrica che tiene conto della pompa di calore, del carico elettrico dell'edificio, della produzione fotovoltaica e del sistema di accumulo.

Per ogni passo  $k$ , il bilancio è espresso come:

$$p_{imp,k} - p_{exp,k} + p_{pv,k} - p_{ch,k} + p_{dis,k} = p_{hp,k} + p_{load,k} \quad (35)$$

dove la potenza elettrica assorbita dalla pompa di calore è ricavata dalla potenza termica mediante il coefficiente di prestazione:

$$p_{hp,k} = \frac{u_k}{COP} \quad (36)$$

Questo vincolo consente di collegare direttamente la scelta del controllo termico con i flussi di energia elettrica.

### ***Vincoli sulla rete elettrica***

Qualora specificato, il modello consente di imporre limiti sulla potenza massima scambiabile con la rete elettrica, sia in immissione che in prelievo.

Tali vincoli permettono di rappresentare realisticamente le limitazioni dell'allacciamento elettrico dell'edificio.

### ***Dinamica dello stato di ricarica della batteria***

La dinamica dello stato di carica della batteria è modellata tramite l'equazione discreta:

$$SOC_{k+1} = SOC_k + \eta_{ch}\alpha p_{ch,k} - \frac{1}{\eta_{dis}}\alpha p_{dis,k} \quad (37)$$

dove  $\alpha = \frac{\Delta t}{3.6 \cdot 10^6}$  è il fattore di conversione da potenza (W) a energia (kWh).

Lo stato di carica è vincolato tra zero e la capacità massima della batteria.

### ***Funzione obiettivo***

La funzione obiettivo del problema MPC combina criteri economici e di comfort termico.

In particolare, essa include:

- il costo di acquisto dell'energia elettrica dalla rete;
- il ricavo derivante dalla vendita di energia alla rete;
- una penalizzazione sulle violazioni dei vincoli di comfort, proporzionali alle variabili di slack.

Il problema risultante è formulato come un problema di ottimizzazione convessa, in quanto è caratterizzato da una funzione obiettivo lineare e vincoli lineari. Tale formulazione garantisce l'unicità della soluzione ottima e un'elevata affidabilità numerica; il problema viene quindi risolto minimizzando il costo totale sull'orizzonte di predizione.

### ***Risoluzione del problema MPC***

La seconda funzione si occupa di risolvere il problema MPC a ogni passo di campionamento. Una volta costruito il modello di ottimizzazione, il solver calcola la traiettoria ottimale delle variabili; tuttavia, secondo il principio dell'orizzonte mobile, viene applicata al sistema sola la prima azione di controllo  $u_0$ .

In conclusione, l'approccio adottato consente di integrare in un'unica formulazione la dinamica termica dell'edificio, la gestione dei vincoli di comfort e l'ottimizzazione dei flussi energetici elettrici, rendendo il controllo predittivo particolarmente adatto alla gestione energetica avanzata degli edifici.

### 2.2.2. *Deep Reinforcement Learning*

Il Reinforcement Learning (RL), o apprendimento per rinforzo, è un paradigma di apprendimento automatico e di controllo decisionale in cui un agente impara a interagire in un ambiente dinamico attraverso un processo di trial and error (prova ed errore) [20]. L'obiettivo dell'agente è individuare una strategia di controllo capace di massimizzare nel tempo una misura di prestazione che viene definita tramite un segnale numerico chiamato reward (ricompensa).

A differenza dei metodi di controllo tradizionali, basati su modelli matematici espliciti, il RL non richiede necessariamente una descrizione analitica completa della dinamica del sistema. L'agente apprende le conseguenze delle proprie azioni e aggiornando progressivamente le decisioni sulla base dell'esperienza accumulata. Questo approccio risulta molto adatto a sistemi complessi, non lineari e soggetti a incertezze, nei quali la modellazione accurata può risultare difficile [20].

Il funzionamento del RL si basa su uno schema di interazione ciclica tra agente e ambiente. A ogni istante discreto di tempo, l'agente osserva lo stato corrente del sistema, seleziona un'azione di controllo e la applica all'ambiente. L'ambiente evolve secondo la propria dinamica interna e restituisce all'agente un nuovo stato e un valore di ricompensa che quantifica l'efficienza dell'azione intrapresa. La sequenza di esperienze così generata costituisce la base del processo di apprendimento [20]

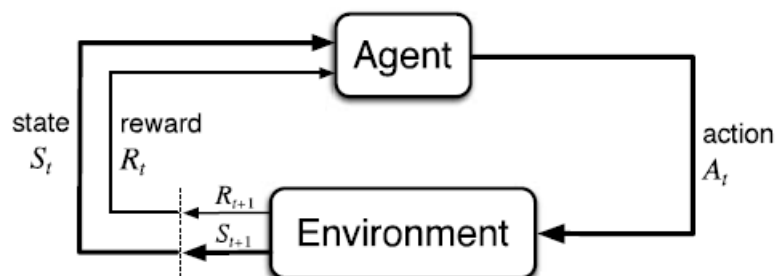


Figura 4: L'interazione agente-ambiente in un RL [20].

L'approccio RL, quindi, si basa su tre elementi chiave: un agente decisionale, responsabile della scelta delle azioni di controllo; un ambiente dinamico, che rappresenta un sistema fisico o simulato con cui l'agente interagisce; una funzione di ricompensa, che fornisce una misura della qualità delle azioni intraprese.

Per descrivere questo processo di interazione e apprendimento, il RL viene modellato attraverso il Markov Decision Process (MDP), che rappresenta il quadro matematico di riferimento per i problemi decisionali sequenziali in ambienti dinamici e imprevedibili.

### **2.2.2.1. Implementazione del Reinforcement Learning (RL)**

Il RL viene formulato mediante il concetto di Markov Decision Process (MDP) [20], è definito dai valori:

$$(S, A, P, R, \gamma)$$

Dove:

- $S$  rappresenta lo spazio degli stati del sistema;
- $A$  è lo spazio delle azioni disponibili all'agente;
- $P(S_{t+1} | S_t, A_t)$  descrive la dinamica di transizione dell'ambiente;
- $R(S_t, A_t)$  è la funzione di ricompensa;
- $\gamma \in [0,1]$  è il fattore di sconto che pesa l'importanza delle ricompense future.

L'integrazione agente-ambiente può essere quindi essere descritta come una sequenza temporale:

$$S_t \rightarrow A_t \rightarrow R_{t+1}, S_{t+1}$$

in cui lo stato futuro dipende esclusivamente dallo stato corrente e dall'azione applicata.

#### ***Stato***

Lo stato  $S_t$  rappresenta l'insieme delle informazioni disponibili all'agente per prendere decisioni[20]. Deve contenere tutte le variabili necessarie affinché, una volta noto lo stato corrente, l'evoluzione futura del sistema non dipenda dagli stati passati.

In questo lavoro, lo stato è rappresentato da un vettore di osservazioni costituito da nove variabili che descrivono le condizioni energetiche dell'edificio, i disturbi termo-

ambientali, il contesto economico e informazioni temporali. In particolare, lo stato include:

- la domanda energetica dell'edificio;
- la produzione fotovoltaica;
- la temperatura esterna;
- i contributi solari;
- il prezzo di acquisto dell'energia elettrica per fasce orarie;
- l'ora del giorno;
- il giorno della settimana.

### ***Azione***

L'azione  $A_t$  rappresenta la decisione di controllo applicata dall'agente al sistema. Nel caso considerato, lo spazio delle azioni è continuo e bidimensionale. Le azioni includono la potenza termica fornita dalla pompa di calore che regola il riscaldamento dell'edificio e la potenza di carica e scarica del sistema di accumulo.

### ***Dinamica di transizione***

Nel RL, la funzione di transizione:

$$P(S_{t+1} | S_t, A_t)$$

descrive l'evoluzione dell'ambiente a seguito dell'azione applicata [20]. Questa dinamica non è necessariamente nota in forma esplicita all'agente, ma viene appresa implicitamente attraverso l'interazione con il sistema.

Nel sistema edificio-impianto considerato, la dinamica è il risultato dell'interazione tra fenomeni termici ed energetici. La potenza termica fornita dalla pompa di calore agisce come ingresso di controllo nel modello dinamico dell'edificio, influenzando l'evoluzione della temperatura interna. Parallelamente, la gestione della batteria determina la variazione dello stato di carica del sistema di accumulo. La decisione dell'agente si riflettono sul bilancio elettrico, che collega domanda edificio, produzione fotovoltaica, pompa di calore e sistema di accumulo. A completare il quadro intervengono le condizioni esterne, quali temperatura esterna, irraggiamento solare e carichi interni, che introducono variabilità e incertezza nel comportamento del sistema. In questo modo, le azioni dell'agente influenzando sia la dinamica termica che quella energetica dell'edificio.

### ***Reward (ricompensa)***

La reward (ricompensa)  $R_t$  rappresenta una misura numerica della qualità dell'azione intrapresa al tempo  $t$  [1] e traduce gli obiettivi di controllo in un segnale scalare utilizzato per guidare l'apprendimento dell'agente.

Nel nostro modello, la reward è definita in modo da orientare l'agente verso una gestione energetica che bilanci comfort termico ed efficienza economica. La ricompensa è formulata come una combinazione pesata di tre contributi: errore di comfort termico, utilizzo della batteria e costo energetico, costruita in questo modo:

$$R_t = (k_{HP}|u_{T,t}| + k_{battery}|u_{battery,t}| + C_t) \cdot k_R \quad (38)$$

dove  $|u_{T,t}|$  rappresenta l'errore di temperatura interna rispetto alla banda di comfort,  $|u_{battery,t}|$  è l'azione relativa al sistema di accumulo e  $C_t$  è il costo energetico istantaneo, mentre  $k_{HP}$ ,  $k_{battery}$ ,  $k_R$  sono coefficienti di pesatura che regolano l'importanza relativa dei diversi contributi, rispettivamente HP\_penalty, BatteryPenalty e RewardMultiplier nello script.

Questa formulazione del reward (38) induce l'agente a ridurre il costo dell'energia acquistata, mantenere la temperatura interna entro i limiti di comfort e limitare l'impiego energetico della pompa di calore e del sistema di accumulo.

### ***Obiettivo dell'agente***

L'agente non massimizza la ricompensa immediata, ma il ritorno cumulativo scontato

$$G_t = R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} + \dots \quad (39)$$

dove  $\gamma \in [0,1]$  è il fattore di sconto [20]. Questo parametro regola l'importanza relativa attribuita alle ricompense future rispetto a quelle immediate: valori di  $\gamma$  prossimi a 1 rendono l'agente maggiormente orientato al lungo termine, mentre valori più bassi privilegiano benefici immediati.

### ***Funzione di valore (q-value) e politica di controllo***

Per valutare quanto siano efficaci le decisioni dell'agente nel tempo, il RL introduce il concetto di funzione di valore [20]. In particolare, la funzione valore-azione, nota come Q-value, è definita come il ritorno cumulativo atteso quando l'agente si trova in uno stato  $S_t$ , esegue l'azione  $A_t$  e successivamente continua a operare seguendo una politica  $\pi$  [20]:

$$Q^\pi(S_t, A_t) = E[G_t | S_t, A_t] \quad (40)$$

Il Q-value fornisce quindi una misura di quanto sia conveniente compiere una certa azione in uno specifico stato, considerando non solo l'effetto immediato, ma anche le conseguenze future [20]. Quindi, permette di confrontare direttamente azioni alternative sulla base del loro impatto a lungo termine.

Il comportamento dell'agente, invece, è descritto dalla politica di controllo, indicata con  $\pi(s|a)$ , che rappresenta la regola con cui vengono selezionate le decisioni dell'agente [20]. Nel nostro caso, le azioni non sono scelte tra poche alternative, ma possono essere valori continui (ad esempio una certa potenza della pompa di calore). Per questo motivo, la politica associa direttamente a ogni stato un preciso valore dell'azione di controllo.

La politica rappresenta l'elemento operativo del controllore RL, poiché effettivamente è quella utilizzata durante il funzionamento del sistema. Il processo di apprendimento consiste proprio nell'aggiornare i parametri che definiscono questa politica, in modo che le azioni selezionate conducano a un miglioramento delle prestazioni complessive [20].

L'obiettivo dell'apprendimento è determinare, quindi, una politica che massimizzi il rendimento atteso nel lungo periodo. Questo concetto è espresso dalla funzione valore-azione ottimale [20]:

$$\pi^*(s) = \arg \max Q^*(s, a) \quad (41)$$

ossia, in ogni stato, l'agente seleziona l'azione che garantisce il rendimento atteso più elevato. Il problema del Reinforcement Learning può essere visto come la ricerca di una funzione di valore che consente di individuare la politica capace di massimizzare il ritorno cumulativo.

### ***Deep Reinforcement Learning***

Nella gestione energetica edificio-impianto, sia lo spazio degli stati sia quello delle azioni sono continui e caratterizzati da più variabili fisiche, energetiche ed economiche. In queste condizioni, i metodi di RL tradizionali basati su rappresentazioni tabellari non sono applicabili, poiché gli stati e le azioni possono assumere infiniti valori e non si può costruire una tabella che contenga tutte le combinazioni possibili. Per questo motivo è necessario utilizzare modelli che sappiano approssimare queste funzioni, come le reti neurali.

Adottiamo, quindi, un approccio di Deep Reinforcement Learning (DRL), nel quale le funzioni fondamentali del RL, la politica di controllo e la funzione di valore, vengono approssimate mediante reti neurali profonde. In questo contesto, la politica diventa una funzione parametrica della forma:

$$A_t = \pi(S_t; \theta) \quad (42)$$

dove  $\theta$  rappresenta l'insieme dei pesi della rete neurale. La rete realizza quindi una mappatura diretta tra lo stato del sistema e l'azione di controllo, consentendo di gestire relazioni non lineari complesse che caratterizzano la dinamica termo-energetica dell'edificio.

Nel caso studio considerato, viene adottata una struttura di tipo Actor-Critic. Una prima rete neurale, detta critic, stima il valore-azione  $Q(S_t, A_t)$ , che misura quanto sia vantaggiosa un'azione in un determinato stato considerando gli effetti futuri. Una seconda rete, detta actor, apprende invece la politica di controllo, ossia la regola che associa ad ogni stato l'azione da applicare. L'actor aggiorna progressivamente i propri parametri in modo da selezionare azioni che massimizzino il valore stimato di critic. Questo schema è alla base dell'algoritmo Deep Deterministic Policy Gradient (DDPG), utilizzato per l'addestramento dell'agente [139].

### ***Implementazione dell'agente DRL nel caso studio***

L'approccio di Deep Reinforcement Learning, descritto in precedenza, è stato implementato utilizzando l'ambiente MATLAB/Simulink, impiegando il Reinforcement

Learning Toolbox. Il modello dinamico edificio-impianto è rappresentato in Simulink, che svolge il ruolo di ambiente di interazione, mentre l'agente di controllo è definito in MATLAB.

L'interazione agente-ambiente avviene con passo temporale di 15 minuti. A ogni passo, l'ambiente fornisce all'agente un vettore di osservazioni composto da nove variabili, che descrivono lo stato energetico, termico ed economico del sistema. Queste osservazioni costituiscono l'ingresso della rete neurale dell'agente.

Lo spazio delle azioni è continuo e bidimensionale. Le due componenti dell'azione sono rispettivamente la potenza termica fornita dalla pompa di calore e la potenza di carica/scarica del sistema di accumulo. Le azioni sono normalizzate nell'intervallo  $[-1,1]$  e successivamente scalate nel modello fisico per ottenere i valori effettivi.

Come detto in precedenza, l'agente viene addestrato mediante l'algoritmo Deep Deterministic Policy Gradient (DDPG). L'addestramento avviene tramite experience replay, che consente di memorizzare le esperienze passate e riutilizzarle per l'aggiornamento delle reti neurali, migliorando stabilità e convergenza. Per favorire l'esplorazione dello spazio delle azioni durante il training, viene aggiunto un rumore gaussiano all'azione selezionata dall'agente.

Il fattore di sconto è 0,99, un valore prossimo a 1, rendendo l'agente orientato al lungo termine, importante nei sistemi energetici per l'evoluzione futura delle variabili di stato.

Al termine dell'addestramento, la rete neurale che rappresenta la politica di controllo viene integrata nel modello Simulink e utilizzata direttamente per calcolare le azioni dell'agente. Ad ogni istante, il controllore riceve lo stato del sistema in ingresso e restituisce immediatamente i segnali di controllo per la pompa di calore e per la batteria.

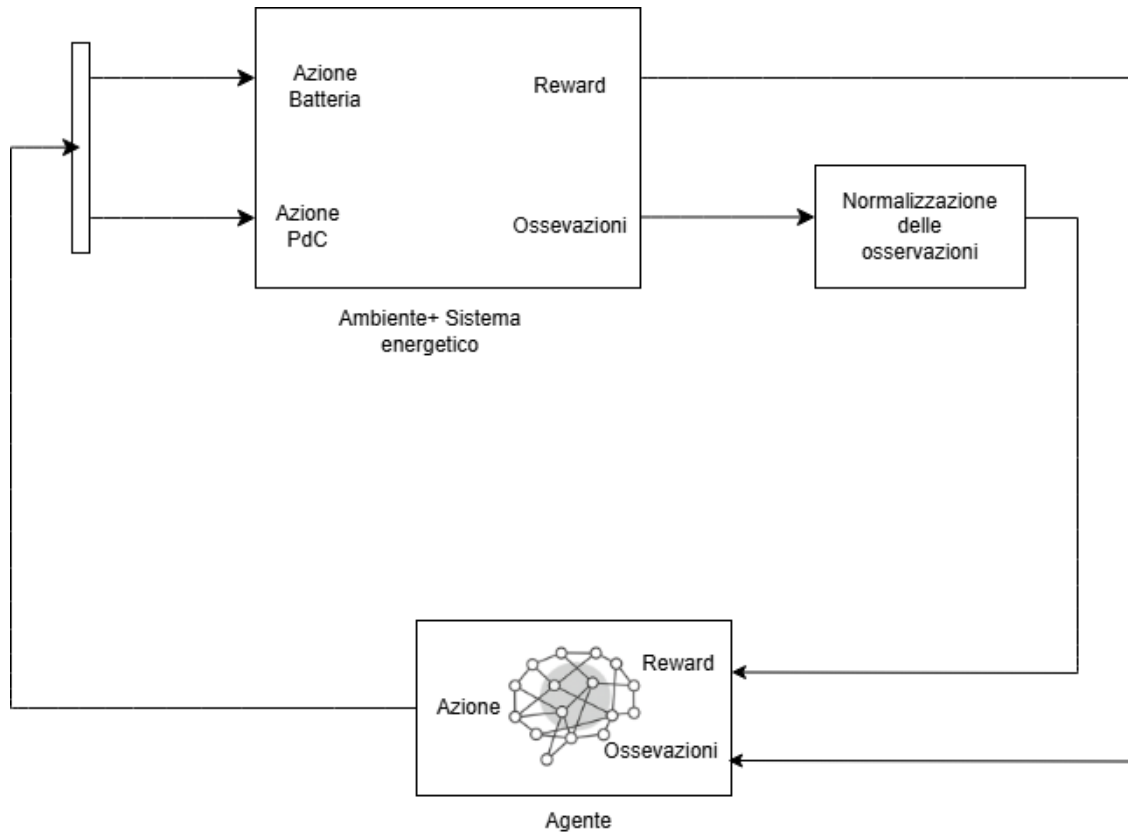


Figura 5: Integrazione ambiente-agente nel nostro modello in Simulink

Dopo la fase di addestramento non viene utilizzato automaticamente l'ultimo agente ottenuto, ma viene effettuata una selezione basata sulle prestazioni effettive delle diverse politiche apprese. Durante il training, MATLAB salva periodicamente diversi agenti intermedi. Per ognuno di essi viene calcolata una evaluation statistic, che rappresenta la ricompensa media ottenuta in episodi di prova eseguiti.

Questa valutazione permette di misurare quanto l'agente sia realmente efficace nel controllo del sistema, in condizioni operative stabili e ripetibili. Gli agenti con ricompensa media più alta sono considerati più performanti, poiché riescono a ottimizzare meglio l'obiettivo del controllo.

Infine, determinato l'agente tra questi con comportamento più stabile e coerente, viene integrato nel modello Simulink e utilizzato come controllore definitivo per le simulazioni del caso studio in periodi differenti da quello di training.

### ***2.2.3. Transfer Learning***

Nel Reinforcement Learning tradizionale, un agente costruisce progressivamente la propria politica a partire da un'inizializzazione priva di conoscenza a priori, basandosi esclusivamente sull'esperienza raccolta attraverso l'interazione con l'ambiente. Questo approccio non richiede informazioni preliminari sul problema, ma comporta tempi di addestramento elevati. L'agente deve infatti esplorare a lungo prima di individuare una strategia efficace. Nelle prime fasi di apprendimento il comportamento risulta inevitabilmente subottimale, con un accumulo limitato di ricompensa.

Nel contesto del controllo energetico degli edifici mediante DRL, questo aspetto assume particolare rilevanza. Addestrare un agente per ogni nuova tipologia edilizia può richiedere un numero significativo di interazioni con l'ambiente simulato, con un conseguente costo computazionale. Diventa quindi naturale interrogarsi sulla possibilità di riutilizzare l'esperienza maturata su un edificio per facilitare l'apprendimento su un altro edificio con caratteristiche affini.

Il Transfer Learning risponde a questa esigenza. L'idea è che, se due problemi condividono una struttura simile, la conoscenza acquisita nel primo possa risultare utile anche nel secondo. Nell'ambito del DRL, ciò significa sfruttare l'esperienza appresa in un ambiente sorgente per migliorare l'apprendimento in un ambiente target correlato. Una trattazione sistematica del Transfer Learning nel RL è proposta da Matthew E. Taylor e Peter Stone nel loro lavoro[22], dove vengono analizzate le modalità con cui la conoscenza può essere trasferita tra omini differenti.

Come detto in [22], il trasferimento può produrre benefici misurabili lungo la curva di apprendimento. L'agente può iniziare il training nel dominio target con una performance superiore rispetto a un'inizializzazione casuale, riducendo la fase di esplorazione inefficiente. Inoltre, può diminuire il numero di episodi necessari per raggiungere una determinata soglia di prestazioni, migliorando la velocità di convergenza. Questo si traduce in un incremento del reward cumulativo durante il training e, in alcuni casi, in un miglioramento della performance finale.

Nel DRL, la conoscenza trasferita coincide tipicamente con i parametri della rete neurale che rappresenta la policy. Durante l'addestramento nel dominio sorgente, la rete apprende una rappresentazione interna delle relazioni tra stati, azioni e reward. Se questa

rappresentazione cattura aspetti strutturali di un edificio, ad esempio le dinamiche generali del comportamento termico di un edificio, essa può risultare utile anche in ambienti differenti ma concettualmente simili.

Tuttavia, il trasferimento non è sempre vantaggioso. Come evidenziato in [22] può verificarsi il fenomeno del *negative transfer*, quando la conoscenza riutilizzata peggiora le prestazioni rispetto all'apprendimento da zero. Questa accade quando la differenza tra dominio sorgente e dominio target è troppo marcata. Se le dinamiche dell'ambiente cambiano in modo significativo o se la distribuzione degli stati osservati è molto diversa, la rete può inizialmente applicare strategia non più adeguate al nuovo contesto, rallentando l'adattamento.

Nel seguente lavoro, il Transfer Learning viene applicato tra edifici appartenenti a tipologie differenti. L'agente viene inizialmente addestrato su un edificio residenziale, che costituisce il dominio sorgente. Successivamente, la conoscenza appresa viene riutilizzata in due domini target: un edificio scolastico e un edificio industriale. Sebbene la struttura del problema di controllo di riscaldamento rimanga invariato, le condizioni operative cambiano in modo significativo. Il passaggio dal residenziale alla scuola introduce differenze nella distribuzione temporale dei carichi interni, mentre il carico industriale presenta dinamiche ancora più distanti, con carichi termici elevati e pattern operativi specifici.

All'interno di questo studio vengono considerate due modalità di trasferimento. Nel caso dello *zero-shot transfer*, la policy appresa nel dominio residenziale viene applicata direttamente ai domini target senza ulteriore addestramento, al fine di valutare la capacità di generalizzazione pura della rete neurale. Nel caso del *fine-tuning*, invece, i parametri appresi nel dominio sorgente vengono utilizzati come inizializzazione per un nuovo processo di addestramento del dominio target, consentendo alla rete di adattarsi progressivamente alle nuove dinamiche mantenendo la struttura già acquisita.

Il confronto tra addestramento *zero-transfer* e *fine-tuning* permette quindi di analizzare in modo sistematico sia i benefici sia i limiti del transfer learning nel controllo energetico di edifici eterogenei, fornendo indicazioni sulla scalabilità del DRL in contesti reali caratterizzati da elevata variabilità tra domini.

#### ***2.2.4. Controllo tradizionale PID + Rule-Based***

Tra le strategie di controllo considerate è stato incluso anche un approccio di riferimento basato sulla combinazione di un controllore PID e di logiche rule-based (PID + RB), utilizzato come termine di confronto rispetto ai metodi avanzati.

Il controllore PID (Proportional-Integral-Derivative) è una tecnica di controllo classica ampiamente utilizzata in ambito industriale per la regolazione di grandezze fisiche. Il principio di funzionamento si basa sulla correzione dell'errore tra il valore desiderato (set-point) e la variabile controllata attraverso tre contributi: il termine proporzionale, che reagisce all'errore istantaneo, il termine integrale, che tiene conto dell'errore accumulato nel tempo, e il termine derivativo, che anticipa l'andamento dell'errore migliorando la stabilità del sistema.

Nel caso in esame, il controllore PID viene utilizzato per regolare la temperatura interna dell'edificio, modulando la potenza della pompa di calore in funzione dello scostamento rispetto al set-point di comfort.

A questo controllo si affiancano logiche Rule-Based per la gestione del sistema energetico, in particolare per il controllo della batteria e degli scambi con la rete. Tali logiche si basano su regole predefinite, ad esempio privilegiando l'autoconsumo dell'energia fotovoltaica, la ricarica della batteria in presenza di surplus e l'utilizzo dell'energia accumulata in caso di deficit.

Questo approccio rappresenta una soluzione semplice e robusta, ma non ottimizzata, che consente di evidenziare i vantaggi delle strategie di controllo avanzate analizzate nel lavoro.

### 3. *Casi studio e analisi dei risultati*

Il presente capitolo riporta i risultati delle simulazioni condotte sui casi studio introdotti nel capitolo precedente, con l'obiettivo di confrontare le prestazioni delle varie strategie di controllo.

Le simulazioni sono state condotte con passo temporale pari a 15 minuti. Per ciascun edificio è stato considerato lo stesso scenario climatico e lo stesso profilo tariffario dell'energia elettrica, in modo da garantire un confronto omogeneo tra le diverse strategie di controllo.

L'analisi dei risultati è condotta esaminando in modo integrato gli effetti delle diverse strategie di controllo sia dal punto di vista energetico sia economico, considerando il mantenimento delle condizioni di comfort, i consumi elettrici e il comportamento complessivo dell'impianto, inclusa la gestione dell'impianto di fotovoltaico e del sistema di accumulo.

Le simulazioni sono state condotte su un periodo invernale, compreso tra il 1° (mercoledì) e il 7 febbraio (martedì), caratterizzato da temperatura esterna variabile. La temperatura di set-point è fissata a 20°C per i casi studio residenziali e scolastico, con un range di comfort compreso fra i 20°C e i 22°C. Per il caso industriale, invece, è stata adottata una banda di comfort compresa tra 18°C e 20°C, coerente con le condizioni tipicamente adottate negli ambienti produttivi, dove i requisiti di comfort risultano generalmente meno restrittivi.

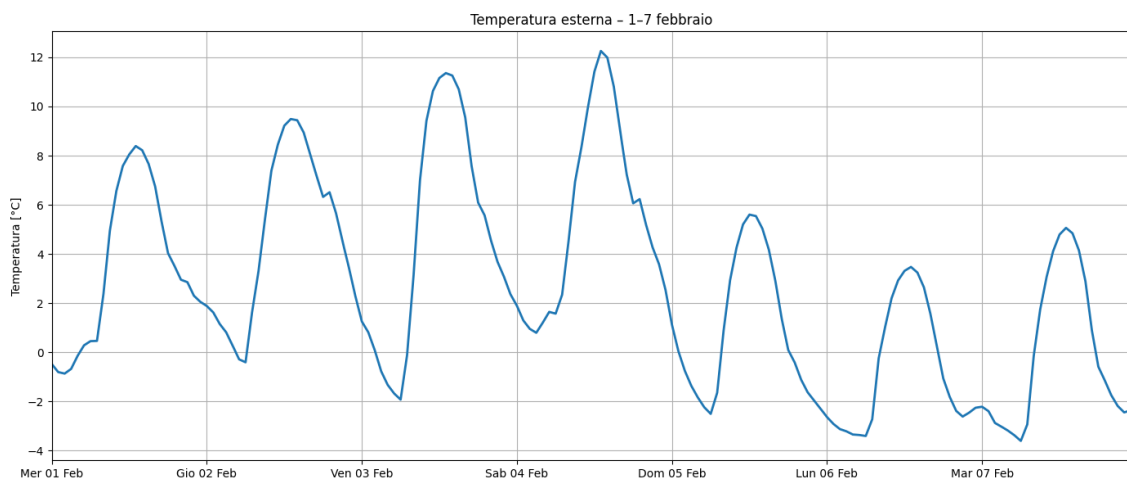


Figura 6: *Temperatura esterna del periodo considerato*

Come si vede dalla Figura 6), la temperatura esterna in quel periodo presenta un andamento fortemente variabile, con escursioni termiche giornaliere significative. Nei primi giorni si registrano i valori massimi superiori ai 10°C, mentre nella parte finale si osserva un progressivo abbassamento dei valori medi, con minimi che raggiungono anche i -3°C. Questa evoluzione evidenzia condizioni climatiche non stazionarie quindi una buona base per l'analisi delle strategie di controllo, poiché il sistema si deve adattare a richieste termiche differenti nel corso della settimana.

Nei paragrafi successivi vengono analizzati i tre casi studio considerati presentando per ciascuno i risultati ottenuti con le diverse strategie di controllo.

### **3.1. Caso studio residenziale**

Per il nostro caso studio è stato preso inizialmente come riferimento un edificio semplice, a scopo residenziale. L'edificio ha una planimetria di lunghezza 15 metri, larghezza 8 metri e altezza 3 metri localizzato nel Nord Italia.

Dal punto di vista costruttivo, l'involucro edilizio è caratterizzato da un'unica stratigrafia opaca, composta da un singolo strato di laterizi, adottata sia per le superfici verticali opache esterne sia per le superfici orizzontali opache, come solaio e pavimento. Gli elementi trasparenti sono costituiti da un serramento con doppio vetro basso emissivo e telaio in alluminio a taglio termico. Le proprietà termo-fisiche sono state determinate tramite configuratore del produttore e secondo UNI EN ISO 10077-1. La trasmittanza risultante dell'infisso è pari a 1,7 W/(m²K).

Avendo a disposizione tutte le stratigrafie otteniamo:

$$U_{opaco} = 0,49 \frac{W}{m^2K} \quad (43)$$

$$U_{win} = 1,7 \frac{W}{m^2K} \quad (44)$$

La ventilazione è stata determinata in conformità alla norma UNI EN 16798, considerando l'edificio ad uso residenziale. Il calcolo è stato effettuato stimando

preliminarmente il numero ideale di occupanti in funzione della superficie calpestabile, ricavando così il parametro  $n_p$  (persona/m<sup>2</sup>). Sulla base nella stessa norma, sono stati assunti i valori di riferimento della portata d'aria per persona  $q_p$  (l/perosna/s) e la portata per area calpestabile  $q_b$  (l/s/m<sup>2</sup>).

Pertanto, il flusso di ventilazione risulta:

$$Q_{ve} = q_p \cdot n_p \cdot A_p + q_b \cdot A_p \left[ \frac{l}{s} \right] \quad (45)$$

Il valore di  $q_p$  è stato prelevato dalla Categoria II della Tabella A.11 della norma (uso: Homes apartments), mentre  $q_b$  dallo stesso documento con riferimento al tipo di spazio unità residenziale senza bagno.

Otteniamo il coefficiente globale di ventilazione:

$$H_{ve} = \frac{Q_{ve}}{1000} \cdot \rho \cdot c_p = 76,8 \frac{W}{K} \quad (46)$$

Sono stati considerati i guadagni interni, determinati anch'essi secondo UNI EN 16798. Il contributo degli occupanti è stato valutato come somma delle componenti sensibili convettive, radiative e latenti, mentre il contributo delle apparecchiature elettriche è stato espresso in funzione dell'area calpestabile. Il guadagno interno totale è quindi dato dalla somma dei contributi dovuti agli occupanti e alle apparecchiature.

Sulla base delle caratteristiche dell'involucro, delle perdite per ventilazione edei guadagni interni, è stato calcolato il carico termico invernale di progetto. Considerando anche il rendimento complessivo dell'impianto, la potenza termica necessaria per il dimensionamento della pompa di calore risulta pari a:

$$\tilde{\Phi}_{P,finale}^* = 9,5 \text{ kW} \quad (47)$$

Questo valore rappresenta la potenza termica necessaria per il dimensionamento della pompa di calore da utilizzare del nostro edificio esaminato.

Paramenti	Edificio residenziale
Volume [m <sup>3</sup> ]	360
Area calpestabile [m <sup>2</sup> ]	120
Dimensione PdC [kW]	9,5
Dimensione PV [kW]	10
Dimensione accumulo [kWh]	18

*Tabella 1: Parametri del modello energetico - caso residenziale*

### **3.1.1. Controllo con MPC**

Il controllo predittivo, attraverso l'utilizzo dell'MPC, è stato implementato utilizzando il modello dinamico a tre nodi descritto nel capitolo precedente, discretizzato con passo temporale di 15 minuti. L'orizzonte predittivo è stato fissato pari a 24 ore, consentendo al controllore di disporre delle previsioni climatiche e del profilo tariffario dell'intera giornata successiva. Questa scelta permette di sfruttare al meglio la natura predittiva dell'MPC, ottimizzando la gestione della pompa di calore e del sistema di accumulo sulla base dell'evoluzione delle condizioni esterne e dei costi dell'energia.

La funzione obiettivo minimizza il costo netto dell'energia elettrica scambiata con la rete, tenendo conto sia del prezzo di acquisto sia del prezzo di vendita, garantendo allo stesso tempo il rispetto dei vincoli di comfort termico imposti sulla temperatura interna.

Il profilo tariffario adottato distingue tre fasce orarie (F1, F2, F3), caratterizzate rispettivamente da prezzi a 0,25 €/kWh, 0,22 €/kWh, 0,19€/kWh. Il prezzo di vendita dell'energia immessa in rete è stato assunto pari a 0,05 €/kWh, significativamente inferiore rispetto ai prezzi di acquisto. Questa differenza incentiva strategie orientate all'autoconsumo e all'accumulo piuttosto che alla vendita diretta dell'energia prodotta.

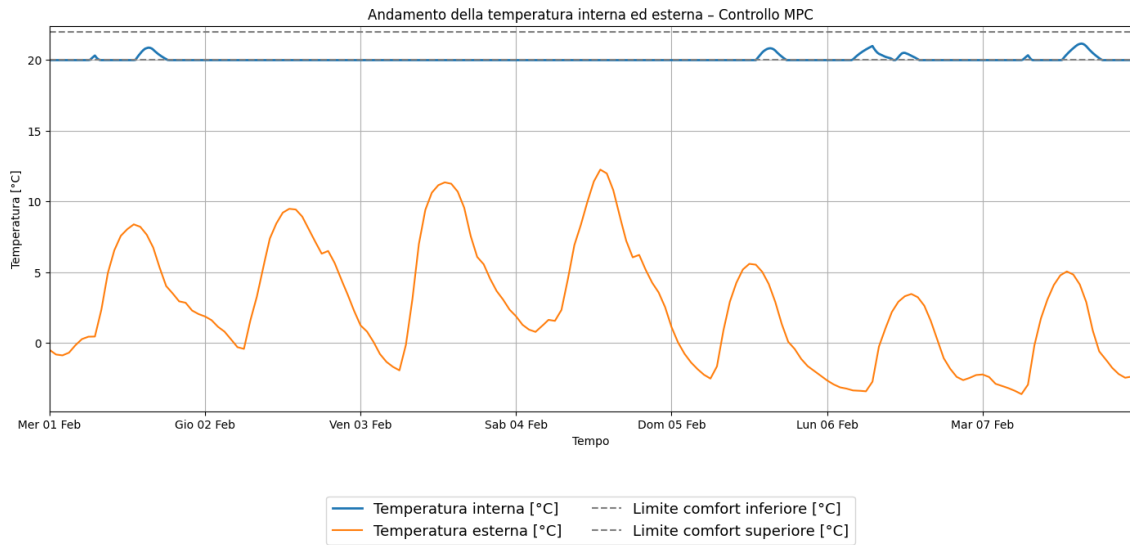


Figura 7: Andamento temperatura interna ed esterna con controllore MPC - caso studio residenziale

Osservando la Figura 7) l'MPC riesce a mantenere la temperatura interna all'interno della banda di comfort per l'intero periodo di simulazione, nonostante la temperatura esterna varia molto durante il periodo con un progressivo abbassamento degli ultimi giorni. Possiamo notare come la temperatura interna si mantiene all'interno della banda di comfort, questo conferma l'efficacia della strategia di controllo adottata. Si nota, inoltre, un leggero incremento della temperatura nelle ore centrali di alcune giornate più fredde, interpretabili come un preriscaldamento controllato, finalizzato a sfruttare la disponibilità di energia elettrica prodotta localmente dall'impianto fotovoltaico.

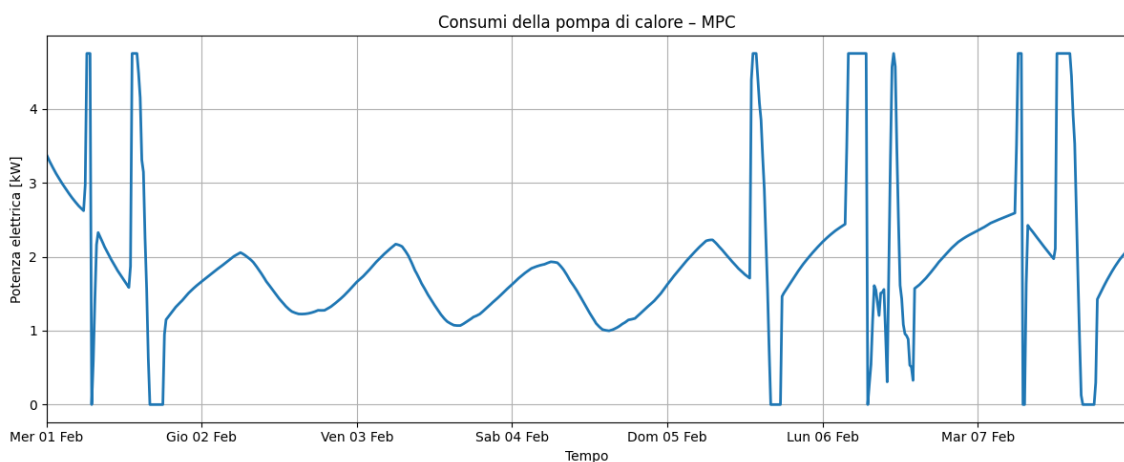


Figura 8: Andamento potenza elettrica consumata dalla pompa di calore con controllo MPC - caso studio residenziale

La Figura 8) mostra l'andamento della potenza termica erogata dalla pompa di calore. Si osserva una modulazione regolare dell'azione di controllo, con valori compresi tra 1 e 2,5 kW e picchi più elevati nei giorni caratterizzati da temperature esterne più rigide.

Nei giorni 5 e 6 febbraio, in corrispondenza delle temperature esterne minime, si registrano incrementi della potenza erogata, a conferma della correlazione tra carico termico e condizioni climatiche. Si osserva inoltre l'assenza di oscillazioni prolungate, ad eccezione delle giornate caratterizzate da temperature esterne più rigide. Questo indica un comportamento stabile del controllore, che sfrutta l'inerzia dell'edificio per distribuire l'energia nel tempo.

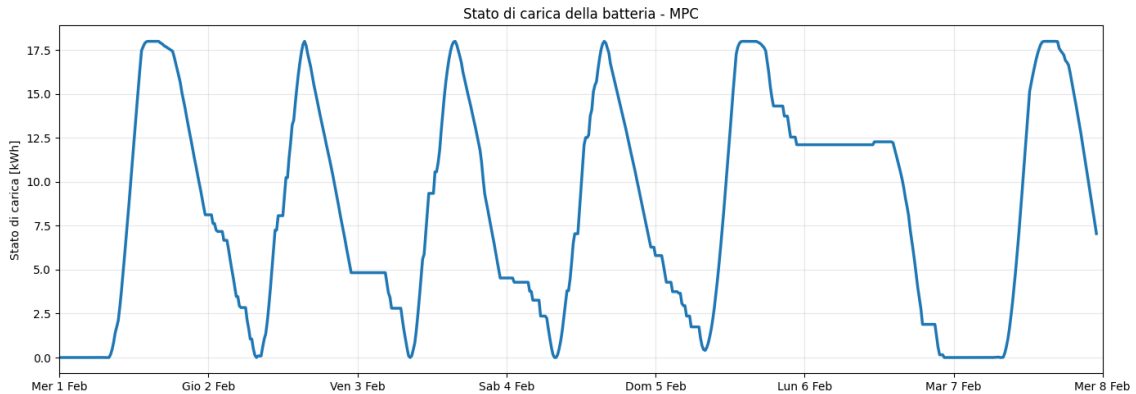


Figura 9: Andamento stato di carica della batteria con controllo MPC - caso studio residenziale

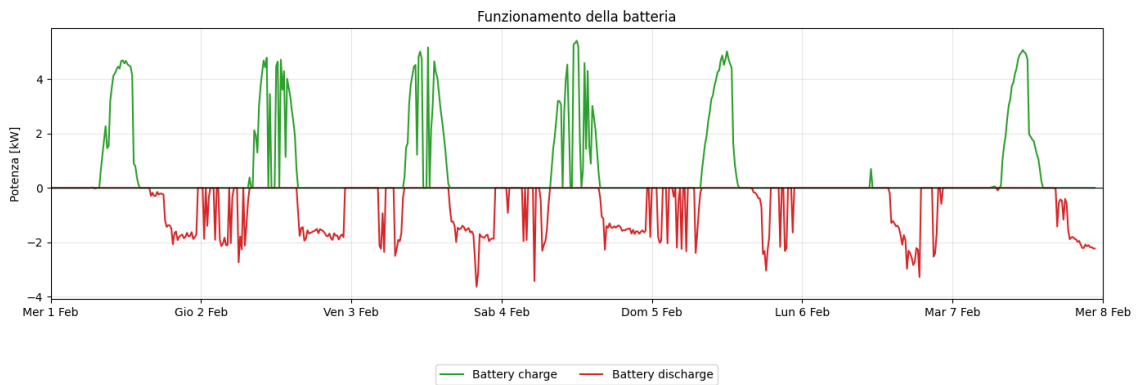


Figura 10: Andamento degli stati di carica e scarica della batteria con controllo MPC - caso studio residenziale

Nella Figura 10) è riportato l'andamento dello stato di carica (SOC) della batteria. Si nota una dinamica ciclica giornaliera, con fasi di carica prevalentemente nelle ore centrali quando abbiamo il surplus di energia prodotta dall'impianto fotovoltaico e di scarica nelle ore serali e notturne.

Anche in questo caso, osserviamo il comportamento tra il 5 e 6 febbraio, in cui possiamo notare come il giorno 5 la batteria viene caricata fino al valore prossimo di capacità massima (18 kWh), mentre il giorno 6 si registra una progressiva scarica. Questa dinamica evidenzia una strategia di controllo di accumulo inter-giornaliero, coerente con

l'orizzonte temporale predittivo di 24 ore, che permette al controllore di pianificare l'utilizzo dell'energia in funzione delle condizioni previste.

Questo comportamento suggerisce un coordinamento tra accumulo elettrico e accumulo termico, l'inerzia dell'edificio, entrambi sfruttati per perseguire l'obiettivo economico senza compromettere il comfort.

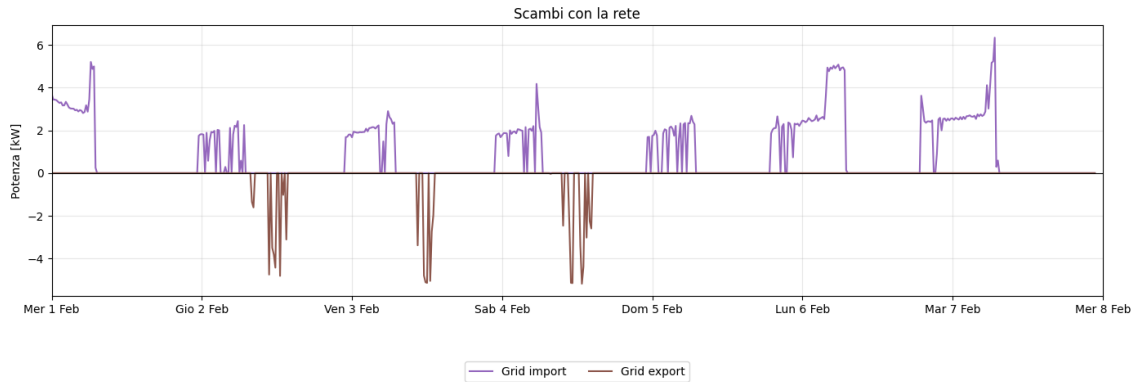


Figura 11: Andamento degli import ed export dalla rete con controllo MPC - caso studio residenziale

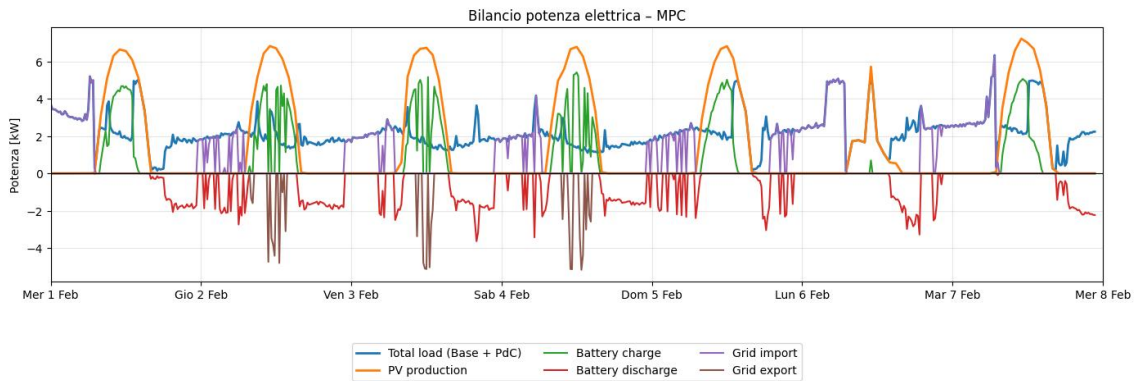


Figura 12: Bilancio della potenza elettrica con controllo MPC - caso studio residenziale

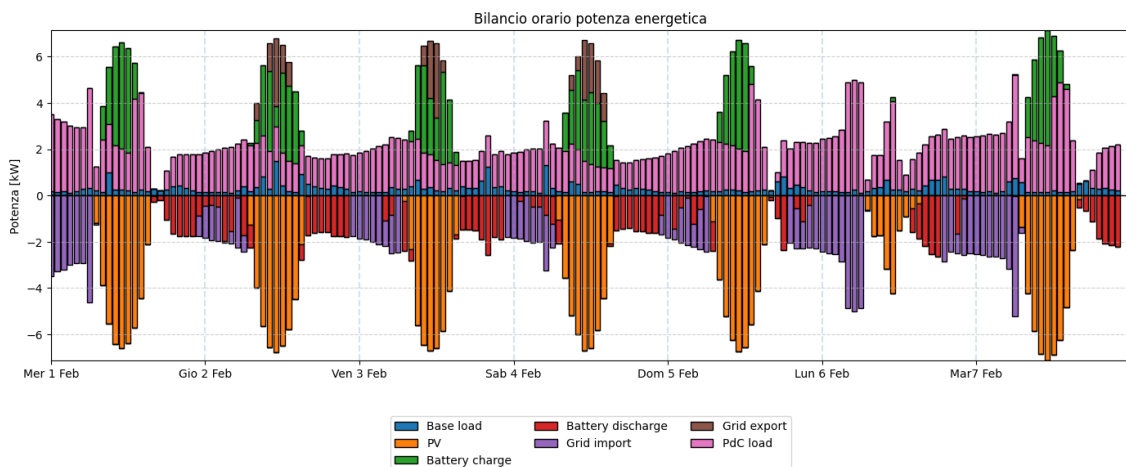


Figura 13: Bilancio orario della potenza elettrica con controllo MPC - caso studio residenziale

In Figura 13), è riportato il bilancio elettrico complessivo del sistema, comprensivo di carico totale richiesto dall'edificio, compreso sia del carico base che del fabbisogno della pompa di calore, della produzione fotovoltaica, scambi con la rete e flussi di carica e scarica della batteria. In particolare, le potenze positive rappresentano consumi elettrici dell'edificio e flussi di energia entranti nel sistema, inclusa la produzione dell'impianto fotovoltaico. I valori negativi indicano, invece, i flussi uscenti: nel caso della batteria corrispondono alla fase di scarica, quindi l'energia fornita al sistema, mentre per la rete rappresentano l'energia esportata verso l'esterno. In questo modo riusciamo a visualizzare in modo chiaro l'equilibrio istante per istante tra produzione, consumo, accumulo e scambi con la rete.

La produzione fotovoltaica raggiunge valori massimi superiori a 6 kW nelle ore centrali delle giornate più favorevoli, mentre il 6 febbraio, caratterizzato da condizioni di minore irraggiamento, si osserva una riduzione significativa della generazione.

In presenza di produzione solare elevata, il sistema privilegia l'autoconsumo diretto riducendo il prelievo dalla rete. Nelle ore serali e notturne, l'energia accumulata viene utilizzata per coprire parte del fabbisogno, limitando l'importazione dalla rete nelle fasce tariffarie meno conveniente.

Si osserva, inoltre, che nelle ore caratterizzate da tariffa F1, economicamente più conveniente rispetto alle altre fasce, il sistema tende a preferire l'importazione dalla rete anziché lo scaricamento della batteria, preservando così l'energia accumulata per le fasce orarie con costo maggiore. Questo comportamento evidenzia una gestione coerente con l'obiettivo di minimizzare il costo complessivo dell'energia.

Dal punto di vista economico, l'analisi degli scambi con la rete consente di quantificare l'efficacia della strategia di controllo adottata. Nel periodo considerato, l'energia complessivamente importata dalla rete ammonta a 140 kWh, mentre l'energia esportata in rete risulta pari a 23 kWh.

Il costo totale sostenuto per l'acquisto di energia elettrica dalla rete è pari a 27,01 €, a fronte di un ricavo derivante dalla vendita in rete di 1,16€. Il costo netto complessivo dell'energia risulta pertanto pari a 25,85€ sull'intero periodo di simulazione.

Questo risultato è coerente con la strategia orientata a privilegiare l'autoconsumo e l'utilizzo del sistema di accumulo rispetto all'immissione diretta dell'energia in rete.

La produzione fotovoltaica totale è pari a 268,83 kWh, di questi 132,24 kWh vengono consumati istantaneamente dall'edificio, corrispondenti a un autoconsumo diretto pari al 49,2% della generazione. Considerando anche l'energia fotovoltaica che viene utilizzata per ricaricare il sistema di accumulo, il livello complessivo di autoconsumo totale (91%) risulta superiore al valore diretto, evidenziando il ruolo strategico della batteria e ridurre l'energia scambiata con la rete.

Nel complesso, il controllo predittivo riesce a coordinare in modo efficiente la produzione fotovoltaica, l'accumulo elettrico e la gestione del carico termico, garantendo il rispetto dei vincoli di comfort e la minimizzazione nel costo energetico netto attraverso una strategia di autoconsumo.

### ***3.1.2. Controllo con Deep Reinforcement Learning***

Nel presente caso studio è stata implementata una strategia di controllo basata su DRL, la cui struttura generale è stata descritta del Capitolo 2.3. L'agente è stato addestrato utilizzando l'ambiente dinamico dell'edificio come ambiente di interazione, considerando l'intero mese di gennaio come periodo di training. La settimana compresa tra il 1° e il 7 febbraio è stata invece utilizzata esclusivamente come fase di test, con l'obiettivo di valutare la capacità di generalizzazione della politica appresa su condizioni climatiche non incluse nel processo di addestramento.

Durante la fase di training, l'agente ha interagito con l'ambiente con passo temporale di 15 minuti, osservando lo stato energetico, termico ed economico del sistema e determinando le azioni di controllo relative alla pompa di calore e al sistema di accumulo. L'addestramento è stato condotto per 300 episodi, in modo da consentire all'agente di esplorare adeguatamente le possibili configurazioni. Al termine del processo non è stato automaticamente scelto l'ultimo agente generato ma viene effettuata una selezione tra gli agenti salvati durante il training, considerando le prestazioni medie ottenute negli episodi di valutazione. L'agente che mostra il comportamento più stabile e la ricompensa media più elevata viene utilizzato per le simulazioni relative al periodo di test.

A differenza dell'MPC, che utilizza un modello esplicito e un orizzonte predittivo finito di 24 ore, il DRL apprende una politica di controllo attraverso l'esperienza accumulata durante l'addestramento, senza disporre di una previsione esplicita delle

grandezze future. Il comportamento del controllore è quindi maggiormente reattivo, poiché l'agente reagisce alle condizioni osservate istante per istante e non si basa su una pianificazione esplicita su base giornaliera.

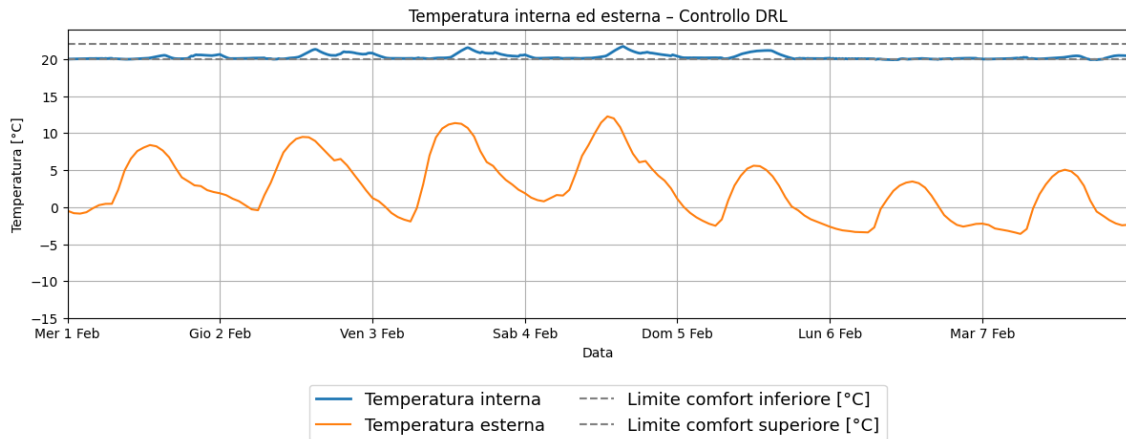


Figura 14: Andamento temperatura interna ed esterna con controllore DRL - caso studio residenziale

Osservando l'andamento della temperatura interna ed esterna durante il periodo di test, si nota come l'agente riesce a mantenere la temperatura all'interno della banda di comfort. Anche nei giorni caratterizzati da temperature estere più basse, in particolare il 5 e 6 febbraio, il sistema conserva una buona stabilità termica.

Il comportamento è coerente con la politica appresa, che reagisce alle condizioni istantanee del sistema. Tuttavia, l'agente sembra sfruttare in modo implicito l'inerzia termica dell'edificio. Infatti, in diverse giornate si osservano leggeri incrementi della temperatura nelle ore centrali, coincidente con la maggiore disponibilità di produzione fotovoltaica. Questo suggerisce che la politica abbia appreso, attraverso l'esperienza di training, la convenienza di anticipare parzialmente il fabbisogno termico quando l'energia è economicamente vantaggiosa.

Il DRL riesce quindi a determinare un comportamento che, pur non sebbene fondato su una previsione esplicita delle grandezze future, risulta coerente con una logica di sfruttamento del surplus e di utilizzo dell'accumulo termico dell'edificio.

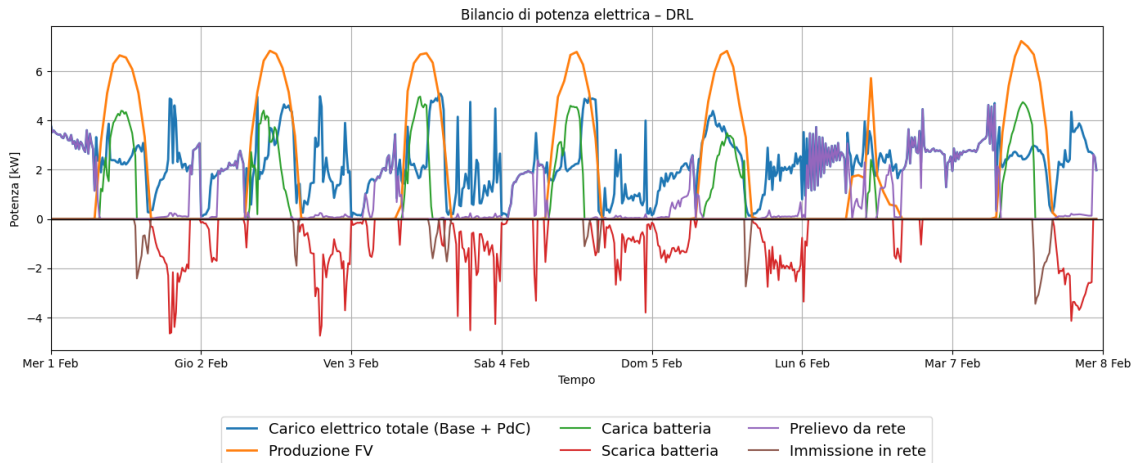


Figura 15: Bilancio della potenza elettrica con controllo DRL - caso studio residenziale

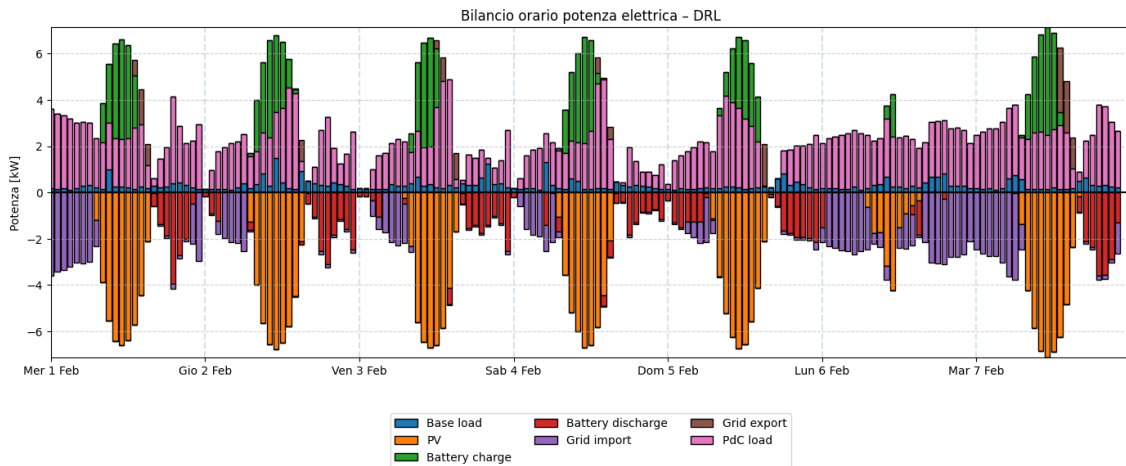


Figura 16: Bilancio orario della potenza elettrica con controllo DRL - caso studio residenziale

L'analisi del bilancio energetico consente di approfondire ulteriormente la logica seguita dall'agente. Nel periodo di test considerato, la produzione fotovoltaica complessiva risulta pari a 268,83 kWh, a fronte di un'energia esportata in rete di solo 15,92 kWh. Ne deriva un livello di autoconsumo totale molto alto che evidenzia un forte utilizzo dell'energia prodotta localmente, invece l'autoconsumo diretto ammonta a 147,4 kWh, corrispondente al 54,82% della produzione complessiva mentre la restante parte viene in larga misura destinata alla ricarica del sistema di accumulo.

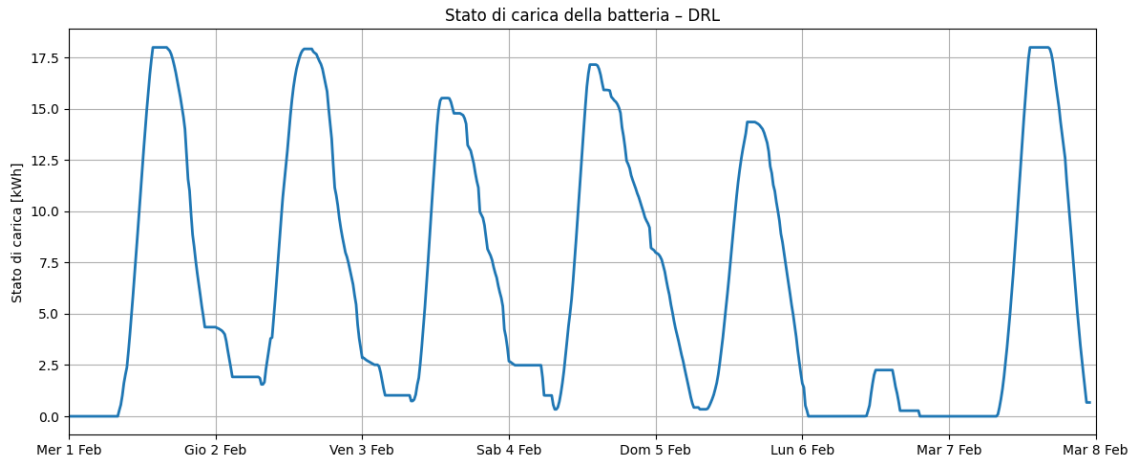


Figura 17: Andamento stato di carica della batteria con controllo DRL - caso studio residenziale

Il comportamento della batteria è ulteriormente chiarito dall'andamento dello stato di carica riportato in Figura 17). Si osserva anche qui una dinamica ciclica giornaliera, con cariche nelle ore centrali della giornata, in corrispondenza dei picchi della produzione fotovoltaica, e fasi di scarica nelle ore serali e notturne. Nel periodo di test si registra un'energia di carica pari a 105,5 kWh e di scarica pari a 94,6 kWh. Questo conferma l'utilizzo attivo della batteria come elemento di flessibilità, con l'obiettivo di ridurre l'energia scambiata con la rete e spostare nel tempo l'impiego della produzione fotovoltaica.

Dal punto di vista economico, il costo complessivo di acquisto è pari a 27,12€, a fronte di un ricavo dalla vendita in rete pari a 0,80€, determinando un costo netto totale nel periodo di test di 26,33€. Anche in questo contesto di evidenza come l'agente è in grado di adottare una strategia energeticamente ed economicamente efficace.

Nel modello implementato è stato inoltre considerata una strategia di controllo tradizionale basata su regolazione PID per la pompa di calore e su logiche rule-based per la gestione del sistema di accumulo. Questa configurazione, oltre a costruire un termine di confronto nelle analisi di prestazioni, è stata utilizzata anche come riferimento operativo durante la fase di sviluppo e di addestramento dell'agente DRL. La presenza di un controllore funzionante ha consentito di operare in un contesto dinamico già fisicamente coerente e stabile, evitando che l'agente si trovasse a esplorare uno spazio decisionale completamente casuale fin dalle prime interazioni. L'apprendimento non avviene per imitazione del PID, ma si sviluppa in un ambiente già strutturato, nel quale le variabili termiche ed energetiche evolvono secondo comportamenti realistici.

### 3.1.3. Applicazione del Transfer Learning

Al fine di valutare la capacità di generalizzazione e adattamento del controllore basato su DRL, è stato considerato un ulteriore scenario operativo in cui lo stesso edificio è analizzato in una diversa localizzazione geografica, identificata in Livigno. In questo scenario restano invariate le caratteristiche geometriche e la configurazione impiantistica, mentre cambiano esclusivamente le condizioni climatiche esterne.

Livigno è caratterizzata da un clima alpino con temperature invernali mediamente più rigide e maggiori escursioni termiche rispetto al caso precedente. Tali condizioni comportano un incremento del fabbisogno termico dell'edificio e una diversa dinamica della produzione fotovoltaica, rendendo la gestione energetica più impegnativa.

L'obiettivo di questa analisi è verificare se la politica appresa in fase di training sia in grado di mantenere prestazione adeguate anche in un contesto climatico più severo, senza modifiche strutturali al modello. A tal fine sono stati considerati due approcci distinti: nel primo caso la politica appresa è applicata direttamente al nuovo scenario (zero-transfer), mentre nel secondo è stata effettuata una fase di riaddestramento nel nuovo ambiente (fine-tuning), al fine di valutare la capacità di adattamento dell'agente.

#### *Transfer zero-shot*

In questo scenario, come già spiegato in precedenza, il migliore agente ottenuto dal DRL del caso residenziale viene applicato direttamente al nuovo ambiente, senza ulteriore fase di addestramento.

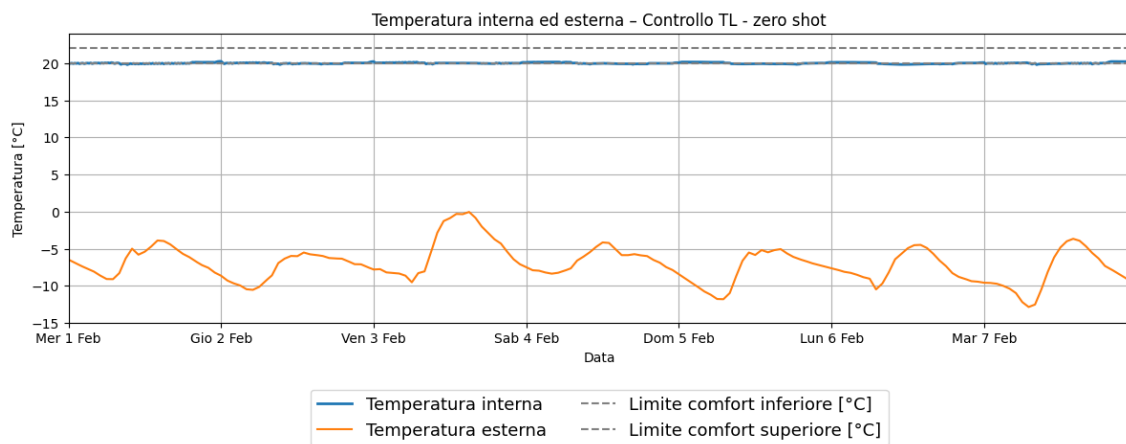


Figura 18: Andamento temperatura interna ed esterna con TL zero-shot - caso studio residenziale Livigno

Il nuovo contesto è caratterizzato da temperature esterne frequentemente inferiori a  $-5^{\circ}\text{C}$  e con valori minimi ancora più rigidi rispetto allo scenario di training.

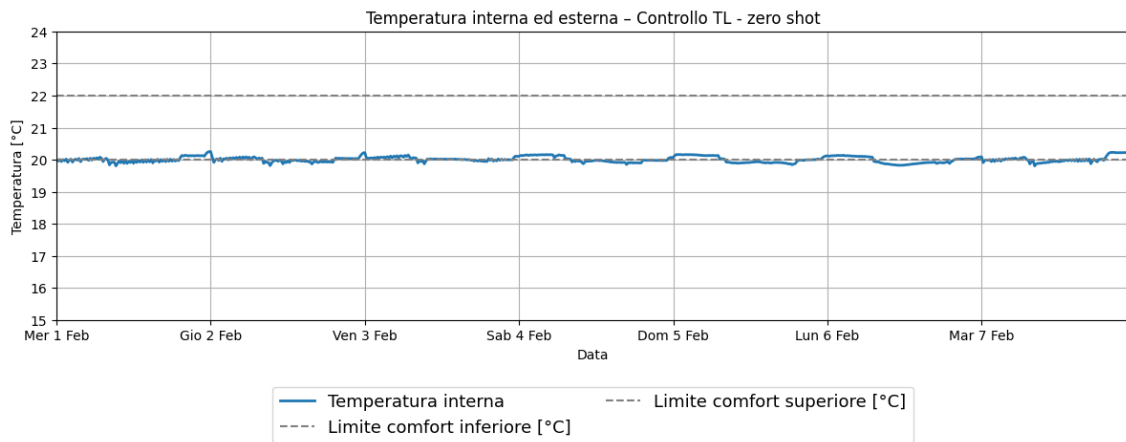


Figura 19: Andamento temperatura interna con TL zero-shot - caso studio residenziale Livigno

Per valutare con maggiore precisione il comportamento termico nel caso zero-shot, è utile osservare il dettaglio della sola temperatura interna rispetto alla banda di comfort. Il profilo evidenzia una dinamica regolare e priva di instabilità, con oscillazione contenuta intorno al set-point. L'errore medio rispetto al valore di riferimento risulta pari a  $-0,066$ , confermando una regolazione comunque accurata anche in un contesto climatico più severo. La temperatura tende a posizionarsi nella parte inferiore della banda di comfort, scelta coerente con un andamento del carico termico e con una strategia che privilegia il contenimento dei consumi senza compromettere in modo significativo le condizioni interne

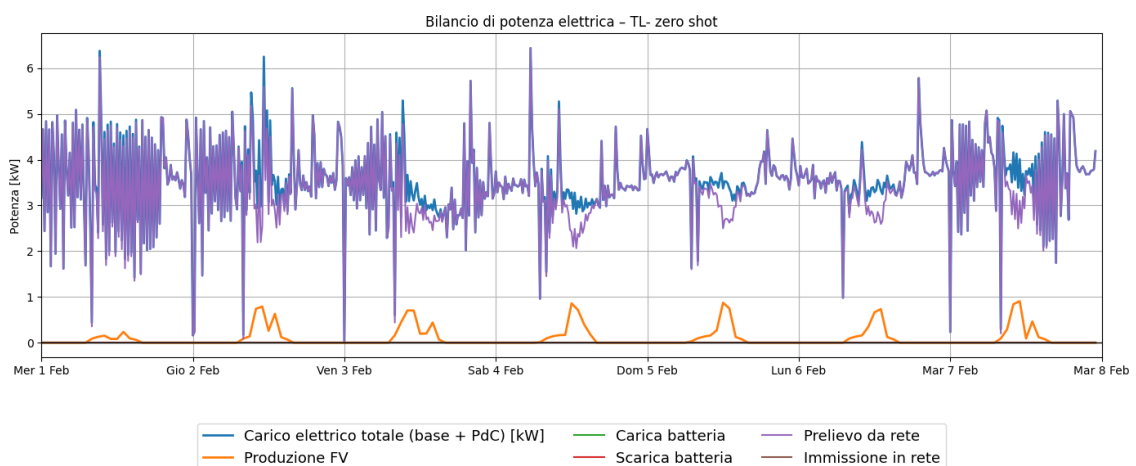


Figura 20: Bilancio della potenza elettrica con TL zero-shot - caso studio residenziale Livigno

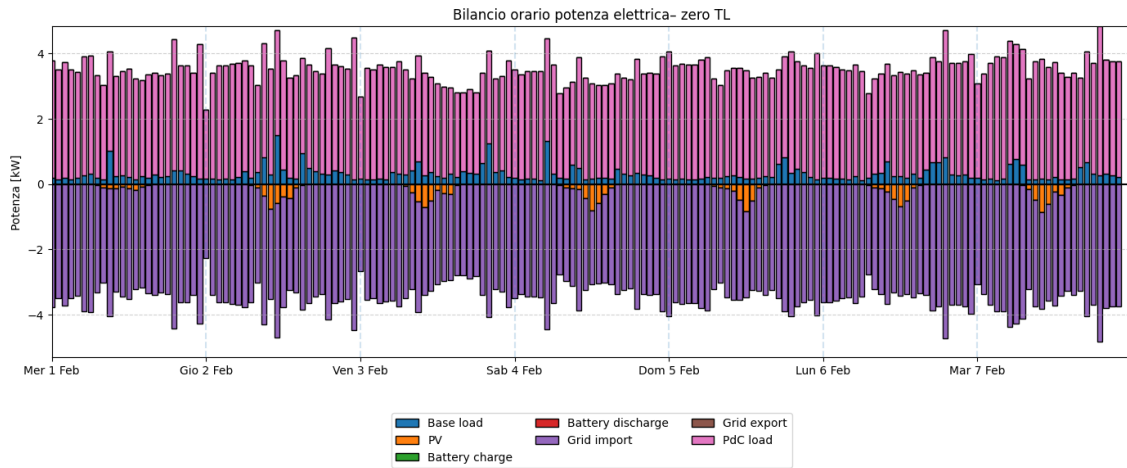


Figura 21: Bilancio orario della potenza elettrica con TL zero-shot - caso studio residenziale Livigno

Dal punto di vista energetico, il bilancio risulta diverso a causa della drastica riduzione della produzione fotovoltaica, pari a soli 17,14 kWh nell'intero periodo considerato. L'assenza di surplus energetico comporta la mancata attivazione del sistema di accumulo e un ricorso quasi completamente esclusivo della rete per la copertura del fabbisogno elettrico. L'aumento del costo complessivo non è quindi impotabile a un malfunzionamento della politica di controllo, ma al cambiamento strutturale del contesto energetico, caratterizzato da maggiore domanda termica e minore disponibilità di energia rinnovabile.

L'andamento del carico totale evidenzia, inoltre, un funzionamento della pompa di calore più intenso e meno regolare rispetto al caso precedente. Questo comportamento è coerente con l'aumento del fabbisogno termico, ma mostra una maggiore variazione dell'azione di controllo.

Nel complesso, il caso zero-shot conferma la robustezza della politica appresa dal punto di vista termico. Tuttavia, la maggiore irregolarità nel funzionamento della pompa di calore suggerisce che una fase di adattamento specifico possa contribuire a rendere l'azione di controllo più stabile nel nuovo dominio climatico.

### *Transfer fine-tuning*

Nel secondo scenario di trasferimento, l'agente precedentemente addestrato è stato ulteriormente riaddestrato nel nuovo contesto climatico. L'obiettivo non è modificare la logica di controllo, ma consentire alla politica di adattarsi in modo più mirato alle nuove dinamiche termiche ed energetiche introdotte dalle nuove condizioni climatiche più

rigide.

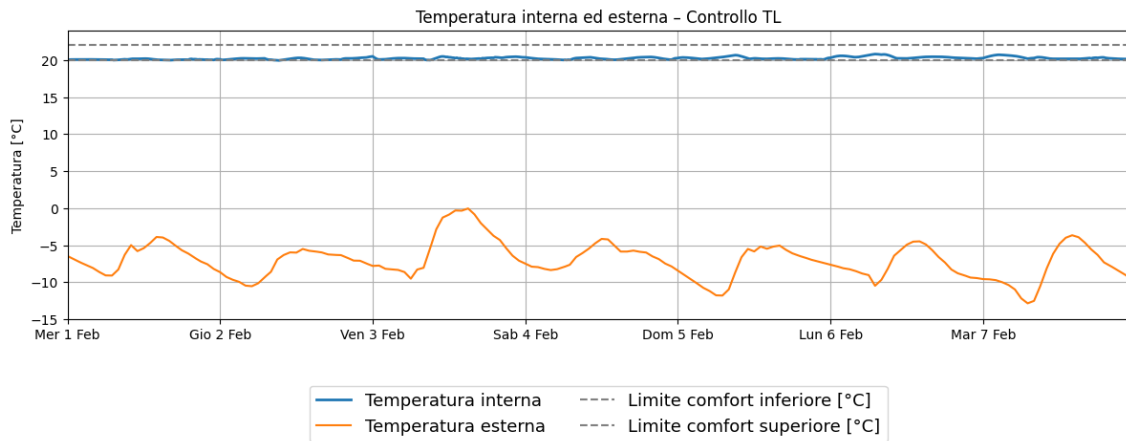


Figura 22: Andamento temperatura interna ed esterna con TL fine-tuning - caso studio residenziale Livigno

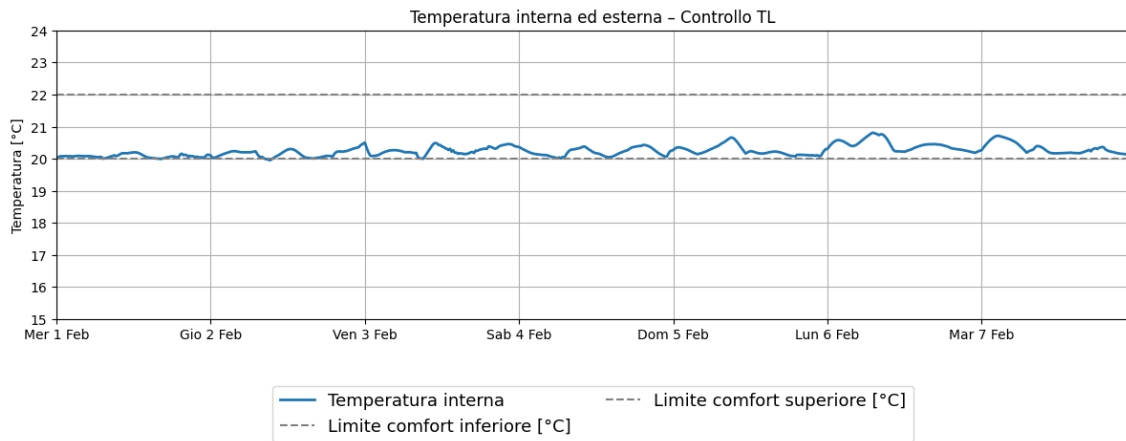


Figura 23: Andamento temperatura interna ed esterna con TL fine-tuning - caso studio residenziale Livigno

L'analisi della temperatura interna evidenzia un andamento leggermente più uniforme, rispetto al caso zero-shot. Le oscillazioni risultano più regolari e distribuite nel tempo, pur rimanendo all'interno della banda di comfort. In particolare, la regolazione rimane più stabile, ciò indica che l'agente ha affinato la propria risposta alle condizioni ambientali locali.

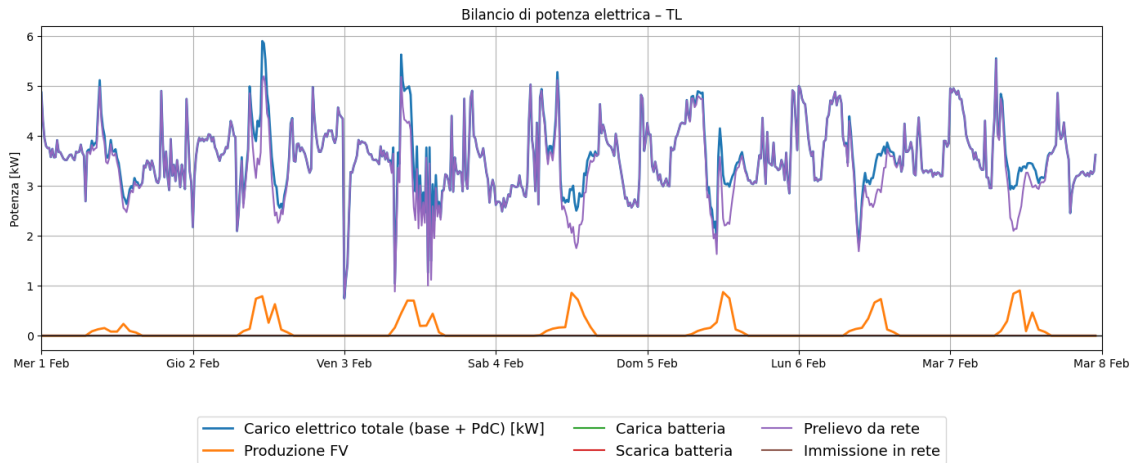


Figura 24: Bilancio della potenza elettrica con TL zero-shot - caso studio residenziale Livigno

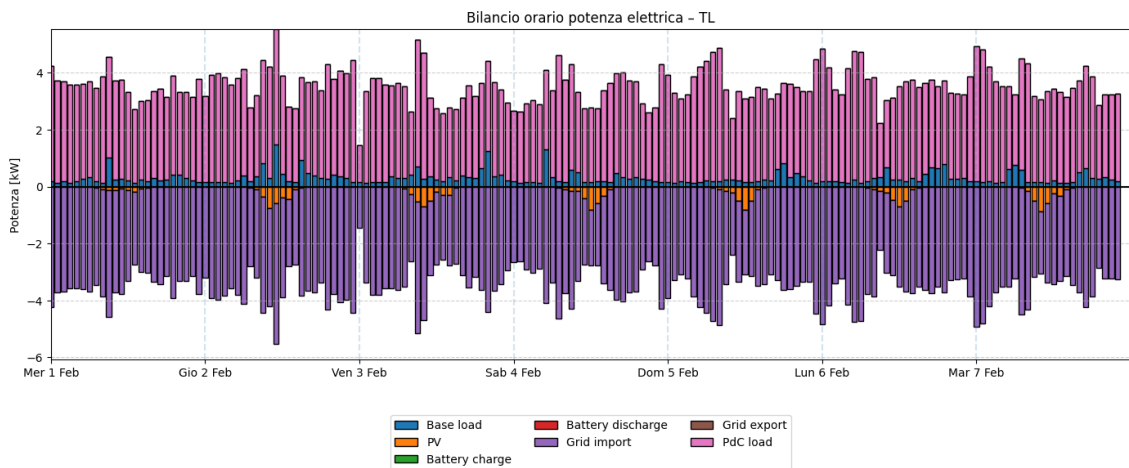


Figura 25: Bilancio orario della potenza elettrica con TL fine-tuning - caso studio residenziale Livigno

Il miglioramento risulta ancora più evidente osservando il carico totale, fortemente influenzato dal contributo della pompa di calore. Nel caso fine-tuning, il suo funzionamento appare meno irregolare e maggiormente coerente con l'andamento del fabbisogno termico. La modulazione della potenza risulta più graduale e meno impulsiva, suggerendo un utilizzo più efficace dell'inerzia termica dell'involucro edilizio, che riesce ad attuare le variazioni e ridurre le fluttuazioni del carico.

Dal punto di vista energetico, il quadro complessivo non subisce variazioni sostanziali rispetto al caso zero-shot, poiché la produzione fotovoltaica non subisce variazioni e di conseguenza il fabbisogno elettrico continua a essere coperto quasi totalmente dal prelievo dalla rete, con valori di energia e costo complessivamente uguali al caso precedente zero-shot.

Il fine-tuning, quindi, non modifica la quantità di energia disponibile, ma migliora la qualità della regolazione. La fase di addestramento consente all'agente di adattare con maggiore coerenza la propria azione di controllo al nuovo dominio climatico, rendendo il comportamento dell'impianto più stabile e meglio allineato alle caratteristiche termiche dell'edificio.

### 3.1.4. Controllo tradizione PID + RB

Analizziamo ora il controllo tradizionale, al fine di osservare le differenza rispetto ai controlli avanzati analizzati.

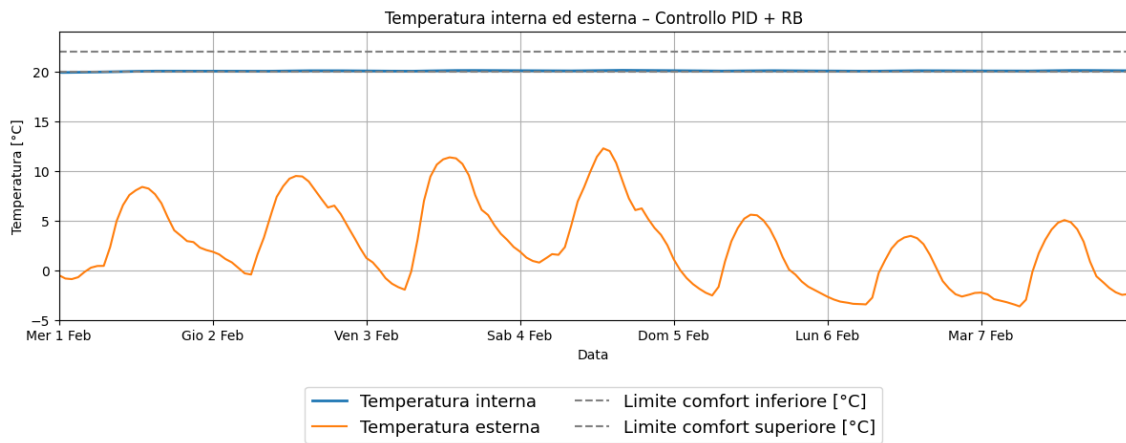


Figura 26: Andamento temperatura interna ed esterna con controllore PID + RB - caso studio residenziale

Come si osserva dal grafico, il controllo tradizionale è in grado di mantenere la temperatura interna costante sul valore di set-point, garantendo quindi il rispetto delle condizioni di comfort. Tuttavia, questo approccio non sfrutta in modo efficace l'inerzia termica dell'edificio, poiché non sfrutta il sistema di generazione in funzione e la disponibilità energetica, ad esempio in presenza di surplus fotovoltaico.

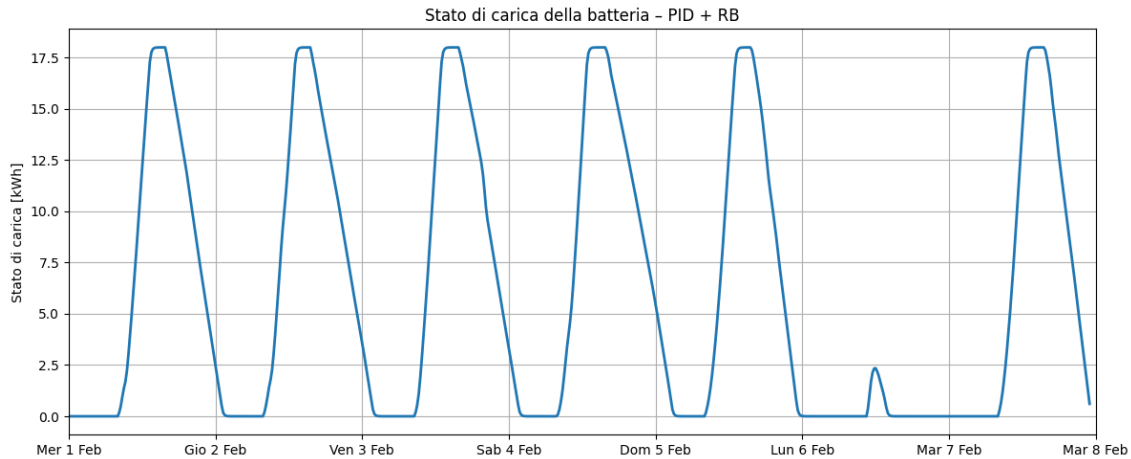


Figura 27: Andamento stato di carica della batteria con controllo PID + RB - caso studio residenziale

L'andamento dello stato di carica della batteria evidenzia cicli giornalieri regolari di carica e scarica, guidati principalmente dalla disponibilità di energia fotovoltaica. Durante le ore diurne, la batteria viene caricata fino alla capacità massima, mentre nelle ore serali e notturne viene progressivamente scaricata per soddisfare il fabbisogno dell'edificio. Tuttavia, questo comportamento risulta tuttavia poco strategico, in quanto la gestione della batteria è basata su regole fisse e non su una logica di ottimizzazione.

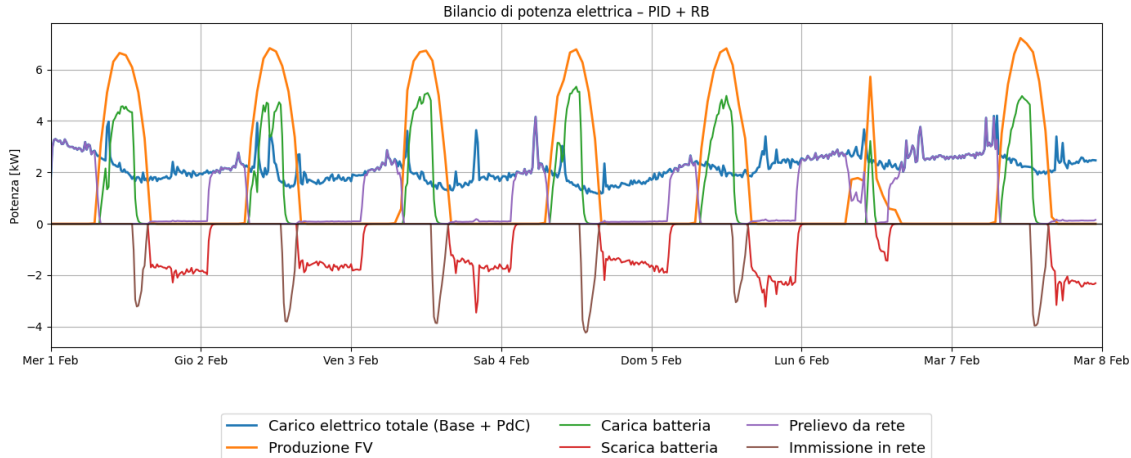


Figura 28: Bilancio della potenza elettrica con controllo PID + RB - caso studio residenziale

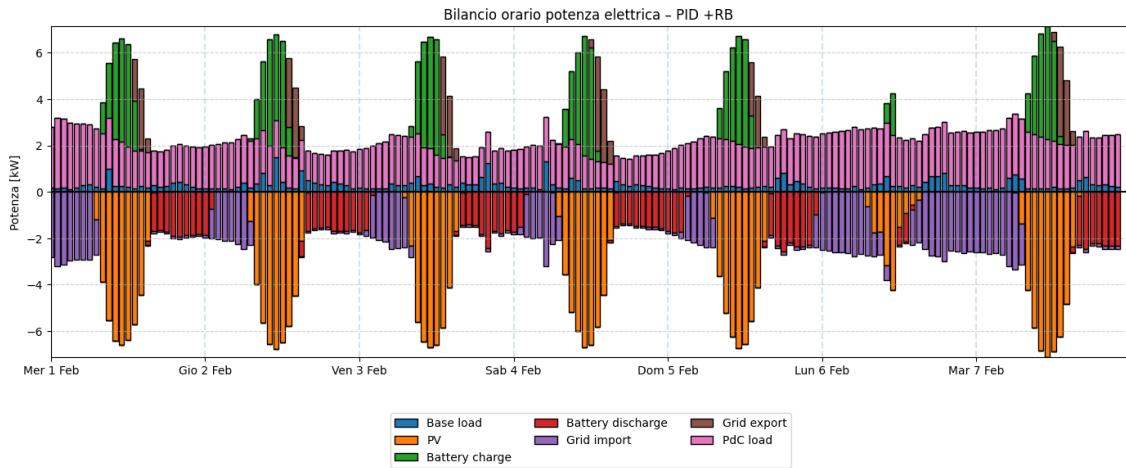


Figura 29: Bilancio orario della potenza elettrica con controllo PID + RB - caso studio residenziale

Dal punto di vista energetico, il comportamento del sistema risulta fortemente influenzato dalle logiche rule-based adottate. Durante le ore centrali della giornata, in presenza di produzione fotovoltaica, il sistema privilegia l'autoconsumo diretto e la ricarica della batteria. Tuttavia, l'assenza di una pianificazione su orizzonte temporale porta a una gestione meno efficiente dell'energia rispetto ai controlli predittivi.

Infatti, come evidenziano i risultati, abbiamo un valore di energia importata dalla rete di 148 kWh ed esportata in rete di 39 kWh, entrambi i valori superiori rispetto alle strategie avanzate. Questo è legato alla minore capacità del controllore di coordinare in modo efficace produzione, accumulo e carico dell'edificio. Nel complesso, l'autoconsumo diretto dell'energia prodotta è del 42 %, valore inferiore rispetto agli altri approcci considerati. Dal punto di vista economico, il controllo PID + RB presenta quindi prestazioni inferiori. Abbiamo un costo d'acquisto dell'energia di 30 € ed un ricavo dalla vendita di 2 €, per un costo netto di circa 28 €. Il costo netto più elevato è dovuto principalmente alla maggiore dipendenza dalla rete e alla minore capacità di ottimizzare l'utilizzo dell'energia prodotta localmente.

### 3.1.5. Confronto delle strategie di controllo

Per sintetizzare in modo chiaro e immediato i risultati ottenuti nei diversi scenari analizzati, si riporta di seguito una tabella comparativa che raccoglie i principali indicatori energetici ed economici. Il confronto include le strategie applicate al caso residenziale di riferimento (MPC e DRL) e le due configurazioni di TL nel contesto climatico di Livigno.

Indicatore	MPC (Residenziale)	DRL (Residenziale)	PID+ Rule-Based (RB)
Comfort rispettato	Sì	Sì	Sì
Energia importata [kWh]	136,66	134,24	148,56
Energia esportata [kWh]	23,13	15,92	39,29
Consumi PdC [kWh]	317	325	316
Produzione FV [kWh]	268,83	268,83	268,83
Autoconsumo diretto [%]	49,2	54,82	42,07
Energia batteria carica [kWh]	113,46	105,52	116,44
Energia batteria scarica [kWh]	95,7	94,59	105,2
Costo netto [€]	25,85	26,33	27,97

Tabella 2: Confronto fra le diverse strategie di controllo - caso studio residenziale

Dall'analisi dei risultati emerge innanzitutto che tutte le strategie garantiscono il rispetto dei vincoli di comfort termico, mostrando che le scelte di controllo risultano coerenti con gli obiettivi prefissati. Le differenze si manifestano invece sul piano energetico ed economico.

Considerando le sole strategie avanzate, il DRL registra un'energia importata leggermente inferiore rispetto all'MPC, con un incremento dell'autoconsumo diretto. Si tratta di piccole differenze, che evidenziano come entrambe le soluzioni siano in grado di gestire efficacemente l'interazione tra produzione fotovoltaica, accumulo e domanda dell'edificio. Dal punto di vista economico, infatti, i costi netti risultano molto simili, confermando prestazioni complessivamente allineate.

Il confronto con la configurazione PID + RB mette maggiormente in evidenza il contributo dei controlli avanzati. In questo caso, il controllo PID applicato alla pompa di calore, combinato con una gestione rule-based della batteria, presenta un valore più elevato di energia importata rispetto a MPC e DRL, segno di una minore capacità di coordinare in modo integrato i flussi energetici del sistema. Anche l'autoconsumo diretto risulta inferiore rispetto alle strategie avanzate, mentre l'energia esportata assume un valore più elevato.

L'analisi del sistema di accumulo conferma queste differenze. Nel caso PID + rule-based si osservano valori di energia caricata e scaricata leggermente superiori rispetto a MPC e DRL, ma senza che ciò traduca in una riduzione significativa dei prelievi dalla

rete. Al contrario, le strategie avanzate mostrano un bilanciamento più coerente tra energia di carica e scarica, segno di una gestione più strategica e orientata alla minimizzazione dei costi energetici.

Sebbene le prestazioni globali di MPC e DRL risultino simili, le due strategie si distinguono nella modalità operativa. L'MPC adotta una logica esplicitamente pianificata, basata su un orizzonte predittivo di 24 ore, mentre il DRL si fonda su una politica appresa che reagisce alle condizioni osservate istante per istante.

Questa differenza emerge in modo più evidente osservando l'andamento dello stato della batteria nei giorni 5 e 6 febbraio, caratterizzati da temperature più rigide. Nel caso MPC si nota una carica significativa il 5 febbraio, seguita da una scarica progressiva il giorno successivo, in assenza di produzione fotovoltaica. Il comportamento appare coerente con la capacità del controllore predittivo di anticipare condizioni meno favorevoli e di pianifica l'utilizzo dell'accumulo. Nel caso DRL, invece, la batteria tende a scaricarsi nella notte stessa, evidenziando una dinamica più reattiva. Questo andamento è coerente con la politica appresa che si basa sulle condizioni osservate istante per istante, senza una pianificazione esplicita su base giornaliera. Questa differenza non compromette il comfort né altera in modo sostanziale il bilancio energetico, ma evidenzia come le due strategie si distinguano nella modalità con cui gestiscono l'accumulo elettrico in base alle condizioni climatiche più critiche.

Per quando riguarda il trasferimento nel contesto climatico di Livigno, si riporta di seguito il confronto tra le configurazioni di transfer learning zero-shot e fine-tuning.

Indicatore	TL Zero-shot (Livigno)	TL Fine-tuning (Livigno)
Comfort rispettato	Sì	Sì
Energia importata [kWh]	574,79	580,38
Energia esportata [kWh]	0	0
Consumi PdC [kWh]	541	547
Produzione FV [kWh]	17,14	17,14
Autoconsumo diretto [%]	100	100
Energia batteria carica [kWh]	0	0
Energia batteria scarica [kWh]	0	0
Costo netto [€]	127,72	129,06

Tabella 3: Confronto fra le diverse strategie di controllo - caso studio residenziale Livigno

Anche in questo caso il comfort termico risulta pienamente garantito. Tuttavia, il cambiamento delle condizioni climatiche incide in modo significativo sul bilancio energetico. L'energia importata dalla rete supera i 570 kWh in entrambe le configurazioni, a fronte di una produzione fotovoltaica estremamente ridotta. In queste condizioni, l'autoconsumo diretto raggiunge il 100%, ma il contributo dell'accumulo risulta praticamente nullo, non essendovi sufficiente energia in eccesso da immagazzinare.

Le differenze tra zero-shot e fine-tuning sugli indicatori globali risultano contenute. Il fine-tuning non determina variazioni sostanziali nei valori di energia importata o nel costo netto; tuttavia, la leggera differenza economica osservata è principalmente legata alla regolazione della temperatura interna. Nel caso fine-tuning, infatti, la temperatura media risulta leggermente più elevata e pienamente all'interno della banda di comfort, mentre nello scenario zero-shot si osservano occasionalmente valori leggermente inferiori al limite inferiore della banda. Questo comporta un fabbisogno energetico leggermente maggiore nel caso fine-tuning e, di conseguenza, un costo complessivo leggermente superiore.

I risultati mettono in luce come in uno scenario climatico fortemente penalizzante per la produzione solare, la disponibilità di risorse energetiche locali incida più della specifica modalità di trasferimento della politica di controllo. Il contributo del fine-tuning emerge principalmente sul piano di stabilità e dell'adattamento, più che sugli indicatori energetici aggregati.

### **3.2. *Caso studio scolastico***

Per ampliare l'analisi delle strategie di controllo e valutare il comportamento dell'agente in un contesto energetico differente rispetto al caso residenziale, è stato introdotto un secondo caso studio rappresentato da un edificio ad uso scolastico. L'obiettivo è disporre di un dominio caratterizzato da dimensioni e carichi energetici maggiori rispetto all'edificio residenziale, mantenendo tuttavia dinamiche operative comparabili. Questo consente di analizzare la capacità dell'agente di adattarsi a condizioni di funzionamento differenti e di valutare con maggiore gradualità la capacità di generalizzazione della politica di controllo.

È stato considerato un edificio con pianta quadrata di 30 m x 30 m e un'altezza di 3 m. La stratigrafia dell'involucro opaco è stata assunta identica a quella del caso residenziale, così come le superfici trasparenti, scalate in funzione delle nuove dimensioni geometriche. Di conseguenza, i valori di trasmittanza termica risultano invariati rispetto al caso residenziale.

Per quanto riguarda la ventilazione, è sempre stata adottata la medesima metodologia normativa del caso residenziale (UNI EN 16798), incrementando solo il numero di occupanti in funzione della maggiore superficie disponibile.

$$H_{ve} = 2,014 \frac{kW}{K} \quad (48)$$

Anche i guadagni interni sono stati determinati coerentemente al caso residenziale, mantenendo lo stesso criterio normativo ma aumentando il numero di occupanti in modo proporzionale alla nuova area calpestabile.

Una volta determinati i contributi di trasmittanza ventilazione è stato calcolato il carico termico invernale di progetto pari a:

$$\tilde{\Phi}_{P,finale}^* = \frac{\tilde{\Phi}_P^*}{\eta_{impianto}} = 100 kW \quad (49)$$

Questo valore evidenzia la diversa scala dell'edificio scolastico rispetto al caso residenziale e introduce un contesto energetico caratterizzato da carichi più elevati.

Definite le caratteristiche dell'edificio scolastico, si procede ora all'analisi delle strategie di controllo applicate a questo nuovo caso studio, come fatto nel caso precedente del residenziale.

Di seguito i parametri a confronto del caso studio che andremo ad analizzare e del precedente per analizzarlo rispetto al Transfer.

Parametri	Caso residenziale	Caso scolastico
Volume [m <sup>3</sup> ]	360	2700
Area calpestabile [m <sup>2</sup> ]	120	900
Dimensione PdC [kW]	9,5	100
Dimensione PV [kW]	10	70
Dimensione accumulo [kWh]	18	70

Tabella 4: Parametri a confronto del modello energetico - caso residenziale e caso scolastico

### 3.2.2. Controllo con MPC

Per il caso studio scolastico è stata mantenuta la stessa struttura di controllo predittivo descritta per l'edificio residenziale, senza modifiche nella formulazione della funzione obiettivo né nell'orizzonte di predizione, fissato a 24 ore con passo temporale di 15 minuti. Ciò che cambia in modo sostanziale è il modello fisico dell'edificio che presenta dimensioni e carico termico di progetto maggiore e, di conseguenza, un sistema di generazione, sistema fotovoltaico e accumulo elettrico opportunamente dimensionati su una scala energetica superiore.

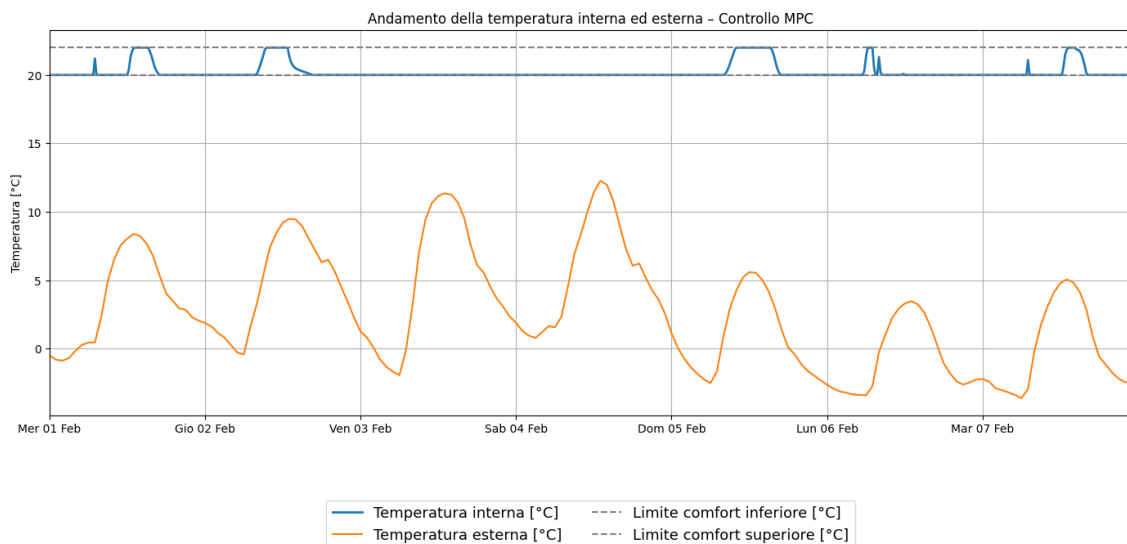


Figura 30: Andamento temperatura interna ed esterna con controllore MPC - caso studio scolastico

L'andamento della temperatura interna, durante la settimana di simulazione, evidenzia come il controllore riesca a mantenere le condizioni all'interno della banda di comfort. La temperatura tende a rimanere per lunghi intervalli vicino al limite inferiore della

banda, mentre in alcuni momenti cresce rapidamente fino a valori prossimi al limite superiore. Questo comportamento è legato alla logica di ottimizzazione economica del controllore, che tende ad attivare la pompa di calore nei momenti più favorevoli dal punto di vista energetico, portando temporaneamente la temperatura verso il limite superiore della banda. Successivamente l'edificio sfrutta la propria inerzia termica, permettendo alla temperatura di scendere gradualmente senza compromettere le condizioni di comfort.

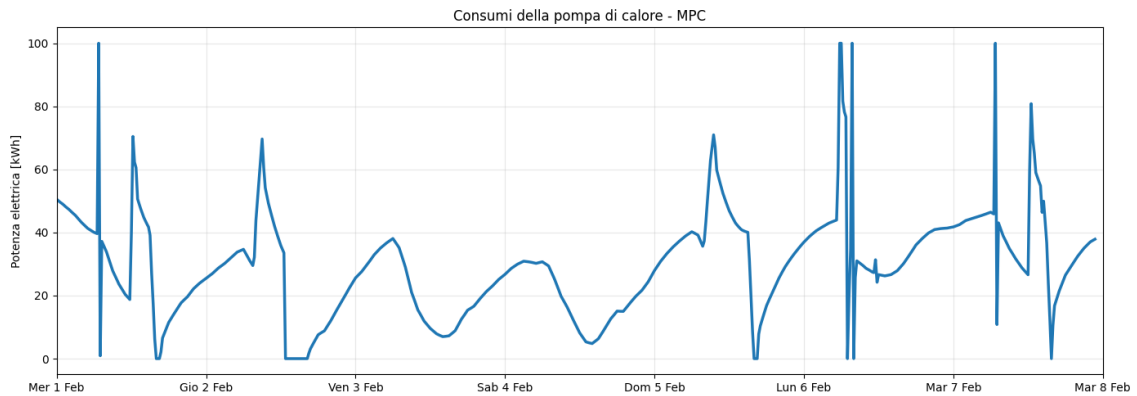


Figura 31: Andamento potenza elettrica consumata dalla pompa di calore con controllo MPC- caso studio scolastico

Analizzando l'andamento della potenza della pompa di calore si osserva una modulazione coerente con il fabbisogno termico dell'edificio e con le condizioni climatiche esterne. Nei giorni caratterizzati da temperature più basse la potenza media risulta maggiore. In generale, l'MPC tende a distribuire nel tempo l'energia fornita dalla pompa di calore, sfruttando l'inerzia dell'edificio per limitare variazioni troppo ripide della temperatura interna.

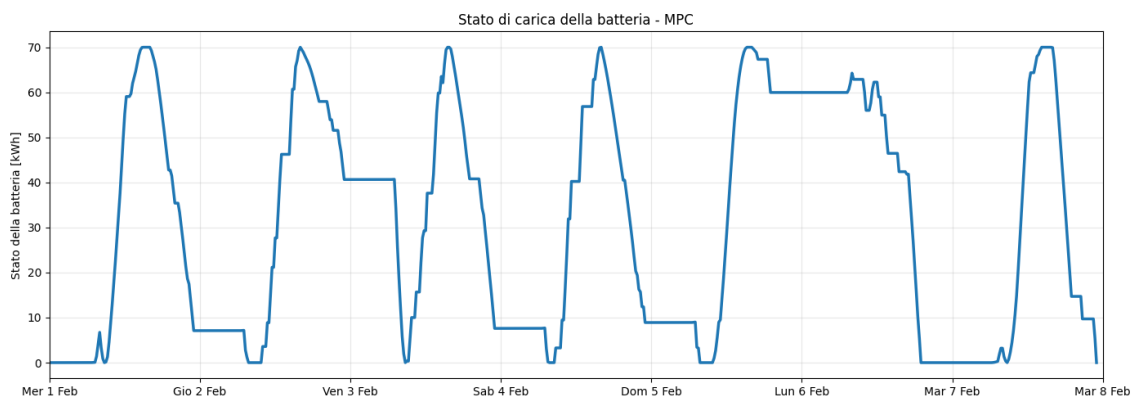


Figura 32: Andamento stato di carica della batteria con controllo MPC - caso studio scolastico

La gestione della batteria riflette la stessa logica predittiva già osservata nel caso residenziale. In alcune situazioni, quando il prezzo dell'energia nelle fasce notturne risulta relativamente conveniente, il sistema tende a preferire l'importazione della rete piuttosto

che l'utilizzo immediato dell'energia accumulata. In questo modo l'MPC preserva l'energia della batteria per le ore successive potenzialmente più costose dal punto di vista tariffario. Questo comportamento è visibile anche nel grafico successivo osservando i momenti in cui il sistema ricorre al prelievo dalla rete nonostante la presenza di energia accumulata

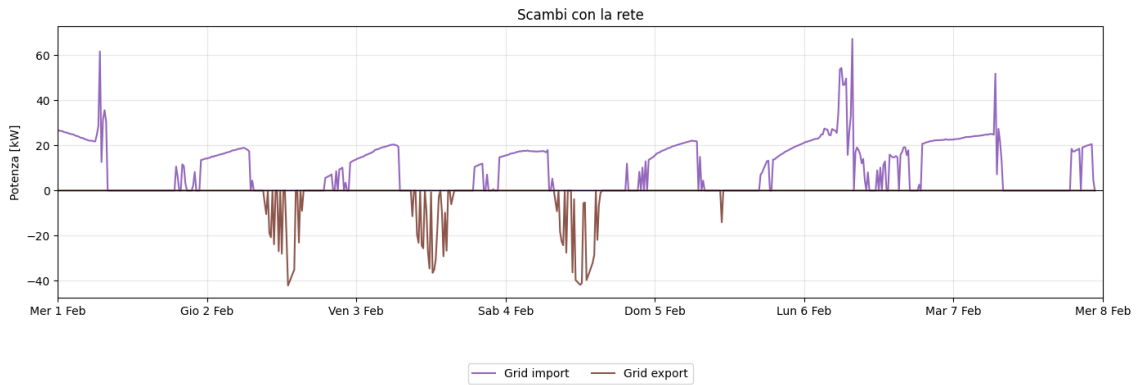


Figura 33: Andamento degli import ed export dalla rete con controllo MPC - caso studio scolastico

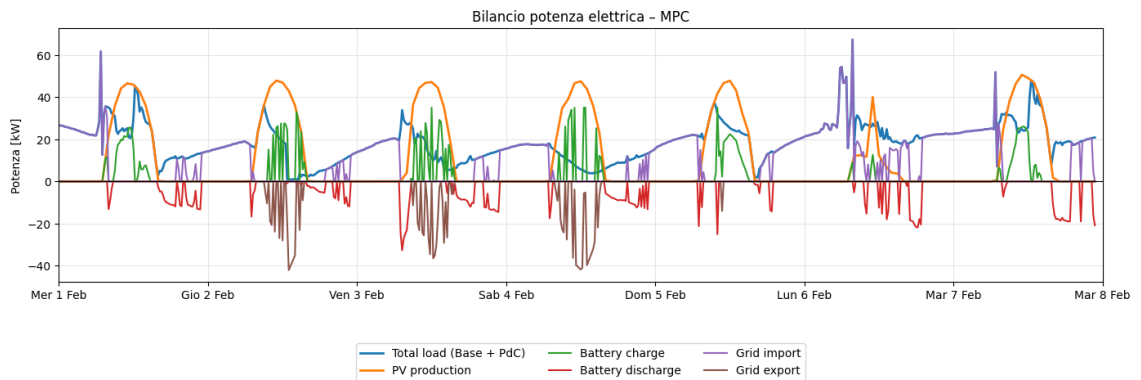


Figura 34: Bilancio della potenza elettrica con controllo MPC - caso studio scolastico

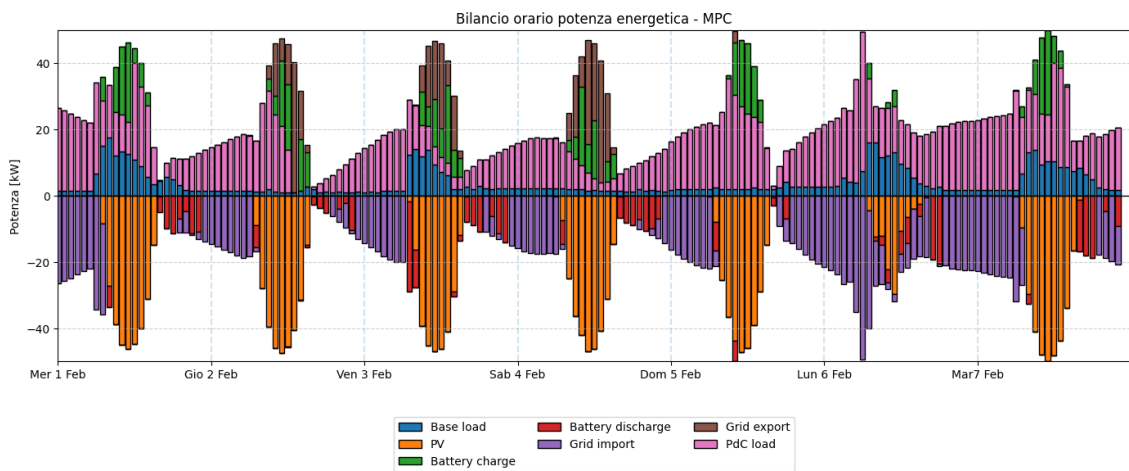


Figura 35: Bilancio orario della potenza elettrica con controllo MPC - caso studio scolastico

Il bilancio elettrico complessivo evidenzia chiaramente l'interazione tra produzione fotovoltaica, carichi elettrici dell'edificio, sistema di accumulo e scambi con la rete. Nelle ore centrali della giornata la produzione fotovoltaica raggiunge i 45-50 kW e, nel periodo considerato, la produzione totale risulta paria a circa 1882 kWh. In queste condizioni il sistema utilizza principalmente l'energia rinnovabile disponibile per alimentare il carico dell'edificio e ricaricare la batteria. Durante le ore serali e notturne, in assenza di produzione solare, il fabbisogno elettrico viene coperto attraverso una combinazione di energia accumulata e prelievo dalla rete.

Complessivamente, l'energia importata dalla rete ammonta a 1617 kWh, mentre quella esportata è pari a 353 kWh, con un autoconsumo diretto del 56,57%. Considerando gli scambi con la rete, il costo netto totale nel periodo analizzato risulta pari a 307,93€. Questo evidenzia come il controllore riesca a sfruttare in modo efficace la produzione fotovoltaica, il sistema di accumulo e il sistema di generazione attraverso una logica predittiva, mantenendo al tempo stesso le condizioni di comfort termico e gestendo in modo economicamente efficiente gli scambi di energia con la rete.

### ***3.2.3. Controllo con Deep Reinforcement Learning***

Nel caso dell'edificio scolastico è stata implementata la stessa architettura di controllo basata sul DRL già descritta per il caso residenziale. L'agente utilizza la medesima struttura di rete neurale e la stessa formulazione della funzione di ricompensa.

La principale differenza riguarda il dominio fisico dell'ambiente di interazione, caratterizzato da dimensioni dell'edificio e carichi energetici significativamente superiori. Di conseguenza, le grandezze energetiche e costi coinvolte nel processo di apprendimento risultano di ordine maggiore rispetto al caso residenziale. Per mantenere stabile il processo di addestramento è stata quindi applicata una riscaldatura della funzione di ricompensa.

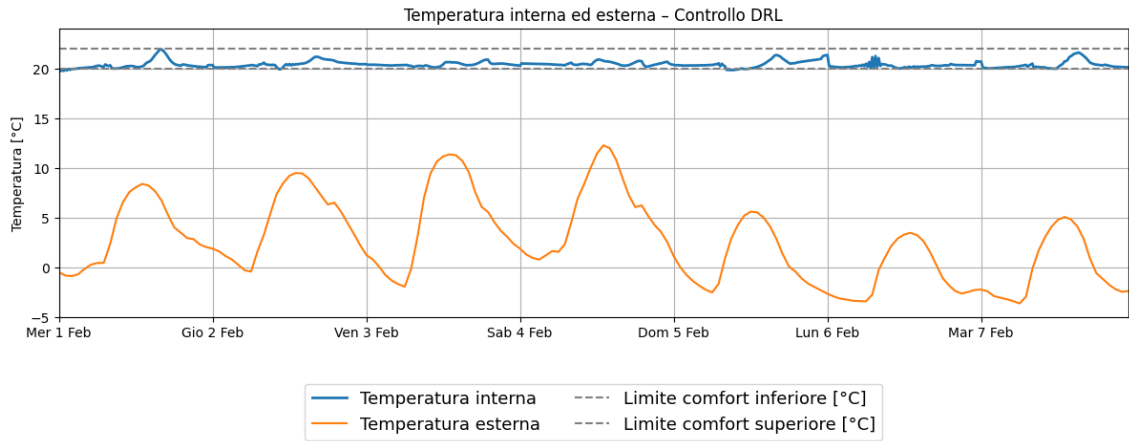


Figura 36: Andamento temperatura interna ed esterna con controllore DRL - caso studio scolastico

L'analisi dell'andamento della temperatura interna mostra che l'agente riesce a mantenere il comfort termico per l'intero periodo di test. Rispetto al comportamento osservato con il controllo MPC, l'azione di controllo appare più regolare e meno soggetta a variazioni brusche della potenza termica. La temperatura tende infatti a rimanere stabile attorno al valore di set-point, suggerendo una gestione più progressiva del carico termico che sfrutta l'inerzia termica dell'edificio senza generare variazioni oscillazioni marcate.

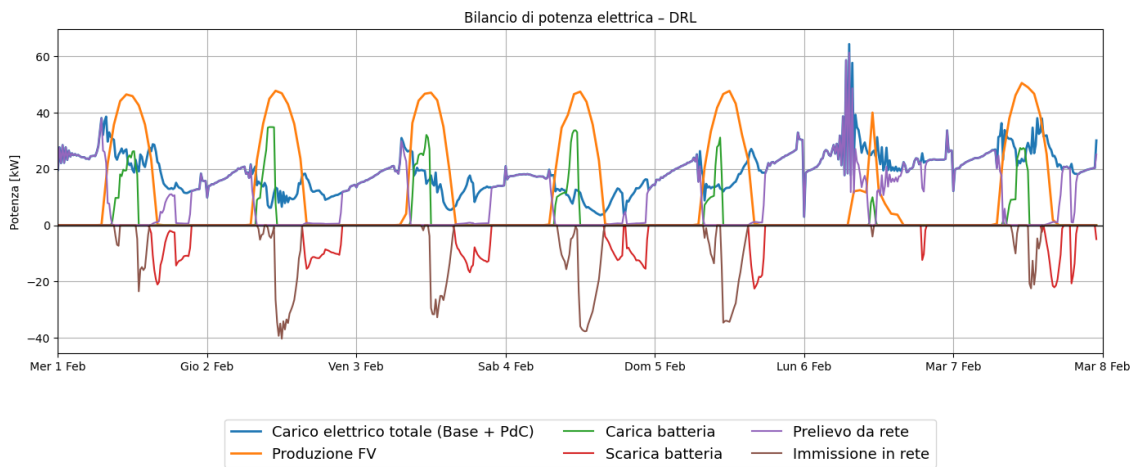


Figura 37: Bilancio della potenza elettrica con controllo DRL - caso studio scolastico

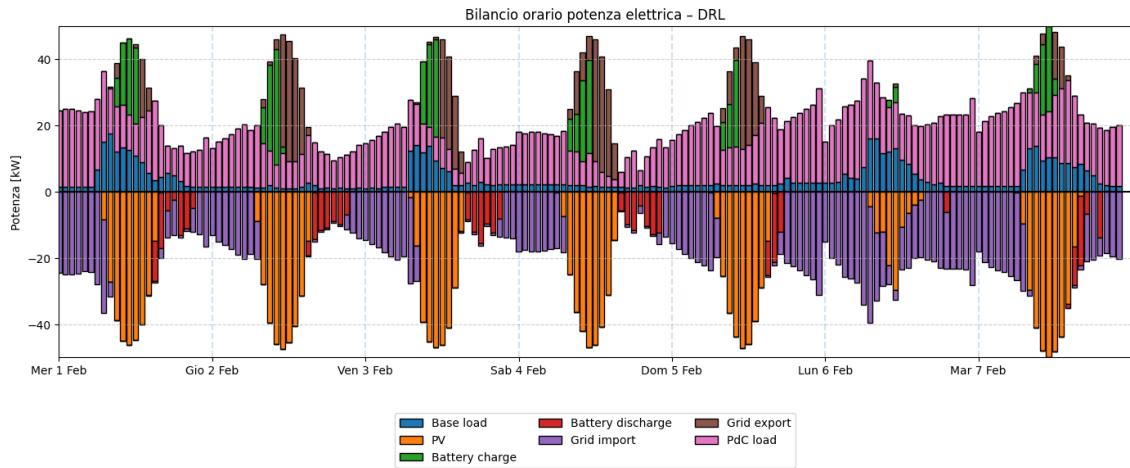


Figura 38: Bilancio orario della potenza elettrica con controllo DRL - caso studio scolastico

Il bilancio elettrico evidenzia una gestione ottima sia del sistema fotovoltaico sia del sistema di accumulo. Nelle ore centrali della giornata, in presenza di produzione solare significativa, l'agente tende a privilegiare l'utilizzo diretto dell'energia rinnovabile disponibile, riducendo il prelievo dalla rete e attivando la ricarica della batteria. Durante le ore serali e notturne, l'energia accumulata viene progressivamente utilizzata per coprire parte del fabbisogno elettrico dell'edificio. Una volta raggiunto il limite minimo di stato di carica, il sistema torna a ricorrere al prelievo dalla rete per soddisfare il carico residuo.

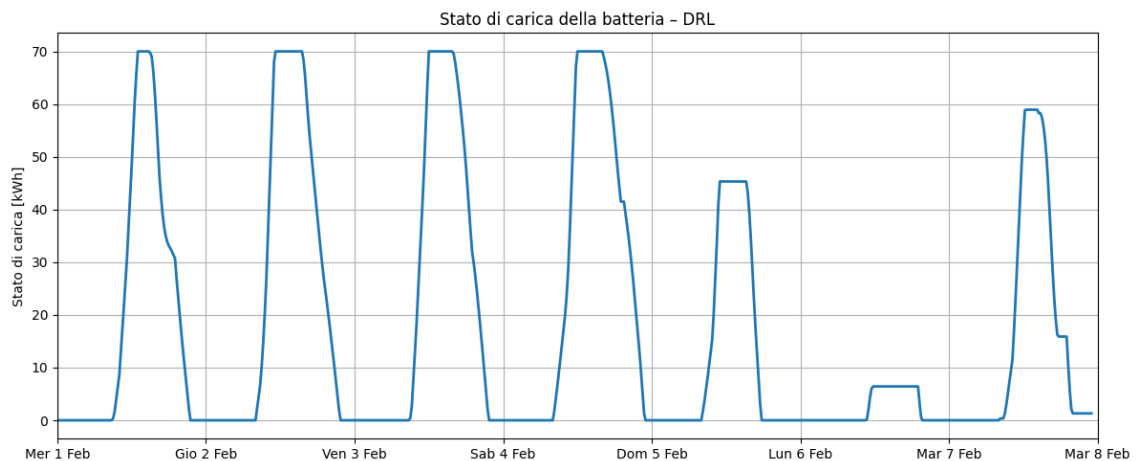


Figura 39: Andamento stato di carica della batteria con controllo DRL - caso studio scolastico

L'andamento dello stato di carica evidenzia una dinamica ciclica giornaliera piuttosto marcata. Nei primi giorni la batteria viene caricata fino alla capacità massima disponibile, per poi scaricarsi progressivamente nelle ore successive. Nelle giornate con minore produzione fotovoltaica, come nel caso del 6 febbraio, il livello di carica è molto basso e l'edificio ricorre in misura maggiore al prelievo dalla rete. Questo comportamento è

coerente con la politica appresa che reagisce alle condizioni istantanee di produzione e carico, come lo era nel caso residenziale precedente.

Dal punto di vista quantitativo, nel periodo analizzato l'energia importata dalla rete ammonta a 1880 kWh, mentre l'energia esportata a 537 kWh. La produzione fotovoltaica è di 1882 kWh, permettendo un autoconsumo diretto pari a 933 kWh corrispondente al 49.6%. Il sistema di accumulo svolge un ruolo significativo nel bilancio energetico con un valore di carica di 411 kWh così come di scarica. Questo evidenzia l'utilizzo attivo della batteria come elemento flessibile all'interno dell'edificio.

Dal punto di vista economico, il costo complessivo di acquisto dell'energia dalla rete è pari a 388 €, a fronte di un ricavo derivante dalla vendita in rete di 27€, determinando un costo netto finale di 361€ sull'intero periodo di simulazione. Anche questi valori rispecchiano la maggiore scala dell'edificio, caratterizzato da fabbisogni energetici significativamente superiori rispetto al caso residenziale.

Il comportamento del DRL nel caso scolastico risulta coerente e stabile. L'agente riesce a coordinare i diversi sistemi presenti all'interno dell'edificio, mantenendo le condizioni di comfort e adattando la propria azione alle variazioni climatiche e di disponibilità di energia rinnovabile.

#### ***3.2.4. Applicazione del Transfer Learning***

In questo scenario è stato considerato l'uso del Transfer Learning per analizzare la capacità dell'agente basato sul DRL addestrato nel caso residenziale di trasferire la politica di controllo appresa su un edificio caratterizzato da condizioni operative differenti.

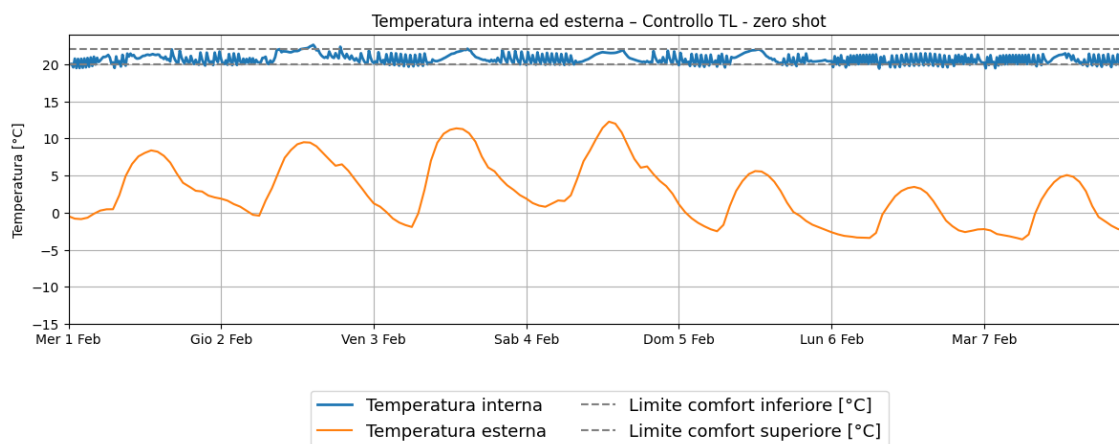
Inizialmente, l'obiettivo dell'analisi prevedeva l'applicazione diretta dell'agente addestrato nel caso residenziale a un edificio industriale, caratterizzato da una scala energetica significativamente più elevata. Tuttavia, la differenza molto marcata tra i due domini avrebbe reso difficile valutare in modo graduale la capacità di adattamento della politica appresa. Per questo motivo è stato introdotto un caso studio intermedio, rappresentato da un edificio scolastico, che rappresenta caratteristiche dimensionali e carichi energetici maggiori al caso residenziali ma meno estremi rispetto al contesto industriale.

Questo approccio consente di analizzare in modo progressivo il comportamento dell'agente quando viene applicato a un dominio con caratteristiche differenti rispetto a quello utilizzato durante l'addestramento.

Nell'analisi del trasferimento è stata condotta considerando due diverse modalità di applicazione della politica appresa. In un primo scenario è stato analizzato il caso di transfer zero-shot, nel quale l'agente addestrato nel caso residenziale viene applicato direttamente all'edificio scolastico senza effettuare ulteriori fasi di addestramento o adattamento dei parametri della rete neurale. Successivamente è stato considerato su uno scenario di transfer fine-tuning, in cui la politica precedentemente appresa viene ulteriormente riaddestrata nel nuovo ambiente di interazione, consentendo all'agente di adattarsi alle diverse condizioni operative dell'edificio scolastico.

### ***Transfer zero-shot***

Nel caso zero-shot, la politica appresa nel dominio residenziale viene applicata direttamente all'edificio scolastico senza ulteriori fasi di addestramento.



*Figura 40: Andamento temperatura interna ed esterna con TL zero-shot - caso studio scolastico*

L'andamento della temperatura interna mostra che l'agente riesce comunque a mantenere le condizioni di comfort all'interno della banda prevista, nonostante la presenza di oscillazioni più frequenti rispetto al caso in cui l'agente viene addestrato direttamente sul dominio scolastico. Questo comportamento è legato al fatto che la politica di controllo è stata appresa in un ambiente caratterizzato da dinamiche termiche e carichi energetici differenti. Nonostante questa differenza tra i due domini, il comfort termico non viene

compromesso, evidenziando una discreta capacità di generalizzazione della politica appresa.

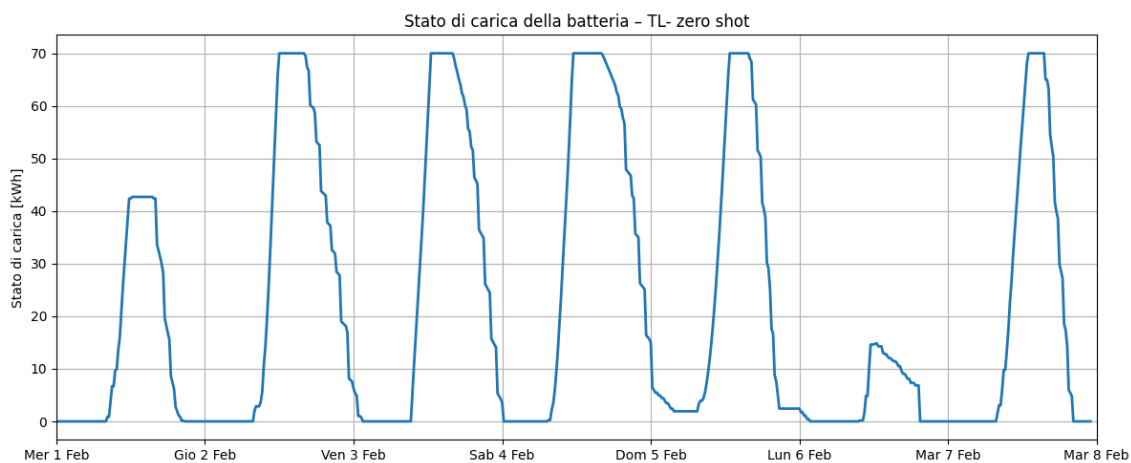


Figura 41: Andamento stato di carica della batteria con TL zero-shot - caso studio scolastico

L'analisi dello stato di carica della batteria mostra una dinamica giornaliera simile a quella osservata nel caso di addestramento DRL per il caso scolastico. Nei giorni in cui è presente del surplus di produzione fotovoltaica, la batteria viene ricaricata completamente, per poi essere scaricata nelle ore serali. Nei giorni caratterizzati da una minore produzione solare, il livello di carica rimane più contenuto.

Una differenza si osserva nel primo giorno di simulazione, dove lo stato di carica presenta un comportamento leggermente diverso rispetto agli altri giorni. Questo effetto è legato alle condizioni iniziali del sistema e alle differenti caratteristiche del dominio applicativo rispetto al caso di addestramento.

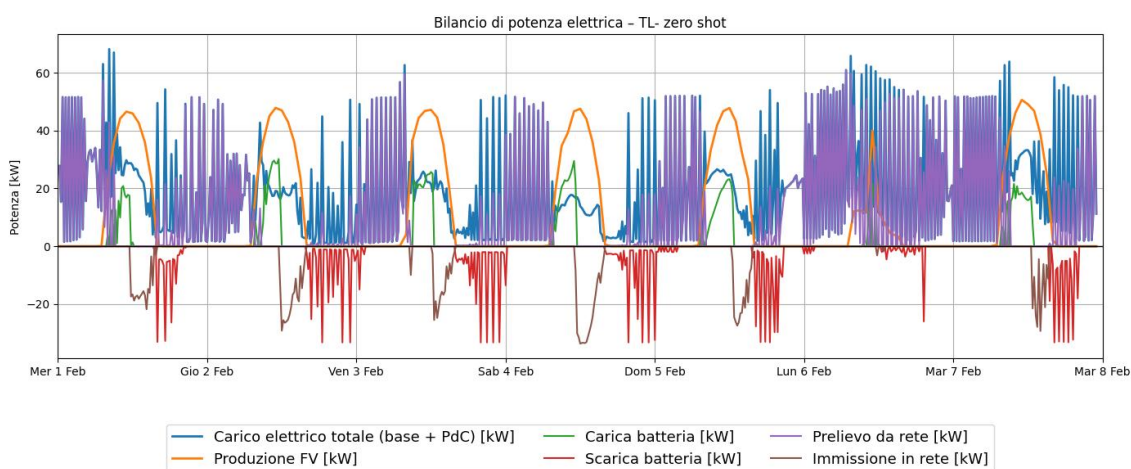


Figura 42: Bilancio della potenza elettrica con TL zero-shot - caso studio scolastico

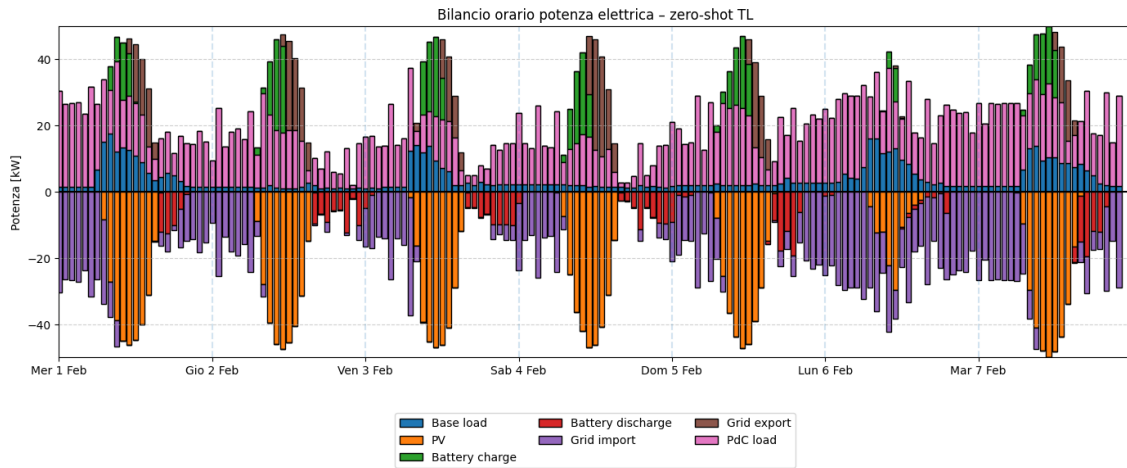


Figura 43: Bilancio orario della potenza elettrica con TL zero-shot - caso studio scolastico

Il bilancio elettrico complessivo ci fa notare un comportamento coerente con quanto osservato nell'andamento della temperatura interna. Nelle ore centrali della giornata, in presenza di produzione fotovoltaica, il sistema utilizza principalmente l'energia rinnovabile disponibile per alimentare il carico dell'edificio e ricaricare la batteria. Durante le restanti ore, invece, si osservano variazioni più frequenti nella potenza della pompa di calore, che si riflettono anche nelle oscillazioni della temperatura interna. Questo comportamento è dovuto alla politica di controllo applicata ad un carico base differente dal caso residenziale. In un edificio scolastico, infatti, il carico elettrico è generalmente più elevato nelle ore centrali della giornata e diminuisce nelle ore serali e notturne, determinando una dinamica operativa diversa rispetto all'ambiente di addestramento.

Nel periodo di simulazione, l'energia importata dalla rete risulta pari a circa 1841 kWh, mentre quella esportata ammonta a 404 kWh. La produzione fotovoltaica è pari a 1882 kWh, con un autoconsumo diretto di 55,86%. Considerando anche il sistema di accumulo, che raggiunge circa 1478 kWh, si arriva ad un autoconsumo totale pari a 78,54%. Questi valori portano ad un costo complessivo di acquisto dalla rete pari a 380€, mentre il ricavo dalla vendita di energia pari a 20€, determinando un costo totale netto di 360€.

Il trasferimento zero-shot mostra come la politica di controllo appresa nel caso residenziale è in grado di operare anche nel dominio scolastico mantenendo le condizioni di comfort nella banda ottimale. Tuttavia, le oscillazioni sulla temperatura e sulla potenza della pompa di calore indicano che la politica non è completamente adattata alle nuove

caratteristiche del sistema, evidenziando il potenziale beneficio di una successiva fase di addestramento tramite fine-tuning.

### *Transfer fine-tuning*

Nel secondo scenario viene analizzato il caso fine-tuning, nel quale la politica appresa del dominio residenziale viene ulteriormente riaddestrata utilizzando l'ambiente dell'edificio scolastico.

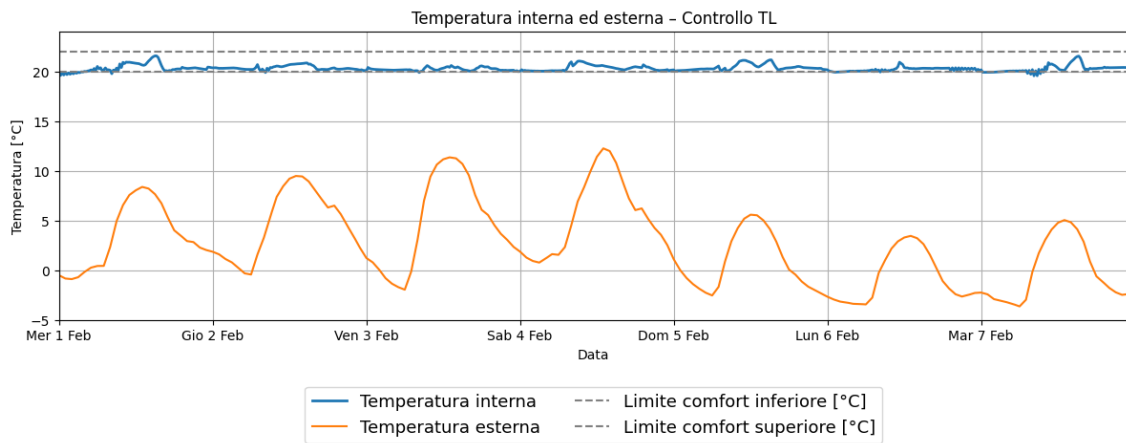


Figura 44: Andamento temperatura interna ed esterna con TL fine-tuning - caso studio scolastico

La temperatura interna mostra un comportamento stabile e ben controllato, rimanendo all'interno dei limiti di comfort imposti. Rispetto allo scenario zero-shot, il contributo risulta più regolare e meno soggetto a variazioni rapide della potenza termica. Questo indica che la fase di addestramento ha permesso alla politica di adattarsi meglio alle dinamiche dell'edificio scolastico.

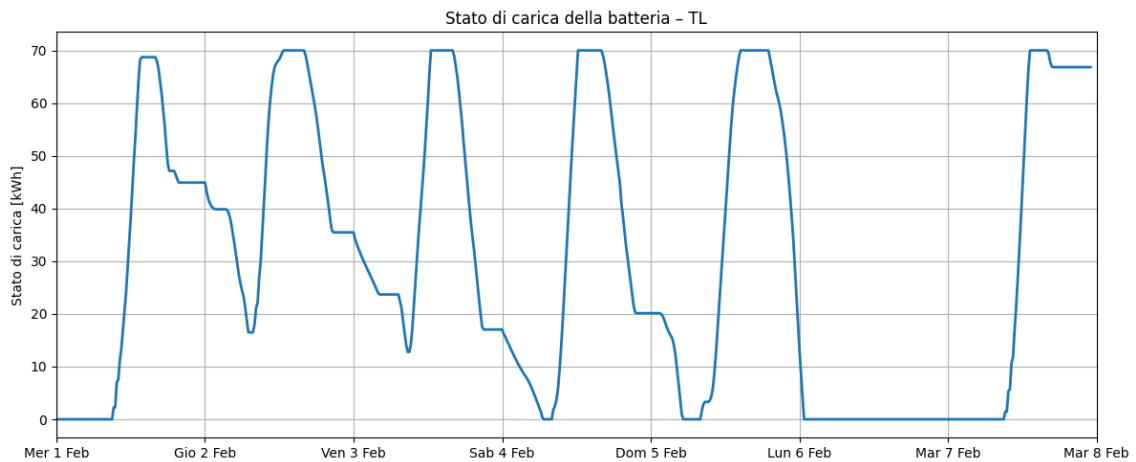


Figura 45: Andamento stato di carica della batteria con TL fine-tuning - caso studio scolastico

Lo stato di carica della batteria mostra una dinamica differente. Le fasi di carica si concentrano nelle ore centrali della giornata, quando la produzione solare supera il fabbisogno dell'edificio. Tuttavia, a differenza di quanto osservato nelle altre strategie del caso studio, la fase di scarica risulta meno efficace. I valori energetici evidenziano infatti come l'energia immagazzinata non viene successivamente utilizzata per coprire il la richiesta energetica dell'edificio.

Questo comportamento suggerisce che, nonostante la fase di fine-tuning abbia migliorato la stabilità complessiva del controllo, la politica appresa non riesce ancora a sfruttare in modo ottimale il sistema di accumulo. In particolare, la batteria tende a rimanere parzialmente carica per lunghi intervalli di tempo, invece di essere scaricata in modo più sistematico nelle ore di maggiore richiesta energetica o nelle fasce tariffarie meno convenienti.

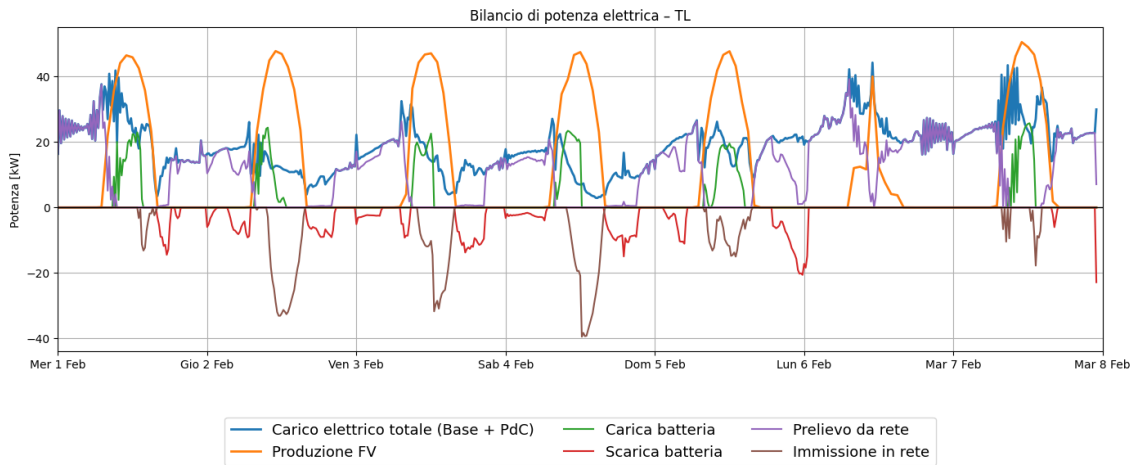


Figura 46: Bilancio della potenza elettrica con TL fine-tuning - caso studio scolastico

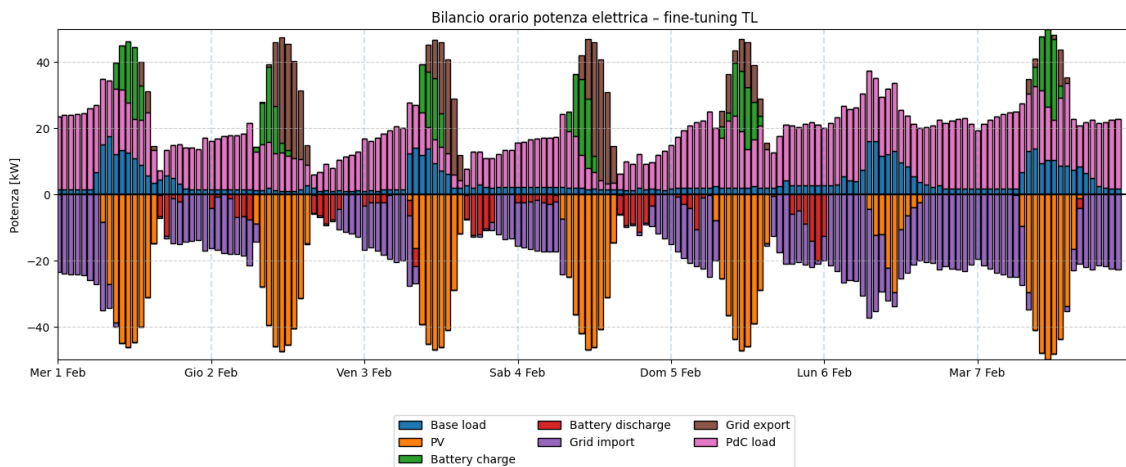


Figura 47: Bilancio orario della potenza elettrica con TL fine-tuning - caso studio scolastico

Anche l'analisi del bilancio energetico evidenzia un comportamento coerente con quello osservato nel caso di addestramento diretto tramite DRL. Nelle ore centrali la produzione fotovoltaica contribuisce in modo significativo alla coperta energetica dell'edificio. Durante le ore serali e notturne, invece, il fabbisogno residuo viene coperto attraverso una combinazione di energia accumulata e prelievo dalla rete.

Nel periodo considerato per la simulazione, l'energia totale importata dalla rete risulta pari a 1854 kWh e quella esportata in rete ammonta a circa 472 kWh. La produzione fotovoltaica totale paria a 1882 kWh, permettendo un autoconsumo diretto di 53,1%. Dal punto di vista economico, abbiamo un costo di acquisto dalla rete di 388€, e di vendita pari a 24€, per un costo netto totale di circa 364€.

### 3.2.5. *Controllo tradizione PID + RB*

Osserviamo il comportamento del controllo tradizionale nel caso dell'edificio scolastico, in modo da analizzare le differenze rispetto alle strategie di controllo avanzate.

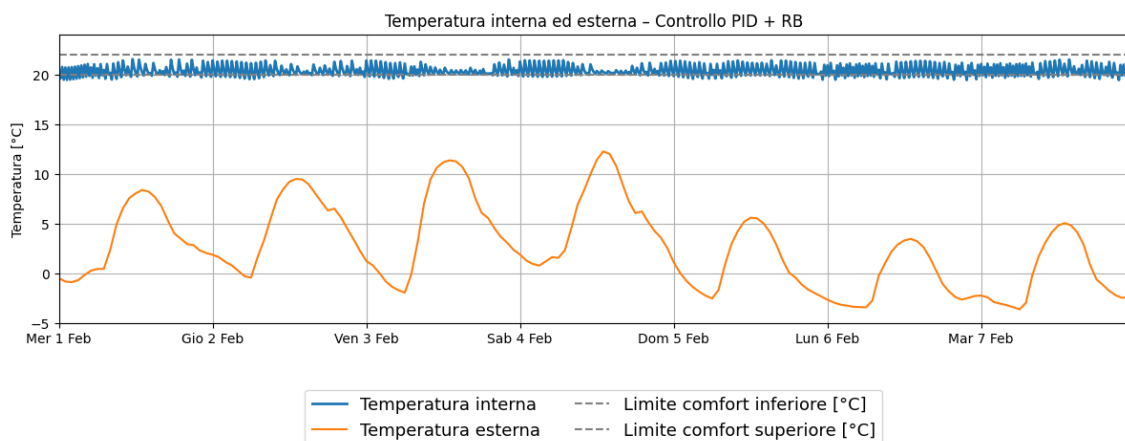


Figura 48: Andamento temperatura interna ed esterna con controllore PID + RB - caso studio scolastico

Dal punto di vista termico, il controllore tradizionale è in grado di mantenere la temperatura interna all'interno della banda di comfort prevista, con oscillazioni all'interno di questo range. Tuttavia, rispetto al caso residenziale, il controllo risulta meno stabile e meno lineare. L'andamento della temperatura interna mostra infatti un comportamento simile a quello osservato nel caso di transfer zero-shot, evidenziando una regolazione non ottimale del sistema di generazione.

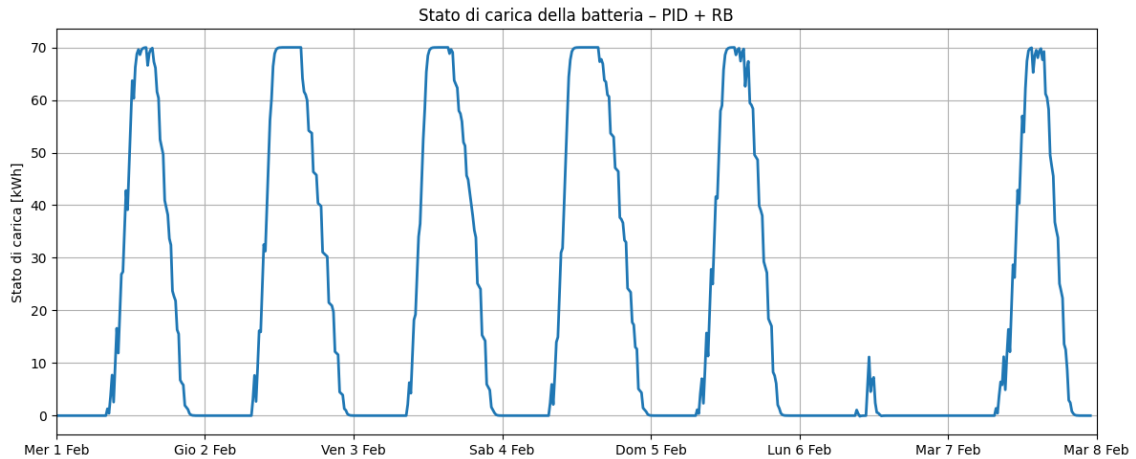


Figura 49: Andamento stato di carica della batteria con controllo PID + RB - caso studio scolastico

L'andamento dello stato di carica della batteria evidenzia anche in questo caso cicli regolari di carica e scarica, guidati principalmente dalla produzione fotovoltaica. Durante le ore diurne la batteria viene caricata fino alla capacità massima, mentre nelle ore successive viene scaricata per coprire il fabbisogno dell'edificio. Anche in questo caso, la gestione dell'accumulo è determinata da regole fisse imposte e non da una logica di ottimizzazione.

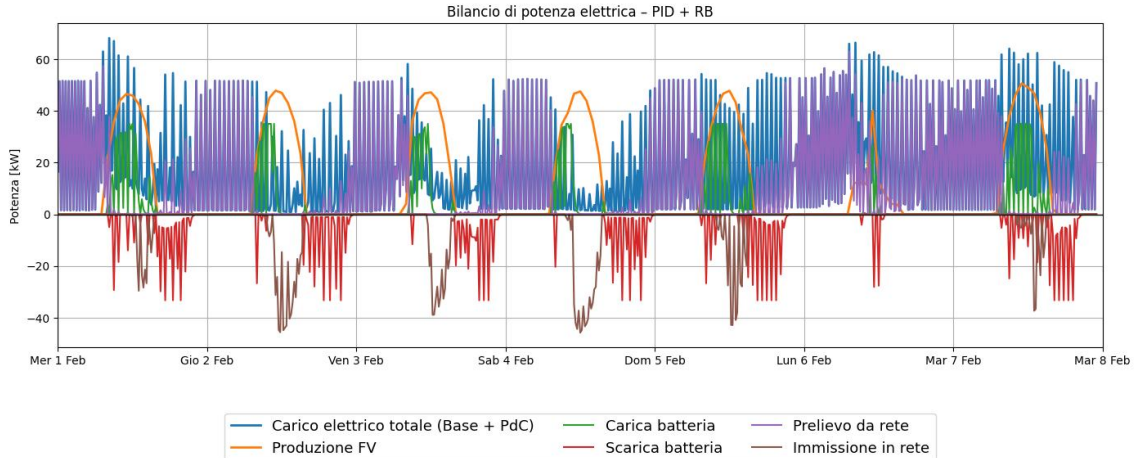


Figura 50: Bilancio della potenza elettrica con controllo PID + RB - caso studio scolastico

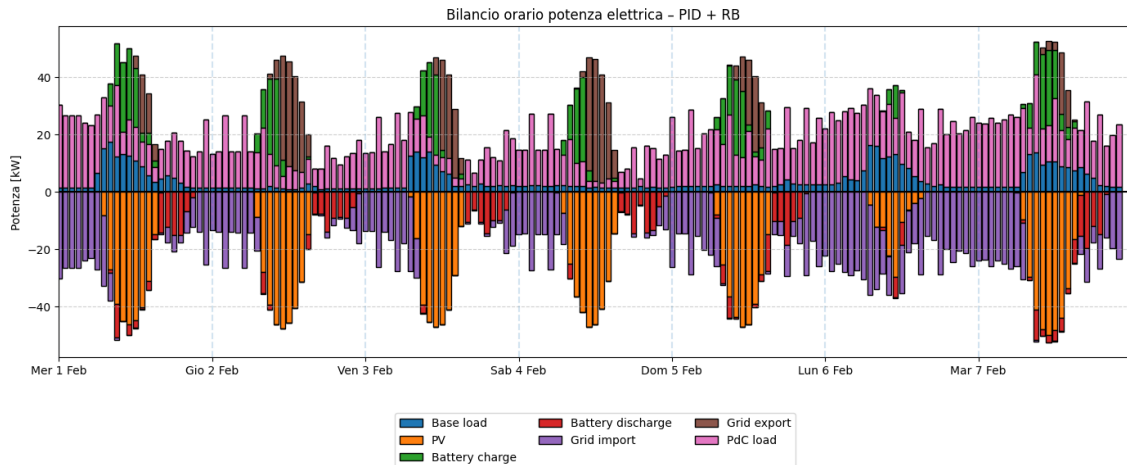


Figura 51: Bilancio orario della potenza elettrica con controllo PID + RB - caso studio scolastico

Il bilancio energetico evidenzia chiaramente la presenza di un carico concentrato nelle ore centrali della giornata, tipico dell'edificio scolastico. Tuttavia, il controllo del sistema di generazione non risulta ottimale, come si può osservare dalle oscillazioni della potenza lungo l'intero arco della giornata.

Dal punto di vista energetico, i risultati mostrano un'energia importata dalla rete pari a circa 1917 kWh e un'energia esportata pari a circa 570 kWh, valori entrambi elevati e indicativi di una gestione non ottimale dei flussi energetici rispetto ai controllori avanzati. L'autoconsumo diretto si attesta intorno al 40%, evidenziando come una parte significativa dell'energia prodotta non venga utilizzata immediatamente per soddisfare i carichi dell'edificio. Questo comportamento è legato alla mancanza di coordinamento tra produzione, accumulo e domanda energetica. Dal punto di vista economico i valori sono molto simili al caso DRL, abbiamo infatti il costo di acquisto dell'energia è pari a circa 394 €, a fronte di un ricavo dalla vendita di circa 28 €, per un costo netto complessivo di circa 365 €.

Quindi, anche nel caso scolastico, il controllo PID + rule-based si dimostra una soluzione semplice e robusta dal punto di vista del comfort, ma poco efficiente nella gestione energetica. L'assenza di una capacità di previsione o di apprendimento limita infatti la possibilità di adattare il comportamento del sistema alle specifiche dinamiche dell'edificio, risultando in prestazioni inferiori rispetto alle strategie basate su MPC e DRL.

### 3.2.6. Confronto delle strategie di controllo

Per analizzare le prestazioni delle diverse strategie applicate all'edificio scolastico, la Tabella 5) riporta i principali indicatori energetici ed economici.

Indicatore	MPC	DRL	PID+RL	Transfer zero-shot	Transfer fine-tuning
Comfort rispettato	Sì	Sì	Sì	Sì	Sì
Energia importata [kWh]	1617	1880	1917	1841	1854
Energia esportata [kWh]	353	537	570	404	472
Consumi PdC [kWh]	2465	2530	2526	2512	2623
Produzione FV [kWh]	1882	1882	1882	1882	1882
Autoconsumo diretto [%]	56,6	49,6	39,9	55,9	53,1
Energia batteria carica [kWh]	464	411	561	427	410
Energia batteria scarica [kWh]	419	371	512	385	312
Costo netto [€]	309	361	365	360	364

Tabella 5: Confronto fra le diverse strategie di controllo - caso studio scolastico

Dall'analisi dei risultati emerge innanzitutto che tutte le strategie di controllo riescono a garantire il rispetto delle condizioni di comfort termico, mantenendo la temperatura interna all'interno della banda prevista. Le differenze tra i controllori riguardano quindi principalmente la gestione dei flussi energetici del sistema e gli scambi con la rete.

Il controllore MPC presenta i valori più bassi sia di energia importata sia di energia esportata. Questo comportamento è legato alla natura predittiva della strategia di controllo, che consente di pianificare nel tempo l'utilizzo delle diverse risorse energetiche disponibili. Grazie alla conoscenza delle condizioni future, l'MPC riesce infatti a coordinare in modo più efficace la produzione fotovoltaica, il sistema di accumulo e il fabbisogno energetico dell'edificio. Di conseguenza, una quota maggiore dell'energia prodotta localmente viene utilizzata direttamente per soddisfare i carichi interni, riducendo complessivamente gli scambi con la rete. Come si può notare, questo controllore presenta il valore di autoconsumo diretto maggiore. Tale comportamento si riflette anche sul piano economico: il costo netto risulta infatti inferiore rispetto alle altre strategie. Questa differenza è principalmente legata alla minore quantità di energia importata dalla rete, resa possibile da un utilizzo più efficace della produzione fotovoltaica e da una gestione più coordinata del sistema di accumulo. Inoltre, anche i

consumi della pompa di calore risultano leggermente inferiori rispetto agli altri controllori, contribuendo ulteriormente alla riduzione del costo complessivo dell'energia.

Le altre strategie mostrano invece valori più elevati sia di energia importata sia di energia esportata. Nel caso DRL, il controllore reagisce alle condizioni osservate istante per istante e non dispone di una pianificazione esplicita su orizzonte temporale futuro. Questo porta a una gestione meno ottimizzata della produzione fotovoltaica e dell'accumulo elettrico. Il controllore PID con logica rule-based, pur utilizzando in modo più intenso il sistema di accumulo, non riesce comunque a ridurre in modo significativo l'energia importata dalla rete, segno di una minore capacità di coordinare i flussi energetici complessivi del sistema.

Per quanto riguarda il Transfer Learning, entrambi gli scenari di trasferimento mostrano prestazioni complessivamente simili tra loro e comparabili con quelle ottenute dal controllore DRL addestrato direttamente sul dominio scolastico. In particolare, il caso zero-shot mostra come la politica appresa nel dominio residenziale riesce a operare in modo efficace anche nel dominio scolastico, mantenendo il comfort termico e un comportamento energetico complessivamente stabile. In alcune fasce orarie si osservano tuttavia oscillazioni leggermente più marcate della temperatura, riconducibili alle diverse dinamiche di carico e alla differente destinazione d'uso dell'edificio scolastico rispetto al caso residenziale. Tuttavia, il comportamento del sistema evidenzia una buona capacità di generalizzazione della politica appresa.

Il trasferimento tramite fine-tuning conferma la capacità dell'agente di adattare progressivamente la politica di controllo alle nuove condizioni operative dell'edificio scolastico. Rispetto allo scenario zero-shot, il controllo della temperatura risulta più stabile e coerente con i vincoli di comfort imposti. Tuttavia, rimangono alcune criticità nella gestione del sistema di accumulo. Questo comportamento suggerisce che, sebbene il fine-tuning migliori la stabilità complessiva del controllo, la politica non ha ancora appreso una strategia pienamente efficace di gestione dell'accumulo elettrico.

I risultati evidenziano come il controllore MPC rappresenti la soluzione più efficiente dal punto di vista energetico ed economico grazie alla capacità di pianificare nel tempo l'utilizzo delle risorse disponibili. Il DRL e le strategie di Transfer Learning mostrano comunque prestazioni comparabili e riescono a mantenere il comfort termico dell'edificio, dimostrando una buona capacità di adattamento a domini caratterizzati da

dinamiche energetiche simili. Rimangono tuttavia alcune differenze nella gestione del sistema di accumulo, che suggeriscono possibili margini di miglioramento nelle strategie basate su apprendimento.

### **3.3. *Caso studio industriale***

Il terzo caso studio analizzato riguarda un edificio ad uso industriale, rappresentato da capannone produttivo di media dimensione, localizzato al Nord Italia. Rispetto ai casi precedenti analizzati, questo scenario introduce un contesto energetico significativamente diverso, caratterizzato da volumi maggiori e da carichi termici interni legati alle attività produttive. L'analisi di questo edificio permette quindi di valutare il comportamento delle strategie di controllo in presenza di dinamiche energetiche più complesse e su una scala di potenze nettamente superiori rispetto ai casi residenziale e scolastico.

L'edificio considerato presenta una pianta rettangolare di 60 m di lunghezza, 20 m di larghezza e un'altezza interna pari a 6 m. Il volume interno netto risulta pari a 5822 m<sup>3</sup> valore superiore rispetto ai casi studio precedenti e che contribuisce ad aumentare l'inerzia termica complessiva dell'edificio.

Dal punto di vista costruttivo, l'involucro edilizio è realizzato mediante pannelli prefabbricati in calcestruzzo, una soluzione comunemente adottata negli edifici industriali per la robustezza e velocità di realizzazione. Le proprietà termo-fisiche dei materiali, quali la conducibilità termica, densità, calore specifico e spessori dei diversi strati, sono state ricavate dalla scheda tecnica fornita dal produttore. A partire da questi parametri è stata determinata la trasmittanza termica complessiva dell'elemento opaco, pari a:

$$U_{opaco} = 0,16 \frac{W}{m^2K}$$

Nel caso considerato non sono state introdotte superfici vetrate, ipotesi che rappresentativa di molti edifici industriali destinati ad attività produttive nelle quali l'illuminazione naturale è risulta limitata.

Per quanto riguarda la ventilazione, negli edifici industriali il ricambio d'aria non è determinato esclusivamente da criteri di comfort, come avviene degli edifici residenziali, ma dipende principalmente dalle esigenze legate alla sicurezza e alla qualità dell'aria nei luoghi di lavoro. Il principale riferimento normativo è rappresentato Decreto Legislativo

81/2008, che richiede di garantire un'adeguata aerazione naturale o artificiale nei luoghi di lavoro.

In assenza di una caratterizzazione dettagliata del processo produttivo, nel lavoro si adotta un approccio semplificato, assumendo un numero di ricambi d'aria costante pari a 6 vol/h, rappresentativo di un ambiente industriale con ventilazione forzata moderata. Tale valore risulta coerente con quanto riportato nell'ASHRAE Handbook, HVAC Applications [15], che per capannoni industriali e ambienti produttivi generici suggerisce tassi di ventilazione tipicamente compresi tra 3 e 10 ricambi d'aria orari.

Consideriamo un volume interno dell'edificio, il tasso di ventilazione adottato porta a una portata d'aria pari a circa 9,7 m<sup>3</sup>/s, corrispondente a un coefficiente di perdita termica per ventilazione pari a:

$$H_{ve} = \frac{Q_{ve}}{1000} \cdot \rho \cdot c_p = 11,7 \frac{kW}{K} \quad (50)$$

Un'altra caratteristica distintiva degli edifici industriali è rappresentata dalla presenza di guadagni termici interni rilevanti, generalmente associati ai macchinari e ai processi produttivi. In questo contesto il contributo termico dovuto alla presenza degli occupanti risulta trascurabile rispetto alla potenza dissipata dalle apparecchiature.

Nel caso studio considerato si ipotizza la presenza di vasche di processo con una potenza termica installata pari a 200 kW. Tuttavia, sola una parte di questa potenza viene effettivamente ceduta all'ambiente interno. Assumendo un fattore di rilascio prudenziale pari a 0,3, il contributo termico interno risulta pari a:

$$\dot{Q}_{int} = 60 \text{ kW} \quad (51)$$

Questo valore viene considerato costante nel tempo è introdotto nel bilancio energetico dell'edificio.

Applicando anche in questo caso le relazioni scritte del sottocapitolo 2.4. per la determinazione del carico termico invernale di progetto, avendo a disposizione le perdite per trasmissione e ventilazione, otteniamo una potenza termica di progetto pari a:

$$\tilde{\Phi}_{P,finale}^* \approx 430 \text{ kW} \quad (52)$$

Questo valore rappresenta la potenza termica necessaria per il dimensionamento del sistema di generazione costituito dalla pompa di calore a servizio della batteria dell'Unità di Trattamento Aria (UTA), soluzione comunemente adottata negli edifici industriali caratterizzato da grandi volumi e da sistemi la ventilazione meccanica.

Per le simulazioni è stato assunto un set-point termico pari a 18°C, con una banda di comfort compresa tra 18°C e 20°C. Questa scelta riflette le condizioni tipicamente adottate negli ambienti industriali, dove i requisiti di comfort risultano generalmente meno restrittivi rispetto agli edifici residenziali o scolastici e sono spesso subordinati alle esigenze del processo produttivo.

Per questo caso studio sono stati fatti due confronti differenti rispetto le dimensioni dell'impianto fotovoltaico. Nel primo caso è stata adottata una configurazione più bilanciata tra produzione e domanda elettrica, invece per il secondo rimane una configurazione tra surplus e dimensionamento della batteria coerente con i casi precedenti.

Di seguito i parametri relativi ai modelli energetici dell'edificio in esame rispetto a quello residenziale per confrontare la differente scala utilizzata per il Transfer.

<b>Paramenti</b>	<b>Caso residenziale</b>	<b>Caso Industriale 1</b>	<b>Caso Industriale 2</b>
Volume [m <sup>3</sup> ]	360	7200	7200
Area calpestabile [m <sup>2</sup> ]	120	1200	1200
Dimensione PdC [kW]	9,5	430	430
Dimensione PV [kW]	10	600	1000
Dimensione accumulo [kWh]	18	500	500

*Tabella 6: Paramenti a confronto del modello energetico - caso residenziale e caso industriale*

### 3.3.1. Caso industriale 1

Il primo caso studio industriale analizzato è caratterizzato da una configurazione energetica in cui la produzione fotovoltaica risulta complessivamente bilanciata rispetto al fabbisogno dell'edificio. In questo scenario, una quota significativa dell'energia prodotta viene direttamente autoconsumata dai carichi dell'impianto produttivo, riducendo la disponibilità di surplus per il sistema di accumulo.

L'obiettivo di questo caso studio è quindi valutare il comportamento delle diverse strategie di controllo in condizioni operative in cui il margine di ottimizzazione legato alla gestione dell'accumulo risulta più limitato, permettendo di analizzare l'efficacia dei controllori in presenza di un equilibrio tra produzione e domanda energetica.

#### 3.3.1.1. Controllo con MPC

La strategia di controllo predittivo utilizzata nei casi precedenti è stata adottata anche per questo caso studio industriale. La struttura dell'MPC, l'orizzonte di predizione di 24 ore, il passo temporale di discretizzazione di 15 minuti rimangono invariati, così come la formulazione della funzione obiettivo, orientata alla minimizzazione del costo netto e del comfort all'interno dell'impianto. La maggiore differenza, rispetto i casi precedenti, è rappresentata dalla scala energetica del sistema che risulta superiore, di conseguenza, avremo delle potenze molto più elevate e dal diverso dimensionamento dell'impianto fotovoltaico,

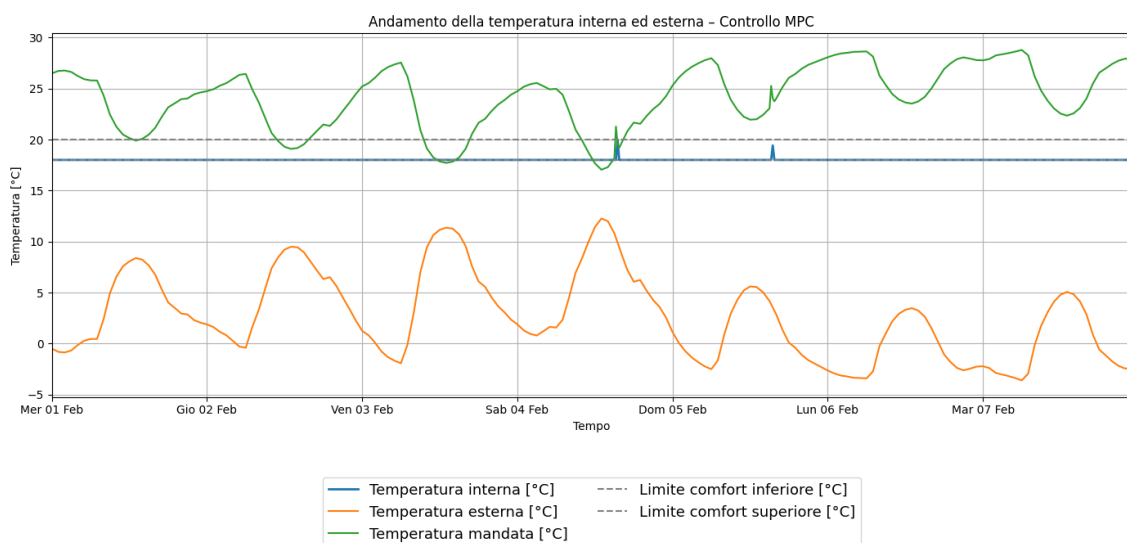


Figura 52: Andamento temperatura interna, esterna e di mandata UTA con controllore MPC - caso studio industriale 1

L'andamento della temperatura interna mostra come il controllore riesce a mantenere la temperatura dell'ambiente fissa sul valore di set-point per l'intero periodo di simulazione. A differenza dei casi precedenti, in cui la temperatura tendeva a sfruttare l'intera banda di comfort per ottimizzare il costo energetico, nel caso industriale il comportamento risulta molto più stabile. Questo è dovuto principalmente alla presenza di elevati guadagni interni associati ai processi produttivi, che contribuiscono in modo significativo al bilancio termico dell'edificio.

Nel grafico è inoltre riportata la temperatura di mandata dell'UTA, che rappresenta la variabile di controllo attraverso cui il sistema modula l'apporto termico all'ambiente. Questa variabile segue un andamento modulante nel tempo, adattandosi alle variazioni delle condizioni climatiche esterne e al fabbisogno termico dell'edificio.

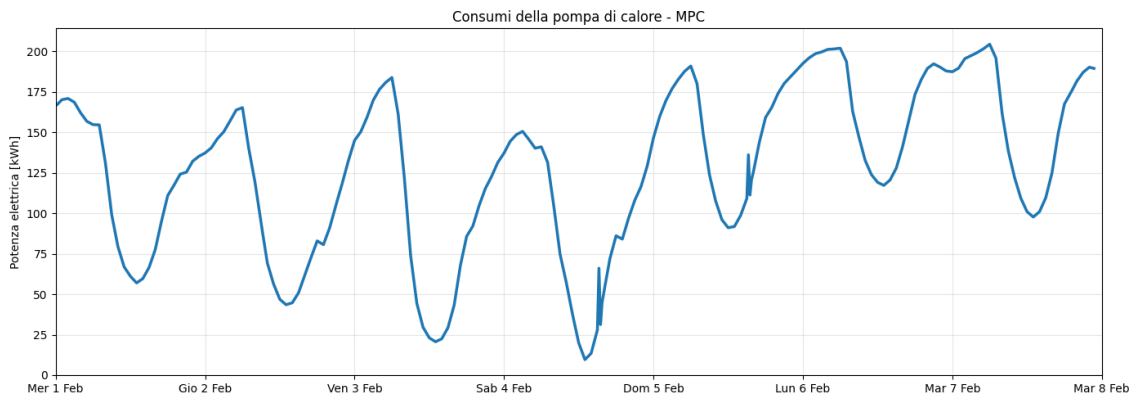


Figura 53: Andamento potenza elettrica consumata dalla pompa di calore con controllo MPC- caso studio industriale

Osservando la potenza della pompa di calore, possiamo notare come i valori siano decisamente più elevati rispetto ai casi precedenti, con picchi che superano i 100 kW. Anche in questo caso si osserva una modulazione coerente con le condizioni climatiche e con l'andamento della temperatura esterna.

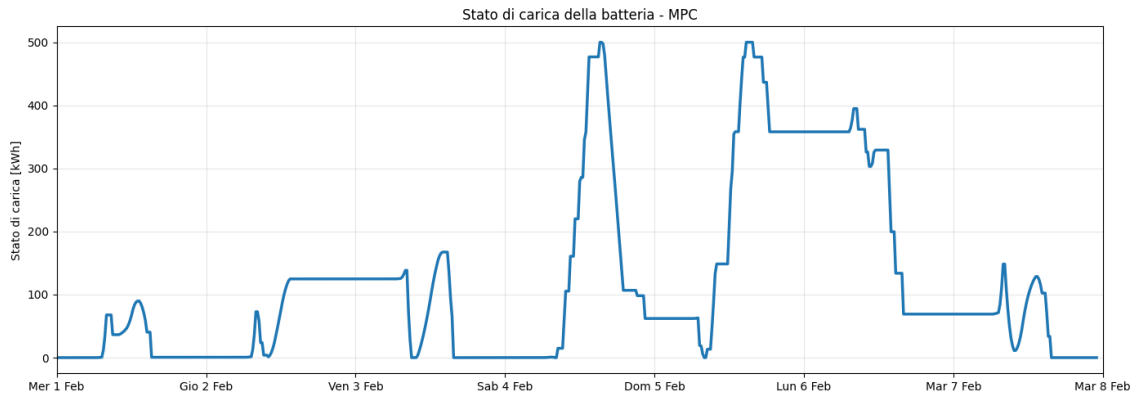


Figura 54: Andamento stato di carica della batteria con controllo MPC - caso studio industriale1

La Figura 54) riporta invece lo stato di carica della batteria. A differenza dei casi residenziale e scolastico, il sistema di accumulo mostra nei giorni lavorativi dei livelli di carica relativamente contenuti. Questo comportamento è legato alla presenza del carico di base industriale, che nelle ore di funzionamento dell'impianto produttivo assorbe gran parte dell'energia dall'impianto fotovoltaico.

Nei giorni in cui l'attività industriale è ridotta o assente, una quota maggiore della produzione fotovoltaica diventa disponibile per la ricarica della batteria. Questo fenomeno è particolarmente evidente nei giorni 4 e 5 febbraio, che corrispondono al fine settimana. Inoltre, l'MPC, anche in questo caso, tende a preservare parte dell'energia accumulata per utilizzarla nelle ore successive in cui la produzione fotovoltaica è minore, come avviene tra il 5 e 6 febbraio.

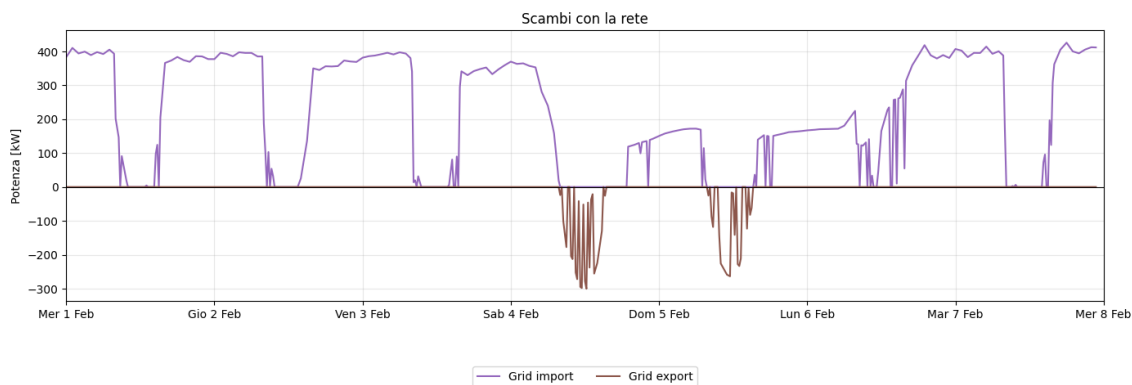


Figura 55: Andamento degli import ed export dalla rete con controllo MPC - caso studio industriale1

Questo comportamento è visibile anche nel grafico degli scambi con la rete, dove si osserva una riduzione del prelievo medio della rete nonostante il giorno 6 sia un giorno lavorativo.

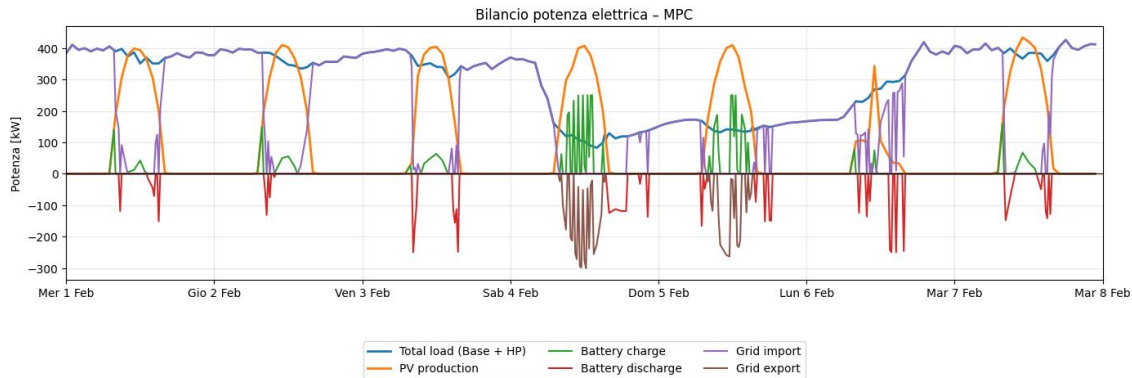


Figura 56: Bilancio della potenza elettrica con controllo MPC - caso studio industriale1

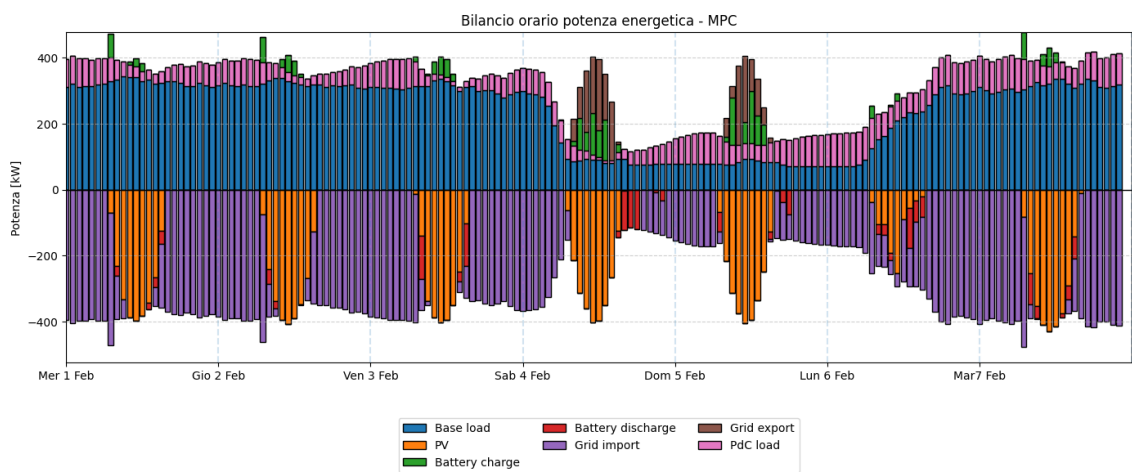


Figura 57: Bilancio orario della potenza elettrica con controllo MPC - caso studio industriale1

Dal bilancio energetico è possibile osservare le interazioni tra i diversi componenti dell'impianto. La produzione fotovoltaica raggiunge valori superiori ai 400 kW nelle ore centrali della giornata e contribuisce in modo significativo alla copertura del fabbisogno dell'edificio.

Nel periodo considerato, l'energia complessivamente importata dalla rete risulta pari a circa 36226 kWh e quella esportata ammonta a 1863 kWh. Con una produzione fotovoltaica totale di 16130 kWh, l'edificio riesce a utilizzare direttamente circa il 77% dell'energia generata. Economicamente, invece, otteniamo un costo netto pari a 7596 €, valore coerente con la scala significativamente superiore del sistema industriale rispetto ai casi studio precedenti.

Il controllo predittivo dimostra una buona capacità di coordinare i diversi componenti del sistema energetico dell'edificio. Nonostante il diverso dimensionamento dell'impianto fotovoltaico e del sistema di accumulo rispetto ai casi precedenti, il

controllore riesce a garantire il mantenimento delle condizioni di comfort e una gestione economicamente efficiente degli scambi energetici con la rete.

### 3.3.1.2. Controllo con Deep Reinforcement Learning

Anche per il caso studio industriale è stata applicata la strategia di controllo basata su DRL descritta nei capitoli precedenti. La struttura dell'agente e la formulazione generale della funzione di ricompensa rimangono invariate.

Le principali differenze riguardano la scala energetica del sistema, l'edificio industriale presenta infatti carichi termici ed elettrici significativamente più elevati rispetto ai casi residenziale e scolastico. Per questo motivo è stato necessario riscalarne opportunamente i pesi della funzione di ricompensa, adattandoli alla nuova scala energetica del sistema.

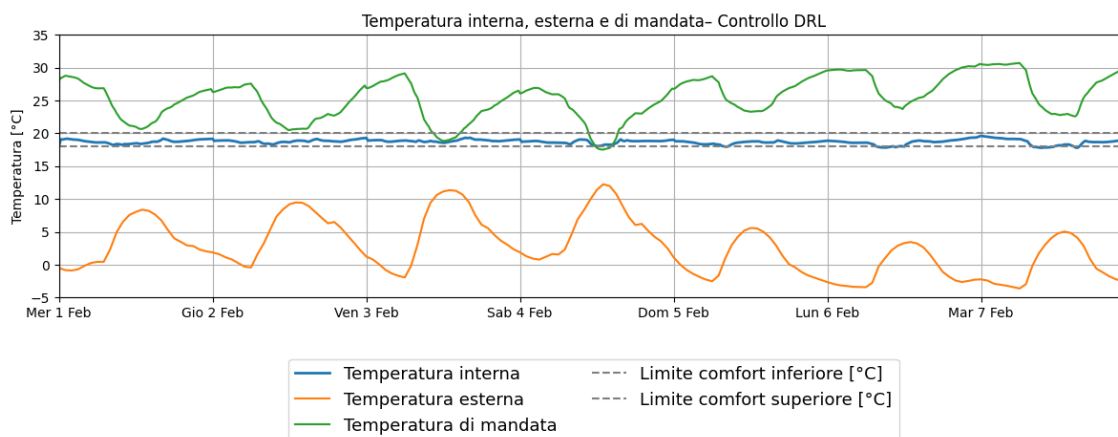


Figura 58: Andamento temperatura interna, esterna e di mandata UTA con controllore DRL - caso studio industriale I

Dal grafico si osserva come l'agente riesca a mantenere la temperatura interna sempre all'interno della banda di comfort compresa tra i 18°C e 20°C, con oscillazioni contenute al suo interno. Questo comportamento evidenzia la capacità del controllore di fruttare l'inerzia dell'edificio, permettendo alla temperatura di variare gradualmente senza uscire dai limiti di comfort.

Nella stessa figura è inoltre riportata anche la temperatura di mandata dell'UTA, la quale mostra un andamento modulante del tempo, adattandosi sia alle variazioni delle condizioni climatiche esterne sia alle esigenze termiche dell'ambiente interno. In particolare, nelle giornate caratterizzate da temperatura esterne più basse si osserva un

incremento medio della temperatura di mandata, necessario per compensare le maggiori perdite termiche dell'edificio.

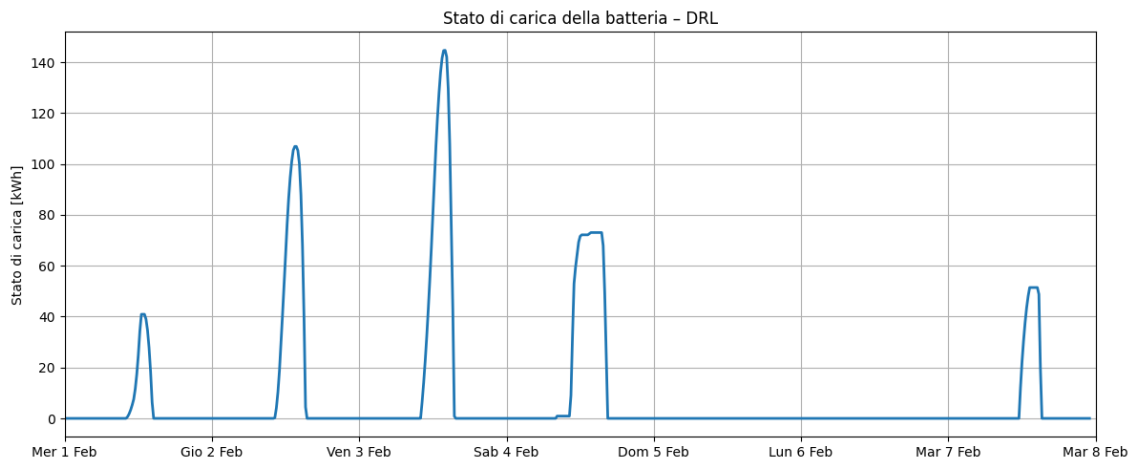


Figura 59: Andamento stato di carica della batteria con controllo DRL - caso studio industriale 1

A differenza dei casi studio precedenti, in cui la batteria viene utilizzata in modo più intensivo grazie alla frequente presenza di surplus fotovoltaico, nel contesto industriale il sistema di accumulo svolge un ruolo marginale. Il livello di carica rimane marginale generalmente contenuto e si osservano cicli di carica e scarica meno pronunciati. Di conseguenza, la disponibilità di energia in eccesso per la ricarica del sistema di accumulo risulta più limitata rispetto ai casi studio precedenti.

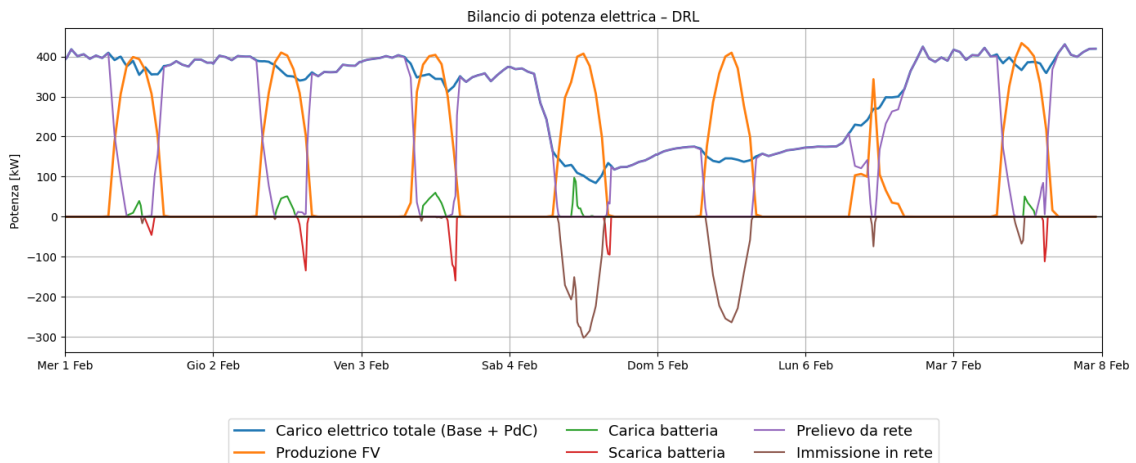


Figura 60: Bilancio della potenza elettrica con controllo DRL - caso studio industriale 1

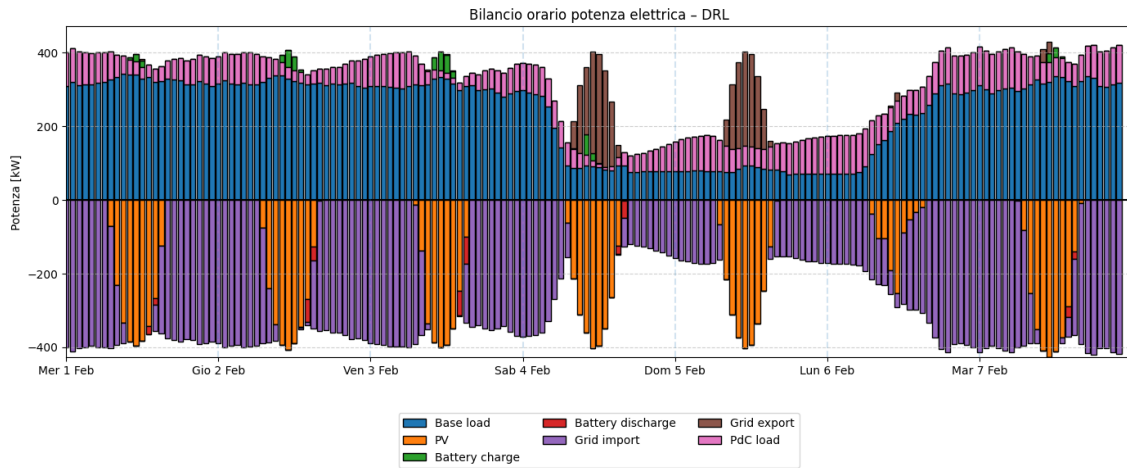


Figura 61: Bilancio orario della potenza elettrica con controllo DRL - caso studio industriale 1

Dal bilancio energetico si osserva come la produzione fotovoltaica riesca a coprire gran parte dei carichi dell'impianto durante i giorni lavorativi. Nel fine settimana, quando l'attività produttiva è ridotta, si genera invece una maggiore disponibilità di energia che, nel caso del controllo DRL, viene principalmente immessa in rete anziché essere utilizzata per la ricarica della batteria, a differenza di quanto osservato con il controllore predittivo.

Complessivamente, il carico elettrico totale dell'edificio, comprensivo sia del carico base dell'impianto sia del consumo della pompa di calore, risulta essere stabile nel tempo, coerente con il profilo operativo dell'edificio industriale. I valori di potenza risultano naturalmente più elevati rispetto ai casi studio precedenti, ma rimangono coerenti con la destinazione d'uso e con la scala energetica dell'impianto analizzato.

Abbiamo un'energia complessiva immessa in rete pari a circa 37976 kWh, mentre l'energia esportata è circa 2899 kWh. A fronte di una produzione fotovoltaica totale di circa 16130 kWh, l'edificio riesce ad autoconsumare direttamente il 79.31% dell'energia prodotta. Dal punto di vista economico, il costo complessivo per l'acquisto di energia elettrica dalla rete risulta pari a 8102€, mentre il ricavo dalla vendita è di 145€, determinando un costo netto totale di 7957€ nel periodo analizzato. Questi valori risultano nettamente più elevati rispetto ai casi studio precedenti, ma sono coerenti con la maggiore scala energetica dell'edificio industriale e con i livelli di potenza coinvolti nel sistema.

### 3.3.1.3. Applicazione del Transfer Learning

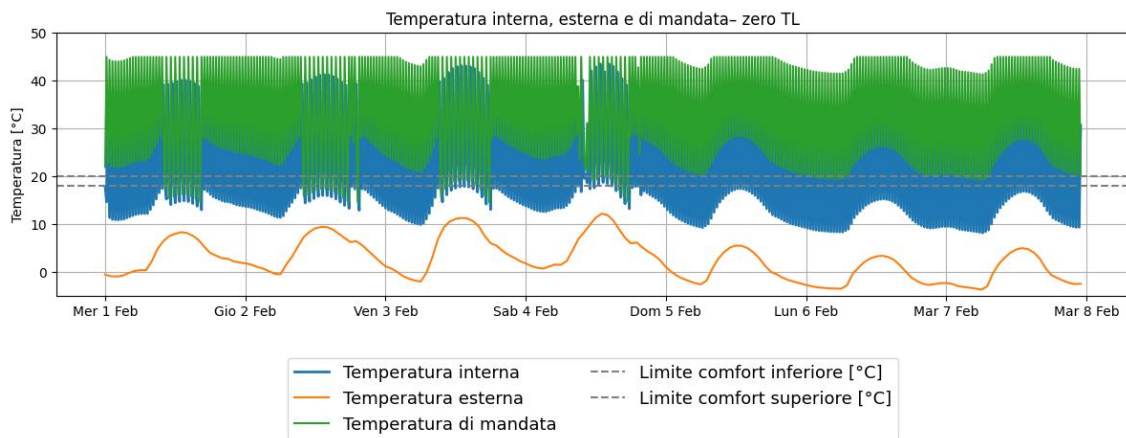
Il caso industriale presenta una scala energetica significativamente superiore rispetto al caso residenziale, con carichi termici ed elettrici molto più elevati e dinamiche operative

legate anche al dimensionamento dell'impianto fotovoltaico differente. L'obiettivo di questa analisi è valutare la capacità dell'agente basato su DRL, addestrato nel caso residenziale, di adattarsi ad un dominio con caratteristiche fisiche ed energetiche molto diverse rispetto a quelle incontrate durante la fase di addestramento.

Analogamente a quanto fatto nel caso scolastico, l'analisi di trasferimento è stata condotta considerando due differenti modalità di applicazione della politica di controllo. In un primo scenario viene analizzato il transfer zero-shot, nel quale l'agente addestrato nel caso residenziale viene applicato direttamente all'edificio industriale. Successivamente viene considerato il transfer fine-tuning, in cui la politica appresa viene ulteriormente riaddestrata nel nuovo ambiente di interazione, consentendo all'agente di adattarsi progressivamente alle diverse dinamiche energetiche del sistema industriale.

### *Transfer zero-shot*

In questo scenario viene applicata la politica appresa nel caso residenziale direttamente all'edificio industriale.



*Figura 62: Andamento temperatura interna, esterna e di mandata UTA con TL zero-shot - caso studio industriale 1*

La figura evidenzia chiaramente i limiti di questo trasferimento diretto. La temperatura interna presenta oscillazioni molto ampie e frequenti, con valori che si discostano spesso dalla banda di comfort prevista. Questo comportamento indica che la politica appresa nel caso residenziale non è in grado di gestire le dinamiche termiche dell'edificio industriale. La causa principale è la forte differenza tra i due domini, infatti, nel caso industriale la scala dei carichi termici ed elettrici significativamente elevati. Di conseguenza, le azioni

apprese durante la fase di training non risultano adeguate per mantenere un controllo stabile della temperatura.

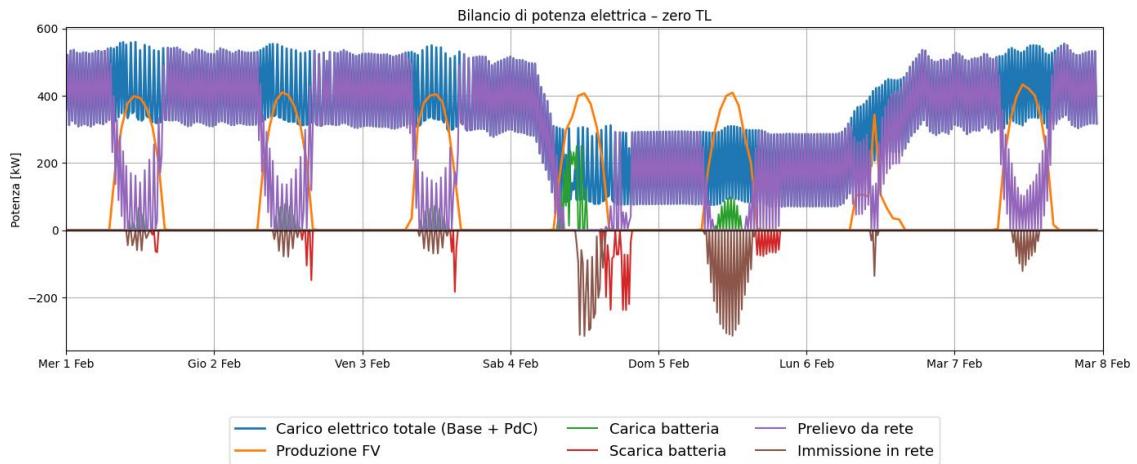


Figura 63: Bilancio della potenza elettrica con TL zero-shot - caso studio industriale 1

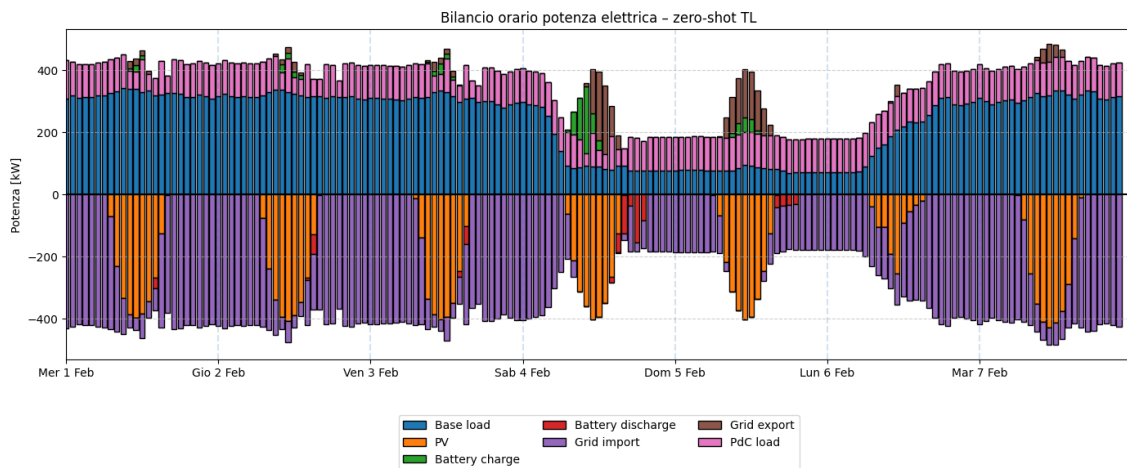


Figura 64: Bilancio orario della potenza elettrica con TL zero-shot - caso studio industriale 1

La stessa dinamica si osserva anche nel bilancio energetico dell'edificio. La potenza associata alla pompa di calore e agli scambi con la rete mostrano variazioni molto marcate nel tempo, quindi un'azione di controllo instabile e poco coerente con il fabbisogno energetico dell'edificio.

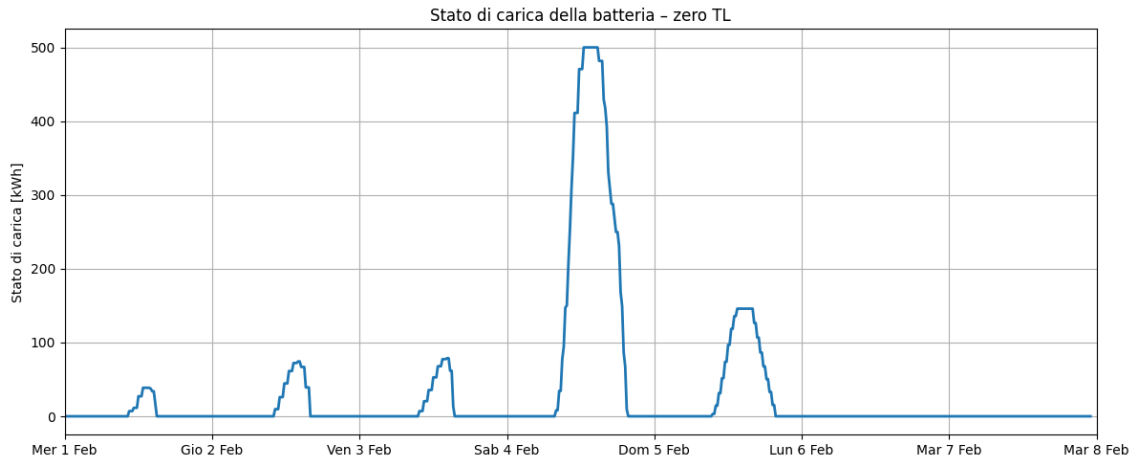


Figura 65: Andamento stato di carica della batteria con TL zero-shot - caso studio industriale 1

L'andamento dello stato di carica della batteria mostra invece comportamento relativamente più coerente. Il sistema di accumulo viene caricato principalmente nelle ore centrali della giornata, quando la produzione fotovoltaica supera temporaneamente il fabbisogno dell'edificio. Tuttavia, anche in questo caso l'utilizzo della batteria rimane limitato, poiché la maggior parte dell'energia prodotta dall'impianto fotovoltaico viene direttamente assorbita dai carichi dell'impianto industriale. Una differenza rispetto al caso DRL addestrato direttamente sul dominio industriale si osserva durante il fine settimana, dove la batteria tende a ricaricarsi maggiormente invece di esportare l'energia in rete.

In questo caso, nel periodo di simulazione l'energia complessivamente importata dalla rete risulta pari a circa 42485 kWh, mentre l'energia esportata ammonta a circa 2071 kWh. La produzione fotovoltaica totale risulta sempre uguale a 16130 kWh che permette un autoconsumo diretto di 81,7%. Questi valori portano ad un costo di acquisto dalla rete di 9147€ e di vendita in rete di 104€, per un costo netto totale di 9044€. Tutto questo conferma che utilizzando questa politica di controllo il peggioramento delle prestazioni energetiche ed economiche.

Questo risultato è coerente con quanto riportato in letteratura sul transfer learning applicato al DRL, dove il trasferimento zero-shot tra domini caratterizzati da dinamiche molto differenti possa portare a prestazioni instabili o sub-ottimali, soprattutto quando cambiano significativamente la scala delle variabili di stato e delle azioni [22]. In tali condizioni, una fase di adattamento tramite fine-tuning risulta generalmente necessaria per permettere alla politica di controllo di adattarsi alle nuove dinamiche del sistema.

### *Transfer fine-tuning*

In questo secondo scenario di trasferimento, la politica di controllo appresa nel caso residenziale viene ulteriormente riaddestrata nel dominio industriale attraverso una fase di fine-tuning. L'obiettivo è consentire all'agente di adattare progressivamente i parametri della politica alle nuove dinamiche energetiche del sistema.

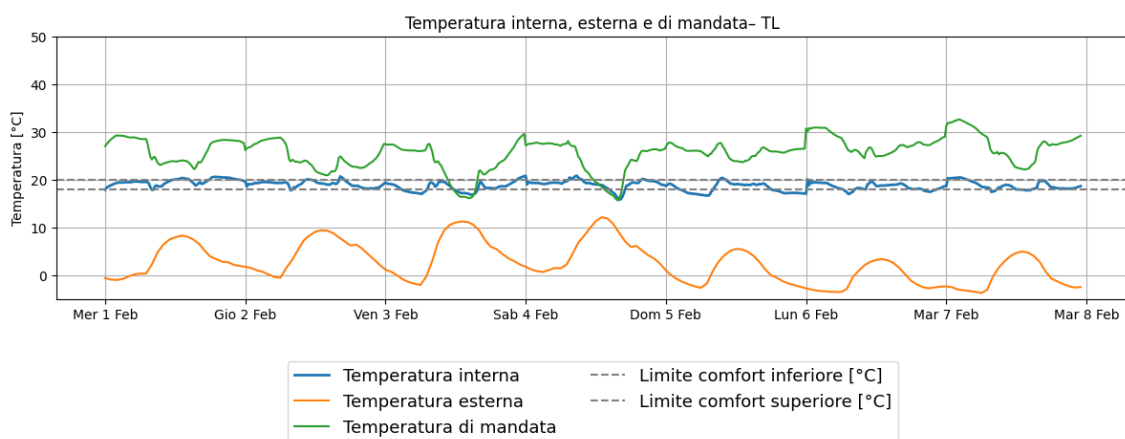


Figura 66: Andamento temperatura interna, esterna e di mandata UTA con TL fine-tuning - caso studio industriale I

L'andamento della temperatura interna evidenzia un miglioramento significativo rispetto allo scenario zero-shot. Le oscillazioni risultano infatti più contenute e, per gran parte del periodo di simulazione, la temperatura rimane all'interno della banda di comfort compresa tra i 18° e i 20°C. In alcuni intervalli si osservano comunque lievi superamenti dei limiti di banda, indice del fatto che la politica non è ancora completamente ottimizzata per il nuovo dominio. Tuttavia, il controllo termico risulta decisamente più stabile e coerente con le dinamiche dell'edificio industriale.

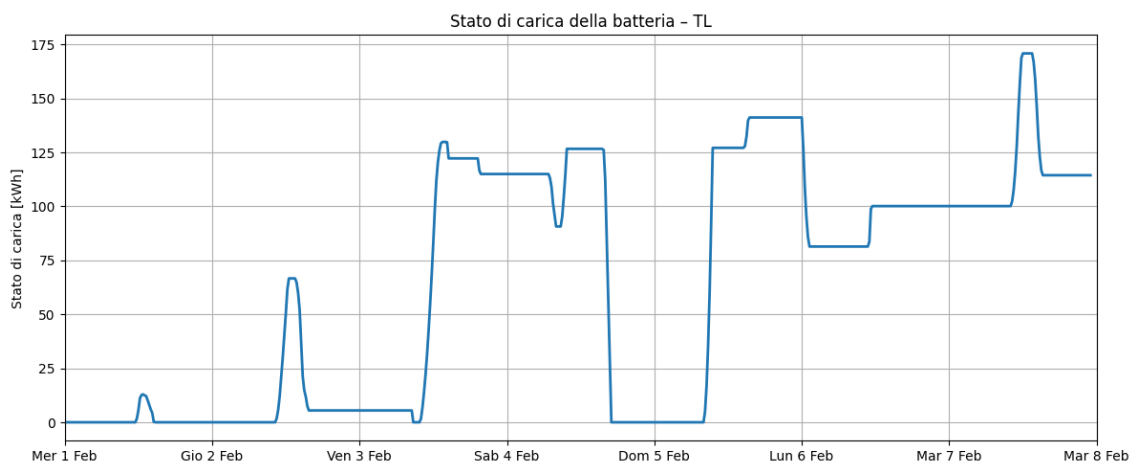


Figura 67: Andamento stato di carica della batteria con TL fine-tuning - caso studio industriale I

L'andamento dello stato di carica della batteria evidenzia un comportamento meno regolare. Il livello di carica tende infatti a rimanere elevato per lunghi intervalli di tempo, senza seguire una dinamica ciclica giornaliera ben definita tra fasi di carica e scarica. Questo comportamento suggerisce che, nonostante la fase di fine-tuning abbia migliorato la stabilità complessiva del sistema rispetto lo scenario zero-shot, la politica di controllo non ha ancora appreso una gestione ottimale del sistema di accumulo.

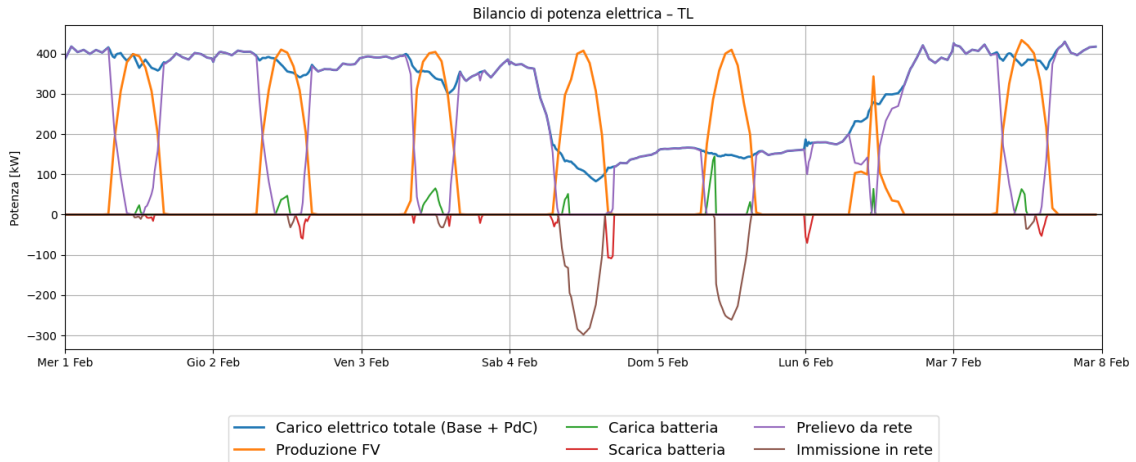


Figura 68: Bilancio della potenza elettrica con TL fine-tuning - caso studio industriale 1

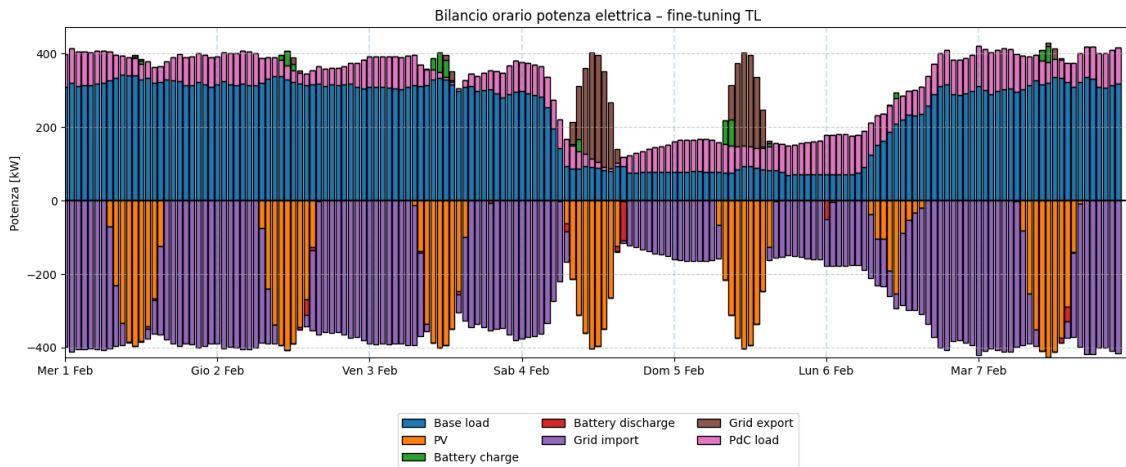


Figura 69: Bilancio orario della potenza elettrica con TL fine-tuning - caso studio industriale 1

Il bilancio elettrico mostra un comportamento complessivamente più regolare rispetto allo scenario precedente zero-shot e simile a quello osservato nel caso DRL addestrato direttamente sul dominio industriale. In particolare, nel fine settimana, quando la produzione fotovoltaica supera il fabbisogno dell'edificio, il sistema tende a esportare l'energia in rete piuttosto che accumularla nella batteria, comportamento coerente con quanto osservato anche nel caso DRL.

In totale, nel periodo di simulazione, l'energia importata dalla rete risulta pari a 38120 kWh, mentre l'energia esportata ammonta a circa 2738 kWh. La produzione fotovoltaica con i suoi 16130 kWh porta ad un autoconsumo diretto di 79,92%. Mentre dal punto di vista economico, il costo complessivo di acquisto dell'energia elettrica dalla rete risulta pari a 8142€, mentre dalla vendita si ricavano 137€, per un costo netto totale di 8005€ nel periodo analizzato.

Il confronto con lo scenario zero-shot evidenzia quindi un miglioramento delle prestazioni energetiche ed economiche. L'energia importata dalla rete si riduce infatti di circa 4300 kWh, mentre il costo netto diminuisce di 1000€. Questo risultato conferma come una fase di addestramento tramite fine-tuning permetta alla politica di controllo di adattarsi alle nuove dinamiche del sistema industriale, migliorando la stabilità del controllo della pompa di calore.

### 3.3.1.4. Controllo tradizionale PID + RB

Analizziamo infine il comportamento del controllore tradizionale nel caso dell'edificio industriale, caratterizzato da carichi elettrici elevati e relativamente continui nel tempo.

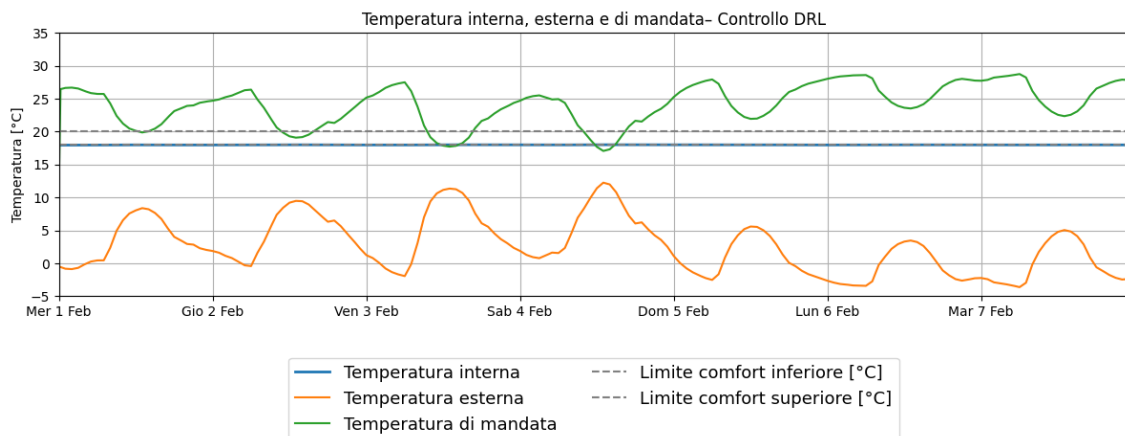


Figura 70: Andamento temperatura interna ed esterna con controllore PID +RB - caso studio industriale 1

Osservando l'andamento della temperatura interna, il controllore PID riesce a garantire una temperatura costante di set-point. Tuttavia, a differenza dei controlli avanzati, il sistema non sfrutta la banda di comfort, mantenendo un comportamento più rigido e meno flessibile.

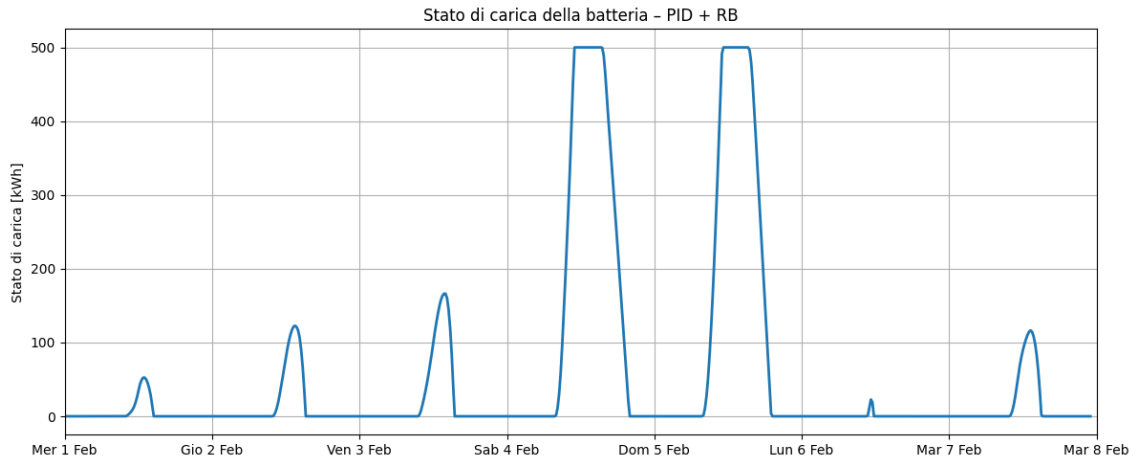


Figura 71: Andamento stato di carica della batteria con controllo PID + RB - caso studio industriale 1

L'andamento dello stato di carica della batteria evidenzia un utilizzo molto limitato del sistema di accumulo. Il comportamento risulta simile a quello osservato nel caso MPC, in quanto la batteria viene ricaricata principalmente in presenza di surplus fotovoltaico. Tuttavia, in assenza di strategia predittiva, il controllore non è in grado di utilizzare l'utilizzo della batteria in funzione delle condizioni future o delle fasce orarie, ma si limita a seguire le regole operative fissate.

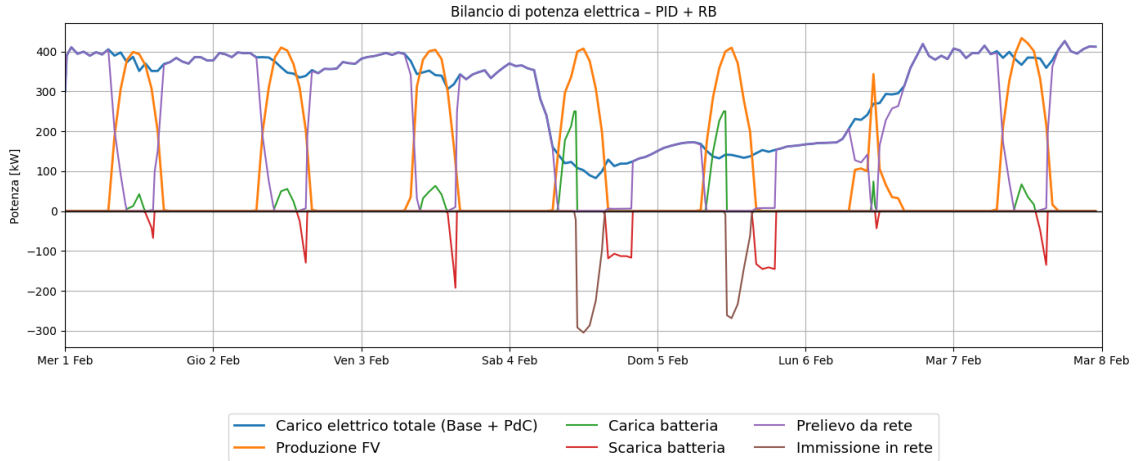


Figura 72: Bilancio della potenza elettrica con controllo PID + RB - caso studio industriale 1

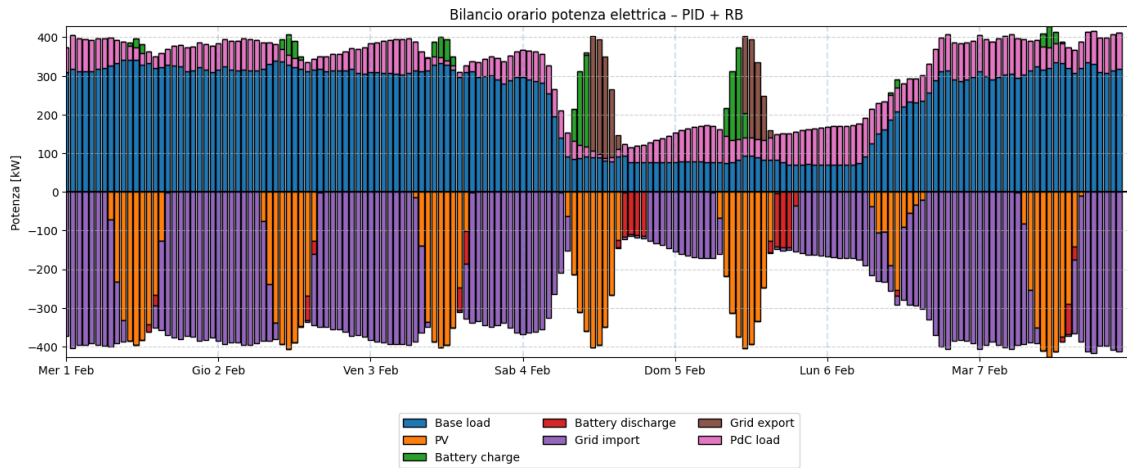


Figura 73: Bilancio orario della potenza elettrica con controllo PID + RB - caso studio industriale 1

Dal punto di vista energetico, il bilancio mostra chiaramente come una quota significativa della produzione fotovoltaica venga direttamente autoconsumata dai carichi dell'edificio. Questo è dovuto al profilo di consumo elevato e continuo tipico del caso industriale, che riduce la disponibilità di energia in eccesso. Sebbene il comportamento possa apparire simile a quello del controllo MPC in termini di distribuzione dei flussi energetici, il risultato è ottenuto in modo meno efficiente e meno flessibile, in quanto guidato unicamente da regole fisse e non da un processo di ottimizzazione.

I risultati evidenziano un'energia importata dalla rete pari a circa 36300 kWh e un'energia esportata pari a circa 1860 kWh, valori che riflettono una forte dipendenza dalla rete elettrica. Tuttavia, l'autoconsumo diretto risulta elevato, pari a circa 79%, proprio grazie alla buona coincidenza tra produzione e domanda energetica.

Dal punto di vista economico, il controllo PID + RB presenta un costo di acquisto dell'energia pari a circa 7713 €, a fronte di un ricavo dalla vendita di circa 93 €, per un costo netto complessivo di circa 7619 €.

Nel caso industriale, il controllo PID + RB mantiene la sua semplicità e robustezza, ma risulta ancora una volta limitato nella gestione ottimale del sistema energetico. In particolare, l'assenza di una strategia di pianificazione o di adattamento riduce la capacità di sfruttare il sistema di accumulo e di ottimizzare gli scambi con la rete, Sebbene il comportamento possa apparire simile a quello del controllo MPC, il risultato è ottenuto in modo meno efficiente e meno flessibile, in quanto guidato unicamente da regole fisse e non da una reale ottimizzazione del sistema.

### 3.3.1.5. Confronto fra strategie di controllo

Per valutare le prestazioni delle diverse strategie di controllo applicate all'edificio industriale, nella Tabella 7) sono riportati i principali indicatori energetici ed economici relativi ai tre controllori analizzati: MPC, DRL e PID con gestione rule-based della batteria.

Indicatore	MPC	DRL	PID+RL	Transfer zero-shot	Transfer fine-tuning
Comfort rispettato	Sì	Sì	Sì	No	Parzialmente
Energia importata [kWh]	36226	37966	36295	42485	38120
Energia esportata [kWh]	1863	2899	1863	2071	2738
Consumi PdC [kWh]	10669	11432	10660	16735	11631
Produzione FV [kWh]	16130	16130	16130	16130	16130
Autoconsumo diretto [%]	77	79	79	82	80
Energia batteria carica [kWh]	1851	439	1571	882	501
Energia batteria scarica [kWh]	1671	396	1393	796	343
Costo netto [€]	7596	7957	7713	9044	8005

Tabella 7: Confronto fra le diverse strategie di controllo - caso studio industriale 1

Dall'analisi dei risultati si osserva innanzitutto che tutte le strategie di controllo riescono a garantire il rispetto delle condizioni di comfort termico, mantenendo la temperatura interna all'interno della banda prevista. Le principali differenze tra i controllori emergono invece nella gestione energetica ed economia del sistema.

Considerando l'energia importata dalla rete, i valori risultano piuttosto simili tra le tre strategie. Il controllore MPC presenta il valore più basso, seguito dal PID con logica rule-based, mentre il DRL mostra un valore leggermente superiore. Anche dal punto di vista economico le differenze risultano principalmente tra il caso MPC e DRL, con una differenza di 361 €, mentre il caso PID si colloca in una posizione intermedia.

Questa differenza economica può essere ricondotta alla gestione del sistema di accumulo. Il controllore MPC mostra infatti un utilizzo più marcato della batteria, segno di una gestione più attiva dell'accumulo finalizzata a ottimizzare gli scambi energetici nel tempo. Il controllore DRL, al contrario, utilizza la batteria in modo più limitato, con una differenza di circa 1400 kWh rispetto all'MPC, privilegiando maggiormente l'utilizzo diretto dell'energia prodotta o l'interazione con la rete. La strategia PID + rule-based

mostra, dal confronto numerico, un comportamento più vicino a quello del controllore MPC.

La produzione fotovoltaica, identica in tutto il caso industriale, porta a valori di autoconsumo diretto molto simile tra le diverse strategie di controllo, compresi tra 77% e 79%.

Questi risultati evidenziano come, nel caso industriale analizzato, una quota significativa della produzione fotovoltaica venga direttamente autoconsumata. Di conseguenza, il ruolo del sistema di accumulo risulta meno determinante rispetto a quanto osservato nei casi residenziale e scolastico, dove la presenza di surplus energetico rende l'utilizzo della batteria più strategico per la gestione dei flussi energetici.

L'analisi complessiva dei due scenari evidenzia differenze significative nelle prestazioni del sistema. Nel caso zero-shot, l'applicazione diretta della politica appresa nel dominio residenziale non consente di gestire correttamente le dinamiche energetiche dell'edificio industriale, caratterizzato da carichi elettrici e termici più elevati. Questo si riflette in un comportamento instabile del sistema, con forti oscillazioni della temperatura interna e dell'utilizzo del sistema di generazione.

Con l'introduzione della fase di fine-tuning, l'agente riesce invece a adattare progressivamente la politica di controllo alle nuove condizioni operative. Dal punto di vista energetico si osserva una riduzione dell'energia importata dalla rete di circa 4300 kWh, accompagnata da una diminuzione del costo netto di circa 1000 € rispetto lo scenario zero-shot. Questo miglioramento si riflette anche in una regolazione della temperatura più stabile e in un bilancio energetico più coerente con quello osservato nel caso DRL addestrato direttamente sul dominio industriale.

Un'ulteriore analisi si ottiene confrontando le prestazioni di transfer fine-tuning e del controllore DRL. Questo confronto consente di valutare quanto il processo di adattamento tramite transfer learning riesca ad avvicinarsi alle prestazioni di un agente addestrato specificamente per il nuovo ambiente.

Dal punto di vista energetico, i risultati mostrano valori complessivamente simili tra i due approcci. Nel caso del fine-tuning l'energia importata dalla rete risulta pari a circa 38120 kWh, valore molto vicino a quello osservato nel caso DRL industriale (circa 37966

kWh). Anche i valori di autoconsumo diretto risultano comparabili, con circa 80% nel caso fine-tuning e 79% nel caso DRL.

Dal punto di vista economico si osserva una differenza relativamente contenuta. Il costo netto nel caso fine-tuning risulta pari a circa 8005 €, mentre nel caso DRL addestrato direttamente sul dominio industriale è pari a 7957 €. Questo evidenzia come, nonostante il miglioramento significativo rispetto allo scenario zero-shot, la politica trasferita non riesca a raggiungere una completa ottimalità.

Le differenze emergono anche dal punto di vista operativo. In particolare, la gestione del sistema di accumulo nel caso fine-tuning risulta meno strutturata rispetto al controllore DRL industriale, con uno stato di carica che tende a rimanere elevato per periodi prolungati e senza una chiara dinamica ciclica tra carica e scarica.

Nel complesso, il confronto evidenzia come il fine-tuning rappresenti un passaggio fondamentale per rendere applicabile il Transfer Learning in contesti caratterizzati da dinamiche energetiche differenti. Pur non raggiungendo le prestazioni di un agente addestrato direttamente sul dominio target, la politica trasferita riesce a ridurre significativamente il divario rispetto allo scenario zero-shot, confermando l'efficacia dell'approccio ma evidenziando al tempo stesso la necessità di un adattamento più approfondito per ottenere prestazioni pienamente ottimali.

### ***3.3.2. Caso Industriale 2***

Il secondo caso studio industriale introduce una configurazione caratterizzata da un maggiore dimensionamento dell'impianto fotovoltaico, tale da generare condizioni di surplus energetico nelle ore centrali della giornata. In questo scenario, la produzione fotovoltaica eccede frequentemente il fabbisogno dell'edificio, rendendo il sistema di accumulo un elemento centrale nella gestione dei flussi energetici.

L'obiettivo di questa analisi è valutare come le strategie di controllo si adattino a condizioni operative caratterizzate da un'elevata disponibilità di energia rinnovabile, analizzando in particolare il ruolo della batteria nell'aumentare l'autoconsumo e nel ridurre gli scambi con la rete. Questo consente inoltre di confrontare il comportamento dei controllori con quello osservato negli altri casi studio, caratterizzati anch'essi da condizioni di surplus.

### 3.3.2.1. Controllo con MPC

La strategia di controllo predittivo utilizzata nei casi precedenti è stata adottata anche per questo caso studio industriale.

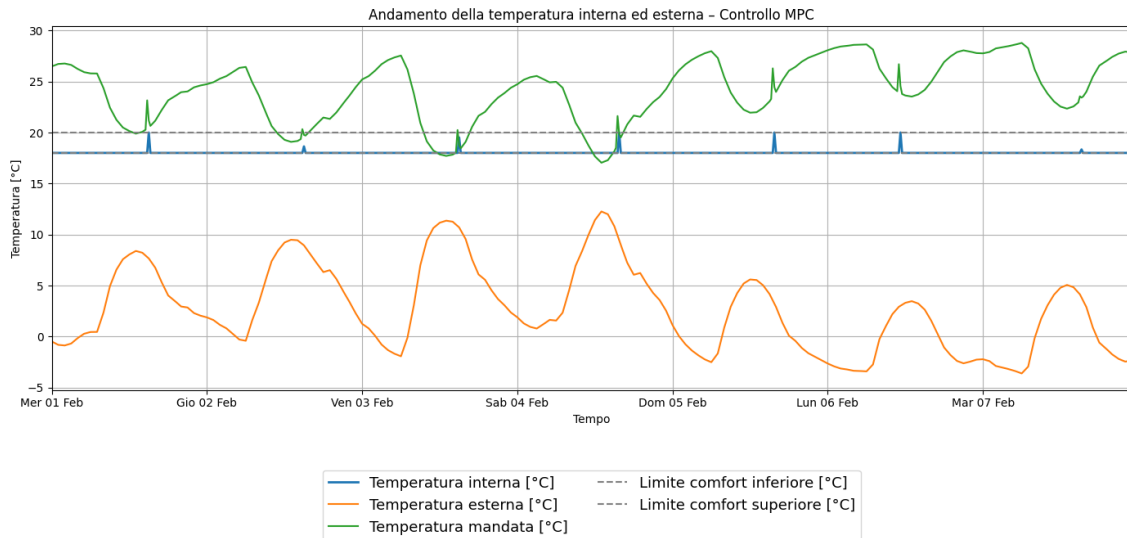


Figura 74: Andamento temperatura interna, esterna e di mandata UTA con controllore MPC - caso studio industriale 2

L'andamento della temperatura interna evidenzia come l'MPC riesce a mantenere l'ambiente in modo pressoché costante sul valore di set-point per l'intero periodo di simulazione. A differenza dei casi precedenti, in cui la temperatura tendeva a sfruttare l'intera banda di comfort per ottimizzare il costo energetico, nel caso industriale il comportamento risulta molto più stabile.

Nel medesimo grafico è inoltre riportata la temperatura di mandata dell'UTA, che rappresenta la variabile di controllo attraverso cui il sistema modula l'apporto termico all'ambiente. Questa variabile segue un andamento modulante nel tempo, adattandosi alle variazioni delle condizioni climatiche esterne e al fabbisogno termico dell'edificio.

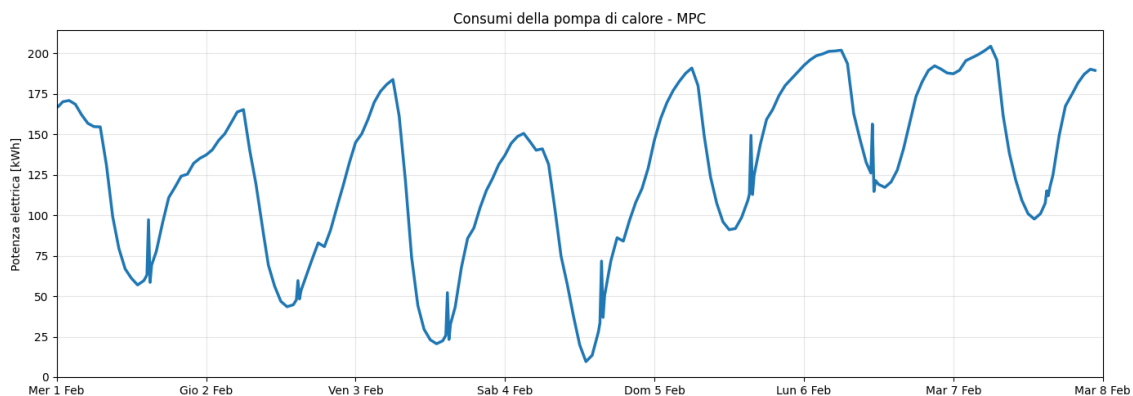


Figura 75: Andamento potenza elettrica consumata dalla pompa di calore con controllo MPC - caso studio industriale 2

Osservando la potenza della pompa di calore, possiamo notare come i valori siano decisamente più elevati rispetto ai casi precedenti, con picchi che superano i 100 kW. Anche in questo caso si osserva una modulazione coerente con le condizioni climatiche e con l'andamento della temperatura esterna.

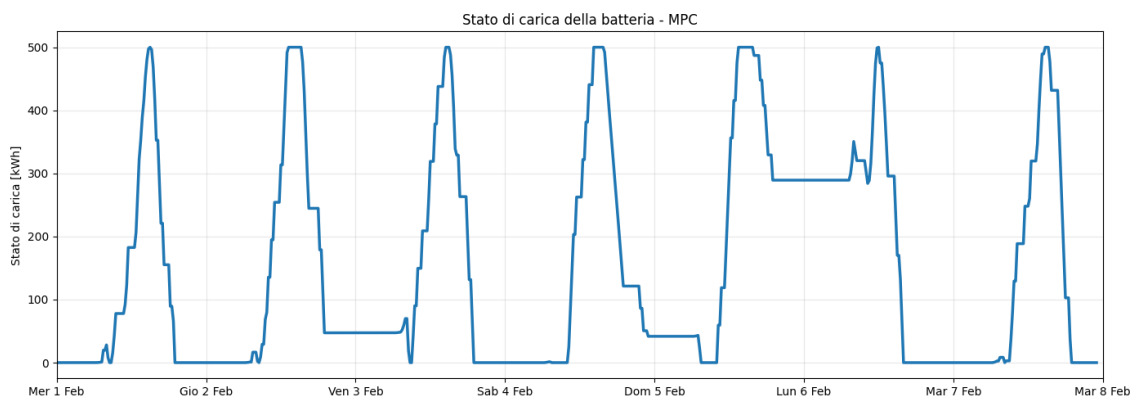


Figura 76: Andamento stato di carica della batteria con controllo MPC - caso studio industriale 2

La Figura 76) riporta invece lo stato di carica della batteria. Notiamo come l'andamento ha lo stesso comportamento dei casi precedenti con carica e scarica cicliche ma che sfruttano le previsioni climatiche che delle fasce orarie per permettere la minimizzazione dei costi.

Inoltre, l'MPC, anche in questo caso, tende a preservare parte dell'energia accumulata per utilizzarla nelle ore successive in cui la produzione fotovoltaica è minore, come avviene tra il 5 e 6 febbraio.

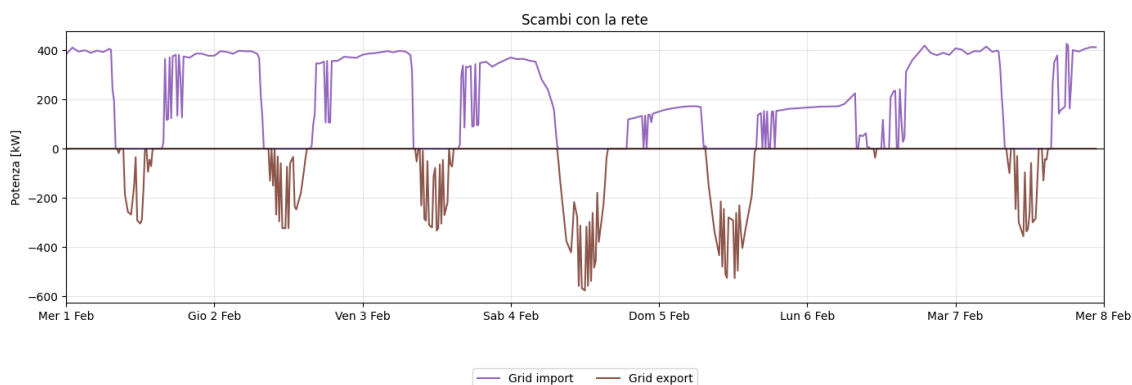


Figura 77: Andamento degli import ed export dalla rete con controllo MPC - caso studio industriale 2

Questo comportamento è visibile anche nel grafico degli scambi con la rete, dove si osserva una riduzione del prelievo medio della rete nonostante il giorno 6 corrisponda a

un giorno lavorativo e quindi a una condizione potenzialmente più gravosa dal punto di vista dei consumi.

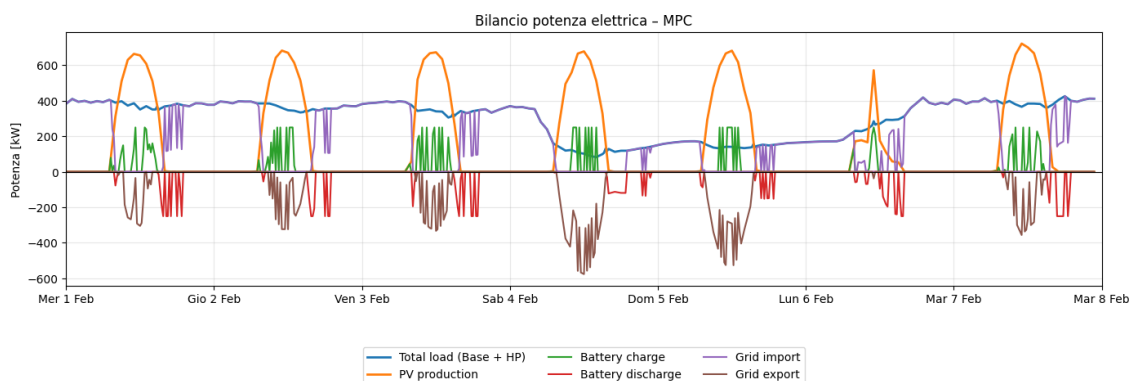


Figura 78: Bilancio della potenza elettrica con controllo MPC - caso studio industriale 2

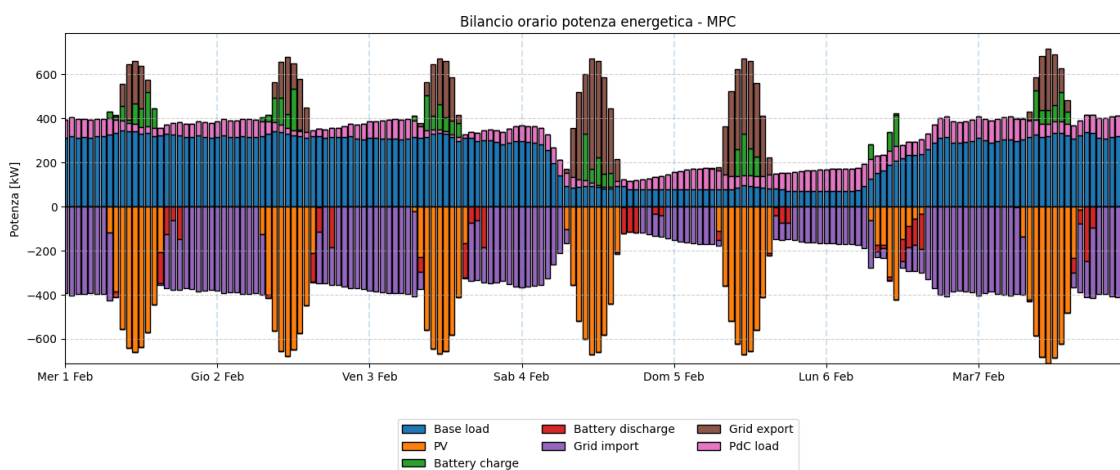


Figura 79: Bilancio orario della potenza elettrica con controllo MPC - caso studio industriale 2

Il bilancio energetico complessivo consente di leggere in modo più completo l'interazione tra i diversi componenti del sistema. La produzione fotovoltaica raggiunge nelle ore centrali della giornata valori superiori ai 600 kW, contribuendo in maniera significativa alla copertura del fabbisogno dell'edificio. Nel periodo considerato, l'energia complessivamente importata dalla rete risulta pari a circa 32638 kWh, mentre quella esportata ammonta a 8853 kWh. A fronte di una produzione fotovoltaica totale di 26883 kWh, l'edificio riesce a utilizzare direttamente circa il 54% dell'energia generata. Dal punto di vista economico, il costo netto risulta pari a 6362 €, un valore coerente con la scala molto più elevata del caso industriale rispetto ai casi studio precedenti.

Il comportamento del controllore MPC in questo scenario risulta coerente con quanto osservato nel caso industriale precedente. Nonostante il diverso dimensionamento

dell'impianto fotovoltaico e la maggiore presenza di surplus energetico, il controllore mantiene una strategia di gestione simile, basata su un utilizzo pianificato del sistema di accumulo e su una riduzione degli scambi con la rete. Questo evidenzia la capacità dell'MPC di adattarsi a condizioni operative differenti mantenendo una logica di ottimizzazione coerente.

### 3.3.2.2. Controllo con Deep Reinforcement Learning

Analizziamo ora il comportamento del controllore basata su DRL nel caso industriale caratterizzato da una configurazione con surplus di produzione energetica.

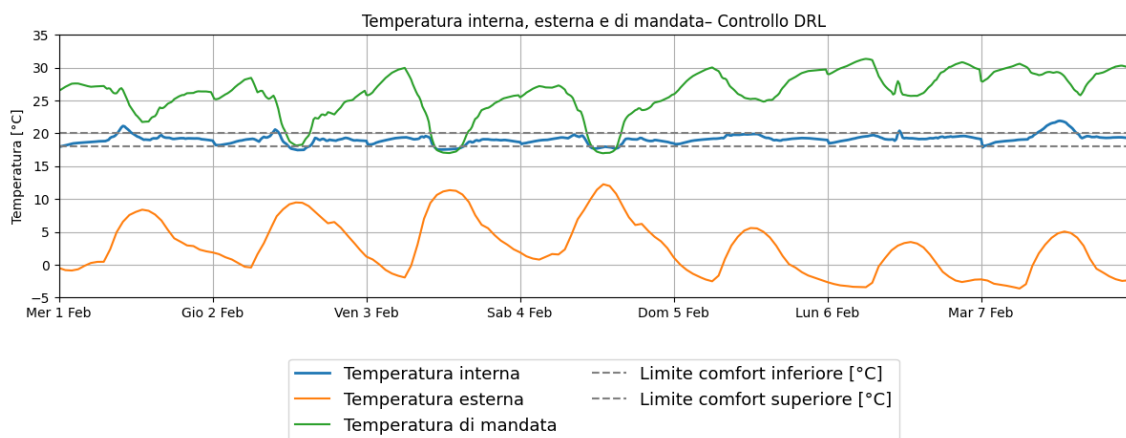


Figura 80: Andamento temperatura interna, esterna e di mandata UTA con controllore DRL - caso studio industriale 2

L'andamento delle temperature interne, esterne e di mandata evidenzia come il controllo DRL riesca a mantenere condizioni operative complessivamente stabili. La temperatura interna si mantiene prevalentemente entro i limiti di comfort, con oscillazioni contenute attorno al valore di riferimento. La temperatura di mandata mostra un comportamento dinamico, adattandosi alle variazioni climatiche esterne e alle esigenze dell'edificio, mentre la temperatura esterna segue il tipico andamento giornaliero.

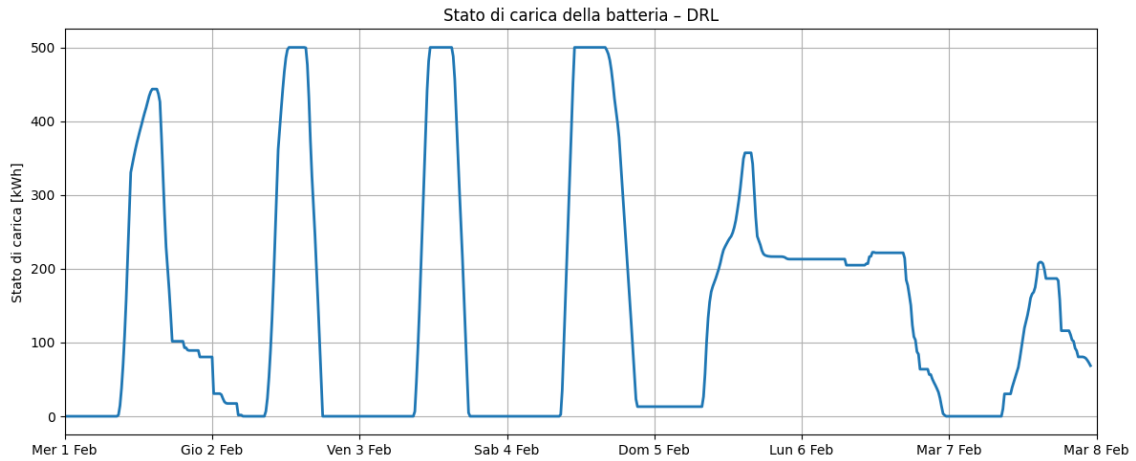


Figura 81: Andamento stato di carica della batteria con controllo DRL - caso studio industriale 2

Lo stato di carica della batteria evidenzia cicli di carica e scarica ben definiti. Durante le ore centrali della giornata, in presenza di elevata produzione fotovoltaica, la batteria viene caricata fino alla capacità massima, mentre nelle ore serali o nei momenti di ridotta produzione viene progressivamente scaricata.

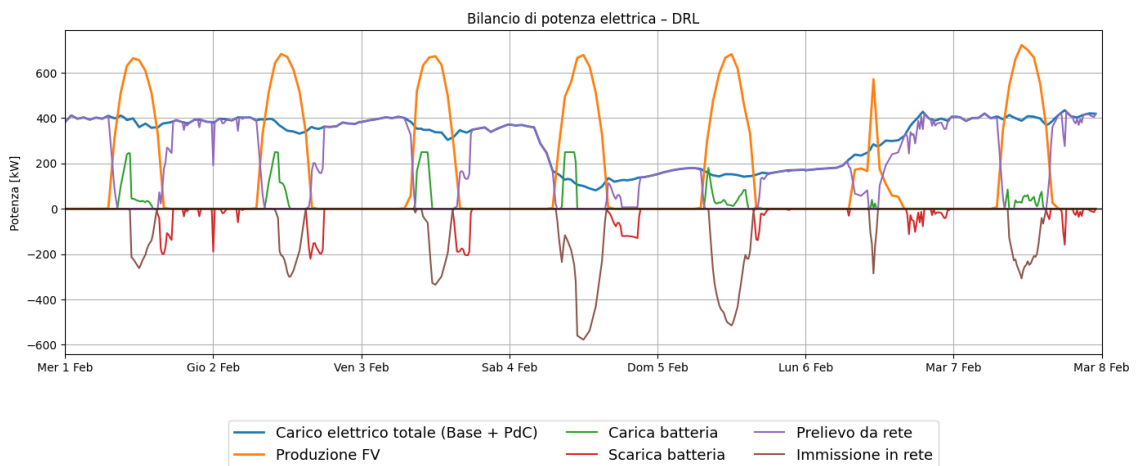


Figura 82: Bilancio della potenza elettrica con controllo DRL - caso studio industriale 2

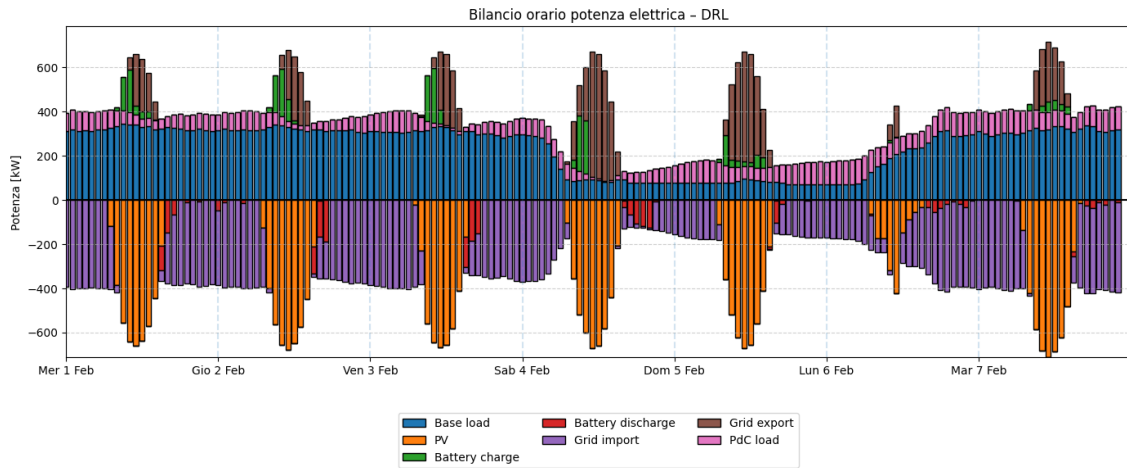


Figura 83: Bilancio orario della potenza elettrica con controllo DRL - caso studio industriale 2

L'analisi del bilancio di potenza mostra chiaramente il ruolo dominante della produzione fotovoltaica nelle ore diurne. In tali intervalli, la produzione eccede il fabbisogno elettrico dell'edificio, determinando flussi di energia verso la rete. La batteria contribuisce a ridurre parzialmente questi eccessi, caricandosi durante i picchi di produzione e scaricandosi nei momenti di maggiore domanda. Tuttavia, la capacità limitata del sistema di accumulo non consente di assorbire completamente il surplus, rendendo necessaria l'immissione in rete. Il prelievo dalla rete si concentra principalmente nelle ore notturne o nei periodi di bassa produzione fotovoltaica, evidenziando una dipendenza residua dalla rete elettrica.

Nel periodo considerato, l'energia complessivamente importata dalla rete risulta pari a circa 34247 kWh, mentre quella esportata ammonta a 9274 kWh. A fronte di una produzione fotovoltaica totale di 2883 kWh, l'edificio riesce ad ottenere un autoconsumo diretto di circa 56%. Dal punto di vista economico, il costo di acquisto dell'energia dalla rete risulta pari a 7194 €, a fronte di un ricavo dalla vendita di circa 464 €, determinando un costo netto complessivo di 6 731 €. Questo valore risulta coerente con la scala del caso industriale e con i livelli di consumo considerati.

Il controllo basato su DRL conferma anche in questo scenario una buona capacità di coordinare i diversi componenti del sistema energetico. La strategia riesce infatti a mantenere condizioni operative adeguate e, allo stesso tempo, a gestire in modo efficace l'interazione tra produzione fotovoltaica, sistema di accumulo e rete elettrica. Si osserva tuttavia come il controllo della temperatura interna possa presentare margini di miglioramento, in particolare nella riduzione di alcune oscillazioni, suggerendo possibili

sviluppi futuri nella definizione della politica di controllo. Il comportamento osservato evidenzia come l'approccio DRL sia in grado di adattarsi a contesti industriali caratterizzati da elevata produzione rinnovabile e da dinamiche energetiche articolate.

### ***3.3.2.3. Applicazione del Transfer Learning***

Applicando gli stessi ragionamenti sviluppati per il precedente caso industriale, si analizza qui un secondo scenario caratterizzato da un diverso dimensionamento dell'impianto fotovoltaico. In questo caso, analogamente a quanto osservato nei casi residenziale e scolastico, la produzione fotovoltaica risulta frequentemente superiore al fabbisogno elettrico dell'edificio nelle ore centrali della giornata, generando condizioni di surplus energetico.

In questo contesto, una quota dell'energia prodotta non può essere direttamente assorbita dai carichi e viene quindi destinata alla ricarica del sistema di accumulo o immessa in rete, rendendo la batteria un elemento centrale nella gestione dei flussi energetici. Questo scenario consente inoltre di osservare come il comportamento dell'agente sia influenzato dal bilanciamento tra produzione, domanda e capacità di accumulo in presenza di condizioni energetiche simili a quelle incontrate durante l'addestramento, il controllore tende infatti a replicare strategie operative analoghe.

L'obiettivo è quindi valutare come un agente addestrato in un dominio differente riesca a adattarsi a queste condizioni operative, caratterizzate da una significativa disponibilità di energia rinnovabile rispetto al fabbisogno.

Anche in questo caso, l'analisi viene condotta considerando entrambe le modalità di trasferimento. In un primo momento si applica il transfer zero-shot, trasferendo direttamente la politica appresa nel caso residenziale al caso industriale. Successivamente si analizza il transfer fine-tuning, in cui l'agente viene ulteriormente riaddestrato nel nuovo dominio, così da adattare progressivamente la politica di controllo alle specifiche dinamiche energetiche del sistema.

### ***Transfer zero-shot***

In questo scenario, la politica di controllo appresa nel caso residenziale viene trasferita direttamente all'edificio industriale, senza alcuna fase intermedia di riaddestramento o adattamento al nuovo dominio.

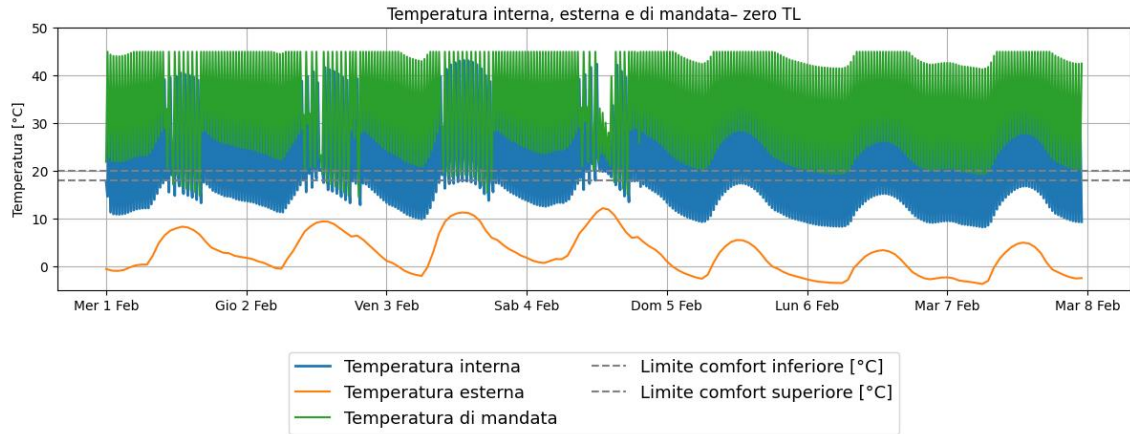


Figura 84: Andamento temperatura interna, esterna e di mandata UTA con TL zero-shot - caso studio industriale 2

L'andamento della temperatura interna mette in evidenza in modo piuttosto chiaro i limiti di questo trasferimento diretto. Si osservano infatti oscillazioni ampie e frequenti, con valori che in diversi momenti si discostano sensibilmente dalla banda di comfort prevista. Questo comportamento mostra come la politica appresa nel contesto residenziale non sia in grado di governare in maniera efficace le dinamiche termiche dell'edificio industriale. La ragione principale risiede nella forte differenza tra i due domini, infatti, nel caso industriale la scala dei carichi termici ed elettrici è molto più elevata e le dinamiche del sistema risultano più complesse. Di conseguenza, le azioni apprese durante il training nel dominio sorgente non risultano adeguate a garantire un controllo stabile della temperatura nel nuovo contesto.

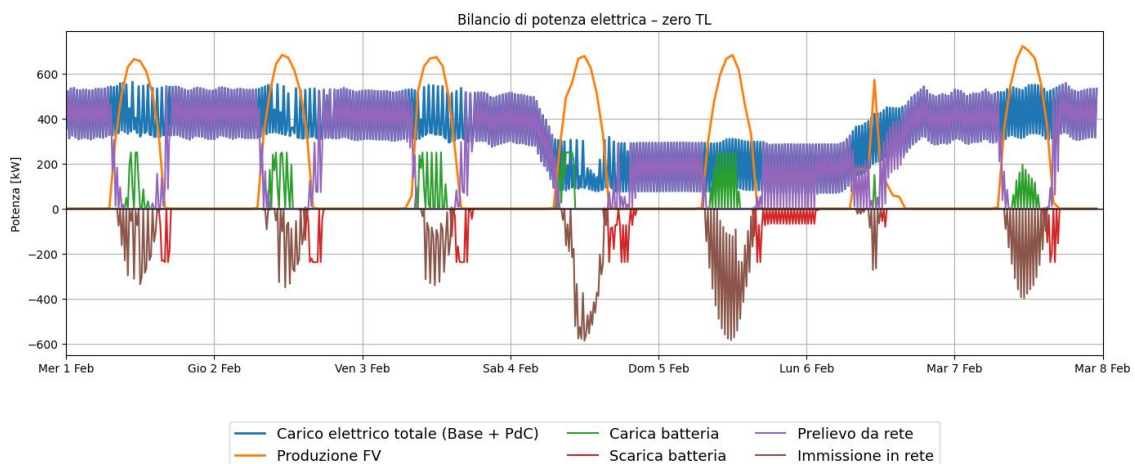


Figura 85: Bilancio della potenza elettrica con TL zero-shot - caso studio industriale 2

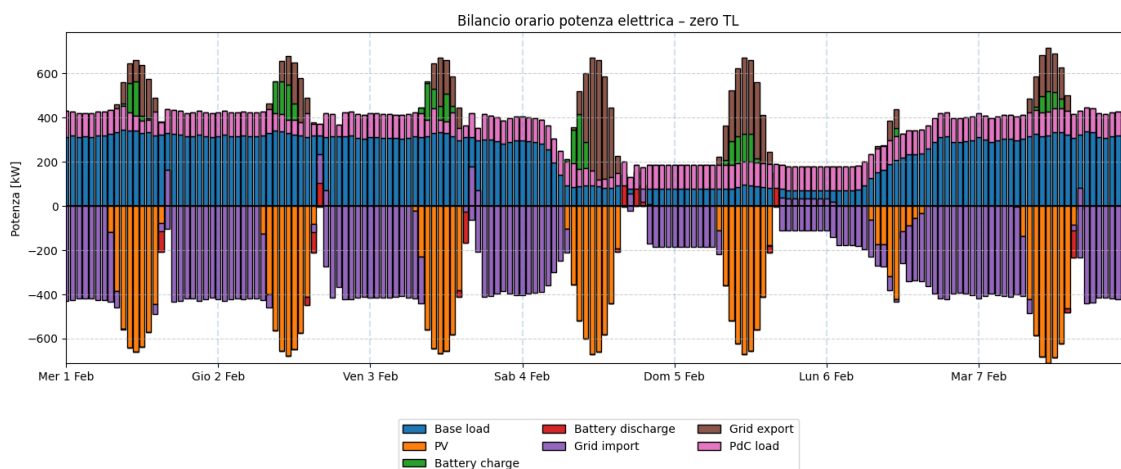


Figura 86: Bilancio orario della potenza elettrica con TL zero-shot - caso studio industriale 2

Le stesse criticità emergono anche osservando il bilancio energetico dell'edificio. In particolare, la potenza assorbita dalla pompa di calore e gli scambi con la rete mostrano variazioni molto marcate nel tempo, segnale di un'azione di controllo poco stabile e non pienamente coerente con il reale fabbisogno energetico dell'edificio. Il comportamento del controllore appare quindi poco adattato alle caratteristiche del sistema industriale, con effetti negativi sia sulla stabilità operativa sia sulle prestazioni complessive.

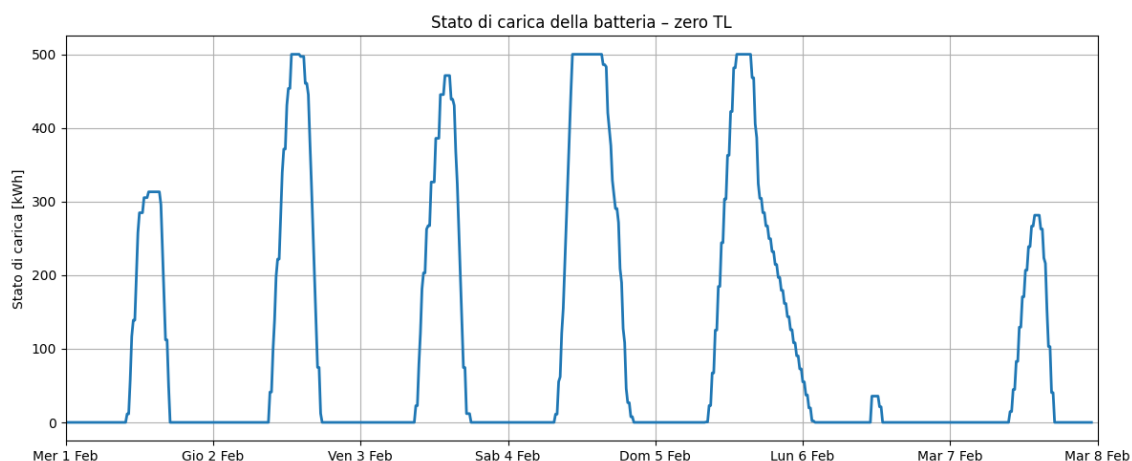


Figura 87: Andamento stato di carica della batteria con TL zero-shot - caso studio industriale 2

L'andamento dello stato di carica della batteria mostra invece comportamento relativamente più coerente. In tali condizioni, il sistema di accumulo viene caricato in modo più significativo durante le ore diurne, quando la produzione fotovoltaica supera temporaneamente il fabbisogno dell'edificio. Questo comportamento risulta analogo a quanto osservato nel caso di transfer zero-shot applicato all'edificio scolastico.

Questa somiglianza suggerisce che la politica appresa dall'agente sia influenzata in misura rilevante dal rapporto tra produzione fotovoltaica e capacità di accumulo. Quando tale rapporto si mantiene simile tra scenari diversi, il controllore tende infatti a replicare strategie energetiche analoghe, indipendentemente dalla tipologia di edificio considerata. Tuttavia, questa coerenza nella gestione della batteria non è sufficiente a compensare le difficoltà del controllore nel governare il sistema termico complessivo.

Nel periodo di simulazione, l'energia complessivamente importata dalla rete risulta pari a circa 37562 kWh, mentre quella esportata ammonta a circa 7736 kWh. La produzione fotovoltaica totale rimane pari a 26883 kWh, consentendo un autoconsumo diretto del 61%. Dal punto di vista economico, questi valori si traducono in un costo di acquisto dalla rete pari a 7929 € e in un ricavo derivante dalla vendita di energia pari a 387 €, con un costo netto complessivo di 7542 €.

Anche in questo caso, una fase di fine-tuning risulta necessaria per addestrare l'agente nel nuovo dominio con caratteristiche fisiche differenti dell'edificio per permettere di migliorare il controllo del sistema di generazione e di conseguenza permettere il confort della temperatura interna.

### ***Transfer fine-tuning***

In questo secondo scenario di trasferimento, la politica di controllo appresa nel caso residenziale viene ulteriormente riaddestrata nel dominio industriale caratterizzato da una significativa presenza di surplus fotovoltaico. L'obiettivo è consentire all'agente di adattare progressivamente la propria politica alle nuove condizioni operative, mantenendo al contempo le strategie apprese in fase di training.

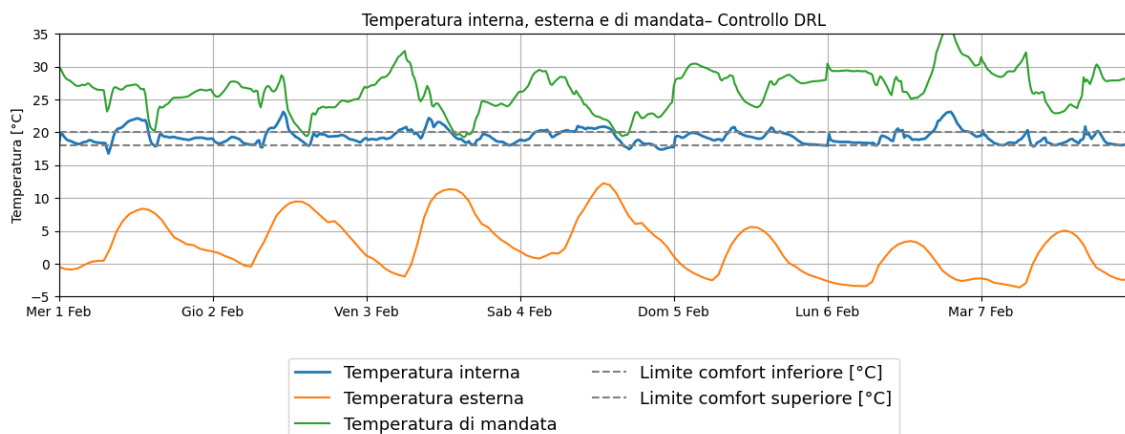


Figura 88 Andamento temperatura interna, esterna e di mandata UTA con TL fine-tuning-caso studio industriale 2

L'andamento della temperatura interna evidenzia un comportamento complessivamente stabile, con valori che si mantengono prevalentemente all'interno della banda di comfort. Le oscillazioni risultano contenute rispetto al caso di trasferimento diretto, indicando una buona capacità dell'agente di adattarsi alle dinamiche termiche dell'edificio industriale.

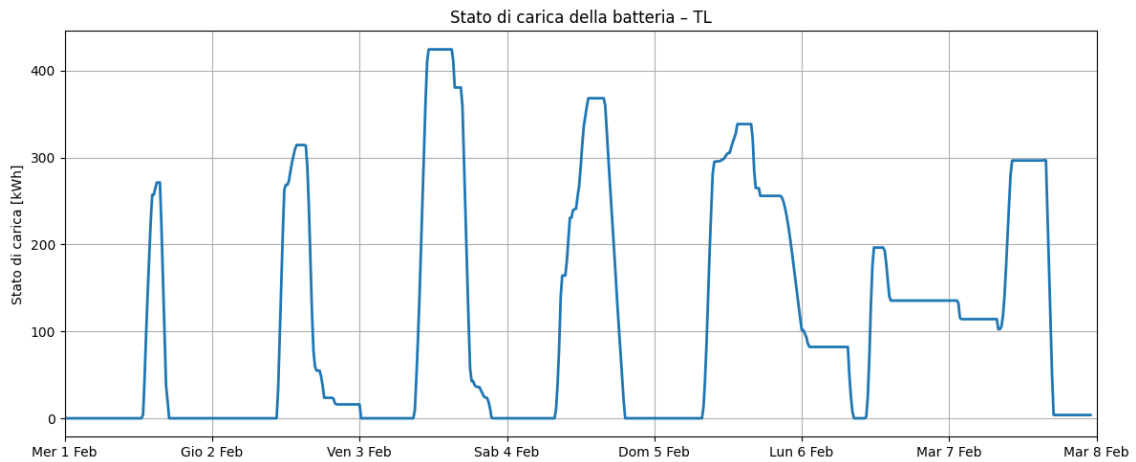


Figura 89: Andamento stato di carica della batteria con TL fine-tuning - caso studio industriale 2

La batteria viene caricata nelle ore centrali della giornata in presenza di elevata produzione fotovoltaica e scaricata nelle fasi successive, seguendo una dinamica più coerente con il profilo energetico del sistema. Tuttavia, si osserva che lo stato di carica non raggiunge la capacità massima, a differenza di quanto avveniva nel caso zero-shot. Questo indica che l'agente ha modificato la propria strategia, ma senza riuscire a sfruttare completamente il sistema di accumulo in presenza di surplus energetico.

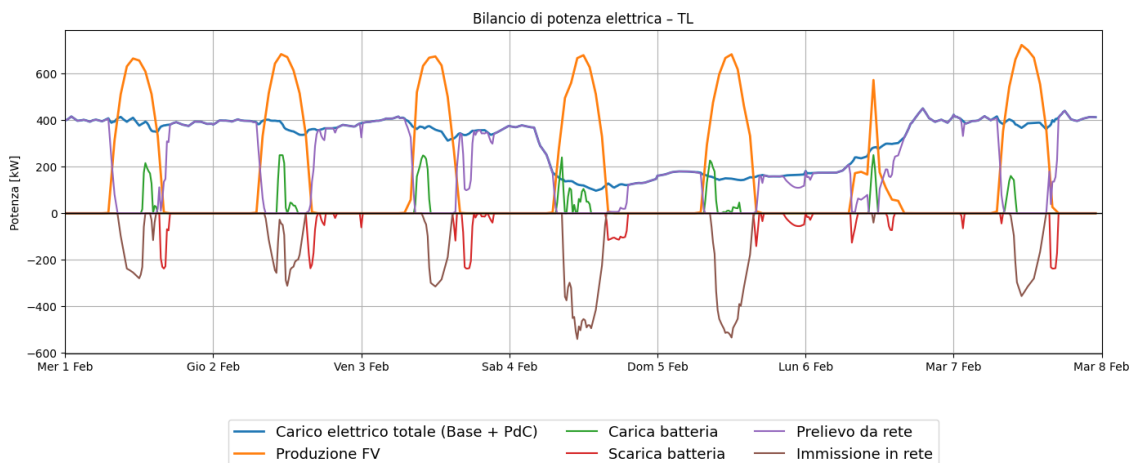


Figura 90: Bilancio della potenza elettrica con TL fine-tuning - caso studio industriale 2

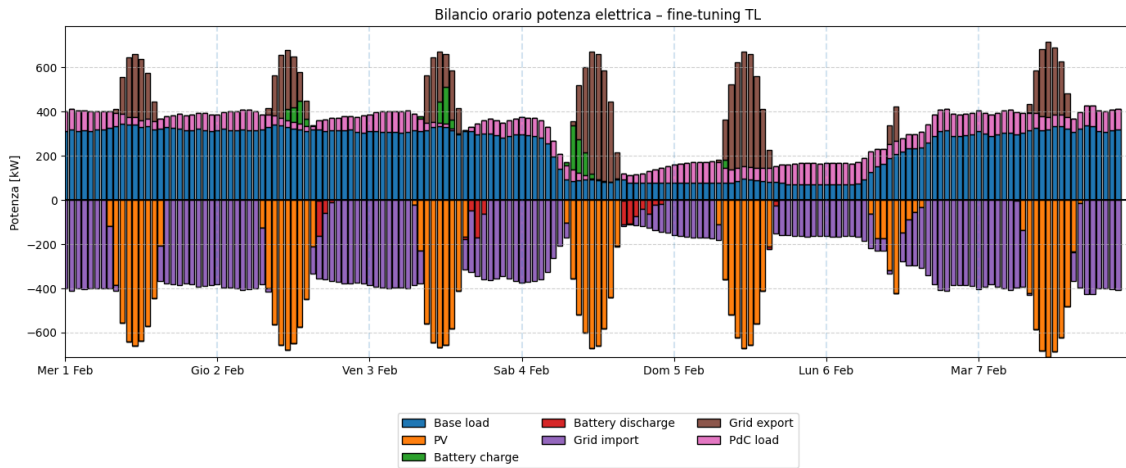


Figura 91: Bilancio orario della potenza elettrica con TL fine-tuning - caso studio industriale 2

Dal punto di vista energetico, il bilancio di potenza evidenzia un comportamento complessivamente più regolare rispetto al caso zero-shot, ma ancora non pienamente ottimizzato. Nelle ore centrali della giornata, la produzione fotovoltaica supera frequentemente il fabbisogno dell'edificio, generando condizioni di surplus. In queste fasi, la batteria contribuisce solo parzialmente all'assorbimento dell'energia in eccesso, mentre una quota significativa viene comunque immessa in rete. Questo indica che il sistema di accumulo non viene sfruttato in modo completo per massimizzare l'autoconsumo.

Nel periodo di simulazione, l'energia importata dalla rete risulta pari a circa 34588 kWh, mentre l'energia esportata ammonta a circa 7492 kWh. A fronte di una produzione fotovoltaica totale di 26883 kWh, l'autoconsumo diretto si attesta intorno al 56 %, valore in linea con gli altri controllori ma ottenuto con un utilizzo meno efficace del sistema di accumulo. Dal punto di vista economico, il costo di acquisto dell'energia dalla rete è pari a circa 7275 €, a fronte di un ricavo di circa 475 €, per un costo netto complessivo di circa 6800 €.

Il confronto con il caso DRL addestrato direttamente sul dominio industriale evidenzia come il fine-tuning, pur consentendo un adattamento della politica di controllo, non riesca a raggiungere prestazioni ottimali in questo scenario. Questo risultato suggerisce che, in presenza di condizioni operative fortemente influenzate dal rapporto tra produzione rinnovabile e capacità di accumulo, la politica trasferita tende a mantenere comportamenti appresi nel dominio sorgente che non risultano completamente adeguati al nuovo contesto. Il fine-tuning migliora la stabilità rispetto a un trasferimento diretto, ma non è

sufficiente a garantire una gestione ottimale del sistema energetico, evidenziando la necessità di un adattamento più approfondito della politica di controllo.

### 3.3.2.4. Controllo tradizionale PID + Rule-Based

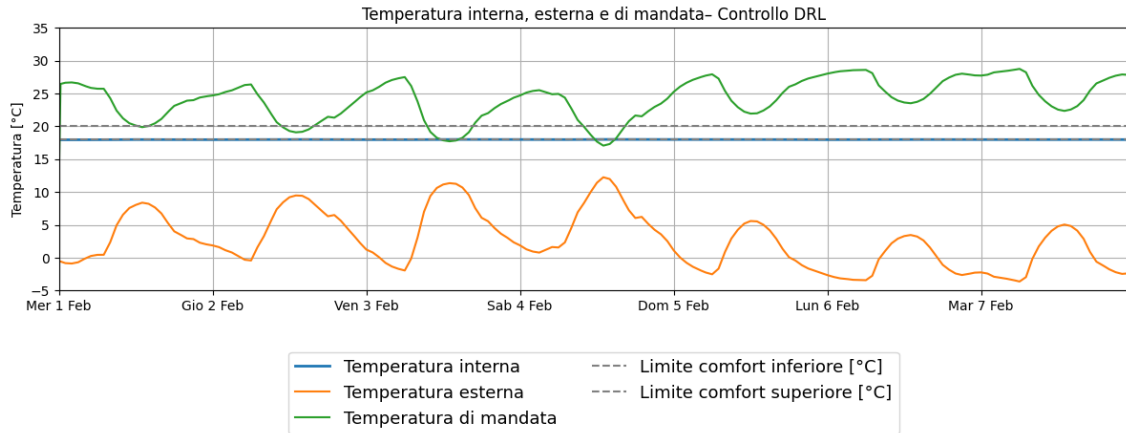


Figura 92: Andamento temperatura interna, esterna e di mandata UTA con controllore PID +RB -caso studio industriale 2

L'andamento della temperatura interna evidenzia come il controllore riesca a mantenere condizioni generalmente stabili sul valore di set-point. Questo comportamento è tipico di una regolazione PID con set-point fisso, in cui il sistema tende a mantenere la temperatura il più possibile costante, senza sfruttare dinamicamente l'intera banda di comfort. A differenza delle strategie avanzate, non si osservano azioni di anticipazione rispetto alle variazioni delle condizioni esterne, ma una risposta prevalentemente reattiva.

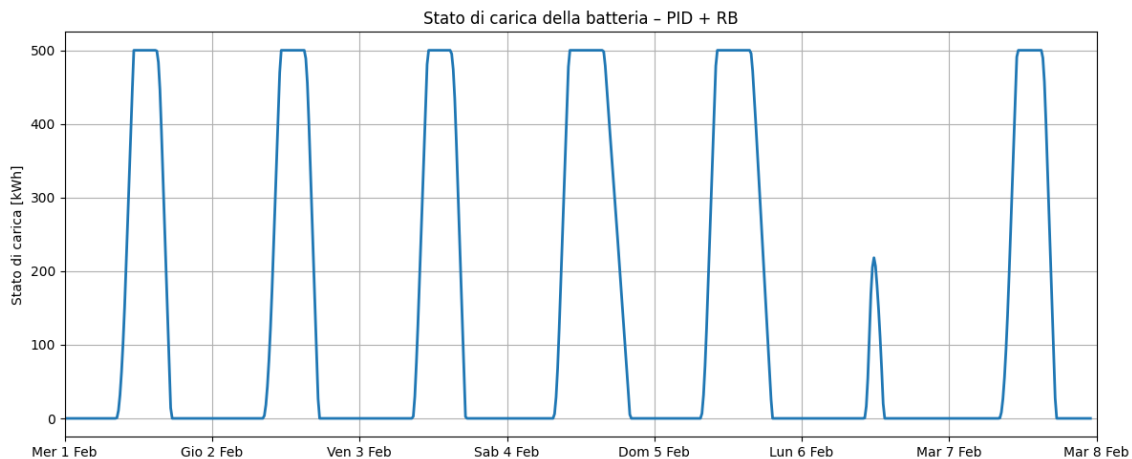


Figura 93: Andamento stato di carica della batteria con controllo PID +RB - caso studio industriale 2

Lo stato di carica della batteria mostra un comportamento fortemente deterministico, caratterizzato da cicli di carica e scarica ben definiti ma poco flessibili. In presenza di

produzione fotovoltaica, la batteria viene rapidamente caricata fino alla capacità massima, mentre in assenza di produzione viene scaricata per coprire il fabbisogno. Questo approccio non tiene conto né delle previsioni né delle condizioni economiche, limitando l'efficacia complessiva della gestione energetica.

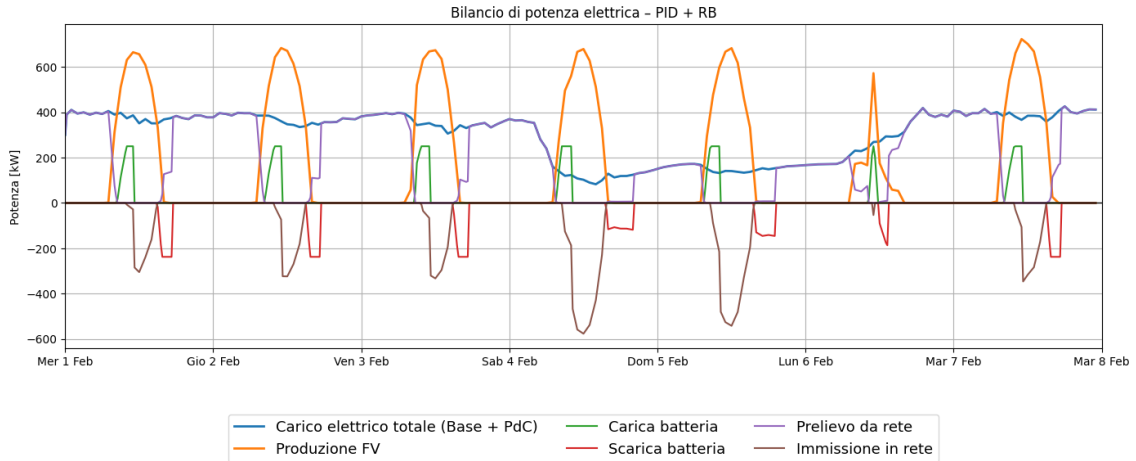


Figura 94: Bilancio della potenza elettrica con controllo PID +RB - caso studio industriale 2

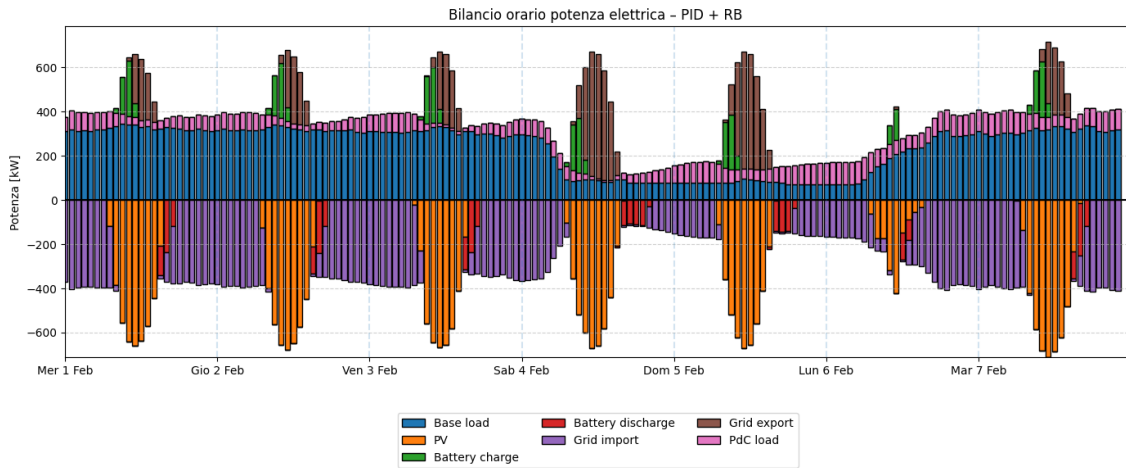


Figura 95: Bilancio orario della potenza elettrica con controllo PID +RB- caso studio industriale2

L'analisi del bilancio di potenza evidenzia come il sistema si comporta in modo tradizionale in base alle regole imposte senza nessun tipo di ottimizzazione strategica.

Nel periodo considerato, l'energia complessivamente importata dalla rete risulta pari a circa 32639 kWh, mentre quella esportata ammonta a 8846 kWh. A fronte di una produzione fotovoltaica totale di 26883 kWh, l'edificio riesce a raggiungere un autoconsumo diretto di circa 54%. Economicamente invece, il costo di acquisto dell'energia dalla rete risulta pari a 6807 €, mentre il ricavo derivante dalla vendita di

energia è pari a circa 442 €, determinando un costo netto complessivo di 6365 €. Anche in questo caso, il valore risulta coerente con la scala del sistema analizzato.

### 3.3.2.5. Confronto fra strategie di controllo

La Tabella 8) riassume i principali indicatori energetici ed economici ottenuti per le diverse strategie di controllo nel caso studio industriale.

Indicatore	MPC	DRL	PID+RL	Transfer zero-shot	Transfer fine-tuning
Comfort rispettato	Sì	Sì	Sì	No	Parzialmente
Energia importata [kWh]	32638	34247	32639	37562	34588
Energia esportata [kWh]	8853	9274	8846	7736	7492
Consumi PdC [kWh]	10680	11811	10660	16720	12037
Produzione FV [kWh]	26883	26883	26883	26883	26883
Autoconsumo diretto [%]	54	56	54	61	56
Energia batteria carica [kWh]	3532	2646	3412	2738	2219
Energia batteria scarica [kWh]	3188	2323	3119	2471	1999
Costo netto [€]	6362	6731	6365	7542	6800

Tabella 8: Confronto fra le diverse strategie di controllo - caso studio industriale 2

Dal punto di vista del comfort termico, tutte le strategie, ad eccezione del transfer zero-shot, riescono a garantire il mantenimento della temperatura interna all'interno della banda prevista. Il trasferimento diretto della politica, infatti, mostra ancora criticità legate alla mancata adattabilità alle dinamiche del sistema industriale, mentre il fine-tuning consente di migliorare significativamente la stabilità del controllo, pur non raggiungendo le prestazioni degli altri approcci.

Analizzando gli indicatori energetici, si osserva come i controllori MPC e PID + rule-based presentino i valori più contenuti di energia importata dalla rete, seguiti dal DRL, mentre le strategie basate su TL mostrano prestazioni peggiori. In particolare, il transfer zero-shot evidenzia il valore più elevato di energia importata, mentre il fine-tuning consente una riduzione parziale di tale valore, senza tuttavia raggiungere i livelli dei controllori progettati direttamente per il dominio industriale.

Per quanto riguarda l'energia esportata in rete, i valori risultano elevati in tutte le strategie a causa delle condizioni di surplus fotovoltaico. Tuttavia, il fine-tuning non

rappresenta il caso peggiore, ma mostra un comportamento intermedio tra le altre strategie. Questo indica che, pur migliorando rispetto al trasferimento diretto, la gestione del surplus energetico non risulta ancora completamente ottimizzata.

L'analisi dell'utilizzo della batteria conferma questa interpretazione. Il controllore MPC mostra i valori più elevati di energia caricata e scaricata, evidenziando una gestione attiva dell'accumulo finalizzata a ottimizzare i flussi energetici nel tempo. Il DRL utilizza la batteria in modo meno marcato, mentre il fine-tuning presenta un utilizzo non ancora ottimo rispetto agli altri controllori. Questo comportamento si traduce in una capacità parziale di assorbire il surplus fotovoltaico.

Dal punto di vista economico, il controllo MPC risulta il più performante, seguito dal PID + rule-based, mentre il DRL presenta un costo leggermente superiore. Le strategie basate su TL mostrano invece prestazioni peggiori, con il fine-tuning che, pur migliorando rispetto allo scenario zero-shot, non riesce a raggiungere i risultati ottenuti dagli altri controllori.

Nel complesso, questo caso studio evidenzia come, in presenza di condizioni di surplus energetico, una gestione efficace del sistema di accumulo risulti fondamentale per ottimizzare le prestazioni del sistema. I controllori progettati direttamente per il dominio industriale riescono a sfruttare meglio questa caratteristica, mentre le politiche trasferite, anche dopo una fase di fine-tuning, mostrano ancora difficoltà nell'adattarsi completamente a queste dinamiche.

## ***4. Discussione***

In questo capitolo vengono discussi i principali risultati emersi dalle simulazioni presentate nel capitolo precedente. L'obiettivo è interpretare il comportamento delle diverse strategie di controllo analizzate, mettendo in evidenza i fenomeni energetici che influenzano le prestazioni del sistema. L'analisi prende in considerazione sia gli aspetti energetici ed economici sia il comportamento dinamico dei controllori nei diversi casi studio.

### ***4.1. Prestazioni delle strategie di controllo***

Il confronto tra le diverse strategie di controllo mostra come gli approcci avanzati, in particolare il Model Predictive Control (MPC) e il Deep Reinforcement Learning (DRL), siano in grado di gestire in modo più efficace il sistema energetico degli edifici rispetto a strategie di controllo tradizionali.

In tutti i casi studio analizzati, le strategie considerate riescono a mantenere la temperatura all'interno della banda di comfort prevista. Le differenze tra i controllori emergono invece nel modo in cui viene gestito il sistema energetico dell'edificio.

Il controllore MPC mostra un comportamento più pianificato grazie alla sua natura predittiva. Utilizzando un modello dell'edificio e informazioni sulle condizioni future, il controllore è in grado di programmare nel tempo l'utilizzo delle diverse risorse energetiche disponibili. Questo aspetto risulta particolarmente evidente nella gestione della batteria, dove il controllore tende a preservare l'energia accumulata quando il costo dell'energia è più basso per utilizzarla nelle fasce orarie economicamente meno convenienti.

Il controllore basato su DRL mostra, invece, un comportamento più reattivo. La politica di controllo viene infatti appresa attraverso l'interazione con l'ambiente e si basa principalmente sulle condizioni osservate istante per istante. Questo approccio consente al controllore di adattarsi alle diverse condizioni operative dell'edificio, anche se non dispone della stessa capacità di pianificazione su orizzonte temporale futuro tipica del controllo predittivo.

I risultati mostrano come il controllo predittivo offra una maggiore capacità di pianificazione delle risorse energetiche, mentre il controllo basato su apprendimento

risulti più flessibile e adattivo alle diverse condizioni operative. Questo evidenzia il tipico compromesso tra approcci modellistici e approcci basati su apprendimento automatico nel controllo energetico degli edifici.

È importante osservare che le prestazioni del controllore MPC dipendono fortemente dall'accuratezza del modello dell'edificio e dalla qualità delle previsioni utilizzate nel processo di ottimizzazione, come ad esempio le previsioni meteorologiche o di carico. Errori nel modello o nelle previsioni possono infatti ridurre l'efficacia della pianificazione delle azioni di controllo. Questo aspetto rappresenta uno dei principali limiti dei controllori predittivi, poiché la qualità della soluzione ottimizzata dipende direttamente dalla capacità del modello di rappresentare correttamente le dinamiche del sistema reale. Al contrario, gli approcci basati su Deep Reinforcement Learning non richiedono un modello esplicito del sistema e apprendono direttamente dall'interazione con l'ambiente, risultando potenzialmente più robusti in presenza di incertezze modellistiche.

Un ulteriore aspetto interessante riguarda il modo in cui i controllori sfruttano l'inerzia termica dell'edificio. Nei casi residenziale e scolastico si osserva che, quando le condizioni energetiche risultano favorevoli, entrambi i controllori tendono ad aumentare la potenza della pompa di calore, portando temporaneamente la temperatura verso il limite superiore della banda di comfort. Questo comportamento si verifica, ad esempio, nelle ore in cui è disponibile una maggiore produzione fotovoltaica o quando il costo dell'energia risulta più conveniente.

In queste condizioni l'edificio può accumulare parte dell'energia termica nella propria massa, costituita da pareti, pavimento o strutture interne. Il calore immagazzinato viene poi rilasciato gradualmente nel tempo, contribuendo al mantenimento del comfort anche nelle ore successive. Dal punto di vista energetico, questo fenomeno può essere interpretato come una forma di accumulo virtuale di energia (virtual energy storage). In questo caso l'energia non viene immagazzinata in un sistema di accumulo dedicato, come una batteria o un serbatoio termico, ma direttamente nella massa dell'edificio sfruttandone l'inerzia termica. L'edificio stesso diventa quindi un elemento di accumulo temporaneo dell'energia, permettendo di spostare nel tempo parte dei consumi energetici senza l'utilizzo di dispositivi di storage aggiuntivi.

Nei casi industriali questo fenomeno risulta invece molto meno evidente. Nel modello considerato la capacità termica dell'edificio è rappresentata principalmente dal volume

dell'aria interno, mentre la massa delle strutture edilizie non contribuisce in modo significativo all'accumulo termico. Questa configurazione è legata alla tipologia di impianto considerato, basato su una Unità di Trattamento Aria (UTA) con batteria di riscaldamento alimentata da pompa di calore, caratterizzata da elevati ricambi d'aria tipici degli ambienti industriali. In queste condizioni l'aria interna viene continuamente rinnovata, riducendo la possibilità di accumulare energia termica nel volume interno dell'edificio. Poiché l'aria possiede una capacità termica relativamente bassa, la quantità di energia che può essere immagazzinata nell'edificio risulta limitata e le strategie di sfruttamento dell'inerzia termica risultano quindi meno efficaci. Questo risultato evidenzia come le prestazioni delle strategie di controllo dipendano non solo dall'algoritmo utilizzato, ma anche dalle caratteristiche fisiche dell'edificio, dal sistema impiantistico adottato e dalla sua capacità di accumulo termico.

#### **4.2. *Ruolo del sistema di accumulo e dell'autoconsumo***

Un altro aspetto rilevante riguarda il ruolo del sistema di accumulo nel bilancio energetico degli edifici.

Nel caso residenziale, scolastico e nel secondo caso industriale, la produzione fotovoltaica supera frequentemente il carico elettrico dell'edificio nelle ore centrali della giornata. In queste condizioni la batteria assume un ruolo importante, poiché consente di aumentare l'autoconsumo dell'energia prodotta localmente e di ridurre gli scambi con la rete elettrica.

In questi scenari, il controllo MPC riesce generalmente a sfruttare in modo più efficace il sistema di accumulo. Grazie alla capacità di prevedere l'evoluzione delle condizioni operative nelle ore successive, il controllore può pianificare i cicli di carica e scarica delle batterie in modo più strategico. Questo permette, ad esempio, di conservare parte dell'energia accumulata quando il costo energetico è più basso e utilizzarla successivamente nelle fasce orarie economicamente meno convenienti.

Nel caso del controllore DRL, invece, l'utilizzo della batteria risulta maggiormente legato alle condizioni energetiche osservate nel momento corrente. La politica di controllo appresa tende quindi a reagire alle condizioni operative istantanee del sistema,

portando a cicli di carica e scarica meno pianificati dal punto di vista economico, pur mantenendo prestazioni energetiche complessivamente comparabili.

Nei casi industriali, il ruolo della batteria dipende in modo significativo dal bilanciamento tra produzione fotovoltaica e domanda energetica. Nel caso più bilanciato, il contributo dell'accumulo risulta limitato, poiché una quota significativa dell'energia prodotta viene direttamente autoconsumata dai carichi dell'impianto produttivo, riducendo la disponibilità di surplus per la ricarica del sistema di accumulo.

Nel caso caratterizzato da surplus, invece, la batteria viene utilizzata più frequentemente per gestire l'energia in eccesso e aumentare l'autoconsumo. Tuttavia, a causa dell'elevato livello dei carichi elettrici dell'edificio, l'energia accumulata viene rapidamente riutilizzata, determinando cicli di carica e scarica frequenti ma con una capacità limitata di copertura prolungata del fabbisogno.

Questo comportamento evidenzia come il contributo del sistema di accumulo sia strettamente legato sia al rapporto tra produzione fotovoltaica e domanda energetica, sia alle caratteristiche del carico dell'edificio. In presenza di consumi elevati e continui, infatti, anche in condizioni di surplus la batteria tende a essere rapidamente scaricata, riducendo la sua efficacia come sistema di accumulo a lungo termine e rendendo fondamentale un corretto dimensionamento rispetto alle condizioni operative reali.

### ***4.3. Capacità di generalizzazione e Transfer Learning***

L'analisi di Transfer Learning permette di valutare la capacità delle politiche di controllo basate su apprendimento di adattarsi a edifici con caratteristiche differenti rispetto a quelle utilizzate durante la fase di addestramento.

Nel caso dell'edificio scolastico, il trasferimento zero-shot della politica appresa nel dominio residenziale mostra risultati complessivamente positivi. La temperatura interna rimane infatti all'interno della banda di comfort gran parte del periodo analizzato. Tuttavia, analizzando più nel dettaglio il comportamento del sistema, si osserva che il controllo risulta particolarmente efficace nelle ore centrali della giornata.

Una possibile spiegazione è legata alla diversa schedule di utilizzo dell'edificio scolastico rispetto al caso residenziale. Negli edifici residenziali il consumo energetico tende a essere uniforme nell'arco della giornata, mentre negli edifici scolastici il profilo

di carico è caratterizzato da un aumento significativo nelle ore centrali, quando sono presenti occupanti e sono attivi i principali sistemi dell'edificio. In queste condizioni, il controllore riesce a gestire in modo più stabile il sistema energetico, poiché il fabbisogno termico ed elettrico segue una dinamica più regolare e prevedibile. Al contrario, nelle ore serali e notturne, quando l'edificio è meno utilizzato e il carico base diminuisce, si osservano oscillazioni leggermente più marcate del sistema. Questo suggerisce che la politica appresa nel dominio residenziale riesce ad adattarsi meglio alle condizioni operative che presentano una maggiore somiglianza con quelle incontrate durante la fase di addestramento.

Un altro elemento interessante emerge anche dal confronto con il caso climatico di Livigno analizzato nel caso residenziale. In quello scenario le condizioni climatiche risultano molto più rigide e la produzione fotovoltaica è fortemente ridotta. Questo porta a oscillazioni intorno alla temperatura di set-point e mostra come la capacità di generalizzazione della politica appresa dipenda non solo dalle caratteristiche dell'edificio ma anche dalle condizioni climatiche del dominio di applicazione.

Nei casi industriali, invece, il trasferimento zero-shot mostra prestazioni significativamente peggiori. L'edificio industriale presenta infatti una scala energetica molto diversa rispetto al caso residenziale, con carichi termici ed elettrici più elevati e dinamiche operative differenti.

La politica appresa durante la fase di addestramento risulta quindi poco adatta a rappresentare correttamente il nuovo sistema, portando a un comportamento meno stabile e a prestazioni energetiche peggiori. Le azioni apprese nel dominio residenziale sono infatti calibrate su un intervallo di potenze e su una dinamica termica tipica di edifici di piccola scala, mentre nel caso industriale il sistema è caratterizzato da volumi molto più grandi e da carichi energetici significativamente superiori. La differenza di scala tra i due domini rende quindi difficile l'applicazione diretta della politica appresa, portando a un comportamento meno stabile e a prestazioni energetiche peggiori.

In entrambi gli scenari industriali, il comportamento della politica trasferita risulta coerente con quanto osservato negli altri casi studio esclusivamente per quanto riguarda la gestione del sistema di accumulo. In presenza di energia in eccesso, il controllore tende infatti a replicare correttamente le strategie di carica e scarica della batteria apprese durante la fase di addestramento. Tuttavia, questa coerenza si limita a questo specifico

aspetto. Il comportamento complessivo del sistema, in particolare dal punto di vista termico e della stabilità del controllo, rimane non ottimale.

Questo risultato è coerente con quanto riportato in letteratura sul Transfer Learning applicato al DRL, dove il trasferimento diretto tra domini caratterizzati da dinamiche molto diverse può portare a prestazioni instabili o sub-ottimali. In questi casi, una fase di adattamento tramite fine-tuning risulta generalmente necessaria per permettere alla politica di controllo di adattarsi alle nuove caratteristiche del sistema.

L'introduzione della fase di fine-tuning consente infatti di adattare progressivamente la politica di controllo al nuovo dominio. Nei casi in cui il dominio target presenta caratteristiche simili a quello di addestramento, come nel passaggio dal caso residenziale a quello scolastico, questa fase risulta particolarmente efficace, permettendo di correggere le imperfezioni del trasferimento zero-shot e di ottenere prestazioni comparabili a quelle di un agente addestrato direttamente.

Nei casi industriali, invece, caratterizzati da differenze più marcate in termini di scala energetica e dinamiche operative, il fine-tuning consente un miglioramento rispetto al trasferimento diretto, ma non è sufficiente a raggiungere prestazioni ottimali. In particolare, si osservano ancora criticità nella gestione del sistema di accumulo e nello sfruttamento del surplus energetico. Questo evidenzia come il fine-tuning consenta un adattamento parziale della politica di controllo, migliorando la stabilità del sistema ma non garantendo necessariamente una piena ottimalità operativa.

Il Transfer Learning si conferma quindi uno strumento promettente per ridurre i tempi di addestramento, ma la sua efficacia dipende fortemente dalla similarità tra dominio sorgente e dominio target.

## ***5. Conclusioni e sviluppi futuri***

In questo lavoro è stata analizzata l'applicazione di diverse strategie di controllo per la gestione energetica degli edifici, con l'obiettivo di confrontare le prestazioni dal punto di vista energetico ed economico. L'analisi è stata condotta su tre tipologie di edificio (residenziale, scolastico e industriale), considerando per quest'ultimo due diverse configurazioni energetiche, al fine di valutare il comportamento dei controllori in contesti operativi differenti. In particolare, sono state confrontate strategie basate su Model Predictive Control (MPC), Deep Reinforcement Learning (DRL) e approcci di Transfer Learning (TL).

I risultati delle simulazioni mostrano che tutte le strategie considerate sono in grado di garantire il rispetto delle condizioni di comfort termico. Le differenze principali emergono invece nella gestione del sistema energetico dell'edificio, in particolare per quanto riguarda l'utilizzo del sistema di accumulo e l'interazione con la produzione fotovoltaica.

Il controllore MPC ha evidenziato una maggiore capacità di pianificazione dell'utilizzo delle risorse energetiche. Grazie alla possibilità di prevedere l'evoluzione delle condizioni operative, questo approccio consente infatti di coordinare in modo più efficiente la produzione rinnovabile, il sistema di accumulo e i consumi dell'edificio. Nei casi analizzati, questo si traduce in una riduzione dell'energia importata e dei costi energetici rispetto ai controlli tradizionali. Nel caso residenziale, ad esempio, il costo netto si riduce di circa l'8% rispetto al controllo PID + Rule-Based, mentre nel caso scolastico supera il 15%, evidenziando un miglior sfruttamento delle risorse energetiche disponibili.

Il controllore basato su DRL ha invece mostrato un comportamento adattativo e reattivo alle condizioni operative osservate durante le simulazioni. Le prestazioni risultano complessivamente comparabili a quelle del controllo predittivo in termini di comfort, sebbene con una gestione meno pianificata del sistema di accumulo. Nel caso residenziale, le differenze tra MPC e DRL risultano contenute, mentre nel caso scolastico il controllo predittivo mostra un vantaggio più evidente, con una riduzione dei costi di circa il 14% rispetto al DRL. Tuttavia, il DRL dimostra una buona capacità di adattamento e, ad esempio nel caso residenziale, riesce a gestire in modo efficace l'autoconsumo

dell'energia prodotta localmente, raggiungendo valori superiori al 50%, rispetto a circa il 42% ottenuto con il controllo tradizionale PID +RB.

Un aspetto rilevante riguarda l'influenza delle caratteristiche fisiche dell'edificio. Nei casi residenziale e scolastico, la maggiore capacità di accumulo termico consente di sfruttare l'inerzia dell'edificio come forma di accumulo energetico passivo. Nel caso industriale, invece, la presenza di un sistema basato su UTA con elevati ricambi d'aria riduce significativamente la capacità di accumulo termico, limitando l'efficacia di queste strategie. Questo evidenzia come le prestazioni del controllo dipendano non solo dall'algoritmo utilizzato, ma anche dalla configurazione fisica ed energetica dell'edificio.

Nei casi industriali, le differenze tra le strategie risultano complessivamente meno marcate rispetto agli altri casi studio e riguardano principalmente le prestazioni economiche. Il controllo MPC mantiene comunque le prestazioni migliori, seguito dal controllo PID con logica rule-based e infine dal controllore DRL.

Questo comportamento è legato alle caratteristiche del sistema industriale, in cui il carico elettrico elevato e relativamente costante porta a un elevato livello di autoconsumo diretto, come nel primo caso industriale analizzato. In queste condizioni, il funzionamento dei controllori risulta molto simile, poiché gran parte dell'energia prodotta viene utilizzata direttamente dai carichi e il ruolo del sistema di accumulo è limitato.

Quando si presentano condizioni di surplus, il funzionamento dei controllori rimane sostanzialmente invariato, ma aumenta il ruolo del sistema di accumulo. In questo scenario, la gestione della batteria influisce principalmente sugli scambi con la rete e, di conseguenza, sul costo netto complessivo, senza modificare in modo significativo la dinamica operativa del sistema.

Un ulteriore contributo del lavoro riguarda l'analisi della potenzialità del Transfer Learning. I risultati mostrano che il trasferimento zero-shot può fornire prestazioni soddisfacenti quando i domini presentano caratteristiche relativamente simili, come nel passaggio dal caso residenziale a quello scolastico o nel cambiamento delle condizioni climatiche nel caso residenziale. In questi casi, il controllore riesce a mantenere il comfort e a garantire prestazioni energetiche comparabili, con differenze economiche contenute.

Quando invece le differenze tra i domini diventano più marcate, come nel caso dell'edificio industriale, la politica di controllo appresa non riesce a rappresentare

correttamente le dinamiche del nuovo sistema, portando a un peggioramento delle prestazioni. In particolare, il trasferimento zero-shot comporta un aumento del costo energetico di circa il 19% rispetto al controllo MPC, evidenziando i limiti del trasferimento diretto tra domini molto diversi.

L'introduzione della fase di fine-tuning consente di adattare progressivamente la politica di controllo al nuovo ambiente. Nei casi in cui le differenze tra i domini sono limitate, il fine-tuning si dimostra particolarmente efficace, permettendo di correggere le imperfezioni del trasferimento zero-shot e di raggiungere prestazioni molto vicine a quelle di un agente addestrato direttamente nel dominio target. Nei casi industriali, invece, caratterizzati da dinamiche energetiche significativamente diverse, il fine-tuning migliora le prestazioni rispetto al trasferimento diretto, ma non risulta sufficiente a colmare completamente il divario con un agente addestrato specificamente per il nuovo contesto. Questo evidenzia come l'efficacia del Transfer Learning dipenda fortemente dal grado di similarità tra il dominio sorgente e quello di destinazione.

Il lavoro evidenzia il potenziale delle tecniche di controllo avanzate e degli approcci basati su apprendimento per migliorare la gestione energetica degli edifici. In particolare, l'integrazione tra modelli fisici dell'edificio, sistema di accumulo e algoritmi di controllo intelligente rappresenta una direzione promettente per aumentare l'efficienza energetica e favorire una maggiore integrazione delle fonti rinnovabili.

Tra i possibili sviluppi futuri, sarebbe interessante estendere l'analisi al funzionamento estivo degli edifici, considerando anche il controllo dei sistemi di raffreddamento. Inoltre, ulteriori ricerche potrebbero approfondire l'applicazione del Transfer Learning in scenari caratterizzati da differenze ancora più marcate tra i domini rispetto a quelle analizzate in questo lavoro. Questo permette di valutare la capacità di generalizzazione della politica di controllo basata sul Deep Reinforcement Learning e la loro applicabilità in contesti energetici sempre più diversificati.

## 6. Riferimenti

- [1] T. Ito, K. Iwata, C. McKenzie, S. Urata, “Energy and environment: Editors’ overview”, *Asian Economic Policy Review*, vol. 16, no. 1, pp. 1-21, 2021.
- [2] International Energy Agency (IEA), *Buildings - Energy System Overview*, 2023.
- [3] E. Boronbaev, “Concept of energy-saving architecture: low-energy and low-emission buildings in Kyrgyzstan”, *E3S Web of Conferences*, vol. 405, art. 04039, 2023.
- [4] European Commission, *Energy Performance of Buildings Directive (EPBD)*, 2018/2024.
- [5] A. Karbasforoushha et al., “Building energy optimization: A review”, *Energy and Built Environment*, 2021.
- [6] X. Ma et al., “A review of influencing factors of building energy consumption”, *Energy*, AIMS Press, 2025.
- [7] K. Ogata, *Modern Control Engineering*, 5th ed. Prentice Hall, 2010.
- [8] F. Oldewurtel, D. Sturzenegger, G. Andersson, M. Morari, and R. S. Smith, “Towards a standardized building assessment for model predictive control performance evaluation,” in *Proc. IEEE Conf. Decision and Control (CDC)*, 2013, pp. 3265–3270.
- [9] T. Wei, Y. Wang, and Q. Zhu, “Deep reinforcement learning for building HVAC control,” in *Proc. 54th ACM/EDAC/IEEE Design Automation Conf. (DAC)*, 2017, pp. 1–6.
- [10] J. R. Vázquez-Canteli and Z. Nagy, “Reinforcement learning for demand response: A review of algorithms and modelling techniques”, *Applied Energy*, vol. 235, pp. 1072–1089, 2019.
- [11] Z. Zhang, A. Chong, and K. P. Lam, “Whole building energy model transfer learning for HVAC control,” *Energy and Buildings*, vol. 209, 109699, 2020.
- [12] M. Maivel, A. Ferrantelli, and J. Kurnitski, “Emission losses of radiator, underfloor and air heating: Measurements, simulations and operative temperature corrections,” *Energy and Buildings*, vol. 149, pp. 228–240, 2017.

- [13] J. Weniger, T. Tjaden, and V. Quaschnig, “Sizing of residential PV battery systems,” *HTW Berlin*, 2014.
- [14] U. Acar and Ö. Kaska, “Comparative analysis of PV-battery energy storage system sizing strategies for grid-connected public facilities: A case study from Türkiye,” *Journal of Energy Storage*, 145, 119875, 2025.
- [15] ASHRAE, *ASHRAE Handbook – HVAC Applications*. Atlanta, GA, USA: ASHRAE, 2019.
- [16] G. Serale, M. Fiorentini, A. Capozzoli, D. Bernardini, and A. Bemporad, “Model predictive control (MPC) for enhancing building and HVAC system energy efficiency,” *Energies*, vol. 11, no. 3, 631, 2018.
- [17] P. Peng and Z. Zhang, *Advanced Industrial Control Technology*. Elsevier, 2010.
- [18] M. E. Meral and D. Çelik, “A comprehensive review of control strategies in energy systems,” *Annual Reviews in Control*, 2019.
- [19] S. Bououden et al., “Control Strategy for Time-Delay Systems”, *ISA Transactions*, 2021.
- [20] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, MIT Press, 1998.
- [21] T. P. Lillicrap et al., “Continuous control with deep reinforcement learning,” 2015.
- [22] M. E. Taylor and P. Stone, “Transfer learning for reinforcement learning domains: A survey,” *Journal of Machine Learning Research*, vol. 10, pp. 1633–1685, 2009.