



UNIVERSITA' POLITECNICA DELLE MARCHE

FACOLTA' DI INGEGNERIA

Corso di Laurea triennale INGEGNERIA MECCANICA

**OTTIMIZZAZIONE DI UN MODELLO PER LA MISURA DEL
COMFORT INDOOR:**

SVILUPPO E VALIDAZIONE TRAMITE UN CASO STUDIO REALE

**OPTIMIZATION OF A MODEL FOR INDOOR COMFORT
MEASUREMENT:**

DEVELOPMENT AND VALIDATION THROUGH A REAL CASE STUDY

Relatore: Chiar.mo/a

Prof. Revel Gian Marco

Correlatrici:

Dott.ssa Cipollone Vittoria

Dott.ssa Morresi Nicole

Tesi di Laurea di:

Mazzucchelli Elisa

A.A. 2023/2024

ABSTRACT:

Questo lavoro di tesi si concentra sull'ottimizzazione di un modello semplificato presente in letteratura per il calcolo del comfort termico attraverso l'indice Predicted Mean Vote (PMV). Il modello analizzato, elaborato da C. Buratti, impiega come parametri la temperatura indoor e l'umidità relativa, e prevede la suddivisione dell'isolamento termico del vestiario (*Icl*) in tre intervalli (0.25-0.5, 0.51-1.0, 1.01-1.65); ad ognuno di questi corrispondono i tre coefficienti necessari per il calcolo dell'sPMV. L'obiettivo è di migliorare la performance del modello, includendo una nuova fascia di *Icl* e differenziando i coefficienti tra zone climatiche distinte. La ricerca ha analizzato il dataset "*ASHRAE Global Thermal Comfort Database II*", estraendone tre Paesi europei rappresentanti di tre diverse zone climatiche. Con l'obiettivo di individuare quattro nuove fasce del parametro *Icl*, ovvero l'isolamento termico del vestiario, ne è stata studiata la dipendenza dai fattori ambientali. Il Generalized Additive Model (GAM) è stato utilizzato per ricavare un'equazione che approssima l'isolamento termico in funzione della temperatura indoor, della temperatura outdoor e dell'umidità. Le nuove fasce di *Icl* individuate (0.25-0.5, 0.51-0.8, 0.81-0.99, 1-1.65) hanno consentito di ricavare i coefficienti per il calcolo semplificato del PMV tramite regressione lineare. I risultati ottenuti sono stati confrontati con il PMV standard: l'analisi statistica evidenzia una riduzione percentuale degli errori medi assoluti (MAE) e quadratici (MSE), rispettivamente di 5.31% e 10.20% per la Francia, 8.52% e 15.25% per la Svezia, 17.19% e 31.40% per il Portogallo, suggerendo che il modello proposto può migliorare l'efficienza nel test degli impianti HVAC e nella valutazione del comfort termico degli occupanti.

SOMMARIO:

1. INTRODUZIONE	4
1.1 Obiettivo dello studio.....	6
2. STATO DELL'ARTE	7
2.1 Il comfort termico: definizioni e parametri di riferimento.....	7
2.1.1 Termoregolazione e bilancio energetico.....	9
2.2 Modelli di misura del comfort termico.....	11
2.2.1 PMV E PPD.....	11
2.2.2 Modelli di comfort semplificati.....	14
2.2.3 Limiti dello stato dell'arte.....	16
3. MATERIALI E METODI	17
3.1 Database ASHRAE.....	18
3.2 Analisi e processamento dati.....	18
3.2.1 K-Means clustering.....	24
3.2.2 Generalized Additive Model (GAM).....	31
3.2.3 Definizione del nuovo modello.....	40
3.3 Valutazione della performance.....	41
4. RISULTATI	41
4.1 Metriche di validazione e coefficienti ottimizzati.....	42
4.2 sPMV ottimizzato.....	43
5. DISCUSSIONE	47
6. CONCLUSIONI	48
7. BIBLIOGRAFIA	50

1. INTRODUZIONE

Il comfort termico [1] è concetto centrale nel campo dell'ingegneria meccanica applicata agli edifici e agli impianti di climatizzazione, con un impatto significativo sia sulla salute e sul benessere delle persone, sia sull'efficienza energetica dei sistemi HVAC (Heating, Ventilation, and Air Conditioning). In particolare, il comfort termico si riferisce a quella condizione in cui un individuo esprime soddisfazione rispetto all'ambiente termico circostante [2], un concetto che, pur essendo altamente soggettivo, può essere quantificato attraverso parametri e modelli fisico-matematici che ne permettono una valutazione più oggettiva e replicabile.

Infatti, è fondamentale fornire un ambiente percepito dalla maggioranza degli occupanti come termicamente neutro. Il concetto è ancor più importante se si parla di luoghi di lavoro come uffici e scuole, poiché il comfort termico ha effetti tangibili sulla produttività individuale e sulla salute.

Il corpo umano possiede un sistema di regolazione della temperatura molto complesso, che riceve segnali dalla pelle ed è regolato dall'ipotalamo. [3] Si percepisce neutralità termica se i segnali inviati dall'esterno e dall'interno si equivalgono, e se si verifica il bilancio energetico tra il calore prodotto dal metabolismo e quello dissipato dal corpo. Un ambiente termico non confortevole può generare stress termico, portando a disagio e, in alcuni casi, a problemi di salute come affaticamento, mal di testa o, nei casi estremi, colpi di calore o ipotermia. Nelle moderne progettazioni di edifici, il raggiungimento di un comfort termico ottimale è un obiettivo primario, da bilanciare con la necessità di minimizzare i consumi energetici, garantendo al contempo la sostenibilità delle costruzioni.

Uno dei modelli più utilizzati per misurare il comfort termico è l'indice PMV (Predicted Mean Vote), sviluppato da Fanger [1] nel 1970. Il PMV rappresenta una formalizzazione della percezione soggettiva del comfort termico, basata su dati empirici raccolti in studi su un ampio numero di soggetti. Questo indice, che verrà approfondito nei capitoli successivi, permette di prevedere la media delle valutazioni di comfort termico di un gruppo di persone, su una scala che va da -3 (molto freddo) a +3 (molto caldo), con 0 che rappresenta il comfort termico neutro. Il PMV tiene conto di sei parametri: temperatura dell'aria, temperatura media radiante, umidità relativa, velocità dell'aria, attività metabolica e isolamento termico dell'abbigliamento.

Il comfort termico è anche strettamente regolato da normative internazionali che definiscono standard di progettazione e gestione dei sistemi HVAC, nonché criteri per la misurazione e

valutazione del benessere termico negli edifici. Tra le norme più importanti a livello globale vi sono la ISO 7730 [2] e gli standard ASHRAE.

La ISO 7730 stabilisce metodi per la determinazione del comfort termico e specifica i limiti per gli ambienti interni, facendo riferimento proprio all'indice PMV e all'indice PPD (Predicted Percentage of Dissatisfied), che stima la percentuale di persone insoddisfatte delle condizioni termiche. Questa norma si applica non solo agli edifici civili, ma anche ad ambienti particolari come ospedali, uffici e stabilimenti industriali.

D'altro canto, l'ASHRAE Standard 55 [4], pubblicato dalla American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE), è uno dei riferimenti chiave per la progettazione degli impianti HVAC negli Stati Uniti e in gran parte del mondo. Questo standard definisce i requisiti per il comfort termico negli edifici, fornendo linee guida pratiche per la progettazione e la gestione di sistemi di climatizzazione.

Secondo studi basati su queste normative, per ottenere un comfort termico adeguato, la temperatura dell'aria in ambienti interni dovrebbe essere compresa tra i 20°C e i 24°C durante l'inverno e tra i 23°C e i 26°C in estate, con un'umidità relativa compresa tra il 30% e il 60%. Questi valori sono indicativi e possono variare a seconda dell'attività svolta all'interno dell'edificio e del tipo di abbigliamento indossato dagli occupanti. Per quanto riguarda la velocità dell'aria, essa dovrebbe essere mantenuta al di sotto di 0,2 m/s per evitare la sensazione di correnti d'aria fastidiose.

L'applicazione pratica del modello PMV risulta complessa in contesti reali e quotidiani a causa dei numerosi parametri ambientali richiesti per la sua misurazione. Per questo, sono stati elaborati modelli semplificati per il calcolo del PMV, utili per stimare il comfort termico sulla base di un minor numero di parametri. Tra quelli che verranno descritti nel capitolo seguente, è importante citare quello di C. Buratti [5] pubblicato nel 2012, il quale costituisce punto di partenza per questo studio.

I capitoli successivi analizzeranno in dettaglio i principi fisici e matematici alla base del PMV e del PPD, nonché gli strumenti disponibili per la progettazione degli impianti HVAC. Porremo attenzione, in particolare, sui modelli per il calcolo del PMV semplificato, poiché costituisce l'argomento centrale dello studio.

1.1 Obiettivo dello studio

Questo lavoro di tesi intende modificare il modello semplificato proposto da Buratti et al. nella stima del PMV, avendo come riferimento il PMV standard calcolato secondo normativa ISO 7730. Per apportare miglioramenti, si propone di tenere conto della diversità tra zone climatiche in termini di adattamento al clima degli individui, e di suddividere il range di Icl in quattro fasce. Il metodo ottenuto è stato validato su tre diverse zone climatiche, utilizzando per ognuna di esse quattro set di coefficienti a , b , c per il calcolo del PMV semplificato.

Partendo, dunque, dall'attuale stato dell'arte e valutandone i limiti, lo studio si è incentrato sull'analisi delle interazioni tra i parametri ambientali per individuare una relazione matematica che esprimesse Icl , da verificare sui dataset a disposizione. In questo modo, sono state individuate quattro fasce di clo , unità di misura che quantifica l'isolamento termico degli indumenti, ognuna con i rispettivi coefficienti a , b , c utilizzati per il calcolo del PMV semplificato. Il modello di Buratti segue la stessa logica, ma prevede solo tre fasce di Icl , corrispondenti a stagione fredda, stagione calda e mezza stagione. Nella discussione finale dei risultati verrà confrontato con il nuovo modello ottimizzato e dimostrato il miglioramento mediante le metriche Mean Absolute Error (MAE) e Mean Squared Error (MSE). Per offrire un modello indipendente dalla stagionalità delle condizioni ambientali esterne, la ricerca si è orientata sulla dipendenza di Icl da tre parametri ambientali: temperatura interna, temperatura esterna e umidità relativa interna.

2. STATO DELL'ARTE

2.1 Il comfort termico: definizioni e parametri di riferimento

Il concetto di comfort termico è definito come *"quella condizione mentale che esprime soddisfazione riguardo l'ambiente termico"* dall' ASHRAE Standard 55 [4]. Questa definizione implica una componente soggettiva intrinseca, poiché la percezione del comfort varia tra gli individui in base a fattori personali e ambientali. Tuttavia, negli ultimi decenni, sono stati sviluppati modelli e metodologie per quantificare tale percezione in termini oggettivi, così da permettere la progettazione di sistemi ambientali e HVAC ottimali. La gestione del comfort termico negli edifici, infatti, è fondamentale non solo per il benessere degli occupanti, ma anche per il controllo dei consumi energetici e l'efficienza dei sistemi di climatizzazione.

Per definire il comfort termico, è necessario considerare una combinazione di fattori ambientali e personali. Da un lato, abbiamo i parametri ambientali misurabili come la temperatura dell'aria, l'umidità relativa, la velocità dell'aria e la radiazione termica. Dall'altro, vi sono le variabili individuali, come il metabolismo e il livello di isolamento termico fornito dall'abbigliamento, che variano da persona a persona. Questo rende la valutazione del comfort termico un campo complesso e multidimensionale, in cui devono essere considerati sia parametri fisici sia percezioni soggettive.

Parametri ambientali

- **Temperatura dell'aria:** Si riferisce alla temperatura dell'aria che circonda gli individui. Un intervallo tipico di comfort termico varia tra i 20°C e i 26°C, con variazioni in funzione delle condizioni esterne e delle attività svolte. La misurazione della temperatura dell'aria segue le indicazioni della norma UNI EN ISO 7726, che definisce i requisiti per gli strumenti di misura.
- **Temperatura media radiante:** Rappresenta la temperatura media percepita dalle superfici circostanti (muri, finestre, soffitto, pavimento). La radiazione termica può avere un impatto significativo sul comfort percepito, soprattutto in presenza di superfici fredde o calde. Anche in questo caso, la misurazione deve essere effettuata secondo la norma UNI EN ISO 7726.
- **Umidità relativa:** È la percentuale di vapore acqueo presente nell'aria rispetto alla quantità massima che potrebbe essere contenuta a quella temperatura. L'umidità relativa ideale per il comfort umano varia tra il 30% e il 60%. Valori troppo bassi possono provocare secchezza delle mucose e irritazione, mentre livelli elevati possono causare

una riduzione dell'efficienza termica di evaporazione del sudore, aumentando la sensazione di calore.

- Velocità dell'aria: La velocità dell'aria influenza la percezione del comfort poiché un'aria troppo ferma può causare una sensazione di calore stagnante, mentre correnti d'aria eccessive possono generare disagio. In base agli standard internazionali, la velocità dell'aria dovrebbe essere mantenuta inferiore a 0,2 m/s in condizioni di comfort normale (UNI EN ISO 7726).

Parametri personali

- Metabolismo (*Met*): L'attività metabolica di una persona influisce significativamente sulla quantità di calore generato dal corpo. L'ISO 8996 definisce le metodologie per la determinazione del livello di attività metabolica, espressa in "met", dove 1 met corrisponde al consumo energetico di un individuo a riposo (58.2 W/m²). Un adulto normale ha una superficie corporea di 1,7 m²; quindi, una persona in comfort termico con un livello di attività di 1 Met perderà approssimativamente 100W di calore. Il nostro metabolismo è al suo livello più basso quando dormiamo (0,8 Met) e raggiunge il livello più alto durante le attività sportive, dove frequentemente si arriva a 10 Met. Un tasso Met comunemente usato è 1,2, che corrisponde a un normale lavoro da seduti in ufficio.
- Isolamento termico dell'abbigliamento (*Icl*): L'isolamento termico fornito dai vestiti è un altro fattore cruciale. Questo parametro viene misurato in "Clo", dove 1 Clo equivale all'isolamento termico di un completo da uomo in lana. La norma ISO 9920 fornisce indicazioni dettagliate per la valutazione dell'isolamento termico fornito dall'abbigliamento in diverse condizioni ambientali e di attività.

L'insieme dei parametri sopra citati viene regolamentato da una serie di norme tecniche internazionali, che forniscono metodologie e limiti operativi per la valutazione del comfort termico e la progettazione dei sistemi HVAC:

- ASHRAE Standard 55 [4]: La norma ASHRAE (American Society of Heating, Refrigerating and Air-Conditioning Engineers) Standard 55 definisce le condizioni ambientali che consentono di ottenere un comfort termico accettabile per la maggior parte degli occupanti. Lo standard si basa sull'utilizzo di modelli predittivi come il PMV e il PPD (Predicted Percentage of Dissatisfied), che verranno trattati in dettaglio nella sezione successiva. L'ASHRAE 55 prevede anche un approccio adattivo per ambienti

non climatizzati, permettendo una maggiore flessibilità in funzione delle condizioni climatiche esterne e delle abitudini degli occupanti.

- ISO 7730 [2]: La ISO 7730 è lo standard internazionale per la determinazione del comfort termico, basato su calcoli del PMV e del PPD. Essa descrive metodologie per valutare l'insieme dei parametri sopra elencati e garantire che le condizioni termiche in un ambiente siano soddisfacenti per almeno l'80% degli occupanti, in linea con gli obiettivi definiti dal PPD.
- ISO 8996 [6] e ISO 9920 [7]: Queste norme trattano la valutazione dei parametri personali. L'ISO 8996 fornisce una classificazione dettagliata dell'attività metabolica, mentre l'ISO 9920 riguarda il calcolo dell'isolamento termico dell'abbigliamento. La corretta combinazione di questi parametri con quelli ambientali è cruciale per garantire il comfort termico.
- UNI EN ISO 7726 [8]: Questa norma stabilisce le specifiche per gli strumenti utilizzati per la misurazione dei parametri ambientali, garantendo che i dati raccolti siano accurati e utilizzabili per il calcolo del comfort termico.

2.1.1 Termoregolazione e bilancio energetico

La termoregolazione [9] è il processo attraverso il quale gli organismi mantenendo la temperatura corporea all'interno di un intervallo ottimale, garantendo così il corretto funzionamento delle funzioni fisiologiche. Negli esseri umani, questo meccanismo è particolarmente sofisticato, coinvolgendo vari sistemi fisiologici e comportamentali per rispondere alle variazioni ambientali. Come detto nel primo capitolo, il corpo umano è dotato di un centro di termoregolazione situato nell'ipotalamo, che agisce come un termostato biologico. Questo centro riceve informazioni da recettori termici distribuiti in tutto il corpo e nell'ambiente circostante, e coordina il sistema nervoso simpatico e parasimpatico per mantenere la temperatura corporea. Il sistema nervoso simpatico regola le risposte attive allo stress termico inducendo, in condizioni di freddo, vasocostrizione periferica, ovvero la riduzione del flusso sanguigno alle estremità per conservare il calore nel nucleo corporeo, e la termogenesi, che aumenta la produzione di calore metabolico. In condizioni di caldo, invece, il sistema simpatico regola vasodilatazione cutanea, che aumenta il flusso sanguigno alla superficie della pelle per favorire la dispersione del calore, e la sudorazione, per perdere calore attraverso l'evaporazione. [9]

Il sistema nervoso parasimpatico, sebbene meno direttamente coinvolto nella gestione delle situazioni di stress termico acuto, gioca un ruolo importante nel mantenimento di uno stato di

riposo e recupero. Durante periodi di stabilità termica, il parasimpatico contribuisce al mantenimento di un metabolismo basale efficiente, evitando sprechi energetici e permettendo al corpo di rimanere in uno stato di comfort a lungo termine.

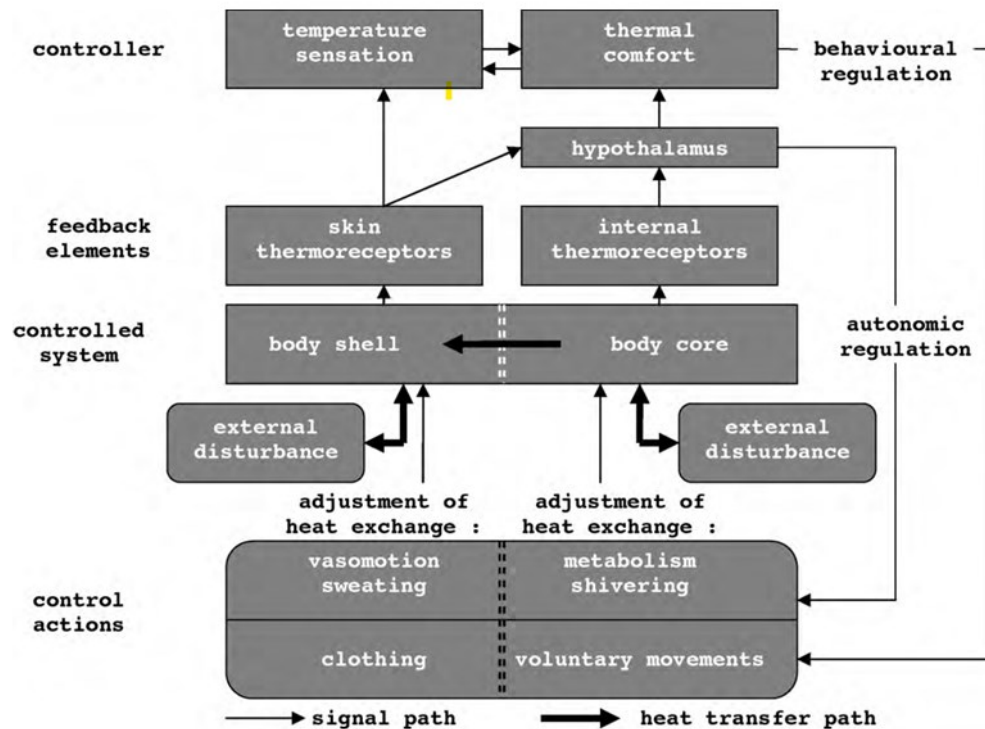


Figura 1: diagramma della regolazione autonoma della temperatura [10]

Il bilancio energetico [10] è un concetto fondamentale nella termoregolazione, poiché rappresenta la relazione tra l'energia assunta attraverso l'alimentazione e l'energia spesa attraverso il metabolismo e l'attività fisica. In condizioni di comfort termico, l'energia assunta e l'energia spesa si bilanciano, mantenendo la temperatura corporea all'interno di un intervallo ottimale. L'equazione di bilancio è così composta:

$$M = C + R + E + P \quad (1)$$

dove:

M = potenza generata dall'attività metabolica [W]

P = potenza meccanica scambiata con l'esterno [W]

C = flusso termico fornito all'esterno per convezione [W]

R = flusso termico all'esterno per irraggiamento corpo-oggetti esterni [W]

E = flusso termico impegnato per evaporare acqua [W]

Le condizioni ambientali, come temperatura, umidità e velocità del vento, hanno un impatto significativo sulla termoregolazione e sul bilancio energetico. Ambienti caldi e umidi possono ostacolare la capacità del corpo di dissipare il calore, portando a un aumento della temperatura corporea e a un maggiore rischio di colpi di calore. Al contrario, in ambienti freddi, il corpo deve aumentare la termogenesi per mantenere la temperatura corporea, richiedendo un maggiore apporto energetico.

2.2 Modelli di misura del comfort termico

Ad oggi per i modelli relativi alla misura del comfort termico esistono due metodi diversi: l'approccio del bilancio termico e l'approccio adattivo. [10] Il primo, anche noto come approccio razionale, è utilizzato per gestire i dati delle camere climatiche, mentre il secondo è generalmente collegato agli studi sul campo. Fondamentale, all'interno del primo gruppo, è il modello di Fanger [1]: il comfort è determinato dalla soddisfazione di tutto il corpo, considerando anche il disagio locale. Il calcolo prevede l'impiego di sei fattori e gli indici principali sono il Predicted Mean Vote (PMV) e la Percentage of People Dissatisfied (PPD).

2.2.1 PMV E PPD

L'equazione completa di Fanger associa le condizioni termiche alla scala di sensazione termica ASHRAE [4] a sette punti (indice PMV).

+ 3	Hot
+ 2	Warm
+ 1	Slightly warm
0	Neutral
- 1	Slightly cool
-2	Cool
- 3	Cold

Figura 2: scala a sette punti - ISO 7730 [2]

Il PMV è determinato dai sei parametri, di cui quattro ambientali e due personali, già elencati in precedenza. Nella normativa UNI EN ISO 7730 [2], viene espresso in questo modo:

$$PMV = [0,303 \cdot \exp(-0,036 \cdot M) + 0,028] \cdot$$

$$\left\{ \begin{array}{l} (M - W) - 3,05 \cdot 10^{-3} \cdot [5733 - 6,99 \cdot (M - W) - p_a] - 0,42 \cdot [(M - W) - 58,15] \\ -1,7 \cdot 10^{-5} \cdot M \cdot (5867 - p_a) - 0,0014 \cdot M \cdot (34 - t_a) \\ -3,96 \cdot 10^{-8} \cdot f_{cl} \cdot [(t_{cl} + 273)^4 - (\bar{t}_r + 273)^4] - f_{cl} \cdot h_c \cdot (t_{cl} - t_a) \end{array} \right\}$$

$$t_{cl} = 35,7 - 0,028 \cdot (M - W) - I_{cl} \cdot \left\{ 3,96 \cdot 10^{-8} \cdot f_{cl} \cdot [(t_{cl} + 273)^4 - (\bar{t}_r + 273)^4] + f_{cl} \cdot h_c \cdot (t_{cl} - t_a) \right\}$$

$$h_c = \begin{cases} 2,38 \cdot |t_{cl} - t_a|^{0,25} & \text{for } 2,38 \cdot |t_{cl} - t_a|^{0,25} > 12,1 \cdot \sqrt{v_{ar}} \\ 12,1 \cdot \sqrt{v_{ar}} & \text{for } 2,38 \cdot |t_{cl} - t_a|^{0,25} < 12,1 \cdot \sqrt{v_{ar}} \end{cases}$$

$$f_{cl} = \begin{cases} 1,00 + 1,290 I_{cl} & \text{for } I_{cl} \leq 0,078 \text{ m}^2 \cdot \text{K/W} \\ 1,05 + 0,645 I_{cl} & \text{for } I_{cl} > 0,078 \text{ m}^2 \cdot \text{K/W} \end{cases}$$

Figura 3: equazione PMV standard - ISO 7730 [2]

I parametri impiegati sono i seguenti, affiancati dagli intervalli previsti dalla normativa.

Tabella 1: intervalli accettabili da normativa ISO 7730 [11]

PARAMETRO	INTERVALLO
Metabolismo M	$46W / m^2 < M < 232W / m^2$
Isolamento termico abbigliamento I_{cl}	$0 < I_{cl} < 0.31m^2K / W$ (0 – 2clo)
Temperatura dell'aria t_a	$10^\circ C < t_a < 30^\circ C$
Temperatura media radiante \bar{t}_r	$10^\circ C < \bar{t}_r < 40^\circ C$
Velocità relativa dell'aria v_{ar}	$0 < v_{ar} < 1m / s$
Pressione parziale di vapore acqueo p_a	$0 < p_a < 2700Pa$

La ISO 7730 individua 3 classi di comfort termico:

1. CLASSE A: $-0.2 \leq PMV \leq 0.2$;
2. CLASSE B: $-0.5 \leq PMV \leq 0.5$;
3. CLASSE C: $-0.7 \leq PMV \leq 0.7$.

Strettamente legato al PMV è la PPD (Predicted Percentage of Dissatisfied), che esprime la percentuale prevista di utenti insoddisfatti

$$PPD = 100 - 95 \cdot \exp(-0,033\ 53 \cdot PMV^4 - 0,217\ 9 \cdot PMV^2) \quad (2)$$

Il grafico che esprime la relazione tra PMV e PPD è il seguente: si nota che il valore del PPD non scende mai sotto al 5%.

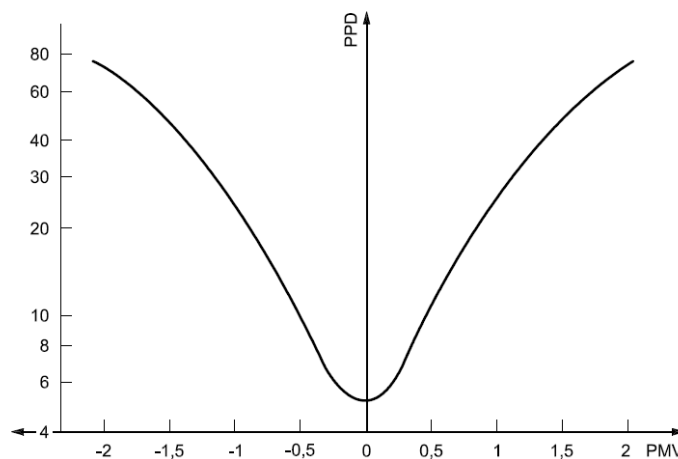


Figura 4: grafico PMV - PPD

Il riscontro con ambienti reali, a differenza delle prove in laboratorio, fa emergere delle criticità nel calcolo del PMV [12]:

1. Questo metodo necessita un ampio campione di utenti, poiché rischia di perdere accuratezza su gruppi più ristretti.
2. Nonostante le normative che specifichino le modalità di misurazione, risulta difficile ottenere dati esatti per i parametri ambientali. Infatti, l'indice metabolico e il *clo* possono variare nel corso della giornata.
3. In edifici comuni, non sempre sono disponibili i sensori in grado di rilevare tutti questi parametri. In mancanza di dati di input, l'unica soluzione può essere un approccio semplificato.

2.2.2 Modelli di comfort semplificati

Il primo a formulare un'espressione di calcolo per il PMV semplificato fu Frederick Rohles (Institute of Environmental Research della State University of Kansas) nel 1971, con l'obiettivo di ottenere un modello che richiedesse parametri facilmente campionabili nell'ambiente: temperatura dell'aria e umidità relativa. [5]

Rohles ha condotto esperimenti su 1600 studenti universitari (800 uomini e 800 donne) in camere climatiche, chiedendo loro di valutare le sensazioni in riferimento a una scala da -3 a +3. La formula semplificata del PMV risultante da questo studio è:

$$PMV = a \cdot T_a + b \cdot P_v - c \quad (3)$$

dove:

- T_a è la temperatura dell'aria ($^{\circ}C$)
- P_v è la pressione di vapore acqueo (kPa) legata all'umidità relativa RH dalla seguente relazione:

$$P_v = (RH \cdot P_s) / 100 \quad (4)$$

in cui P_s è la pressione di vapore saturo

- a, b, c sono coefficienti sperimentali

Tabella 2: coefficienti di Rohles

a	b	c
0.245	0.248	6.475

Tuttavia, il modello di Rohles rimane limitato a ipotesi stringenti e stazionarie: è circoscritto solo all'attività sedentaria (tasso metabolico = 1,2 met) e a un intervallo fisso di isolamento termico degli indumenti ($I_{cl} = 0,6$ clo). Inoltre, è valido solo per valori di velocità dell'aria minori di 0.2 m/s ed è basato su esposizioni temporali di 1, 2 e 3 ore. Al contrario, le condizioni di un ambiente reale sono dinamiche.

Un approccio semplificato alla valutazione del comfort termico è necessario a causa del crescente interesse per le prestazioni energetiche degli edifici, legate alle caratteristiche dell'involucro e al comfort degli occupanti. Il vantaggio di un modello semplificato è quello di stabilire un legame tra l'approccio del bilancio termico stazionario e l'approccio adattivo [13]: i parametri misurati sono ridotti, gli strumenti sono meno invasivi ed è anche possibile verificare le condizioni di comfort fornite dai sistemi HVAC sul campo, al fine di rispettare il comportamento e le abitudini degli occupanti. I modelli adattivi di comfort, infatti, tengono

conto della capacità degli utenti di adattarsi alle condizioni ambientali. [14] Dunque, in quest'ottica, non occorre mantenere i parametri ambientali costantemente fissi come nel modello razionale di Fanger, a beneficio del consumo energetico. Pertanto, un modello semplificato per il calcolo del PMV basato solo sulle misure di temperatura dell'aria e umidità relativa potrebbe essere molto utile nel test finale dell'impianto. [10]

Tra i successivi modelli semplificati, è stato preso come punto di partenza il modello di Buratti [5], che ambisce a migliorare il modello di Rohles, ampliando il range dei casi possibili e, dunque, cercando una prospettiva più dinamica, mantenendo la semplicità dell'equazione del PMV. Il lavoro citato ha analizzato dati disponibili presso l'Università di Perugia e di Pavia, consistenti in oltre 450 misurazioni strumentali e 1100 risultati di questionari.

I dati sono riferiti a edifici con impianti HVAC in diverse stagioni e in aule e uffici open space, dove gli occupanti potevano essere considerati in condizioni sedentarie. È stata registrata una variabilità dell'isolamento termico degli indumenti degli occupanti nell'intervallo 0,25-1,65, diviso in tre range di I_{cl} al fine di adattare il modello all'intera gamma di isolamento termico. Sono stati ottenuti nove diversi gruppi di coefficienti e diagrammi per gli intervalli 0,25-0,50 clo, 0,51-1,00 clo e 1,01-1,65 clo e per maschio, femmina ed entrambi i sessi.

Table 1
Number of surveys, Rohles and new coefficients a , b , and c , and percentage of situations in which data of the two models were more reliable, in the different I_{cl} ranges.

	I_{cl} (clo)	Sex	No. of surveys	a	b	c	% More reliable results	
							Rohles	New model
Rohles (exposure time 1 h [3,21])	0.60	Male	–	0.220	0.233	5.673	–	–
		Female	–	0.272	0.248	7.245	–	–
		Both	–	0.245	0.248	6.475	–	–
New coefficients	0.25–0.50	Male	31	0.2630	0.3027	6.8066	23	77
		Female	80	0.2658	0.1072	6.7232	28	72
		Both	111	0.2803	0.1717	7.1383	37	63
	0.51–1.00	Male	142	0.1162	–0.1338	2.2011	12	88
		Female	97	0.2424	0.0614	5.5869	10	90
		Both	239	0.1383	0.0269	3.0190	13	87
	1.01–1.65	Male	70	0.1500	–0.1668	2.5121	4	96
		Female	43	0.1494	–0.1056	2.6408	0	100
		Both	113	0.1478	–0.1371	2.5239	0	100
Total	–	–	463	–	–	–	–	–

Figura 5: tabella coefficienti di Buratti et al. [5]

Nella Figura 5 si può vedere il confronto tra il modello di Rohles e il modello di Buratti. I risultati hanno mostrato che l'applicazione del modello di Rohles per valori di I_{cl} superiori a 0,6 clo non è adatta. Inoltre, Buratti afferma che sia più corretto utilizzare equazioni diverse per uomini e donne solo nell'intervallo 0,51-1,00 clo, mentre negli altri intervalli di I_{cl} dovrebbe essere utilizzata l'equazione ottenuta per entrambi. Infine, sono stati calcolati i Valori Medi (MV) delle differenze tra dati sperimentali e calcolati ($PMV_{meas} - PMV_{calc}$) e la relativa

Deviazione Standard (SD). Il valore medio della Deviazione Standard è 0.22 applicando il nuovo modello; pertanto, la variazione del PMV semplificato è, in media, ± 0.2 .

2.2.3 Limiti dello stato dell'arte

Come già accennato in precedenza, il modello standard per il calcolo del PMV negli anni ha presentato delle difficoltà implementative dovute dalla dipendenza della funzione da sei parametri che richiede numerosi sensori di acquisizione, non presenti in edifici comuni. Inoltre, risulta meno accurato per misure relative a un gruppo più ristretto di utenti.

Questi limiti hanno spinto Rohles a formulare un modello in grado di calcolare il PMV semplificato con l'Equazione 3. [5] Dunque, da sei parametri si passa a due, ovvero temperatura indoor e umidità relativa. Tuttavia, le ipotesi necessarie non sono molto realistiche:

- $I_{cl} = 0.6$ clo;
- Indice metabolico = 1.2 met;
- Velocità dell'aria (v_a) < 0.2 m/s;
- Tempi di esposizione pari a 1, 2, o 3 h.

Buratti, in continuità con il lavoro di Rohles, ha esteso il modello semplificato a un maggior numero di casi, con I_{cl} compreso tra 0.25 e 1.65.

D'altra parte, il modello semplificato proposto da Buratti et al. non sempre risulta accurato in termini di errore medio (MAE= 0.22), soprattutto per la fascia di I_{cl} intermedia ($0.51 < clo < 1.00$). Un altro aspetto che il modello di Buratti non considera è la diversità di adattamento alle condizioni ambientali tra una zona climatica e l'altra: è dimostrato, infatti, che individui provenienti da regioni climatiche differenti presentino differenze di adattamento all'ambiente [13]. Lo studio presentato in questo articolo intende ottimizzare il modello di Buratti tenendo conto dei limiti appena discussi.

3. MATERIALI E METODI

La prima fase della metodologia sviluppata in questo studio è un'analisi preliminare del Database online ASHRAE. A partire da questo, si è effettuato il processamento e il filtraggio dei dati, mantenendo solo quelli relativi ai Paesi Europei. Dopo un'analisi riguardo l'andamento dei parametri ambientali, sono stati scelti alcuni Paesi campione, ognuno rappresentante di una diversa zona climatica (clima oceanico temperato, mediterraneo temperato e clima freddo continentale), sulla base della qualità e della quantità dei dati.

Da qui, l'indagine si è rivolta alla ricerca di un metodo per individuare quattro fasce di isolamento termico, corrispondenti a precisi intervalli di temperatura outdoor che tipicamente variano in diverse zone climatiche. Come vedremo, questo passaggio è stato caratterizzato da notevole difficoltà poiché la distribuzione dei dati risulta molto complessa, a dimostrazione del carattere intrinsecamente soggettivo dell'*Icl*.

Il clustering [15] con K-Means costituisce un primo approccio che ha indirizzato la ricerca verso una visione più globale e complessa dell'interazione tra parametri ambientali. Per analizzare l'andamento di *Icl* e le sue caratteristiche è stato impiegato un algoritmo di apprendimento supervisionato, chiamato GAM [16] (Generalized Additive Model), che verrà spiegato dettagliatamente in seguito.

Quello che è emerso dall'indagine è un'equazione di calcolo che, a partire da temperatura indoor, umidità relativa e temperatura outdoor, permette di ottenere un valore di *Icl* approssimato per la misura in questione. Una volta trovato il modo per suddividere il range in quattro intervalli, dunque, è stato possibile ricavare i nuovi coefficienti a, b, c per ogni fascia di *clo* tramite regressione lineare.

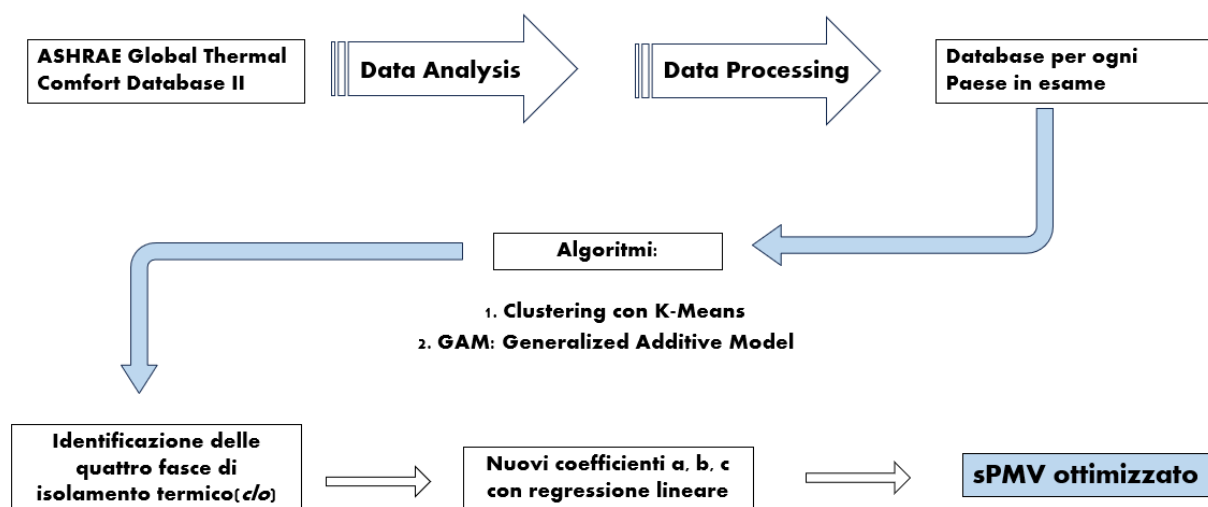


Figura 6: schema delle fasi del lavoro

3.1 Database ASHRAE

Il database si può trovare nel sito GitHub [17] con il nome: “*ASHRAE Global Thermal Comfort Database II*”, ed è fornito dall’ente ufficiale ASHRAE. Contiene misurazioni ambientali provenienti da tutto il mondo, insieme a valutazioni personali degli utenti. Il file è in formato CSV (Comma-Separated Values), ed è stato analizzato nell’IDE (ambiente di sviluppo integrato) Spyder, che utilizza Python come linguaggio di programmazione.

Si compone di due parti:

- “db_measurements_v2.1.0”, che contiene i dati relativi alle misurazioni ambientali

building_id	timestamp	season	subject	age	gender	ht	wt	ta	ta_h	ta_m	ta_l	top	tr	tg	tg_h	tg_m	tg_l	rh	vel	vel_h	vel_m	vel_l	met	clo	ctivity_1	ctivity_2	ctivity_3	ctivity_6	ial_sens	thermal_acc
1	1995-05-1...	winter	10	nan	female	nan	nan	22.3	nan	22.3	nan	22.6...	22.95	nan	nan	nan	nan	61	0.03	nan	0.03	nan	1.70648	0.95	nan	nan	nan	nan	2	acceptable
1	1995-05-1...	winter	14	nan	female	nan	nan	23	nan	23	nan	23.5...	24.1...	nan	nan	nan	nan	59	0.08	nan	0.08	nan	1.10922	1.07	nan	nan	nan	nan	2	unacceptable
1	1995-05-1...	winter	15	nan	female	nan	nan	22	nan	22	nan	22.0...	22.0...	nan	nan	nan	nan	61	0.04	nan	0.04	nan	1.2116	0.88	nan	nan	nan	nan	0	acceptable
1	1995-05-1...	winter	17	nan	female	nan	nan	23.2	nan	23.2	nan	23.7...	24.3...	nan	nan	nan	nan	58	0.03	nan	0.03	nan	1.77474	0.59	nan	nan	nan	nan	0	acceptable
1	1995-05-1...	winter	21	nan	male	nan	nan	22.2	nan	22.2	nan	22.6...	23.0...	nan	nan	nan	nan	62	0.15	nan	0.15	nan	1.48464	0.83	nan	nan	nan	nan	0	acceptable
1	1995-05-1...	winter	22	nan	male	nan	nan	21.2	nan	21.2	nan	21.1...	21.0...	nan	nan	nan	nan	65	0.15	nan	0.15	nan	1.50171	0.66	nan	nan	nan	nan	-1	acceptable

Figura 7: db_measurements. Dataset estratto dal DB ASHRAE

- “db_metadata”, con informazioni descrittive utili per identificare gli edifici e gli utenti

building_id	timestamp	season	subject	age	gender	ht	wt	ta	ta_h	ta_m	ta_l	top	tr	tg	tg_h	tg_m	tg_l	rh	vel	vel_h	vel_m	vel_l	met	clo	ctivity_1	ctivity_2	ctivity_3	ctivity_6	ial_sens	thermal_acc
1	1995-05-1...	winter	10	nan	female	nan	nan	22.3	nan	22.3	nan	22.6...	22.95	nan	nan	nan	nan	61	0.03	nan	0.03	nan	1.70648	0.95	nan	nan	nan	nan	2	acceptable
1	1995-05-1...	winter	14	nan	female	nan	nan	23	nan	23	nan	23.5...	24.1...	nan	nan	nan	nan	59	0.08	nan	0.08	nan	1.10922	1.07	nan	nan	nan	nan	2	unacceptable
1	1995-05-1...	winter	15	nan	female	nan	nan	22	nan	22	nan	22.0...	22.0...	nan	nan	nan	nan	61	0.04	nan	0.04	nan	1.2116	0.88	nan	nan	nan	nan	0	acceptable
1	1995-05-1...	winter	17	nan	female	nan	nan	23.2	nan	23.2	nan	23.7...	24.3...	nan	nan	nan	nan	58	0.03	nan	0.03	nan	1.77474	0.59	nan	nan	nan	nan	0	acceptable
1	1995-05-1...	winter	21	nan	male	nan	nan	22.2	nan	22.2	nan	22.6...	23.0...	nan	nan	nan	nan	62	0.15	nan	0.15	nan	1.48464	0.83	nan	nan	nan	nan	0	acceptable
1	1995-05-1...	winter	22	nan	male	nan	nan	21.2	nan	21.2	nan	21.1...	21.0...	nan	nan	nan	nan	65	0.15	nan	0.15	nan	1.50171	0.66	nan	nan	nan	nan	-1	acceptable

Figura 8: db_metadata. Dataset estratto dal DB ASHRAE

3.2 Analisi e processamento dati

Tramite il *building_id* si è estratto un dataframe completo dei dati europei, associando le misure del db_measurements_v2.1.0. ad ogni edificio e filtrando le colonne utili. I parametri selezionati sono quindici:

Index	building_id	timestamp	season	ta	clo	rh	t_out	pmv	pmv_ce	region	country	city	lat	lon	climate
0	740	1998-07-31	summer	28.4	0.52	33.9	25.36	1.09	0.74	europe	france	lyon	45.764	4.83566	temperate oceanic
1	740	1998-09-11	summer	26.2	0.56	48.6	16.8	0.58	0.53	europe	france	lyon	45.764	4.83566	temperate oceanic
2	740	1998-09-11	summer	25.2	0.6	55.4	16.8	0.42	0.32	europe	france	lyon	45.764	4.83566	temperate oceanic
3	740	1998-09-11	summer	26.1	0.61	50.2	16.8	0.46	0.46	europe	france	lyon	45.764	4.83566	temperate oceanic

Figura 9: esempio di dataframe finale - Francia

Tabella 3: parametri selezionati nel dataset

Parametri selezionati dal Dataset estratto dal DB ASHARE.							
building_id	timestamp	season	ta	clo	rh	t_out	climate
Edificio	Data e ora	Stagione	Temperatura interna	Isolamento termico	Umidità relativa	Temperatura esterna	Zona climatica
pmv	pmv_ce	region	country	city	lat	lon	
Pmv (ISO 7730)	Pmv (ASHRAE 55)	Continente	Paese	Città	Latitudine	Longitudine	

Le zone climatiche seguono la classificazione dei climi di Köppen [18]:

Europe map of Köppen climate classification

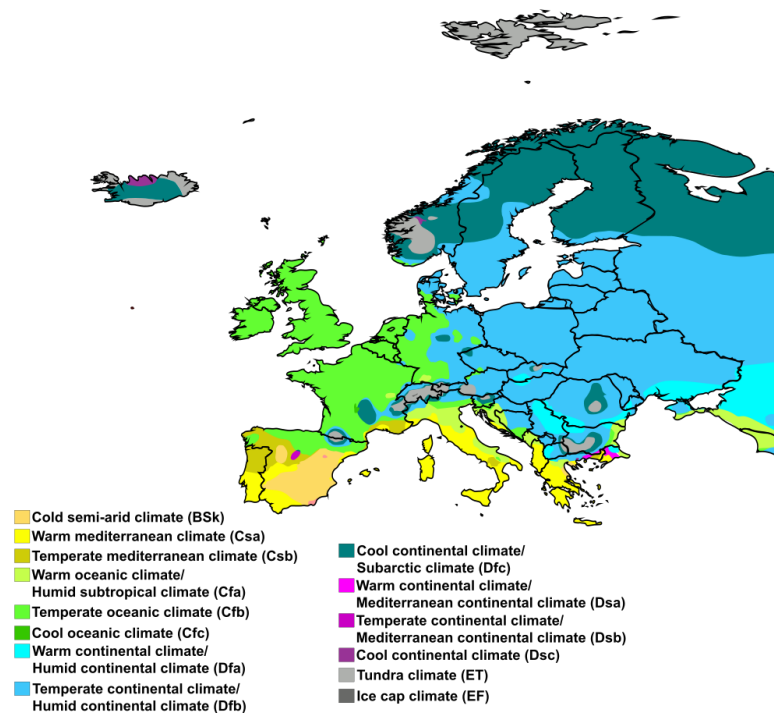


Figura 10: classificazione zone climatiche Köppen

Tra i Paesi Europei presenti nel database sono stati scelti Francia (Figura 11a), Svezia (Figura 11b) e Portogallo (Figura 11c), poiché presentavano un numero consistente di dati.

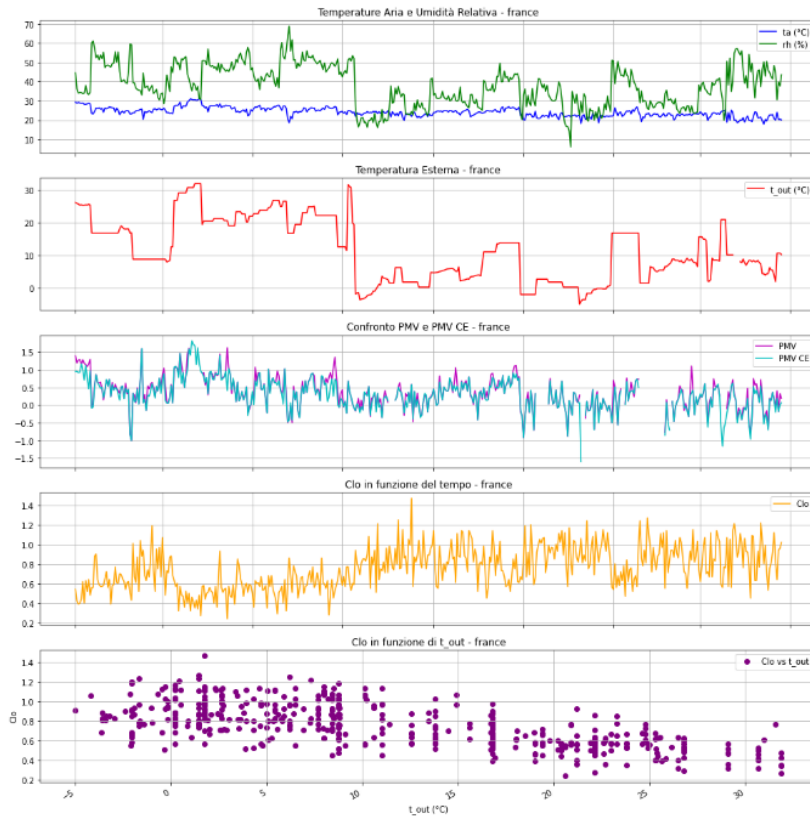


Figura 11a: plot dei parametri - Francia

Svezia

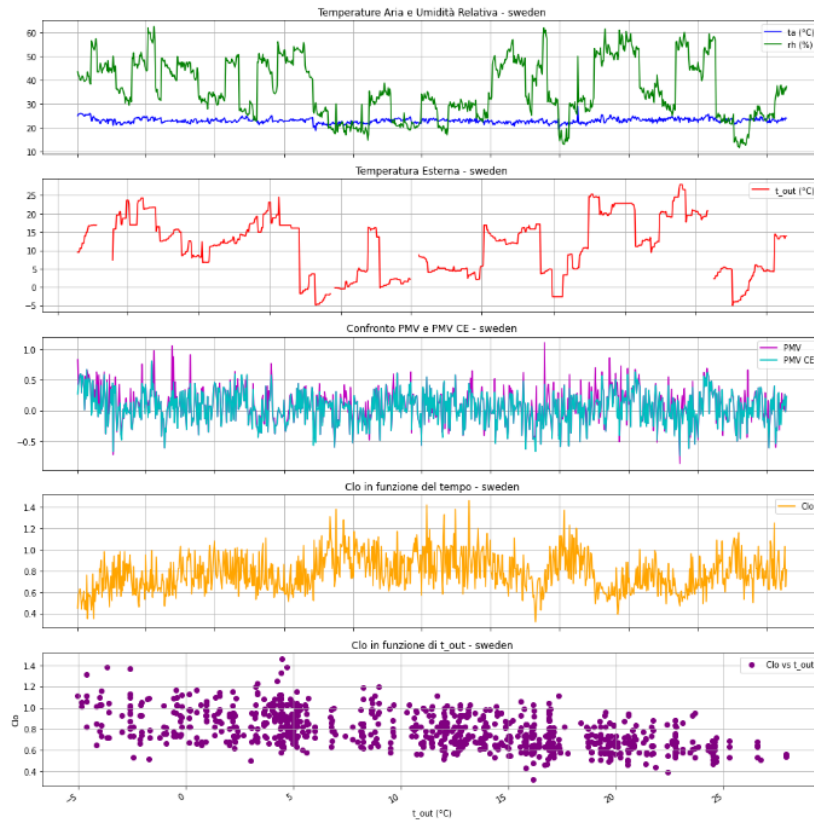


Figura 11b: plot dei parametri - Svezia

Portogallo

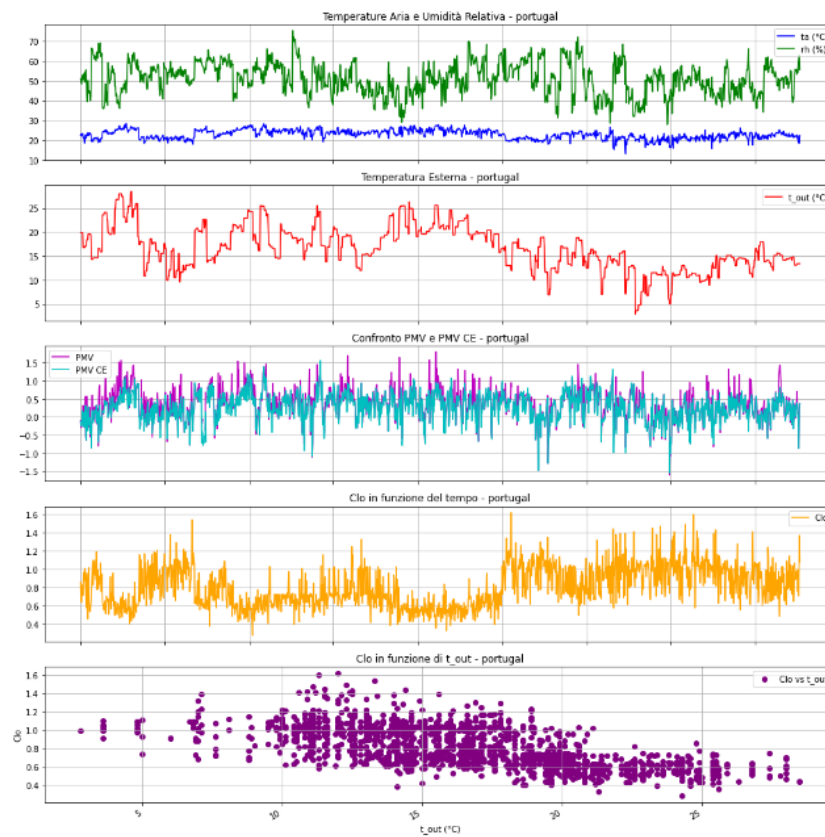


Figura 11c: plot dei parametri - Portogallo

Le zone climatiche di Köppen [18] corrispondenti a tali paesi sono i seguenti: clima oceanico temperato, mediterraneo temperato e clima freddo continentale.

La fase successiva è il processamento dei dati, fondamentale per garantire che l'analisi sia condotta su dati affidabili e consistenti. Questo processo include tecniche di data filtering, come la gestione di valori mancanti (NaN), l'interpolazione dei dati e l'eliminazione di outlier. Queste operazioni, implementate utilizzando il linguaggio di programmazione Python e librerie specifiche come *pandas* e *numpy*, permettono di filtrare il dataset, migliorando l'accuratezza e l'affidabilità dei modelli sviluppati nel contesto di questa tesi.

Durante la raccolta dei dati, possono verificarsi casi in cui alcune osservazioni siano incomplete, generando valori NaN ("Not a Number"). I valori NaN rappresentano celle vuote o corrotte che, se non trattate, possono alterare i risultati delle analisi. Per affrontare questo problema, è stata applicata l'interpolazione lineare. In questo modo, si garantisce che i valori stimati seguano la tendenza naturale dei dati senza introdurre deviazioni significative,

supposizione ragionevole considerando che si tratta di misurazioni in ordine cronologico distanziate da un intervallo di tempo.

L'individuazione degli outlier è stata effettuata utilizzando il metodo dell'Interquartile Range (IQR), uno dei più comuni per la rilevazione di valori anomali. Gli outlier sono valori che deviano significativamente dalla distribuzione complessiva dei dati, che possono essere dovuti a errori di misura o a variazioni estreme nei fenomeni osservati, rischiando che possano distorcere i risultati delle analisi statistiche e predittive.

L'IQR è definito come la differenza tra il terzo quartile (Q3) e il primo quartile (Q1) della distribuzione dei dati. Gli outlier da rimuovere sono i valori che cadono al di fuori del range definito come:

$$\text{IQR Range} = [Q1 - 1.5 \times \text{IQR}, Q3 + 1.5 \times \text{IQR}] \quad (5)$$

Questa metodologia è stata applicata a ciascuna variabile del dataset, identificando e rimuovendo i valori estremi che potevano influire negativamente sulle analisi successive.

Seguono grafici di confronto tra dati originali e dati processati per ogni Paese: nelle Figure 12, 13 e 14 sono riportati in alto i plot di ta (temperatura indoor, in blu) e rh (umidità relativa, in verde), mentre nei grafici inferiori sono visibili i plot di t_{out} (temperatura outdoor, in rosso).

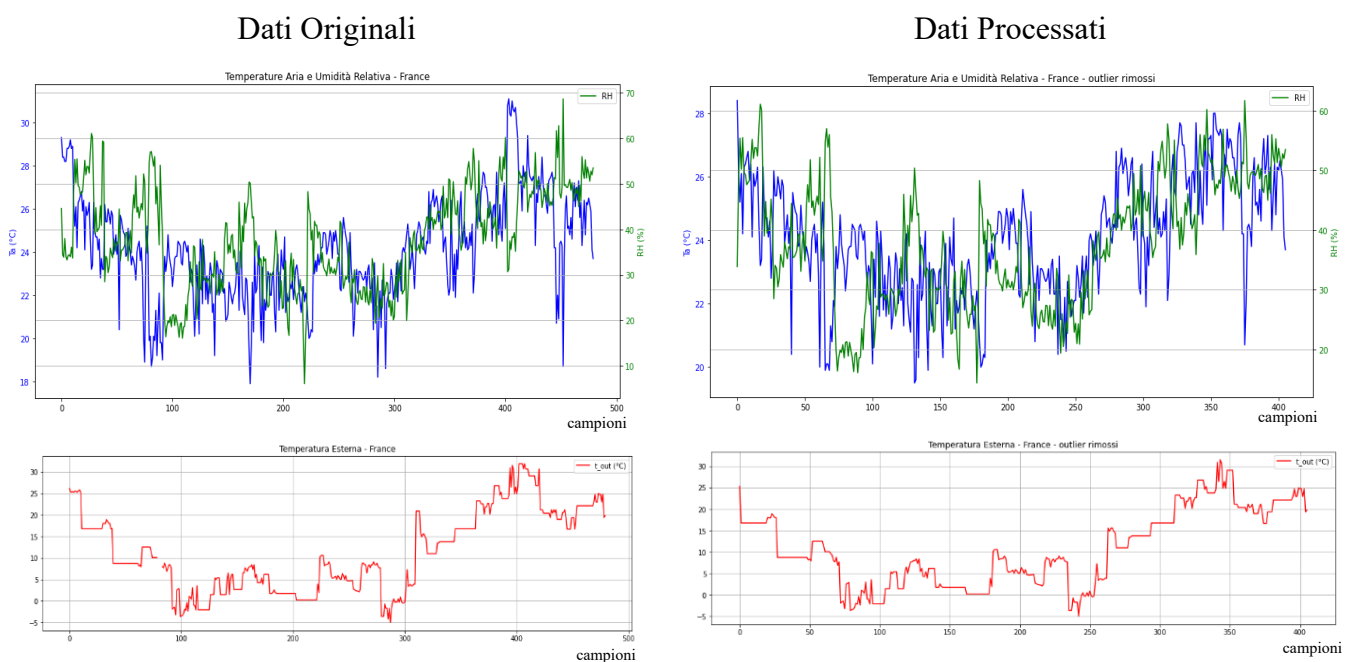
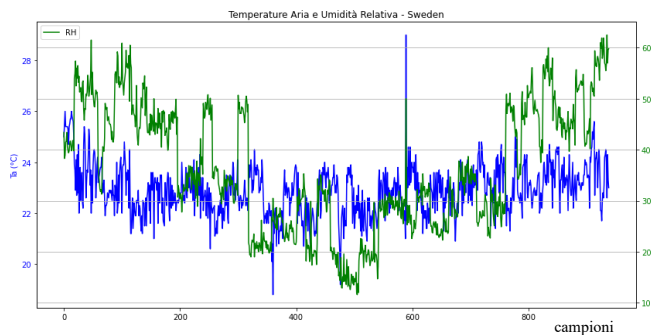
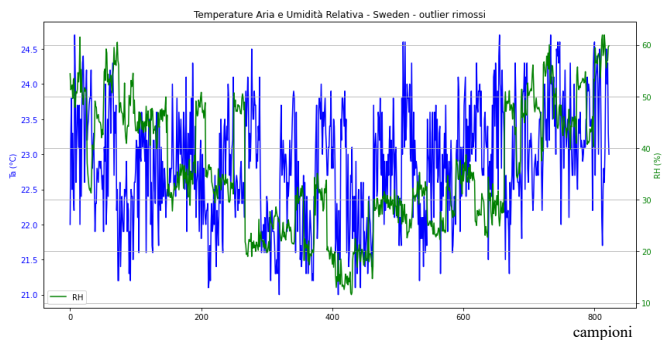


Figura 122: analisi e processamento dati - Francia

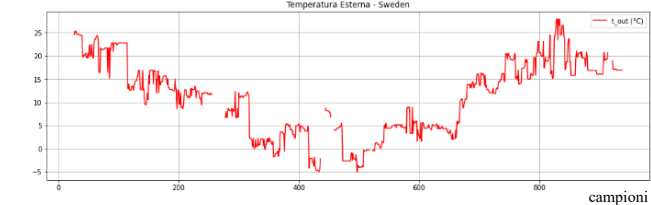
Dati Originali



Dati Processati



Temperatura Esterna - Sweden



Temperatura Esterna - Sweden - outlier rimossi

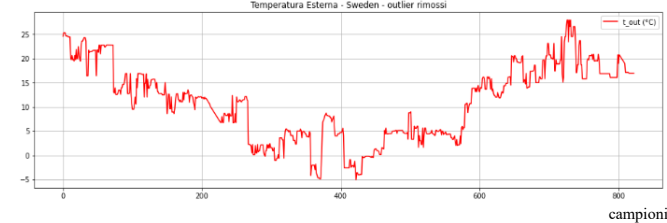
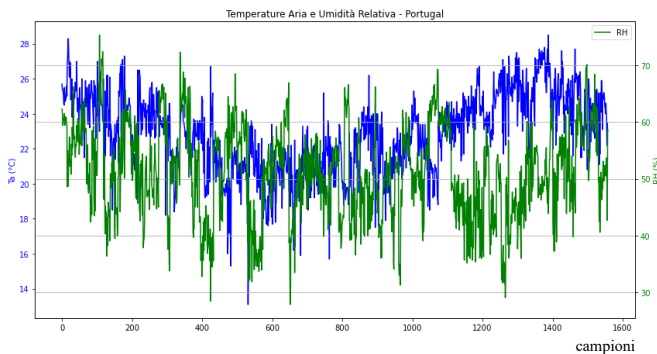


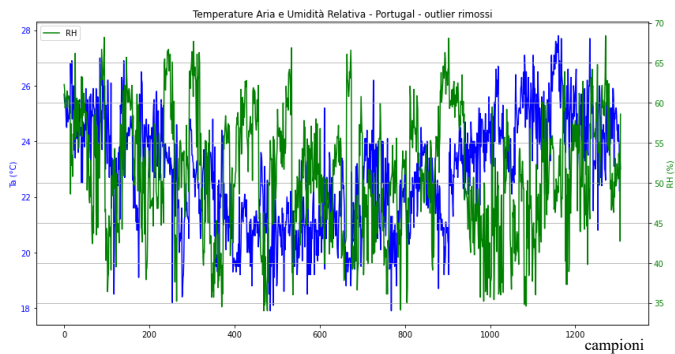
Figura 133: analisi e processamento dati - Svezia

Portogallo:

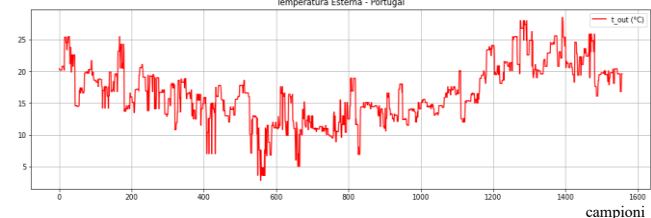
Dati Originali



Dati Processati



Temperatura Esterna - Portugal



Temperatura Esterna - Portugal - outlier rimossi

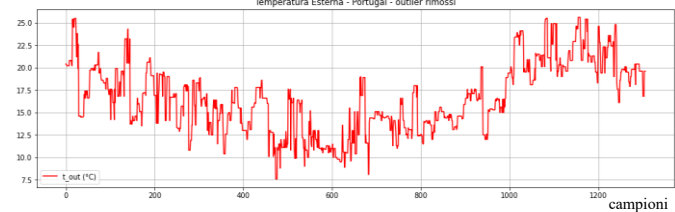


Figura 1414: analisi e processamento dati – Portogallo

3.2.1 K-Means clustering

Ricordando l'obiettivo dello studio, si è cercato un metodo per individuare quattro fasce di Icl corrispondenti ad altrettante fasce di temperatura outdoor. Per fare ciò la prima scelta è stato il clustering con K-Means [19]: una tecnica di machine learning non supervisionato [20], utilizzata per identificare gruppi omogenei all'interno di un dataset. Si basa sulla suddivisione dei dati in insiemi chiamati cluster, che contengono osservazioni simili tra loro in termini di una o più variabili, rilevando pattern nascosti.

In questo studio, il clustering è stato applicato per analizzare la relazione tra la temperatura esterna (t_{out}) e il livello di isolamento termico degli indumenti, espresso in clo , con l'obiettivo di identificare fasce di temperatura che corrispondessero a specifici intervalli di Icl . Tra i vari algoritmi di clustering disponibili, è stato scelto il K-Means, un metodo diffuso e particolarmente efficace per dati che presentano una struttura lineare. Il metodo K-Means suddivide i dati in k gruppi, assegnando ogni punto al cluster il cui centroide (un punto rappresentativo del cluster) è il più vicino, in base alla distanza euclidea. I centroidi vengono ricalcolati ad ogni iterazione, fino a quando non si raggiunge la convergenza, cioè quando i centroidi e le assegnazioni dei punti non cambiano più in modo significativo.

K-Means è stato utilizzato nel nostro studio per identificare gruppi di osservazioni omogenee in relazione a t_{out} e clo . L'obiettivo era verificare se fosse possibile ottenere una suddivisione chiara del livello di isolamento termico in base alla temperatura esterna, utile per identificare fasce di temperatura associate a comportamenti specifici nel vestiario. Una volta identificati i cluster, l'aspettativa era di ottenere intervalli di temperatura ben definiti, che potessero corrispondere a diverse esigenze di clo .

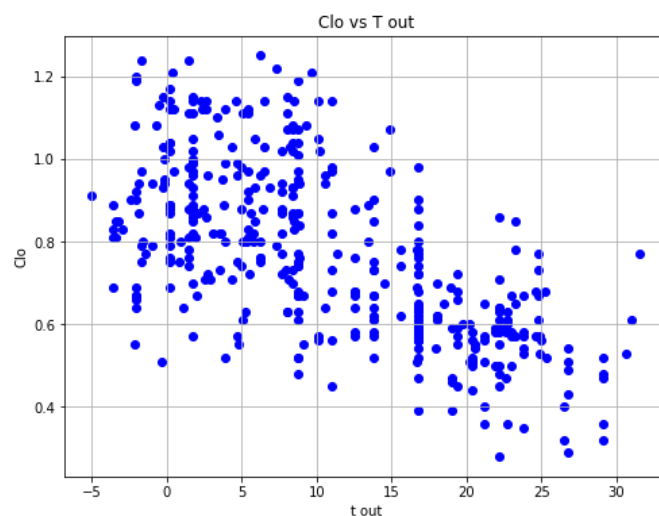


Figura 15: scatter clo vs t_{out} - Francia

Nella Figura 15 è riportata come esempio la distribuzione di *Icl* in funzione della temperatura outdoor per la Francia. Prima di applicare K-Means, è stato necessario procedere alla standardizzazione del dataset, un passaggio fondamentale poiché l'algoritmo si basa su calcoli di distanza. In Python, uno dei metodi più comuni per effettuare la standardizzazione è tramite la classe `StandardScaler` della libreria `scikit-learn`. Questa classe normalizza i dati sottraendo la media e dividendo per la deviazione standard per ogni feature, applicando la seguente formula:

$$z = \frac{x - \mu}{\sigma} \quad (6)$$

Dove:

- x è il valore di un dato osservato,
- μ è la media dei valori della feature,
- σ è la deviazione standard.

Variabili con scale differenti potrebbero avere un peso eccessivo, influenzando il risultato del clustering in modo distorto. Ad esempio, il valore di *t_out* potrebbe variare su un intervallo ampio, mentre i valori di *clo* si trovano su una scala più ristretta, portando l'algoritmo a dare più importanza alla temperatura che all'isolamento termico. La standardizzazione ha trasformato le variabili in una scala comune, con una media pari a zero e una deviazione standard pari a uno, assicurando che tutte le variabili avessero lo stesso peso nel processo di clustering.

Un aspetto cruciale del K-Means è la scelta del numero di cluster k . Dal principio l'obiettivo fissato era di individuare quattro clusters, ma si è comunque fatta un'indagine preliminare per determinare il valore ottimale di k . L'Elbow Method si basa sull'analisi del SSE (Sum of Squared Errors), che misura la distanza tra i punti dati e i rispettivi centroidi all'interno di ciascun cluster. Aumentando il numero di cluster, il SSE diminuisce, ma dopo un certo punto l'incremento di k non porta miglioramenti significativi, e si osserva un "gomito" (elbow) nel grafico del SSE. Questo punto segna il numero ottimale di cluster, oltre il quale i benefici della suddivisione ulteriore diventano trascurabili. grafico del SSE. Questo punto segna il numero ottimale di cluster, oltre il quale i benefici della suddivisione ulteriore diventano trascurabili.

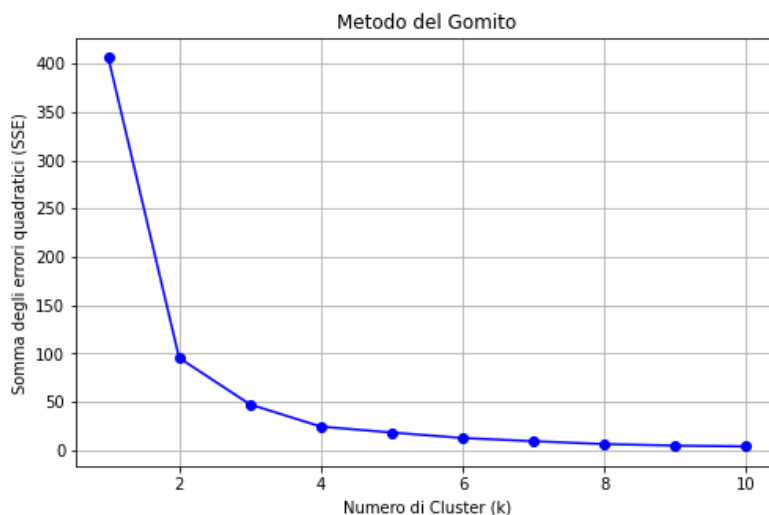


Figura 1615: Elbow Method - Francia

Nel nostro caso, il grafico risultante dall'Elbow Method ha indicato un valore ottimale di k intorno a tre cluster, suggerendo che i dati potevano essere ragionevolmente suddivisi in tre gruppi principali, riflettendo possibili differenze comportamentali in termini di isolamento termico in relazione a diverse fasce di temperatura. In particolare, in Figura 16 si può osservare il plot dell'Elbow Method relativo alla Francia.

In secondo luogo, è stato utilizzato il Silhouette Score per valutare la qualità dei cluster ottenuti. Il Silhouette Score è una metrica che misura quanto ciascun punto, che rappresenta un dato, è vicino agli altri punti del proprio cluster rispetto ai punti degli altri cluster. Il punteggio varia tra -1 e 1, dove un valore vicino a 1 indica che i punti sono ben allineati con il loro cluster, un valore vicino a 0 indica che i punti si trovano ai margini tra due cluster, e un valore negativo indica che i punti sono assegnati al cluster sbagliato.

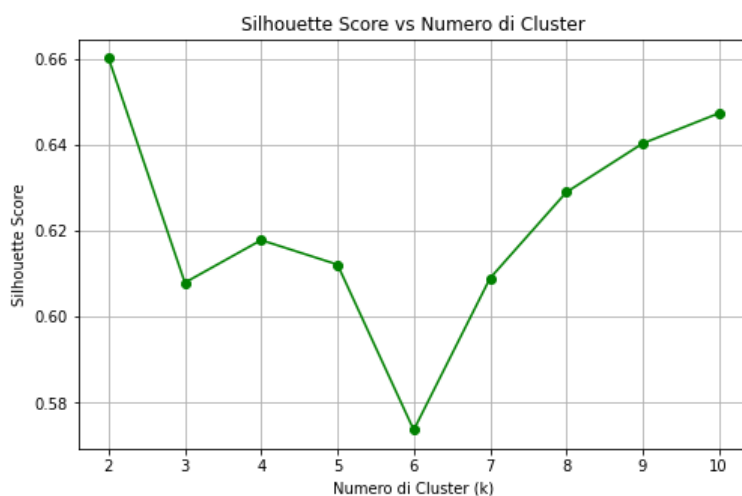


Figura 1716: Silhouette Score - Francia

Nel nostro studio, il Silhouette Score è stato utilizzato per verificare che i cluster ottenuti fossero ben separati tra loro e che le osservazioni fossero correttamente raggruppate. Il punteggio risultante mostrato nella Figura 17 evidenzia che la suddivisione in quattro cluster per la Francia si potrebbe considerare di buona qualità, indicando una chiara separazione tra gruppi di dati. Nonostante le aspettative iniziali, il clustering ha rivelato un risultato inaspettato. L'idea di base era quella di identificare fasce ben definite di I_{cl} in relazione a intervalli di temperatura esterna (t_{out}). Tuttavia, i risultati del clustering hanno mostrato che i valori di clo si sovrapponevano tra i diversi cluster di temperatura.

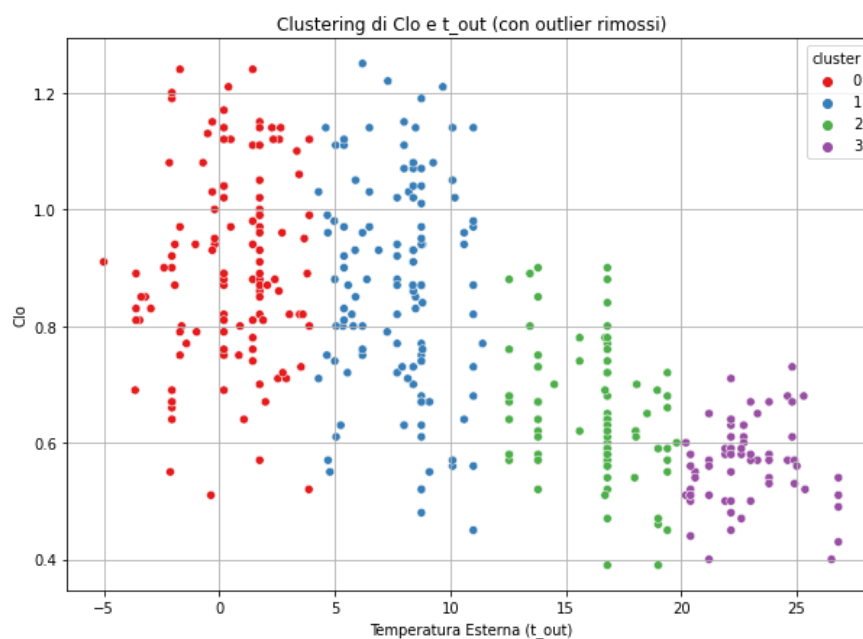


Figura 1817: scattering del clustering su t_{out} - Francia

Ciò significa che non è emersa una chiara correlazione tra l'aumento della temperatura e una riduzione proporzionale dell'isolamento termico. In altre parole, i dati indicavano che il valore di I_{cl} non segue una progressione lineare o facilmente segmentabile sulla base di t_{out} , suggerendo che altre variabili o fattori personali potrebbero influenzare maggiormente la scelta dell'abbigliamento. Data la sovrapposizione osservata con K-Means, è stato tentato un approccio di clustering combinato [20], che ha coinvolto l'utilizzo di variabili aggiuntive, come la temperatura indoor o l'umidità relativa, per arricchire l'analisi.

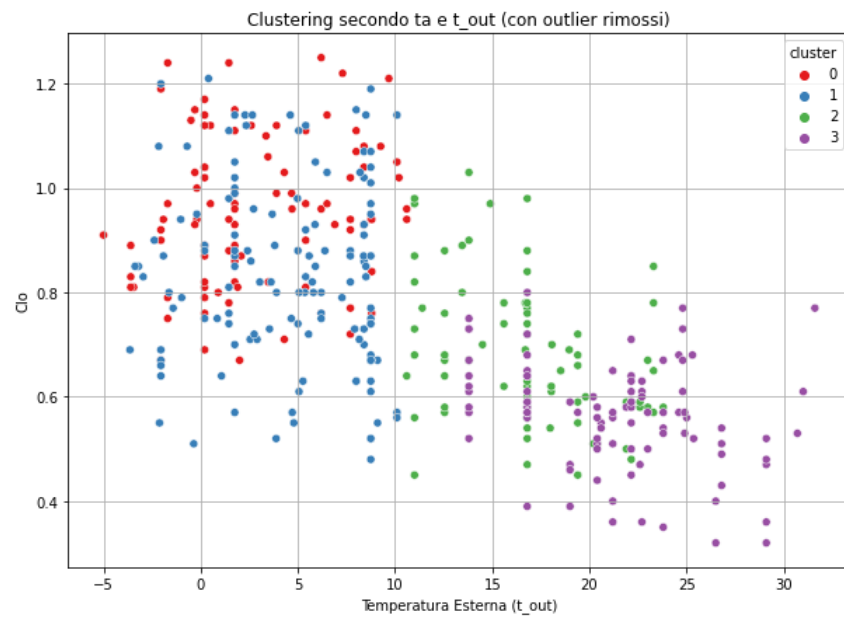


Figura 1918: scattering clustering combinato - Francia

Tuttavia, anche includendo queste variabili, non è stato possibile identificare fasce ben definite di *clo* basate su *t_out*, al contrario si osserva solo un abbassamento notevole del Silhouette Score.

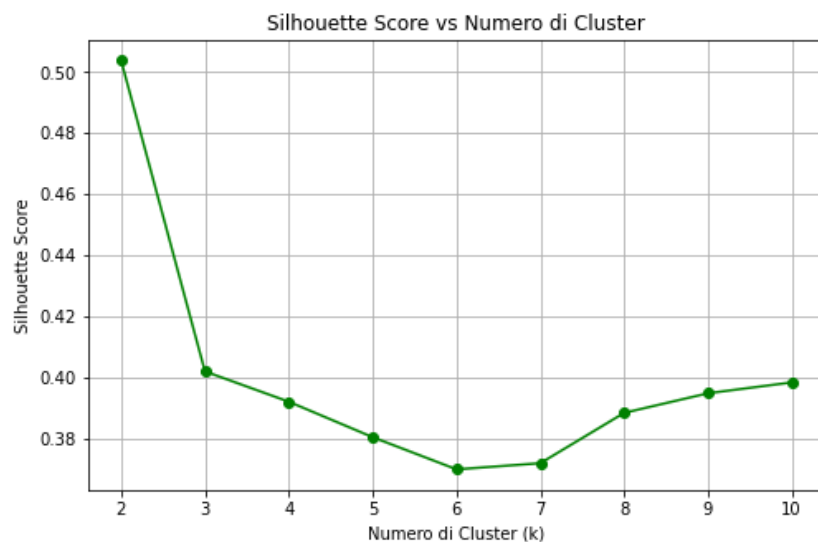


Figura 2019: Silhouette Score del clustering combinato - Francia

In un ulteriore tentativo, è stato sperimentato il clustering gerarchico [21], un metodo che differisce dal K-Means in quanto non richiede di specificare il numero di cluster in anticipo. Il clustering gerarchico costruisce una dendrogramma, visibile in Figura 21, che rappresenta le relazioni tra le osservazioni, permettendo una fusione graduale dei cluster fino a formare

un'unica struttura gerarchica. Questo approccio offre una visione più flessibile dei dati, consentendo di esplorare diversi livelli di raggruppamento.

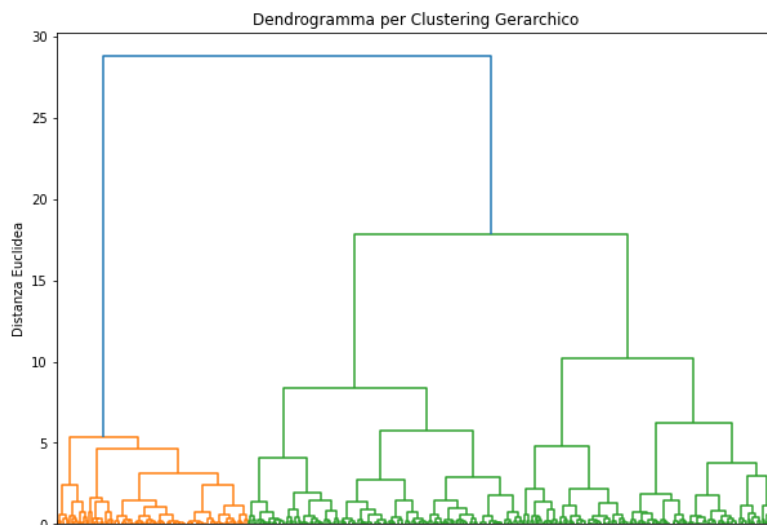


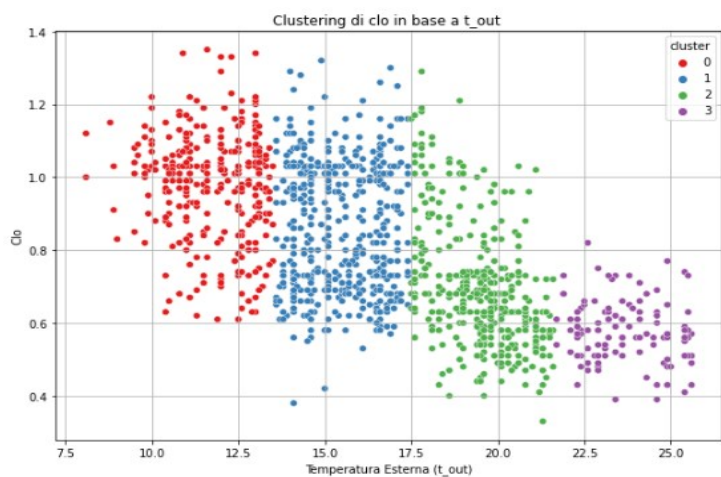
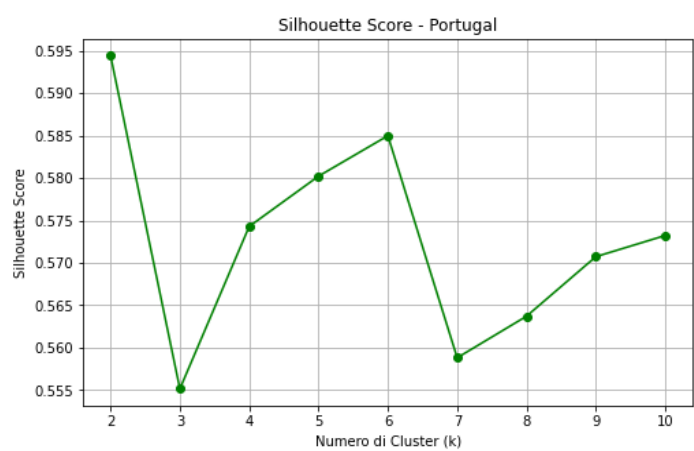
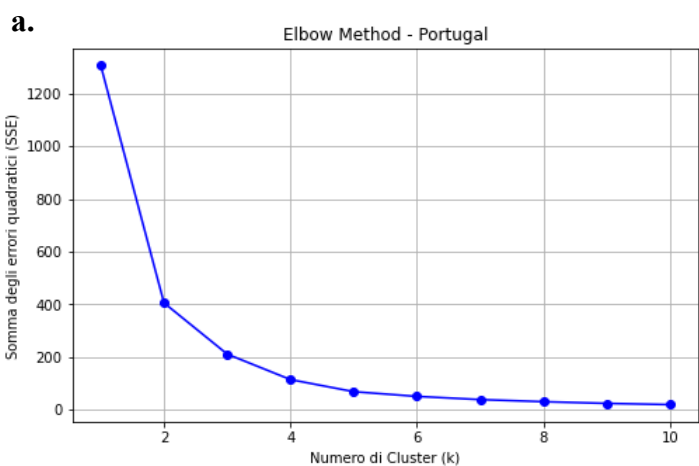
Figura 2120: dendrogramma clustering gerarchico - Francia

Tuttavia, anche il clustering gerarchico non ha prodotto le fasce di *clo* attese. I risultati hanno continuato a mostrare una forte sovrapposizione tra i cluster, suggerendo che la relazione tra *Icl* e *t_out* non sia abbastanza semplice da essere descritta con i modelli di clustering utilizzati.

Il clustering applicato nel presente studio ha permesso di esplorare la relazione tra la temperatura esterna e il livello di isolamento termico indossato, ma non ha prodotto i risultati attesi. L'analisi ha rivelato che non esiste una correlazione così diretta tra *t_out* e *clo*, e che i valori di isolamento termico si sovrappongono tra i diversi intervalli di temperatura, suggerendo che la scelta dell'abbigliamento sia influenzata da una combinazione maggiore di parametri ambientali, oltre che dalle preferenze personali.

I grafici in Figura 22 mostrano i risultati del clustering su *t_out* anche per la Svezia e il Portogallo.

Portogallo



Svezia

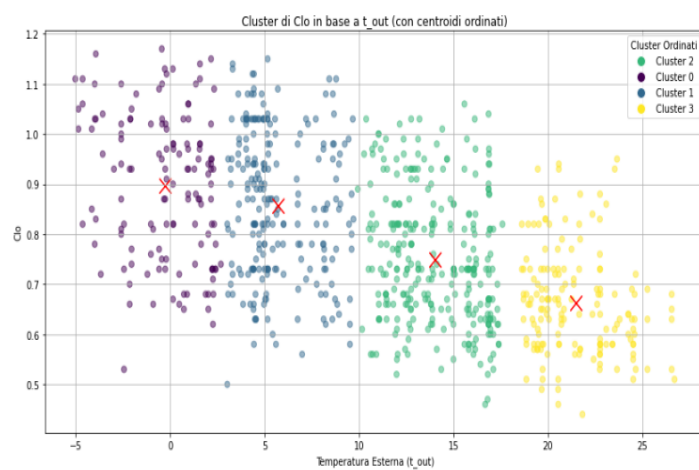
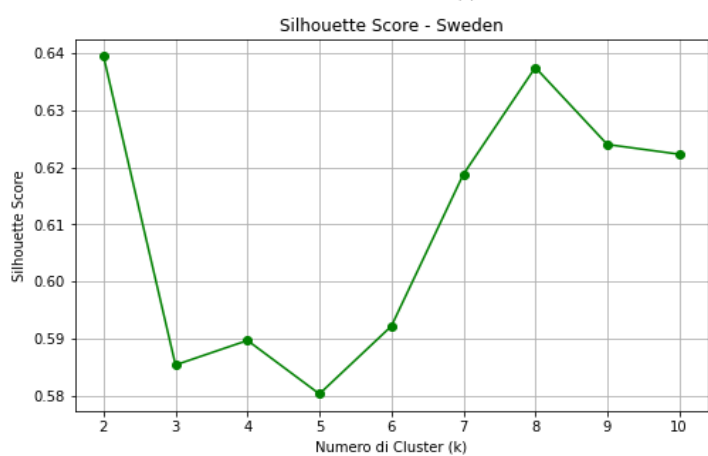
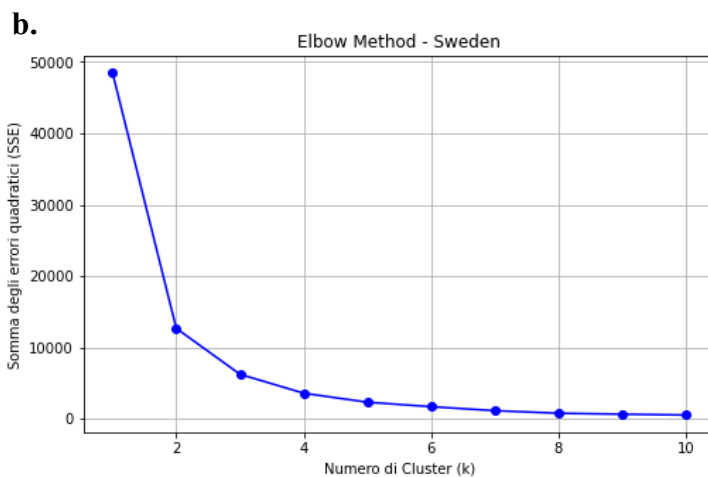


Figura 212: fasi di clustering per Portogallo e Svezia

3.2.2 Generalized Additive Model (GAM)

Alla luce di quanto emerge dal processo di clustering, lo studio si è orientato verso un approccio più complesso e generale, intuendo che *Icl* sia, in realtà, influenzato da un maggior numero di fattori. Per esplorare questa prospettiva, è stato impiegato un algoritmo supervisionato: il Generalized Additive Model [22] (GAM), modello statistico che consente relazioni non lineari tra le variabili indipendenti e la variabile dipendente. Mentre nei modelli lineari classici si assume che ogni variabile indipendente influenzi la variabile dipendente in modo lineare, il GAM permette di modellare tale relazione con maggiore flessibilità, utilizzando funzioni non lineari, chiamate *smooth functions*, per ciascuna variabile predittiva. Dunque, è stato scelto perché la relazione tra le variabili ambientali, come la temperatura outdoor (*t_out*), la temperatura dell'aria (*ta*), l'umidità relativa (*rh*), e il livello di isolamento termico (*clo*) non è necessariamente lineare. [23] Questo modello permette di esplorare e quantificare meglio la complessità e la non linearità di queste relazioni. Inoltre, riesce a quantificare l'importanza relativa di ciascun predittore (*t_out*, *ta*, *rh*) e di capire come queste variabili influenzano *Icl* in maniera separata, offrendo così una visione chiara della dinamica del fenomeno. Il GAM si basa su un concetto chiave: le variabili indipendenti influenzano la variabile dipendente attraverso funzioni additive che possono essere non lineari.

La formula generale di un GAM è la seguente:

$$g(y) = \beta_0 + f_1(x_1) + f_2(x_2) + \dots + f_k(x_k) + \epsilon \quad (7)$$

Dove:

- $g(y)$ è la funzione di link, che collega la variabile dipendente y (in questo caso, il *clo*) alla somma dei termini che rappresentano gli effetti delle variabili indipendenti. In pratica, la funzione di link trasforma il risultato del modello in modo che possa essere adattato al tipo di distribuzione dei dati;
- β_0 è l'intercetta;
- $f_1(x_1)$, $f_2(x_2)$, ... sono le funzioni non lineari (*smooth functions*) che descrivono come ciascuna variabile indipendente influisce sulla variabile dipendente;
- ϵ è l'errore residuo.

Nel nostro caso, la variabile dipendente è *Icl*, mentre le variabili indipendenti sono *t_out*, *ta* e *rh*. Il GAM, grazie alle smooth functions, permette di modellare la dipendenza del clo da ciascuna di queste variabili in modo molto più flessibile rispetto a un modello lineare, rilevando andamenti non lineari che potrebbero sfuggire a un'analisi meno sofisticata. La funzione *summary* fornisce un resoconto di tutte le informazioni estrapolate dal modello riguardo il set di dati:

```

LinearGAM
-----
Distribution:          NormalDist Effective DoF:          61.7124
Link Function:        IdentityLink Log Likelihood:       -7511.1045
Number of Samples:    406 AIC:                          15147.6339
                                     AICc:                       15170.9801
                                     GCV:                          0.027
                                     Scale:                         0.0197
                                     Pseudo R-Squared:             0.6141
-----
Feature Function      Lambda          Rank      EDoF      P > x      Sig. Code
-----
s(0) T_OUT           [0.6]          40        23.9      3.44e-03    **
s(1) TA              [0.6]          40        20.1      5.93e-04    ***
s(2) RH              [0.6]          40        17.7      6.33e-02    .
intercept            1              1         0.0       1.11e-16    ***
-----
Significance codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Figura 223: summary del GAM - Francia

I parametri riportati in Figura 23 sono spiegati di seguito:

- **GCV**: Generalized Cross-Validation, per valutare la qualità del modello;
- **Scale**: stima della varianza;
- **Pseudo R-Squared**: indicazione della percentuale di varianza spiegata dal modello: superiore a 0.5 può considerarsi un buon modello;
- **Lambda**: controlla la regolarizzazione della funzione;
- **Rank**: numero di parametri nella funzione;
- **EDoF**: gradi di libertà effettivi associati a ciascuna funzione;
- **P > x**: significatività statistica di ciascun predittore.

Ogni parametro associato a una variabile ha un p-value, che misura la probabilità che l'effetto osservato sia dovuto al caso. I Significance Codes sono una codifica sintetica di questi p-value, secondo la seguente convenzione standard (in R e in molti altri software statistici):

- *******: p-value < 0.001 (molto significativo)
- ******: p-value < 0.01 (altamente significativo)
- *****: p-value < 0.05 (significativo)
- **.**: p-value < 0.1 (tendenzialmente significativo)
- Nessun simbolo: p-value >= 0.1 (non significativo)

In questo caso, emerge una forte dipendenza anche dalla temperatura indoor, che prima non era stata presa in esame. Con i grafici di dipendenza parziale (Partial Dependence Plots, PDP) si può visualizzare la relazione tra ciascuna variabile indipendente e il clo , mantenendo costanti tutte le altre variabili. Questi grafici, mostrati in Figura 24, permettono di capire visivamente come la variabile predittiva influenzi il valore atteso di Icl , rappresentando le funzioni non lineari che il GAM ha stimato.

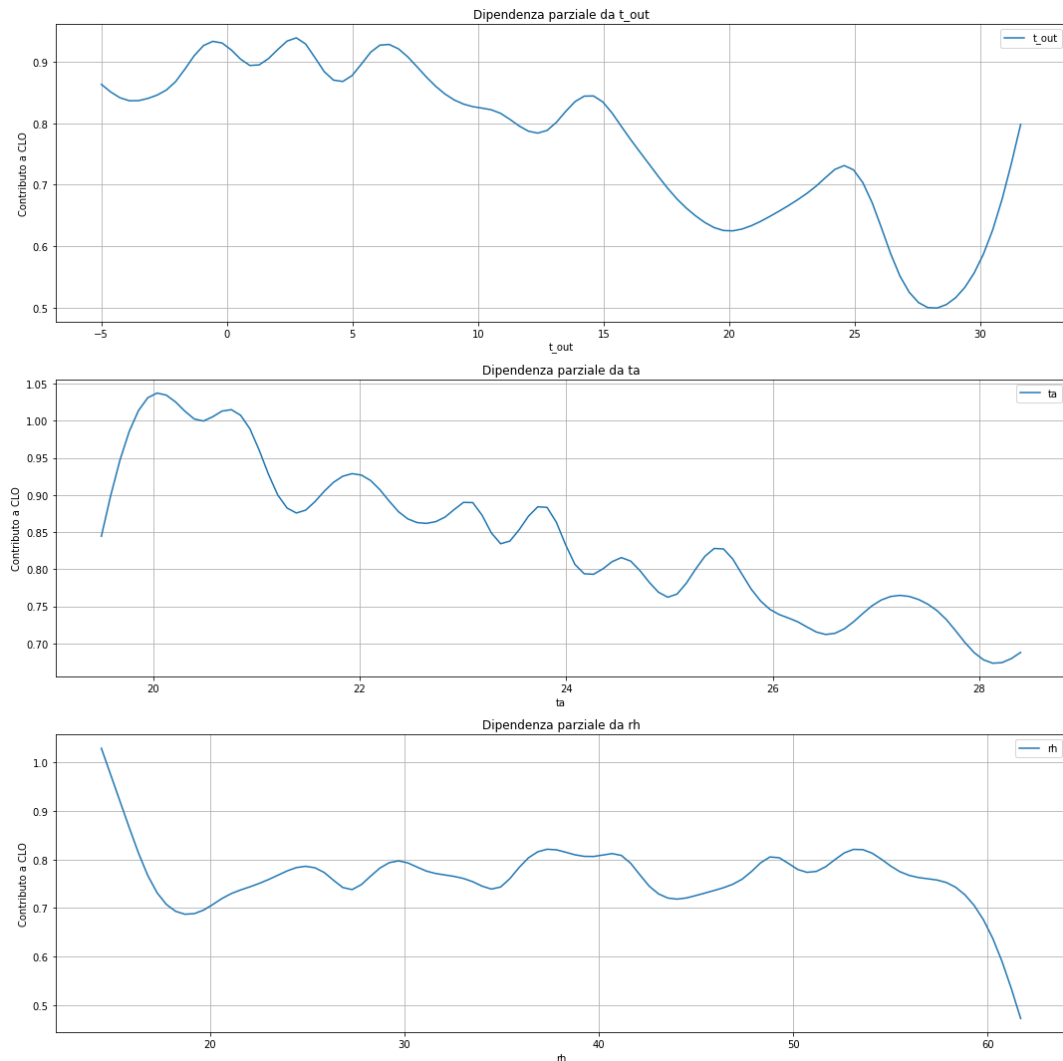


Figura 234: grafici di dipendenza parziale – Francia

La funzione da cui sono stati estratti questi grafici ha la seguente forma:

$$g(clo) = f_1(t_{out}) + f_2(t_a) + f_3(rh) + \beta_0 + \epsilon \quad (8)$$

Il numero di splines imposte a priori, ovvero il numero di funzioni di smoothing impiegate per esprimere il contributo di ciascun parametro, pari a 40. Meno funzioni vengono imposte, più il modello si avvicina alla linearità; aumentando il numero, il modello riesce a cogliere relazioni più complesse, con un conseguente incremento del parametro Pseudo R-Squared. È fondamentale imporre il giusto numero di splines: deve permettere di avere un modello non troppo rigido, ma allo stesso tempo senza rischiare overfitting, ovvero l'eccessivo adattamento ai dati di addestramento, penalizzando la performance su nuovi dati.

L'accuratezza di approssimazione dell'isolamento termico è stata misurata mediante l'errore assoluto medio (MAE) = 0.104 e l'errore quadratico medio (MSE) = 0.017.

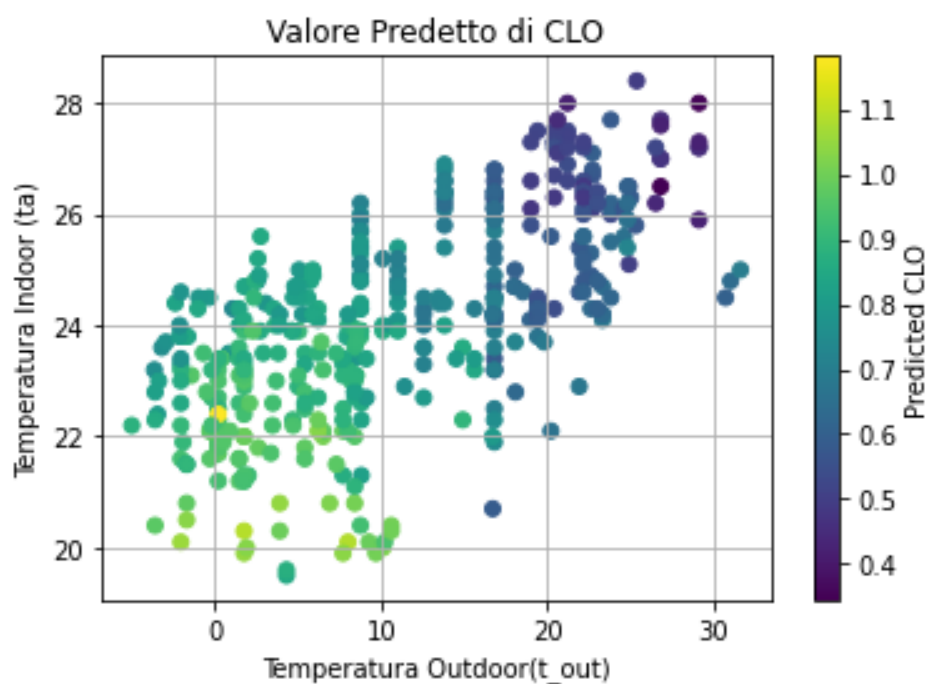


Figura 245: scatter del clo approssimato dal GAM – Francia

La Figura 25 mostra la distribuzione dei dati in funzione di t_{out} e t_a , colorato secondo i valori di clo approssimati dal modello GAM. L'implementazione è avanzata andando a semplificare ulteriormente le variabili (splines = 10).

Le funzioni di dipendenza parziale assumono la forma visibile in Figura 26: si può intuire che queste curve siano rappresentabili da funzioni polinomiali.

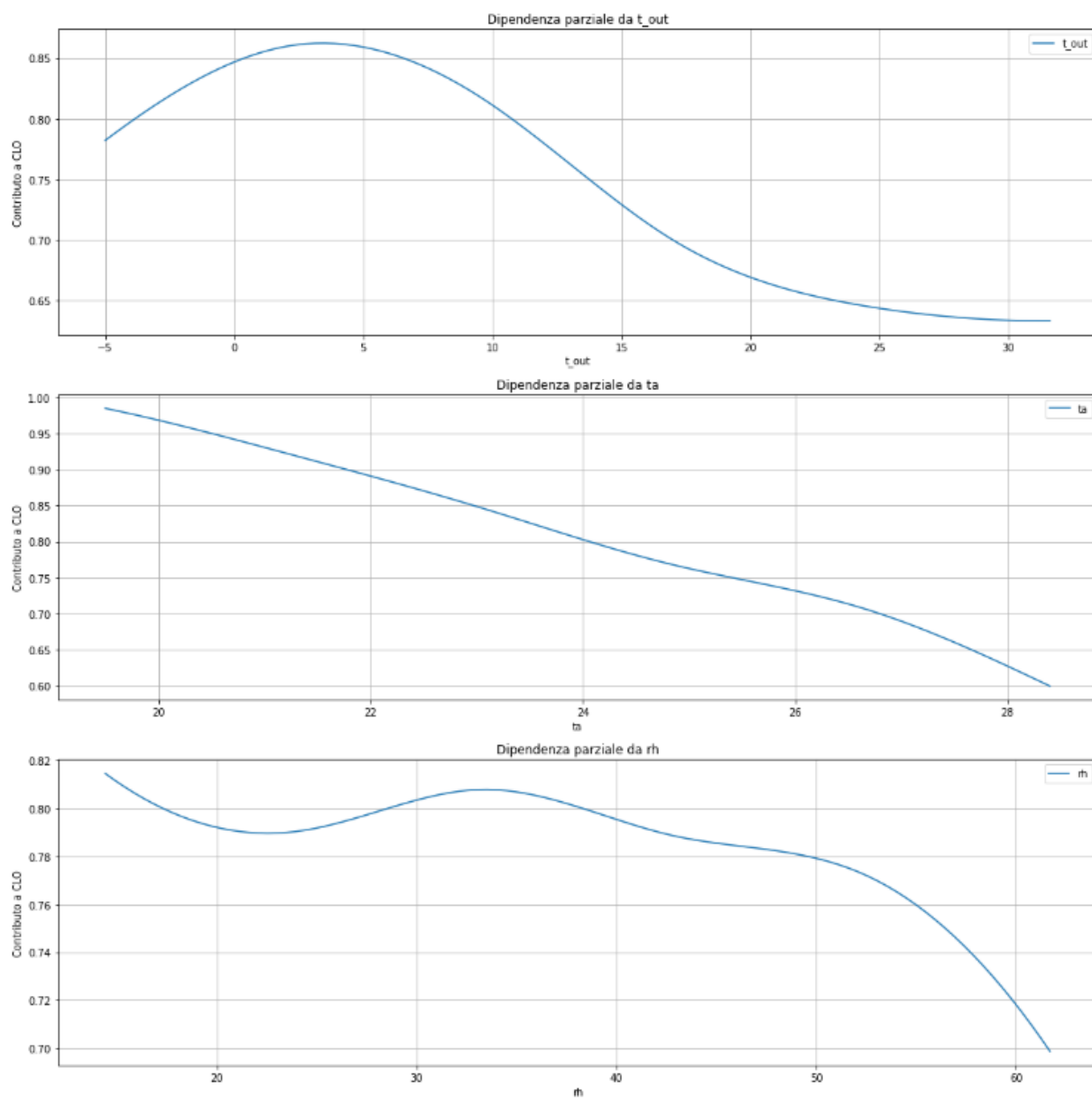
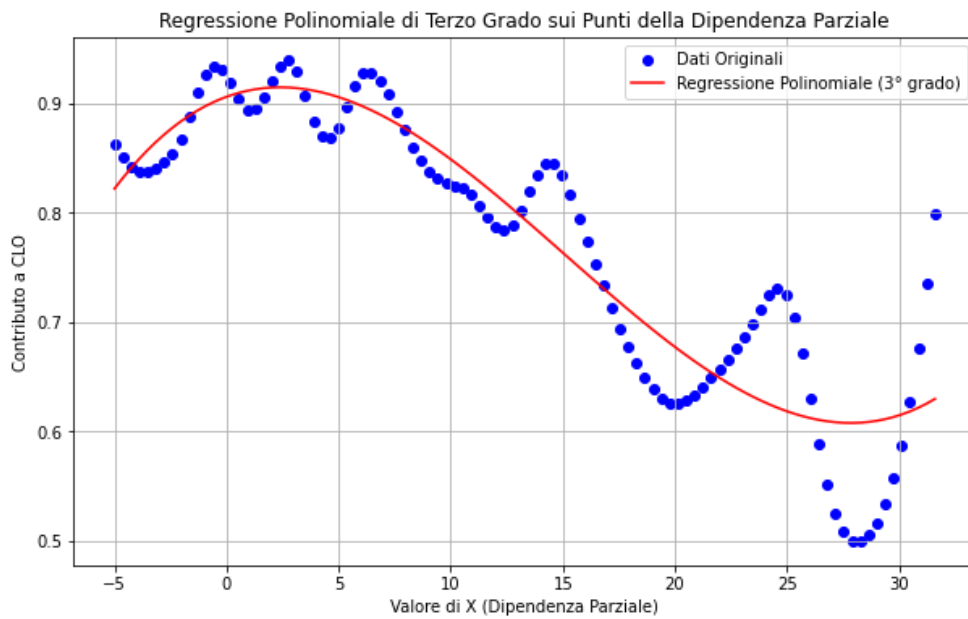


Figura 256: PDP, splines =10, Pseudo R-Squared = 0.55 - Francia

Effettuando una regressione con termini fino al terzo grado su ogni PDP¹, si verifica che delle curve di terzo grado si adattano bene ad approssimare la dipendenza tra i parametri. Queste sono mostrate nelle Figure 27, 28 e 29.

¹ Partial Dependence Plot

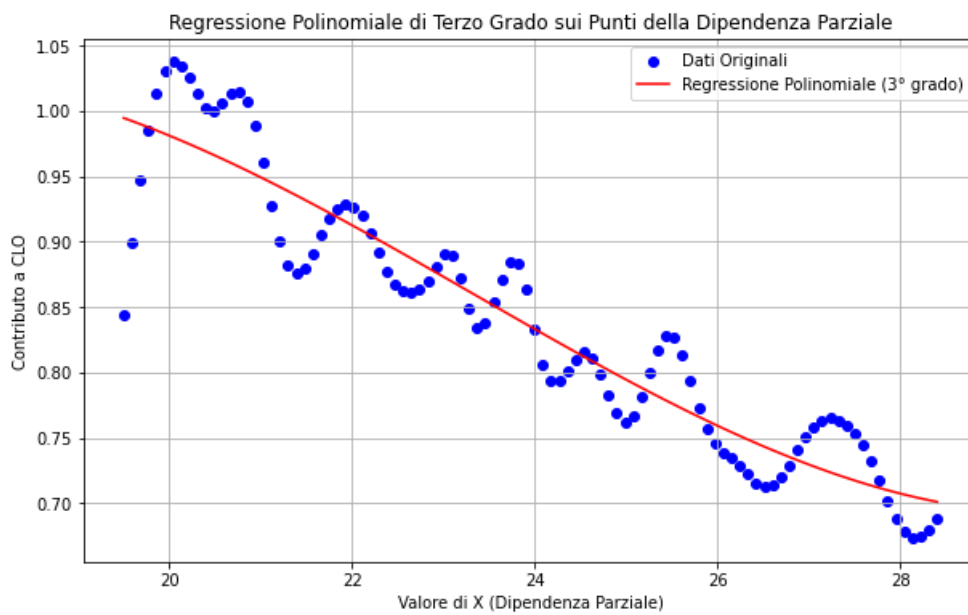
T_OUT



MSE = 0.0025

Figura 27: plot regressione polinomiale sul PDP di t_{out} - Francia

TA



MSE = 0.0012

Figura 28: plot regressione polinomiale sul PDP di ta - Francia

RH

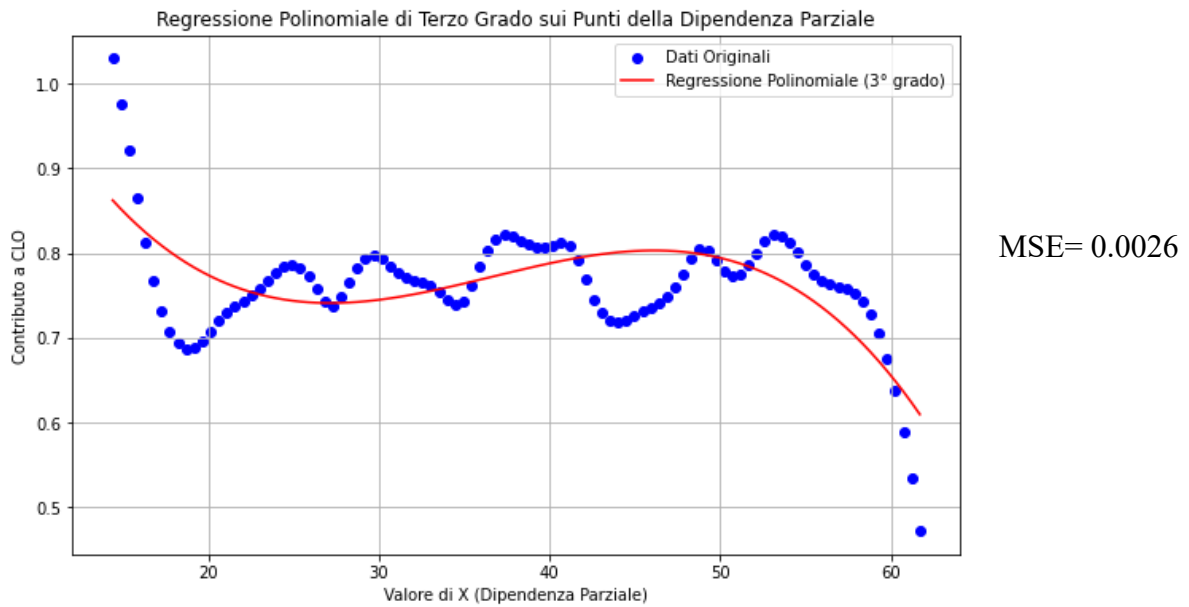


Figura 29: plot regressione polinomiale sul PDP di rh - Francia

Per ogni approssimazione polinomiale è stato calcolato il Mean Squared Error (MSE), ovvero la media degli errori quadrati tra i valori predetti e quelli osservati. Si calcola come:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (9)$$

n è il numero di osservazioni, e $(y_i - \hat{y}_i)^2$ è il quadrato della differenza tra i valori osservati e predetti. Essendo il MSE sufficientemente basso, si conclude che una regressione polinomiale con termini di terzo grado potrebbe approssimare in modo accettabile la complessa dipendenza di Icl dai parametri ambientali analizzati. La regressione lineare è un metodo statistico utilizzato per modellare la relazione tra una variabile dipendente (target) e una o più variabili indipendenti (predittori) attraverso una linea retta. Essa assume che ci sia una relazione lineare tra le variabili. Nel nostro caso, la dipendenza è modellata con una curva di terzo grado.

Seguendo questo ragionamento, sono stati ricavati otto coefficienti per calcolare un clo approssimato con la seguente equazione:

$$\begin{aligned} clo_{app} = & 0.0109 * T_{out} - 0.0019 * T_{out}^2 + 0.00004 * T_{out}^3 - 0.241 * indoorT \\ & + 0.0078 * indoorT^2 - 0.0001 * indoorT^3 - 0.0011 * indoorRH + 3.5530 \end{aligned} \quad (10)$$

Per indicare l'errore tra il clo standard e clo_{app} si usa il Mean Absolute Error (MAE), ovvero la media degli errori assoluti tra i valori predetti (\hat{y}_i) e i valori osservati (y_i). La formula è la seguente:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (11)$$

Dove:

- n è il numero di osservazioni;
- y_i è il valore osservato;
- \hat{y}_i è il valore predetto.

Il MAE misura la grandezza media degli errori nelle previsioni, senza considerare la direzione (non distingue tra valori positivi o negativi). È interpretabile facilmente poiché esprime la deviazione media in unità della variabile osservata. In questo caso si mantiene intorno a 0.1 per tutte le zone climatiche, dimostrando una buona affidabilità a fronte di una distribuzione reale molto complessa, che non permetterebbe di classificare i dati con sufficiente sicurezza in alcun modo. Infatti, si può confrontare lo scatter di questa approssimazione con quello dei dati originali:

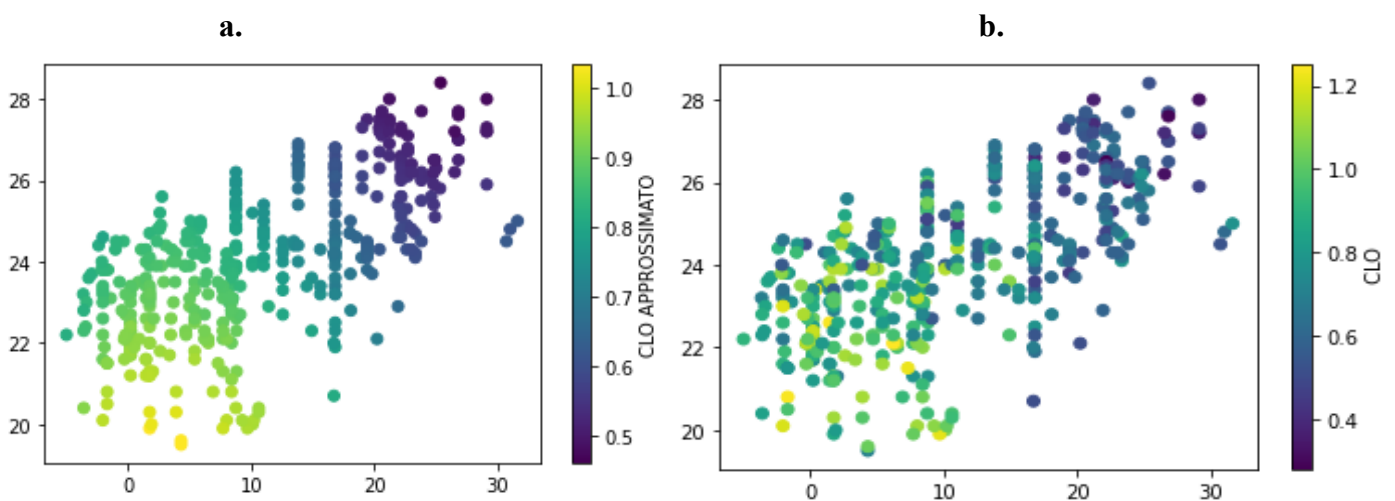


Figura 29: scatter clo approssimato dal GAM vs reale - Francia

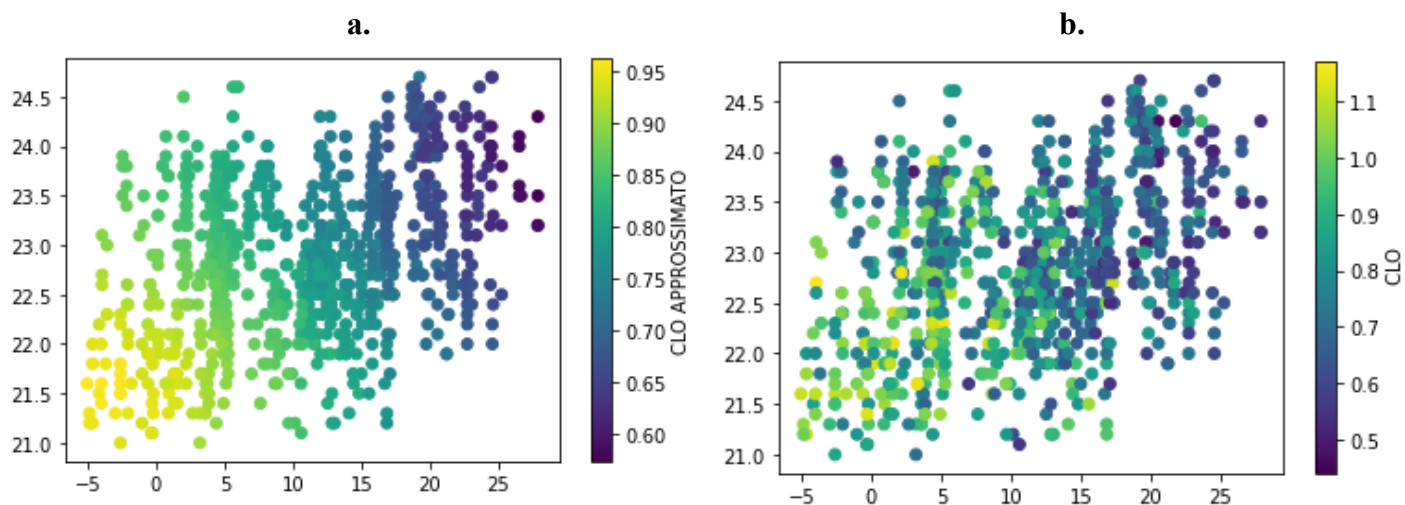


Figura 301: scatter clo approssimato dal GAM vs reale - Svezia

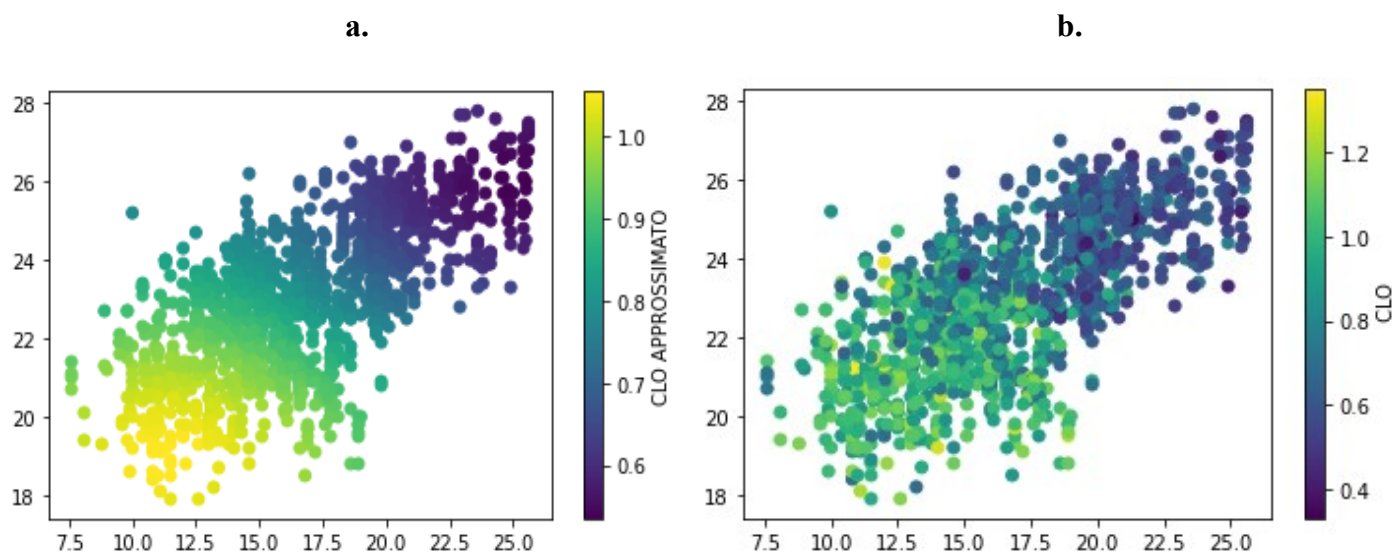


Figura 312: scatter clo approssimato dal GAM vs reale - Portogallo

Considerando le mappature del *clo* approssimato mostrate nelle Figure 30, 31 e 32, si può dividere l'isolamento termico in quattro nuove fasce come riportato nella tabella 4.

Tabella 4: nuove fasce di Icl

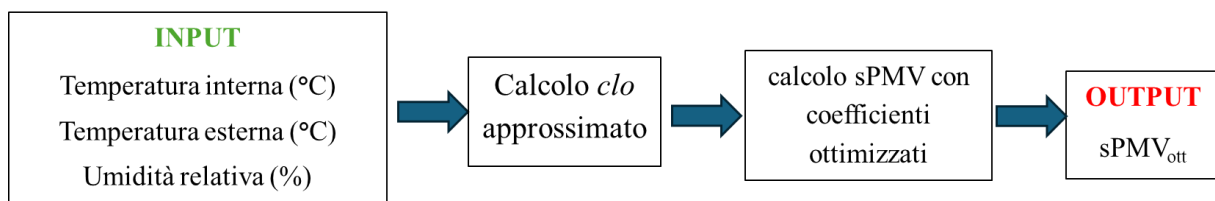
Fasce	In letteratura	Nuovo modello
1	0.25 – 0.50	0.25 – 0.50
2	0.51–1.00	0.51 – 0.8
3	1.01–1.65	0.81 – 0.99
4	/	1 – 1.65

3.2.3 Definizione del nuovo modello

A partire da queste quattro nuove fasce, il passaggio successivo è impostare la funzione di regressione lineare con la quale sono stati ricavati i coefficienti a , b , c per ogni intervallo, da impiegare nella funzione di calcolo del PMV semplificato (Equazione 3). Il metodo identifica una retta che meglio descriva come la variabile dipendente cambia in funzione delle variabili indipendenti, stimando i coefficienti in modo che la somma dei quadrati degli errori (differenza tra i valori osservati e quelli predetti) sia minimizzata.

Dunque, per ogni zona climatica si è ricavato il set di coefficienti che permette di ottenere il valore del PMV semplificato avendo come input i tre parametri ambientali t_a , t_{out} e rh .

La funzione ottenuta, con la quale calcolare il PMV semplificato, segue il seguente schema:



Il codice Python è analogo a quello di partenza basato sul modello Buratti, ma in questo caso non discrimina tramite il mese relativo alla misurazione, bensì calcolando il clo approssimato con l'Equazione 10 e associandolo alla fascia corrispondente. Nel capitolo seguente, verranno mostrati i risultati tramite grafici di confronto, seguiti dalle metriche di valutazione già citate.

3.3 Valutazione della performance

L'affidabilità del modello è valutata secondo le metriche di MAE e MSE per ogni fascia di *clo*, calcolati tra il sPMV ottimizzato e quello standard PMV già presente nel DB ASHRAE. Il MAE e il MSE sono due metriche comunemente utilizzate per valutare le prestazioni dei modelli di previsione o regressione. Misurano quanto i valori predetti differiscono da quelli reali, aiutando a capire la precisione del modello. Il MAE si calcola con l'Equazione 11, il MSE con l'Equazione 9.

Il MSE penalizza maggiormente gli errori grandi rispetto al MAE perché eleva al quadrato le differenze, aumentando l'influenza degli errori più grandi: dunque, aiuta a comprendere se siano presenti errori estremi e quanto questi impattino sui risultati. Questa proprietà rende il MSE particolarmente utile quando si vuole dare maggiore importanza a previsioni molto imprecise. Inoltre, ha una base teorica solida, essendo derivato dalla funzione di perdita quadratica, spesso utilizzata nelle tecniche di regressione lineare e modelli di machine learning. Non si può prescindere affermare quali valori di queste metriche siano accettabili; ovviamente si tiene conto del fatto che più ci si avvicina allo zero, più il modello risulta accurato. A livello applicativo, per questo studio sono state considerati accettabili valori uguali o minori del MAE e del MSE calcolati con i coefficienti a, b, c di Buratti, già presenti in letteratura.

Per evidenziare la differenza e il miglioramento dei modelli a confronto, è stato anche calcolato la differenza percentuale [24] tra le metriche relative al modello di Buratti e quelle appena calcolate. La formula impiegata è la seguente:

$$\text{Differenza percentuale} = \frac{\text{Errore del vecchio modello} - \text{Errore del nuovo modello}}{\text{Errore del vecchio modello}} \times 100 \quad (12)$$

Se il valore è positivo, significa che il nuovo modello è migliorato rispetto al precedente.

4. RISULTATI

Come anticipato nel capitolo precedente, verranno riportate le metriche di validazione e saranno mostrati i grafici relativi ai valori di PMV, confrontando tre set di dati:

- PMV standard, fornito nel dataframe e assunto come riferimento, visualizzato in blu;
- sPMV di Buratti, punto di partenza dello studio, visualizzato in verde;
- sPMV ottimizzato, calcolato con i nuovi coefficienti, visualizzato in rosso;

I risultati verranno, come sempre, separati per ogni zona climatica.

4.1 Metriche di validazione e coefficienti ottimizzati

Nella tabella di seguito si possono osservare le metriche di valutazione per ogni Paese confrontando tra loro tutti i PMV presi in esame, con i rispettivi coefficienti ottimizzati.

Tabella 5: miglioramento percentuale di MAE e MSE

Differenza %	Francia	Svezia	Portogallo
MAE	5.31%	8.52%	17.19%
MSE	10.20%	15.25%	31.40%

Tabella 6: MAE, MSE e coefficienti dei modelli a confronto

	sPMV _{OTT} - PMV _{standard}		PMV _{Buratti} - PMV _{standard}		PMV _{Buratti} - sPMV _{OTT}		Coefficienti Ottimizzati			
	MAE	MSE	MAE	MSE	MAE	MSE	Icl: 1.5-1.0	Icl: 0.99-0.81	Icl: 0.8-0.51	Icl: 0.5-0.2
Francia	0.18	0.05	0.20	0.06	0.05	0.005	a (ta): 0.1025 b (pv): 0.5029 c : -1.448	a (ta): 0.1017 b (pv): 0.425 c : 2.5045	a (ta): 0.15168 b (pv): -0.1787 c : -3.098	a (ta): 0.1664 b (pv): 0.0360 c : -3.780
Svezia	0.18	0.05	0.20	0.06	0.11	0.02	a (ta): 0.076 b (pv): 0.2769 c : 1.7139	a (ta): 0.1253 b (pv): 0.1953 c : 2.8667	a (ta): 0.1391 b (pv): 0.1207 c : 3.3579	a (ta): -0.2851 b (pv): 0.5620 c : -6.2674
Portogallo	0.22	0.08	0.28	0.12	0.14	0.04	a (ta): 0.1390 b (pv): 2.5912 c : 5.4318	a (ta): 0.0983 b (pv): 0.2883 c : 2.1689	a (ta): 0.1185 b (pv): 0.2192 c : 2.7955	a (ta): 0.1552 b (pv): 0.3268 c : 3.8501
Media	0.19	0.06	0.23	0.08	0.10	0.02				

I risultati riportati suggeriscono che gli errori tra il modello Buratti [5] e il nuovo modello ottimizzato siano calati: per esprimere questo cambiamento, senza dubbio positivo, si è usato il miglioramento percentuale secondo l'Equazione 12. Nella tabella 6 vengono visualizzate queste ulteriori metriche.

4.2 sPMV ottimizzato

Seguono i grafici di confronto tra il PMV standard calcolato secondo ISO 7730, il PMV ottimizzato appena studiato, e il PMV di Buratti.

Francia

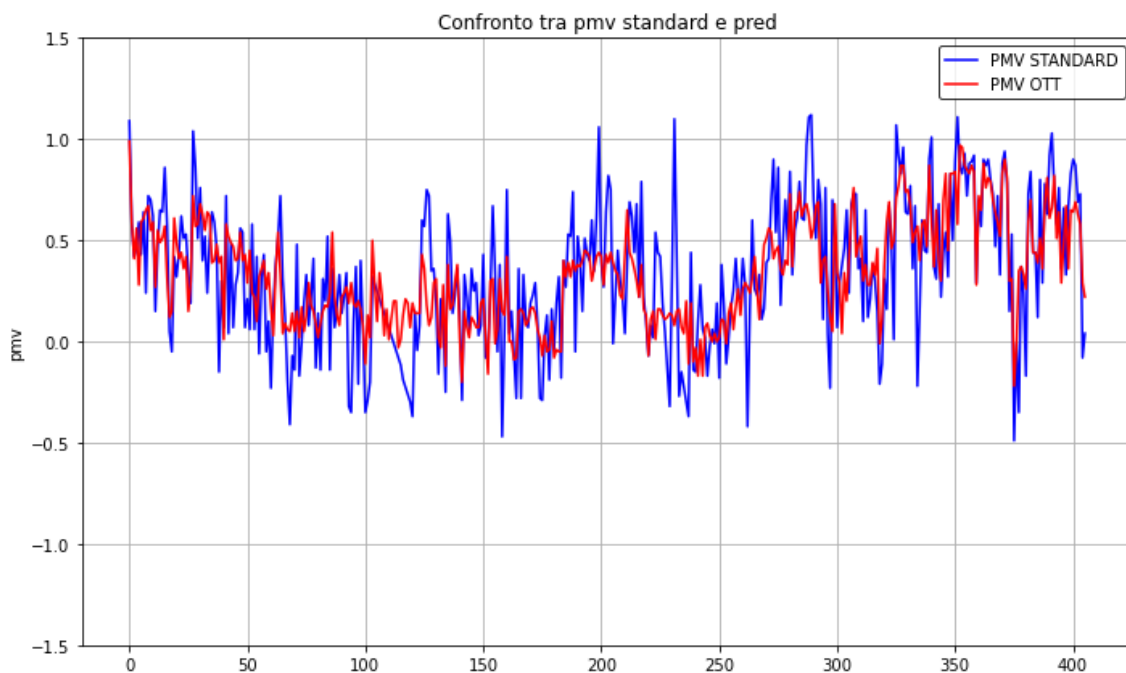


Figura 323: Francia - Confronto PMV_{OTT} – PMV standard - MAE: 0.18, MSE: 0.05

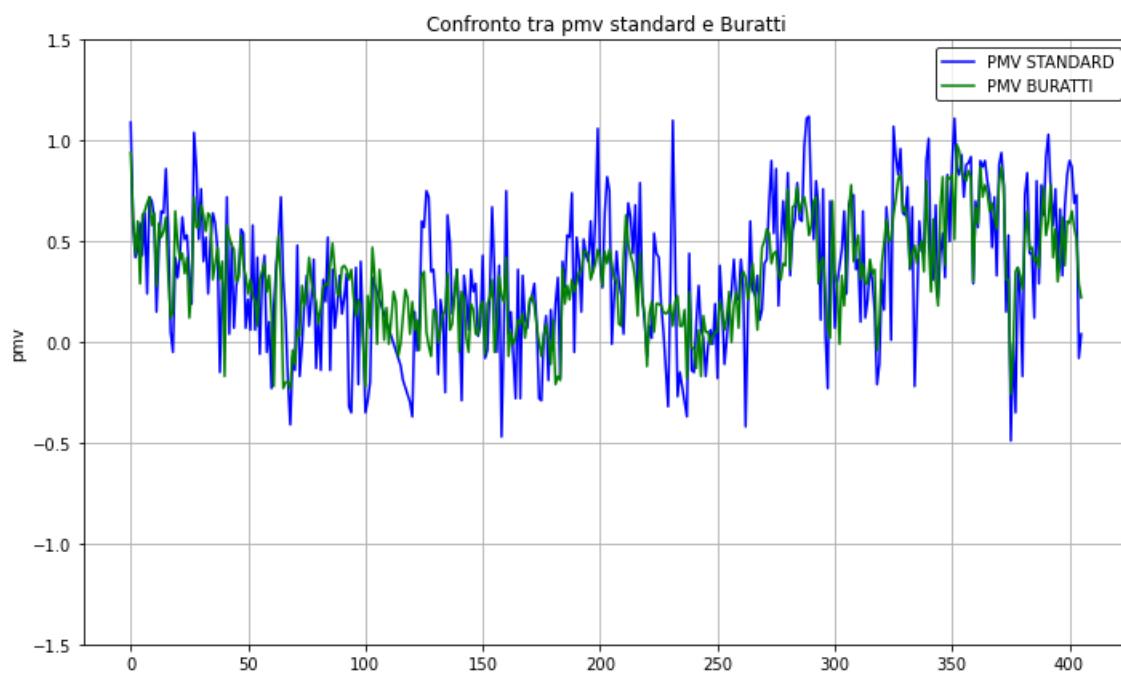


Figura 334: Francia - Confronto PMV Buratti – PMV standard - MAE: 0.20, MSE: 0.06

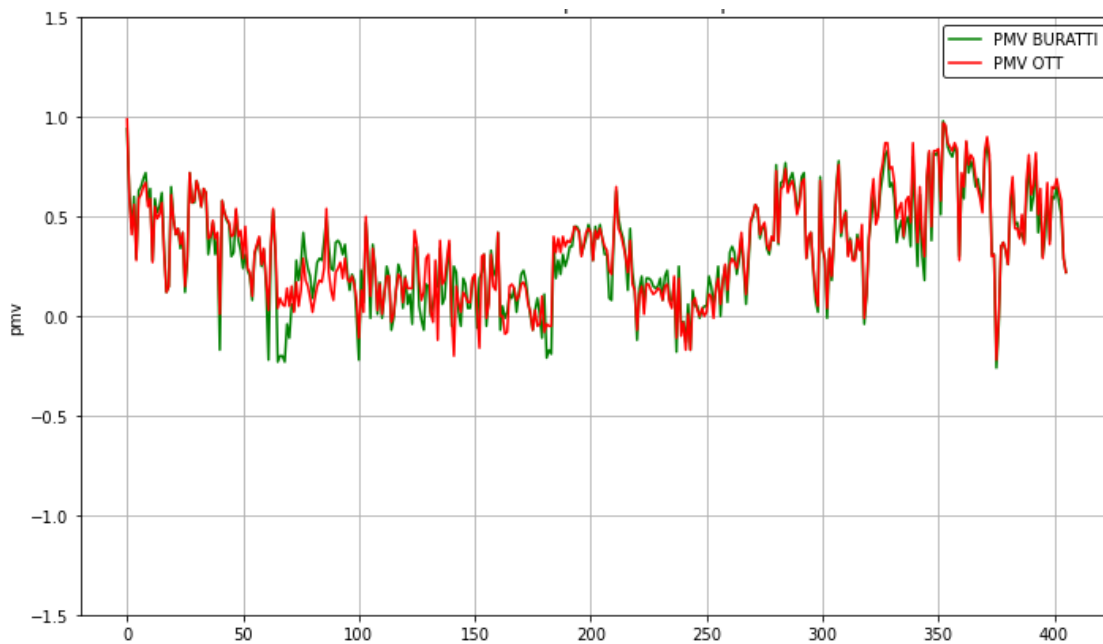


Figura 345: Francia - Confronto PMV Buratti – PMV_{OTT} - MAE: 0.05, MSE: 0.005

Svezia

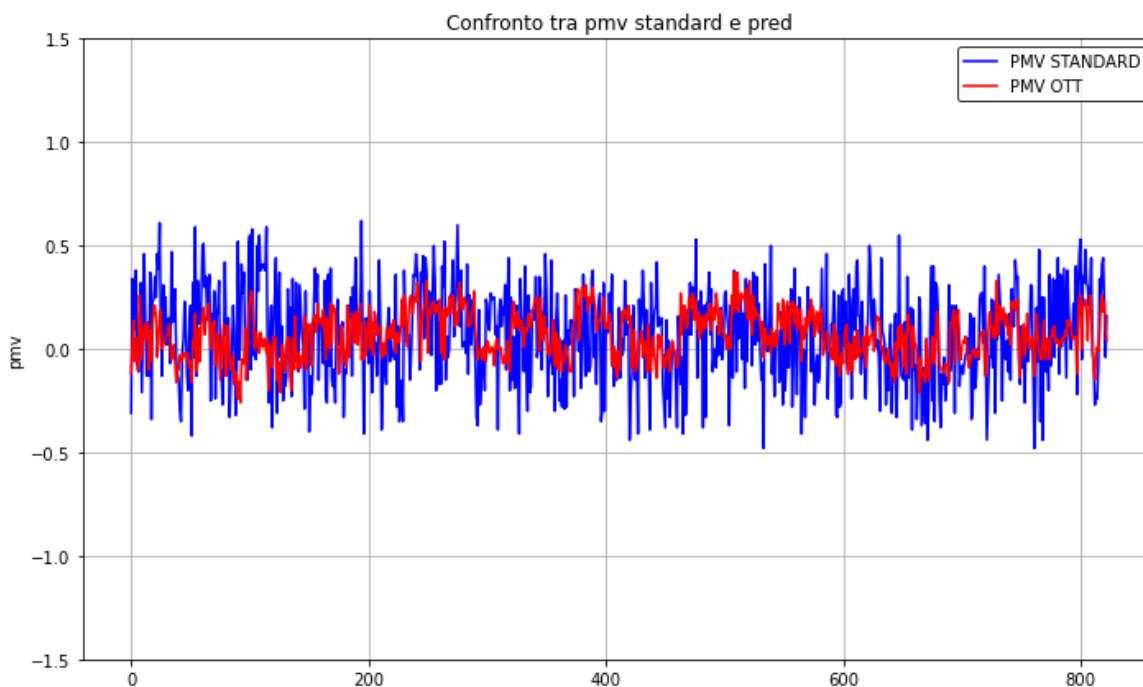


Figura 36: Svezia - Confronto PMV_{OTT} – PMV standard - MAE: 0.18, MSE: 0.05

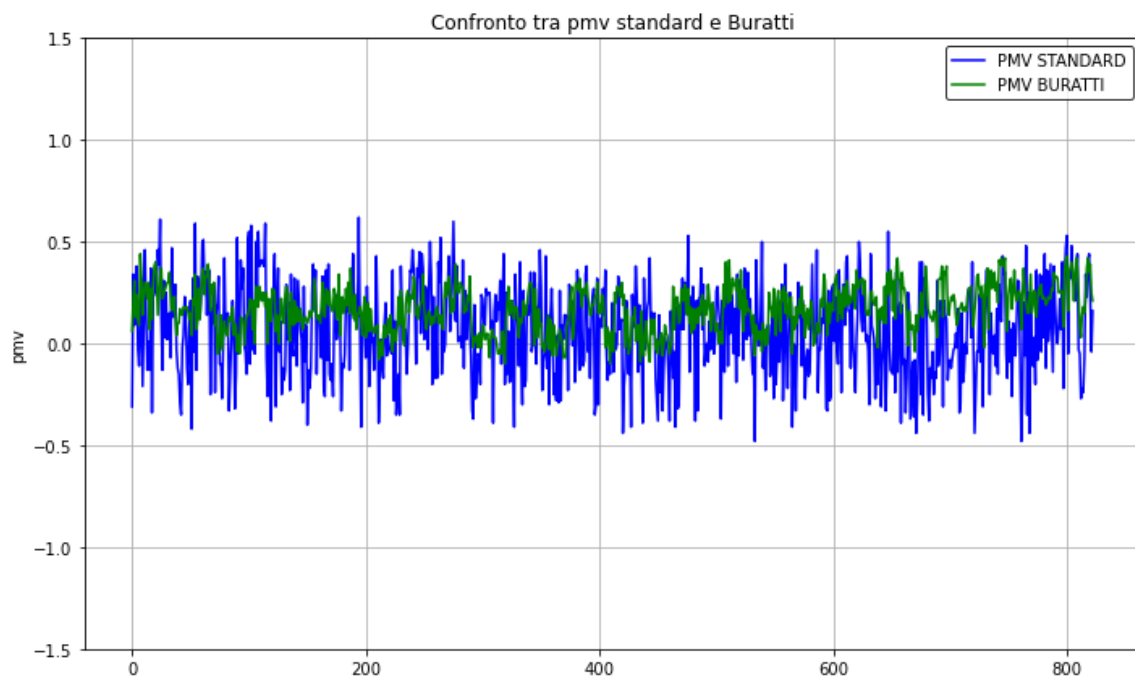


Figura 357: Svezia - Confronto PMV Buratti – PMV standard - MAE: 0.20, MSE: 0.06

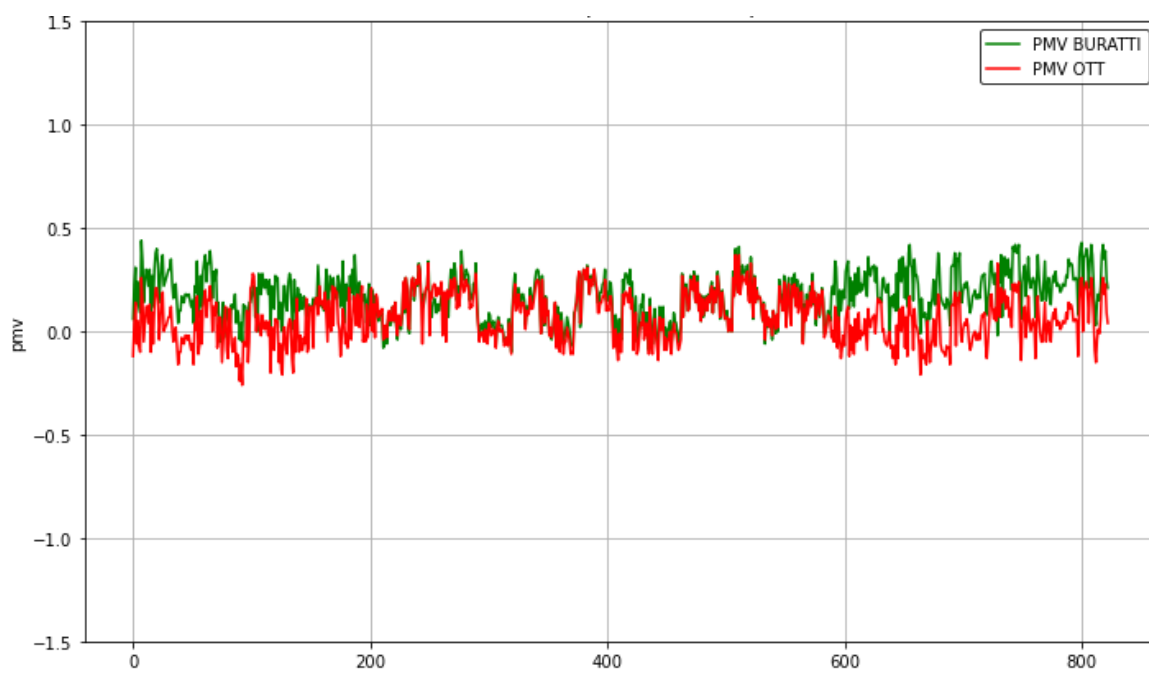


Figura 3836: Svezia - Confronto PMV Buratti – PMV_{OTT} - MAE: 0.11, MSE: 0.02

Portogallo

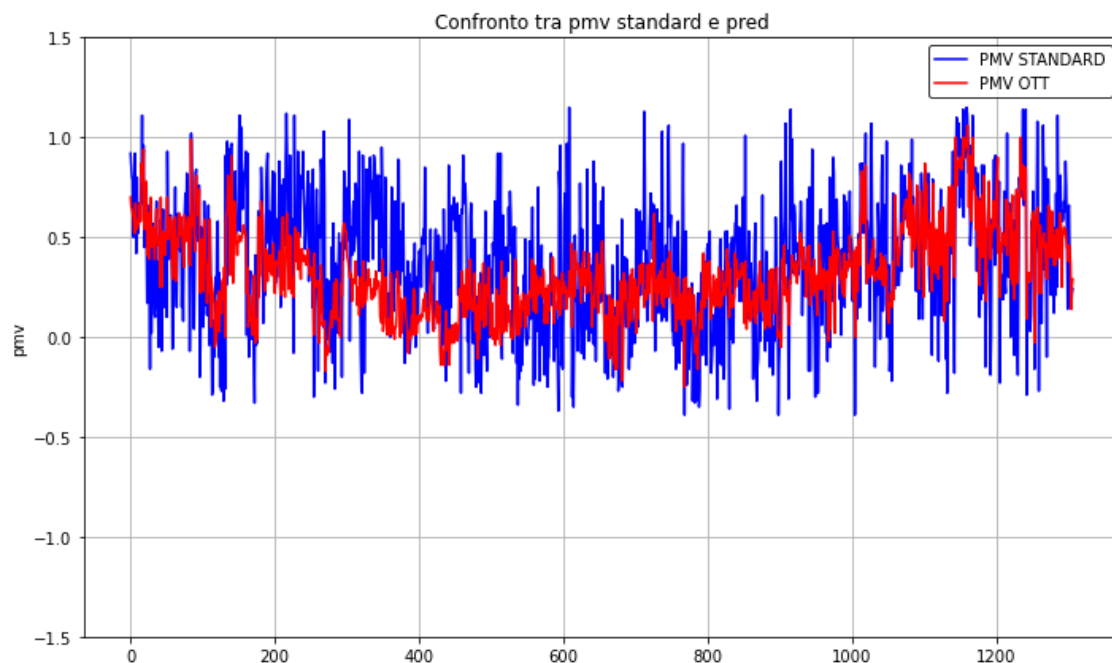


Figura 39: Portogallo - Confronto PMV_{OTT} - $PMV_{standard}$ - MAE: 0.22, MSE: 0.08

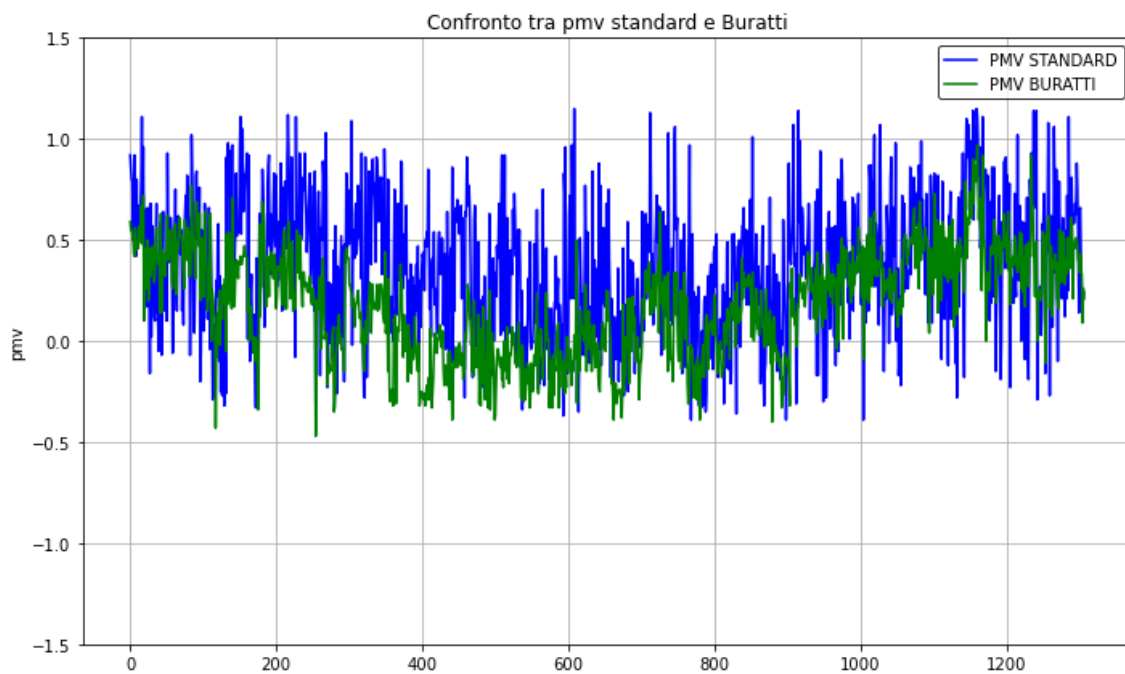


Figura 40: Portogallo - Confronto $PMV_{Buratti}$ - $PMV_{standard}$ - MAE: 0.28, MSE: 0.12

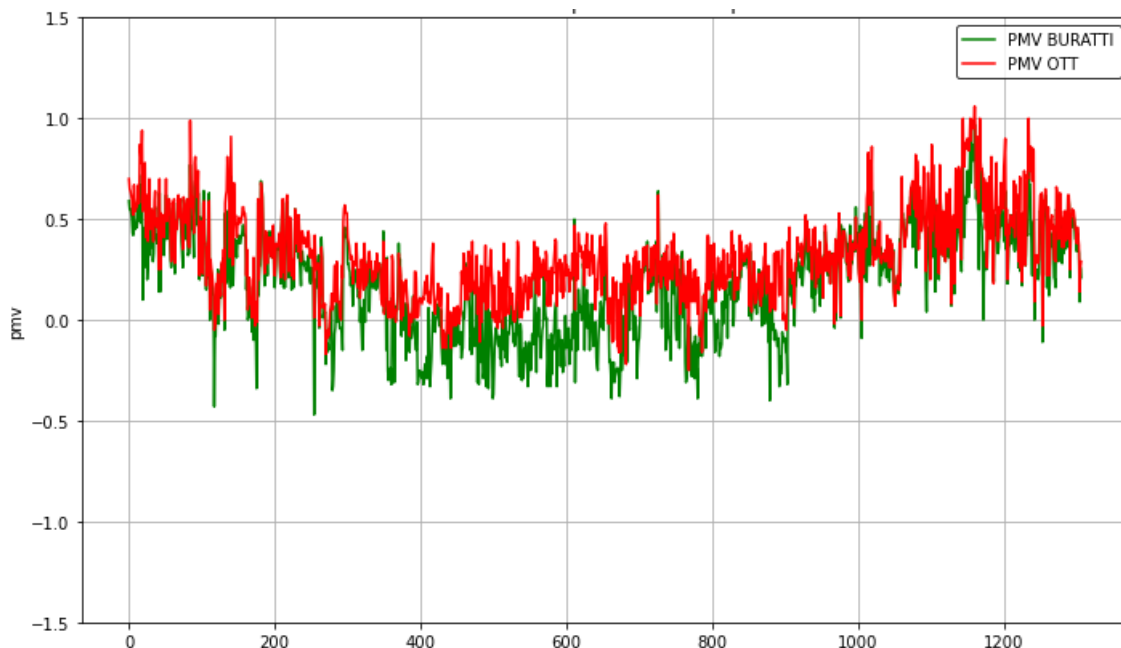


Figura 41: Portogallo - Confronto PMV Buratti – PMV_{OTT} - MAE: 0.11, MSE: 0.02

5. DISCUSSIONE

Avendo analizzato i risultati dello studio, è evidente che il modello sPMV ottimizzato apporti miglioramenti significativi nella misura del comfort rispetto al precedente proposto da Buratti et al. Le metriche di valutazione, ossia il MAE e il MSE, mostrano incrementi percentuali importanti in diversi contesti climatici, a seconda del Paese e del numero di dati disponibili. Questi miglioramenti offrono spunti interessanti da discutere in termini di efficacia del modello e della sua applicabilità in condizioni reali.

I miglioramenti percentuali osservati nelle metriche MAE e MSE confermano la maggiore precisione del nuovo modello rispetto al precedente basato sui coefficienti di Buratti [5]. Nello specifico, il miglioramento del MAE è del 5.31% per la Francia, dell'8.52% per la Svezia e del 17.19% per il Portogallo. Per quanto riguarda il MSE, i miglioramenti sono ancora più rilevanti: 10.20% in Francia, 15.25% in Svezia e 31.40% in Portogallo.

Le differenze nei miglioramenti percentuali tra i diversi Paesi possono essere in parte attribuite alla quantità di dati disponibili per ogni Paese. Come osservato, il Portogallo presenta il miglioramento maggiore, con una riduzione del MAE e del MSE rispettivamente del 17.19% e del 31.40%. Questo potrebbe essere legato al fatto che per il Portogallo si dispone di 1307

misurazioni, il che rende il dataset più ampio e robusto, permettendo al modello di affinare meglio l'implementazione dell'sPMV.

La Francia, con un dataset di 406 dati, mostra miglioramenti più modesti, ma comunque significativi. La quantità di dati più ridotta potrebbe aver limitato la capacità del modello di apprendere correttamente tutte le variabili climatiche e ambientali influenti sul PMV. Questo suggerisce che un volume maggiore di dati consenta una rappresentazione più accurata delle condizioni climatiche locali e quindi una migliore calibrazione del modello.

6. CONCLUSIONI

Attualmente, i modelli disponibili per il calcolo del PMV presentano alcuni limiti. Il modello standard non tiene conto dell'adattamento degli utenti [13] e risulta difficile da applicare in edifici comuni a causa della mancanza di sensori adeguati. D'altra parte, i modelli semplificati spesso non risultano sufficientemente accurati, poiché non considerano la diversità presente tra le zone climatiche. [24] Con l'obiettivo di voler migliorare la performance degli attuali modelli di calcolo del PMV semplificato, mantenendo allo stesso tempo la semplicità e il carattere generale della formula di Rohles [3], il lavoro di tesi appena esposto si è orientato verso l'analisi dell'isolamento termico (*clo*). Difatti, questo parametro, che possiede un carattere intrinsecamente soggettivo, in quanto regolato dalle sensazioni del singolo utente, risulta influenzato da numerosi parametri ambientali. Visto il ruolo cruciale che questo svolge in relazione al comfort termico, è stata fondamentale l'analisi di questo parametro mediante il modello di apprendimento supervisionato GAM [22], che ha permesso di individuare tre parametri indipendenti (t_a , t_{out} , rh) con i quali modellare l'andamento di *Icl*. La temperatura interna e l'umidità relativa erano già stati presi in considerazione nella formulazione semplificata del PMV; a questi, è stata aggiunta la temperatura esterna, parametro di partenza della ricerca. Dalla prima analisi tramite il clustering con K-Means [20], è emerso che solo quest'ultimo, come già discusso, non è sufficiente per definire una corrispondenza diretta tra intervalli di temperatura outdoor e di isolamento termico. Grazie all'equazione di terzo grado ricavata tramite regressione lineare (Equazione 10) su tre parametri (t_a , t_{out} , rh), si è ottenuta una modellazione di *Icl* sufficientemente accurata, fondamentale per risolvere il problema relativo alla suddivisione in quattro nuove fasce, una in più rispetto a quanto presente in letteratura.

Considerando i limiti del PMV standard, l'accuratezza del calcolo secondo normativa si scontra con la difficoltà reale di acquisire sei parametri, per i quali un edificio comune non possiede gli adeguati sensori. Al contrario, questo modello impiega solamente tre variabili di input, con le quali prima effettua una stima di I_{cl} in modo tale da collocare la misura nella fascia corrispondente, poi calcola l'sPMV con gli opportuni coefficienti.

A differenza del modello di Buratti, inoltre, il modello ottimizzato presenta coefficienti diversi per le varie zone climatiche, tenendo conto anche dell'adattamento dell'individuo al clima tipico del luogo in cui vive. [13]

Paragonato al PMV standard, il modello proposto dal lavoro di tesi presenta MAE medio pari a 0.19 e MSE medio pari a 0.06, con un miglioramento percentuale rispetto al modello di Buratti di MAE e MSE rispettivamente di 5.31% e 10.20% per la Francia, 8.52% e 15.25% per la Svezia, 17.19% e 31.40% per il Portogallo.

Nell'ottica di migliorare ulteriormente il modello, studi futuri potrebbero analizzare dataset più ampi, espandendo la ricerca a un maggior numero di zone climatiche. In questo modo, si darebbe ancora più importanza all'adattamento degli utenti al clima caratteristico della propria area, assicurando, dunque, un calcolo più accurato.

7. BIBLIOGRAFIA

- [1] P. O. Fanger, «FUNDAMENTALS OF THERMAL COMFORT,» 1988.
- [2] E. C. F. STANDARDIZATION, «UNI EN ISO 7730,» 2005.
- [3] I. A. Instruments, «Thermal Comfort Booklet,» 2002.
- [4] A. S. o. H. R. a. A.-C. E. Inc., «ASHRAE 55,» 2010.
- [5] C. Buratti, P. Ricciardi e V. M., «HVAC systems testing and check: A simplified model to predict thermal comfort conditions in moderate environments,» 2013.
- [6] E. C. F. STANDARDIZATION, «UNI EN ISO 8996,» 2004.
- [7] E. C. F. STANDARDIZATION, «ISO 9920,» 2007.
- [8] E. C. O. STANDARDIZATION, «UNI EN ISO 7726,» 2001.
- [9] E. A. Arens e H. Zhang, «The skin's role in human thermoregulation and comfort,» 2006.
- [10] N. Djongyang, R. Tchinda e D. Njomo, «Thermal comfort: A review paper,» 2010.
- [11] V. Molinaro e S. Del Ferraro, «Microclima Valutazione confort termico - Dipartimento Medicina del Lavoro».
- [12] N. Morresi, S. Casaccia, M. Sorcinelli, M. Arnesano, A. Uriarte, J. I. Torrens-Galdiz e G. M. Revel, «Sensing physiological and environmental quantities to measure human thermal comfort through machine learning techniques,» 2021.
- [13] R. de Dear e G. S. Brager, «Developing an Adaptive Model of Thermal Comfort and Preference,» 1998.
- [14] P. Ricciardi e C. Buratti, «Thermal comfort in open plan offices in northern Italy: An adaptive approach,» 2012.
- [15] K. P. Sinaga e M. S. Yang, «Unsupervised K-Means Clustering Algorithm,» 2020.
- [16] I. Charalampopoulos, «A comparative sensitivity analysis of human thermal comfort indices with generalized additive models,» 2019.
- [17] «GitHub - CenterForTheBuiltEnvironment_ashrae-db-II,» [Online]. Available: <https://github.com/CenterForTheBuiltEnvironment/ashrae-db-II>.
- [18] Britannica, «Köppen climate classification,» [Online]. Available: <https://www.britannica.com/science/climate-change>.

- [19] K. P. Sinaga e M. S. Yang, «Unsupervised K-means clustering algorithm,» 2020.
- [20] M. Suyal e S. Sharma, «A Review on Analysis of K-Means Clustering Machine Learning Algorithm based on Unsupervised Learning,» 2024.
- [21] A. M. Bueno, I. Mendes da Luz, I. L. Niza e E. E. Broday, «Hierarchical and K-means clustering to assess thermal dissatisfaction and productivity in university classrooms,» 2023.
- [22] B. Mullie, «EXPLAINABLE GAM MODELS PREDICTING SEPSIS IN ICU PATIENTS,» 2018.
- [23] Y. Hou, B. Cao, Y. Zhu, H. Zhang, L. Yang e L. Duanmu, «Temporal and spatial heterogeneity of indoor and outdoor temperatures and their relationship with thermal sensation from a global perspective,» 2023.
- [24] Y. Boutahri e A. Tilioua, «Machine learning-based predictive model for thermal comfort and energy optimization in smart buildings,» 2024.