

UNIVERSITÀ DEGLI STUDI DI PADOVA
FACOLTÀ DI INGEGNERIA



Finito di scrivere il giorno 19 aprile 2012 utilizzando L^AT_EX 2_ε

UNIVERSITÀ DEGLI STUDI DI PADOVA
FACOLTÀ DI INGEGNERIA

—
DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE

—
TESI DI LAUREA MAGISTRALE IN INGEGNERIA
DELL'AUTOMAZIONE

TECNICHE DI MACHINE
LEARNING PER LA RICERCA
GUASTI IN IMPIANTI HVAC

RELATORE: CH.MO PROF. ING. ALESSANDRO BEGHI

CORRELATORE: ING. MIRCO RAMPAZZO

LAUREANDO: LORENZO CORSO

ANNO ACCADEMICO 2011-2012

Ai miei genitori...

“ Non trovo un rimpianto, non riesco ad arrendermi a tutti i miei sbagli...”

SUBSONICA - MICROCHIP EMOZIONALE, 1999

Indice

Sommario	XI
Introduzione	XIII
1 Impianti di aereazione VAV	1
1.1 Generalità	1
1.2 Componenti principali	4
2 Modello e controllo di impianti VAVAC	11
2.1 Modellizzazione	12
2.1.1 Caratterizzazione delle zone	12
2.1.2 Batteria di raffreddamento	14
2.1.3 Ventilatori	15
2.1.4 Ipotesi semplificative	17
2.1.5 Modello in spazio di stato	18
2.2 Controllo	21
3 FDD mediante tecniche di machine learning	25
3.1 K-Nearest Neighbors	27
3.2 Support vector machine	28
4 Estensioni delle support vector machine	37
4.1 Kernel Principal Component Analysis	38
4.2 Particle swarm optimization	41

4.3	Probabilità a posteriori	46
5	Risultati sperimentali	51
5.1	Dati di allenamento e di validazione	52
5.2	Risultati sperimentali	53
5.2.1	Malfunzionamento in condizioni operative costanti	54
5.2.2	Malfunzionamento in condizioni operative variabili	62
5.2.3	Malfunzionamento cassette di zona VAV	64
6	Conclusioni e sviluppi futuri	67
A	Metodo di ottimizzazione aurea	69
B	Pseudocodice algoritmo di Platt	73
	Bibliografia	77

Sommario

Il corretto funzionamento degli impianti di tipo HVAC (Heating Ventilating and Air Conditioning) di taglia medio/grande, risulta cruciale al fine di garantire un adeguato livello di comfort agli occupanti degli edifici ed evitare inutili sprechi energetici. Oltre ad un buon sistema di controllo, risulta assai utile un sistema in grado di rilevare e classificare possibili guasti o malfunzionamenti. Questo ambito di ricerca é noto come *fault detection and diagnosis* (FDD). L'obiettivo che si intende perseguire in questo elaborato é quello di ricavare un modello di un semplice sistema HVAC di tipo VAV (*Variable Air Volume*), multi-zona, che permetta di generare dei dati da utilizzare per la valutazione di tecniche di machine learning per la rilevazione di anomalie. Una volta ricavato il modello ed il relativo sistema di controllo si progetteranno dei sistemi di fault detection and diagnosis basati su *k-nearest neighbor* (K-NN) e *support vector machine* (SVM).

Introduzione

Negli ultimi decenni i sistemi di condizionamento di tipo HVAC hanno riscontrato una notevole diffusione ed un numero sempre più crescente di edifici ne prevedono l'installazione. Con *HVAC systems* si intendono quei sistemi che si occupano di riscaldamento (Heating), ventilazione (Ventilating) e condizionamento dell'aria (Air Conditioning). Essi hanno lo scopo di regolare le condizioni dell'aria all'interno di strutture solitamente di grandi dimensioni come industrie, ospedali e sedi di uffici [1]. Sono anche fondamentali nel caso di particolari applicazioni che richiedono controlli accurati di temperatura ed umidità, come nel caso della conservazione di derrate alimentari o laboratori



Figura 1: Rappresentazione di un sistema di condizionamento HVAC

di biologia con organismi viventi. Per questi motivi è importante assicurare il corretto funzionamento di tali impianti, in modo da garantire il desiderato livello di comfort ambientale, sviluppando un opportuno sistema di controllo e integrandolo con un apparato di rilevazione guasti che sia in grado di segnalare prontamente possibili guasti o malfunzionamenti. Vale la pena notare che con l'invecchiamento questa tipologia di impianti, anche in assenza di malfunzionamenti critici, comportano una riduzione del comfort nonché una riduzione delle prestazioni con sprechi energetici dell'ordine del 15 – 30% [2]. Un dato che assume ancora maggior valenza se si pensa che, per esempio, in Cina il 12% del consumo energetico ad uso domestico sia dovuto agli impianti di tipo HVAC; negli USA il dato sale al 44% [3].

Relativamente alla FDD, secondo Isermann [4] si possono individuare tre livelli di supervisione dei sistemi:

- a) monitoraggio: vengono tenute sotto controllo alcune delle variabili e il guasto è semplicemente segnalato all'operatore con degli allarmi quando escono dalle soglie di tolleranza;
- b) protezione automatica: in caso di grandi anomalie il sistema di supervisione reagisce in modo automatico con delle routine di protezione dell'impianto;
- c) supervisione con diagnosi: in base al monitoraggio di alcune variabili vengono elaborate delle *features*, analizzate, e infine in base a delle regole di decisione si interviene nel modo più opportuno.

Nei primi due casi c'è il vantaggio della semplicità e della affidabilità, mentre il terzo permette un'analisi più approfondita, e in particolare consente di:

- evidenziare piccole anomalie di carattere improvviso o progressivo;
- stabilire se la parte interessata è un attuatore, il processo o un sensore;
- localizzare il problema in sistemi a anello chiuso;
- monitorare il sistema anche nei regimi transitori.

In letteratura l'ambito di rivelazione di anomalie nel funzionamento di un sistema è noto con l'espressione *fault detection*; nel paragrafo seguente ne è data una panoramica.

Fault Detection

Facendo riferimento a [5], si possono dare le seguenti definizioni:

- *Fault* (letteralmente difetto) si intende un cambiamento inaspettato nel comportamento del sistema considerato, un malfunzionamento che in generale può anche non avere ripercussioni gravi;
- *Detection* si intende la rilevazione di anomalie che possono verificarsi nel sistema a cui tale monitoraggio è applicato;
- *Isolation* si riferisce alla localizzazione di dove l'anomalia si è verificata;
- *Identification* riguarda la determinazione dell'entità e della causa che l'ha provocata;
- *Diagnosis* si può intendere l'insieme di *isolation* ed *identification*.

In Figura 2 è schematizzato il processo a tre stadi della fault diagnosis.

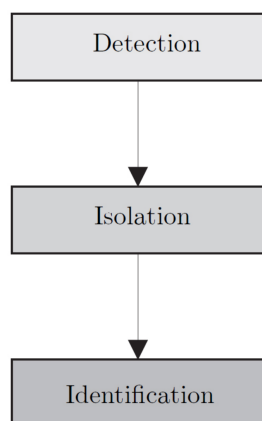


Figura 2: Fault diagnosis a tre stadi.

Tra gli obiettivi della *fault detection and diagnosis* rivestono particolare importanza quelli di aumentare la sicurezza (fondamentale ad esempio per sistemi quali centrali nucleari e veicoli aerei), allungare la vita ed incrementare l'efficienza dei singoli componenti dei sistemi in esame.

Storicamente i primi studi e risultati sulla fault detection risalgono agli anni '70, ma la ricerca riguardante la FDD in riferimento a sistemi di tipo

HVAC ha iniziato a svilupparsi solo verso la fine degli anni '80 - inizio anni 90'. Uno dei principali oggetti di studio consisteva, e consiste tutt'ora, nelle unità di trattamento aria [6] [7] [8], integrate in un impianto più o meno complesso, magari suddiviso in sottoimpianti [7] o interagenti con altre unità in parallelo. A volte lo studio è rivolto maggiormente alla parte sensoristica [9] [10] o componentistica [11], altre volte l'ottica è più generale, come ad esempio nella supervisione di impianto. Nei primi tempi i criteri per la rilevazione di anomalie erano semplicemente degli allarmi, poi si è passati a considerare dei valori di soglia (dapprima fissi, in seguito anche adattabili) per le grandezze più significative, fino a arrivare all'uso di regole di decisione sofisticate che tengono conto di più parametri e talvolta anche di vincoli.

Una classificazione accurata delle varie tecniche di fault detection, accompagnata dalla descrizione dei principali aspetti positivi e negativi, è esposta in [2] (da cui è tratto anche lo schema in Figura 3). A un livello più alto esse possono essere suddivise in *model-based* [3], [12], quando si basano su un modello deciso e costruito in modo più o meno astratto, ma legato alla fisica che caratterizza il sistema, e *process history-based* (o *data-driven*), quando il modello si costruisce direttamente dai dati senza imporre particolari strutture. Nel primo caso si ha una struttura legata alla realtà quindi meglio interpretabile, con lo svantaggio tuttavia che si devono accettare delle approssimazioni di vario tipo, ad esempio perché i moderni sistemi sono molto complessi e spesso fortemente non lineari. Gli esempi sono rappresentazioni in spazio di stato o tramite funzione di trasferimento, alle quali possono essere applicate tecniche di fault detection basate sul filtro di Kalman, come si fa in [11] e [13], oppure inserite in uno schema con un controllore PID con azione feed-forward [14]. Nel secondo caso non si hanno questi problemi, ma spesso si deve ripiegare su rappresentazioni di tipo black-box [8], ad esempio mediante l'impiego di reti neurali che possono richiedere di avere a disposizione una considerevole quantità di dati. Un aspetto della analisi FDD è la fault detection che considera il manifestarsi di più anomalie contemporaneamente (*multi-fault*) [15]. Infine, conseguentemente allo sviluppo dei sistemi di intelligenza artificiale, nuove tecniche per la FDD sono state sviluppate [5]: oltre alle già citate reti neurali, si possono menzionare l'impiego di logiche fuzzy, algoritmi genetici (GA), support vector machine e k-nearest neighbor [16].

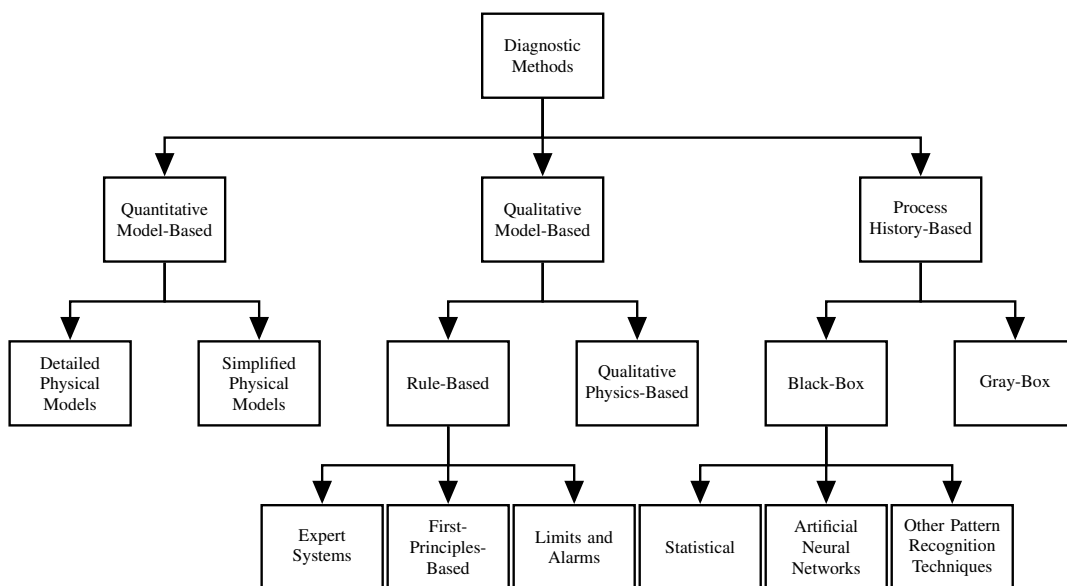


Figura 3: Schematizzazione delle tecniche di fault detection.

Di seguito si espone la struttura della tesi. Nel primo capitolo si descrivono gli impianti a volume d'aria variabile per il condizionamento (VAVAC) e le loro componenti fondamentali. Nel secondo capitolo si modella, attraverso un modello non lineare, un impianto VAVAC multizona nonché si progetta il relativo sistema di controllo giovandosi dell'impiego di una tecnica di *direct feedback linearization*. Nel terzo capitolo si analizzano tecniche di FDD basate sull'apprendimento automatico di tipo black-box concentrandosi in particolare modo sulle support vector machine e k-nearest neighbor. Il capitolo successivo è dedicato alla taratura delle SVM, analizzando sia la fase di pre-processo dei dati mediante il *kernel principal component analysis* (KPCA), che permette sia di scegliere in maniera automatica le variabili di maggior interesse nell'analisi di uno specifico guasto, che di mappare i dati di train in modo da agevolare il compito della support vector machine. Successivamente si presenta un algoritmo euristico di ottimizzazione appartenente alla classe degli algoritmi evolutivi, noto come *particle swarm optimization* (PSO) per tarare opportunamente i parametri delle SVM. Si propone altresì un metodo per associare un'informazione probabilistica alla classificazione fornita dal classifi-

catore SVM. Nel sesto capitolo, a partire da una serie di accurate simulazioni, si valutano le prestazioni della FDD basata sull'apprendimento automatico sfruttando il modello costruito nel secondo capitolo per la generazione dei dati. Considerazioni conclusive e spunti sui possibili sviluppi futuri sono infine riportati.

Impianti di aereazione VAV

1.1 Generalità

Gli impianti a tutta aria a portata variabile (*variabile air volume*, abbreviato VAV) prevedono l'immissione di aria a temperatura costante e portata variabile in funzione dei carichi termici delle varie zone servite [17]. Si parla di zone intendendo uno o più locali (uffici, laboratori, ecc.) che sono simili dal punto di vista degli scambi termici (ad esempio in base a dimensione, utilizzo, esposizione al sole, ecc) e nei quali si desiderano condizioni termoigrometriche simili. La variazione di portata è ottenuta mediante apposite cassette termali di zona regolate tramite un apposito sistema di controllo. Il m più economico prevede ventilatori a velocità variabile, unitamente alla regolazione dell'apertura di serrande dislocate opportunamente nella rete aeraulica. Di norma, il sistema è progettato per mantenere costante la pressione statica nel condotto di mandata a opera del ventilatore sempre di mandata. A sua volta il ventilatore di ripresa ha il compito di mantenere costante la pressione statica nella stanza.

Di seguito si illustrano sinteticamente i principali tipi di impianti VAV.

- Semplice: l'aria viene interamente trattata in una unità centrale detta unità di trattamento aria (UTA o AHU, dall'inglese Air Handling Unit) e poi immessa nella rete aeraulica costituita da condotti semplici (*single duct*) grazie al ventilatore di mandata. Nelle cassette termali di zona si ha solo una variazione della portata. È la classe che si considera in questa tesi.

- Con post riscaldamento locale: nelle cassette termali è presente una batteria di riscaldamento. Questa soluzione è adottata, ad esempio, per le zone perimetrali di un edificio, le quali sono più soggette a variazione dei carichi in base alle condizioni esterne.
- Con unità di ricircolo locale: l'unità centrale tratta solamente aria primaria (aria esterna) che viene inviata a speciali cassette di zona, dette FAT (*fan assisted terminal*), dotate di un piccolo ventilatore che permette di ricircolare aria di ripresa della zona, di miscelarla con aria primaria e di reimmetterla in ambiente. Anche in questo caso è possibile prevedere una batteria di post-riscaldamento.
- A doppio condotto (*dual duct*): sono costituiti da una rete a doppio canale in cui circolano separatamente aria fredda e aria calda; la miscelazione avviene in apposite cassette in prossimità delle zone. Gli impianti *single duct* permettono di risparmiare fino al 30% di energia rispetto agli impianti *dual duct*, sebbene questi ultimi consentano un miglior controllo della qualità dell'aria.

Una schematizzazione di un sistema HVAC a volume d'aria variabile di tipo *single duct* è proposta in Figura 1.1 (tratta da una pubblicazione dell'ABB SACE s.p.a; altri schemi simili sono presenti in [3] [18] [19] [20] [21]).

L'aria esterna (detta di rinnovo) e l'aria di ripresa (o di ricircolo) vengono convogliate verso l'unità di trattamento dell'aria (3.) dove è possibile regolare la temperatura e l'umidità dell'aria di ingresso (8.). I principali componenti di tale unità sono il ventilatore di mandata (2.), la batteria di riscaldamento (5.), la batteria di raffreddamento (6.), il filtro (7.) ed il dispositivo di controllo dell'umidità (non presente in figura). La temperatura dei singoli locali è misurata tramite termostati (11.) che regolano direttamente le cassette termali di ciascuna stanza. La velocità del ventilatore di mandata, che convoglia l'aria nelle singole stanze di tutto l'edificio attraverso gli appositi condotti, è controllata tramite un azionamento in c.a. (4.). Esso regola la portata dell'aria mantenendo costante la pressione statica, misurata dal sensore (13.). Il ventilatore dell'aria di ripresa (1.) espelle aria all'esterno dell'edificio (9.), oppure, una parte di essa (10.) viene ricircolata verso l'unità di trattamento aria. Il ventilatore di ripresa viene controllato tramite un azionamento in c.a. (14.)

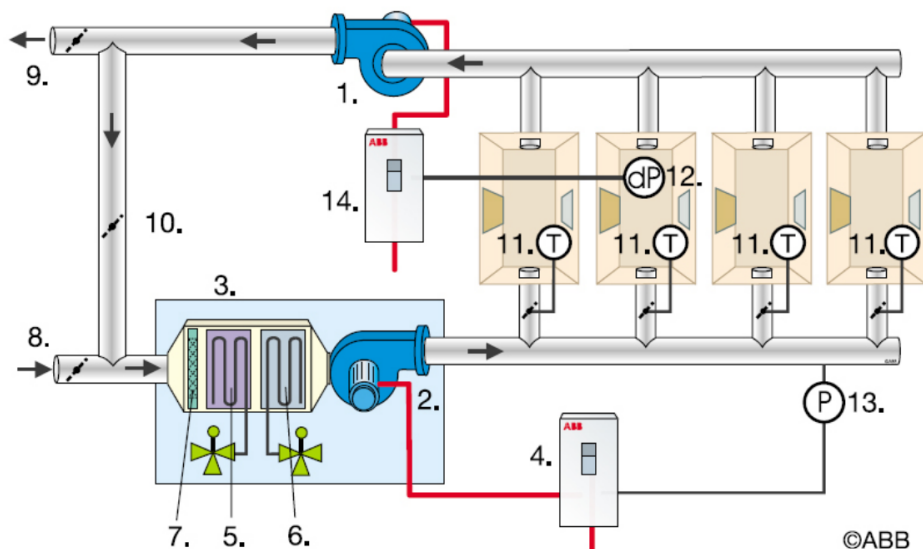


Figura 1.1: Schema di un sistema HVAC a volume variabile d'aria.

che mantiene costante il differenziale di pressione (12.) rispetto alla pressione esterna.

Gli impianti VAV presentano i seguenti aspetti positivi:

- ottimo controllo della temperatura;
- possibilità di utilizzare il *free-cooling* nelle mezze stagioni (consiste nel raffreddare gli ambienti senza utilizzare la pompa di calore ma bensì sfruttando l'aria esterna quando presenta una temperatura un po' più bassa di quella interna);
- ridotti ingombri;
- ridotti consumi energetici (nella versione standard) e costi di esercizio;
- minor portata di aria trattata rispetto ad altre soluzioni;
- manutenzione in zone non occupate;
- possibilità di regolazione locale da parte dell'utente.

Uno svantaggio di tali impianti è la difficoltà di esercizio a carichi ridotti e controllo della qualità ed umidità dell'aria.

1.2 Componenti principali

Prima di procedere alla descrizione di ogni singolo componente dell'impianto se ne fornisce una classificazione d'insieme. Un sistema di tipo VAV, come evidenziato in Figure 1.2, è composto da quattro sottosistemi principali: psicrometrico, alimentazione, zone termiche e scarico.

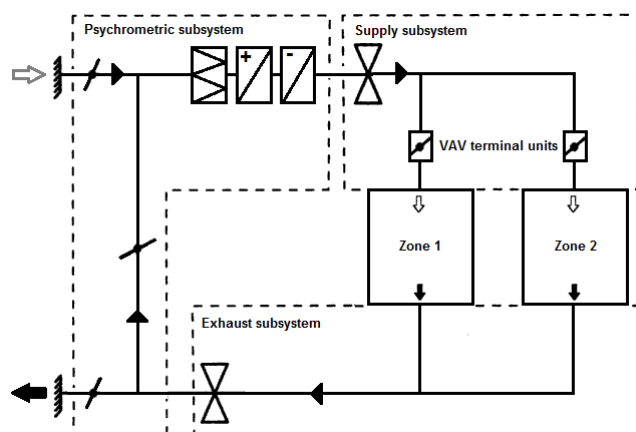


Figura 1.2: Sistema VAV a 2 stanze e singolo condotto

Gli elementi del sottosistema psicrometrico sono i componenti dell'impianto coinvolti nel trattamento dell'aria: chiller, filtri dell'aria, batteria di raffreddamento, valvole per il ricircolo dell'aria, ecc.

Il sistema di alimentazione comprende il ventilatore a velocità variabile, le condotte di mandata e le cassette termali VAV. Il sottosistema di zona include il controllo di ogni singolo locale. Infine, lo scarico estrae aria da ogni stanza mediante il ventilatore di ritorno, l'insieme delle condotte e le cassette termali. Di seguito viene esposta una descrizione accurata di ogni componente.

Chiller

Un chiller (refrigeratore) è una macchina che rimuove calore da un liquido attraverso un ciclo di compressione di vapore, o attraverso un ciclo ad assorbimento. Il liquido che viene raffreddato è di solito acqua, anche se possono essere presenti altre sostanze in percentuale variabile, come inibitori di corrosione o altri additivi. L'acqua fredda in uscita dal chiller viene utilizzata per raffreddare e deumidificare l'aria in strutture commerciali, industriali e istituzionali di medio-grandi dimensioni.

In questo studio non si entrerà nel dettaglio del suo funzionamento, ma si assumerà che operi correttamente, con regolazione sull'acqua di mandata¹, in modo che fornisca acqua fredda alla temperatura nominale di 7°C.

Unità di trattamento aria

Una unità di trattamento aria è di norma costituita da un grande involucro di metallo che, come detto poco sopra, contiene al suo interno un ventilatore, batterie di raffreddamento e riscaldamento, filtri, limitatori di rumore, valvole [1]. Questo sistema è collegato ai condotti per la distribuzione dell'aria condizionata, che trasportano quest'ultima nelle zone, facendola poi ritornare nell'unità stessa. Nel caso vengano utilizzate per il riscaldamento, le unità possono contenere al loro interno un bruciatore oppure semplicemente delle serpentine riscaldate usando acqua o vapore caldi, forniti da un boiler centrale. Altre possibili soluzioni prevedono l'impiego di resistenze elettriche o pompe di calore. Nel caso in cui l'unità venga usata per il raffreddamento, si possono trovare al suo interno degli evaporatori o semplicemente delle serpentine percorse da acqua fredda proveniente dal chiller. Naturalmente possono essere previsti entrambi gli utilizzi. All'interno di una UTA vengono inoltre predisposti un dispositivo per il filtraggio dell'aria ed uno per l'umidificazione della stessa. Infine, le UTA permettono l'introduzione di aria esterna nell'edificio e l'estrazione di aria esausta dall'interno: si parla rispettivamente di portata di rinnovo e portata di espulsione.

Batteria di raffreddamento (*cooling coil*)

All'interno della UTA, l'aria proveniente dal miscelatore attraversa la batteria di raffreddamento, tipicamente una serpentina alettata percorsa da acqua fredda proveniente dal chiller, con lo scopo di estrarre calore dall'aria e di conseguenza raffreddarla. Il calore è dovuto principalmente ai carichi sensibili ed in secondo luogo ai carichi latenti, relativi all'aria umida. Per quanto concerne l'umidità va sottolineato che la regolazione igrometrica è una procedura

¹La regolazione sulla mandata è da preferire alla regolazione sul ritorno non solo perché garantisce una temperatura media in uscita costante e senza elevati scostamenti, ma anche in quanto permette tarature più fini nella progettazione del gruppo frigorifero. Per contro, è necessario assicurare la corretta installazione della sonda in uscita (cosa non sempre facile perché va rispettata una certa distanza dall'uscita dell'evaporatore) e una buona qualità di tale componente.

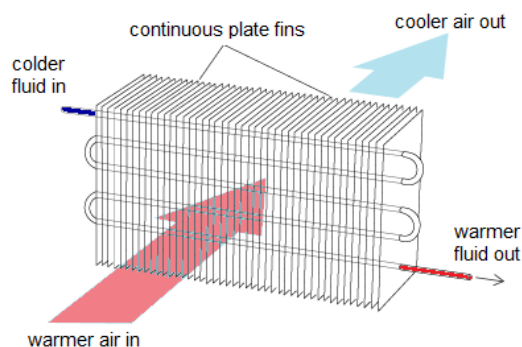


Figura 1.3: Batteria di raffreddamento.

abbastanza delicata quando coinvolge aria fredda: basti pensare al rischio di formazione di condensa. Questo inconveniente non si presenta invece quando si considera la batteria in riscaldamento. In questa tesi, in prima approssimazione, non si considerano gli effetti legati all'umidità ma solo quelli relativi alla temperatura dell'aria.

Ventilatori

Sono macchine in grado di trasferire all'aria l'energia necessaria per percorrere il circuito aeraulico, compensando le perdite di carico. Aumentando la pressione totale dell'aria dall'ingresso all'uscita, ne modificano la velocità e perciò generano un flusso in una direzione prefissata. Sono comandati da inverter per la regolazione di velocità, che viene effettuata tramite l'impiego di sonde di pressione poste nei canali. Esistono tre categoria principali di ventilatori: assiali (in cui il flusso d'aria è parallelo all'asse di rotazione delle pale), centrifughi (in cui il percorso dell'aria segue una direzione assiale in entrata e parallela ad un piano radiale in uscita) e tangenziali (in cui il percorso dell'aria nella girante è normale rispetto all'asse sia in entrata che in uscita). Nel contesto considerato i ventilatori sono due di tipo centrifugo: il primo all'interno della UTA, che rifornisce aria alle zone climatizzate tramite i condotti di mandata principale e secondari; il secondo che estrae aria dalle zone stesse e la immette nel condotto di ritorno (parte di quest'aria può essere rimessa in ricircolo). Le prestazioni di questi dispositivi sono descritte da un grafico che mette in relazione il flusso d'aria con la variazione di pressione statica o totale generata (detta prevalenza).



Figura 1.4: Ventilatore centrifugo.

Condotti

Il sistema di condotti permette di veicolare l'aria trattata nella UTA verso le zone climatizzate e viceversa; permette inoltre il collegamento con l'esterno per il prelievo dell'aria di rinnovo e l'espulsione dell'aria esausta. All'interno del condotto di mandata deve essere mantenuto un livello minimo di pressione statica per far sì che l'aria possa accedere alle stanze attraverso le cassette termali. Come descritto in [22], un valore troppo basso di pressione statica non garantirebbe un flusso sufficiente nemmeno con le bocchette aperte al 100%, mentre un valore troppo elevato comporterebbe aspetti negativi come spreco di energia (per il funzionamento del ventilatore), rumorosità (poiché le bocchette opererebbero a percentuali di apertura più basse), difficoltà di controllo (in quanto le bocchette agirebbero in un intervallo ristretto di valori). Il valore di pressione statica consente tra l'altro di classificare i sistemi di condotti: ad esempio fino a 500Pa sono a bassa pressione, con 1000Pa a media pressione, oltre i 2500Pa ad alta pressione; oppure li si distingue per uso residenziale (fino a 250Pa), commerciale (fino a 2500Pa) e industriale. Il posizionamento del sensore di pressione è un aspetto delicato: teoricamente andrebbe posto dove la pressione statica è minore, ma il punto in cui ciò si verifica potrebbe non essere sempre lo stesso; inoltre le geometrie della rete di canali potrebbero ulteriormente complicare la determinazione della posizione più opportuna per il sensore.. Un criterio di scelta è di porlo a due terzi della lunghezza del canale principale; un'altra soluzione è posizionarne più di uno. Per approfondimenti

di carattere generale o su design e progettazione si veda [1] (capitolo 18) e [23] (capitolo 35).

Serrande (*dampers*) Sono dei regolatori di portata, hanno il compito di mantenere il valore assegnato di portata d'aria nel canale in cui sono collocati indipendentemente dalle oscillazioni di pressione statica a monte di essi. Possono essere a portata costante o variabile, nel secondo caso la portata d'aria è determinata attraverso una grandezza misurata da un regolatore (in generale la temperatura, ma anche la pressione statica in ambiente o il differenziale tra l'ambiente e un volume di riferimento). Le serrande qui considerate sono a portata d'aria variabile in cui il controllo dell'apertura regola i flussi di aria di rinnovo, di espulsione e di ricircolo. La prima è posta in corrispondenza della presa d'aria esterna, la seconda comunica anch'essa con l'esterno ma per espellere aria mentre la terza è collocata nel condotto di ricircolo dalle zone alla UTA. Le serrande possono essere a sezione rettangolare o circolare. Le più comuni sono le seconde, con farfalla di chiusura all'interno, un loro schema è riportato in Figura 1.5. Con riferimento a [21], dalla Figura 1.6 si nota che la relazione tra angolo di apertura e flusso d'aria è non lineare, però nella stessa fonte viene consigliato di approssimare tale andamento con una spezzata. Un ulteriore modello viene proposto in [24], con una relazione di tipo esponenziale in cui i coefficienti devono essere determinati dai dati del dispositivo forniti dal produttore. Infine in [25] il flusso d'aria volumetrico è espresso come $\dot{V}_{dam} = \sqrt{\frac{\Delta P_{dam}}{R_{dam}}}$ in cui ΔP_{dam} è la differenza di pressione tra i due lati e il coefficiente di resistenza R_{dam} dipende dall'angolo di apertura θ .

L'attuatore che determina il movimento della farfalla è di tipo elettrico e si può assumere proporzionalità diretta tra angolo e comando di azionamento in Volt. In tal modo si avrebbe un'attuazione di tipo continuo, mentre in [21] si consiglia di comandare il movimento della farfalla con degli impulsi di tipo apri/chiedi (con riferimento ad un angolo fisso, ad esempio 1.5% di 90°).

Cassette terminali VAV

Sono un particolare tipo di regolatore di portata, presenti nel canale di mandata (secondario) di ciascuna delle zone climatizzate. Essenzialmente sono costituite da un involucro di lamiera contenente un dispositivo di misura di portata,

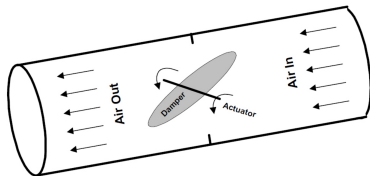


Figura 1.5: Semplice schematizzazione di una serranda a sezione circolare

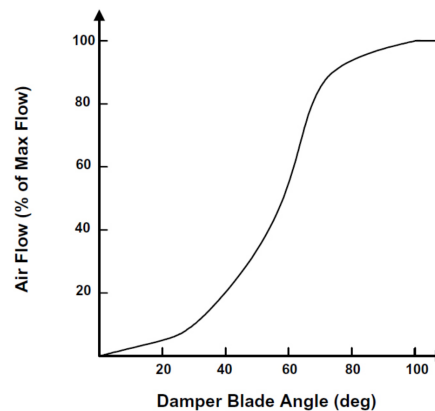


Figura 1.6: Flusso d'aria nella serranda in funzione dell'angolo di apertura.

un regolatore di portata, e un attenuatore acustico: un'immagine si ha in Figura 1.7. Possono essere *pressure-dependent*, nel caso in cui il flusso d'aria sia dipendente dalla pressione statica nel canale di mandata principale, oppure *pressure-independent*, nel caso in cui esso non ne dipenda (in questo caso è presente il dispositivo di misura di portata che rivela la velocità dell'aria o la pressione differenziale e la comunica al controllore, il quale ne tiene conto nella regolazione dell'apertura). Di norma si preferisce la seconda classe in quanto si riesce a ottenere un miglior controllo di portata. Tali cassette agiscono in base ad un segnale proveniente da sonde ambiente (sensori di temperatura) posti nei relativi locali o zone. A seguito, ad esempio, della riduzione di portata della cassetta si verifica un aumento della pressione nel canale principale di mandata, un pressostato provvede quindi a ridurre la portata dei ventilatori di mandata e di ripresa (per ogni zona è infatti presente anche una cassetta VAV nel canale di ripresa, la quale, assieme al ventilatore, permette di mantenere una determinata pressione all'interno di essa; in questo caso il controllo è solo di pressione e non di temperatura). In questo studio le cassette saranno modellate come serrande.

Miscelatore (*mixing box*)

È una camera di espansione in cui confluiscono, tramite i condotti, flussi d'aria con diverse caratteristiche (temperatura, umidità, velocità) e da cui esce un flusso omogeneo ottenuto dalla loro miscelazione. Quella che qui interessa è



Figura 1.7: Cassetta terminale VAV.

posta prima della UTA e riceve aria di rinnovo e di ricircolo.

Sensori di temperatura

Per il processo di regolazione servono dei sensori che misurano la temperatura nella stanza.

Essi possono essere modellizzati, in prima approssimazione, come un sistema del primo ordine, con una costante di tempo di qualche secondo. La funzione di trasferimento dalla temperatura reale T_z a quella che fornisce il sensore T_s ha la forma:

$$T_s(s) = \frac{\tau}{\tau s + 1} T_z(s).$$

Capitolo 2

Modello e controllo di impianti VAVAC

In questo capitolo, dopo aver introdotto la simbologia utilizzata, si ricava un modello di un impianto VAVAC a due zone e il relativo controllo basato su una tecnica di *direct feedback linearization* facendo riferimento a quanto presentato in [26].

Simbologia

A_z	superficie di una stanza a contatto con l'esterno [m ²]
C_{pa}	calore specifico dell'aria [1005J/kgK]
C_{pw}	calore specifico dell'acqua [4187J/kgK]
\dot{m}_a	flusso d'aria di mandata totale [kg/s]
\dot{m}_w	flusso d'acqua nella batteria [kg/s]
M_c	massa della batteria di raffreddamento [kg]
r	percentuale di apertura serranda di ricircolo
\dot{Q}_1	carichi interni ed esterni stanza 1 [W]
\dot{Q}_2	carichi interni ed esterni stanza 2 [W]
T_{ai}	temperatura aria entrante nella batteria [°C]
T_{ao}	temperatura aria uscente dalla batteria [°C]
T_m	temperatura aria miscelata [°C]
T_{ext}	temperatura aria esterna [°C]
T_r	temperatura aria di ricircolo [°C]
T_{wi}	temperatura acqua entrante batteria [7°C]
T_{wo}	temperatura acqua uscente batteria [°C]
T_{z1}	temperatura della stanza 1 [°C]
T_{z2}	temperatura della stanza 2 [°C]
$(UA)_c$	fattore di trasmittanza superficiale complessiva della batteria [W/K]
U_z	trasmittanza termica superficiale delle stanze [0.3W/m ² K]
V_z	volume di una stanza [m ³]
ρ_a	densità dell'aria [1.25kg/m ³]
ρ_w	densità dell'acqua [998kg/m ³]

2.1 Modellizzazione**2.1.1 Caratterizzazione delle zone**

Oltre all'impianto di distribuzione dell'aria è essenziale considerare anche le zone dove questa viene distribuita. Ci sono infatti molti fattori che intervengono nella caratterizzazione degli scambi termici: collocazione geografica, esposizione al sole, isolamento delle pareti esterne e della copertura, attività e affollamento degli interni, dimensione degli ambienti, ecc. Esistono software avanzati

che permettono di modellizzare sotto questi punti di vista un edificio o parti di esso, come a esempio DESIGNBUILDER™.

In questo studio non si intende entrare nel dettaglio della geometria e delle caratteristiche di uno specifico edificio o della singola zona (stanza), si considera invece un modello di una certa generalità, semplice quanto basta ma sufficientemente accurato al fine di poterlo impiegare per gli scopi di FDD precedentemente illustrati.

Ricordando che:

- secondo il principio di conservazione dell'energia la somma delle energie entranti ed uscenti da un sistema è pari all'energia accumulata dal sistema stesso (dato che questo non compie lavoro),
- il calore posseduto da un gas si esprime come $Q = MC_p\Delta T$, in cui M rappresenta la massa, C_p il calore specifico e ΔT la variazione di temperatura,
- la massa può essere espressa come prodotto tra volume e densità,

per ogni singola zona risulta:

$$Q_a + Q_{int} + Q_{est} + Q_{out} + Q_d = V_z \rho_a C_{pa} \Delta T \quad (2.1)$$

dove:

- Q_a energia dall'aria di mandata [J]
- Q_{int} energia da carichi interni (persone, dispositivi vari) [J]
- Q_{est} energia da carichi esterni (sole) [J]
- Q_{out} energia dell'aria uscente [J]
- Q_d energia dispersa attraverso le pareti [J]
- V_z volume della zona [m³]
- T_z temperatura della zona [°C]

Essendo inoltre:

$$Q_{out} = M_{aria\ uscente} C_p (T_{aria\ uscente} - T_z) \ e \ T_{aria\ uscente} = T_z \quad (2.2)$$

allora:

$$Q_{out} = 0 \quad (2.3)$$

mentre per lo scambio termico attraverso le pareti viene preso in considerazione solo quello con l'esterno (U_z è la trasmittanza termica superficiale delle pareti e A_z l'aria di contatto con l'esterno):

$$Q_d = U_z A_z (T_{ext} - T_z). \quad (2.4)$$

Altresì:

$$Q_a = M_a C_{pa} (T_{sa} - T_z) \quad (2.5)$$

e unendo carichi interni e esterni $Q_{int,est} = Q_{int} + Q_{est}$ l'equazione 2.1 diviene

$$Q_{int,est} + M_a C_{pa} (T_{sa} - T_z) + U_z A_z (T_{ext} - T_z) = V_z \rho_a C_{pa} \Delta T. \quad (2.6)$$

Considerando, infine, le variazioni nel tempo delle grandezze, ΔT diviene $\frac{dT_z}{dt}$, il calore diventa potenza (indicata con il simbolo \dot{Q} [W]), M_a è sostituito dal flusso \dot{m}_a ; dopo qualche semplice manipolazione si ottiene la seguente dinamica per la temperatura della stanza:

$$\boxed{\frac{dT_z}{dt} = \frac{\dot{Q}_{int,est}}{V_z \rho_a C_{pa}} + \frac{\dot{m}_a}{V_z \rho_a} (T_{sa} - T_z) + \frac{U_z A_z}{V_z \rho_a C_{pa}} (T_{ext} - T_z).} \quad (2.7)$$

Si noti che non si è tenuto conto dell'effetto capacitivo, cioè del fatto che le pareti e le altre parti dell'edificio assorbono e rilasciano calore con una certa legge temporale, a esempio prolungando lo scambio di energia anche dopo che una fonte esterna ha smesso di fornirne.

2.1.2 Batteria di raffreddamento

La batteria di raffreddamento (UTA) è la componente del sistema in cui l'aria viene raffreddata cedendo calore all'acqua. Per la sua modellizzazione si è preso a riferimento il modello trattato in [27]:

$$\boxed{M_c C_{pc} \frac{dT_c}{dt} = \dot{m}_w C_{pw} (T_{wi} - T_{wo}) + (UA)_c (T_{ai} - T_c)} \quad (2.8)$$

$$\boxed{\dot{m}_a C_{pa} (T_{ao} - T_{ai}) = (UA)_c (T_c - T_{ai})} \quad (2.9)$$

in cui si suppone che l'acqua fredda entrante nella batteria abbia temperatura costante $T_{wi} = 7^\circ\text{C}$. Il fattore di trasmittanza superficiale del coil $(UA)_c$ e la sua massa M_c sono calcolati supponendo che lo scambiatore in rame sia in grado di soddisfare un carico massimo nominale di 2000W: $U_c = 50\text{W}/\text{m}^2\text{K}$, $A_c = 3.8\text{m}^2$, $M_c = 2.1\text{kg}$. È inoltre ragionevole porre la temperatura dell'acqua in uscita dal coil pari alla temperatura del coil $T_{wo} = T_c$.

Dato che, è di interesse come variano acqua e aria in uscita dalla UTA, fatte le ulteriori e ragionevoli ipotesi che T_{ai} , \dot{m}_a , e \dot{m}_w rimangano costanti mentre avviene lo scambio termico (quindi derivate nulle), derivando la (2.9) si ha:

$$\dot{m}_a C_{pa} dT_{ao}/dt = (UA)_c dT_c/dt = (UA)_c dT_{wo}/dt \quad (2.10)$$

e con alcune manipolazioni si giunge alle seguenti equazioni dinamiche:

$$\frac{dT_{ao}}{dt} = \left[\frac{\dot{m}_w C_{pw}}{\dot{m}_a C_{pa}} (T_{wi} - T_{wo}) + (T_{ai} - T_{ao}) \right] \frac{(UA)_c}{M_c C_{pc}} \quad (2.11)$$

$$\frac{dT_{wo}}{dt} = \left[\dot{m}_w C_{pw} (T_{wi} - T_{wo}) + \dot{m}_a C_{pa} (T_{ai} - T_{ao}) \right] \frac{1}{M_c C_{pc}} \quad (2.12)$$

Analizzando la 2.11 si nota che non è ammesso un flusso d'aria totale \dot{m}_a nullo. Inoltre, quando $\dot{m}_w = 0$, è opportuno sostituire la 2.12 con:

$$dT_{wo}/dt = K(T_{wi} - T_{wo}) \quad \text{se} \quad \dot{m}_w = 0 \quad (2.13)$$

ad esempio con $K = 0.5$ in modo che la temperatura dell'acqua si porti a T_{wi} quando non ne sta circolando, evitando che aumenti ulteriormente a causa di $\dot{m}_a \neq 0$.

Si ricordi che l'attenzione è posta solo sui carichi sensibili e non su quelli latenti, ossia si trascura ciò che riguarda effetti e regolazioni dell'umidità.

2.1.3 Ventilatori

Un modo per controllare la portata d'aria alle bocchette utilizzando ventilatori a velocità variabile è quello di regolare questi ultimi in modo da assicurare un determinato valore di pressione statica nelle condotte, che, come detto in precedenza, non deve essere né troppo elevato né troppo basso. La pressione

statica è funzione delle perdite di carico, che a loro volta dipendono dal flusso d'aria circolante, ma non è facile stabilire tale dipendenza. Inoltre, data la dipendenza della pressione statica dalla tipologia e dalla topologia della rete aureulica, una corretta analisi richiederebbe la caratterizzazione dello specifico impianto VAV. Per mantenere un certo livello di generalità, e non precludere lo studio dei metodi di FDD a questo tipo di componenti, in questo elaborato ci si limita a completare il modello del sistema con una mappa statica che legghi i flussi d'aria alle cassette VAV, la pressione interna al condotto e la velocità del ventilatore. Per ottenere questa relazione si fa riferimento ai dati sperimentali presenti in [28], le cui tabelle, riportate di seguito, mostrano la variazione della pressione statica in funzione della portata dell'aria di alimentazione alle stanze quando la frequenza dei ventilatori è pari a 32, 36 e 40 Hz.

\dot{m}_{a2} Kg/s	\dot{m}_{a1} Kg/s				
	0,0295	0,059	0,0885	0,118	0,1475
0,0295	149	145	139	132	120
0,059	146	138	130	118	108
0,0885	140	131	118	107	95

\dot{m}_{a2} Kg/s	\dot{m}_{a1} Kg/s				
	0,0295	0,059	0,0885	0,118	0,1475
0,0295	197	191	187	181	168
0,059	192	185	177	164	153
0,0885	185	177	166	150	138

\dot{m}_{a2} Kg/s	\dot{m}_{a1} Kg/s					
	0,0295	0,059	0,0885	0,118	0,1475	0,2065
0,0295	243	238	233	228	219	203
0,0885	235	227	216	204	188	177
0,1475	219	206	190	175	164	149

Interpolando linearmente i dati sopra riportati, sfruttando il software DATAFIT, si ottiene:

$$p = 11,95v - 323,38(\dot{m}_a)_1 - 357,29(\dot{m}_a)_2 - 205,38 \quad (2.14)$$

dove p identifica la pressione statica nelle condotte, v la frequenza di lavoro del ventilatore, mentre ma_1 e ma_2 il flusso d'aria di mandata rispettivamente nella prima zona termica e nella seconda; in Figura 2.1 si riporta una rappresentazione dei dati e della curva di regressione individuata.

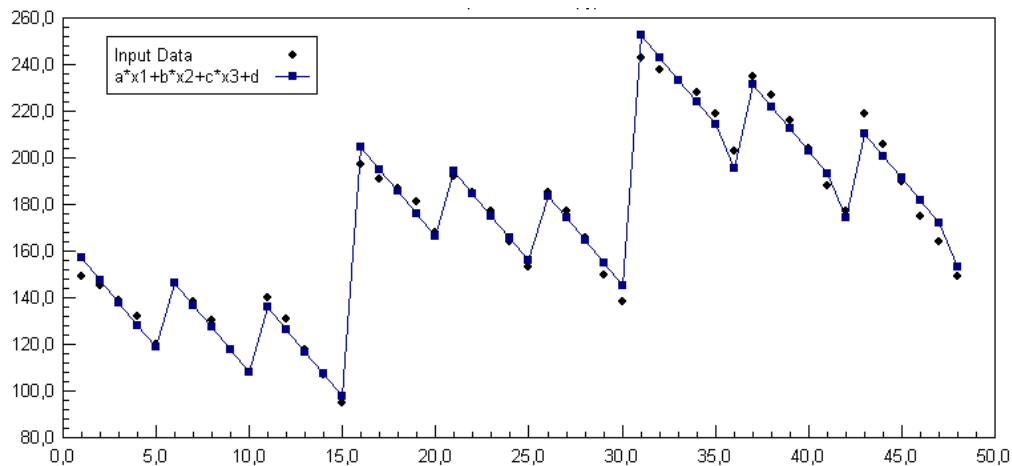


Figura 2.1: Legame tra pressione statica (Pa) e frequenza lavoro del ventilatore (Hz).

2.1.4 Ipotesi semplificative

Come spesso accade quando si deve descrivere una realtà fisica tramite un modello matematico, è utile, se non addirittura necessario, apportare delle semplificazioni, in modo da selezionarne uno sufficientemente accurato e al tempo stesso non più complesso di quanto strettamente necessario.

Oltre a quanto osservato per la batteria di raffreddamento, si possono considerare le seguenti ipotesi:

- le dinamiche dei dispositivi elettro/meccanici possono essere trascurate rispetto a quelle di scambio termico, infatti le prime hanno tempistiche al più di qualche secondo, mentre le ultime sono dell'ordine dei minuti;
- per quanto riguarda l'aria entrante nel miscelatore si suppone che la proporzione tra l'aria di rinnovo e di ricircolo sia di 1 : 3: ciò equivale a considerare che la serranda che preleva aria dall'esterno sia aperta per il 75%, mentre quella di ricircolo per il 25%;

- tutti i parametri sono tempo invarianti;

- si considerano due zone termiche schematizzate a parametri concentrati in cui temperatura, pressione e velocità dell'aria sono uniformemente distribuite;

- le cassette VAV di ripresa e il ventilatore di ripresa sono ignorati in quanto si suppone un comportamento simmetrico rispetto ai corrispondenti di mandata: tanta aria entra in ogni stanza quanta ne esce e la pressione all'interno di essa viene mantenuta costante con un opportuno controllo su questi componenti.

2.1.5 Modello in spazio di stato

Tralasciando per il momento il controllo di pressione, per pervenire alle equazioni che descrivono la dinamica del sistema, si devono considerare le seguenti variabili di stato e i seguenti ingressi di controllo:

$$\begin{aligned}x_1 &= T_{z1} && \text{temperatura aria nella stanza 1} \\x_2 &= T_{z2} && \text{temperatura aria nella stanza 2} \\x_3 &= T_{sa} && \text{temperatura aria di mandata} \\x_4 &= p && \text{pressione statica nel condotto di mandata} \\u_1 &= (\dot{m}_a)_1 && \text{portata d'aria verso la stanza 1} \\u_2 &= (\dot{m}_a)_2 && \text{portata d'aria verso la stanza 2} \\u_3 &= \dot{m}_w && \text{flusso d'acqua nella batteria di raffreddamento} \\u_4 &= v && \text{velocità ventilatore di mandata}\end{aligned}$$

ai quali si aggiungono uno stato per l'acqua nelle batteria (che non viene controllato perché, come illustrato in precedenza, si suppone il chiller in grado

di garantire una temperatura nominale T_{wi} di 7°C) e i disturbi:

$$\begin{aligned} x_5 &= T_{wo} && \text{temperatura acqua uscente dalla batteria} \\ u_5 &= (\dot{Q}_{int,est})_1 && \text{carichi totali stanza 1} \\ u_6 &= (\dot{Q}_{int,est})_2 && \text{carichi totali stanza 2} \\ u_7 &= T_{ext} && \text{temperatura esterna} \end{aligned}$$

Si assume inoltre il vincolo per cui $\dot{m}_a = u_2 + u_3$. L'uscita si considera coincidente con lo stato:

$$y = x. \quad (2.15)$$

Per quanto riguarda le zone è immediato, a partire dell'equazione (2.7), ottenere

$$\dot{x}_1 = a_1 u_1 (x_3 - x_1) + a_2 u_4 + a_3 (u_6 - x_1) \quad (2.16)$$

$$\dot{x}_2 = b_1 u_2 (x_3 - x_2) + b_2 u_5 + b_3 (u_6 - x_2) \quad (2.17)$$

in cui si assume:

$$\begin{aligned} a_1 &= \frac{1}{V_{z1} \rho_a} && b_1 = \frac{1}{V_{z2} \rho_a} \\ a_2 &= \frac{1}{V_{z1} \rho_a C_{pa}} && b_2 = \frac{1}{V_{z2} \rho_a C_{pa}} \\ a_3 &= \frac{U_{z1} A_{z1}}{V_{z1} \rho_a C_{pa}} && b_3 = \frac{U_{z2} A_{z2}}{V_{z2} \rho_a C_{pa}} \end{aligned}$$

Per avere anche la dinamica della batteria di raffreddamento in funzione delle variabili introdotte occorre tener presente quanto riportato in [29]. É innanzi tutto necessario ricavare la temperatura dell'aria che entra nella batteria:

$$\dot{m}_m = \dot{m}_r + \dot{m}_{ext} \quad (2.18)$$

$$\dot{m}_m C_{pa} T_m = \dot{m}_r C_{pa} T_r + \dot{m}_{ext} C_{pa} T_{ext} \quad (2.19)$$

$$\boxed{T_m = \frac{\dot{m}_r T_r + \dot{m}_{ext} T_{ext}}{\dot{m}_r + \dot{m}_{ext}}} \quad (2.20)$$

	\dot{m}_m	flusso di massa d'aria miscelata [kg/s]
	\dot{m}_r	flusso di massa d'aria ricircolata [kg/s]
in cui	\dot{m}_{ext}	flusso di massa d'aria esterna [kg/s]
	T_m	temperatura aria miscelata [°C]
	T_r	temperatura aria di ricircolo [°C]
	T_{ext}	temperatura aria di rinnovo [°C]

Segue dunque che $T_{ai} = T_m$. In particolare, dall'ipotesi sulle proporzioni d'aria entrante nel miscelatore, parametrizzando l'apertura della serranda di ricircolo con r (inizialmente pari a 0.25), la 2.20 diventa semplicemente

$$T_m = rT_r + (1 - r)T_{ext}. \quad (2.21)$$

Per calcolare T_r si considera ancora una miscelazione (sostanzialmente ad opera del ventilatore di ritorno) tra le masse d'aria uscenti delle due stanze:

$$T_r = \frac{(\dot{m}_r T_z)_1 + (\dot{m}_r T_z)_2}{\dot{m}_r} \quad \Rightarrow \quad T_r = \frac{u_1 x_1 + u_2 x_2}{u_1 + u_2}. \quad (2.22)$$

Osservato infine che $T_{ao} = T_{sa}$, la dinamica della batteria diventa:

$$\dot{x}_3 = \left[\frac{C_{pw}}{C_{pa}} \frac{u_3}{u_1 + u_2} (T_{wi} - x_4) + (rT_r + (1 - r)u_6 - x_3) \right] \frac{(UA)_c}{M_c C_{pc}} \quad (2.23)$$

$$\dot{x}_5 = [C_{pw} u_3 (T_{wi} - x_4) + C_{pa} (u_1 + u_2) (rT_r + (1 - r)u_6 - x_3)] / (M_c C_{pc}) \quad (2.24)$$

$$\dot{x}_5 = K(T_{wi} - x_4) \quad \text{se } u_3 = 0 \quad (2.25)$$

Va sottolineato che in questo modello non hanno senso valori negativi per gli ingressi u_1 , u_2 , u_3 e u_5 .

Il modello complessivo, su cui si progetta il controllo nel paragrafo successivo, si riassume come:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 = 0 \end{bmatrix} = \begin{bmatrix} a_1 u_1 (x_3 - x_1) + a_2 u_4 + a_3 (u_6 - x_1) \\ b_1 u_2 (x_3 - x_2) + b_2 u_5 + b_3 (u_6 - x_2) \\ \left[\frac{C_{pw}}{C_{pa}} \frac{u_3}{u_1 + u_2} (T_{wi} - x_4) + (rT_r + (1 - r)u_6 - x_3) \right] \frac{(UA)_c}{M_c C_{pc}} \\ 11,95u_4 - 323,38u_1 - 357,29u_2 - x_4 - 205,38 \end{bmatrix} \quad (2.26)$$

2.2 Controllo

Il sistema ottenuto al paragrafo precedente risulta essere non lineare e accoppiato (questo è dovuto al fatto di aver considerato più di una stanza).

Per controllare il sistema, prendendo spunto da [30], dapprima si disaccoppiano gli ingressi mediante la *direct feedback linearization*, al fine di ottenere un singolo anello e un ingresso di controllo per ognuna delle variabili di stato che si desiderano controllare e, successivamente, si progetta un controllore PID per la regolazione d'anello.

Il motivo per cui si disaccoppiano gli ingressi anziché linearizzare il sistema in un intorno del punto di lavoro, risiede nel fatto che questi sistemi operano in un range elevato di condizioni operative.

Applicando la *direct feedback linearization*, che si basa sul teorema della funzione implicita di Dini, è possibile riportare il modello originale in una forma equivalente ma più semplice.

Consideriamo un sistema controllabile e non lineare espresso mediante equazioni alle differenze:

$$y^{(n)} + a_1 y^{(n-1)} + \dots + a_{n-1} \dot{y} + a_n y = f(y^{(n-1)}, \dots, \dot{y}, y, u, t) \quad (2.27)$$

dove $u(t)$ e $y(t)$ sono rispettivamente gli ingressi e le uscite del sistema. Ponendo:

$$f(y^{(n-1)}, \dots, \dot{y}, y, u, t) = V(t) \quad (2.28)$$

,dove $V(t)$ è il nuovo ingresso, sfruttando ora il teorema di Dini, in un intorno di $Y_0 = (y_0^{(n-1)}, \dots, \dot{y}, y, u, t)$, sotto l'ipotesi che $\frac{\partial f}{\partial u}|_{(y_0)} \neq 0$, è possibile esprimere $u(t)$ come:

$$u(t) = g(y^{(n-1)}, \dots, \dot{y}, y, v, t). \quad (2.29)$$

Quindi nel nuovo ingresso di controllo, l'originale sistema, che ricordiamo essere non lineare, sarà invece lineare rispetto a $V(t)$:

$$y^{(n)} + a_1 y^{(n-1)} + \dots + a_{n-1} \dot{y} + a_n y = V(t). \quad (2.30)$$

Questa tecnica è conosciuta come *direct feedback linearization*.

Nel caso in esame le uscite del sistema coincidono con gli stati, perciò il nuovo ingresso di controllo si pone :

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix} = \begin{bmatrix} \dot{y}_1 \\ \dot{y}_2 \\ \dot{y}_3 \\ \dot{y}_4 \end{bmatrix} = \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \end{bmatrix} \quad (2.31)$$

in cui il v_1 e v_2 rappresentano l'ingresso di controllo per la bocchetta delle zone, v_3 il flusso d'acqua che percorre la batteria di raffreddamento e v_4 la velocità del ventilatore.

Si può ottenere il sistema lineare e disaccoppiato:

$$\begin{bmatrix} \dot{y}_1 \\ \dot{y}_2 \\ \dot{y}_3 \\ \dot{y}_4 \end{bmatrix} = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix} \quad (2.32)$$

Invertendo le relazioni (2.16), (2.17), (2.26) e la relazione del controllo di pressione (2.14) si ottengono gli ingressi di controllo:

$$u_1 = \frac{v_1 - a_2 u_4 - a_3 (u_6 - x_1)}{a_1 (x_3 - x_1)} \quad (2.33)$$

$$u_2 = \frac{v_2 - b_2 u_5 - b_3 (u_6 - x_2)}{b_1 (x_3 - x_2)} \quad (2.34)$$

$$u_3 = \left[\frac{M_c C_{pc}}{(UA)_c} v_3 - (T_{ai} - x_3) \right] \frac{C_{pa} (u_1 + u_2)}{C_{pw} (T_{wi} - x_4)} \quad (2.35)$$

$$u_4 = (v_4 + 323.38v_1 + 357.27v_2 + 205.38)/11.95 \quad (2.36)$$

i quali perdono di significato nel caso in cui $x_3 = x_1$, $x_3 = x_2$) e $x_4 = T_{wi}$. Avendo disaccoppiato gli ingressi, si individuano quattro anelli di regolazione indipendenti che producono i v_i ; si possono di conseguenza progettare anche quattro controllori indipendenti. I primi due saranno responsabili del mantenimento delle temperature delle stanze a prefissati setpoint tramite la variazione della portata d'aria a esse; il terzo lavorerà per mantenere T_{sa} al valore desiderato intervenendo sulla portata d'acqua alla batteria di raffreddamento. Per

questi tre casi si è scelto un controllore proporzionale a guadagno unitario. Per il controllo del ventilatore si è optato per un controllo PI. In Figura 2.2 si riporta una schematizzazione di modello e controllo.

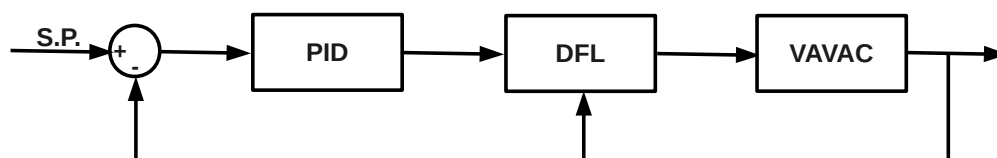


Figura 2.2: Schema controllo impianto VAVAC.

Capitolo 3

FDD mediante tecniche di machine learning

L'*apprendimento automatico*, dall'inglese *machine learning*, rappresenta una delle branche dell'intelligenza artificiale e si occupa di realizzare sistemi e algoritmi che si basano su osservazioni del comportamento di un dato sistema per sviluppare una nuova conoscenza propria. L'apprendimento può avvenire catturando caratteristiche provenienti da esempi, strutture dati o sensori, per analizzarle e valutarne le relazioni tra le variabili osservate.

Uno degli obiettivi principali di ricerca sull'apprendimento automatico è quello di imparare a riconoscere autonomamente modelli complessi e prendere decisioni intelligenti basate su dati osservati; la difficoltà principale risiede nel fatto che l'insieme di tutti i possibili comportamenti, dati tutti gli ingressi possibili, è troppo grande per essere coperto da insiemi di esempi osservati. Gli esempi che servono a far sì che l'algoritmo di apprendimento automatico acquisisca una propria coscienza prendono il nome di *data train* (dati di allenamento). Per l'impossibilità di fornire tutti gli esempi relativi a tutti i possibili comportamenti del sistema, è necessario l'utilizzo di tecniche che generalizzino gli esempi di *train* in modo che l'algoritmo riesca a interpretare anche casi estranei al set di addestramento. Al giorno d'oggi non si è in grado di riprodurre sistemi di apprendimento automatico simili a quello umano ma in varie applicazioni la ricerca ha prodotto notevoli risultati, come ad esempio per il riconoscimento vocale.

Gli algoritmi di machine learning si classificano in tre principali tipologie:

- **Apprendimento supervisionato:** in questo caso il data training è composto da una coppia di esempi determinata da un oggetto di input e un valore di output desiderato. Un algoritmo di apprendimento supervisionato genera una funzione di inferenza (classificatore) che dovrebbe essere in grado di predire il corretto valore di output per ogni input valido.
- **Apprendimento non supervisionato:** in cui il problema è trovare strutture nascoste in strutture di dati non pre-classificate da cui non è possibile valutare una possibile soluzione. L'apprendimento non supervisionato è strettamente legato al problema di densità in statistica.
- **Apprendimento con rinforzo:** questa tecnica si basa sul presupposto di potere ricevere degli stimoli dall'esterno a seconda delle scelte dell'algoritmo. Gli algoritmi per il reinforcement learning tentano di determinare una politica tesa a massimizzare gli incentivi cumulati ricevuti dall'agente nel corso della sua esplorazione del problema. L'apprendimento con rinforzo differisce da quello supervisionato perché non sono mai presentate coppie input-output di esempi di train, né si procede alla correzione esplicita di azioni sub-ottimali.

In relazione agli obiettivi di questa tesi, si è scelto di risolvere il problema della fault detection and identification utilizzando tecniche di machine learning. Il principale vantaggio di questo tipo di approccio è di non necessitare di un modello dell'impianto del quale è prevista la supervisione (rientra infatti nelle tecniche black-box). Nel caso in esame si sfruttano algoritmi di apprendimento supervisionato in cui la prima fase del processo è dedicata alla raccolta dei dati di train, sia in funzione di normale funzionamento che in presenza di guasti, la seconda fase stabilisce come scegliere questi dati. Vale la pena notare che, le prestazioni delle SVM sono influenzate dalla scelta dei campioni di allenamento ma la robustezza dell'algoritmo di decisione non impone rigide scelte sulle condizioni alle quali allenare il sistema.

L'algoritmo decisionale risultante opererà in linea durante il funzionamento nominale dell'impianto. Per quanto concerne le tecniche da utilizzare si è scelto di fare riferimento a quanto suggerito nella letteratura ([3],[?]). In particolare

si mette a confronto l'impiego della tecnica KNN (K-nearest neighbors) e SVM (Support vector machine).

3.1 K-Nearest Neighbors

Il k-nearest neighbors classifica un nuovo campione analizzando la distanza tra esso ed i dati di train a lui più prossimi. In altre parole, le regole di classificazione del K-NN si basano sulla similarità con i campioni del training set, un po' come il test dell'anatra: "if it walks like a duck, quacks like a duck, then it is probably a duck" [31].

Per classificare un dato senza etichetta x il classificatore valuta una distanza, tipicamente quella euclidea, e secondo tale metrica individua i k elementi dei dati di allenamento più vicini. Dato un punto e trovati i suoi vicini, vi sono due possibili scelte decisionali: classificarlo in base alla maggioranza di voti oppure alla somma pesata dei voti. Nel primo caso il dato viene assegnato alla classe che compare con maggior frequenza tra i vicini, nel secondo si assegna un peso ai vicini, ad esempio il reciproco del quadrato della distanza, e se ne valuta la somma ponderata delle distanze, in tal modo si attribuisce maggior importanza ai vicini più prossimi.

La scelta del valore k si effettua secondo logiche euristiche valutando le prestazioni a diversi k su dei dati di validazione. Nel caso di classificazione binaria il numero intero k è consigliabile sia dispari in modo da non aver problemi decisionali.

Il problema principale di questo semplice classificatore è il calcolo delle distanze che lo rende computazionalmente oneroso soprattutto all'aumentare del set dei dati di train. Inoltre, per ottenere buone prestazioni, è necessario un numero elevato di misure per l'allenamento. In Figura 3.1 si riporta un semplice esempio: si vuole classificare il pallino verde secondo la metodologia della maggioranza voti. Se si considera $k = 3$, ovvero i tre elementi più prossimi, allora il pallino verde viene inserito nella categoria dei triangolini verdi perché sono presenti due triangolini a fronte di un solo quadratino. Se però si sceglie $k = 5$ il pallino viene classificato come quadratino blu perché sono presenti tre quadratini e due triangolini.

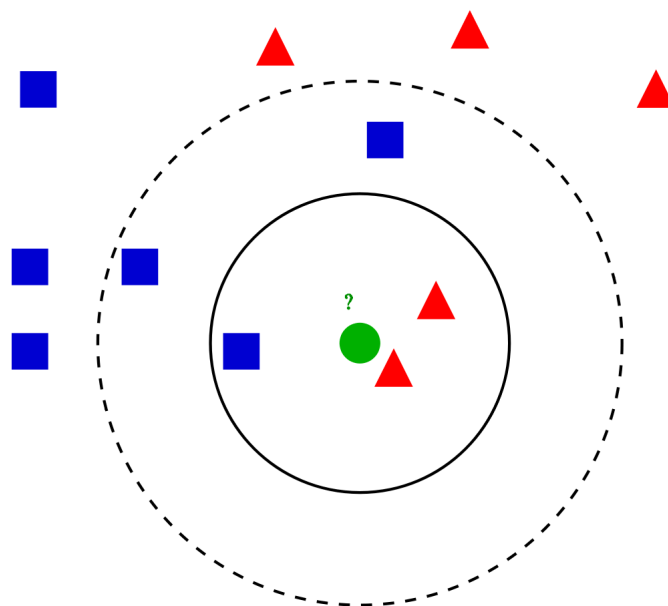


Figura 3.1: Esempio classificatore K-NN

3.2 Support vector machine

Le macchine a vettori di supporto (SVM, dall'inglese Support Vector Machines), sono un insieme di metodi di apprendimento supervisionato per la regressione e la classificazione di pattern, sviluppati negli anni '90 da Vladimir Vapnik. Questo metodo di classificazione permette di classificare tra due possibili gruppi. L'idea alla base delle SVM è di mappare i dati in ingresso in uno spazio caratteristico, *higher-dimensional feature space*, e quindi individuare un iperpiano che massimizzi i margini di classificazione, separando lo spazio caratteristico in due zone e valutando, nel caso di nuovi dati, la posizione rispetto a tale delimitazione.

Il set di dati di allenamento è composto da N coppie $(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)$, dove $x_i \in \mathbb{R}^p$ è il vettore dei dati di train mentre $y_i \in \{-1, 1\}$ rappresenta la classe associata al relativo dato di allenamento.

Un iperpiano si definisce mediante:

$$x : f(x) = x^T \beta + \beta_0 = 0, \quad \|\beta\| = 1 \quad (3.1)$$

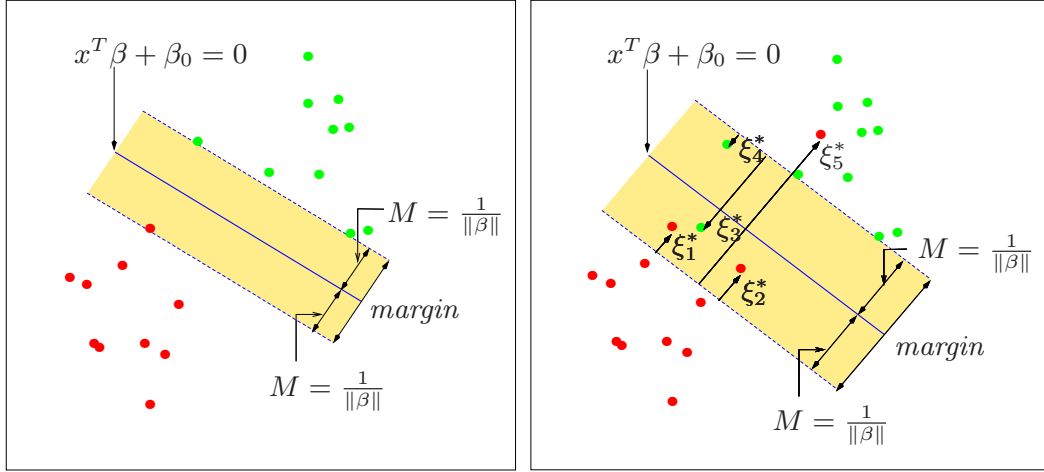


Figura 3.2: Support vector machine: a sinistra il caso di classi separabili, a destra il caso di classi non separabili

La funzione descritta dall'iperpiano genera una regola per classificare i nuovi dati osservati infatti $f(x)$ rappresenta la distanza con segno tra il punto x e l'iperpiano, per cui la classificazione si ottiene semplicemente come:

$$G(x) = \text{sign}[x^T \beta + \beta_0]. \quad (3.2)$$

Nel caso in cui si sia in presenza di classi separabili (Figura 3.2, a sinistra), è possibile individuare una funzione $f(x) = x^T \beta + \beta_0$ tale che $y_i f(x_i) > 0 \forall i$. Quindi è possibile individuare l'iperpiano che massimizza il margine tra i dati di train della classe 1 e quelli di -1. Il problema di ottimo si può esprimere come:

$$\begin{aligned} & \max_{\beta, \beta_0, \|\beta\|=1} M \\ & \text{vincolato a } y_i(x_i^T \beta + \beta_0) \geq M, \quad i = 1, \dots, N \end{aligned} \quad (3.3)$$

Avendo ipotizzato che le classi sono separabili, tutti i punti saranno distanti dall'iperpiano definito da β e β_0 almeno M . La banda illustrata in figura è M unità distante dall'iperpiano in entrambi i lati, cioè di ampiezza $2M$, ed è chiamato *margine*. È possibile rilassare l'ipotesi $\|\beta\|=1$ rimpiazzando la

precedente condizione di vincolo con:

$$\frac{1}{\|\beta\|} y_i (x_i^T \beta + \beta_0) \geq M, \quad (3.4)$$

e quindi equivalentemente:

$$y_i (x_i^T \beta + \beta_0) \geq M \|\beta\|. \quad (3.5)$$

Dato che per ogni β e β_0 che soddisfano la (3.5) anche una loro riscalatura ad opera di moltiplicatori positivi ne manterrà la validità, è possibile porre $\|\beta\| = 1/M$. Quindi la 3.3 è esprimibile come il problema di minimo:

$$\begin{aligned} & \min_{\beta, \beta_0} \frac{1}{2} \|\beta\|^2 \\ & \text{vincolato a } y_i (x_i^T \beta + \beta_0) \geq M, \quad i = 1, \dots, N. \end{aligned} \quad (3.6)$$

Per costruzione si ha una regione vuota attorno al confine di decisione di spessore $1/\|\beta\|$, si scelgono dunque β e β_0 affinché sia massimizzato tale margine. Questo è un problema di programmazione convessa ed è risolvibile sfruttando il metodo di Lagrange, la cui funzione lagrangiana associata risulta:

$$L_P = \frac{1}{2} \|\beta\|^2 - \sum_{i=1}^N \alpha_i [y_i (x_i^T \beta + \beta_0) - 1]. \quad (3.7)$$

Annullando le derivate si ottiene:

$$\beta = \sum_{i=1}^N \alpha_i y_i x_i \quad (3.8)$$

$$0 = \sum_{i=1}^N \alpha_i y_i \quad (3.9)$$

che sostituite in (3.7) portano al cosiddetto duale di Wolfe:

$$\begin{aligned} L_D = \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{k=1}^N \alpha_i \alpha_k y_i y_k x_i^T x_k \\ \text{con } \alpha_i \geq 0. \end{aligned} \quad (3.10)$$

La soluzione si ottiene massimizzando L_D , ovvero risolvendo il semplice pro-

blema di ottimizzazione convessa; soluzione che si può ottenere mediante un qualunque software di ottimizzazione. Il risultato ottenuto deve soddisfare la condizione di Karush-Kuhn-Tucker (KKT), che comprende (3.8), (3.9), $\alpha_i > 0$ e

$$\alpha_i [y_i(x_i^T \beta + \beta_0) - 1] = 0 \forall i. \quad (3.11)$$

Da quest'ultima equazione è utile osservare che:

- se $\alpha_i > 0$, allora $y_i(x_i^T \beta + \beta_0) = 1$ e dunque x_i appartiene al confine della zona di separazione ed è denominato *support point*;
- se $y_i(x_i^T \beta + \beta_0) > 1$, allora x_i non giace sul confine e $\alpha_i = 0$.

Dalla (3.8) si vede come il vettore delle soluzioni β si ottenga dalla combinazione lineare dei support point x_i . Allo stesso modo β_0 si ottiene risolvendo (3.11) per ognuno dei punti di supporto.

Il caso più interessante, in quanto più spesso presente nei problemi reali, è quando le classi non siano separabili ma si sovrappongono nello spazio caratteristico (Figura 3.2, a sinistra).

Un modo per affrontare questo problema è di continuare a massimizzare M , ma permettendo ad alcuni campioni di essere nella parte sbagliata del margine. Definite delle variabili di approssimazione (o slack) $\xi = (\xi_1, \xi_2, \dots, \xi_N)$, vi sono due strade per modificare i vincoli dell'equazione (3.3) che descriveva il problema di ottimo da risolvere nel caso precedente:

$$y_i(x_i^T \beta + \beta_0) \geq M - \xi_i, \quad (3.12)$$

oppure

$$y_i(x_i^T \beta + \beta_0) \geq M(1 - \xi_i), \quad (3.13)$$

$\forall i, \xi_i \geq 0, \sum_{i=1}^N \xi_i \leq \text{costante}$. Le due scelte portano a risultati differenti. La prima sembra più naturale in quanto misura la sovrapposizione nell'attuale distanza dal margine; la seconda scelta misura la sovrapposizione in distanza relativa, la quale cambia con la larghezza M del margine. Nonostante questo la prima opzione risulta sconsigliata perché porta ad un problema di ottimizzazione non convesso mentre nel secondo caso è convesso; per questo motivo le support vector machine applicano la condizione (3.13). Il parametro ξ_i nel

vincolo $y_i(x_i^T \beta + \beta_0) \geq M(1 - \xi_i)$ fornisce una stima di quanto la classificazione $f(x_i) = x_i^T \beta + \beta_0$ sia nella parte sbagliata del margine. Quindi, maggiorando la somma $\sum \xi_i$, si limita in maniera proporzionale il valore delle predizioni sbagliate. Gli errori di classificazione si hanno per $\xi_i > 1$, perciò limitare $\sum \xi_i$ ad un valore K equivale a stabilire il numero massimo di campioni di train che apparterranno alla parte sbagliata dell'iperpiano.

Come nel caso delle classi separabili siamo in grado di eliminare il vincolo su β definendo $M = 1 / \|\beta\|$ e di esprimere (3.6) come:

$$\min \|\beta\| \quad \text{con} \quad \begin{cases} y_i(x_i^T \beta + \beta_0) \geq 1 - \xi_i \forall i, \\ \xi_i \geq 0, \sum \xi_i < \text{costante} \end{cases} \quad (3.14)$$

Analizzando quest'ultimo risultato si osserva come i punti ben all'interno del confine della loro classe non giocano un ruolo importante nella determinazione dei margini. Anche in questo caso si è in presenza di un problema di ottimizzazione convesso risolubile mediante i moltiplicatori di Lagrange. A tal fine è conveniente riscrivere (3.14) nell'equivalente forma quadratica

$$\begin{aligned} & \min_{\beta, \beta_0} \frac{1}{2} \|\beta_0\|^2 + C \sum_{i=1}^N \xi_i \\ & \text{vincolato a } \xi_i \geq 0, y_i(x_i^T \beta + \beta_0) \geq 1 - \xi_i \forall i, \end{aligned} \quad (3.15)$$

dove il parametro C rappresenta la costante introdotta precedentemente ($\sum \xi_i < C$); il caso di classi separabili corrisponde a $C = \infty$.

La lagrangiana associata al problema di ottimizzazione è:

$$L_P = \frac{1}{2} \|\beta\|^2 + C \sum_{i=1}^N \xi_i - \sum_{i=1}^N \alpha_i [y_i(x_i^T \beta + \beta_0) - (1 - \xi_i)] - \sum_{i=1}^N \mu_i \xi_i \quad (3.16)$$

annullando le derivate:

$$\beta = \sum_{i=1}^N \alpha_i y_i x_i, \quad (3.17)$$

$$0 = \sum_{i=1}^N \alpha_i y_i, \quad (3.18)$$

$$\alpha_i = C - \mu_i, \quad \forall i, \quad (3.19)$$

in cui le costanti $\alpha_i, \mu_i, \xi_i \geq 0 \forall i$.

Sostituendo (3.17)-(3.19) in (3.16) si perviene a

$$L_D = \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{i'=1}^N \alpha_i \alpha_{i'} y_i y_{i'} x_i^T x_{i'} \quad (3.20)$$

Si vuole dunque massimizzare L_D con $0 \leq \alpha_i \leq C$ e $\sum_{i=1}^N \alpha_i y_i = 0$. Oltre alle (3.17)-(3.19), le condizioni di Karush-Kuhn-Tucker impongono:

$$\alpha_i [y_i (x_i^T \beta + \beta_0) - (1 - \xi_i)] = 0, \quad (3.21)$$

$$\mu_i \xi_i = 0, \quad (3.22)$$

$$y_i (x_i^T \beta + \beta_0) - (1 - \xi_i) \geq 0, \quad (3.23)$$

per $i = 1, \dots, N$. Le equazioni (3.17)-(3.23) caratterizzano univocamente la soluzione del problema di minimizzazione sia nel caso primale che nel duale. Dall'equazione (3.17) si osserva che β ha la forma:

$$\hat{\beta} = \sum_{i=1}^N \hat{\alpha}_i y_i x_i, \quad (3.24)$$

con i coefficienti α_i diversi da zero solo per i dati che soddisfano esattamente il vincolo in (3.23), questi valori sono chiamati *support vector*, vettori di supporto, poiché β è rappresentata solo in loro termini. Ottenute le soluzioni ottime $\hat{\beta}$ e $\hat{\beta}_0$ la funzione decisionale per un nuovo dato x è

$$\hat{G}(x) = \text{sign}[\hat{f}(x)] = \text{sign}[x^T \hat{\beta} + \hat{\beta}_0]. \quad (3.25)$$

La support vector machine illustrata fino ad ora permette di individuare dei confini lineari nel feature space. Come per altri metodi lineari è possibile rendere più flessibile la procedura modificando lo spazio caratteristico mediante funzioni che espandono lo spazio, come ad esempio funzioni polinomiali o esponenziali. In generale i confini nel nuovo spazio permettono una migliore separazione dei dati di addestramento e corrispondono a limiti non lineari nello

spazio originario. Le macchine a vettori di supporto generalmente permettono questa mappatura mediante funzioni kernel e i casi illustrati corrispondono alla scelta di un kernel lineare.

Una volta scelte le funzioni $h_m(x)$, $m = 1, \dots, M$ per mappare i dati in ingresso nel nuovo spazio caratteristico, la procedura di sintesi è la medesima: si individuano i support vector nel feature space $h(x_i) = (h_1(x_i), h_2(x_i), \dots, h_M(x_i))$, $i = 1, \dots, N$, e si ottiene la funzione di decisione non lineare $\hat{f}(x) = h(x)^T \hat{\beta} + \hat{\beta}_0$. Di conseguenza la classificazione sarà semplicemente $\hat{G} = \text{sign}(\hat{f}(x))$.

La funzione di Lagrange duale (3.20) si può esprimere come:

$$L_D = \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{i'=1}^N \alpha_i \alpha_{i'} y_i y_{i'} \langle h(x_i), h(x_{i'}) \rangle \quad (3.26)$$

dove $\langle \cdot, \cdot \rangle$ rappresenta l'operazione di prodotto interno. Dalla (3.17) si vede che la soluzione $f(x)$ si può scrivere:

$$f(x) = h(x)^T \beta + \beta_0 = \sum_{i=1}^N \alpha_i y_i \langle h(x), h(x_i) \rangle + \beta_0. \quad (3.27)$$

Come analizzato nel caso di classi separabili e non, α_i , β_0 si ottengono risolvendo $y_i f(x_i) = 1$ in (3.27) per ogni x_i con $0 < \alpha_i < C$. In entrambe le relazioni (3.26) e (3.27) ci si riferisce a $h(x)$ solo mediante il prodotto scalare per cui non è necessario esplicitare $h(x)$ ma è sufficiente conoscere la *funzione kernel* così definita:

$$K(x, x') = \langle h(x), h(x') \rangle \quad (3.28)$$

la quale valuta il prodotto interno nel nuovo spazio generato da $h(x)$. Si osserva facilmente che K è una funzione simmetrica e semidefinita positiva.

I kernel comuni utilizzati nell'ambito delle SVM sono:

- *dth-Degree polynomial*: $K(x, x') = (1 - \langle x, x' \rangle)^d$,
- *Radial basis*: $K(x, x') = \exp(-\gamma \|x - x'\|^2)$,
- *Neural network*: $K(x, x') = \tanh(k_1 \langle x, x' \rangle + k_2)$.

Se per esempio si considera un feature space con due ingressi X_1 e X_2 , e un kernel polinomiale di grado 2, questo risulta semplicemente esplicitato come:

$$\begin{aligned}K(X, X') &= (1 + \langle X, X' \rangle)^2 \\ &= (1 + X_1 X'_1 + X_2 X'_2)^2 \\ &= 1 + 2X_1 X'_1 + 2X_2 X'_2 + (X_1 X'_1)^2 + (X_2 X'_2)^2 + 2X_1 X'_1 X_2 X'_2.\end{aligned}$$

Dalla (3.27) si osserva che la soluzione si esprime, in termini della funzione kernel, come:

$$\hat{f}(x) = \sum_{i=1}^N \hat{\alpha}_i y_i K(x, x_i) + \hat{\beta}_0 \quad (3.29)$$

Capitolo 4

Estensioni delle support vector machine

Le support vector machine sono uno strumento molto versatile che trova innumerevoli applicazioni. La loro sintesi è tuttavia non banale e le prestazioni delle SVM dipendono fortemente dalle scelte che si operano in fase di definizione e dalla scelta dei parametri. In questo capitolo si fornisce una breve descrizione di alcune tecniche che possono giovare al processo di sintesi delle macchine a vettori di supporto. In primo luogo risulta cruciale la scelta delle variabili da considerare per istruire una data SVM, in tal senso si è deciso di utilizzare il *kernel principal component analysis* (KPCA).

Una volta scelto il kernel delle support vector nasce il problema di tarare opportunamente sia i parametri di tale kernel che la costante C che penalizza i margini delle SVM; si può pensare di determinare questi parametri in maniera del tutto automatica attraverso la soluzione di un problema di ottimo in cui la funzione obiettivo restituisce un indice legato alle prestazioni della SVM.

E' altresì importante notare che le SVM forniscono semplicemente la classe di appartenenza di una osservazione valutando il segno della funzione $f(x)$, (eq. (3.25)); tuttavia in molti casi si è più interessati ad un livello di confidenza che alla mera classificazione; a quest'ultima può quindi risultare utile associare un qualche tipo di informazione probabilistica.

4.1 Kernel Principal Component Analysis

La *kernel principal component analysis* (KPCA) è l'estensione non lineare, mediante una pre-mappatura dei dati ad opera di una funzione kernel, dell'analisi delle componenti principali (PCA). La PCA è una tecnica per la semplificazione dei dati utilizzata nell'ambito della statistica multivariata. Fu ideata nel 1901 da Karl Pearson e sviluppata da Harold Hotelling nel 1933. Lo scopo primario di questa tecnica è la riduzione di un numero più o meno elevato di variabili (rappresentanti altrettante caratteristiche del fenomeno analizzato) in alcune variabili latenti (feature reduction).

In questo modo si ottiene una riduzione della grande mole di dati riducendo la complessità computazione della sintesi della support vector machine. Il PCA cerca di minimizzare la perdita di informazione dovuta a tale semplificazione. La selezione delle variabili avviene considerando quelle a massima covarianza, per questo il problema del PCA si può esprimere come un problema agli autovalori dominanti applicato alla matrice di covarianza.

Dato un set di dati di ingresso $x_1, x_2, \dots, x_m \in \mathbb{R}^n$ (generalmente $n < m$) e tali che $\sum_{t=1}^m x_t = 0$, la cui matrice di covarianza associata è:

$$C = \frac{1}{m} \sum_{t=1}^m x_t x_t^T \quad (4.1)$$

Le componenti principali (PC) si ottengono risolvendo il problema agli autovalori relativo alla matrice C ,

$$\lambda_t u_t = C u_t \quad (4.2)$$

dove λ_t sono gli autovalori di C che andranno ordinati in ordine decrescente, mentre u_t sono i corrispondenti autovettori. Scegliere i primi k autovalori coincide con l'approssimare i dati di origine con le k componenti che più descrivono il dataset. Per la scelta del k è possibile introdurre una soglia θ che denota la precisione dell'approssimazione corrispondente a considerare i primi k autovettori:

$$\frac{\sum_{t=1}^k \lambda_t}{\sum_{t=1}^m \lambda_t} \geq \theta. \quad (4.3)$$

Data la precisione desiderata θ , il parametro k risulta deciso. Ponendo $U = [u_1, u_2, \dots, u_k]$, $\Lambda = [\lambda_1, \lambda_2, \dots, \lambda_k]$, il nuovo feature space, di dimensioni minori al precedente, si ottiene come:

$$S = U^T x_t. \quad (4.4)$$

Le nuove componenti sono dette principal component.

Il PCA funziona bene per problemi lineari mentre per il caso non lineare la questione è più delicata.

Il kernel principal component analysis può efficacemente estrarre la relazione non lineare dagli ingressi originali; questo mappando il dato di ingresso x in un nuovo feature space F utilizzando una mappa non lineare φ , e successivamente applica il PCA in F . Tuttavia, è difficile ottenere questo direttamente perché la dimensione h dello spazio caratteristico F può essere arbitrariamente elevata o persino infinita. Infatti, nelle usuali implementazioni il vettore nello spazio caratteristico F non viene calcolato esplicitamente ma si considera il prodotto scalare tra due vettori in F sfruttando una funzione kernel (come avviene nelle kernel SVM).

I campioni della matrice di covarianza \hat{C} di $\varphi(x_i)$ si ottengono come:

$$\hat{C} = \frac{1}{m} \sum_{i=1}^m \varphi(x_i) \varphi(x_i)^T. \quad (4.5)$$

Ogni autovettore u_i appartiene allo spazio generato da $\varphi(x_1), \dots, \varphi(x_m)$, perciò, u_i si può esprimere come combinazione lineare:

$$u_i = \sum_{i=1}^m \alpha_i \varphi(x_i). \quad (4.6)$$

Moltiplicando ambo i membri dell'equazione (4.2) per $\varphi(x_i)$, si perviene a:

$$\varphi(x_i) C u = \lambda (\varphi(x_i) \cdot u). \quad (4.7)$$

Sostituendo la (4.5) e la (4.6) in (4.7) si ottiene l'equivalente problema agli autovalori:

$$K \alpha_i = \lambda \alpha_i, \quad i = 1, \dots, m \quad (4.8)$$

dove K è la matrice kernel $m \times m$. Il valore di ogni elemento di K è pari al prodotto interno tra due vettori x_i e x_j nel spazio caratteristico cioè $\varphi(x_i)$ e $\varphi(x_j)$; perciò

$$K(x_i, x_j) = \varphi(x_i) \cdot \varphi(x_j). \quad (4.9)$$

L'introduzione della funzione kernel, come visto per le SVM, permette di non dover conoscere né la forma di $\varphi(x)$ né di dover calcolare il prodotto scalare nel feature space. Ogni funzione che soddisfa la condizione di Mercer¹ può essere utilizzata come K . Se $\tilde{\lambda}_i$ è un autovalore di K esso soddisfa $\tilde{\lambda}_i = m\lambda_i$. Inoltre per ortonormalizzare gli autovettori di $\varphi(x_i)$, $v_i \cdot v_i = 1$, ogni α_i è normalizzato grazie al corrispondente autovalore con $\tilde{\alpha}_i = \alpha_i/\tilde{\lambda}_i$. Si può dunque concludere che la componente principali per un vettore x_i in ingresso si calcola come:

$$s(i) = u_i^T \varphi(x_i) = \sum_{j=1}^n \tilde{\alpha}_i(j) \cdot K(x_i, x_j), \quad i, j = 1, 2, \dots, m. \quad (4.10)$$

In realtà i dati non soddisfano sempre la condizione $\sum_{k=1}^n \varphi(x_k) = 0$, e quindi K nell'equazione (4.8) viene rimpiazzato da

$$\bar{K} = K - 1_n \cdot K - K \cdot 1_n + 1_n K 1_n \quad (4.11)$$

dove la matrice $1_n = (1/n)_{n \times n}$.

Oltre alla riduzione delle dimensioni dei dati di train, basata sulla selezione delle variabili che portano maggiori informazioni, il KPCA permette di ottenere una migliore separazione dei dati appartenenti alle due classi, avvantaggiando la successiva sintesi delle SVM. Tale divisione è opera della funzione kernel scelta e di come tali parametri vengano tarati. In Figura 4.1 si illustra un esempio: in alto sono riportati i dati originali, in cui sono evidenziate le due classi di appartenenza con simboli e colori differenti, mentre in basso la mappatura mediante il KPCA. In figura si è utilizzato un kernel rbf. Come si può osservare, nel secondo caso, la divisione fra le classi è visibilmente più semplice

¹Sia $K(x, x')$ un kernel continuo e simmetrico definito nell'intervallo chiuso $a \leq x, x' \leq b$. Il kernel $K(x, x')$ può essere espanso nella serie $K(x, x') = \sum_{i=1}^{\infty} \lambda_i \varphi_i(x) \varphi_i(x')$ con $\lambda_i \geq 0$. Affinché tale espansione sia valida e per la sua convergenza assoluta ed uniforme, è necessario e sufficiente che la condizione $\int_b^a \int_b^a K(x, x') \psi(x) \psi(x') dx dx' \geq 0$ sia vera per tutte le $\psi(\cdot)$ che soddisfano $\int_b^a \psi^2(x) dx < \infty$.

da effettuare. Nel seguito dell'elaborato si considera il radial basis kernel e per ottenere la migliore separazione possibile è essenziale scegliere opportunamente il parametro γ . A tal fine si introduce un indice di costo opportuno e, come proposto in [32], si vuole massimizzare:

$$F_j = \frac{\text{tr}(S_w)}{\text{tr}(S_b)} \quad (4.12)$$

in cui:

$$S_w = \frac{1}{m} \sum_{i=1}^2 m_i (\bar{y}_i - \bar{y})(\bar{y}_i - \bar{y})^T. \quad (4.13)$$

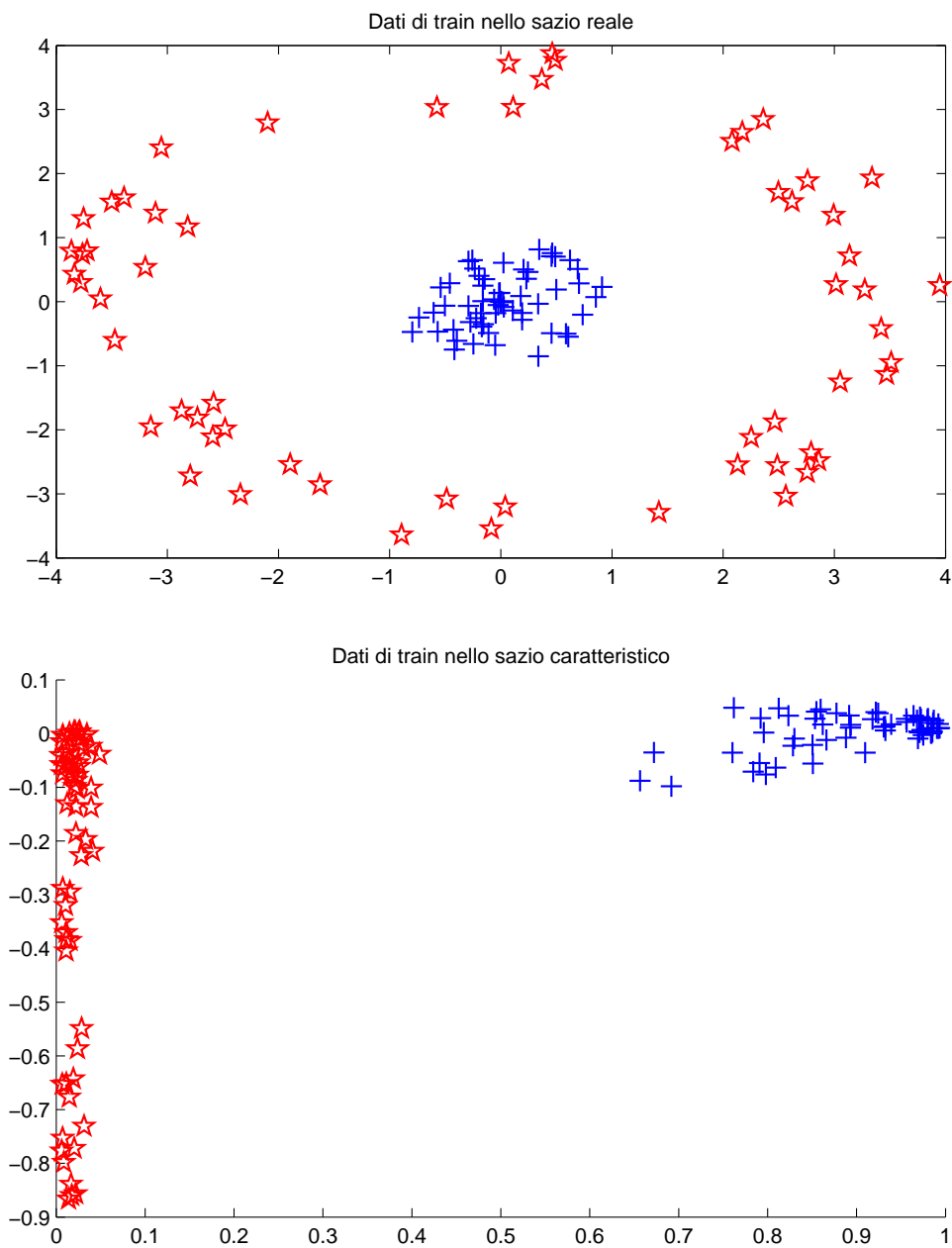
$$S_b = \frac{1}{m} \sum_{i=1}^2 \sum_{j=1}^{m_i} (y_j^i - \bar{y}_i)(y_j^i - \bar{y}_i)^T. \quad (4.14)$$

dove m_1 è il numero degli elementi appartenenti alla prima classe mentre m_2 alla seconda; il numero totale dei campioni è dunque $m = m_1 + m_2$. Per mantenere più compatta e leggibile la scrittura si è posto $\{y_i\}_{i=1}^M$ pari all'immagine dei dati nel feature space. Con \bar{y} , \bar{y}_1 e \bar{y}_2 sono indicati rispettivamente il centro di tutti i dati di train, il centro di quelli appartenenti alla classe 1 e il centro della classe 2. Infine, y_j^i è il j-esimo dato della i-esima classe ($i = 1, 2$).

Il valore dell'indice varia dunque in funzione della scelta del parametro γ del kernel rbf. Massimizzare tale indice porta ad una miglior separazione dei dati, per risolvere tale problema di ottimo si è utilizzato il metodo della sezione aurea (Appendice A).

4.2 Particle swarm optimization

Al fine effettuare il tuning della SVM e determinare i valori più opportuni per i parametri più sinigicativi si risolve unproblema di ottimo. In questa tesi si impiega una tecnica euristica, basata sul *particle swarm optimisation* (PSO). L'ottimizzazione a sciami di particelle appartiene a quella classe di algoritmi che si basano sullo studio del comportamento collettivo di gruppi di animali ed in particolare proprio gli stormi di uccelli hanno ispirato tale ricerca. Le regole delle iterazione tra i componenti della popolazione si basano su metriche classiche quali distanze, momenti e differenti velocità. L'idea di base è legata

**Figura 4.1:** Esempio KPCA

al concetto che ogni individuo (una possibile soluzione nell'insieme di ricerca) di una popolazione sciame vada alla ricerca di una soluzione ottima con una propria dinamica legata alla propria posizione migliore individuata (dal punto di vista della valutazione della funzione fitness), e della posizione migliore trovata dello sciame. Oltre ad un comportamento individuale, quello che emerge è un comportamento collettivo che ben si presta alla risoluzione di diversi problemi di ottimizzazione. Reynolds [33] propose un modello comportamentale nel quale ogni agente segue tre semplici regole:

1. Separazione: ogni agente tenta di muoversi lontano dai propri vicini se questi sono troppo prossimi,
2. Allineamento: ogni agente si orienta nella direzione mediata dei propri vicini,
3. Coesione: ogni agente cerca di muoversi verso il baricentro della configurazione dei propri vicini.

Dal punto di vista algoritmico, queste idee sono riprese nel modello che descrive la dinamica dell'individuo nella versione di metodo PSO più semplice e primitiva. Data una popolazione di particelle in uno spazio n -dimensionale, alla ricerca di una (sub)-ottima soluzione di un problema di ottimizzazione di tipo:

$$\begin{aligned} \min f(\mathbf{x}) \\ \mathbf{x} \in X \end{aligned} \tag{4.15}$$

ogni individuo $x(t) = [x_1(t), x_2(t), \dots, x_n(t)]^T$ iterativamente si muove nello spazio di ricerca come:

$$\begin{aligned} \mathbf{x}(\mathbf{t} + \mathbf{1}) &= \mathbf{x}(\mathbf{t}) + \mathbf{v}(\mathbf{t}) \\ \mathbf{v}(\mathbf{t} + \mathbf{1}) &= \mu\mathbf{v}(\mathbf{t}) + \rho_p(\mathbf{x}_{\mathbf{p},\mathbf{best}} - \mathbf{x}(\mathbf{t})) + \rho_g(\mathbf{x}_{\mathbf{g},\mathbf{best}} - \mathbf{x}(\mathbf{t})) \end{aligned} \tag{4.16}$$

Il vettore $v(t) = [v_1(t), v_2(t), \dots, v_n(t)]^T$ rappresenta la velocità e viene anch'esso aggiornato linearmente in modo iterativo, influenzato da tre distinti fattori: la velocità al passo precedente (inerzia), migliore posizioni personale individuata nel passato $\mathbf{x}_{\mathbf{p},\mathbf{best}}$ e dalla miglior posizione globale $\mathbf{x}_{\mathbf{g},\mathbf{best}}$. Quest'ultime vengono aggiornate non appena se ne trovano di migliori, in termini

di valutazione della fitness del problema. μ è il coefficiente che pesa l'inerzia del sistema, mentre ρ_p e ρ_g sono coefficienti aleatori di distribuzione uniforme $U[0, 1]$ delle due componenti personali e globali rispettivamente. Si analizzano di seguito in modo più specifico le componenti di aggiornamento della velocità:

- Velocità precedente $v(t)$: agisce come memoria della precedente direzione di movimento (nell'immediato passato). Questo termine può essere visto come un momento che previene il cambio repentino di direzione della particella; per questo motivo questo termine è spesso detto *termine inerziale*.
- Componente cognitiva $\rho_p(x_{p,best} - x(t))$: essa quantifica la dinamica della particella relativamente alle performance del passato (non immediato). In un certo senso rappresenta la memoria individuale della particella, e modella la tendenza degli individui di tornare indietro verso le proprie posizioni migliori, è anche detto *nostalgia della particella*.
- Componente sociale $\rho_g(x_{g,best} - x(t))$: essa quantifica la dinamica della particella relativamente alle performance delle particelle vicine (o di tutte le particelle dello sciame in generale); ogni particella tende a muoversi anche in direzione della miglior posizione dello sciame. In letteratura sono presenti diverse varianti, sempre ispirate ai comportamenti sociali osservati in natura. Dotando le particelle di un raggio di osservazione, per esempio, si prediligerà il formarsi di sotto gruppi dello sciame originale che potrebbe dar vita ad un virtuoso meccanismo di competizione.
- Coefficiente inerzia μ : viene introdotto come meccanismo di controllo del processo di esplorazione globale e locale dello spazio. In questo senso tale parametro regola il tradeoff della ricerca da globale a locale: alti valori di inerzia alimentano la ricerca globale, valori sempre più bassi facilitano quella locale, fino ad eliminare del tutto la capacità di ricerca dell'individuo. Per $\mu = 1$ la velocità aumenta ad ogni iterazione portando all'instabilità del sistema (lo sciame diverge); per $\mu < 1$ le particelle decelerano finché la velocità non si annulla del tutto. Il problema del tuning del coefficiente di inerzia è molto dibattuto in letteratura, e attualmente si preferisce sostituire alla versione statica, una versione dinamica di

inerzia $\mu(t)$ caratterizzata da valori iniziali elevati (≈ 0.9) e da una dinamica asintoticamente stabile e tendente a valori bassi (≈ 0.4). In questo modo si favorisce una ricerca globale iniziale, e un'esplorazione locale più accurata nelle iterazioni finali.

Il numero di particelle e di iterazioni dell'algoritmo si impostano all'inizio dell'esecuzione. Inizialmente le particelle sono disposte casualmente nello spazio di ricerca. Nel caso di vincoli sul dominio, nel caso una particella tenti di posizionarsi in una zona proibita, esse viene fatta rientrare in zona utile per la ricerca.

Nella progettazione delle support vector machines è possibile sfruttare il PSO per scegliere nel migliore dei modi i parametri di kernel e di penalità sui margini ovvero, nel caso di rbf-kernel, σ e C . Le prestazioni del classificatore dipendono fortemente dalla scelta di tali parametri, e più precisamente dalla loro combinazione e questo rende difficile il problema di sintesi. Vi sono delle linee guida sugli effetti dovuti a questi parametri presi singolarmente ma è difficile, soprattutto in casi di un numero elevato di variabili, pervenire a buone prestazioni senza ricorrere ad algoritmo di ottimizzazione, proprio perché non è semplice determinare la correlazione dell'effetto delle variazioni sui due parametri.

Per definire un indice di costo appropriato si dividono i dati in train in due parti: con la prima si istruisce la support vector machine mentre con la seconda (dati di validazione) se ne valutano le prestazioni. La funzione di fitness, che si vuole minimizzare, è il numero di errori in fase di classificazione dei dati di validazione ovvero:

$$\min T = \frac{1}{l} \sum_{i=1}^l \Phi(-y_i f(x_i)). \quad (4.17)$$

Una tecnica per selezionare i dati di train e di validazione prende il nome di *k cross validation* in cui i campioni si dividono in k gruppi, e successivamente, se ne considero $k - 1$ per l'allenamento e il rimanente per la validazione. Nel seguente capitolo si entra nel dettaglio di come scegliere tali dati, in questa fase ci si limita a riportare in Figure 4.2 l'iperpiano che si ottiene con una scelta non accurata dei parametri, in alto, e con i parametri ottenuti mediante PSO. Si osserva subito come con dati complessi, anche nel più semplice caso

di considerare solo due variabili, come sia essenziale ricorrere ad un'opportuna taratura della support vector machine.

4.3 Probabilità a posteriori

L'idea alla base del voler ricavare una maggiore informazione dalle SVM, oltre alla semplice classe di appartenenza, nasce dalla seguente considerazione: nel caso vi siano delle osservazioni che vengano classificate erroneamente è ipotizzabile che esse siano mappate nelle prossimità della soglia di decisione ovvero dell'iperpiano; allontanandosi da tale margine si ha una maggiore probabilità che la classe di appartenenza stimata sia esatta.

La necessità di definire una misura probabilistica a posteriori della bontà della stima della SVM emerge soprattutto quando si devono definire delle SVM multi-classe. In letteratura, nella rilevazione delle anomalie per impianti HVAC, si utilizza l'approccio *one again all* in cui per ogni fault è definita una SVM che discrimina un solo guasto mentre gli altri sono, per essa, classificati come funzionanti. È frequente che per malfunzionamenti, che occorrono sulla stessa categoria di componenti, si abbiano segnalazioni di guasto in più parti del sistema nonostante si ipotizzi che si possa avere un unico guasto alla volta. Un esempio può riguardare i fault sulle bocchette di zona VAV in cui l'impianto di diagnosi potrebbe, nel caso di molte stanze, segnalare un'anomalia in più zone. Grazie alla probabilità a posteriori è possibile capire in quale ambiente risieda effettivamente il guasto.

Nel 1999, John Platt, propose di approssimare tale probabilità con una funzione a sigmoide da parametrizzare mediante i dati di train utilizzati per costruire la SVM [34].

L'informazione sulla distanza con segno di un campione x dall'iperpiano individuato dai support vector, come visto in precedenza, è data da:

$$f(x) = h(x) + b \quad (4.18)$$

dove

$$h(x) = \sum_{i \in SV} y_i \alpha_i k(x_i, x). \quad (4.19)$$

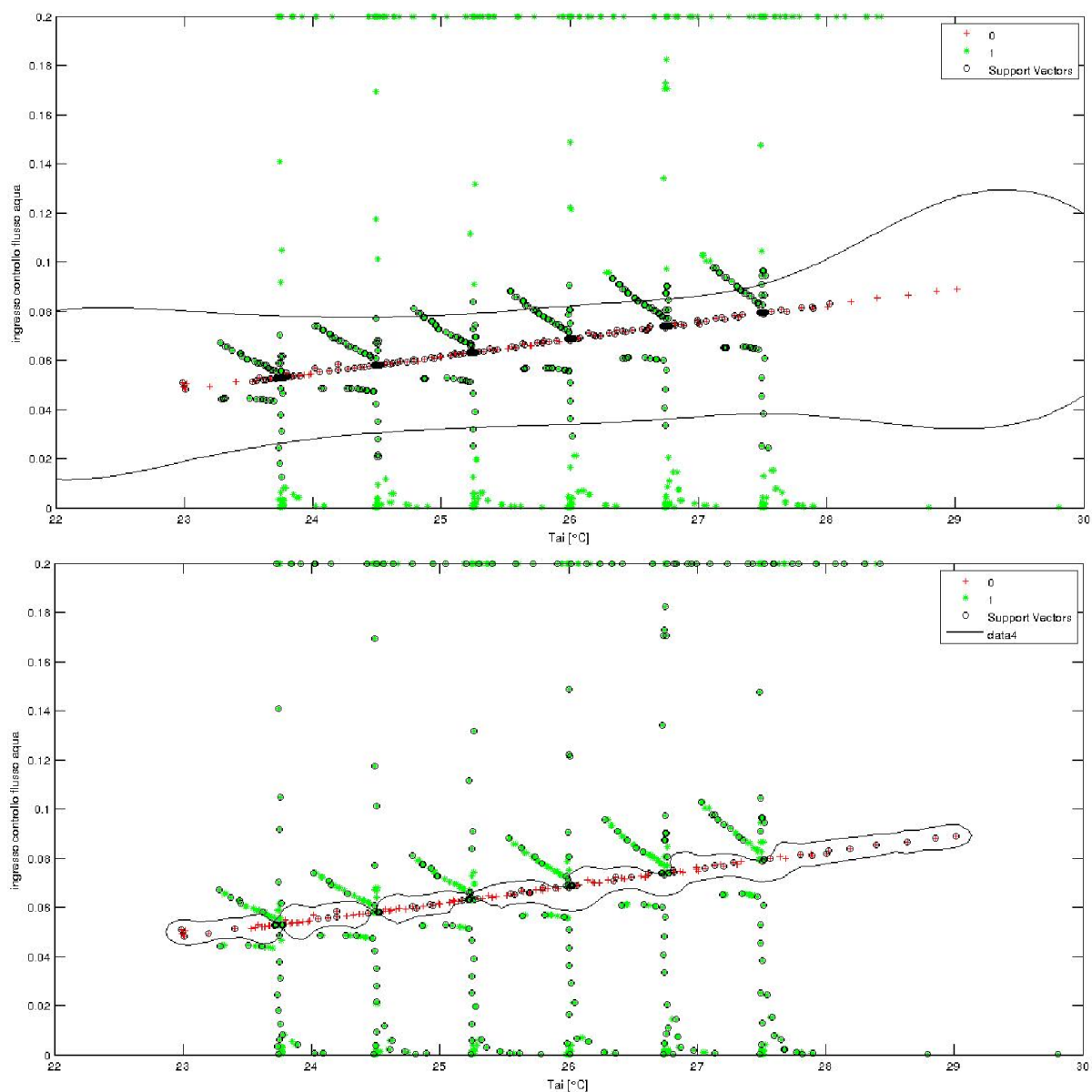


Figura 4.2: Esempio PSO-SVM in cui si vuole classificare un guasto sulla valvola di ricircolo a partire dalla temperatura dell'acqua in uscita dalla batteria di raffreddamento e dalla temperatura dell'aria in ingresso. Nella figura in alto non è stato applicato il PSO e si sono considerati valori standard ovvero $C = 1$ e $\sigma = 1$ mentre in basso è riportato l'iperpiano con l'uso del PSO.



Figura 4.3: Specie di fiori studiate nel Fisher's Iris data set

La funzione probabilistica a sigmoide che si desidera costruire ha forma:

$$P(y = 1|f) \approx P_{A,B}(f) = \frac{1}{1 + \exp(Af + B)}, \text{ dove } f = f(x). \quad (4.20)$$

In questo modello vi sono due parametri da tarare: A e B . I parametri ottimi $z^* = (A^*, B^*)$ si ottengono risolvendo il seguente problema di massima verosimiglianza:

$$\min_{z=(A,B)} F(x) = - \sum_{i=1}^l ((t_i \log(p_i) + (1 - t_i) \log(1 - p_i))), \quad (4.21)$$

dove

$$p_i = P_{A,B}(f_i), \text{ e } t_i = \begin{cases} \frac{N_+ + 1}{N_+ + 2} \text{ se } y_i = +1 \\ \frac{1}{N_- + 2} \text{ se } y_i = -1 \end{cases}, \quad i = 1, \dots, l \quad (4.22)$$

N_+ è il numero di elementi appartenenti alla classe 1 mentre N_- quello della classe -1 .

Nelle simulazioni, che verranno riportate nel proseguo dell'elaborato, si è fatto riferimento all'implementazione proposta in [35] il cui codice matlab viene illustrato in appendice B.

Un esempio esplicativo di quanto visto sin ora è l'applicazione delle SVM e della probabilità di Platt al Fisher's Iris data set, il quale, contiene la dimensione dei sepal di tre varietà di fiori: Setosa, Versicolor e Virginica. Si desidera istruire una support vector machine che riconosca i fiori Setosa; tutti i dati di train corrispondenti alla specie Setosa apparterranno alla classe 1 mentre quelli relativi alle altre due saranno posti in -1 . In questo caso le classi sono separabili e si è ricercato un iperpiano lineare che li dividesse, come esposto

nel caso delle classi separabili in sezione 3.2. La classificazione che ne deriva è illustrata in Figura 4.4 (in alto). I dati relativi alla specie *Setosa* sono riportati come asterischi mentre quelli delle altre specie come +, i dati di train sono di colore verde e rosso mentre quelli di validazione in azzurro e magenta. I support vector che individuano l'iperpiano sono i campioni cerchiati. A partire dai medesimi dati di train è stata istruita la probabilità a posteriori e, come si osserva dalla Figura 4.4 (in basso), vengono rispecchiate le attese: i campioni più lontani dall'iperpiano hanno un'alta probabilità di essere classificati in maniera corretta mentre per quelli prossimi al margine di separazione la probabilità scende anche al 40% di appartenere alla specie *Setosa*.

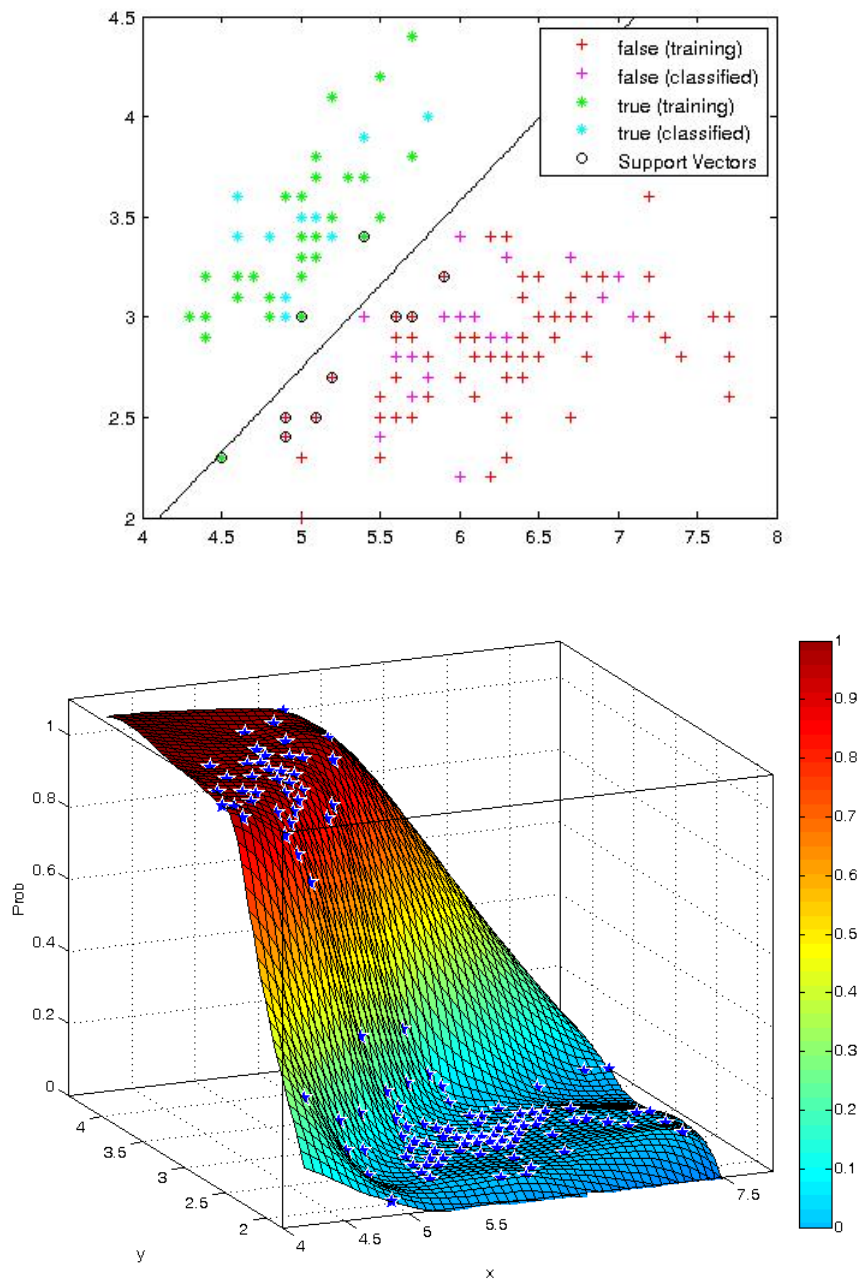


Figura 4.4: Classificatore SVM lineare e probabilità di Platt applicati al Fisher's Iris data set

Capitolo 5

Risultati sperimentali

In questo capitolo si analizzano i risultati ottenuti applicando le tecniche proposte in precedenza allo studio dei malfunzionamenti in impianti di tipo VAVAC. Si sono presi in considerazione le seguenti tipologie di guasto:

- Guasto sulla cassetta terminale: il malfunzionamento inficia il corretto afflusso di aria nella zona interessata.. Questo può portare ad una variazione indesiderata della temperatura della stanza.
- Guasto della pompa del flusso d'acqua della UTA: il flusso d'acqua che percorre le serpentine della UTA, a seguito del guasto, non è quello richiesto dal sistema di controllo. Nel caso in cui, ad esempio, il flusso sia inferiore da quello richiesto, il controllo tenta di bilanciare tale mancanza aprendo di più le bocchette di zona (la temperatura dell'aria di mandata risulta infatti superiore a seguito del malfunzionamento); anche la velocità del ventilatore deve aumentare per mantenere costante la pressione nelle condotte.
- Guasto sul sensore della temperatura della stanza: in questo caso il malfunzionamento è relativo ad un offset sulla misura. Ad esempio, se il sensore rileva una temperatura errata di $+1^{\circ}\text{C}$ allora il controllo porta la temperatura della zona ad un grado inferiore rispetto al set-point desiderato.
- Guasto della valvola di ricircolo: in seguito al malfunzionamento essa non introduce la corretta portata di aria esterna nella UTA. Questo può

determinare una perdita delle prestazioni del sistema, sprechi energetici e non ultimo compromettere le specifiche sul ricircolo d'aria nei locali (se per via di un guasto essa introduce meno aria ossigenata nelle stanze si ha un maggior tasso di anidride carbonica).

5.1 Dati di allenamento e di validazione

La prima fase nella progettazione di un sistema di diagnosi semi-automatica, basato su tecniche black-box, è la raccolta dei dati dai segnali accessibili. Nel caso in esame si ha accesso a tutte le variabili coinvolte nel controllo ovvero:

- i segnali di attuazione del controllo: il flusso d'acqua richiesto al chiller, la velocità del ventilatore di mandata, i flussi d'aria entranti nelle stanze.
- le variabili controllate: la temperatura dell'aria e la pressione statica nel condotto di mandata, le temperature nelle stanze.
- le stime di variabili quali la temperatura dell'aria di ritorno, la temperatura dell'acqua uscente dalla UTA e quella dell'aria in ingresso.

Vale la pena notare che la temperatura esterna e i carichi presenti nelle zone trattate rivestono un ruolo significativo per quando riguarda la raccolta dati. Tali grandezze, pur essendo assimilabili a dei disturbi per il sistema di controllo, non possono essere trascurate in quanto influenzano il punto lavoro ed il sistema di FDD deve essere in grado di discriminare se tale variazione sia imputabile a tali variazioni oppure ad un malfunzionamento per evitare la segnalazione di falsi allarmi. Le simulazioni effettuate evidenziano che da un lato è necessario istruire le SVM in diverse condizioni operative, dall'altro non è necessario coprire l'intero range di funzionamenti possibili. Nelle simulazioni condotte i dati sono stati raccolti per una temperatura esterna di 25, 29 e 33°C e carichi di 450, 500 e 550W.

I dati per il train sono stati raccolti attraverso simulazioni della durata di un'ora con il guasto che si presenta dopo quindici minuti, di tali dati si considera un valore per le variabili di interesse prima del fault e uno dopo, avendo l'accortezza che il sistema raggiunga una condizione di regime. In questa fase il malfunzionamento si manifesta repentinamente a gradino, ovvero il sistema

passa in un intervallo di campionamento dal funzionamento normale al fault. L'ampiezza del gradino si considera in percentuale rispetto al punto di lavoro della variabile interessata dal guasto. Si considerano variazioni tra il 10% ed il 70%. Nel caso specifico dei guasti ai sensori si è imposto un errore di offset da mezzo fino ad un massimo di 4 gradi. Avendo la possibilità di generare agevolmente, grazie al modello, campioni la validazione non è stata effettuata con la tecnica della cross validation ma si è preferito rifettuare la procedura di raccolta con diverse percentuali e con un numero inferiore di simulazioni rispetto ai dati di train. In questo modo si rimuove la componente casuale e si è a conoscenza di quali dati sono utilizzati per il train delle SVM. Inoltre, il data set risulta simmetrico e ugualmente distribuito rispetto all'utilizzo della cross-validation.

5.2 Risultati sperimentali

Si vuole ora riportare e commentare i risultati ottenuti. In un primo momento si tralascia il KPCA sostituendolo, come avviene nella maggior parte delle trattazioni in letteratura, con una tabella decisionale. Tale tabella, riportata di seguito, determina quali variabili sono utilizzate per istruire le SVM e successivamente per effettuare la diagnosi in real-time; si considerano le variabili che sono maggiormente influenzate da un dato malfunzionamento.

Fault	T_{z1}	T_{z2}	T_{sa}	T_{ai}	T_r	\dot{m}_{a1}	\dot{m}_{a2}	\dot{m}_w
Valvola Acqua			X			X	X	X
VAV stanza 1	X					X	X	X
VAV stanza 2						X	X	X
Valvola ricircolo				X	X	X	X	X
Sensore stanza 1	X							X
Sensore stanza 2		X						X

Un'ultima precisazione riguarda la disposizione dei classificatori: in letteratura, nella identificazione di più guasti, è consuetudine istruire una SVM per ogni guasto (identification) più una che discrimini tra il comportamento normale del sistema e la presenza di un guasto (detection), le SVM sono disposte in cascata come illustrato in Figura 5.1. L'ordine delle SVM è scelto in base alla

probabilità del guasto a cui si riferiscono ovvero vengo interrogate per prime quelle associate ai malfunzionamenti che si presentano con maggiore frequenza. Il vantaggio di questo approccio è che si riduce il numero di interrogazioni del classificatore ed inoltre ogni SVM deve saper discriminare solo tra il proprio guasto e quelli analizzati successivamente nella catena agevolando la sintesi dei classificatori e il riducendo numero di support vector (questo si tramuta in un minor numero di calcoli in fase di classificazione). Per contro, l'approccio

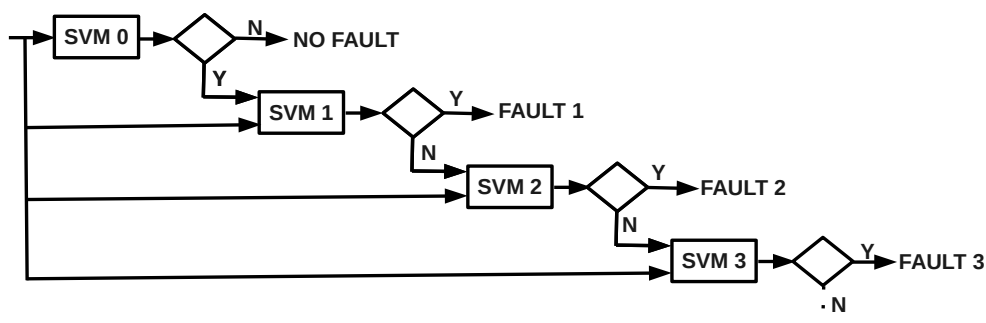


Figura 5.1: Disposizione in serie dei classificatori.

appena analizzato risiede nel fatto che un errore di classificazione da parte di una SVM risulta esiziale ai fini di una buona diagnosi. Si è dunque preferito disporre le SVM in parallelo (Figura 5.2). Questo non comporta un elevato tempo di calcolo in quanto la classificazione si basa su semplici operazioni e l'esito di tutte le SVM può essere calcolato agevolmente; inoltre, vi è la possibilità di associare la probabilità a posteriori di Platt ad ogni classificazione eseguita.

5.2.1 Malfunzionamento in condizioni operative costanti

Si valutano le prestazioni dei classificatori SVM relativamente al sistema VAV in condizioni nominali nell'intorno di un prefissato punto di lavoro con temperatura esterna e carichi interni costanti. Nella sezione successiva si completa l'analisi valutando l'effetto di carichi e temperatura esterna variabili.

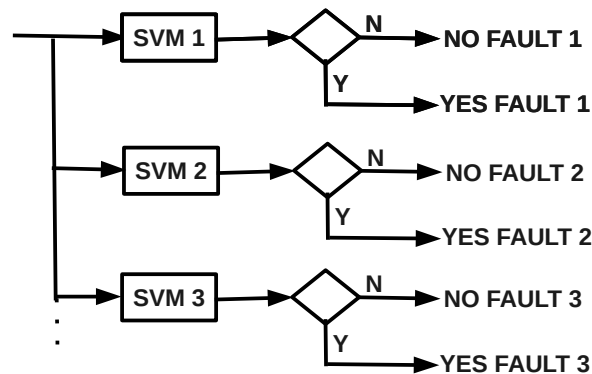


Figura 5.2: Disposizione in parallelo dei classificatori.

Malfunzionamento valvola acqua batteria

Il malfunzionamento si presenta sulla valvola di controllo del flusso d'acqua che percorre la batteria di raffreddamento. Il guasto (Figura 5.3, primo grafico) si manifesta dopo due ore e ha una durata di quattro ore poi, dopo due ore, il sistema ritorna alle condizioni operative nominali. In particolare si impone che l'apertura della valvola sia falsata, nell'arco delle quattro ore, da un incremento del 30% rispetto al valore imposto dal controllore. Il secondo grafico della Figura 5.3 evidenzia come il controllo tenti di contrastare questo inconveniente con un'azione speculare. In questo caso il sistema agendo sulle bocchette VAV delle stanze riesce comunque a mantenere la temperatura desiderata nelle zone mantenendo inalterate le prestazioni di comfort. Nei grafici seguenti è evidenziata la risposta dell'apparato di FDD: per la SVM (terzo grafico) e KNN (quinto grafico) in cui il valore 1 rappresenta il corretto funzionamento mentre 0 identifica il guasto; nel quarto grafico è rappresentata la probabilità (di Platt), che fornisce un valore di confidenza per la classificazione: il valore 1 corrisponde alla classificazione NON guasto mentre 0 a quella di guasto (entrambe con probabilità unitaria). Dai risultati si evince la differenza di prestazione tra K-NN ed SVM: mentre il K-NN riesce a rilevare la presenza del guasto solo per un breve periodo di tempo la risposta delle macchine a vettori di supporto è ottima. Si può verificare che se si aumentano i dati di train le prestazioni del K-NN migliorano sensibilmente ma comunque non superano mai quelle delle SVM. La scelta del K ottimale si è effettuata minimizzando lo

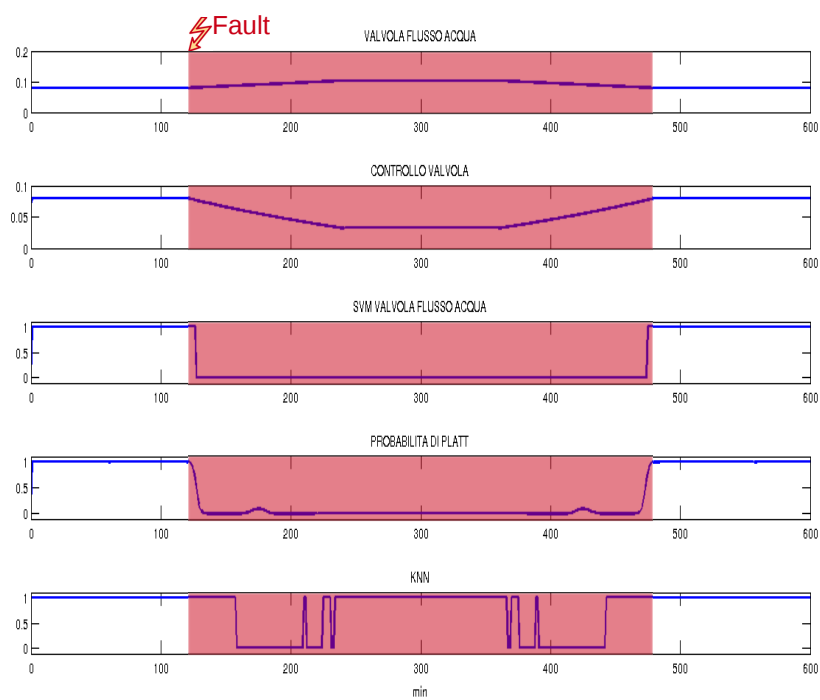


Figura 5.3: Fault del 30% su valvola flusso acqua batteria con temperatura esterna e carichi costanti.

stesso indice di costo presentato nel paragrafo 4.2 per le SVM.

Per confermare la bontà del sistema di FDD, si controlla che le SVM riferite agli altri fault non generino dei falsi allarmi: la risposta di tutte le SVM in riferimento al tipo di guasto appena proposto viene presentata in Figura 5.4.

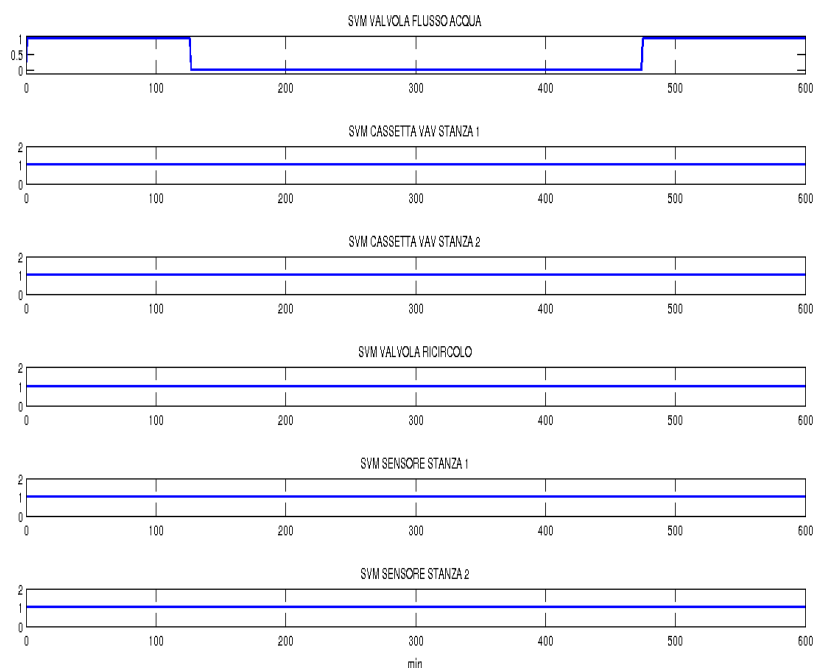


Figura 5.4: Fault del 30% su valvola flusso acqua batteria con temperatura esterna e carichi costanti.

Il risultato ottenuto evidenzia che non sono presenti falsi allarmi generati dalle altre SVM e quella relativa al flusso dell'acqua ha una risposta pronta. Il breve ritardo nella segnalazione è accettabile. Una maggiore prontezza potrebbe comportare una riduzione in termini di robustezza. Le simulazioni condotte indicano come variando alcuni parametri del sistema quali il coefficiente di scambio della batteria, le dimensioni delle stanze, la superficie di esposizione verso l'esterno e la trasmittanza delle zone di una percentuale fino al 10% il sistema di FDD continui a funzionare correttamente.

La situazione speculare a quella appena descritta, con un malfunzionamento

di -30% anziché $+30\%$, è riportata nelle Figure 5.5 e 5.6. Le considerazioni che se ne traggono sono analoghe alle precedenti con l'unica nota che il segnale di controllo tende a saturare. Le differenze di comportamento del sistema è da attribuire alla non simmetria dell'impianto in quanto si tratta di un sistema di condizionamento (raffreddamento dell'aria).

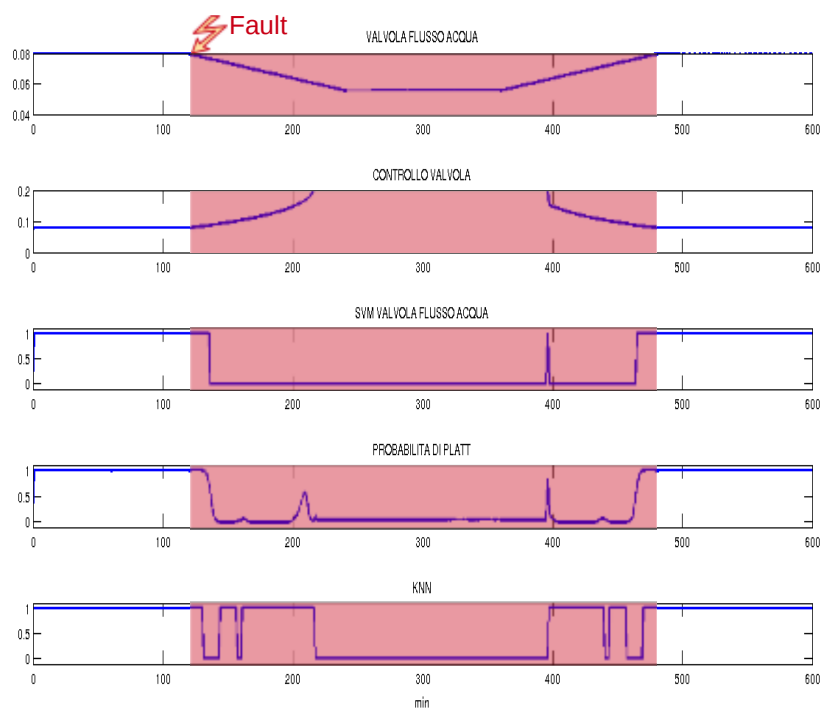


Figura 5.5: Fault di -30% su valvola flusso acqua batteria con temperatura esterna e carichi costanti.

Malfunzionamento cassette di zona VAV

Come per il guasto sul flusso dell'acqua si analizza un degrado del componente che porta ad un blocco della bocchetta in una posizione di $+30\%$ rispetto al punto di lavoro richiesto dal controllo. In Figura 5.7 è riportato il comportamento della bocchetta e delle richieste del sistema di controllo. Si osserva come il guasto abbia ripercussioni anche sul controllo del flusso d'acqua della UTA,

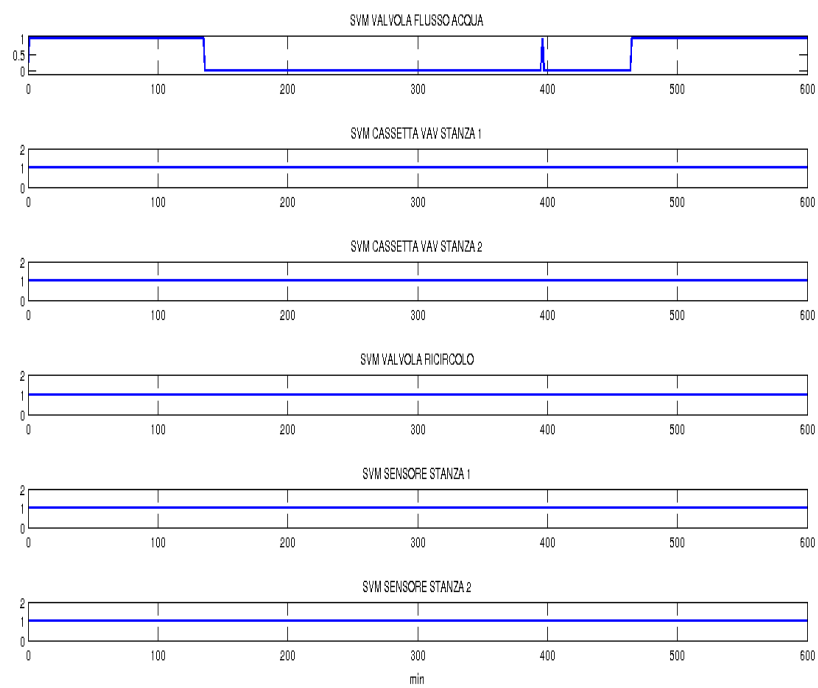


Figura 5.6: Fault di -30% su valvola flusso acqua batteria con temperatura esterna e carichi costanti.

questo perché la temperatura dell'aria di ritorno è maggiore. La temperatura della stanza è destinata a salire a causa dei carichi e della temperatura esterna.

Il sistema di fault detection (Figura 5.8) rileva prontamente il guasto e quindi è possibile, eventualmente, risolvere il problema tempestivamente senza dover attendere che l'effetto dei carichi comprometta il comfort degli utenti.

Il grafico in Figura 5.8 sottolinea l'importanza della probabilità a posteriori: in un intervallo di tempo sia la SVM relativa alla prima che alla seconda zona rilevano un guasto ma la corrispondente indicazione di Platt permette di discriminare; infatti, la probabilità di avere un guasto è sempre superiore per la zona 1 che per la 2.

Il guasto sulla bocchetta VAV della seconda stanza presenta il comportamento simmetrico a quello appena illustrato per la prima stanza.

Malfunzionamento valvola di ricircolo

La valvola di ricircolo concorre a determinare la composizione dell'aria in ingresso alla UTA. Questo componente non è controllato dal sistema e solitamente le si impone in modo da garantire un 25% di aria di rinnovo, proveniente dall'esterno, e il restante 75% di aria di ricircolo dalle condotte di ritorno. In Figura 5.9 si riporta una simulazione con una variazione della posizione del 30%. Anche in questo caso le prestazioni del sistema di rilevamento guasti si possono definire soddisfacenti. Il caso di un malfunzionamento che determina un -30% di variazione rispetto alla posizione nominale è del tutto simile in quanto l'unica differenza è che la risposta è ancora più celere.

Malfunzionamento sensori

In Figura 5.10 viene evidenziata la risposta della FDD ad un malfunzionamento corrispondente alla presenza di un offset, pari a 3°C, sul sensore della prima zona di tre gradi. Si osserva che la temperatura della zona viene diminuita di tre gradi, questo perché il sistema di controllo ha un'informazione sbagliata dai dati di retroazione. Si osserva inoltre che non vi sono problemi di false classificazioni con il sensore dell'altra stanza.

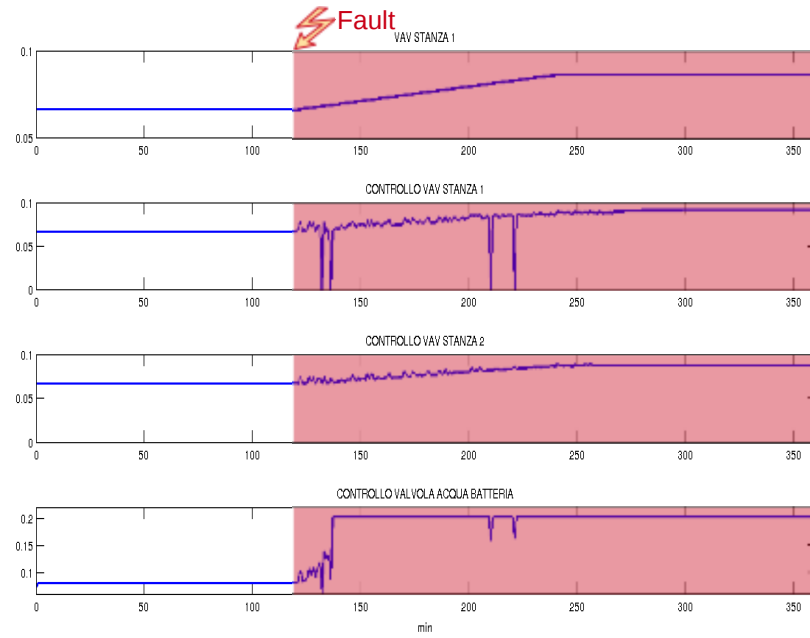


Figura 5.7: Attuatore VAV stanza 1 e segnali di controllo con fault del 30%.

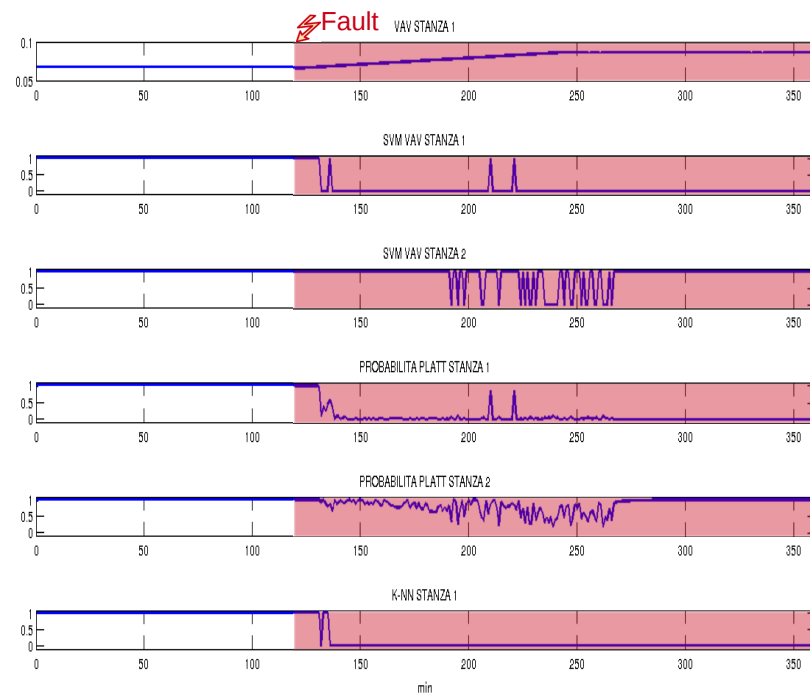


Figura 5.8: FDD malfunzionamento bocchetta VAV zona 1.

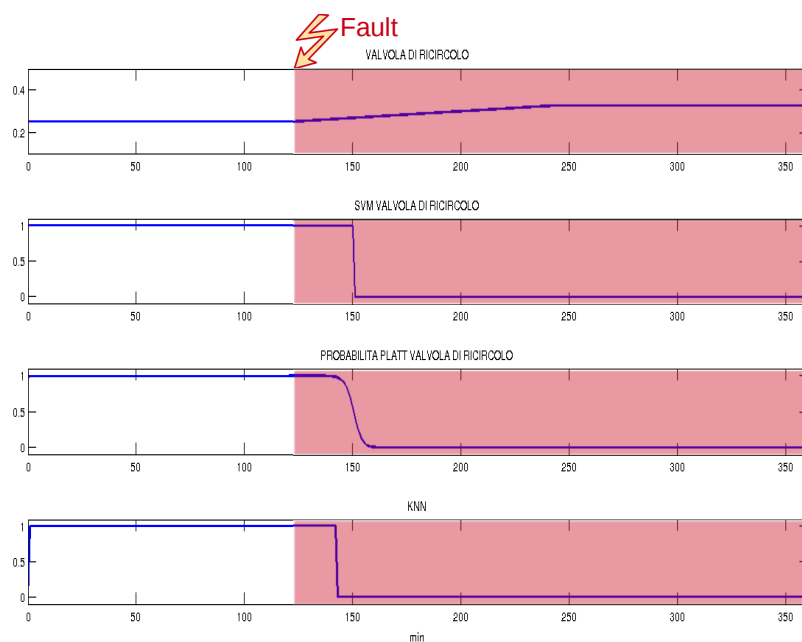


Figura 5.9: Fault di +30% su valvola ricircolo con temperatura esterna e carichi costanti.

5.2.2 Malfunzionamento in condizioni operative variabili

Il caso più realistico corrisponde nel considerare una temperatura esterna e dei carichi variabili del tempo. Nelle simulazioni riportate di seguito si è considerata una temperatura esterna che varia sinusoidalmente nel corso di 24 ore tra una minima di 25°C ed un massima di 33°C mentre i carichi variano tra i 440 ed 560W (Figura 5.11). I risultati ottenuti non sono ottimi come per il caso costante ma evidenziano la possibilità di usare utilmente per la FDD le tecniche SVM. Per migliorare le prestazioni del sistema è necessario porre particolare attenzione alla scelta dei dati di allenamento e prevedere una tecnica per la scelta dei dati di maggior interesse.

Malfunzionamento valvola acqua batteria

In Figura 5.12 è riportato il risultato di una simulazione in cui si presenta un malfunzionamento sulla pompa dell'acqua pari ad una variazione di portata del 20% dopo 2 ore di corretto funzionamento. Con la linea tratteggiata è evi-

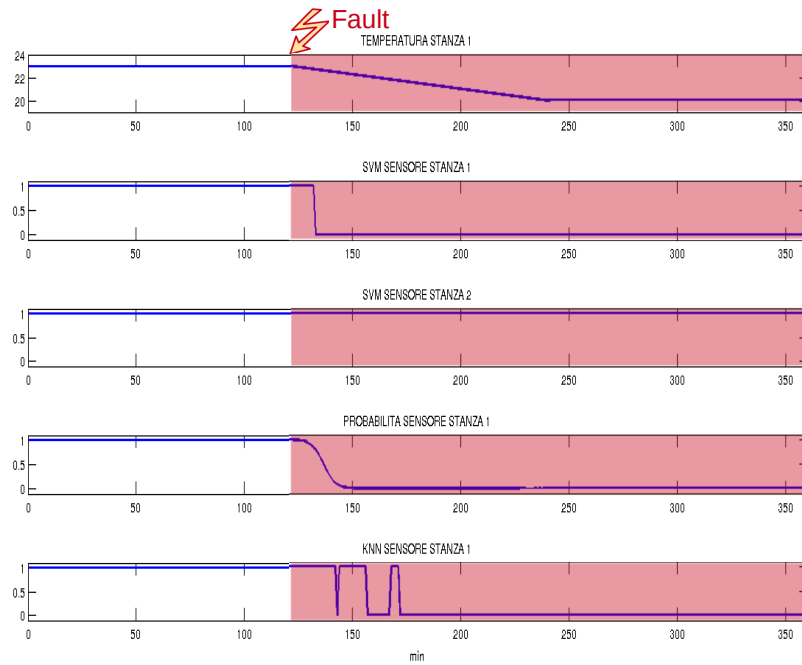


Figura 5.10: Fault di $+3^{\circ}\text{C}$ su sensore stanza 1 con temperatura esterna e carichi costanti.

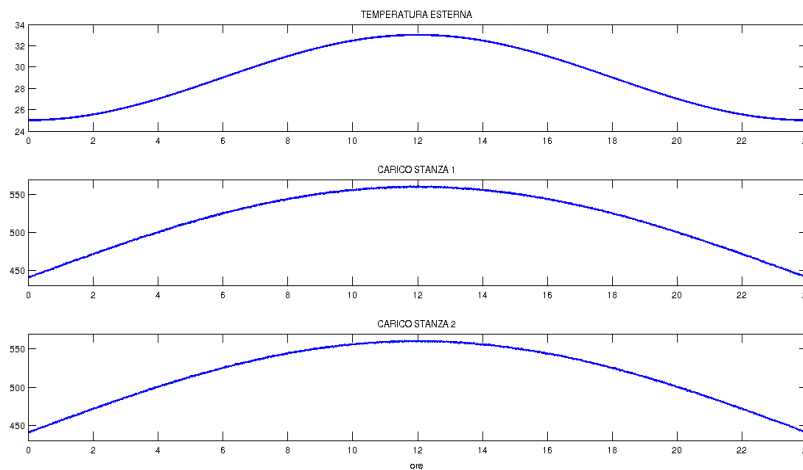


Figura 5.11: Variazioni di carichi e temperatura esterna nel tempo.

denziato l'andamento che avrebbe seguito il sistema se non si fosse presentato il guasto (linea continua). La SVM riesce ad individuare prontamente la presenza del guasto ma non mantiene la segnalazione di guasto in modo costante. La probabilità non riesce a fugare i dubbi ma comunque, in certi tratti in cui la classificazione indica che non vi sia guasto, suggerisce di dubitare della classificazione. È probabile che una taratura con maggiori valori di temperature esterne e carichi oltre che una scelta più accurata delle variabili di train potrebbe migliorare i risultati.

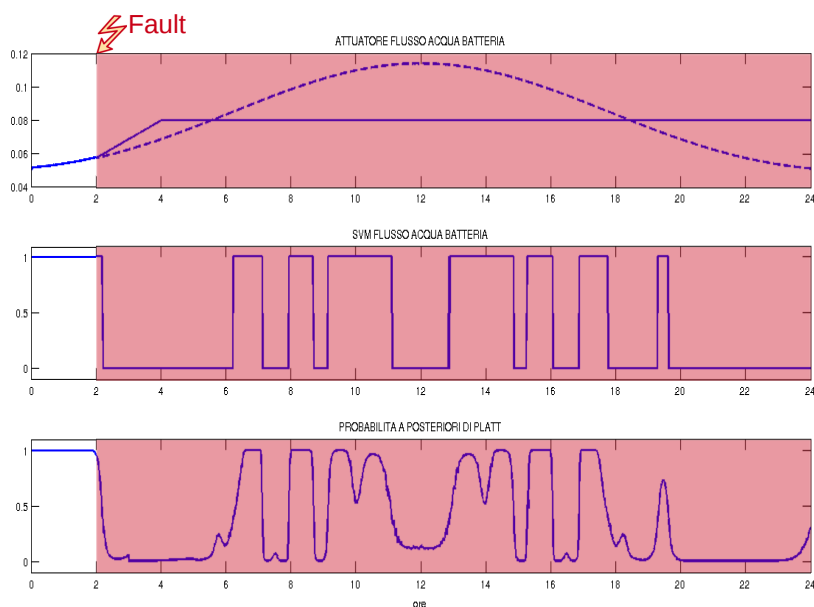


Figura 5.12: Fault su flusso acqua batteria di raffreddamento.

5.2.3 Malfunzionamento cassette di zona VAV

Nel caso di guasto sulla bocchetta VAV la SVM riesce a rilevare il guasto e mantenere la segnalazione per tutta la durata della simulazione a meno delle ultime ore. È importante notare come la probabilità a posteriori indichi una possibile errata classificazione (infatti la probabilità di non avere malfunzionamenti rimane inferiore al 60% (Figura 5.13)).

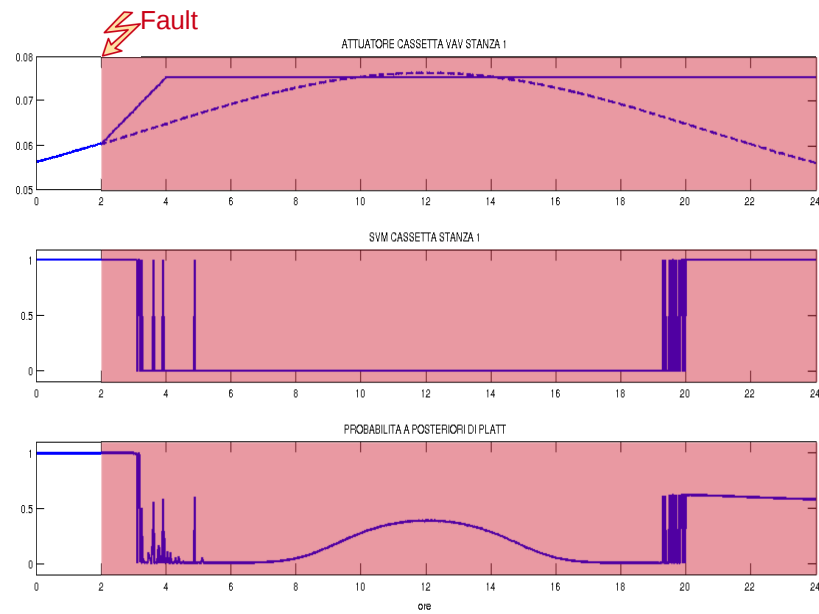


Figura 5.13: Fault su bocchetta VAV stanza 1.

Malfunzionamento sensori

La FDD sui sensori non sembra risentire delle variazioni di carico e temperatura esterna e si mantiene ottima come nel caso di grandezze costanti presentato in precedenza come si vede in Figura 5.14.

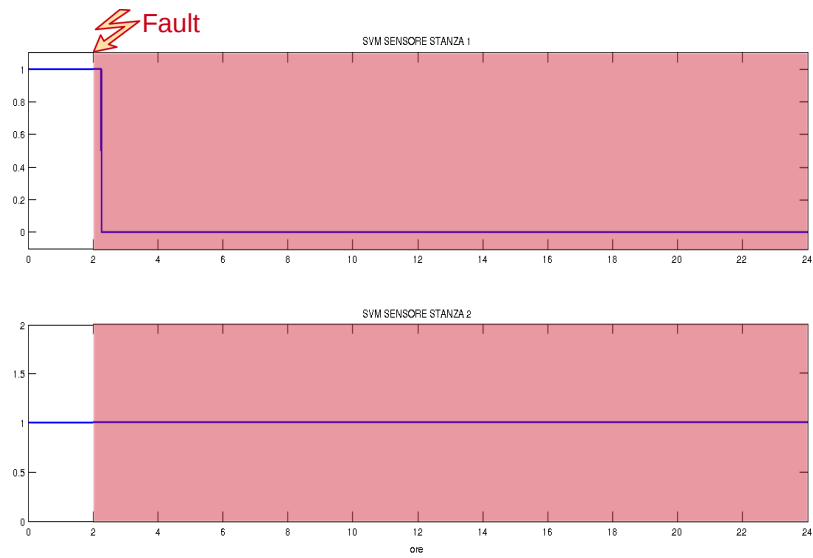


Figura 5.14: Fault su sensore stanza 1.

Capitolo 6

Conclusioni e sviluppi futuri

In questa tesi si è analizzato il problema della FDD applicato a sistemi VAVAC. Dopo aver presentato un modello non lineare che permette la generazione di dati per progettare e testare tecniche di fault detection si è scelto di focalizzarsi sugli approcci di diagnosi di tipo black-box ed in particolare quelli che impiegano support vector machine. Si è posta particolare attenzione alla progettazione delle SVM riscontrando quanto sia essenziale procedere ad una accurata taratura dei suoi parametri e quelli del kernel scelto. A tal fine l'utilizzo di una tecnica euristica come il PSO abbinata ad un indice di costo che valuta gli errori di classificazione su un data set di validazione si è rilevata vincente. Un altro aspetto analizzato con particolare attenzione è stata l'informazione fornita dai classificatori SVM che è stata completata con l'aggiunta di una probabilità a posteriori che costituisce un valido ausilio alla classificazione dei malfunzionamenti. I risultati ottenuti nel caso di temperatura esterna e carico costanti sono ottimi: non sono presenti falsi allarmi e le classificazioni sono accurate e repentine. Nel caso di temperatura esterna e carico variabile gli esiti non sono stati completamente soddisfacenti ma lasciano comunque intuire che con degli accorgimenti queste tecniche permettano di ottenere un apparato di FDD che dia ottime prestazioni. Per migliorare i risultati, come sviluppo futuro, è necessario uno studio accurato sulla scelta dei dati di train sia da un punto di vista delle entità e numero di simulazioni per ogni guasto che di scelta delle temperature esterne e carichi ai quali trainare la SVM. Un altro aspetto interessante è la fase di pre-processo dei dati raccolti per permettere

alle SVM ed eventualmente ai K-NN di lavorare con i dati più significativi ai fini della classificazione; in tal senso si ritiene che il KPCA, già presentato in linea teorica in questo elaborato, possa portare ad un miglioramento delle prestazioni e possa essere un ausilio per la scelta delle variabili maggiormente influenti ai fini di individuare un determinato malfunzionamento. Anche per il KPCA è necessaria una fase di tuning in cui si sceglie al meglio il parametro del kernel che si utilizza, per questo si suggerisce di valutare l'indice di costo presentato nel capitolo 4.1: tale indice cerca di ottenere la maggiore separazione possibile, in base alla classe di appartenenza, dei dati train nel feature space individuato dal kernel. Se si considera un kernel rbf l'indice di costo dipende da un solo parametro (in questo caso il problema di ottimizzazione monodimensionale corrispondente può essere risolto efficacemente impiegando il metodo della sezione aurea riportato in appendice A). Concludendo, in questo elaborato si è discussa l'importanza della FDD e sono state presentate delle tecniche di machine learning che permettono di individuare e classificare i malfunzionamenti senza l'aggiunta di sensori o modifiche sull'impianto.

Metodo di ottimizzazione aurea

Data una funzione $f(x)$ che dipenda da un'unica variabile, si vuole trovare il punto in corrispondenza del quale f assume valore massimo o minimo. I due problemi sono equivalenti in quanto il massimo (minimo) di f è il minimo (massimo) di $-f$. Esistono due tipi di metodi per la ricerca di massimi o minimi: quelli che richiedono il calcolo delle derivate e quelli in cui non è richiesto. Il metodo della sezione aurea appartiene a quella seconda classe. Prima di descrivere il metodo di ottimizzazione si espone il concetto di *intrappolamento*: Un minimo di una funzione $f(x)$ si dice intrappolato se esistono tre punti a, b, c con $a < b < c$ tali che $f(b) < f(a)$ e $f(b) < f(c)$. In questo caso, se $f(x)$ è continua, il minimo si troverà in un punto dell'intervallo $[a, c]$ Figura A.1.

Nel metodo della sezione aurea si parte da un minimo inizialmente intrappolato. Detti a, b, c i punti che intrappolano il minimo si ha:

$$a < b < c \text{ tali che } f(b) < f(a) \wedge f(b) < f(c). \quad (\text{A.1})$$

Analogamente a quanto avviene nel metodo della bisezione, si cerca un nuovo punto x , compreso tra a e b oppure tra b e c , che restringa l'intervallo. Supponendo di considerare $x \in [b, c]$, se $f(x) > f(b)$ il nuovo tripletto di punti sarà a, b, x altrimenti sarà b, x, c Figura A.2.

Il processo di intrappolamento viene arrestato quando la distanza $c - a$ è sufficientemente piccola. Se ϵ è la precisione della macchina, si potrebbe pensare di fermare il processo quando $a = b(1 - \epsilon)$ e $c = b(1 + \epsilon)$; in realtà conviene

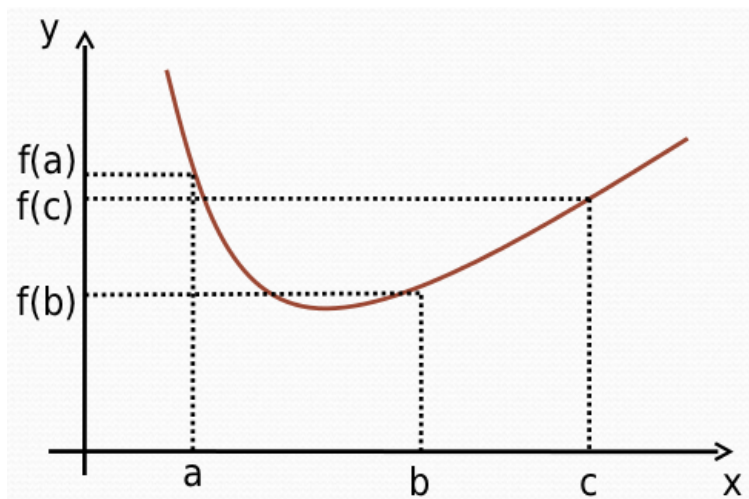


Figura A.1: Intrappolamento del minimo.

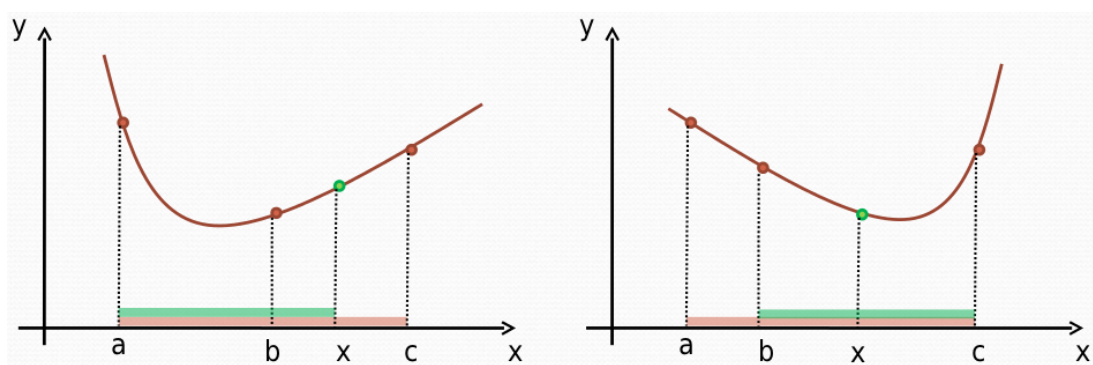


Figura A.2: Scelta del triplitto per il passo successivo.

fermarsi prima per evitare troppi calcoli.

Per scegliere il nuovo punto x , ad ogni iterazione, si pone

$$W = \frac{b-a}{c-a} \implies 1-W = \frac{c-b}{c-a} \quad (\text{A.2})$$

e supponendo che il punto x successivo si trovi tra b e c :

$$Z = \frac{x-b}{c-a}. \quad (\text{A.3})$$

Se x si trova tra a e b si ragiona analogamente, in questo caso si ottiene $Z < 0$. A seconda del valore di $f(x)$ si sceglierà il nuovo tripletto di punti:

- se $f(x) > f(b)$ i nuovi tre punti da usare sono a, b, x , il nuovo intervallo $[a, x]$ ha lunghezza $(W+Z)(c-a)$,
- se $f(x) < f(b)$ i nuovi tre punti da usare sono b, x, c , il nuovo intervallo $[b, c]$ ha lunghezza $(1-W)(c-a)$.

Conviene scegliere Z in maniera tale che, qualunque condizione si verifichi, l'intervallo finale abbia sempre la stessa lunghezza:

$$W+Z = 1-W \implies Z = 1-2W \quad (\text{A.4})$$

Con questa scelta $|b-a| = |x-c|$:

$$\begin{aligned} b-a &= W(c-a) \\ x-c &= (x-b) - (c-b) = Z(c-a) - (1-W)(c-a) = (Z-1+W)(c-a) = -W(c-a) \end{aligned}$$

Il punto x è il simmetrico di b nell'intervallo $[a, c]$ e si trova sempre all'interno del più lungo tra i segmenti $[a, b]$ (se $Z < 0$) e $[b, c]$ (se $Z > 0$).

Considerando i tripletto di punti a, b, c e b, x, c :

$$\frac{b-a}{c-a} \frac{x-b}{c-b} = \frac{Z(c-a)}{(1-W)(c-a)} = \frac{Z}{1-W} \quad (\text{A.5})$$

Se gli intervalli vengono divisi sempre allo stesso modo, allora i due rapporti devono essere uguali e quindi deve essere:

$$\frac{Z}{1-W} = W \implies Z = W(1-W) \quad (\text{A.6})$$

$$1 - 2W = W - W^2 \implies W = \frac{3 - \sqrt{5}}{2} = 0.38197 \quad (\text{A.7})$$

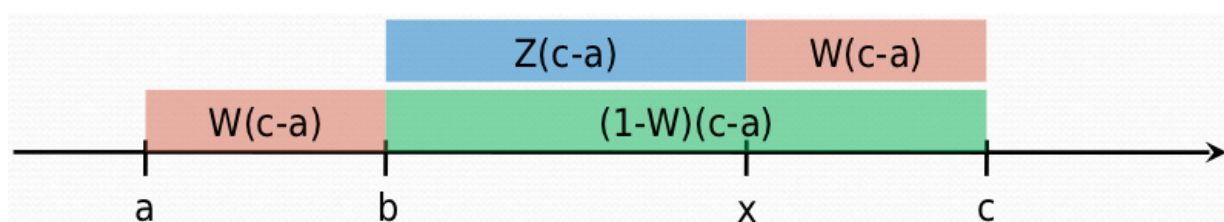


Figura A.3: Partizione del dominio di ricerca.

Si possono trarre alcune conclusioni sull'algoritmo appena illustrato:

- L'intrappolamento ottimale porta a triplette di punti in cui il punto centrale si trova ad una distanza frazionaria $W = 0.38197$ da uno dei due e a $1 - W = 0.61803$ dall'altro ovvero le sezioni auree.
- Dato un tripletto di punti a, b, c , il punto successivo x in cui calcolare il valore della funzione si trova alla distanza frazionaria $W=0.38197$ dal punto di mezzo del tripletto, nel più lungo dei due intervalli $[a, b]$ o $[b, c]$.
- Se gli intervalli di partenza non rispettano i rapporti aurei non è un problema perchè la procedura iterativa converge rapidamente verso intervalli ottimali.
- La dimensione dell'intervallo ottenuto alla n -esima iterazione è pari a 0.61803 volte la dimensione dell'intervallo ottenuto alla $(n-1)$ -esima iterazione.

Appendice **B**

Pseudocodice algoritmo di Platt

Di seguito viene riportato il codice Matlab dell'algoritmo per determinare i parametri A e B della funzione a sigmoide che rappresenta la funzione a probabilità a posteriori.

```
function [A,B]=platt(out,target,prior0,prior1)
% find the coefficients A and B such that the posterior probability
% of  $P(y=1|x) = 1/(1+\exp(A*f(x)+B))$ , where  $f(x)$  is the output
% of the SVM
%
% out: vector of outputs of the SVM on a validation set
% target: validation labels
% prior0: number of negative points
% prior1: number of positive points
%
% If no validation set is available, one might use the training
% set to do a leave-one-out procedure. Using the span, this means
% replacing out by something like out-target.*alpha.*span

%Parameter setting
maxiter=100;      %Maximum number of iterations
minstep=1e-10;   %Minimum step taken in line search
sigma=1e-12;     %Set to any value sufficient to make  $H' = H + \text{sigma} I$  always PD
%Construct initial values: target support in array t, initial function value in f
hiTarget=(prior1+1.0)/(prior1+2.0);
loTarget=1/(prior0+2.0);
t = (target>=0)*hiTarget + (target<0)*loTarget;
```

```

A=0.0;
B=log((prior0+1.0)/(prior1+1.0));

%% A vector version that calculates
%% fval = sum m_i
%% m_i = t_i fApB_i + log(1.0 + exp(-fApB_i)) for fApB_i >= 0
%% m_i = (t_i - 1) fApB_i + log(1.0 + exp(fApB_i)) for fApB_i < 0
fApB=out*A+B;
logF=log(1.0+exp(-abs(fApB)));
fval = sum((t - (fApB < 0)) .* fApB + logF);

% Save repeated calculation in the main loop
out2=out.*out;

for it = 1:maxiter
    %Update Gradient and Hessian (use H' = H + sigma I)

    %% A vector version that calculates
    %% d1_i = t_i - p_i
    %% d2_i = p_i q_i
    %% p_i = exp(-fApB) / (1.0 + exp(-fApB)), q_i = 1.0 / (1.0 + exp(-fApB)) for fApB_i >= 0
    %% p_i = 1.0 / (1.0 + exp(fApB)), q_i = exp(fApB) / (1.0 + exp(fApB)) for fApB_i < 0

    expF = exp(-abs(fApB));
    oneexpFinv = (1.0+expF).^(-1);
    d2 = expF.*oneexpFinv.*oneexpFinv;
    d1 = t - max(expF, (fApB<0)).*oneexpFinv;
    h11 = sigma + sum(out2.*d2);
    h22 = sigma + sum(d2);
    h21 = sum(out.*d2);
    g1 = sum(out.*d1);
    g2 = sum(d1);

    if (abs(g1)<1e-5 & abs(g2)<1e-5) %Stopping criteria
        break;
    end;

    detinv=(h11*h22-h21*h21).^(-1);
    dA=-(h22*g1-h21*g2) .* detinv;
    dB=(-h21*g1+h11*g2) .* detinv; %Modified Newton direction
    gd=g1*dA+g2*dB;

```

```
stepsize=1;
while (stepsize >= minstep) %Line search
    newA=A+stepsize*dA;
    newB=B+stepsize*dB;

    %% A vector version that caculates
    %% newf = sum m_i
    %% m_i = t_i fApB_i + log(1.0 + exp(-fApB_i)) for fApB_i >= 0
    %% m_i = (t_i - 1) fApB_i + log(1.0 + exp(fApB_i)) for fApB_i < 0
    fApB=out*newA+newB;
    logF=log(1+exp(-abs(fApB)));
    newf = sum((t - (fApB < 0)) .* fApB + logF);

    if (newf<fval+0.0001*stepsize*gd) %Check sufficient decrease
        A=newA; B=newB; fval=newf;
        break
    else
        stepsize=stepsize*0.5;
    end;

    if (stepsize < minstep) %Line search fails
        break
    end;
end;
end;
```


Bibliografia

- [1] *HVAC Systems and Equipment*, ASHRAE - Handbook, 2008, www.ashrae.org.
- [2] S. Katipamula and M. R. Brambley, "Methods for Fault Detection, Diagnostics, and Prognostics for Building Systems - A Review, Part I," *International Journal of HVAC&R Research*, vol. 11, no. 1, January 2005, www.ashrae.org.
- [3] J. Liang and R. Du, "Model-based Fault Detection and Diagnosis of HVAC systems using Support Vector Machine method," *International Journal of Refrigeration*, no. 30, pp. 1104–1114, 2007.
- [4] R. Isermann, "Model-based fault detection and diagnosis. Status and applications," 2004.
- [5] V. Palade, C. D. Bocaniala, and L. Jain, *Computational Intelligence in Fault Diagnosis*. London: Springer, 2006.
- [6] J. M. House, W. Y. Lee, and D. R. Shin, "Classification Techniques for Fault Detection and Diagnosis of an Air-Handling Unit," *ASHRAE Transaction: Symposia*, pp. 1087–1097, May 1999.
- [7] W. Y. Lee, J. M. House, and N. H. Kyong, "Subsystem level fault diagnosis of a building's air-handling unit using general regression neural networks," *Applied Energy*, no. 77, pp. 153–170, 2004, www.elsevier.com/locate/apenergy.

-
- [8] H. Yoshida, S. Kumar, and Y. Morita, "Online fault detection and diagnosis in VAV air handling unit by RARX modeling," *Energy and Buildings*, no. 33, pp. 391–401, 2001, www.elsevier.com/locate/enbuild.
- [9] Z. Hou, Z. Lian, Y. Yao, and X. Yuan, "Data mining based sensor fault diagnosis and validation for building air conditioning system," *Energy Conversion and Management*, no. 47, pp. 2479–2490, 2006, www.elsevier.com/locate/enconman.
- [10] Z. Du and X. Jin, "Detection and diagnosis for sensor fault in HVAC systems," *Energy Conversion and Management*, no. 48, pp. 693–702, 2007, www.elsevier.com/locate/enconman.
- [11] N. Tudoroiu and M. Zaheeruddin, "Fault detection and diagnosis of valve actuators in HVAC systems," in *2005 IEEE Conference on Control Applications*, August 2005.
- [12] S. Li, "A Model-Based Fault Detection and Diagnostic Methodology for Secondary HVAC Systems," Ph.D. dissertation, Drexel University, August 2009.
- [13] N. Tudoroiu, M. Zaheeruddin, E.-R. Tudoroiu, and V. Jeflea, "Fault Detection and Diagnosis (FDD) in Heating Ventilation Air Conditioning Systems (HVAC) Using an Interactive Multiple Model Augmented Unscented Kalman Filter (IMMAUKF)," *IEEE*, May 2008.
- [14] T. I. Salsbury and R. C. Diamond, "Fault detection in HVAC systems using modelbased feedforward control."
- [15] S.-H. Cho, Y.-J. Hong, W.-T. Kim, and M. Zaheeruddin, "Multi-fault detection and diagnosis of HVAC systems: an experimental study," *International Journal of Energy Research*, no. 29, pp. 471–483, 2005.
- [16] C. Lo, P. Chan, Y. Wong, A. Rad, and K. Cheung, "Fuzzy-genetic algorithm for automatic fault detection in HVAC systems," *Applied Soft Computing*, no. 7, pp. 554–560, 2007, www.elsevier.com/locate/asoc.
- [17] L. Stefanutti, *Manuale degli impianti di climatizzazione*. Tecniche Nuove, 2008.

- [18] J. Liang and R. Du, "Thermal Comfort Control Based on Neural Network for HVAC Application," in *2005 IEEE Conference on Control Applications*, August 2005.
- [19] X. Li, Z. Shi, and S. Hu, "A Novel Control Method of a Variable Volume Air Conditioning System for Indoor Thermal Environment," *2010 2nd International Conference on Computer Engineering and Technology*, vol. 2, 2010.
- [20] B. Arguello-Serrano and M. Vélez-R, "Nonlinear Control of a Heating, Ventilating, and Air Conditioning System with Thermal Load Estimation," *IEEE Transaction on Control System Technology*, vol. 7, no. 1, January 1999.
- [21] N. Ben-Aissa, *Heating, Ventilation, and Air Conditioning (HVAC) Controls: Variable Air Volume (VAV) Systems*, VisSim Tutorial Series, 1997.
- [22] R. Moul, "Control Strategies for VAV Air Handling Units - Part 2," *Air Conditioning and Refrigeration Journal*, www.ishrae.in/journals/1999apr/article03.html.
- [23] *Fundamentals*, ASHRAE - Handbook, 2005, www.ashrae.org.
- [24] N. Nassif, S. Kajl, and R. Sabourin, "Modeling and validation of existing VAV system components."
- [25] W. Chen and S. Deng, "Development of a dynamic model for a DX VAV air conditioning system," *Energy Conversion and Management*, no. 47, pp. 2900–2924, 2006, www.elsevier.com/locate/enconman.
- [26] A. Beghi, L. Cecchinato, F. Paggiaro, and M. Rampazzo, "Vavac systems modeling and simulation for fdd applications," in *Proceedings of the International Conference on Control Applications ICCA 2011*, Settembre 2011, pp. 800–805.
- [27] B. R. Sørensen, "Modelling and simulation of a HVAC system," 1999.

- [28] J. Wang, Y. Wang, and H. Shao, "Performance improvement of vav air conditioning control system through diagonal matrix decoupling and lonworks technology," *Energy and Buildings*, vol. 37, no. 9, pp. 911 – 919, 2005. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0378778804003056>
- [29] B. Tashtoush, M. Molhim, and M. Al-Rousan, "Dynamic model of an HVAC system for control analysis," *Energy*, no. 30, pp. 1729–1745, 2005, www.elsevier.com/locate/energy.
- [30] L. Gao, L. Chen, Y. Fan, and H. Ma, "A non-linear control design for power systems," *Automatica*, vol. 28, no. 5, pp. 975 – 979, 1992. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/000510989290150E>
- [31] Q. He and J. Wang, "Fault detection using the k-nearest neighbor rule for semiconductor manufacturing processes," *Semiconductor Manufacturing, IEEE Transactions on*, vol. 20, no. 4, pp. 345 –354, nov. 2007.
- [32] H. Xiong, M. Swamy, and M. Ahmad, "Optimizing the kernel in the empirical feature space," *Neural Networks, IEEE Transactions on*, vol. 16, no. 2, pp. 460 –474, march 2005.
- [33] C. W. Reynolds, "Flocks, herds and schools: A distributed behavioral model," in *Proceedings of the 14th annual conference on Computer graphics and interactive techniques*, ser. SIGGRAPH '87. ACM, 1987, pp. 25–34. [Online]. Available: <http://doi.acm.org/10.1145/37401.37406>
- [34] J. C. Platt, "Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods," in *ADVANCES IN LARGE MARGIN CLASSIFIERS*. MIT Press, 1999, pp. 61–74.
- [35] H.-T. Lin, C.-J. Lin, and R. C. Weng, "A note on plat-t's probabilistic outputs for support vector machines," *Mach. Learn.*, vol. 68, no. 3, pp. 267–276, oct 2007. [Online]. Available: <http://dx.doi.org/10.1007/s10994-007-5018-6>