Università degli Studi di Padova

Dipartimento di Ingegneria Industriale

Corso di Laurea in Ingegneria dell'Energia Elettrica

Attuatore elettromagnetico biventricolare per un cuore artificiale totale: ottimizzazione del progetto mediante tecniche di machine learning

Relatore: **Andriollo Mauro**

Correlatore: **Tortella Andrea**

Studente: **Visonà Nicola**

Matricola: 2055495

Anno Accademico: 2022-2023

In ricordo del professor Andrea Tortella

Indice

Introduzione **3**

1 L'apparato circolatorio sanguigno **5**
1.1 Anatomia 5
1.1.1 Il cuore e il sistema vascolare 6
1.2 Fisiologia 7
1.2.1 La Circolazione cardiaca 7
1.2.2 Il ciclo cardiaco 8
1.3 Il lavoro cardiaco 11

2 Il cuore artificiale **15**
2.1 L'insufficienza cardiaca 15
2.2 Sistemi di assistenza alla circolazione sanguigna 16
2.2.1 Sistemi di assistenza ventricolare (VAD) 17
2.2.2 Il cuore artificiale totale (TAH) 17

3 L'attuatore lineare biventricolare MM-LOA **19**
3.1 Progetto 19
3.1.1 Concepimento dell'attuatore 19
3.1.2 Primo design 20
3.1.3 Secondo design 22
3.2 Funzionamento dell'attuatore 24
3.3 Calcolo dei profili di forza 25
3.4 Indici di prestazione 32
3.5 Vincoli 32

4 Il Metodo di Taguchi **35**
4.1 Revisione della letteratura 35
4.2 Definizione input e output 36
4.2.1 Definizione delle variabili di progetto 37
4.2.2 Definizione degli indici di prestazione 41
4.3 Matrici ortogonali e conduzione degli esperimenti 42
4.4 Risultati 43
4.5 Riduzione della dimensionalità 46
4.5.1 Scelta delle variabili da eliminare 46
4.5.2 Selezione del valore delle variabili eliminate 46

5 Predizione degli indici **51**
5.1 Rifinitura dei range 51
5.2 Il machine learning 52
5.2.1 Reti Neurali (ANN) 54
5.2.2 Gaussian Process Regression (GPR) 56
5.3 Costruzione dei modelli predittivi 57
5.4 Scenario $P_{Loss} = 5$ W 58
5.5 Scenario $P_{Loss} = 10$ W 61
5.6 Conclusioni 62

6 Ottimizzazione multi-obiettivo **65**
6.1 L'ottimizzazione 65
6.2 Descrizione dell'algoritmo multi-obiettivo 67
6.2.1 Algoritmo stocastico 68
6.2.2 Algoritmo deterministico 70
6.3 Validazione dei risultati 71
6.3.1 Scenario 5 W 71
6.3.2 Scenario 10 W 74
6.4 Analisi del Fronte di Pareto 75
6.4.1 Fronte scenario 5 W 75
6.4.2 Fronte scenario 10 W 77
6.4.3 Confronto tra i due scenari 78
6.5 Seconda ottimizzazione 82
6.5.1 Restrizione dei range 82
6.5.2 Costruzione dei modelli predittivi 83
6.5.3 Calcolo del fronte di Pareto 84

7 Analisi dell'attuatore **87**
7.1 Architettura 87
7.1.1 Variabili di progetto 87
7.1.2 Costruzione 89
7.2 Analisi delle prestazioni 92

Conclusioni **95**

Sommario

L'insufficienza cardiaca viene definita la piaga dei paesi industrializzati, rappresentandone la prima causa di morte. Negli stadi più avanzati, una delle migliori terapie è il trapianto di cuore, ma il numero di donatori non aumenta di pari passo con l'incidenza della suddetta sindrome. È quindi crescente l'interesse verso i dispositivi di assistenza cardiaca, in particolare per il cuore artificiale totale (TAH).

Lo scopo del presente lavoro è l'ottimizzazione multi obiettivo dell'architettura di un attuatore lineare elettromeccanico oscillante per l'applicazione in un TAH.

Per conseguire tale obiettivo si affiancano tecniche ben consolidate nel campo ingegneristico, come il metodo di Taguchi e l'analisi agli elementi finiti (FEA), ad altre di più recente sviluppo, quali modelli predittivi basati su machine learning, in particolare su reti neurali profonde, gaussian process regression e support vector machine. Tali modelli ricevendo come input le variabili di progetto dell'attuatore ne predicono le prestazioni. Questo consente di ridurre i tempi di calcolo rispetto alla FEA, accelerando l'individuazione del fronte di Pareto mediante un algoritmo ibrido stocastico/deterministico.

Al termine dell'ottimizzazione, le prestazioni dell'attuatore sono migliorate in termini di forza media durante la fase sistolica di circa il 14% sul ventricolo sinistro, e del 40% su quello destro, ottenendo contestualmente una riduzione della potenza dissipata del 54%.

Introduzione

Il cuore è un organo vitale, in quanto garantisce la sopravvivenza di tutte le cellule del corpo e quindi dell'organismo stesso. Per fare ciò si comporta come una pompa, assicurando la circolazione sanguigna in tutto il corpo, la quale porta sostanze nutritive e ossigeno ai vari distretti corporei, rimuovendo i prodotti di scarto.

Quando il cuore non è più in grado di svolgere correttamente il proprio lavoro si parla di insufficienza cardiaca o scompenso cardiaco. Questa è una sindrome molto diffusa, tanto che è spesso definita come la piaga del nostro tempo, la quale miete più vittime di qualsiasi altra malattia nei paesi industrializzati [1]. Per le forme più gravi di scompenso cardiaco, l'unica possibilità di sopravvivenza a lungo termine per un paziente, che garantisce contestualmente un'adeguata qualità di vita è il trapianto di cuore. Recentemente la prevalenza dell'insufficienza cardiaca è in aumento [2], ma la possibilità di effettuare trapianti di cuore è limitata dalla scarsità di donatori [3]. Per questa ragione è sempre maggiore l'interesse verso dispositivi che possano sostituire il cuore nella sua funzione.

In commercio è possibile trovare molti dispositivi di assistenza alla circolazione, in grado di diminuire il lavoro richiesto all'organo cardiaco, che non richiedono l'asportazione chirurgica del cuore. Quello che si vuole ottenere con il lavoro qui presentato è invece un dispositivo che possa sostituire completamente un cuore danneggiato. Si sta parlando di un cuore artificiale totale (TAH, dall'inglese total artificial heart). Il TAH studiato deve riprodurre fedelmente la fisiologia del sistema cardiovascolare, garantendo un flusso ematico pulsante.

Il lavoro presentato si concentra sull'ottimizzazione dell'attuatore lineare elettromeccanico per tale TAH. Esistono degli attuatori lineari di tipo pulsante utilizzati a questo scopo, ma riproducono la funzionalità di un solo ventricolo; il dispositivo che si vuole studiare li dovrà sostituire entrambi: sarà un dispositivo biventricolare totalmente impiantabile nel torace del paziente, al fine di garantire un'adeguata qualità di vita. In letteratura non si sono trovati riscontri di attuatori elettromeccanici simili a quello studiato, utilizzati per questo scopo.

Il lavoro qui presentato è la prosecuzione di un lavoro già avviato presso il LABME, il Laboratorio di Macchine Elettriche, afferente al Dipartimento di ingegneria industriale (DII) dell'Università degli Studi di Padova, di cui si può trovare riscontro in [4, 5].

Per il calcolo delle prestazioni del dispositivo è utilizzato il software agli elementi finiti Ansys; per l'ottimizzazione vengono adottate varie strategie, quali il metodo di Taguchi per individuare le variabili più influenti sulle prestazioni dell'attuatore e un algoritmo di ottimizzazione multi-obiettivo per la ricerca del

fronte di Pareto e la selezione della configurazione migliore. Tale algoritmo valuta le prestazioni dell'attuatore (le funzioni obiettivo), modificando di volta in volta i parametri di progetto. Ogni valutazione di questo tipo è fatta mediante l'analisi agli elementi finiti (FEM), la quale richiede un onere computazionale elevato. Per ridurre il tempo di calcolo del fronte sono state usate tecniche di machine learning, quali reti neurali, gaussian process regression e support vector marchine, al fine di predire le prestazioni a partire dalla geometria del dispositivo senza effettuare analisi FEM.

Il presente lavoro ha anche uno scopo secondario, ovvero quello di creare un processo per l'ottimizzazione che sia generalizzabile ad altri dispositivi o situazioni.

Dopo una panoramica sul funzionamento del cuore e sulle possibili malattie che possono colpirlo e che possono portare un paziente ad aver bisogno di un cuore sostitutivo (capitoli 1 e 2) si introduce l'attuatore lineare oggetto di studio (capitolo 3). Una volta definiti i parametri su cui si può agire e i vincoli che bisogna rispettare, ha inizio il processo di ottimizzazione con il metodo di Taguchi, tramite cui è possibile viene ridotta la dimensionalità del problema (capitolo 4). Vengono poi costruiti i modelli predittivi, in modo da sostituire l'analisi agli elementi finiti (capitolo 5) nel calcolo delle prestazioni. Tali modelli vengono poi utilizzati nell'algoritmo per la ricerca del fronte di Pareto (capitolo 6), con un onere computazionale ridotto. Una volta trovato il fronte, viene scelta la configurazione ritenuta migliore. Si esplora poi più attentamente lo spazio delle decisioni attorno a tale configurazione in modo da determinare la più conveniente. L'architettura ottimizzata viene poi analizzata in dettaglio (capitolo 7).

La procedura seguita è generalizzabile e può essere applicata in altri ambiti, anche completamente diversi. Con tale metodo si è arrivati ad un significativo miglioramento delle prestazioni dell'attuatore rispetto alla configurazione precedente, presentata in [5].

Capitolo 1

L'apparato circolatorio sanguigno

Questo capitolo ha l'obiettivo di esporre brevemente l'anatomia e la fisiologia dell'apparato circolatorio sanguigno con una particolare attenzione sul cuore[1]. Nella prima parte del capitolo verrà illustrata l'anatomia del sistema cardiovascolare, mentre nella seconda parte verrà studiata la circolazione cardiaca e il relativo ciclo, in modo da evidenziare i compiti a cui dovrà assolvere l'attuatore lineare studiato.

1.1 Anatomia

L'apparato circolatorio sanguigno è l'insieme di organi e vasi che garantiscono la corretta circolazione del sangue in tutto l'organismo. Tale apparato è un sistema chiuso composto da un organo motore, il cuore, che svolge la funzione di pompa, e i vasi sanguigni, suddivisibili in arterie, capillari e vene che costituiscono i condotti all'interno dei quali viene trasportato il sangue.

Le funzioni svolte da questo apparato possono essere così schematizzate[6]:

- trasporto di ossigeno e sostanze nutritizie e loro distribuzione a tutte le parti dell'organismo;
- allontanamento dei prodotti del catabolismo cellulare, in particolare anidride carbonica, dalla loro sede di produzione;
- contributo al mantenimento di una temperatura corporea costante (termoregolazione);
- regolazione dell'omeostasi dei fluidi corporei;
- impegno nei processi immunitari mediante il trasporto di cellule dotate di attività fagocitaria, di cellule immunocompetenti e di anticorpi;

[1] Per una trattazione più completa dell'anatomia e fisiologia dell'apparato sanguigno cardiovascolare si rimanda a testi specialistici, come [6, 7], da cui sono state ricavate varie informazioni per la scrittura di questo capitolo.

1.1.1 Il cuore e il sistema vascolare

La funzione del cuore è quella di fornire l'energia necessaria al sangue per farlo fluire all'interno dei vasi sanguigni in una precisa direzione. Ciò viene reso possibile da contrazioni ritmiche di tale organo e da un sistema valvolare.

Il cuore si divide in due parti, non comunicanti tra di loro: parte destra e parte sinistra, chiamate anche cuore destro e cuore sinistro. Questa divisione è necessaria per evitare la mescolanza tra sangue venoso ricco di anidride carbonica e sostanze di scarto con il sangue arterioso ricco di ossigeno e sostanze nutritizie. Ciascuna di queste due sezioni è ulteriormente divisa in due, atrio e ventricolo, comunicanti tra loro mediante l'orifizio atrioventricolare, provvisto di un apparato valvolare (valvola atrioventricolare destra o tricuspide e valvola atrioventicolare sinistra anche chiamata valvola bicuspide o mitrale) che permette il flusso del sangue solamente dagli atri verso i ventricoli. Negli atri giungono le vene e dai ventricoli si dipartono le arterie alla cui origine si trovano le valvole arteriose o semilunari che permettono il passaggio del sangue dai ventricoli alle arterie e non viceversa. La valvola aortica collega il ventricolo sinistro con l'aorta e la valvola polmonare collega il ventricolo destro con l'arteria polmonare. Ogni camera ventricolare è provvista di due valvole che, aprendosi e chiudendosi possono metterla in comunicazione con l'atrio e quindi con le vene, con le arterie oppure possono isolarla completamente, quando entrambe le valvole sono chiuse. Uno schema dell'anatomia del cuore è riportato in figura 1.1 dove è ben visibile la separazione tra il cuore destro e sinistro e la posizione delle valvole.

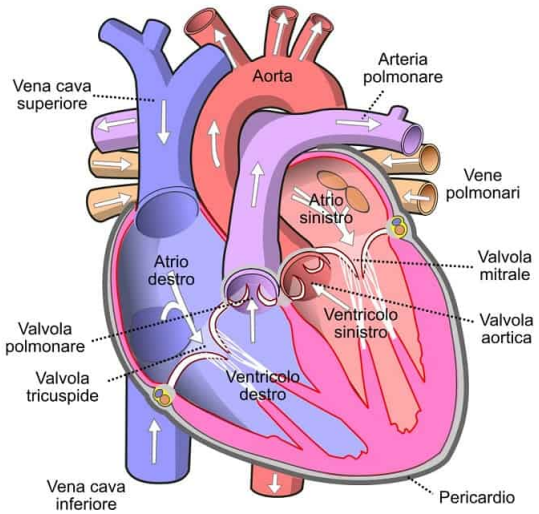


Figura 1.1: Anatomia del cuore

Il cuore, come detto, spinge il sangue in una precisa direzione, in particolare lo preleva dalle vene che si trovano a bassa pressione e lo immette nelle arterie, a pressione più elevata. Queste, a mano a mano che si allontanano dal cuore, si ramificano progressivamente aumentando di numero e riducendosi di diametro, arrivando alle arteriole. Tutte le arterie, indipendentemente dal calibro sono impermeabili e svolgono solamente una funzione di trasporto del sangue in esse

contenuto. Dal segmento terminale delle arteriole traggono origine i capillari, che, a causa della sottigliezza della loro parete sono permeabili e consentono il passaggio dell'acqua e delle sostanze in essa disciolte dal sangue ai tessuti circostanti e viceversa[7], modificando quindi la composizione fisico-chimica del sangue. I capillari confluiscono poi nelle venule che raccolgono il sangue e, riunendosi fra loro, formano le vene il cui calibro diventa via via sempre maggiore fino a raggiungere il cuore. Le vene, così come le arterie, indipendentemente dal loro diametro, possiedono una parete impermeabile, assolvendo solamente alla funzione di trasporto del sangue verso il cuore.

1.2 Fisiologia

Il processo di respirazione cellulare che avviene nelle cellule dell'organismo umano prevede l'utilizzo di ossigeno (introdotto nell'organismo tramite la respirazione polmonare) e glucosio (introdotto tramite l'alimentazione) come reagenti per la produzione di ATP, un composto ad alta energia richiesto dalla quasi totalità delle reazioni metaboliche, necessaria quindi al sostentamento dell'organismo. Nella respirazione cellulare vengono prodotti anche anidride carbonica e acqua come prodotti di scarto. È necessario portare l'ossigeno a tutte le cellule corporee ed eliminare l'anidride carbonica: questo è reso possibile dalla cooperazione tra il sistema circolatorio, che tramite il sangue trasporta ossigeno e anidride carbonica da e verso i polmoni, e il sistema respiratorio che ha il compito di espellere l'anidride carbonica e introdurre l'ossigeno nell'organismo.

1.2.1 La Circolazione cardiaca

Nell'uomo la circolazione cardiaca è doppia e completa: si distinguono un grande circolo o circolazione sistemica, il cui compito è quello di portare a tutti i segmenti corporei ossigeno e sostanze nutritive e un piccolo circolo o circolazione polmonare necessaria per ristabilire nel sangue un'adeguata concentrazione di ossigeno ed espellere l'anidride carbonica. Entrambe le circolazioni hanno origine nei ventricoli del cuore e finiscono negli atri. Uno schema di tale circolazione è riportato in figura 1.2, in cui è distinguibile il sangue venoso ricco di anidride carbonica (in blu) dal sangue arterioso ricco di ossigeno (in rosso).

La circolazione generale ha inizio dal ventricolo sinistro, mediante l'arteria aorta la quale si suddivide in arterie sempre più piccole fino ad arrivare alle arteriole e successivamente ai capillari, i quali raggiungono tutti i segmenti corporei. Nei capillari il sangue cede ossigeno e sostanze nutritizie alle cellule dei tessuti irrorati e si carica di anidride carbonica e cataboliti. Qui il sangue modifica la sua composizione fisico-chimica passando da sangue arterioso, di colore scarlatto a sangue venoso, di colore bluastro. Dalle reti capillari, tramite il sistema venoso, il sangue viene raccolto e ritorna al cuore, più precisamente nell'atrio destro, dove sboccano la vena cava superiore e quella inferiore; da qui il sangue passa al ventricolo destro dove comincia il piccolo circolo.

La circolazione polmonare ha inizio nel ventricolo destro da cui si diparte il tronco polmonare che, dividendosi nelle due arterie polmonari trasporta sangue

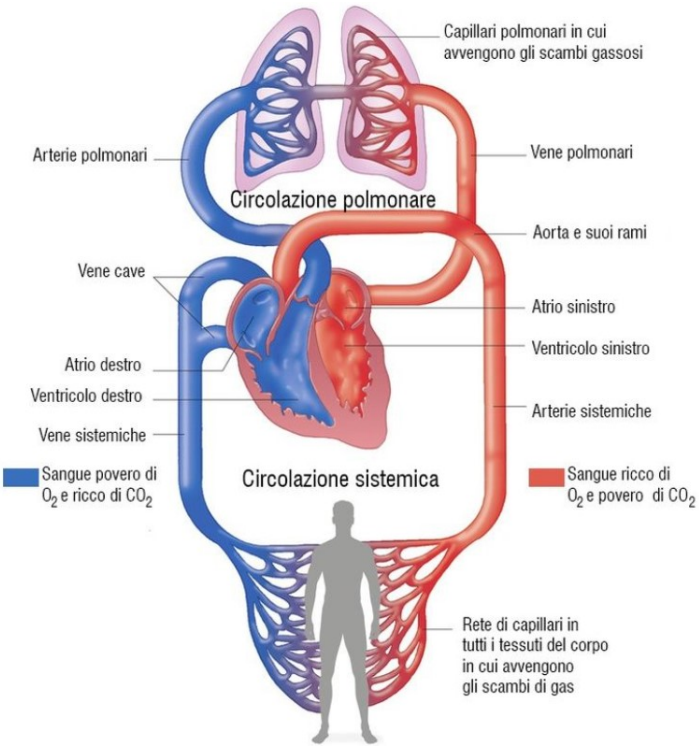


Figura 1.2: Schema semplificato della circolazione cardiaca.

venoso ai polmoni. Le arterie polmonari hanno fine in una rete di capillari addossata agli alveoli polmonari, dove il sangue cede anidride carbonica, che viene espulsa tramite espirazione, e si arricchisce di ossigeno, introdotto per inspirazione, diventando così sangue arterioso. Tramite le vene polmonari il sangue ritorna al cuore, più precisamente all'atrio sinistro dove confluisce poi nel ventricolo sinistro da cui ha inizio il grande circolo.

Dal punto di vista funzionale il cuore è quindi costituito da due pompe separate, cuore destro e cuore sinistro, poste in serie, la cui gittata deve essere identica. Uno scompenso, anche minimo, tra le due gittate, può portare nel tempo a problematiche legate ad alterazioni di volumi e pressioni nei vari distretti del corpo. Il cuore destro pompa il sangue venoso proveniente dalla vena cava superiore e inferiore nell'arteria polmonare, alimentando la piccola circolazione, mentre il cuore sinistro pompa il sangue proveniente dalle vene polmonari nell'arteria aorta, alimentando la grande circolazione.

1.2.2 Il ciclo cardiaco

Atri e ventricoli alternano continuamente fasi di rilassamento, in cui si riempiono di sangue a fasi di contrazione, in cui lo espellono, inviandolo alla sezione successiva, sia essa il ventricolo o l'arteria. Questo processo ritmico è denominato ciclo cardiaco, che si compone di due fasi:

- la fase di sistole che, per definizione, inizia con la chiusura della valvola mitrale e termina con la chiusura della valvola aortica e dura all'incirca 1/3

del ciclo. Questa è la fase in cui il ventricolo si contrae e il sangue in esso contenuto viene immesso nelle arterie;

- la fase di diastole che copre i restanti 2/3 del ciclo, in cui il ventricolo si rilassa e si riempie di sangue che verrà poi espulso nella fase di sistole.

La sistole può essere suddivisa ulteriormente in due fasi: contrazione isovolumetrica ed eiezione; analogamente la diastole può essere divisa in quattro fasi: rilasciamento isovolumetrico, riempimento rapido diastolico, diastasi e sistole atriale[8]. Nel seguito queste fasi verranno analizzate in dettaglio; per una migliore comprensione si faccia riferimento alla figura 1.3 in cui è visualizzato il diagramma di Wiggers per il cuore sinistro. Tale diagramma riporta i valori di pressione delle sezioni del cuore sinistro nel tempo. Il flusso sanguigno tra due sezioni comunicanti è proporzionale al gradiente pressorio tra le due sezioni, fluendo da dove la pressione è più elevata a dove è più bassa. Nel diagramma è anche riportato il volume sanguigno contenuto nel ventricolo sinistro, il grafico che viene riportato nell'elettrocardiogramma e i toni cardiaci (fonogramma). Per semplicità le considerazioni che seguiranno prendono come esempio il cuore sinistro, ma possono essere estese al cuore destro, seppur con alcune differenze legate principalmente ai valori pressori. I valori indicati fanno riferimento ad una persona adulta sana, in condizioni di riposo, con frequenza cardiaca di ca. 70 bpm. Per convenzione

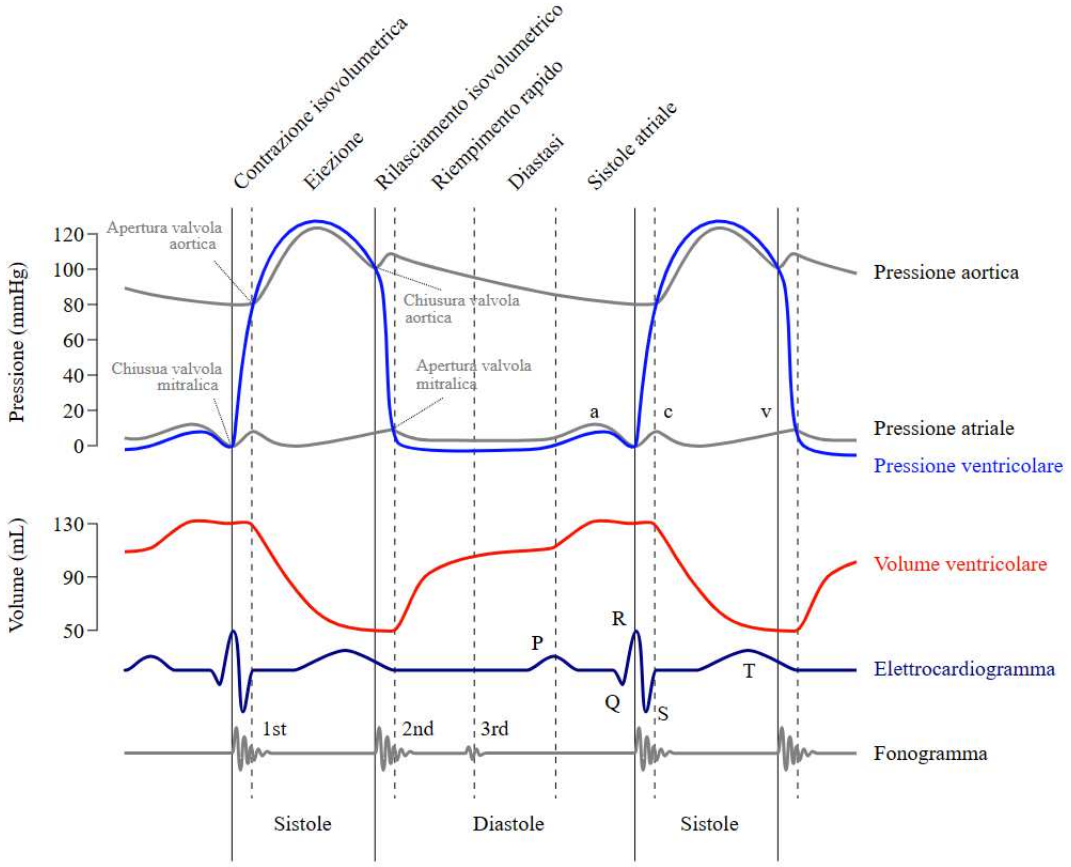


Figura 1.3: Diagramma di Wiggers per il cuore sinistro [9].

il ciclo cardiaco ha inizio con la prima fase di sistole e le fasi che lo compongono sono:

Contrazione isovolumetrica: il ventricolo risulta riempito di sangue e comincia a contrarsi innalzando la pressione al suo interno; l'atrio invece è in fase di rilassamento. Quando la pressione ventricolare supera la pressione atriale si ha la chiusura della valvola mitrale: in questo istante ha inizio la fase di contrazione isovolumetrica. La pressione ventricolare continua ad aumentare fino al raggiungimento della pressione aortica; in tale istante la valvola semilunare si apre e ha inizio la fase di eiezione. Durante la fase di contrazione isolvolumetrica sono chiuse sia la valvola mitrale che quella aortica, quindi il volume di sangue resta costante e pari al volume telediastolico[2] di ca. 130 ml.

Eiezione: nel momento in cui la pressione ventricolare supera quella aortica, si ha l'apertura della valvola semilunare e inizia questa fase che è caratterizzata da un ulteriore aumento di pressione ventricolare e dall'immissione di sangue nell'aorta, causando quindi una diminuzione del volume ventricolare. La pressione ventricolare arriva ad un picco di ca. 120 mmHg. Si noti come non venga espulso tutto il sangue presente nel ventricolo, ma solamente una frazione, corrispondente a ca. 70-80 ml. Il volume di sangue espulso in questa fase prende il nome di gittata sistolica. Quando il ventricolo finisce di contrarsi, la pressione ventricolare inizia a diminuire e nel momento in cui torna a essere inferiore rispetto alla pressione aortica si ha la chiusura della valvola semilunare e la fine della fase di sistole. Si può distinguere la fase di eiezione rapida, fino al raggiungimento del picco pressorio, in cui si ha l'espulsione del 70% della gittata sistolica e la fase di eiezione lenta in cui si ha l'inizio del rilassamento delle fibre ventricolari che porta alla diminuzione della pressione ventricolare.

Rilasciamento isovolumetrico: questa è la prima fase di diastole e si trova tra la chiusura della valvola aortica e l'apertura della valvola mitrale che avviene quando la pressione ventricolare risulta inferiore rispetto a quella atriale. Durante questa fase il ventricolo si sta rilassando e la pressione al suo interno cala drasticamente, ma essendo chiuse entrambe le valvole, è isolato e il volume al suo interno rimane quindi invariato.

Riempimento rapido: questa fase ha inizio con l'apertura della valvola mitrale. Si ha un rapido riempimento del ventricolo a causa del gradiente pressorio tra atrio e ventricolo. Il riempimento rapido garantisce il riempimento ventricolare anche a frequenze cardiache elevate, associate ad una diminuzione della durata della diastole. Atri e ventricoli sono entrambi rilassati.

Diastasi: a seguito del riempimento del ventricolo, la pressione ventricolare e atriale diventano molto simili e il flusso ematico attraverso la valvola mitrale diventa quasi nullo.

[2] Il volume telediastolico è il volume di sangue presente nel ventricolo alla fine della diastole.

Sistole atriale: in questa fase si ha la contrazione dell'atrio. Ciò porta ad un leggero aumento della pressione atriale e quindi al completo riempimento del ventricolo a causa del gradiente pressorio tra atrio e ventricolo.

In figura 1.4 viene mostrato il diagramma di Wiggers per il cuore destro. Se comparato con il rispettivo diagramma riferito al cuore sinistro, in figura 1.3, si nota una somiglianza nell'andamento della pressione ventricolare, con un picco pressorio notevolmente inferiore, dell'ordine di 30 mmHg. Questo è dovuto al fatto che il cuore destro alimenta la circolazione polmonare che, essendo più piccola della circolazione sistemica, richiede minore energia. Sono inoltre presenti delle leggerissime differenze temporali negli eventi di cuore sinistro e destro: l'inizio della sistole dell'atrio destro precede di poco quella del sinistro, mentre l'inizio della contrazione isovolumetrica del ventricolo destro segue quella del ventricolo sinistro. I valori di pressione indicativi dei singoli distretti durante le varie fasi

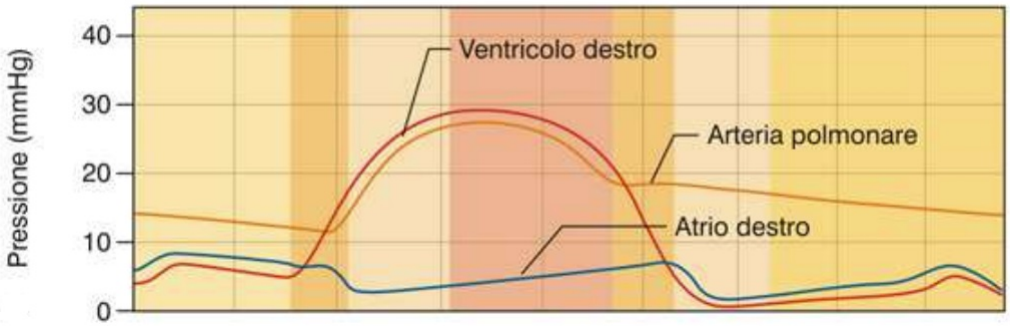


Figura 1.4: Il diagramma di Wiggers per il cuore destro[7]

.

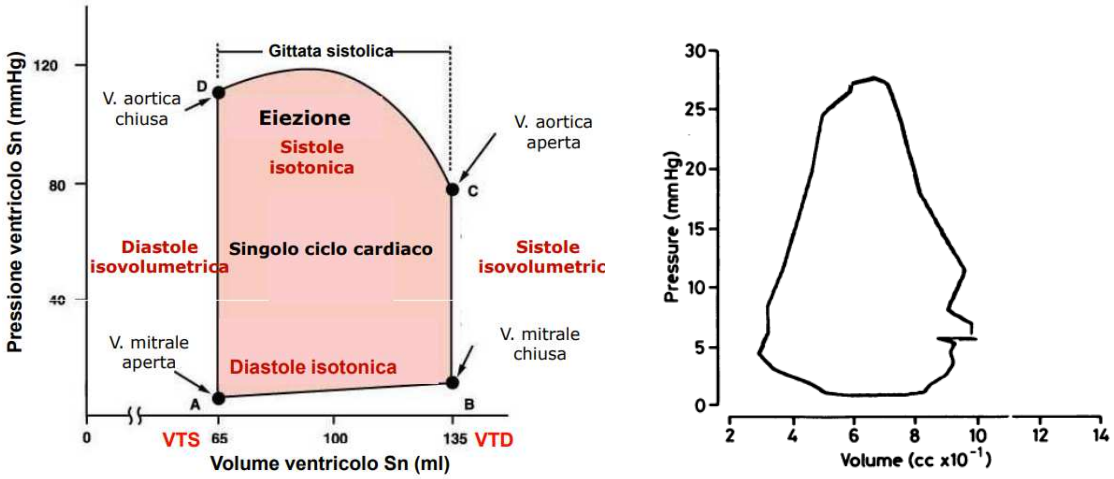
sono riportati in tabella 1.1. Questi valori, così come tutti quelli presenti in questo capitolo, si riferiscono ad una persona adulta media in buone condizioni di salute e potrebbero variare da individuo a individuo.

1.3 Il lavoro cardiaco

A partire dalle informazioni presenti in questo capitolo, è possibile calcolare il lavoro effettuato del cuore durante un ciclo. In figura 1.5 è riportato il ciclo cardiaco nel piano pressione-volume (PV) per entrambi i ventricoli; nel seguito si farà riferimento al ventricolo sinistro, il cui grafico è riportato in figura 1.5a. Si possono riconoscere le varie fasi del ciclo descritte nel paragrafo precedente. In particolare la fase di sistole è quella che va dal punto B (chiusura valvola mitrale) al punto D (chiusura valvola aortica); la fase di contrazione isovolumetrica termina nel punto C, in cui si ha l'apertura della valvola aortica, e comincia la fase di eiezione: la pressione ventricolare sale ancora e diminuisce il volume del ventricolo. La fase di diastole copre il resto del grafico, iniziando nel punto B

Tabella 1.1: Pressioni nei vari distretti dell'apparato cardiovascolare nelle varie fasi del ciclo cardiaco. La notazione [x,y] indica un range di valori, mentre la notazione z-p è indica il valore di picco (p) e il successivo valore minimo (z) di pressione per la fase considerata.

Distretto dell'apparato cardiovascolare	**Pressione [mmHg]**
Atrio sinistro (media)	[6,12]
Ventricolo sinistro sistolica	80-120
Ventricolo sinistro telediastolico	9
Aorta	80-120
Pressione aortica media	100
Atrio destro (media)	[3,8]
Ventricolo destro sistolica	8-30
Ventricolo destro telediastolico	4
Arteria polmonare	8-25
Pressione media arteria polmonare	20



(a) Ventricolo sinistro **(b)** Ventricolo destro

Figura 1.5: Ciclo cardiaco dei due ventricoli del cuore riportato nel piano pressione volume. Sono visibili i grafici del ventricolo sinistro **(a)** e destro **(b)**. Tratti da [10, 11]

e concludendosi nel punto D. Il tratto DA corrisponde alla fase di rilasciamento isovolumetrico, in cui la pressione del ventricolo scende, ma il volume resta invariato. Nel momento in cui si ha l'apertura della valvola mitrale comincia la fase di riempimento del ventricolo, corrispondente al tratto AB, caratterizzata dalle 3 fasi analizzate nel paragrafo precedente. Un discorso analogo si può fare riferendosi al ventricolo destro (figura 1.5b), ma con valori di pressione diversi.

Il lavoro che compie ogni ventricolo durante un ciclo cardiaco può essere suddiviso in due componenti:

1. una quota di lavoro di pressione, individuato dall'area racchiusa all'interno del grafico pressione/volume del ventricolo, che si calcola con la formula

$$L_p = \oint_{ABCD} p \cdot dV \tag{1.1}$$

   Indicando con $p_{m,CD}$ e $p_{m,AB}$ rispettivamente la pressione media ventricolare nella fase di eiezione e la pressione media ventricolare nella fase diastolica di riempimento e con GS la gittata sistolica (circa 70 ml) l'equazione 1.1 può essere sviluppata nel seguente modo:

$$\begin{aligned} L_p &= \int_A^B p \cdot dV + \int_B^C p \cdot dV + \int_C^D p \cdot dV + \int_D^A p \cdot dV \\ &= p_{m,AB} \cdot GS + 0 - p_{m,CD} \cdot GS + 0 \\ &= -(p_{m,CD} - p_{m,AB}) \cdot GS \end{aligned}$$

   Questo lavoro risulta essere negativo, in quanto è un lavoro svolto sul ventricolo da un organo esterno, il cuore. Il lavoro che viene svolto dal cuore è quindi:

$$L_p = (p_{m,CD} - p_{m,AB}) \cdot GS$$

2. Una quota di lavoro cinetico compiuto per imprimere alla massa m di sangue espulso una certa velocità v. Sapendo che $m = \rho_s \cdot GS$, dove ρ_s è la densità del sangue ($\rho_s = 1.062$ g/ml) e GS è la gittata sistolica, il lavoro cinetico si calcola con la formula:

$$L_c = \frac{1}{2} \cdot \rho_s \cdot GS \cdot v^2 \tag{1.2}$$

Sommando questi due contributi, il lavoro che svolto dal ventricolo k durante un ciclo cardiaco vale

$$\begin{aligned} L_k &= L_p + L_c \\ &= (p_{m,CD} - p_{m,AB}) \cdot GS + \frac{1}{2} \cdot \rho_s \cdot GS \cdot v^2 \\ &= GS \cdot \left[(p_{m,CD} - p_{m,AB}) \cdot + \frac{1}{2} \cdot \rho_s \cdot v^2 \right] \end{aligned} \tag{1.3}$$

Inserendo i valori presentati in tabella 1.2, riferiti ad una persona sana, in condizione di riposo, nella formula 1.3, con le opportune conversioni, il lavoro compiuto dai due ventricoli durante ogni ciclo vale:

$$L_{LV} \cong 900 \text{ mJ}$$
$$L_{RV} \cong 200 \text{ mJ}$$

Tabella 1.2: Valori utilizzati per il calcolo del lavoro cardiaco durante un ciclo.

Grandezza	**Valore**
v	0.5 m/s
$p_{m,CD,LV}$	100 mmHg
$p_{m,AB,LV}$	5 mmHg
$p_{m,CD,RV}$	25 mmHg
$p_{m,AB,RV}$	4 mmHg
ρ_s	1.062 g/ml
GS	70 ml

Capitolo 2

Il cuore artificiale

2.1 L'insufficienza cardiaca

Il cuore, come visto nel capitolo 1, è un muscolo il cui compito è quello di garantire l'adeguato afflusso di sangue a tutti i distretti corporei. Per fare ciò agisce come una pompa, prelevando il sangue da una zona a bassa pressione (le vene) e spingendolo in una zona a pressione più elevata (le arterie). Quando il cuore non è più in grado di assolvere al proprio compito si parla di insufficienza cardiaca o scompenso cardiaco.

Secondo le linee guida ESC, l'insufficienza cardiaca non è definita come una singola patologia, quanto piuttosto come una sindrome clinica composta da diversi sintomi principali (dispnea, gonfiore alle caviglie e affaticamento) che possono essere accompagnati da vari segnali (pressione venosa giugulare elevata, rantoli polmonari ed edema periferico). Essa viene causata da un'anomalia strutturale e/o funzionale del cuore che comporta pressioni intracardiache elevate e/o una gittata cardiaca inadeguata a riposo e/o durante l'esercizio [12]. Nella diagnosi di insufficienza cardiaca è importante identificare la causa della disfunzione cardiaca in modo da poter scegliere il percorso terapeutico migliore per il singolo paziente.

L'insufficienza cardiaca è dovuta a disfunzioni nella fase sistolica, diastolica o in entrambe le fasi. Anche patologie delle valvole cardiache, del pericardio, dell'endocardio o anormalità nel ritmo cardiaco possono contribuire alla patologia. È possibile avere insufficienza cardiaca che coinvolge il cuore destro, sinistro oppure entrambi. Esistono dei meccanismi fisiologici, quali il meccanismo di Frank-Starling[1], l'attivazione del sistema neuroumorale o l'ipertrofia miocardica, che riescono a rispondere all'aumentato lavoro cardiaco dovuto all'inizio della sindrome. Questi meccanismi possono inizialmente mantenere un'adeguata pressione sanguigna e perfusione degli organi, per cui il paziente sarà affetto da insufficienza cardiaca senza saperlo, ma a lungo andare si incorrerà in un progressivo deterioramento della funzione contrattile del cuore [13].

L'insufficienza cardiaca è una sindrome che ha inevitabilmente ripercussioni su tutto l'organismo con gravi conseguenze su vari apparati, quali quello locomotore,

[1]Questo è un meccanismo grazie al quale, all'aumentare del volume di riempimento del cuore, aumenta la forza di contrazione delle fibre miocardiche. Grazie a ciò il cuore è in grado di espellere questo volume più elevato di sangue.

respiratorio ed endocrino a causa della minore gittata sanguigna che porta meno ossigeno e sostanze nutritizie ai tessuti. Come conseguenza si ha la progressiva morte delle cellule dei tessuti periferici.

Con il termine scompenso cardiaco si fa riferimento ad una condizione di insufficienza cardiaca più grave ed avanzata. Una classificazione della gravità dello scompenso cardiaco è quella data dalla New York Heart Association (NYHA) che analizzando i sintomi dell'intolleranza all'esercizio fisico, individua quattro classi [14]:

1. Nessun sintomo
2. Sintomi con livelli di attività moderati o intensi
3. Sintomi con attività lieve
4. Sintomi a riposo

Lo scompenso cardiaco, soprattutto se in fase avanzata, è una patologia a prognosi negativa. Questo significa che è una condizione irreversibile, ovvero un cuore scompensato è un cuore che non può più ritornare a funzionare come prima. La prognosi di uno scompenso cardiaco in fase avanzata è infausta, con una mortalità a un anno tra il 25% e il 75%, e si manifesta con un lento e inesorabile declino dell'organo cardiaco costellato da eventi acuti [12].

Esistono però delle terapie farmacologiche che possono ridurre la sintomatologia e rallentare la progressione della malattia. Tuttavia l'unica soluzione per rendere la prognosi positiva è quella di ricorrere al trapianto di cuore, che rappresenta il gold standard[2], con una sopravvivenza a un anno dall'intervento del 90% e una sopravvivenza media di 12,5 anni [12]. Questa è un'operazione molto delicata che consiste nella sostituzione del cuore danneggiato con un cuore sano proveniente da una persona morta di recente. Purtroppo non sempre è possibile ricorrere a questa terapia a causa del basso numero di donatori e dell'elevata incidenza della malattia. Per esempio in Italia nel 2021 il numero di iscritti in lista d'attesa per un trapianto di cuore era di 1041 persone, e sono stati effettuati 252 trapianti di cuore. Il tempo medio di attesa (TMA) è di 3.7 anni in età adulta con un tasso di mortalità in lista d'attesa del $6,9$ %. I dati sono in peggioramento rispetto a quelli del 2020, che ha visto 992 persone in lista d'attesa con un tasso di mortalità del 3.9 % [15]. A livello mondiale si stima che 26 milioni di persone convivano con questa problematica [16] e che 1 persona su 5 nel corso della sua vita possa svilupparla [17]. Negli Stati Uniti d'America circa 6 milioni di persone l'anno soffre di insufficienza cardiaca e si prevede che questa cifra arriverà a quota 8 milioni nel 2030 [18].

2.2 Sistemi di assistenza alla circolazione sanguigna

Data la sempre crescente richiesta di trapianti di cuore e la diminuzione dei donatori, come soluzione ai più gravi disturbi di cuore sono studiati dei dispositivi che

[2]La migliore delle terapie possibili

possano aiutare o sostituire il cuore danneggiato nella sua funzione di pompaggio. Questi dispositivi possono avere due finalità diverse:

- ponte verso il trapianto (BTT, dall'inglese "Bridge to transplantation"), ovvero devono aiutare o sostituire il cuore nella sua funzione per un periodo di tempo limitato, fino a quando non sarà disponibile un cuore trapiantabile per la persona affetta da scompenso cardiaco. In questo modo si permette al paziente di condurre una vita relativamente normale, fuori dall'ospedale;
- terapia definitiva (DT, dall'inglese "Destination therapy"), ovvero devono sostituire o aiutare il cuore nella sua funzione per il resto della vita del paziente e sono utilizzati ad esempio in persone in cui non è possibile effettuare un trapianto [4].

2.2.1 Sistemi di assistenza ventricolare (VAD)

Quando si parla di dispositivi in grado di aiutare il cuore nella sua funzione di pompa si fa riferimento ai dispositivi di assistenza ventricolare (VAD). Questi sono dispositivi che non sostituiscono completamente il cuore, ma vanno a supportarlo, diminuendo il lavoro che esso dovrà compiere, evitandone l'asportazione. Vengono utilizzati in pazienti con scompenso cardiaco grave o avanzato, ma anche in pazienti che si stanno riprendendo da un infarto del miocardio, per diminuire il lavoro richiesto al cuore. A seconda della parte del cuore che vanno a supportare possiamo trovare LVAD (i più comuni, installati nel cuore sinistro), RVAD (installati nel cuore destro) e bi-VAD (assistono entrambi i ventricoli).

Nel corso degli anni sono stati sviluppati diversi dispositivi, che possono produrre un flusso continuo o pulsatile. Vari studi hanno dimostrato come l'utilizzo di questi sistemi aumenti la sopravvivenza dei pazienti in lista d'attesa per un trapianto di cuore rispetto alle terapie tradizionali [19, 20]

2.2.2 Il cuore artificiale totale (TAH)

Quando si parla di un dispositivo che deve rimpiazzare completamente il cuore si parla di un cuore artificiale totale (TAH) che viene utilizzato principalmente quando entrambe le parti del cuore sono gravemente compromesse. Il dispositivo deve essere impiantato all'interno o all'esterno del paziente in base al tipo di tecnologia, rispettando dei vincoli fisiologici e meccanici molto stringenti. Tra i più importanti troviamo: [4]

- l'adempimento dei compiti cardiaci, anche in condizioni diverse da quelle di riposo;
- l'utilizzo di materiali biocompatibili e durabili (è richiesta una durata di almeno 5 anni);
- ridotto consumo energetico e produzione di calore;
- presenza di una fonte energetica di back-up nel caso di guasto dell'unità principale;

- evitare, per quanto possibile, equipaggiamenti troppo pesanti e ingombranti per offrire al paziente un'adeguata qualità di vita.

Il rispetto di tutti questi vincoli è una grande sfida e, allo stato attuale, solamente due dispositivi hanno ottenuto l'approvazione medica dalla Food and Drug Administration (FDA). È necessario notare che gli attuali dispositivi di assistenza alla circolazione sono progettati per una persona adulta; sarà necessario lo sviluppo di dispositivi più piccoli e leggeri anche per pazienti in età pediatrica, in quanto disturbi del cuore possono comparire anche in questo periodo della vita [4].

Oltre al TAH è necessario prevedere una serie di dispositivi ausiliari per il controllo e l'alimentazione del cuore artificiale. Una possibile architettura del cuore artificiale totale è visibile in figura 2.1. Il TAH (A) è alimentato in maniera wireless (sono visibili la bobina trasmittente (F) e ricevente (G)) da una batteria esterna (B). È presente l'interfaccia utente (D) che raccoglie le informazioni sul sistema e ne regola i parametri, la batteria ausiliaria (C) che serve come fonte energetica di back-up, il microprocessore (H) in cui avviene il controllo in catena chiusa dell'attuatore del TAH e la camera di compliance (E), dove il sangue si accumula per evitare sovrapressioni nei ventricoli, facilitando allo stesso tempo l'azione di pompaggio del TAH [4].

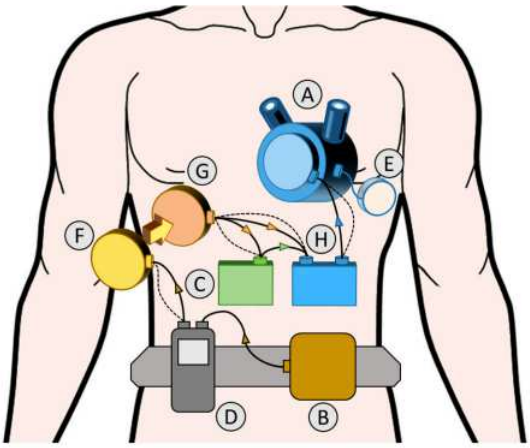


Figura 2.1: Schema di un TAH e dei suoi sistemi ausiliari [4]. Le linee continue indicano un trasferimento di potenza e quelle tratteggiate di segnale. (A): TAH, (B): batteria esterna, (C): batteria interna, (D): interfaccia utente, (E): camera di compliance, (F): bobina trasmittente, (G): bobina ricevente

Capitolo 3

L'attuatore lineare biventricolare MM-LOA

Un attuatore lineare elettromeccanico è un dispositivo in grado di trasformare energia elettrica in energia meccanica, con un moto traslatorio, lungo una linea retta. Questo tipo di dispositivo trova largo impiego nell'industria, in settori come quello dell'automotive e dell'automazione, ovunque sia richiesto un moto lineare. L'applicazione per cui è pensato l'attuatore in esame richiede tuttavia il soddisfacimento di vincoli molto stringenti, rendendo necessaria una progettazione ad hoc.

3.1 Progetto

3.1.1 Concepimento dell'attuatore

L'attuatore lineare analizzato nasce per essere il driver di un cuore artificiale totale, il quale possa riprodurre fedelmente la fisiologia del cuore. Il TAH in questione deve essere in grado di replicare la funzione di pompaggio di entrambi i ventricoli, in maniera sincrona, esattamente come accade in un cuore normale, come visto nel capitolo 1. Il TAH poi deve essere totalmente impiantabile nel torace, in modo tale da garantire un'adeguata qualità della vita del paziente, il che implica dimensioni ridotte. Questo dispositivo deve essere affidabile e durabile nel tempo affinché possa essere impiegato sia come BTT che come DT.

Il TAH si costituisce di due camere ventricolari in materiale biocompatibile, come quello presentato in [21], che hanno il compito di raccogliere il sangue in arrivo dalle vene durante la fase diastolica. Una volta che le camere sono riempite, due equipaggi mobili, detti mover, esercitano una forza su di esse comprimendole, innalzando così la pressione al loro interno; quando la pressione ventricolare supera quella dell'arteria collegata alla camere, il sangue viene spinto in tali arterie. Per garantire l'unidirezionalità del flusso ematico vengono inserite quattro valvole: due vengono interposte tra le vene e la camere ventricolari e le altre due tra le camere ventricolari e le arterie. Le prime fungono da valvole interatriali (v. mitrale e v. tricuspide) evitando il ritorno del sangue nelle vene, una volta che ha inizio la fase sistolica, mentre le seconde fungono da valvole semilunari

(v. aortica e v. polmonare), impedendo il passaggio di sangue dalle arterie alle camere ventricolari dopo la fase di eiezione. Uno schema del TAH appena descritto è visualizzabile in figura 3.1. Questa architettura è in grado di riprodurre il meccanismo pulsante del cuore ed è inoltre possibile avere una separazione tra l'attuatore e il sangue, evitando così dei vincoli troppo stringenti sui materiali utilizzabili per l'attuatore. Questo non è possibile con una pompa rotativa.

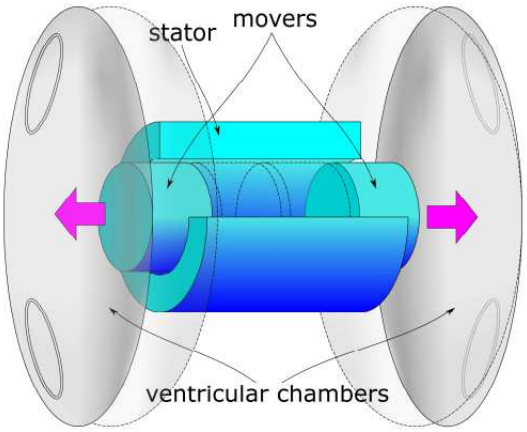


Figura 3.1: Schema di funzionamento del cuore artificiale totale[5]. Due organi mobili, detti mover, comprimono delle camere ventricolari di materiale biocompatibile. Grazie a delle valvole, viene prodotto un flusso pulsante unidirezionale di sangue, analogamente a ciò che fa il cuore.

Ci si è orientati verso un attuatore di tipo elettromeccanico, invece che pneumatico in quanto i primi sono più compatti, efficienti ed affidabili, con una produzione di calore limitata e sono facilmente controllabili, consentendo di adattare il carico di lavoro ai bisogni del paziente. Tra i vari attuatori elettromeccanici lineari è stata scelta la configurazione tubulare a magneti permanenti (PMs) in quanto considerata la più efficiente. Questi attuatori si dividono in due principali categorie: [5]

a) con bobine mobili senza nucleo che scorrono lungo dei magneti permanenti fissi montati su un meteriale ferromagnetico (MC-LOA)

b) con magneti permanenti mobili, con o senza un elemento ferromagnetico posteriore, che si muovono lungo uno statore a forma di C o di E, che racchiude un singolo o un doppio avvolgimento.

Tra le due soluzioni possibili è stata scelta l'opzione b) per varie ragioni, tra cui la presenza dell'effetto del cogging, che aumenta l'efficienza del sistema, la maggiore facilità di controllo e perché i conduttori attivi sono sullo statore e si evitano contatti mobili che sono soggetti a usura.

3.1.2 Primo design

Tra le varie opzioni possibili di attuatori che utilizzano magneti permanenti come mover è stata scelta una configurazione composta da uno statore a forma di E

con due avvolgimenti distinti, alimentati separatamente con correnti di segno opposto, in modo tale da ottenere dei movimenti in direzione opposta sui due mover. Questa configurazione, visibile in 3.2(a), è stata scelta perché ritenuta la più adatta ai vincoli di un TAH ed è di facile costruzione. La figura 3.2(a) mostra solo una sezione trasversale del dispositivo, il quale è ottenuto tramite una rotazione di 360° lungo l'asse z, presentando quindi una simmetria radiale.

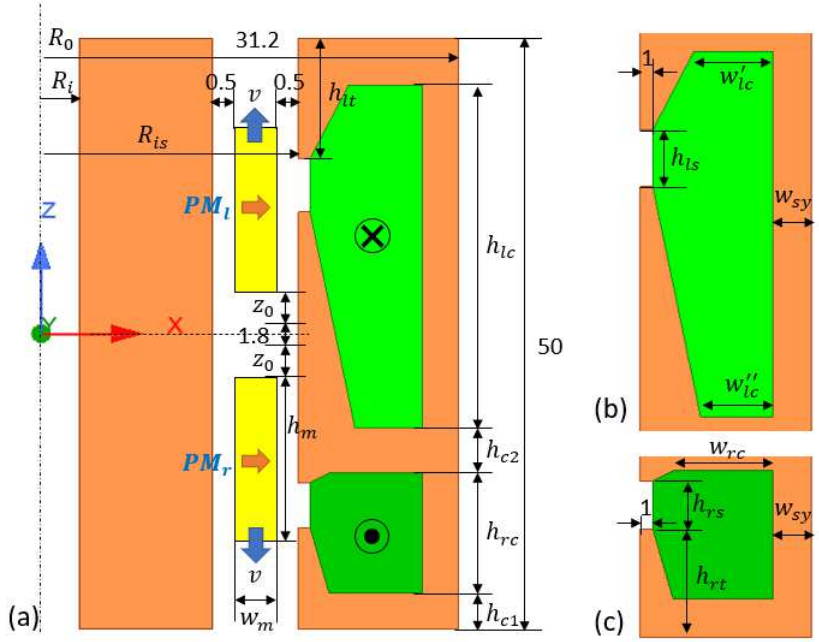


Figura 3.2: (a) Sezione trasversale del primo design del MM-LOA, con un dettaglio sull'avvolgimento sinistro (b) e destro (c). L'attuatore è ottenuto tramite una rotazione di 360° di tale sezione lungo l'asse z. Le quote sono in espresse in millimetri.

L'attuatore si compone di un nucleo interno, necessario per la richiusura del flusso magnetico, e uno statore esterno a forma di E realizzati in materiale ferromagnetico massiccio. Nelle due cave di statore sono inserite due bobine, dotate di alimentazione indipendente, per comandare i due mover in maniera separata[1]. Durante la fase sistolica si può assumere che il rapporto tra le forze massime che devono esplicare i due mover sia di circa 1:3. Questo si riflette sul rapporto tra le ampere spire dei due avvolgimenti, da cui dipendono le forze agenti sui mover. Imponendo una densità di corrente costante per i due avvolgimenti, il rapporto tra le aree trasversali delle due bobine viene impostato di conseguenza a circa 1:2.5 con il mover superiore che va a comprimere la camera ventricolare sinistra e quello inferiore che comprime quella destra. Il materiale ferromagnetico utilizzato è una lega di ferro cobalto (Vacoflux®48) la cui curva BH e permeabilità relativa in funzione dell'induzione magnetica H sono riportate in figura 3.3. È stato scelto questo materiale in quanto è un materiale ad alte prestazioni magnetiche ed è ottenibile per sinterizzazione, un processo tramite cui è possibile realizzare forme complesse mediante uno stampo. Nucleo e statore sono massicci, in quanto la

[1]Ciò non è del tutto possibile, a causa del mutuo accoppiamento tra i due avvolgimenti

frequenza a cui dovrà operare il dispositivo è quella cardiaca, compresa tra 1 e 2 Hz e un campo magnetico variabile a tale frequenza produce perdite trascurabili per correnti parassite nel materiale ferromagnetico. Tra le due parti fisse sono

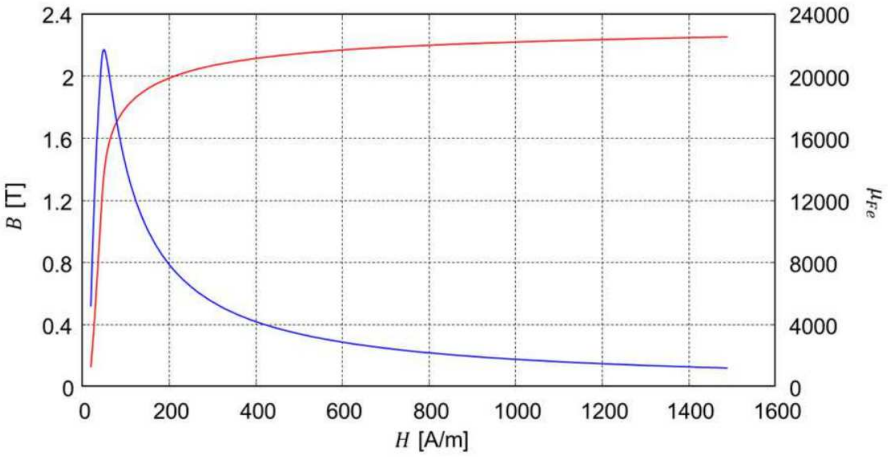


Figura 3.3: Curva BH e permeabilità relativa del materiale Vacoflux®48

inseriti due magneti permanenti a forma di anello, con magnetizzazione radiale verso l'esterno in NdFeB-48[2] ($B_r = 1.42$ T @ 20°C $\mu_{PM} = 1.1$). La magnetizzazione dei due magneti è repulsiva per evitare che si uniscano se si trovano troppo vicini. Le bobine sono realizzate in rame ($\rho_{cu} = 1.68 \cdot 10^{-8}\ \Omega \cdot m$ @ 20°C).

Per garantire la durabilità del dispositivo, si è deciso di evitare l'inserimento di molle o smorzatori meccanici che, essendo soggetti a cicli continui potrebbero usurarsi nel tempo. Le dimensioni del dispositivo sono state scelte affinché possa essere installato all'interno della gabbia toracica di una persona adulta e sono:

- Raggio esterno (R_t) = 31,2 mm
- Lunghezza assiale (L_a) = 50 mm

Come evidenziato in [5] con questa configurazione è possibile ottenere una forza media sui mover agenti sulla camera ventricolare sinistra (PM_l) e destra (PM_r) rispettivamente di 58,4 N e 25,4 N. Questi valori sono ottenuti con una densità di corrente (δ) uniforme nei due avvolgimenti di circa 5.8 A/mm^2, con perdite negli avvolgimenti di 13 W. La massa dell'attuatore risulta inferiore a 1 kg.

3.1.3 Secondo design

A partire dai risultati ottenuti dalla configurazione mostrata nel paragrafo precedente, sono state effettuate delle leggere modifiche alla geometria, che hanno portato a delle migliorie nell'attuatore; da questa nuova architettura ha avuto inizio il lavoro di ottimizzazione qui presentato.

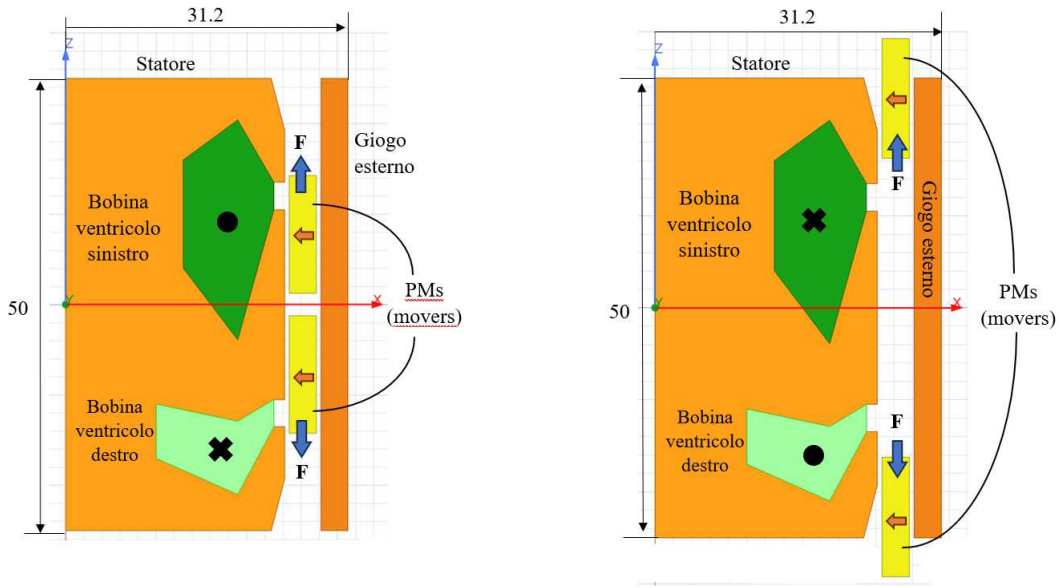
[2]Lega composta da Neodimio, Ferro e Boro

Innanzitutto è stata invertita la posizione di bobine e anelli magnetici, invertendo polo sud e nord di questi ultimi. Il dispositivo risulta così composto da:

- uno statore interno in materiale ferromagnetico massiccio a forma di E con due avvolgimenti in rame alimentati separatamente;
- due magneti permanenti mobili (mover) ad anello con magnetizzazione radiale verso l'interno. Questi dovranno avere un moto opposto e sincronizzato;
- un giogo esterno fisso in materiale ferromagnetico massiccio che ha la funzione di richiusura del flusso dei PMs e degli avvolgimenti.

I materiali utilizzati sono gli stessi introdotti nel paragrafo 3.1.2, in quanto è necessario utilizzare materiali ad alte prestazioni magnetiche, ovvero lega di cobalto e ferro per lo statore e per il giogo e NdFeB-48 per i magneti permanenti.

Viene poi realizzata una cianfrinatura agli estremi dei denti esterni di statore. Questa modifica viene introdotta perché, come verrà mostrato successivamente, porta dei benefici nella corsa di ritorno dei mover, durante l'espansione delle camere ventricolari. Inoltre viene modificata la geometria dei denti di statore sostituendo la geometria rettangolare con una trapezoidale, in modo da ottenere cave di sezione trasversale più grande e quindi, a parità di corrente, ridurre le perdite per effetto Joule sugli avvolgimenti.



(a) Design con i mover a inizio corsa

(b) Design con i mover a fine corsa

Figura 3.4: Rappresentazione di una sezione trasversale dell'attuatore lineare oscillate biventricolare dopo le prime modifiche. L'oggetto ha simmetria radiale lungo l'asse z. Sono indicati il verso delle correnti nelle bobine e la magnetizzazione dei magneti per avere un moto dei mover verso l'esterno nelle fasi di inizio (a) e fine (b) della corsa nella fase diastolica. Le quote sono espresse in mm.

L'attuatore così realizzato è visibile in figura 3.4, mentre la figura 3.5 è una sua rappresentazione in 3D. In queste immagini è anche possibile vedere la magnetizzazione dei magneti e il verso della corrente necessario per avere un moto dei mover verso l'esterno. Come già intuibile da queste immagini, e come riportato anche in [5], per mantenere una forza verso l'esterno durante tutta la corsa di 15 mm dei mover, sarà necessario invertire la polarità della corrente negli avvolgimenti in un punto della corsa.

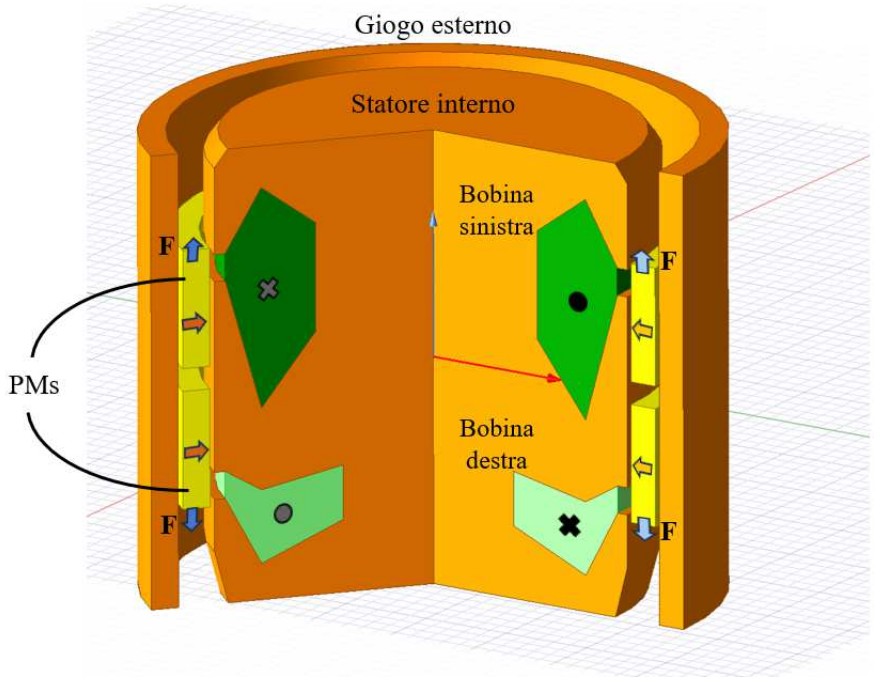


Figura 3.5: Rappresentazione 3D dell'attuatore lineare oscillate biventricolare, con verso della corrente e magnetizzazione dei PMs per ottenere un moto verso l'esterno di questi ultimi.

3.2 Funzionamento dell'attuatore

Una volta definita la geometria dell'attuatore è necessario analizzarla. Il dispositivo avrà il compito di comprimere due camere in materiale biocompatibile, mediante l'azione oscillante dei due mover, riproducendo il ciclo cardiaco. Nel paragrafo 1.2.2, è stato spiegato come un'istante prima che il cuore inizi a contrarsi i ventricoli risultino completamente riempiti di sangue; questo momento corrisponde alla posizione iniziale dell'attuatore, visibile in figura 3.4a e 3.5, in cui i magneti sono in posizione interna. Imponendo una corrente di verso opportuno sulle bobine, viene prodotta una forza magnetomotrice (f.m.m.) che va ad interagire con il campo magnetico prodotto dai PMs, esercitando una forza su questi ultimi diretta verso l'esterno; i mover, non essendo vincolati, sono liberi di muoversi. Questa forza è trasmessa alle camere ventricolari del TAH, le quali vengono compresse, innalzando la pressione al loro interno. Quando la pressione

di queste camere supera la pressione delle vene ad esse collegate, si ha la chiusura delle valvole inserite tra le camere ventricolari e le vene, dando inizio al ciclo cardiaco del cuore artificiale.

I mover, grazie all'effetto della f.m.m. prodotta dalle correnti circolanti negli avvolgimenti, vengono spinti verso l'esterno, comprimendo le camere ventricolari; la corsa prevista dei mover è di 15 mm. Questa fase vuole riprodurre la fase sistolica del ciclo cardiaco di un cuore normale. L'attuatore con i magneti a fine corsa è raffigurato in figura 3.4b. Durante questa fase si prevede che la pressione all'interno delle camere ventricolari salga al punto da superare la pressione delle arterie ad esse collegate, aprendo le valvole interposte tra camere e arterie, spingendo il sangue in queste ultime. Una volta che i mover raggiungono la loro posizione finale, la corrente negli avvolgimenti viene invertita di segno, in modo tale da produrre una forza di richiamo sui magneti: ha inizio la fase diastolica, in cui le camere ventricolari si espandono, riempiendosi di sangue. I mover vengono quindi riportati alla loro posizione iniziale, dove, al momento opportuno verrà iniettata una corrente negli avvolgimenti, in modo da ricominciare il ciclo cardiaco. La fase sistolica, di compressione, corrisponde quindi alla corsa di andata dei mover, mentre la fase diastolica, di rilassamento, corrisponde alla corsa di ritorno.

Da questa analisi risulta evidente come le forze agenti sui due mover durante le corse di andata e di ritorno siano di estrema importanza per capire se l'attuatore è in grado di svolgere il suo compito o meno; è quindi necessario calcolare i profili di forza sui mover al variare della loro posizione per la corsa di andata e di ritorno.

3.3 Calcolo dei profili di forza

Per il calcolo dei profili di forza si utilizza il software agli elementi finiti Ansys Electromagnetics®, conducendo analisi magnetostatiche. Siccome l'attuatore ha una simmetria radiale si possono condurre delle analisi assialsimmetriche 2D con geometria cilindrica lungo l'asse z utilizzando l'ambiente di calcolo "Maxwell 2D" presente all'interno del pacchetto fornito da Ansys®. Viene sfruttata la simmetria in quanto condurre analisi agli elementi finiti in 2D è molto meno oneroso in termini computazionali rispetto ad un'analisi 3D e i risultati che si ottengono sono gli stessi. Data la bassa frequenza di lavoro e siccome si suppone di lavorare a corrente pressoché costante durante la corsa, vengono condotte analisi magnetostatiche variando la posizione dei magneti, in modo da ottenere i profili di forza. La forza agente sui mover è senza dubbio dipendente dalla forza magnetomotrice al traferro e dalla reazione d'indotto prodotta dai PMs. Per calcolare i profili di forza è necessario definire, oltre alla geometria dell'attuatore, anche il valore delle ampere spire circolanti nei due avvolgimenti.

Il valore assoluto della corrente circolante nei due avvolgimenti, quindi anche il valore delle amperespire delle bobine, viene mantenuto costante durante la corsa di andata e di ritorno, in modo da mantenere costanti le perdite per effetto Joule negli avvolgimenti di statore durante un ciclo di lavoro; così facendo è possibile confrontare due geometrie dell'attuatore diverse a parità di perdite. Analogamente a quanto fatto in [5] è previsto un punto in entrambe le corse in cui

viene invertita la polarità della corrente (che verrà chiamato punto di switching), al fine di aumentare la forza media agente sui mover sia nella corsa di andata che in quella di ritorno. Questo punto è identico per entrambi gli avvolgimenti, in modo da diminuire il numero di simulazioni necessarie per la determinazione dei profili di forza. L'utilizzo di punti di switching diversi per i due avvolgimenti comporterebbe un'eccessiva complicazione nel calcolo dei profili di forza, che per il momento si vuole evitare, ma che non si esclude possa essere introdotta in analisi future, se i benefici supereranno gli svantaggi.

Una volta definite la geometria dell'attuatore e le le ampere spire circolanti nei due avvolgimenti ($N_L I_L$ per la bobina di sinistra e $N_R I_R$ per quella di destra), si può procedere al calcolo dei profili di forza. Viene costruito l'attuatore nell'ambiente di calcolo "Maxwell 2D" mettendo i magneti nella posizione iniziale, come ad esempio quella mostrata in figura 3.4a che è esattamente una schermata presa dall'ambiente di calcolo. Sulle bobine di sinistra e di destra (area verde scuro e verde chiaro dell'immagine 3.4a) viene imposta una corrente continua rispettivamente di

$$\begin{cases} I_{LC} = +N_L I_L \cdot pol_{LV} \\ I_{RC} = -N_R I_R \cdot pol_{RV} \end{cases} \tag{3.1}$$

dove pol_{LV} e pol_{RV} possono assumere solamente i valori -1 o +1 e indicano la polarità della corrente. In questo modo, quando $pol_{LV} = pol_{RV}$, sugli avvolgimenti circoleranno correnti di segno opposto. Per convenzione una corrente I_{LC} o I_{RC} di segno positivo è concorde con l'asse y di una terna cartesiana destrogira[3].

A questo punto, dopo aver impostato la polarità delle correnti ($pol_{LV} = pol_{RV} = pol = -1$) e la posizione dei magneti, è sufficiente analizzare la geometria tramite il software agli elementi finiti, con un'analisi magnetostatica 2D assialmmetrica; il software, calcolando la distribuzione del campo magnetico nel dispositivo è in grado di dare diverse informazioni utili, tra cui:

- le forze sui mover, F_{LV} e F_{RV}, utili alla determinazione dei profili di forza;
- le autoinduttanze, L_{LC} e L_{RC}, e il coefficiente di mutua induttanza, M_{LR}, delle due bobine, utili nella progettazione del controllo;
- il valore del flusso magnetico, λ_{LC} e λ_{RC} ,concatenato con le due bobine utili nella progettazione del controllo e dell'alimentazione del dispositivo;
- il valore del campo magnetico minimo lungo l'asse x all'interno dei magneti, $B_{x,min}$, necessario per evitare la smagnetizzazione di questi ultimi[4].

Si introducono ora dei termini che verranno utilizzati nelle seguenti pagine:

- $F_{for,LV}(x)$ e $F_{for,RV}(x)$: profili di forza sui mover in funzione della loro posizione nella corsa di andata, con modulo delle correnti costante;

[3]Con riferimento alla figura 3.4a, una corrente positiva è una corrente entrante nel piano xz visualizzato.

[4]Viene analizzato il campo magnetico minimo a causa della convenzione utilizzata: il magnete ha magnetizzazione opposta rispetto l'asse x, quindi produce campo magnetico negativo lungo quest'asse. Se dovesse essere rilevato un $B_{x,min} > 0$, il magnete andrebbe incontro a smagnetizzazione, cosa che non deve accadere.

- $F_{bck,LV}(x)$ e $F_{bck,RV}(x)$: profili di forza sui mover in funzione della loro posizione nella corsa di ritorno, con modulo delle correnti costante;
- $F^-_{LV}(x)$ e $F^-_{RV}(x)$: profili di forza sui mover in funzione della loro posizione imponendo una corrente pari a quella prevista dall'equazione 3.1, con $pol = -1$;
- $F^+_{LV}(x)$ e $F^+_{RV}(x)$: profili di forza sui mover in funzione della loro posizione imponendo una corrente pari a quella prevista dall'equazione 3.1, con $pol = +1$;

Una volta eseguita l'analisi, si ottengono i valori di forza sui mover in una singola posizione. Per ottenere le funzioni $F^-_{LV}(x)$ e $F^-_{RV}(x)$ è necessario ripetere l'analisi precedente per varie posizioni dei magneti. É stato scelto di dividere la corsa di 15 mm in 15 intervalli da 1 mm, quindi vengono effettuate in totale 16 analisi come quella appena descritta, ogni volta spostando i PMs di 1 mm in più verso l'esterno e mantenendo le correnti costanti, come definite nell'eq. 3.1. Siccome bisogna calcolare il profilo delle forze anche per la corsa di ritorno ed è necessario effettuare l'inversione delle polarità delle correnti in un punto in entrambe le corse, serve effettuare il calcolo di $F^+_{LV}(x)$ e $F^+_{RV}(x)$ analogamente a quanto appena fatto, imponendo nelle bobine sinistra e destra correnti di polarità inversa, impostando $pol_{LV} = pol_{RV} = pol = +1$. Si è giunti quindi al calcolo di:

$$F^-_{LV}(x^*) \quad F^-_{RV}(x^*) \quad F^+_{LV}(x^*) \quad F^+_{RV}(x^*) \quad \text{con } x^* \in \mathbb{N} : 0 \leq x^* \leq 15$$

A questo punto, tramite l'algoritmo "Akima modificato", un algoritmo di interpolazione cubica Hermitiana a tratti, presente nella libreria MatLab e descritto in [22, 23] si può eseguire l'interpolazione dei valori puntuali ottenuti con l'analisi agli elementi finiti (FEA), ottenendo quindi le funzioni:

$$F^-_{LV}(x) \quad F^-_{RV}(x) \quad F^+_{LV}(x) \quad F^+_{RV}(x) \quad \text{con } x \in \mathbb{R} : 0 \leq x \leq 15$$

In figura 3.6 sono riportati gli andamenti ottenuti in questo modo, avendo impostato come geometria dell'attuatore quella illustrata in figura 3.4 e come amperespire 381 A per la bobina sinistra e 200 A per quella destra, corrispondenti a perdite per effetto Joule negli avvolgimenti di 5 W. Una forza positiva è una forza concorde con l'asse z.

Ora si può procedere con la determinazione di $F_{for,LV}(x)$, $F_{for,RV}(x)$, $F_{bck,LV}(x)$ e $F_{bck,RV}(x)$. Ci sono due possibili strategie di controllo: la prima e più semplice prevede un unico punto di switching della corrente per la corsa di andata e di ritorno, mentre la seconda prevede due possibili punti di switching diversi tra le due corse. In entrambe le strategie di controllo il punto di switching è unico per i due avvolgimenti.

Unico punto di switching

Per il calcolo del profilo di forza con questa strategia, si procede come segue (per capire meglio la procedura risulta utile il confronto con la figura 3.6 in cui sono riportate $F^-_{LV}(x)$, $F^-_{RV}(x)$, $F^+_{LV}(x)$ e $F^+_{RV}(x)$):

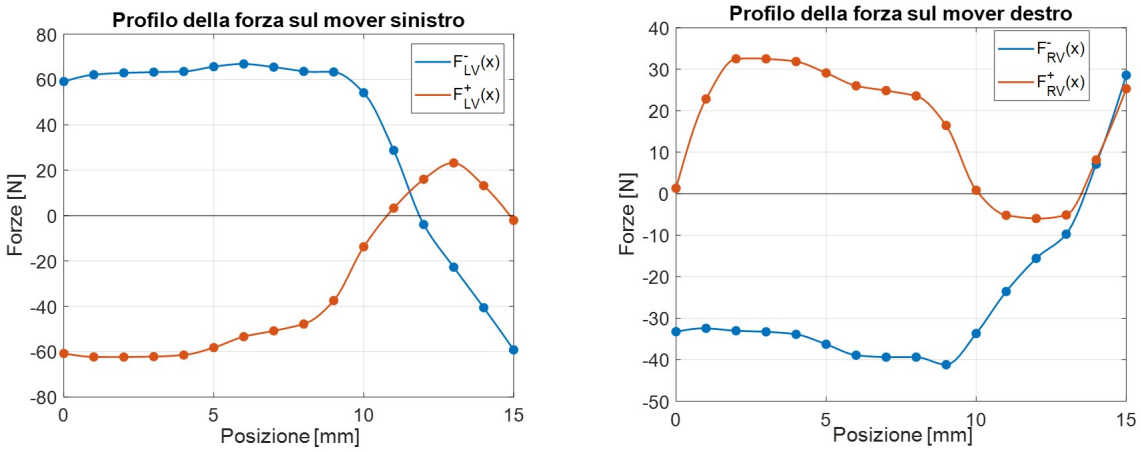


Figura 3.6: Andamenti delle forze sui due mover con Ampere spire sugli avvolgimenti fissate e i due valori di polarità possibili. I marker rappresentano punti campionati con le analisi FEM, le linee rappresentano invece i profili di forza interpolati con l'algoritmo "Akima modificato".

1. si inizia la corsa d'andata con $pol = -1$. In questo modo la forza sul mover sinistro è positiva e sul mover destro è negativa, concorde con il moto che si vuole ottenere[5].

2. Si determinano i punti:

$$\begin{cases} x^*_{kl} : F^-_{LV}(x^*_{kl}) = F^+_{LV}(x^*_{kl}) \\ x^*_{kr} : F^-_{RV}(x^*_{kr}) = F^+_{RV}(x^*_{rl}) \end{cases} \tag{3.2}$$

   questi sono i punti in cui le forze con polarità inversa hanno uguale valore e, tendenzialmente, saranno diversi per i due mover. Da lì in poi è sensato invertire la polarità: così facendo si avrà forza media più elevata in valore assoluto sui due mover.

3. Si determina il punto di switching x^*_{sw}, identico per i due avvolgimenti, nel seguente modo:

$$x^*_{\mathrm{sw}} = \min(x^*_{\mathrm{kl}}; x^*_{\mathrm{kr}}) \tag{3.3}$$

4. Si possono determinare le forze di andata e di ritorno nel seguente modo:

$$F_{\mathrm{for,k}}(x) = \begin{cases} F^-_k(x) & \text{se } x < x^*_{\mathrm{sw}} \\ F^+_k(x) & \text{se } x \geq x^*_{\mathrm{sw}} \end{cases} \tag{3.4}$$

$$F_{\mathrm{bck,k}}(x) = \begin{cases} F^+_k(x) & \text{se } x < x^*_{\mathrm{sw}} \\ F^-_k(x) & \text{se } x \geq x^*_{\mathrm{sw}} \end{cases} \tag{3.5}$$

[5]Si ricorda che il verso della forza è determinato dall'asse z, il cui verso positivo punta verso la camera ventricolare sinistra.

$$\text{con k} = \text{LV, RV}$$

Un esempio dei profili di forza così ottenuti è visualizzabile in figura 3.7, dove è stata analizzata la configurazione visibile in figura 3.4. Con questa configurazione e questa strategia di controllo si è ottenuta una forza media sui due mover nelle corse di andata di 49.3 N sul mover sinistro e di 26.2 N su quello destro, entrambe dirette verso l'esterno. Le forze medie nelle corse di ritorno sono rispettivamente di 41.8 N e 15.4 N dirette entrambe verso l'interno.

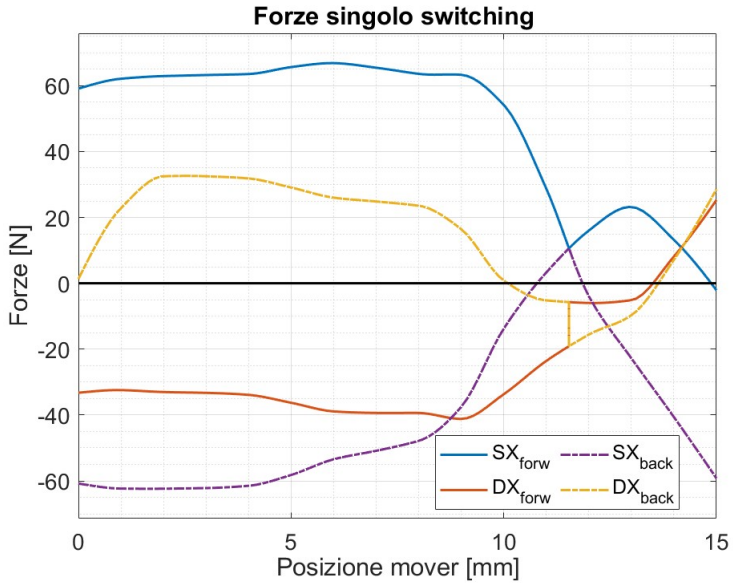


Figura 3.7: Profili delle forze sui mover al variare della loro posizione durante la loro corsa adottando la strategia di un unico punto di switching. Una forza positiva è concorde con l'asse z, ovvero è diretta verso il ventricolo sinistro.

Doppio punto di switching

Nella strategia precedente x^*_{sw} viene scelto in modo da migliorare le prestazioni all'andata e di conseguenza non c'è nessun margine per migliorare le prestazioni al ritorno. Scegliendo un punto di inversione della corrente diverso per le due corse si lascia più libertà in quella di ritorno, potendo così ottenere prestazioni migliori anche in questa fase. Per capire meglio i passaggi di questa nuova strategia è necessario introdurre il concetto di energia minima cumulativa della corsa di ritorno, che si calcola con i seguenti passaggi matematici: innanzitutto viene calcolato il lavoro fatto dalla forza di richiamo sul mover durante la corsa di ritorno, dal punto di fine corsa l (corrispondente a $x = 15$ mm) fino alla coordinata x,

$$e_{\text{k,b}}(x) = p_k \cdot \int_l^x F_{\text{k,bck}}(x)\, dx \tag{3.6}$$

dove k=LV, RV e $p_{LV} = +1$, $p_{RV} = -1$. Sono così identificate due funzioni, una per ogni mover, che possono essere calcolate solo una volta note $F_{\text{k,bck}}(x)$.

Può succedere che, durante il ritorno, la forza di cogging, diretta verso l'esterno, sia più elevata della forza di richiamo che si riesce ad imprimere grazie alle bobine e quindi la forza risultante sul mover si opponga al suo moto. Se il mover ha però acquistato abbastanza energia è in grado di superare questo ostacolo, completando la sua corsa. Per capire se il mover riesce a tornare nella posizione iniziale si guarda l'indice calcolato in 3.6: se $e_{\mathrm{k,b}}(x)$ diventa negativa, significa che l'equipaggio non è in grado di ritornare alla posizione iniziale, in quanto verrà fermato dalla forza di segno opposto al moto, se invece resta sempre positiva, il mover riesce a tornare nella sua posizione iniziale.

Siccome, per costruzione, $e_{\mathrm{k,b}}(15) = 0$, più che la funzione è conveniente guardare il suo valore minimo. Se questo è 0, il mover k riesce a tornare in posizione iniziale, se è negativo non ci riesce. Viene definito un indice unico che unisca le informazioni sui due mover per una maggiore sintesi:

$$E_{m,b} = \sqrt{\sum_k \left[\min\left(e_{k,b}(x)\right)\right]^2} \tag{3.7}$$

Se tale indice è positivo, l'attuatore non riesce a riportare i mover nella loro posizione iniziale e può esserci un problema sul magnete di destra, di sinistra o su entrambi; se invece l'indice è uguale a 0, i magneti possono essere riportati nella loro posizione iniziale. Nella pratica ci si aspetta che il sangue in arrivo dalle vene possa imprimere una forza sui mover durante la fase diastolica che li aiuti a completare la loro corsa di ritorno, vincendo l'ostacolo dato dalla forza di cogging; l'attuatore deve però essere in grado di garantire il corretto ritorno in posizione iniziale dei mover, senza fare affidamento su forza esterne.

Di seguito vengono riportati i passaggi da seguire per la strategia di doppio switching:

1 - 3. Uguale ai passaggi da 1 a 3 della strategia precedente, solo che nel punto 3 viene determinato $x^*_{\mathrm{sw,for}}$ invece di x^*_{sw}; si determinano quindi i profili delle forze di andata

$$F_{\mathrm{for,k}}(x) = \begin{cases} F_k^-(x) & \text{se } x < x^*_{\mathrm{sw,for}} \\ F_k^+(x) & \text{se } x \geq x^*_{\mathrm{sw,for}} \end{cases} \tag{3.8}$$

con k = LV, RV

4. si testano vari valori di $x^*_{\mathrm{sw,bck}}$: $x^*_{sw,bck,min} \leq x^*_{\mathrm{sw,bck}} \leq 15$ con un intervallo tra un valore e l'altro di 0.1 e si calcolano i profili di $F_{bck,k}$ e di $E_{m,b}$ come nelle equazioni 3.5 e 3.7; per l'applicazione in esame si è ritenuto conveniente impostare $x^*_{sw,bck,min} = 8$ mm

5. si sceglie il valore di $x^*_{\mathrm{sw,bck}}$ che minimizzi $E_{m,b}$; se ci sono più valori per cui $E_{m,b} = 0$, si sceglie $x^*_{\mathrm{sw,bck}}$ che massimizzi le forze medie di ritorno dei due mover.

6. si determina quindi il profilo delle forze sulle corse di ritorno

$$F_{\mathrm{bck,k}}(x) = \begin{cases} F_k^+(x) & \text{se } x < x^*_{\mathrm{sw,bck}} \\ F_k^-(x) & \text{se } x \geq x^*_{\mathrm{sw,bck}} \end{cases} \tag{3.9}$$

con k = LV, RV

Utilizzando come esempio la configurazione già analizzata in questo capitolo, i profili di forza che si ottengono sono riportati in figura 3.8. Con tale strategia si ottengono delle forze medie sulle corse di andate del mover sinistro e destro rispettivamente di 49.3 N e 26.2 N, esattamente come nell'altra strategia e delle forze medie durante le corse di ritorno rispettivamente di 40 N e 16.2 N. La stessa configurazione, con la strategia precedente presenta un valore di $E_{m,b}$ = 11.2 mJ, mentre con questa strategia si ottiene un miglioramento di tale indice che viene portato a 0.47 mJ. Entrambe le strategie di switching producono gli

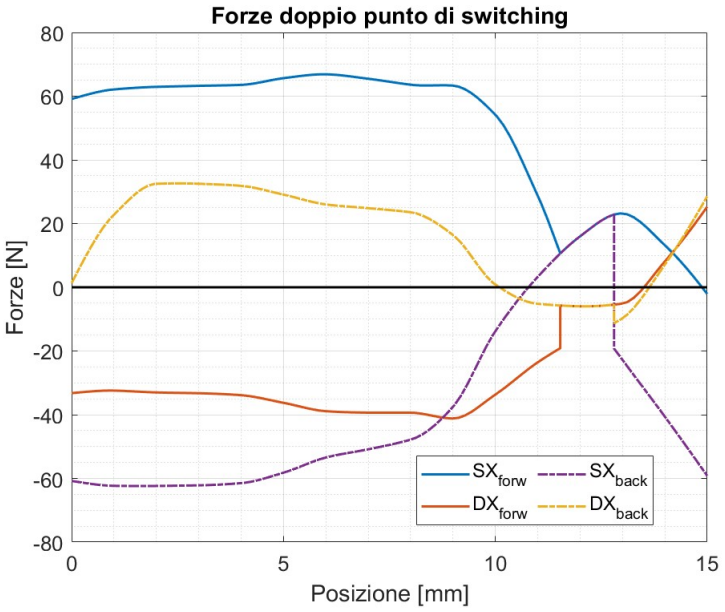


Figura 3.8: Profili delle forze sui mover al variare della loro posizione durante la loro corsa adottando la strategia del doppio punto di switching. Una forza positiva è concorde con l'asse z, ovvero è diretta verso il ventricolo sinistro.

stessi risultati in termini di forze medie sulle corse di andata dei mover; quello per cui differiscono sono le prestazioni nella fase di ritorno. La seconda strategia premette infatti di ottenere prestazioni migliori in questa fase rispetto alla prima, ma è anche più complicata, in quanto prevede l'utilizzo di due sensori per poter capire quando effettuare l'inversione di corrente nelle due corse. La prima prevede invece l'utilizzo di un unico sensore. Nell'esempio analizzato, nessuna delle due strategie permette il ritorno dei mover in posizione iniziale, ma la seconda va decisamente più vicino.

Per il calcolo dei profili di forza è necessario effettuare 32 analisi magnetostatiche agli elementi finiti (16 con $pol = 1$ e 16 con $pol = -1$). Impostando un mesh iniziale di circa 20 000 elementi, forzando il software a fare tra i due e i tre passi adattivi e impostando come criterio di convergenza che l'errore percentuale sull'energia sia inferiore allo 0.1 %, l'esecuzione di una singola analisi FEM dura circa 22 secondi, quindi per calcolare il profilo delle forze per una configurazione sono necessari circa 11 minuti e 44 secondi. Impostando su Ansys il calcolo

parallelo multiprocessore è possibile eseguire 4 simulazioni contemporaneamente, riducendo il tempo di calcolo per le 32 analisi necessarie alla determinazione dei profili di forza per una singola configurazione a circa 3 minuti e 44 secondi.

3.4 Indici di prestazione

Per effettuare un confronto tra varie geometrie dell'attuatore, risulta utile poter definire degli indici che definiscono la bontà di una configurazione. Innanzitutto, dato il profilo di forza in una corsa, risulta immediato il calcolo della forza media e del ripple relativo della forza, anche chiamato ripple in per unità (p.u.), associato alla corsa analizzata. Questi indici, per le corse d'andata dei due mover, sono così calcolati:

$$F_{m,k}^{for} = \int_0^{15} F_{for,k} dx \tag{3.10}$$

$$r_{k,for} = \sqrt{\frac{\frac{1}{15}\int_0^{15} F_{for,k}^2(x)dx}{(F_{m,k}^{for})^2} - 1} \tag{3.11}$$

con k = LV, RV

Per le corse di ritorno le formule sono identiche, solo che sono invertiti gli estremi di integrazione:

$$F_{m,k}^{bck} = \int_{15}^{0} F_{bck,k} dx \tag{3.12}$$

$$r_{k,bck} = \sqrt{\frac{\frac{1}{15}\int_{15}^{0} F_{bck,k}^2(x)dx}{(F_{m,k}^{bck})^2} - 1} \tag{3.13}$$

con k = LV, RV

Forza media e ripple in p.u. sono indici rispettivamente del lavoro che viene svolto dall'attuatore sulle camere ventricolari e di quanto la forza abbia un andamento oscillante attorno al suo valore medio.

Altri indici di cui è necessario tenere conto sono le forze di avviamento dei due mover, $F_{start,LV} = F_{LV}^{-}(0)$ e $F_{start,RV} = F_{RV}^{-}(0)$, nonché i valori dell'energia minima cumulativa delle due corse di ritorno, $e_{LV,min}$, $e_{RV,min}$.

3.5 Vincoli

L'attuatore in esame deve sottostare a dei vincoli per poter assolvere al suo compito. Innanzitutto deve essere compatto per poter essere impiantato nel torace di una persona adulta. Le dimensioni esterne non possono quindi essere modificate (lunghezza assiale di 50 mm e diametro esterno di 62.4 mm). In secondo luogo, la potenza termica generata dal dispositivo deve essere bassa. Il mezzo di raffreddamento del dispositivo è infatti il sangue, che se supera la temperatura di 42°C va incontro a denaturazione. Per non eccedere questa temperatura, si

fissa il limite di potenza termica generata dal dispositivo a 20 W [24, 25], anche se è preferibile limitare la generazione di calore ben al di sotto di questo valore. Anche il peso del TAH non deve essere troppo elevato; viene imposto un limite di 1 kg al peso dell'attuatore[5].

É poi importante che l'attuatore riesca a svolgere il lavoro del cuore. Il TAH è pensato per lavorare alla stessa frequenza a cui lavora un cuore reale, quindi con gittata sistolica di ca. 70 ml. Il lavoro che esplica l'attuatore sui due ventricoli nella fase sistolica è dato dal prodotto della forza media agente sul mover nella corsa di andata per lo spostamento effettuato dal mover, di 15 mm. Questo lavoro deve essere uguale a quello effettuato dal cuore nel tratto BCD della figura 3.9, già introdotta nel capitolo 1. Tale lavoro corrisponde all'area sottesa a tale curva ed è da considerarsi positivo in quanto si è interessati al lavoro svolto sul ventricolo da parte di un organo esterno, a cui va aggiunto il lavoro cinetico: Riprendendo

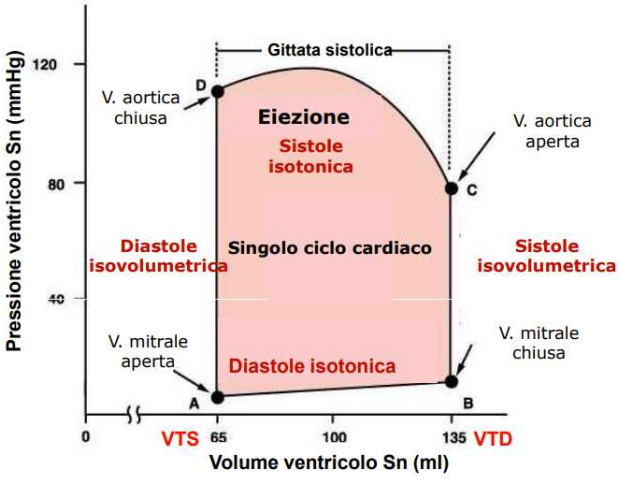


Figura 3.9: Rappresentazione nel piano pressione volume del ciclo cardiaco

ciò che è stato introdotto nel paragrafo 1.3, tale lavoro vale

$$L_k = GS \cdot \left(p_{m,CD} + \frac{1}{2} \cdot \rho_s \cdot v^2 \right) \quad \text{con k = LV, RV} \tag{3.14}$$

Ponendo $L_k = F_{m,h}^{for} \cdot stroke$, dove con $stroke$ si intende la lunghezza della corsa, ovvero 15 mm, è possibile calcolare $F_{m,k}^{for}$ tramite la seguente espressione:

$$F_{m,k}^{for} = \frac{GS}{stroke} \cdot \left(p_{m,CD} + \frac{1}{2} \cdot \rho_s \cdot v^2 \cdot \right) \tag{3.15}$$

la $p_{m,CD}$ corrisponde alla pressione media arteriosa durante la fase di eiezione. In letteratura i valori di pressione che si trovano sono di 100 mmHg per il ventricolo sinistro e 20 mmHg per quello destro, anche se per quest'ultimo è possibile trovare valori anche più bassi, intorno ai 15 mmHg [26]. Con tali valori, grazie alla 3.15, è possibile calcolare le forze medie nominali su cui tarare l'ottimizzazione dell'attuatore, che risultano di

$$F_{m,LV}^{for} = 63 \text{ N} \qquad F_{m,RV}^{for} = 15 \text{ N} \tag{3.16}$$

È poi necessario esplicare delle forze di picco di rispettivamente 70 N e 25 N per poter portare la pressione ventricolare a 120 mmHg e 30 mmHg [5, 25]. I valori qui ottenuti sono in linea con quelli che si trovano in letteratura, anche se si può trovare un valore un po' più elevato sulla forza media nel mover di destra, intorno ai 20 N [24, 27].

Questi valori sono calcolati per una persona sana a riposo e possono variare da persona a persona. La pressione media aortica è fortemente dipendente dal sistema circolatorio del paziente in esame ed è possibile trovare valori più elevati, che richiedono forze medie esplicabili dall'attuatore più elevate. L'attuatore deve anche potersi adattare ad un sovraccarico, ad esempio se il paziente sta effettuando uno sforzo. Risulta quindi necessario che l'attuatore sia in grado di poter esplicare delle forze anche superiori a quelle indicate in 3.16. Inoltre, a parità di forza media esplicata è preferibile che la forza sia più elevata ad inizio corsa e più bassa verso la fine, in modo tale da riuscire a mettere in movimento il sangue fin dai primi istanti.

Una volta terminata la corsa, il dispositivo deve essere in grado di riportare i mover nella loro posizione iniziale senza l'ausilio di forze esterne. Ciò significa che il minimo dell'energia cumulativa di entrambe le corse deve essere zero, altrimenti ci sarà un punto in cui il mover si bloccherà.

C'è anche un aspetto riguardante il controllo: l'attuatore deve poter essere comandato da una batteria di backup impiantabile, che deve garantire la funzionalità del TAH in caso di guasto dell'alimentazione principale. Questa batteria normalmente ha una tensione di 20 V, per cui la tensione di picco di una fase non deve superare tale valore [5].

Si possono identificare altri vincoli sul profilo della forza che deve essere esplicata dall'attuatore sui mover:

- le forze all'avviamento devono essere superiori a 40 N per il mover sinistro e 20 N per quello destro, dirette verso l'esterno, altrimenti i PMs non hanno modo di iniziare la loro corsa [5];

- invertendo la polarità degli avvolgimenti, con i magneti in posizione iniziale, la forza agente su questi deve essere diretta verso l'interno. Questo è necessario affinché si possa trovare una condizione per cui i mover si trovino in posizione iniziale, con forza nulla agente su di essi[6] ed è importante per il controllo dell'attuatore;

- analizzando il profilo della forza di ritorno, è necessario che questa sia diretta verso l'interno per entrambi i mover quando questi si trovano nella posizione più esterna (x=15mm): solo così è possibile richiamare i mover verso l'interno;

- i valori medi delle forze delle 4 corse devono essere concordi con il verso del moto desiderato.

[6]Questa condizione è raggiunta quando la forza repulsiva dei magneti e la forza di cogging che tendono a muovere il mover verso l'esterno vengono bilanciate dalla forza prodotta dall'interazione dei PMs con la f.m.m. prodotta dagli avvolgimenti.

Capitolo 4

Il Metodo di Taguchi

In questo capitolo viene illustrato il metodo di Taguchi e come è stato utilizzato in questo lavoro, insieme a tecniche di analisi della varianza per l'ottimizzazione dell'attuatore lineare MM-LOA.

Il metodo di Taguchi è un metodo standardizzato, sviluppato dal Dr. Genichi Taguchi, largamente utilizzato come tecnica di ottimizzazione per i processi, che consente di trovare la migliore combinazione dei fattori di progetto che forniscono la migliore risposta in termini di fattori controllati, o indici di prestazione, conducendo il minor numero di esperimenti [28]. Tale metodo si può dividere in quattro fasi, che sono condotte in modo tale da ridurre la variabilità del processo rispetto a fattori esterni incontrollabili [29]:

1. identificazione dei parametri di progetto modificabili e dei livelli che può assumere ogni parametro e degli indici di prestazione o di risposta e del target a cui li si vuole far tendere;

2. selezione della matrice ortogonale appropriata;

3. conduzione degli esperimenti seguendo la matrice ortogonale scelta, analizzando i risultati con tecniche di analisi della varianza, per scegliere la miglior combinazione dei parametri al fine di ottenere le prestazioni migliori;

4. condurre il test di conferma usando la combinazione ottimale di parametri scelta.

4.1 Revisione della letteratura

Questo metodo trova molte applicazioni in settori come quello farmaceutico e chimico. Data una serie di parametri che influenzano un processo consente infatti di trovare la combinazione di tali parametri che massimizzi le prestazioni del processo, conducendo il minor numero di esperimenti possibili. Quest'ultimo aspetto è fondamentale in quanto ogni esperimento ha un costo che può essere espresso sia in termini di denaro che di tempo. Un esempio si trova in [30], dove viene utilizzato il metodo di Taguchi al fine di ottimizzare i parametri fisico chimici di incubazione del batterio *Stenotrophomonas maltophilia* responsabile della

crescita biogenica di nanoparticelle di ossido di Tungsteno, avendo come scopo quello di creare un composto omogeneo. Come indici di risposta sono utilizzati la dimensione delle nanoparticelle e l'indice di poli-dispersione delle particelle, indicativo dell'uniformità del composto ottenuto. Dopo aver individuato i parametri che possono influenzare il processo di incubazione, quali la temperatura dell'aria, il pH, il tempo di crescita, il tasso di aerazione, la concentrazione del batterio, vengono individuati i livelli che questi parametri possono assumere. Sono poi condotti gli esperimenti secondo una matrice ortogonale. Una volta ottenuti i risultati è possibile, mediante analisi della varianza (ANOVA) e i signal to noise ratio proposti da Taguchi individuare il livello di ogni variabile per avere le prestazioni migliori in termini degli indici di risposta. Un esperimento può durare anche diversi giorni: da questo fatto si comprende la forza del metodo di Taguchi che consente di ridurre al minimo il numero di esperimenti da condurre.

Numerose sono le applicazioni di tale metodo anche in ambito ingegneristico. Esempi si possono trovare in [31, 32], in cui al metodo di Taguchi viene affiancata la GRA analisi, tecnica utilizzata quando le variabili di risposta sono più di una. In [31] viene utilizzato il metodo di Taguchi per ottimizzare i parametri della saldatura a trasmissione laser per ottenere una giunzione ottimale in termini di resistenza e spessore della saldatura tra due materiali differenti. Questi sono due indici di prestazione che hanno obiettivi differenti: la resistenza della saldatura deve essere massimizzata, mentre lo spessore della giuntura deve essere minimizzata. Con il metodo di Taguchi è possibile perseguire obiettivi diversi per indici di risposta diversi. In [32] il metodo di Taguchi viene applicato per ottimizzare la ventilazione in una stanza di un edificio, scegliendo la combinazioni di 5 parametri: la velocità dell'aria in ingresso, la temperatura dell'aria, l'altezza e lo spessore della stanza e lo spessore dell'apertura di areazione. Per la conduzione degli esperimenti sono utilizzate delle simulazioni fluidodinamiche al calcolatore, invece che esperimenti di laboratorio, come negli altri due casi qui presentati. In [33] il metodo di Taguchi viene integrato con le reti neurali al fine di predire gli stress interni indotti durante la deposizione chimica tramite vapore, aiutata dal plasma, di strati sottili di silicio amorfo idrogenato. In [34] il metodo di Taguchi è integrato con un algoritmo genetico per l'ottimizzazione della costruzione di fibre ottiche ad alto potere trasmissivo.

4.2 Definizione input e output

Punto essenziale del metodo di Taguchi è quello iniziale: è necessario individuare i parametri di progetto modificabili e i valori che possono assumere. Da questi parametri dipenderanno gli indici di prestazione del processo o, come nel nostro caso, del dispositivo che si vuole ottimizzare. Il metodo poi sarà in grado di fornire la combinazione dei parametri di progetto tali per cui gli indici di prestazione siano i migliori, sia che si tenda alla loro massimizzazione, minimizzazione o che li si voglia far convergere verso un valore predefinito.

Lo scopo del lavoro presentato è di trovare la migliore geometria dell'attuatore in modo tale da massimizzare le prestazioni dell'attuatore. Le variabili di input sono quindi quelle che descrivono la geometria del dispositivo, mentre le variabili

di output sono quelle che descrivono i profili di forza sui mover durante le loro corse di andata e di ritorno.

4.2.1 Definizione delle variabili di progetto

Lo scopo di questo paragrafo è quello di trovare una corrispondenza biunivoca tra un numero fisso di variabili iniziali, possibilmente non troppo elevato, e una configurazione ben definita dell'attuatore. Si vuole infatti ottimizzare la geometria dell'attuatore esprimendola in termini di variabili di progetto.

Punto di partenza è la geometria presentata in figura 4.1, in cui viene dato un nome alle varie dimensioni dell'attuatore. Risulta evidente come non si possa imporre una variabile per ogni quota: questo porterebbe sì a una corrispondenza biunivoca tra le variabili e la geometria dell'attuatore, ma si pagherebbe il prezzo di avere troppe variabili. È necessario quindi trovare delle relazioni tra le varie dimensioni in modo che, una volte che ne vengano definite alcune, le altre siano definite di conseguenza.

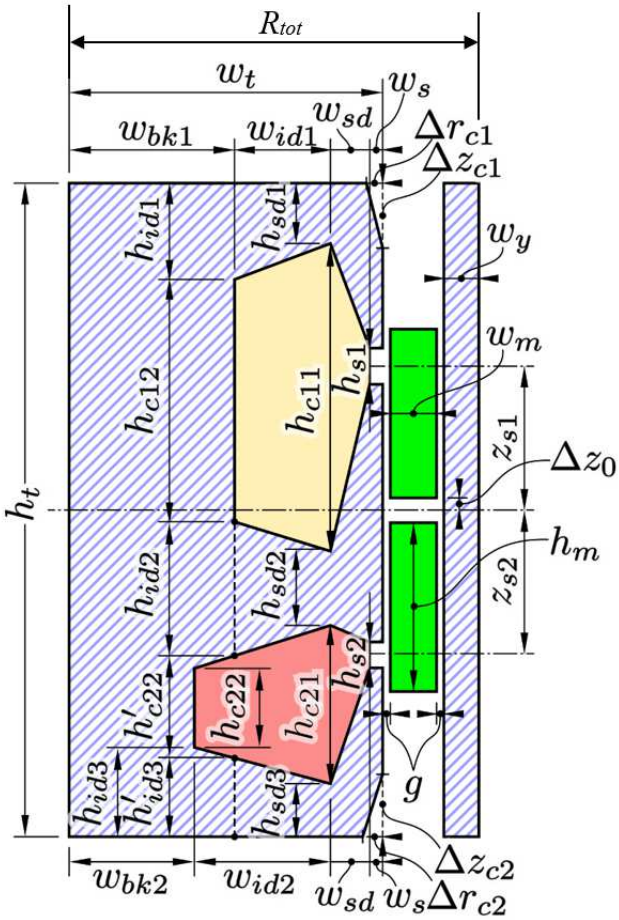


Figura 4.1: Architettura dell'attuatore lineare MM-LOA quotata.

Innanzitutto, come già segnalato nel paragrafo 3.5 le dimensioni esterne sono fissate; ciò significa che $h_t = 50$ mm e $R_{tot} = 31.2$ mm. Per quanto riguarda il traferro, g, questo deve essere il più piccolo possibile, per avere una migliore

interazione tra la f.m.m. e i PMs e viene pertanto fissato a 0.5 mm, sulla base di vincoli costruttivi. La semidistanza tra i magneti a inizio corsa viene impostata a $\Delta z_0 = 1.25$ mm. Vengono poi impostate come variabili indipendenti i parametri geometrici w_y, w_m, h_m, w_{bk1}, w_{bk2}, h_{s1}, h_{s2}, w_{sd}, h_{sd1}, h_{sd3}, h_{c11}, Δr_{c1}, Δr_{c2}, z_{s1}, z_{s2} da cui si ricavano di conseguenza gli altri parametri w_t h_{sd2}, h_{c21}, w_{id1}, w_{id2}, h_{id1}, h_{id2}, h_{id3}, h_{c22}, h_{c12}, w_{id1} e w_{id2} in base alle considerazioni che seguono.

Imponendo $w_s = 1.2$ mm, per differenza tra il raggio esterno e gli altri parametri già fissati, si può calcolare w_t:

$$w_t = R_{tot} - (2 \cdot g + w_y + w_m) \tag{4.1}$$

L'area alla base della testa del dente centrale deve essere pari alla somma delle aree alla base delle teste dei due denti laterali perché in esso viene convogliato un flusso corrispondente grosso modo alla somma dei flussi che percorrono gli altri due. Trovandosi tutte e tre alla stessa quota radiale, questo si traduce nel vincolo:

$$h_{sd2} = h_{sd1} + h_{sd3} \tag{4.2}$$

Si può quindi calcolare h_{c21} grazie a considerazioni geometriche lungo l'asse z:

$$h_{c21} = h_t - (h_{c11} + h_{sd1} + h_{sd2} + h_{sd3}) \tag{4.3}$$

Le altezze delle cave w_{id1} e w_{id2} si ricavano per differenza a partire da w_t, w_{bk1}, w_{bk2}, w_s e w_{sd}:

$$\begin{aligned} w_{id1} &= w_t - (w_{bk1} + w_s + w_{sd}) \\ w_{id2} &= w_t - (w_{bk2} + w_s + w_{sd}) \end{aligned} \tag{4.4}$$

Al fine di mantenere all'incirca costante il valore dell'induzione nel dente, si ipotizza che le larghezze alla base dei denti, h_{id1}, h_{id2}, h_{id3} stiano alle corrispondenti larghezze della base della testa dente in modo inversamente proporzionale alla coordinata radiale, ottenendo:

$$h_{id1} = \left(1 + \frac{w_{id1}}{w_{bk1}}\right) \cdot h_{sd1} \tag{4.5}$$

$$h_{id2} = \left(1 + \frac{w_{id1}}{w_{bk1}}\right) \cdot h_{sd2} \tag{4.6}$$

$$h_{id3} = \left(1 + \frac{w_{id2}}{w_{bk2}}\right) \cdot h_{sd3} \tag{4.7}$$

Imponendo un andamento rettilineo dei fianchi delle cave, si ricava che le larghezze del dente e della cava inferiori corrispondenti al valore della coordinata radiale w_{bk1} (fondo della cava superiore) valgono rispettivamente:

$$h'_{id3} = h_{sd3}\frac{w_{id2} - wid1}{w_{id2}} + h_{id3}\frac{w_{id1}}{w_{id2}} \tag{4.8}$$

$$h'_{c22} = h_{c21}\frac{w_{id2} - wid1}{w_{id2}} + h_{c22}\frac{w_{id1}}{w_{id2}} \tag{4.9}$$

Dato che la somma delle larghezze di denti e cave in corrispondenza di qualsiasi valore della coordinata radiale vale h_t, ed in particolare per la coordinata w_{bk1}, si ottiene il vincolo:

$$h_{id1} + h_{c12} + h_{id2} + h'_{c22} + h'_{id3} = h_t \tag{4.10}$$

Dalla figura 4.1 si possono ricavare le aree trasversali delle due bobine, mediante le formule:

$$A_{LV} = h_{c12}\frac{w_{id1}}{2} + h_{c11}\frac{w_{id1} + w_{sd1}}{2} + h_{s1}\frac{w_{sd1}}{2} \tag{4.11}$$

$$A_{RV} = h_{c22}\frac{w_{id2}}{2} + h_{c21}\frac{w_{id2} + w_{sd2}}{2} + h_{s2}\frac{w_{sd2}}{2} \tag{4.12}$$

Imponendo un rapporto tra le aree trasversali delle due bobine pari a ρ, si ottiene la seguente condizione:

$$\frac{A_{RV}}{A_{LV}} = \rho \tag{4.13}$$

Sostituendo in 4.13 le espressioni ricavate in 4.11 e 4.12 si ottiene l'equazione

$$h_{c22}\frac{w_{id2}}{2} + h_{c21}\frac{w_{id2} + w_{sd2}}{2} + h_{s2}\frac{w_{sd2}}{2} = \rho\left(h_{c12}\frac{w_{id1}}{2} + h_{c11}\frac{w_{id1} + w_{sd1}}{2} + h_{s1}\frac{w_{sd1}}{2}\right) \tag{4.14}$$

in cui le incognite sono h_{c12} e h_{c22}. Mettendo poi a sistema l'equazione 4.14 con il vincolo 4.10 è possibile ricavarsi i valori di h_{c12} e h_{c22} mediante le

$$h_{c12} = \frac{h_a w_{id2}^2 + w_{id1}\left[(h_{c21} + h_b)w_{id2} + h_c w_{sd} - \rho h_{c11} w_{id1}\right]}{den} \tag{4.15}$$

$$h_{c22} = \frac{\rho h_b w_{id1}^2 + w_{id2}\left[\rho(h_{c11} + h_a)w_{id1} - h_{c21} w_{id2} - h_c w_{sd}\right]}{den} \tag{4.16}$$

avendo introdotto, al fine di semplificare la scrittura, le grandezze:

$$\begin{aligned}
h_a &= h_t - (h_{c21} + h_{id1} + h_{id2} + h_{sd3}) \\
h_c &= h_{c21} + h_{s2} - \rho(h_{c11} + h_{s1}) \\
h_b &= h_{c21} - h_{id3} + h_{sd3} \\
den &= \rho w_{id1}^2 + w_{id2}^2
\end{aligned}$$

Rimangono da determinare i valori di Δz_{c1} e di Δz_{c2}. Si vuole evitare che la cianfrinatura intersechi l'apertura di cava che, tradotto a livello di disequazioni diventa:

$$\Delta z_{c1} < \frac{h_t}{2} - \left(z_{s1} + \frac{h_{s1}}{2}\right) \tag{4.17}$$

$$\Delta z_{c2} < \frac{h_t}{2} - \left(z_{s2} + \frac{h_{s2}}{2}\right) \tag{4.18}$$

si preferisce quindi introdurre due nuove variabili α_{c1} e di α_{c2}, sulla base delle quali saranno calcolati Δz_{c1} e Δz_{c2} grazie alle equazioni seguenti:

$$\Delta z_{c1} = \alpha_{c1}\left[\frac{h_t}{2} - \left(z_{s1} + \frac{h_{s1}}{2}\right)\right] \tag{4.19}$$

$$\Delta z_{c2} = \alpha_{c2} \left[\frac{h_t}{2} - \left(z_{s2} + \frac{h_{s2}}{2} \right) \right] \tag{4.20}$$

Così facendo, per valori di α_{c1} e α_{c2} inferiori a 1, le disequazioni 4.17 e 4.18 sono soddisfatte.

Per la determinazione univoca di una configurazione, sono state utilizzate 17 variabili geometriche indipendenti. Il metodo di Taguchi richiede che siano ora determinati i valori, anche chiamati livelli, che ogni variabile può assumere. Per poterli determinare, si è prima individuato il range di variazione di ogni singola variabile, sulla base di considerazioni geometriche e di risultati di lavori precedenti [5], dopodiché si è diviso questo range in N livelli equi-spaziati. Più N è grande e più il range risulta ben rappresentato, al costo però di un maggior numero di simulazioni da dover effettuare. Per la prima analisi con il metodo di Taguchi si è scelto $N = 5$. Le 17 variabili indipendenti con i relativi range di variazione sono riportate nella tabella 4.1. Una volta impostato il valore delle

Tabella 4.1: Le 17 variabili indipendenti utilizzate come input per l'ottimizzazione e relativi range di variazione.

Variabile	Range
w_{bk1}	(10; 13) mm
h_{sd1}	(4; 6) mm
h_{s1}	(2; 5) mm
w_m	(2.4; 3.6) mm
z_{s1}	(9; 15) mm
Δr_{c1}	(0.5; 2.5) mm
α_{c1}	(0.2; 0.7)
h_{c11}	(20; 24) mm
w_y	(2.5; 4) mm

Variabile	Range
w_{bk2}	(10;13) mm
h_{sd3}	(4; 6) mm
h_{s2}	(2; 5) mm
h_m	(12; 17) mm
z_{s2}	(9; 15) mm
Δr_{c2}	(0.5; 2.5) mm
α_{c2}	(0.2; 0.7)
w_{sd}	(3; 5) mm

17 variabili la geometria dell'attuatore è univocamente definita. Per calcolare i profili di forza è ora necessario impostare il valore delle ampere spire delle due bobine. Per un confronto equo tra diverse configurazioni, più che il valore NI per i due avvolgimenti, si impongono le perdite totali del dispositivo P_{Loss}, in quanto, come già affermato nel paragrafo 3.5, esiste un vincolo molto stringente su di esse ($P_{Loss} < 20$ W). P_{Loss} si compone di due termini: le perdite per effetto Joule dovute al passaggio di corrente sulla resistenza di valore finito degli avvolgimenti, P_{Cu} e le perdite nel materiale ferromagnetico, P_{Fe} dovute alle perdite per isteresi magnetica, dipendenti da f e alle perdite per correnti parassite generate dal campo magnetico variabile nel tempo, dipendenti da f^2, dove f corrisponde alla frequenza di lavoro [35]. Essendo tale frequenza estremamente bassa, di ca. 1/2 Hz, le perdite nel ferro possono essere trascurate. Pertanto risulta che

$$P_{Loss} = P_{Cu} \tag{4.21}$$

Le perdite negli avvolgimenti possono essere calcolate mediante la ben nota formula [36]:

$$P_{Cu} = \rho_{Cu} \cdot \delta^2 \cdot V_{Cu} \tag{4.22}$$

dove ρ_{Cu} è la resistività del rame, δ è la densità di corrente nell'avvolgimento e V_{Cu} è il volume del conduttore attivo dell'avvolgimento. Ipotizzando un coefficiente di riempimento $k_f = 0.5$ si può legare il volume delle parti attive in rame al volume della cava, V_c, mediante la formula:

$$V_{Cu} = k_f \cdot V_c \tag{4.23}$$

Grazie alle equazioni 4.22 e 4.23, si può ricavare la densità di corrente degli avvolgimenti (supposta omogenea per le due bobine), una volta definite le perdite totali P_{Loss}:

$$\delta = \sqrt{\frac{P_{Loss}}{k_f \rho_{Cu}(V_{c,LV} + V_{c,RV})}} \tag{4.24}$$

Risulta possibile definire le ampere spire dei due avvolgimenti, mediante la

$$NI_k = k_f \delta A_k \tag{4.25}$$

dove A_k è la superficie trasversale della cava, calcolata in 4.11 e 4.12 e k = LV, RV. Per il calcolo del volume delle cave è stata usata la formula del volume di un solido di rotazione:

$$V_{c,k} = 2\pi \cdot R_{mc,k} \cdot A_k \tag{4.26}$$

dove $R_{mc,k}$ è il raggio medio della cava, calcolato mediante le formule approssimate:

$$R_{mc,LV} = \frac{(w_t - w_s) + w_{bk1}}{2} \tag{4.27}$$

$$R_{mc,RV} = \frac{(w_t - w_s) + w_{bk2}}{2} \tag{4.28}$$

4.2.2 Definizione degli indici di prestazione

Punto molto importante del metodo di Taguchi risulta anche la determinazione dei giusti indici di prestazione, perché saranno quelli che verranno ottimizzati. Gli indici di prestazione che si individuano sono:

- le 4 forze medie agenti sui mover nelle corse di andata e ritorno ($F^{for}_{m,LV}$, $F^{for}_{m,RV}$, $F^{bck}_{m,LV}$, $F^{bck}_{m,RV}$). Imponendo una forza media positiva, come una forza nella direzione del moto considerato, il processo di ottimizzazione mira a massimizzare questi indici;

- i due ripple in p.u. delle forze agenti sui mover destro e sinistro durante le corse di andata (r^{for}_{LV}, r^{for}_{RV}). Questi indici dovranno essere minimizzati;

- le due forze di avviamento agenti sui mover nella corsa di andata, ($F_{start,LV}$, $F_{start,RV}$). Questi indici devono essere superiori a 40 N e 20 N rispettivamente;

- il modulo[1] dei due valori di energia minima cumulativa dei mover per la corsa di ritorno, (e_{min}^{LV}, e_{min}^{RV}). Questi indici devono essere portati a 0 alla fine del processo di ottimizzazione.

4.3 Matrici ortogonali e conduzione degli esperimenti

Uno dei punti di forza di questo metodo è il fatto che consente di ottenere un grande risparmio di tempo. Infatti, una volta determinati i parametri di progetto e i valori che essi possono assumere, per condurre gli esperimenti non si andrà a testare la popolazione composta da tutte le combinazioni possibili, ma se ne sceglierà un campione mediante l'utilizzo di una matrice ortogonale, che garantisce che questo campione sia ben rappresentativo dell'intera popolazione. Il primo è l'approccio tipico del DOE (Design of Experiments), introdotto da Fisher a fine anni '20 del secolo scorso [28] il quale, dati M fattori, ciascuno dei quali può assumere N valori, o livelli, prevede la conduzione di N^M esperimenti, che, per un numero elevato di fattori, richiederebbe un numero enorme di esperimenti. Nel caso analizzato un esperimento corrisponde alla conduzione di 32 analisi FEM, per determinare i profili di forza dell'attuatore, da cui sono ricavati gli indici di prestazione. Per 17 fattori, ognuno dei quali variabile su 5 livelli, sarebbe necessario condurre 762 939 453 152 esperimenti. L'utilizzo di una matrice ortogonale consente invece di ridurre notevolmente il numero di esperimenti da dover condurre, garantendo un'adeguata rappresentatività del campione selezionato.

Una matrice ortogonale (OA), secondo la definizione di Taguchi, è una tabella le cui voci sono i numeri naturali da 1 a N, indicanti i livelli delle variabili; ogni colonna è rappresentativa di un fattore, o parametro iniziale di progetto, e ogni riga corrisponde a un esperimento da condurre impostando ogni variabile al livello indicato nella relativa colonna. Le colonne di tale tabella sono costruite in modo che:

- all'interno di ogni colonna ci sia lo stesso numero di ripetizioni per ogni livello;
- le colonne siano bilanciate tra di loro in modo da formare un numero uguale di tutte le possibili combinazioni tra i livelli.

Per capire meglio tale concetto si fa riferimento alla matrice ortogonale L_8, presentata in figura 4.2, che si riferisce alla conduzione di esperimenti con 8 fattori, ognuno dei quali variabili su 2 livelli. Come si può notare, in ogni colonna si trova un equo numero di 1 e 2 (quattro per specie, prima condizione soddisfatta) e, prendendo due colonne qualsiasi, ognuna delle quattro possibili combinazioni tra il numero 1 e 2 viene ripetuta due volte (ovvero le coppie ordinate (1;1), (1;2), (2;1), (2;2)), soddisfacendo il secondo vincolo. Questa è una matrice ortogonale

[1] Si utilizza il modulo per una mera questione grafica: un valore positivo diventa così indice di un mover che non riesce a tornare in posizione iniziale, mentre un valore pari a 0 è indice di un mover che ci riesce.

che prevede la conduzione di 8 esperimenti invece che dei $2^7 = 128$ esperimenti previsti dall'approccio convenzionale del DOE. Il primo esperimento prevede il settaggio di tutti i parametri al livello 1, il secondo dei parametri da A a C al livello 1 e gli altri al livello 2 e così via.

TRIAL \ FACTOR	A	B	C	D	E	F	G
1	1	1	1	1	1	1	1
2	1	1	1	2	2	2	2
3	1	2	2	1	1	2	2
4	1	2	2	2	2	1	1
5	2	1	2	1	2	1	2
6	2	1	2	2	1	2	1
7	2	2	1	1	2	2	1
8	2	2	1	2	1	1	2

Figura 4.2: Matrice ortogonale L_8

Nel caso dell'attuatore MM-LOA è stata definita una matrice ortogonale composta da 17 fattori, ognuno dei quali variabile su 5 livelli e 200 righe. Di queste 200 configurazioni, 12 sono da scartare in quanto non fisicamente realizzabili, perché le altezze di fondo cava h_{c22} e/o h_{c12} calcolate con le formule presentate in 4.2.1 assumono un valore negativo. I livelli di ogni variabile sono stati definiti nel paragrafo 4.2.1 e sono stati analizzati due scenari: uno imponendo $P_{Loss} = 5$ W e uno con $P_{Loss} = 10$ W, in modo tale da rendere più evidente la saturazione del ferro. La tecnica di switching di polarità di corrente, per questa analisi è quella di singolo switching, in quanto più semplice.

4.4 Risultati

Una volta condotti i 188 esperimenti per entrambi gli scenari di perdite si sono analizzati i risultati ottenuti come presentato in questa sezione. Per la generica variabile x_i, i 188 esperimenti condotti sono stati divisi in N gruppi, corrispondenti al numero di livelli utilizzati per la variabile x_i, 5 nel nostro caso; il gruppo g, con $g = [1, 2, ...N]$ è costituito degli esperimenti condotti con la variabile x_i settata al livello g. A questo punto, per ogni gruppo viene calcolata la media di ogni indice di prestazione, in modo tale da poter ricostruire l'andamento dell'indice di prestazione al variare del parametro di progetto x_i. Questo processo viene poi ripetuto per tutte le 17 variabili di progetto, ottenendo i profili di tutti i 10 indici di prestazione al variare del livello di ognuna delle 17 variabili x_i.

Si procede ora alla determinazione delle variabili meno influenti sulle prestazioni dell'attuatore, in modo da ridurre la dimensionalità del problema. Si vuole ottenere un indice unico che possa riassumere l'influenza della generica variabile x_i su tutti gli indici di prestazione e che possa essere utilizzato per confrontare tra di loro due variabili diverse. Siccome questo indice deve unire indici di prestazioni diversi, con diversi range di variabilità e unità di misura, è necessario

normalizzare gli indici. Per questo passaggio è stata adottata la normalizzazione tipica della "GRA Analysis" spesso utilizzata insieme al metodo di Taguchi, nel caso in cui si voglia riassumere in un unico indice più variabili di risposta i cui range di variazione possono essere molto diversi, di cui si può trovare riscontro in [32, 37].

Le normalizzazioni proposte sono due, in base al fatto che si voglia tendere alla massimizzazione (utilizzo della normalizzazione "larger the better") o alla minimizzazione dell'indice (utilizzo della normalizzazione "smaller the better"). La prima viene utilizzata per le forze medie e le forze di avviamento, mentre la seconda per i ripple e le energie minime; sia $y_{k,i}$ il generico indice di prestazione k dell'esperimento i. L'indice $y_{k,i}$ viene così trasformato nell'indice normalizzato $y'_{k,i}$ secondo le seguenti formule:

- "Larger the better":

$$y'_{k,i} = \frac{y_{k,i} - \min_i y_{k,i}}{\max_i y_{k,i} - \min_i y_{k,i}} \cdot 100$$

- "Smaller the better"

$$y'_{k,i} = \frac{\max_i y_{k,i} - y_{k,i}}{\max_i y_{k,i} - \min_i y_{k,i}} \cdot 100$$

In questo modo il valore più elevato o più basso, a seconda di dove si vuole far tendere l'indice assume il valore 100, mentre l'estremo opposto assume il valore di 0. Tutti gli indici così normalizzati avranno valore compreso tra 0 e 100.

A questo punto, per ogni variabile, si ricalcolano gli andamenti degli indici così normalizzati, come indicato in precedenza. Si ottengono quindi 5 valori per ogni indice normalizzato, per ogni variabile, indicanti come i primi varino al cambiare della variabile analizzata. Per ogni indice normalizzato viene calcolato lo scarto quadratico medio (σ) dei 5 valori. Confrontando il σ dello stesso indice, per due variabili differenti, la variabile che presenta un σ più elevato è quella che produce una variabilità maggiore su quell'indice, quindi è anche la più influente sull'indice. Per riunire poi i diversi indicatori in uno unico, $\bar{\sigma}$, ne viene calcolata la media pesata utilizzando i pesi mostrati in tabella 4.2. I valori di $\bar{\sigma}$ così ottenuto, per ogni variabile e per i due scenari (perdite Joule pari a 5 W e 10 W) sono visualizzabili in tabella 4.3 e nel grafico a barre di figura 4.3. Valori di $\bar{\sigma}$ più elevati sono indicativi di una variabile che ha una grande influenza sulle prestazioni dell'attuatore.

Tabella 4.2: Pesi utilizzati nel calcolo della media pesata degli scarti quadratici medi per la realizzazione di un indice unico.

Indice	$F^{for}_{m,LV}$	$F^{for}_{m,RV}$	$F^{bck}_{m,LV}$	$F^{bck}_{m,RV}$	r^{for}_{LV}	r^{for}_{RV}	F^{start}_{LV}	F^{start}_{RV}	en^{min}_{LV}	en^{min}_{RV}
Peso	0.12	0.12	0.08	0.08	0.05	0.05	0.10	0.10	0.15	0.15

Si può notare come le variabili più influenti nello scenario a 5 W siano le più influenti anche nell'altro scenario.

Tabella 4.3: Valore di $\bar{\sigma}$, indice unico indicante l'influenza delle variabili sulle prestazioni dell'attuatore. Un valore di $\bar{\sigma}$ più elevato è indice di una variabile molto influente sulle prestazioni. Le variabili sono ordinate in ordine decrescente per il valore di $\bar{\sigma}$ per lo scenario con 5 W di perdite Joule.

Variabile	**Scenario 5W**	**Scenario 10W**
h_m	13.71	11.72
z_{slot1}	7.10	8.26
w_m	7.02	6.34
z_{slot2}	3.85	4.19
h_{sd1}	3.62	4.03
w_{bk1}	3.54	3.95
Δ_{rc1}	3.28	3.26
Δ_{rc2}	3.20	3.35
h_{s2}	3.01	3.02
h_{sd3}	2.82	3.26
h_{s1}	2.63	2.98
w_y	2.55	3.00
α_{c2}	2.54	2.51
α_{c1}	2.54	2.62
h_{c11}	2.38	2.59
w_{sd}	2.21	2.44
w_{bk2}	1.70	1.71

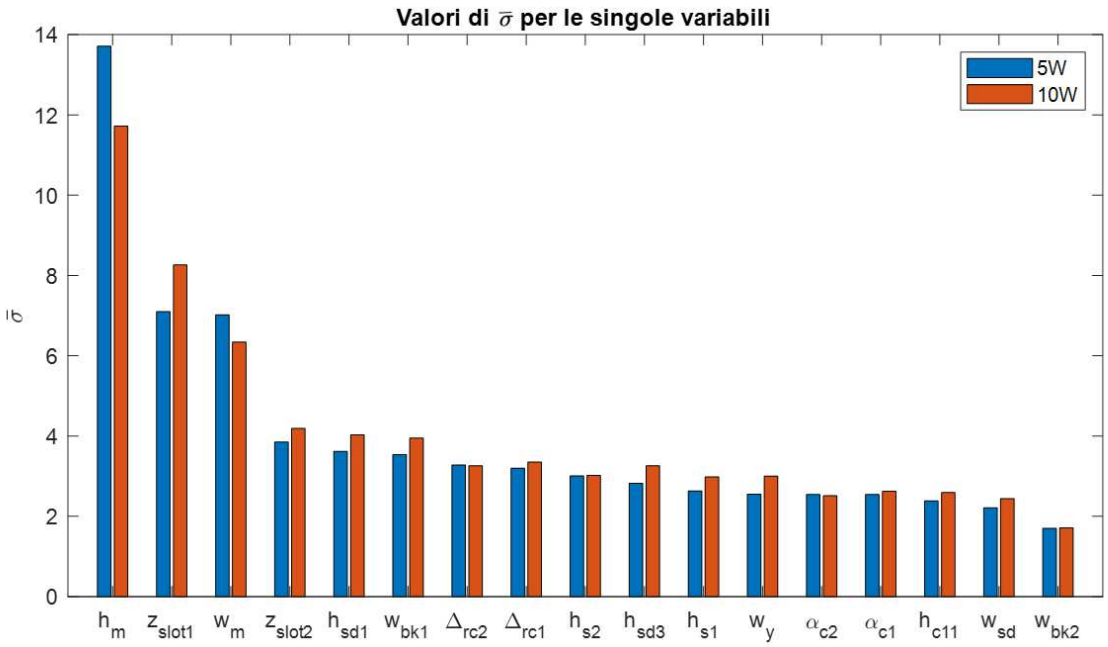


Figura 4.3: Valori di $\bar{\sigma}$ per le diverse variabili di progetto considerate per entrambi gli scenari di 5W e 10W di perdite Joule

4.5 Riduzione della dimensionalità

4.5.1 Scelta delle variabili da eliminare

Si procede ora alla selezione delle variabili meno influenti sulle prestazioni dell'attuatore e alla determinazione del loro valore al fine di ridurre la dimensionalità del problema. Vengono eliminate le variabili con un indice $\bar{\sigma}$ più basso, considerando principalmente quello derivante dallo scenario a 5W, quindi w_{bk2}, w_{sd}, h_{c11} w_y, h_{s1} e h_{sd3}. In questa lista non sono incluse le variabili α_{c1} e α_{c2} in quanto sono importanti al fine della determinazione della geometria della cianfrinatura che è essenziale per poter far tornare il mover nella sua posizione iniziale. La figura 4.4 mostra l'andamento del valore assoluto dell'energia minima cumulativa dei due mover al variare di queste due variabili con i valori calcolati come descritto nel paragrafo precedente; si può notare come α_{c1} e α_{c2} abbiano un ruolo importante nella variazione di questo indice. Si preferisce quindi mantenere queste due variabili nel successivo processo di ottimizzazione per avere più gradi di libertà. Dai grafici di figura 4.4 è già intuibile che il processo di ottimizzazione porterà alla scelta di valori più alti per queste due variabili al fine di minimizzare gli indici sull'energia minima cumulativa.

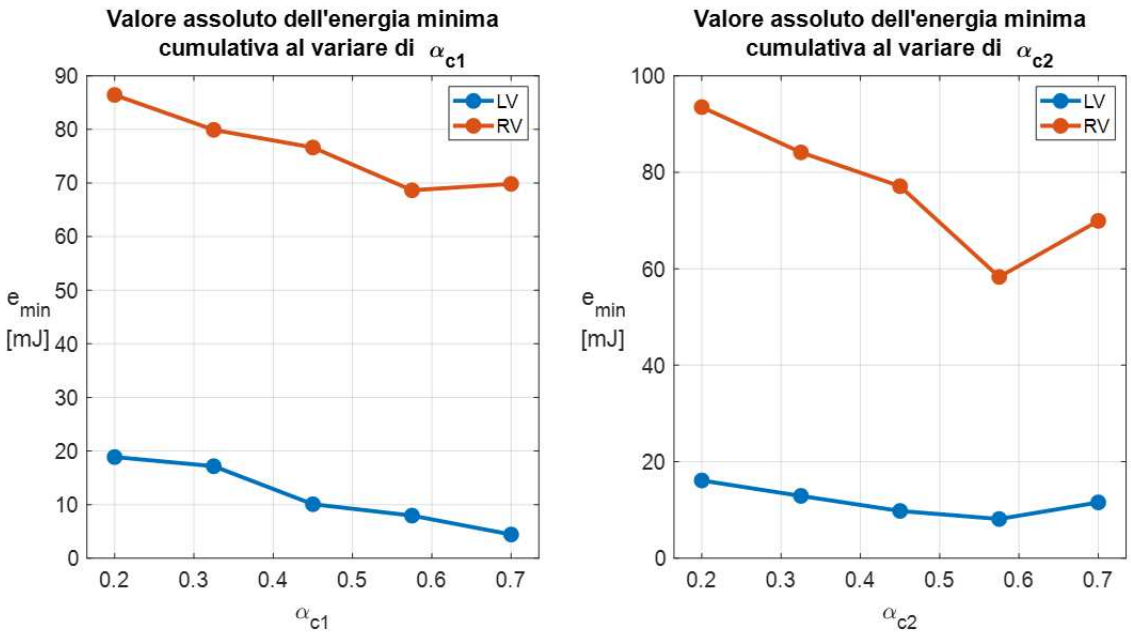


Figura 4.4: Valore assoluto dell'energia minima cumulativa per la corsa di ritorno dei due mover al variare di α_{c1} e α_{c2}. Si nota come per minimizzare quest'indice, si tenderà a stare su valori più elevati delle due variabili

4.5.2 Selezione del valore delle variabili eliminate

Una volta selezionate le variabili da eliminare nel processo di ottimizzazione è necessario impostare il loro valore per procedere con le successive analisi. Nel seguito si analizzeranno le variabili eliminate dalla meno influente alla più influente

secondo il parametro $\bar{\sigma}$, rappresentato in figura 4.3. Si farà riferimento principalmente allo scenario 5 W, considerando che il profilo degli indici di prestazione, non differisce di molto tra i due scenari. Ciò che cambia sono invece i valori che questi indici assumono.

- w_{bk2}: questa è la variabile meno influente di tutte sulle prestazioni dell'attuatore. Viene quindi impostata al valore minimo in modo tale da avere una cava più ampia per la bobina di destra, potendo aumentare così le ampere-spire e di conseguenza la f.m.m.

$$w_{bk2} = 10 \text{ mm}$$

- w_{sd}: come si può notare dalla figura 4.5 questo parametro ha poca influenza sulle forze medie. Un parametro molto importante è quello sull'energia minima. Si decide pertanto di impostare questa variabile in modo da minimizzare l'energia minima cumulativa del mover di destra che è quella più critica. Si imposta pertanto

$$w_{sd} = 4 \text{ mm}$$

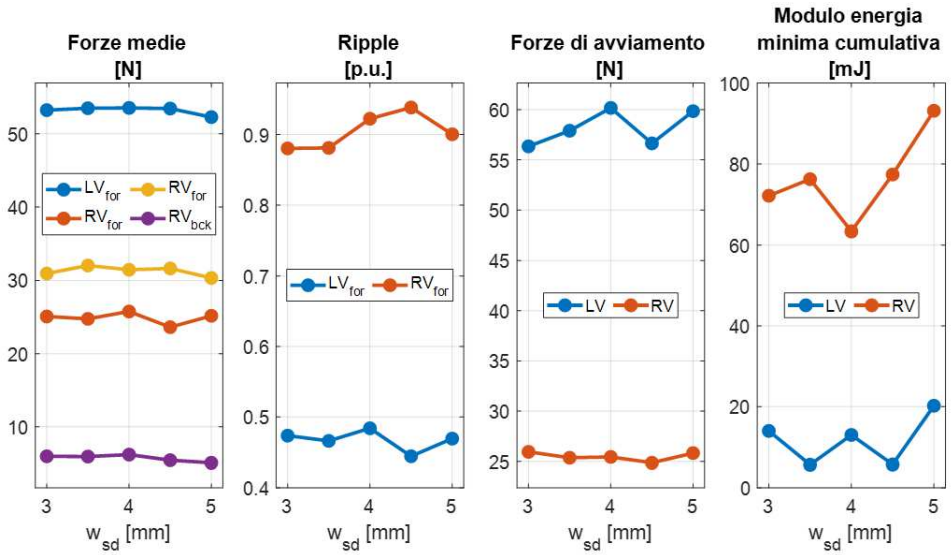


Figura 4.5: Andamenti degli indici di prestazione al variare di w_{sd} nello scenario $P_{Loss=}$ 5 W.

- h_{c11}: si seleziona

$$h_{c11} = 23 \text{ mm}$$

Per tale valore infatti, si trae vantaggio per quasi tutti gli indici di prestazione, soprattutto per l'energia minima cumulativa del mover di destra, come si evince dalla figura 4.6.

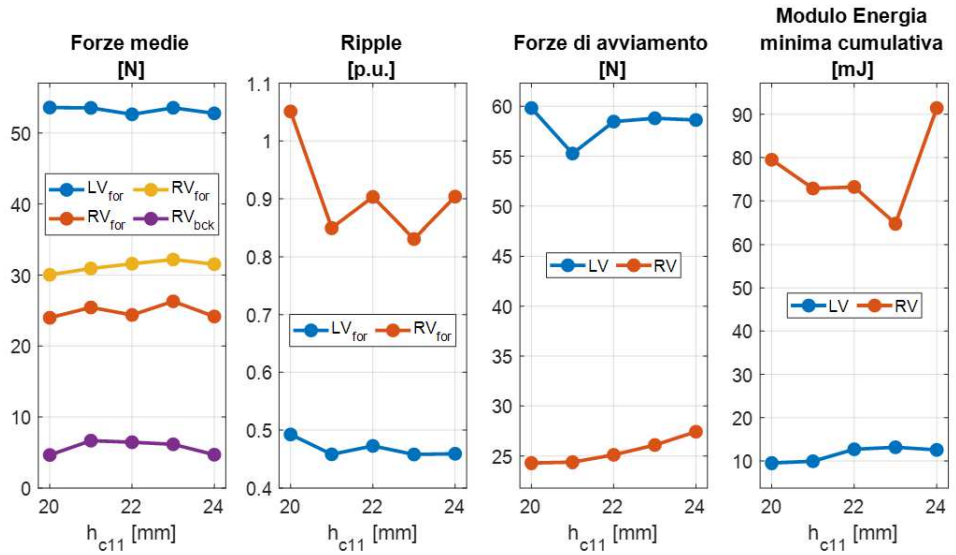


Figura 4.6: Andamenti degli indici di prestazione al variare di h_{c11} nello scenario $P_{Loss=}$ 5 W.

- w_y: questa variabile viene impostata a

$$w_y = 2.5 \text{ mm}$$

Infatti questo valore massimizza le forze medie e di avviamento, minimizzando i ripple delle forze nelle corse di andata, come si può vedere in figura 4.7. Non c'è un'influenza marcata di tale parametro sulle energie minime cumulative.

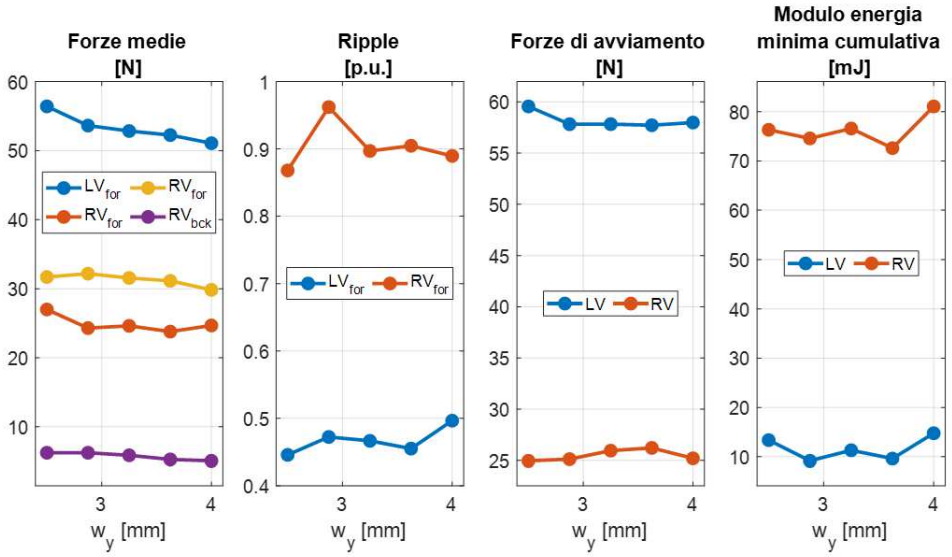


Figura 4.7: Andamenti degli indici di prestazione al variare di w_y nello scenario $P_{Loss=}$ 5 W.

- h_{s1}: questa variabile viene impostata a

$$h_{s1} = 2.75 \text{ mm}$$

Questo valore risulta essere un valore benefico per quanto riguarda i ripple e l'energia minima cumulativa del mover di destra, come indicato in figura 4.8.

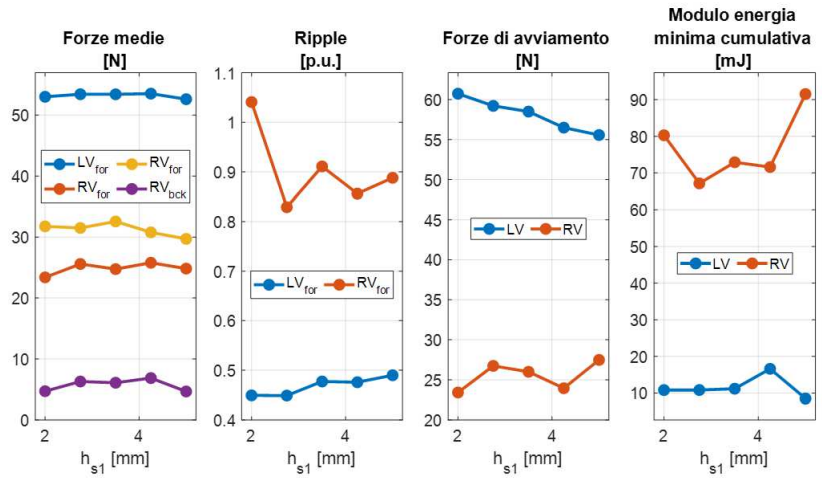


Figura 4.8: Andamenti degli indici di prestazione al variare di h_{s1} nello scenario $P_{Loss=}$ 5 W.

- h_{sd3}: questa variabile viene settata al valore di

$$h_{sd3} = 4 \text{ mm}$$

Tale scelta va a beneficio di tutti gli indici di prestazione, come si può facilmente vedere in figura 4.9

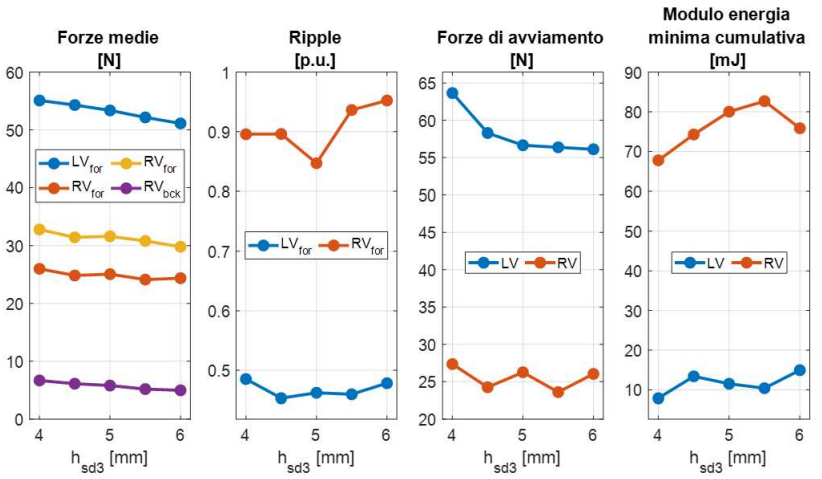


Figura 4.9: Andamenti degli indici di prestazione al variare di h_{sd3} nello scenario $P_{Loss=}$ 5 W.

Si è così riusciti a ridurre la dimensionalità del problema, diminuendo il numero di variabili da 17 a 11, selezionando per le variabili eliminate il valore più conveniente al fine di ottimizzare le prestazioni dell'attuatore.

Capitolo 5

Predizione degli indici

Il processo di ottimizzazione che si vuole portare avanti in questo lavoro è necessariamente un processo di ottimizzazione multiobiettivo. Risulta infatti necessario ottimizzare vari indici e un miglioramento di un indice può portare al peggioramento di un altro. Un esempio è il fatto che per aumentare la forza media dei mover nelle corse di andata si può agire cercando una configurazione che aumenti l'effetto del cogging [5]. Questo effetto, benefico nella corsa di andata, si oppone al movimento dei mover durante la corsa di ritorno, peggiorando le relative forze medie e i valori dell'energia minima cumulativa. Il fine dell'ottimizzazione multiobiettivo è l'individuazione del fronte di Pareto, definito come l'insieme di punti tali per cui non è possibile migliorare uno degli indici di prestazione senza peggiorarne un altro. Come verrà spiegato meglio nel capitolo 6 gli algoritmi disponibili[1] per individuare il fronte si basano sulla valutazione degli indici di prestazioni a partire da diverse combinazioni dei parametri di progetto. Questo processo può richiedere molte valutazioni degli indici e, considerato che per ogni valutazione sono richiesti circa 4 minuti di calcolo, tale processo può richiedere un tempo di calcolo molto elevato. Lo scopo di questo capitolo è quindi quello di costruire modelli di regressione che siano in grado di predire gli indici di prestazione a partire dalle 11 variabili di progetto individuate nel capitolo 4. Per questo scopo ci si è orientati verso tecniche di machine learning supervisionato. In questo modo si riesce a bypassare l'analisi FEM con tali funzioni predittive che richiedono un tempo di calcolo minimo, rendendo il processo di ottimizzazione molto più veloce.

5.1 Rifinitura dei range

Impostati i valori delle sei variabili eliminate come indicato nel capitolo 4, viene eseguita l'analisi di Taguchi imponendo $P_{Loss} = 5$ W e ampliando i range di variazione delle 11 variabili, in modo da non perdere eventuali configurazioni interessanti. La matrice ortogonale utilizzata è costruita con 11 fattori, ognuno variabile su 11 livelli, per un totale di 121 righe. Le variabili sono opportunamente

[1] Come ad esempio MO-PSO (multi-objective particle swarm optimization)[38], algoritmi genetici o di ricerca diretta.

assegnate alle colonne in modo tale da ridurre il numero di configurazioni fisicamente irrealizzabili, riducendo il numero di esperimenti da condurre (o righe) a 112. Successivamente sono costruiti gli andamenti degli indici di prestazione in funzione delle variabili di progetto come descritto nel capitolo 4 dalla cui analisi viene ridefinito il range di ogni variabile, in modo da concentrare lo studio solamente alle zone reputate più interessanti. I range così ridefiniti sono indicati in tabella 5.1.

Tabella 5.1: Nuovi range delle variabili indipendenti utilizzati per l'ottimizzazione.

Variabile	Range	Variabile	Range
w_{bk1}	(9; 13.5) mm	h_{s2}	(1.5; 4) mm
h_{sd1}	(4; 6) mm	h_m	(12; 16) mm
w_m	(2.4; 4) mm	z_{s2}	(9; 13) mm
z_{s1}	(9; 14) mm	Δr_{c2}	(1.2; 3.2) mm
Δr_{c1}	(0.5; 2.5) mm	α_{c2}	(0.3; 0.8)
α_{c1}	(0.3; 0.8)		

5.2 Il machine learning

Il machine learning (ML) è una branca dell'intelligenza artificiale (AI) il cui funzionamento è schematizzato in figura 5.1: a partire da un set di dati noto, di cui si conoscono sia gli input che gli output, il ML genera, tramite un processo di addestramento, un modello che si adatti bene a questi dati. Tale modello può essere poi utilizzato per predire gli output di dati diversi da quelli di training. Questo è il fine ultimo del modello e più viene svolto bene e più il modello è generale.

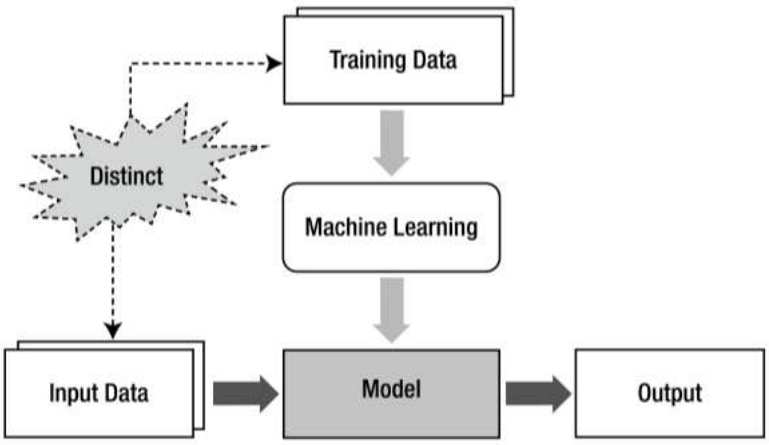


Figura 5.1: Funzione del machine learning [39]

Si possono avere dei problemi di regressione, dove lo scopo è quello di predire una variabile di output continua a partire da dei valori di input, e problemi di

classificazione, in cui lo scopo è quello di assegnare la particolare combinazione di input ad una determinata classe. Un esempio di quest'ultimo problema è quello del riconoscimento della scrittura di una persona, in cui l'input sono i pixel di una foto della lettera scritta a mano e l'output è il riconoscimento del carattere scritto, con l'assegnazione alla classe di lettera corrispondente.

Può accedere che il modello si adatti esattamente ai dati di addestramento, fornendo però scarse predizioni per dati non osservati; il modello è quindi scarsamente generalizzabile, vanificando lo scopo per cui è stato costruito. Quando ciò avviene si parla di overfitting. Un esempio di questo problema è visualizzabile in figura 5.2 in cui viene richiesto di trovare un modello che classifichi i dati in base alle loro coordinate x e y. I dati di training sono visualizzati nel grafico, la forma rappresenta la loro classe, ovvero l'output, la linea rossa rappresenta il modello e la loro posizione nello spazio costituisce l'input. Il modello a) è quello più appropriato per descrivere il fenomeno, mentre il modello b), nonostante si adatti perfettamente ai dati di training può sbagliare facilmente con dei nuovi dati, soprattutto se vicino alla frontiera.

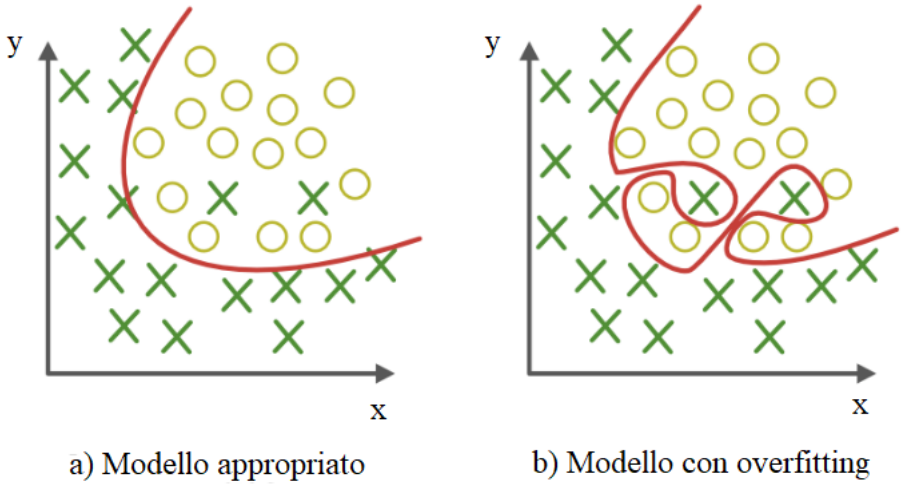


Figura 5.2: Interpretazione grafica dell'overfitting: il modello **(a)** è un buon modello di classificazione, mentre il modello **(b)** soffre del problema di overfitting [40].

L'addestramento del modello segue un processo iterativo, adattandosi sempre di più ai dati di training. Ogni dato di addestramento è composto dalle variabili di input e la relativa variabile di output nota; per avere un addestramento efficiente, i dati di training vanno divisi in 3 parti o set:

- Training set: questi sono i dati su cui il modello viene effettivamente addestrato.
- Set di validazione: questi sono i dati, diversi dal training set, su cui si valutano le prestazioni del modello durante l'addestramente e viene introdotto per evitare l'overfitting. Se infatti, durante l'addestramento, si riscontra un peggioramento per più iterazioni successive delle prestazioni su questo set, nonostante quelle sul trainig set siano in miglioramento, si sta andando

incontro ad overfitting. L'addestramento viene quindi fermato e il modello finale è quello precedente al peggioramento delle prestazioni sul set di validazione.

- Test set: questi sono i dati utilizzati a modello addestrato per valutarne le prestazioni e compararlo con altri modelli. Questi dati non vengono mai forniti in fase di addestramento ed è importante che ciò avvenga in quanto assicura che le valutazioni delle prestazioni siano realistiche e rappresentative della capacità del modello di generalizzare su nuovi dati.

Per la costruzione dei modelli si è sfruttato il pacchetto "Machine Learning and Deep Learning" fornito da MatLab, il quale è in grado di costruire modelli predittivi, sia per problemi di regressione che di classificazione, come reti neurali (ANN), Gaussian Process Regression (GPR), Support Vector Machine (SVM), modelli di regressione lineare, ecc. Tra tutti i modelli disponibili ci si è orientati verso la GPR per i problemi di regressione e le reti neurali per quelli di classificazione; per questa ragione in questa sezione verranno presentati questi modelli.

5.2.1 Reti Neurali (ANN)

Quando si apprende qualcosa il cervello umano immagazzina l'informazione; anche un computer può immagazzinare un'informazione utilizzando la sua memoria interna. Nonostante i due processi portino allo stesso risultato, sono molto diversi tra di loro: il computer immagazzina l'informazione in specifiche locazioni di memoria, mentre il cervello immagazzina le informazioni modificando le associazioni tra i neuroni. Il neurone non ha infatti la capacità di immagazzinare informazioni, ma può semplicemente trasmettere il segnale elettrico che gli arriva ad un altro neurone.

Le reti neurali (ANN) sono reti costruite esattamente su questo principio [39]. Una ANN è costituita da diversi nodi, chiamati neuroni, ordinati su più livelli. La struttura di un neurone è visualizzabile in figura 5.3: diversi segnali di input, x_i, dopo essere stati moltiplicati per un peso, ω_i, confluiscono all'interno del neurone

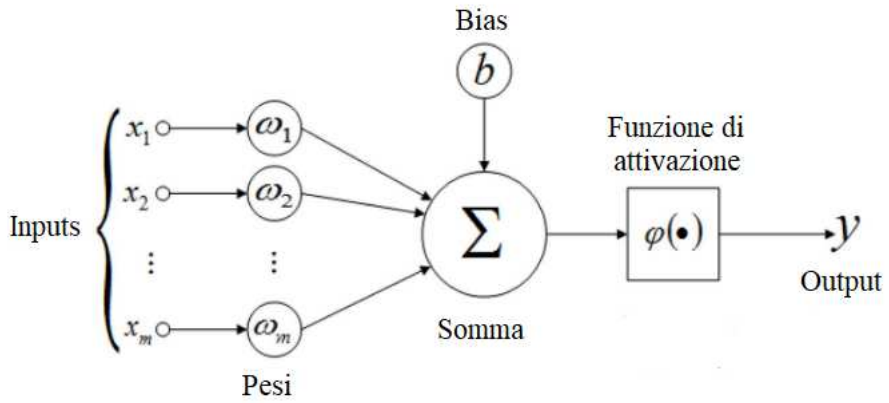


Figura 5.3: Struttura di un neurone di una ANN.

dove vengono sommati tra di loro, insieme ad un valore di bias, b, rappresentativo del potenziale di attivazione del neurone. Il valore così ottenuto passa attraverso una funzione di attivazione ϕ per produrre in uscita l'output del neurone y.

I neuroni vengono organizzati su più livelli, come mostrato in figura 5.4, in modo da formare una rete in cui ogni neurone di un livello riceve in input i segnali di output dei neuroni del livello precedente e trasmette il suo output ai neuroni del livello successivo. Si può riconoscere il livello di input, dove si inseriscono le variabili di input e il livello di output, ovvero l'uscita della rete. Se sono presenti anche dei livelli nascosti tra quello di input e di output si parla di rete neurale profonda (o deep neural network). Questo tipo di reti neurali offre prestazioni superiori rispetto a quelle a singolo strato, ma richiedono algoritmi di addestramento più sofisticati, motivo per cui hanno subito solo recentemente un grande sviluppo a seguito dell'introduzione di questi algoritmi.

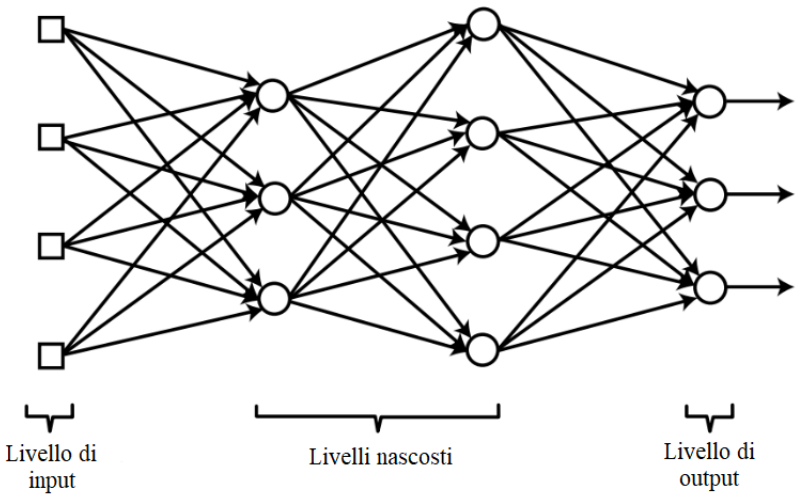


Figura 5.4: Struttura di una rete neurale profonda [39].

Il processo di addestramento di una rete neurale è basato su un processo iterativo, come indicato in figura 5.5: a partire dal set di training si calcolano gli output della rete, che vengono confrontati con gli output corretti calcolando l'errore; sulla base di questo errore vengono modificati i pesi della rete (che rappresentano le forza delle connessioni tra i neuroni). Il processo continua fino al raggiungimento di un errore inferiore ad una certa soglia o fino a quando le prestazioni sul set di validazione non iniziano a peggiorare. La rete quindi immagazzina le informazioni nei pesi w_i, ovvero modificando la forza delle connessioni tra i vari neuroni, esattamente come fa il cervello umano.

Cruciale è la definizione dei parametri che definiscono la struttura della rete, chiamati anche iperparametri. Essi sono il numero di strati nascosti, il numero di neuroni di ogni strato e la relativa funzione di attivazione. In particolare quest'ultima è importante che sia non lineare in modo da cogliere possibili non linearità nei dati. Le reti neurali possono essere utilizzate per problemi di regressione e di classificazione. Per i problemi di classificazione il livello di output deve presentare tanti neuroni quante sono le classi in cui possono essere classificati i dati di input,

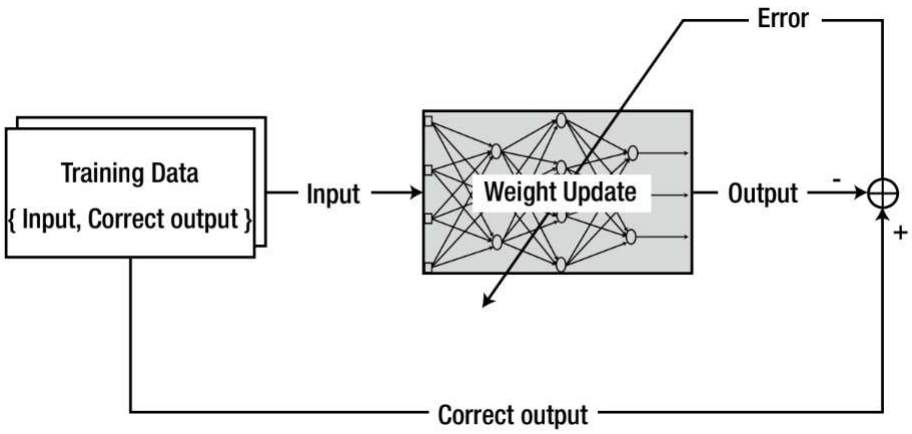


Figura 5.5: Processo iterativo di apprendimento di una rete neurale[39].

in modo da creare una corrispondenza tra i neuroni di questo strato e le classi. Risulta poi importante che la funzione di attivazione di tale strato sia la funzione softmax, espressa nella formula 5.1, in cui K è in numero di neuroni di output.

$$\phi(z)_i = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}} \tag{5.1}$$

In questo modo l'output di ogni neurone corrisponde alla probabilità con cui il dato di input appartiene alla classe individuata da tale neurone. Per una trattazione più approfondita delle reti neurali si rimanda a testi quali [39].

5.2.2 Gaussian Process Regression (GPR)

Il Gaussian Process Regression o GPR è un metodo di modellazione statistica utilizzato per affrontare problemi di regressione e si inserisce tra le tecniche di machine learning supervisionato. Un processo gaussiano (GP) è un processo stocastico, o una distribuzione tra le funzioni. Quando si modella una funzione come un processo Gaussiano si fa l'ipotesi che ogni numero di punti campionato formi una distribuzione normale multivariata tra tutte le funzioni possibili. In altre parole, un GP è una generalizzazione di una distribuzione normale a un numero infinito di variabili casuali ed è completamente definito dalla sua media e dalla sua funzione di covarianza.

A differenza di altri approcci di regressione in cui i dati vengono approssimati con una funzione parametrica specifica, come una retta (regressione lineare) o una curva polinomiale (regressione polinomiale), il GPR utilizza un processo gaussiano come modello probabilistico per la regressione. Questo significa che, invece di cercare una singola funzione deterministica che approssimi i dati, cerchiamo una distribuzione di probabilità sulla funzione stessa. Nel GPR, si assume quindi che i dati siano generati da un processo gaussiano, il che significa che ogni possibile valore di output è descritto da una distribuzione gaussiana. Punto fondamentale del GPR è la determinazione della tipologia di kernel (o funzione di covarianza) che definisce come i dati siano correlati tra loro nello spazio delle caratteristiche.

Una volta selezionato il tipo di kernel è necessario impostare gli iperparametri ad esso associati, come la varianza σ^2 e la lunghezza caratteristica l o scala. Utilizzando il kernel e gli iperparametri scelti si calcolano le medie e le covarianze tra i dati di addestramento. Viene così definita la distribuzione di probabilità delle funzioni possibili, rendendo possibile effettuare previsioni. Si testa poi la bontà del modello sul set di validazione, grazie metriche quali l'errore quadratico medio o MSE sulla base di cui si deciderà se modificare gli iperparametri del modello al fine di ottenere delle predizioni più accurate. Il set di validazione serve, così come per le reti neurali, per prevenire l'overfitting, che può verificarsi se le prestazioni sul set di addestramento sono molto più buone di quelle sul set di validazione. Per una trattazione più approfondita del GPR si rimanda a testi specialistici, come [41, 42].

5.3 Costruzione dei modelli predittivi

Si può procedere ora alla costruzione dei modelli per la predizione dei parametri più importanti che normalmente verrebbero calcolati grazie all'analisi agli elementi finiti tramite il software "Ansys Maxwell". I modelli che si costruiscono non devono solamente essere in grado di predire le prestazioni della configurazione selezionata, ma anche se tale configurazione è in grado di soddisfare i vincoli posti nel capitolo 3: i valori da predire sono quindi:

- le quattro forze medie agenti sui mover nelle corse di andata e ritorno ($F^{for}_{m,LV}$, $F^{for}_{m,RV}$, $F^{bck}_{m,LV}$, $F^{bck}_{m,RV}$);
- i due ripple in p.u. delle forze agenti sui mover destro e sinistro durante le corse di andata (r^{for}_{LV}, r^{for}_{RV});
- le due forze di avviamento agenti sui mover nelle corse di andata ($F_{start,LV}$, $F_{start,RV}$);
- le due forze di avviamento agenti sui mover nelle corse di ritorno, quando i magneti si trovano a fine corsa ($F^{bck}_{start,LV}$, $F^{bck}_{start,RV}$);
- le due forze di avviamento agenti sui mover nelle corse di andata, con polarità inversa ($F^{++}_{LV}(0)$, $F^{++}_{RV}(0)$);
- il modulo tra le energie minime cumulativa delle corse di ritorno dei due mover ($E_{min} = \sqrt{e^2_{min,LV} + e^2_{min,RV}}$).

I primi 6 indici definiscono le prestazioni dell'attuatore, mentre i restanti 7 sono utilizzati per verificare il soddisfacimento dei vincoli. L'indice sull'energia minima cumulativa serve per capire se l'attuatore è in grado di riportare i mover nella loro posizione iniziale. Data l'asimmetria di tale indice[2], che può portare a problemi se si utilizza un modello regressivo, e dato che non è necessario valutare il valore

[2]Se l'attuatore riesce a riportare i mover in posizione iniziale l'indice assume valore nullo, altrimenti è positivo

esatto di e_{min}, si utilizzerà una rete di classificazione. Tale rete dovrà essere in grado di predire se l'attuatore, una volta individuata la geometria impostando le 11 variabili di progetto, è in grado di far tornare i magneti nella posizione iniziale senza l'aiuto di forze esterne dopo che sono arrivati a fine corsa.

Per gli altri indici si è invece optato per la realizzazione di un modello regressivo, in modo da determinare esattamente il valore di output in base ai valori di input. Si è inoltre scelto di costruire un modello diverso per ogni indice, in modo da ottimizzare gli iperparametri di ogni modello sull'indice che deve essere predetto.

5.4 Scenario $P_{Loss} = 5$ W

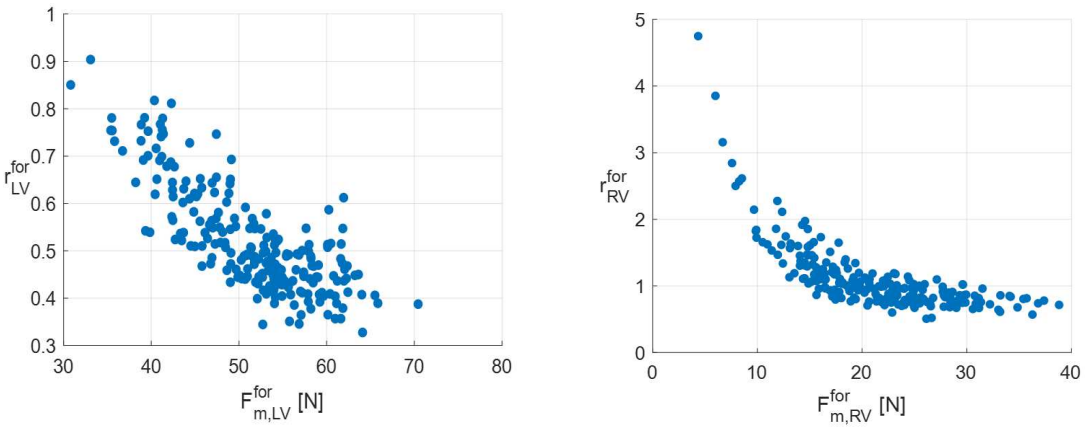
Come dati di training per i modelli dello scenario con $P_{Loss} = 5$ W viene utilizzata una matrice ortogonale con 11 fattori, ognuno dei quali variabile su 11 livelli e come range delle variabili sono utilizzati quelli rifiniti presentati in tabella 5.1. La matrice così costruita, dopo aver tolto 6 configurazioni fisicamente irrealizzabili, è formata da 115 configurazioni. Siccome nel paragrafo 5.1 era stata svolta un'altra analisi di Taguchi con $P_{Loss} = 5$ W e le 11 variabili, utilizzando dei range diversi si ritiene opportuno aggiungere i dati ottenuti mediante quell'analisi al training test, in modo da avere un set di dati più ampio. In totale il set di dati di training si compone di 227 configurazioni, provenienti da due analisi di Taguchi. Per la protezione contro l'overfitting è stata utilizzata la tecnica di validazione incrociata [43] che si compone dei seguenti passaggi:

1. il training set viene diviso in k gruppi o folds ($k = 7$ in questo caso)
2. per ogni gruppo di validazione viene addestrato un modello utilizzando i restanti k-1 gruppi e vengono valutate le performance utilizzando il gruppo di validazione
3. viene calcolata la media dell'errore di validazione tra tutti i gruppi. Questo valore è rappresentativo dell'errore del modello sul test di validazione.

Questo metodo dà una buona stima dell'accuratezza della capacità predittiva del modello addestrato sull'intero set di training. Richiede anche addestramenti multipli, ma utilizza efficientemente tutti i dati, funzionando bene su set di dati piccoli. Come dati di test sono state condotte delle analisi settando le 11 variabili ad un valore casuale compreso nei range di tabella 5.1, utilizzando il generatore di numeri random di MatLab [44]. È buona norma che le dimensioni del test-set sia di almeno il 30% del set di training, quindi vengono utilizzate 78 configurazioni.

Per la predizione degli indici sono stati utilizzati come valori di input le 11 variabili di progetto, eccetto che per i ripple delle forze nelle corse di andata. Infatti nella loro definizione è coinvolto anche il valore della forza media della relativa corsa, quindi è lecito aspettarsi una correlazione tra queste due grandezze. Tale correlazione è confermata dalla figura 5.6 in cui ogni punto rappresenta una configurazione del training set, inserita nel piano formato dalla forza media e dal relativo ripple. Per i due ripple è stato scelto quindi di utilizzare 12 variabili di input, aggiungendo le rispettive forze medie, al fine di ottenere risultati migliori.

Per utilizzare il modello che predice il ripple della forza, sarà prima necessario effettuare una predizione del valore della forza media usando la relativa rete, valore che sarà aggiunto come input alle 11 variabili di progetto.



(a) **(b)**

Figura 5.6: Ripple della forza in p.u. nelle corse di andata del mover sinistro **(a)** e destro **(b)** in funzione delle relative forze medie per le varie configurazioni del training set utilizzato con $P_{Loss} = 5$ W. Si può notare una correlazione tra i due indici.

Le metriche utilizzate per valutare le prestazioni di diversi modelli sono l'RMSE_{std} e l'RMSE, calcolati nel seguente modo:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \bar{y}_i)^2} \tag{5.2}$$

$$RMSE_{\text{std}} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\frac{y_i - \bar{y}_i}{\bar{y}_i}\right)^2} \cdot 100\% \tag{5.3}$$

dove N è il numero di dati utilizzati per il calcolo dell'indice, y_i è il valore di output fornito dal modello utilizzando come input il dato i-esimo e $\bar{y}_i$ è il valore reale dell'indice analizzato del dato $i-$esimo. L'indice RMSE viene utilizzato per confrontare due modelli diversi che devono predire lo stesso output e rappresenta l'errore medio che si compie ad ogni previsione; l'indice $RMSE_{\text{std}}$ è simile e rappresenta l'errore percentuale compiuto su ogni previsione, quindi può essere utile per confrontare fra di loro modelli con output diversi. Bisogna fare molta attenzione nell'utilizzo di quest'indice, in quanto può risultare corrotto se alcuni dati sono prossimi a 0. In tal caso un piccolo errore sulla predizione, che non è eliminabile, si riflette in un $RMSE_{\text{std}}$ molto elevato. Ecco perchè i due indici andrebbero monitorati insieme.

Vengono testati diversi modelli, ottimizzando i parametri con un algoritmo Bayesiano, grazie al "Regression Learner APP", e calcolando per ognuno le prestazioni sul test set, che indica se il modello è capace di generalizzare bene. Tra i vari modelli testati, quello che si è dimostrato più efficace nella maggior parte dei casi è risultato essere il GPR i cui iperparametri sono stati ottimizzati tramite

un algoritmo Bayesiano[45]. I modelli che presentano le migliori prestazioni in termini di RMSE sul set di test per ogni indice di prestazione sono riportati nella tabella 5.2. In tabella viene anche riportata l'impostazione degli iperparametri, indicando con BF il tipo di funzione di base, con Kf il kernel utilizzato e con SD se i dati sono stati standardizzati o meno prima dell'addestramento del modello. Per l'indice $F_{start,LV}$ il modello migliore è risultato essere una rete neurale con uno strato nascosto composto da 175 neuroni, con la funzione di attivazione sigmoide; tale modello ha ottenuto un RMSE_{std} sul test set pari al 3.49% e un RMSE di 1.53 N.

Tabella 5.2: Modelli GPR che offrono le prestazioni migliori per quanto riguarda la predizione sul test set test per ogni indice nello scenario 5 W. Sigle utilizzate: NM = nonisotropic matern; IRQ = isotropic rational quadratic; Im = isotropic matern, SE = squared exponential

INDICE	**BF**	**KERNEL**	$\boldsymbol{\sigma}$	**SD**	$\mathbf{RMSE_{std}}$ (test set)	**RMSE** (test set)
$F^{for}_{m,LV}$	Lineare	NM 5/2	$2.96 \cdot 10^{-3}$	Si	0.77%	0.34 N
$F^{for}_{m,RV}$	Lineare	NM 5/2	$1.031 \cdot 10^{-4}$	Si	1.81%	0.43 N
$F^{bck}_{m,LV}$	Zero	NM 5/2	$1.13 \cdot 10^{-3}$	Si	5.13%	1.21 N
$F^{bck}_{m,RV}$	Costante	NM 5/2	2.5891	Si	4.77%	0.36 N
r^{for}_{LV}	Zero	NM 5/2	0.995	Si	2.63%	0.00064 p.u.
r^{for}_{RV}	Lineare	IRQ	0.0165	No	5.69%	0.0086 p.u.
$F_{start,RV}$	Lineare	NM 5/2	54.77	Si	5.34%	2.03 N
$F^{bck}_{start,LV}$	Costante	SE	0.8437	Si	6.64%	2.87 N
$F^{bck}_{start,RV}$	Costante	NM 5/2	0.1017	Si	4.57%	2.2 N
$F^{++}_{LV}(0)$	Lineare	IM 5/2	0.3045	Si	6.86 %	3.4 N
$F^{++}_{RV}(0)$	Lineare	IM 5/2	0.1245	No	4.32%	3.1 N

Per quanto riguarda il modello per la predizione dell'energia minima, come già accennato in precedenza si opta per un modello di classificazione, invece che di regressione. Dopo il test di vari modelli tramite il "Classification Learner App" fornito da MatLab, il modello con le migliori prestazioni è una rete neurale con uno strato nascosto formato da 100 neuroni e 2 neuroni di output. La funzione di attivazione dello strato nascosto è la funzione sigmoide mentre quella dello strato di output è la funzione Softmax. Il primo neurone di output corrisponde alla classe delle configurazioni che soddisfano il vincolo sull'energia minima (classe positiva), mentre il secondo alla classe di quelle che non soddisfano tale vincolo (classe negativa). Dati gli 11 parametri di input, l'output della rete è costituito da due numeri, la cui somma vale 1 e ognuno rappresenta la probabilità che la configurazione costruita a partire da quelle determinate variabili di input soddisfi o meno il vincolo sull'energia minima cumulativa. L'algoritmo di addestramento utilizzato è lo "Scaled conjugate gradient backpropagation" e sono state necessarie 17 epoche di addestramento. Le prestazioni di tale rete sul set di training e di test sono riportate nelle matrici di confusione in figura 6.4. La precisione della

rete è del 93 % sul set di training e del 92 % su quello di test. Tali prestazioni sono state ritenute accettabili.

Train set		
Classe reale	positiva	negativa
positiva	90	6
negativa	8	123
Classe predetta		

Test set		
Classe reale	positiva	negativa
positiva	36	4
negativa	2	36
Classe predetta		

Figura 5.7: Matrice di confusione sul set di training e di test della rete neurale utilizzata per la predizione della classe di energia delle configurazioni.

5.5 Scenario $P_{Loss} = 10$ W

Nello scenario con $P_{Loss} = 10$ W per la determinazione del set di training viene condotta un'analisi di Taguchi con 11 fattori e 11 livelli, con le variabili comprese nei range indicati in tabella 5.1. Dopo aver eliminato 7 configurazioni fisicamente infattibili, ne sono rimaste 118 come dataset di training. Per questo scenario sono state utilizzate meno configurazioni, ma le prestazioni delle reti non differiscono di molto da quelle dello scenario a 5 W. Come dati di test sono state analizzate 49 configurazioni assegnando ad ogni variabile un valore random all'interno del suo range di variazione, come già fatto per lo scenario precedente. Come fatto in precedenza per i modelli che hanno il compito di predirre i ripple sono state fornite in ingresso le 11 variabili di progetto e la forza media relativa alla corsa di cui si vuole calcolare il ripple. Anche in questo scenario i risultati migliori sono stati ottenuti con il GPR ottimizzato con l'algoritmo Bayesiano fornito dal "Regression learner app" di MatLab. I metodi utilizzati per la costruzione e la valutazione delle prestazioni di questi modelli sono analoghi a quelli utilizzati per lo scenario precedente; i risultati dei migliori modelli con relativi iperparametri sono indicati nella tabella 5.3.

Per quanto riguarda la predizione del vincolo sull'energia sono state usate le classi utilizzate in precedenza: positiva se $E_{min} = 0$, negativa altrimenti. Per la costruzione di tale modello predittivo è stata utilizzata la support vector machine (SVM), in quanto ha fornito prestazioni migliori rispetto alle ANN. La SVM è una tecnica di machine learning che, dato il set di training, cerca di trovare un iperpiano che separi le diverse classi tra di loro. Tale iperpiano, chiamato "iperpiano di separazione ottimale" viene selezionato in modo da massimizzare la distanza degli elementi delle classi da tale iperpiano. Per tale modello è stato impostato un kernel polinomiale di secondo grado. In figura 5.8 è riportata la confusion matrix di tale modello sul set di test e di training. L'accuratezza del modello è del 94.9 % sul set di test e dell'81.6 % su quello di training

Tabella 5.3: Modelli GPR che garantiscono le prestazioni migliori per quanto riguarda la predizione sul test set per ogni indice nello scenario a 10 W di perdite Joule. Sigle utilizzate: NM = nonisotropic matern; RQ = rational quadratic; IM = isotropic matern, SE = squared exponential

INDICE	BF	KERNEL	σ	SD	$\mathbf{RMSE_{std}}$ (test set)	RMSE (test set)
$F^{for}_{m,LV}$	Zero	NM 5/2	$5.53{\cdot}10^{-4}$	Si	0.26%	0.65 N
$F^{for}_{m,RV}$	Zero	NM 5/2	2.50	No	3.60%	0.45 N
$F^{bck}_{m,LV}$	Costante	NM 5/2	34.54	Si	9.57%	2.1 N
$F^{bck}_{m,RV}$	Zero	NM 3/2	50.22	No	8.17%	0.87 N
r^{for}_{LV}	Zero	NM 3/2	0.806	Si	3.32%	0.0076 p.u.
r^{for}_{RV}	Lineare	NM 5/2	0.0139	No	2.06%	0.025 p.u.
$F_{start,LV}$	Costante	NM 5/2	0.0239	Si	13.6%	2.7 N
$F_{start,RV}$	Lineare	RQ	63.1	Si	5.91%	1.9 N
$F^{bck}_{start,LV}$	Lineare	SE	5.69	No	0.57%	1.2N
$F^{bck}_{start,RV}$	Lineare	NM 5/2	96.13	No	12.3%	1.9 N
$F^{++}_{LV}(0)$	Zero	IM 5/2	0.33	Si	10.56%	2.3 N
$F^{++}_{RV}(0)$	Costante	IM 5/2	$1.00 \cdot 10^{-4}$	No	3.94%	1.4 N

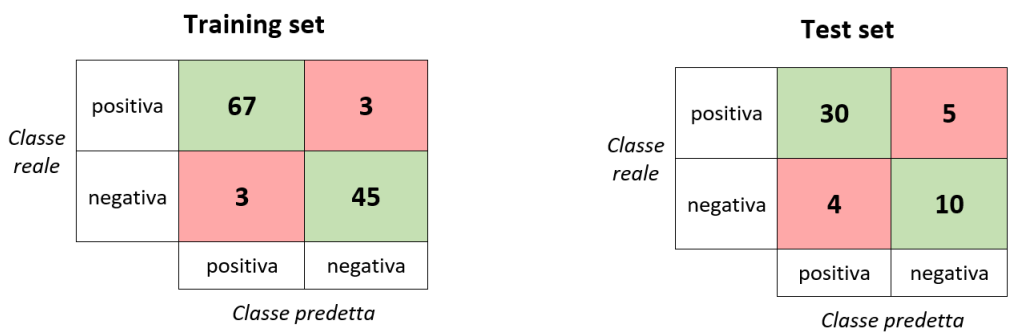


Figura 5.8: Matrice di confusione sul set di training e di test del modello SVM utilizzato per la predizione della classe di energia delle configurazioni nello scenario $P_{Loss} = 10$ W.

5.6 Conclusioni

Le prestazioni dei modelli costruiti sono ritenute soddisfacenti, soprattutto per quel che riguarda gli indici principali, ovvero le forze medie di andata, i relativi ripple e la classificazione sull'energia.

Si può notare come per il problema di regressione il GPR fornisca risultati migliori rispetto alle reti neurali, in quanto il primo riesce a predire bene la variabile di output a partire da un set di dati più limitato, mentre per le ANN è generalmente necessario un set di dati più ampio. Un esempio di tale confronto è visibile in figura 5.9 in cui sono comparate le prestazioni sul test set del GPR e della ANN. La variabile considerata è la forza media della corsa di andata del

mover sinistro nello scenario a 5 W; entrambi sono stati ottimizzati con l'algoritmo Bayesiano grazie al "Regression Learner App" di MatLab ed è visibile a colpo d'occhio come il GPR fornisca previsioni migliori.

Le prestazioni peggiori dei modelli si ottengono quando si vuole predire un valore puntuale della forza, come le forze di avviamento, mentre per quanto riguarda valori integrali, come le forze medie, si ottengono prestazioni migliori.

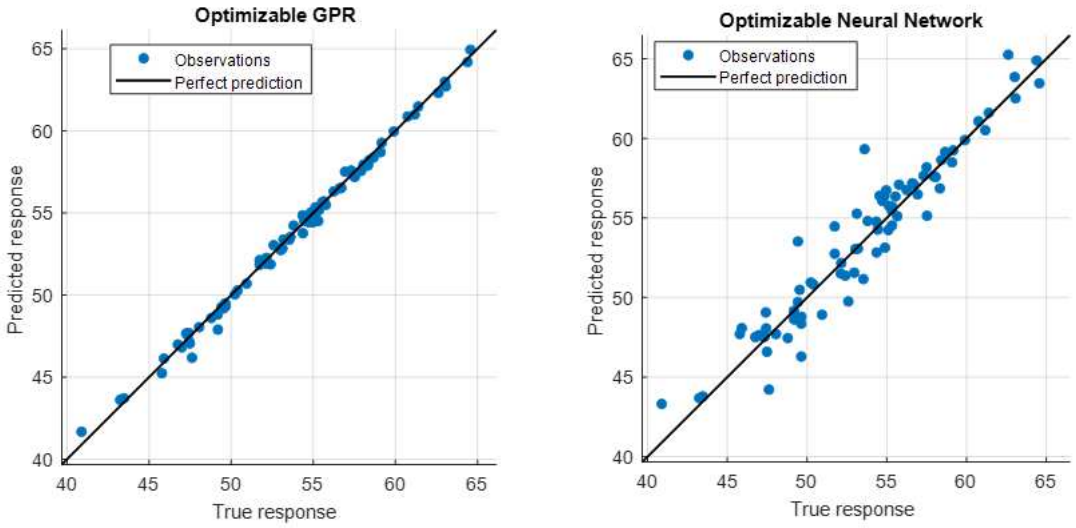


Figura 5.9: Conftonto tra le prestazioni sul test set del GPR e una ANN per l'indice $F_{m,LV}^{for}$ nello scenario $P_{Loss} = 5$ W. Entrambi i modelli sono stati ottimizzati mediante algoritmo Bayesiano con il "Regression Learner App" di MatLab". Il GPR fornisce prestazioni superiori rispetto alla ANN.

Capitolo 6

Ottimizzazione multi-obiettivo

In questo capitolo viene portato a termine il processo di ottimizzazione, obiettivo ultimo di questo lavoro. I modelli predittivi costruiti nel capitolo 5 vengono ora utilizzati in un algoritmo di ottimizzazione multi-obiettivo al posto dell'analisi agli elementi finiti, in modo da ridurre i tempi di calcolo dell'algoritmo stesso. Prima però si ritiene opportuno fornire qualche nozione riguardante l'ottimizzazione multi-obiettivo

6.1 L'ottimizzazione

Un problema di ottimizzazione mono-obiettivo è un problema in cui, data una funzione, chiamata anche funzione obiettivo,

$$f : A \subset \mathbb{R}^n \to B \subset \mathbb{R}, \quad \mathbf{x} \mapsto f(\mathbf{x})$$

viene richiesto di trovare il valore x^* tale per cui $f(\mathbf{x}^*)$ sia minimo. Questo è un problema di semplice risoluzione, ma un problema reale è difficilmente riconducibile a tale impostazione. In un problema reale è più frequentemente richiesta la contemporanea minimizzazione di più funzioni obiettivo.

Un problema in cui si hanno m funzioni obiettivo $f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_m(\mathbf{x})$ e si vuole arrivare alla loro minimizzazione simultanea è un problema di ottimizzazione multi-obiettivo. Tale problema è solitamente definito in termini matematici come in 6.1:

$$\min_{\mathbf{x} \in A} \mathbf{F}(\mathbf{x}) = \min_{\mathbf{x} \in A} [f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_m(\mathbf{x})] \tag{6.1}$$

dove $A \subset \mathbb{R}^n$ è definito lo spazio delle decisioni a cui appartiene $\mathbf{x}$ ed è solitamente limitato, mentre lo spazio $B \subset \mathbb{R}^m$ è definito come spazio delle funzioni obiettivo ed è identificato in termini matematici come

$$B = \{y \in \mathbb{R}^m : y = f(\mathbf{x}), \mathbf{x} \in A\} \tag{6.2}$$

La relazione tra i due spazi è visibile in figura 6.1 dove $n = m = 2$. In un'ottica di ottimizzazione multi-obiettivo la soluzione può non essere univoca: infatti il valore $\mathbf{x_i^*}$ per cui è minima la funzione $f_i(\mathbf{x_i^*})$ può essere diverso dal valore $\mathbf{x_j^*}$ per cui è minima la funzione $f_j(\mathbf{x_j^*})$. Quindi il concetto "scalare" di

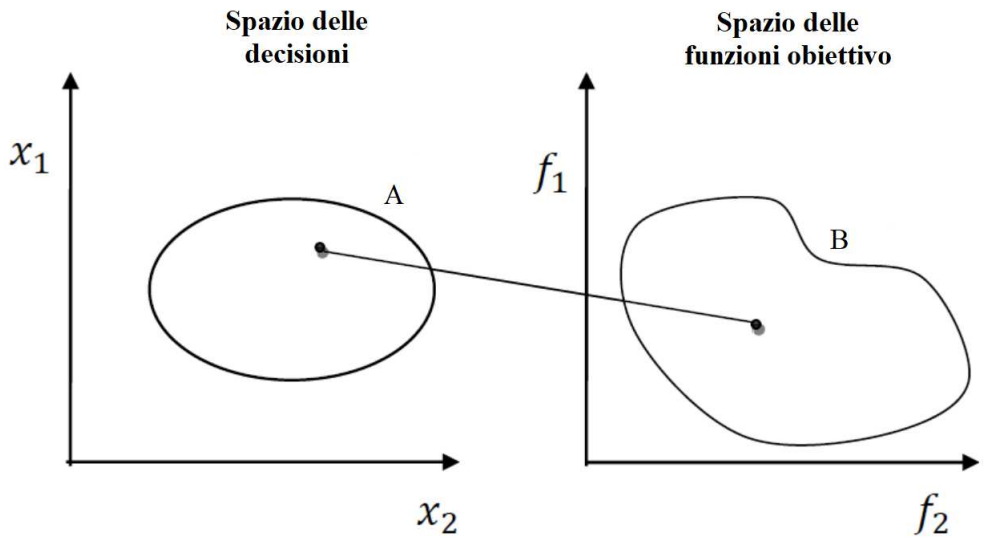


Figura 6.1: Relazione tra lo spazio delle decisioni (a sinistra) e lo spazio delle funzioni obiettivo (a destra) con $n = m = 2$ (entrambi spazi bidimensionali).

ottimizzazione, usato in un contesto mono-obiettivo e definito in 6.1, non si può applicare direttamente in un contesto multi-obiettivo. In tale contesto viene introdotta la nozione di ottimo di Pareto: un punto $\mathbf{x}^* \in A$ si definisce ottimo di Pareto se tutti gli altri punti $\mathbf{x} \in A$ hanno un valore più elevato per almeno una funzione obiettivo. L'insieme dei punti ottimi secondo il concetto di Pareto viene definito come insieme efficiente e la sua immagine viene definita fronte di Pareto. Questo concetto è visibile in figura 6.2 con i due insiemi rappresentati in rosso nei rispettivi spazi. Indicando con E l'insieme efficiente, l'affermazione di cui sopra, tradotta in linguaggio matematico diventa:

$$\mathbf{x}^* \in E \iff \nexists \ \mathbf{x} \in A : f_i(\mathbf{x}) < f_i(\mathbf{x}^*) \text{ con } i \in \{1, 2, ..., m\}$$

Il fronte di Pareto rappresenta quell'insieme di soluzioni per cui non si può migliorare una funzione obiettivo senza peggiorarne un'altra; in tale contesto non esiste un valore di $\mathbf{x}$ migliore di un altro: tutti i punti appartenenti all'insieme efficiente sono ugualmente ottimi. Una volta individuato il fronte di Pareto e l'insieme efficiente sta al "decision maker" scegliere il valore di $\mathbf{x}$ più appropriato al suo scopo, trovando il giusto compromesso tra le varie funzioni obiettivo.

Si può introdurre ora il concetto di punti dominanti e dominati, che tornerà utile nella spiegazione degli algoritmi multi-obiettivo: dati due punti $\mathbf{x_1}$ e $\mathbf{x}_2 \in A \subset \mathbb{R}^n$ e m funzioni obiettivo, $\mathbf{x_1}$ si dice dominante rispetto a $\mathbf{x_2}$ se e solo se

$$f_i(\mathbf{x_1}) < f_i(\mathbf{x_2}) \quad \forall \ i \in \{1, 2, ...m\}$$

In tal caso $\mathbf{x_2}$ risulta dominato da $\mathbf{x_1}$. Esistono coppie di punti $(\mathbf{x_1}, \mathbf{x_2})$ in cui nessuno dei due punti domina l'altro. Ciò accade quando $f_i(\mathbf{x_1}) < f_i(\mathbf{x_2})$ e $f_j(\mathbf{x_1}) > f_j(\mathbf{x_2})$. Se il punto $\mathbf{x_1}$ domina il punto $\mathbf{x_2}$ significa che $\mathbf{x_1}$ presenta prestazioni migliori rispetto a $\mathbf{x_2}$ per ogni funzione obiettivo. Dovendo scegliere tra i due punti ci si orienterà quindi verso il primo. L'insieme efficiente è l'insieme dei punti che non sono dominati da nessun altro punto $\mathbf{x} \in A \subset \mathbb{R}^n$.

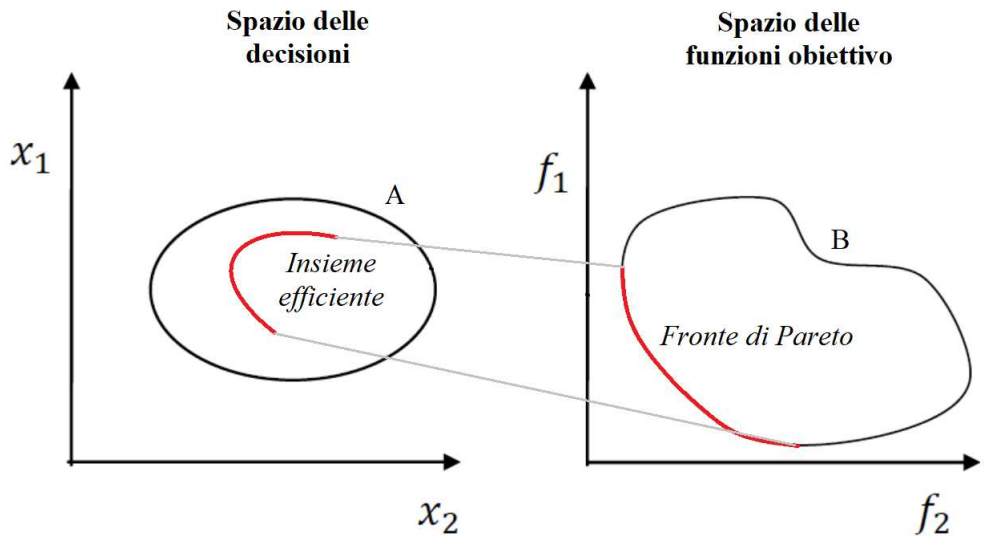


Figura 6.2: Rappresentazione dell'insieme efficiente (a sinistra) e del fronte di Pareto (a destra) nei rispettivi spazi con $n = m = 2$ (entrambi spazi bidimensionali).

Traslando i concetti fin qui esposti all'attuatore lineare in esame, si fanno corrispondere le variabili di progetto alle coordinate $\mathbf{x}$ e gli indici di prestazione con le funzioni obiettivo $f(\mathbf{x})$; così facendo lo spazio delle decisioni A ha 11 dimensioni ed è limitato dai range di variazione di ogni variabile individuati nei capitoli precedenti, mentre lo spazio degli obiettivi ha 6 dimensioni. Per l'individuazione del fronte di Pareto è necessario esplorare questo spazio 11-dimensionale grazie a degli algoritmi di ottimizzazione multi-obiettivo. Per il lavoro qui presentato viene utilizzato un algoritmo ibrido implementato nel programma di calcolo MatLab.

6.2 Descrizione dell'algoritmo multi-obiettivo

Esiste una grandissima varietà di algoritmi di ottimizzazione multi-obiettivo per l'individuazione del fronte di Pareto. Ogni algoritmo si basa sull'esplorazione dello spazio delle decisioni e sulla valutazione delle funzioni obiettivo in tale spazio. Questi algoritmi procedono in maniera iterativa, individuando ad ogni passo un'insieme di punti appartenenti allo spazio delle decisioni e valutando le funzioni obiettivo in tali punti. Una volta fatto ciò, si determinano i punti non dominati e si crea un altro insieme di punti a partire dai migliori trovati in precedenza. Si procede così fino ad un criterio di stop. In questo modo i punti non dominati trovati dall'algoritmo vengono definiti come l'insieme efficiente e viene individuata un'approssimazione del fronte di Pareto. Si capisce subito infatti che viene trovato solo un sottoinsieme finito del fronte e quindi dell'insieme efficiente. Se però i punti trovati nello spazio delle funzioni obiettivo sono abbastanza vicini, ci si può ritenere soddisfatti del fronte così individuato. Gli algoritmi di ottimizzazione multi-obiettivo sono progettati in modo da trovare in fretta le zone più vicine al fronte di Pareto e concentrare le ricerche attorno a quella zona.

Ciò che differenzia i vari algoritmi è il modo in cui esplorano lo spazio delle decisioni. Possono essere individuate due grandi categorie: gli algoritmi stocastici e quelli deterministici. I primi seguono un processo stocastico non determinabile a priori, sono più lenti, anche se è possibile velocizzarli utilizzando il calcolo parallelo, garantiscono una migliore esplorazione di tutto lo spazio delle decisioni $A \subset \mathbb{R}^n$ ed è meno probabile che si fermino su un minimo locale; sono però meno precisi nell'identificare il fronte di Pareto. I secondi sono più rapidi, seguono un processo sequenziale, deterministico, determinabile a priori e sono più precisi nell'identificazione del fronte, ma rischiano di bloccarsi su un minimo locale delle funzioni obiettivo.

Per il problema analizzato viene implementato un algoritmo ibrido, in modo da unire i vantaggi dei due algoritmi: come primo step si utilizza un algoritmo stocastico, dai cui punti trovati è inizializzato un algoritmo di ricerca diretta. In questo modo il primo step assicura l'esplorazione di tutto lo spazio delle decisioni, senza la possibile convergenza ad un minimo locale e il secondo step aumenta la precisione del fronte trovato. Lo spazio delle decisioni è composto da 11 dimensioni, ognuna delle quali corrispondente ad una variabile di progetto, mentre le funzioni obiettivo selezionate, ricordando che una forza positiva è intesa come una forza concorde con il moto della corsa, sono:

- le due forze medie d'andata cambiate di segno $-F^{for}_{m,LV}, -F^{for}_{m,RV}$
- le due forze medie di ritorno cambiate di segno $-F^{bck}_{m,LV}, -F^{bck}_{m,RV}$
- i due ripple in p.u. relativi alle forze medie d'andata $r^{for}_{F,LV}, r^{for}_{F,RV}$

Le prime 4 funzioni obiettivo sono le forze medie cambiate di segno, perché l'algoritmo utilizzato tende alla minimizzazione delle funzioni obiettivo, mentre ciò che si vuole perseguire è una massimizzazione di tali forze. Per quanto rigurda i ripple invece si vuole minimizzarli, in modo da rendere le forze il più piatte possibili, evitando $\frac{dF}{dx}$ troppo elevati che potrebbero causare problemi di natura meccanica sulle camere ventricolari.

6.2.1 Algoritmo stocastico

Come algoritmo stocastico di partenza viene utilizzato l'algoritmo "gamultiobj" presente in MatLab. Questo è un algoritmo genetico elitario (una variante dell'algoritmo NSGA-II[46]) e prima di spiegare il processo iterativo che segue è necessario introdurre dei concetti:

Rango: dato un set di punti $\{\mathbf{x_1}, \mathbf{x_2}, ...\mathbf{x_k}\} \in A$, si definiscono punti di rango 1 quei punti che non sono dominati da altri punti; i punti di rango 2 sono punti dominati solamente da punti di rango 1 e così via. In generale punti di rango i-esimo sono dominati da punti di rango i-1 o minori. Punti con un rango più basso sono punti più vicini al fronte di Pareto.

Distanza di crowding: questo è un parametro utilizzato per misurare la vicinanza di un punto rispetto ai suoi vicini. Si vuole trovare un fronte che

sia il più simile possibile al fronte reale, quindi, a parità di rango e punti è meglio avere punti distribuiti meglio lungo tutto il fronte piuttosto che tutti vicini nella stessa zona, quindi si preferiranno punti con distanza di crowding più bassa.

Spread: è una misura del movimento del fronte di Pareto e viene utilizzata come criterio di arresto dell'algoritmo. Se infatti il fronte non viene modificato di molto tra un'iterazione e l'altra, si può dedurre che sia stato trovato il fronte di Pareto.

Il processo iterativo dell'algoritmo segue i seguenti passaggi:

1. L'algoritmo genera casualmente una popolazione iniziale, formata da k individui $\{\mathbf{x_1}, \mathbf{x_2}, ...\mathbf{x_k}\} \in A$.

2. Vengono valutate le funzioni obiettivo per ogni individuo assegnando loro un punteggio in base alle loro prestazioni. Un punteggio migliore viene assegnato agli individui di rango 1, e via a scalare aumentando il rango. Tra individui dello stesso rango si opera poi un'altra classificazione basata sulla distanza di crowding: individui con tale valore più basso riceveranno un punteggio peggiore rispetto a quelli con un punteggio più elevato.

3. Gli individui con punteggi più elevati vengono selezionati come genitori e vengono generati i figli attraverso operatori genetici come crossover e mutazione. Dati due individui genitori, l'operatore crossover genera un figlio combinando casualmente le coordinate nello spazio delle decisioni dei due genitori, mentre l'operatore mutazione genera un figlio a partire da un genitore introducendo delle piccole variazioni randomiche nelle coordinate di quest'ultimo. I figli devono appartenere all'insieme A.

4. Vengono valutate le funzioni obiettivo dei figli e vengono uniti con i genitori in un'unica grande popolazione.

5. Viene assegnato un punteggio ad ogni individuo della nuova popolazione e la si riduce a k individui, in base al punteggio degli individui, selezionando i migliori.

6. L'algoritmo continua a iterare tra i punti 2 e 5 fino a che non viene raggiunto un criterio d'arresto, che nel caso in esame è lo spread del fronte inferiore ad una soglia. Dalla popolazione finale viene estratta una frazione in base al rango e alla distanza di crowding che rappresenta l'insieme efficiente, la cui immagine è il fronte di Pareto.

Gli algoritmi stocastici di questo tipo non corrono il rischio di arrestarsi su un minimo locale in quanto ad ogni iterazione si assicura l'introduzione di individui generati in maniera randomica che possono andare ad esplorare zone dello spazio A non ancora esplorate.

In tale algoritmo, quando vengono valutate le prestazioni in termini di funzioni obiettivo di un individuo si utilizzano i modelli di regressione costruiti nel capitolo 5. Usare l'analisi FEM infatti richiederebbe un'onere computazionale troppo

elevato, dato che è richiesta l'esplorazione di tanti punti per ottenere il fronte. Nella valutazione delle funzioni obiettivo, ad un individuo che non soddisfa i vincoli imposti viene assegnata una penalità, impostando le forze medie a 0 e i ripple a 10, in modo tale da non considerare quella configurazione nell'ottimizzazione. I vincoli imposti sono:

- forze di avviamento superiori a 40 N per il mover di sinistra e 20 N per quello di destra nella direzione del moto;
- configurazione classificata come positiva per quel che riguarda il vincolo sull'energia minima cumulativa;
- forze medie concordi con la direzione del moto;
- forze di avviamento per la corsa di ritorno dirette verso l'interno per entrambi i mover;
- forze di avviamento nella corsa di andata con polarità delle correnti invertita dirette verso l'interno per entrambi i mover;
- $h_{c12}, h_{c21}, h_{c22} > 0$, il che presuppone una geometria fisicamente realizzabile[1].

6.2.2 Algoritmo deterministico

L'insieme efficiente individuato con l'algoritmo stocastico viene utilizzato, insieme ad altri punti generati casualmente all'interno di A come punti iniziali per l'algoritmo deterministico "paretosearch" presente nella libreria Matlab, in modo tale da avere una valutazione più accurata del fronte. Tale algoritmo si basa anch'esso sulla valutazione delle funzioni obiettivo di una popolazione di individui $\mathbf{x} \in A$, ma questi individui vengono selezionati in maniera diversa rispetto all'algoritmo precedente. Questo algoritmo cerca i punti non dominati a partire dalla popolazione iniziale muovendosi secondo delle direzioni precise e determinate a priori. Se vengono trovati dei punti che dominano i punti di partenza, si prediligerà la ricerca di punti non dominati verso quella direzione. La gestione dei vincoli è identica a quanto descritto per l'algoritmo stocastico: si assegna una penalità se questi non sono rispettati. Le funzioni obiettivo valutate sono le stesse introdotte in 6.2.1. Alla fine del processo viene identificato l'insieme efficiente e il relativo fronte in modo più preciso rispetto ad un algoritmo stocastico. Per approfondimenti riguardo il funzionamento dell'algoritmo, si rimanda alla documentazione MatLab [47].

[1] In realtà sarebbe meglio che le altezze di fondo cava siano di almeno 3 mm, per poter sfruttare al meglio tale cava, ma nell'eventualità in cui si dovesse selezionare una configurazione con altezze di fondo cava inferiori a 3 mm, verranno presi degli accorgimenti in seguito.

6.3 Validazione dei risultati

Inizialmente è stato deciso di impostare una popolazione di output di 50 individui in modo da avere una buona rappresentatività del fronte. Per l'algoritmo genetico viene impostata una popolazione iniziale di 100 individui, il 35% dei quali viene utilizzato dall'algoritmo come output per l'individuazione del primo fronte, quindi 35 individui. Questi individui vengono poi utilizzati come punti di partenza per l'algoritmo di ricerca diretta insieme ad altri 15 individui appartenenti all'insieme A selezionati in maniera randomica, in modo che non siano troppo vicini tra di loro. L'algoritmo "paretosearch" fornisce come output 50 individui rappresentativi dell'insieme efficiente e del fronte di Pareto. Una volta trovati i punti di output è necessario validare i risultati con l'analisi FEM per capire se i modelli costruiti riescono a lavorare bene nella zona più di interesse, ovvero quella del fronte di Pareto. Questo processo è stato fatto sia per lo scenario a 5W che per quello a 10W.

6.3.1 Scenario 5 W

La prima analisi che viene condotta è quella sullo scenario $P_{Loss} = 5$ W. Viene individuato il fronte di Pareto composto da 50 punti e si verificano questi punti con l'analisi FEM. Mentre i modelli di regressione forniscono buone prestazioni, il classificatore presenta delle criticità. Come visibile dalla matrice di confusione sul fronte di Pareto riportata in figura 6.3, il classificatore ha una percentuale di errore del 50 %, assolutamente non accettabile in quanto 25 individui costituiscono una popolazione che difficilmente può rappresentare bene il fronte; si rende necessario migliorare il classificatore. Uno dei motivi per cui il classificatore ha

		Fronte di Pareto scenario 5 W	
Classe reale	positiva	25	0
	negativa	25	0
		positiva	negativa
		Classe predetta	

Figura 6.3: Matrice di confusione del primo classificatore utilizzato sul fronte di Pareto.

basse performance sul fronte di Pareto è che per aumentare le forze medie nelle corse di andata è necessario adottare una geometria che aumenti la forza di cogging, forza che si oppone al moto durante la corsa di ritorno. L'algoritmo di ottimizzazione multi-obiettivo utilizzato cerca di massimizzare le forze di andata, quindi va a trovare quelle configurazioni al limite tra le due classi. È probabile

inoltre che alcune configurazioni che soddisfano il vincolo sull'energia minima, con forze medie elevate, quindi possibili candidate ad essere scelte come configurazioni finali, vengano classificate come negative, quindi scartate dall'algoritmo. Per tale motivo si decide di aggiungere un'ulteriore classe, la classe borderline. Per l'addestramento di tale modello si utilizzano il training set e il test set già presentati nel capitolo 5, riclassificando però le configurazioni secondo il processo seguente:

1. Per ogni configurazione sono calcolati i due profili dell'energia cumulativa al variare dello spazio percorso dai mover, $e_{cum,k}(x)$, dove k è il mover considerato (destro o sinistro) e x è lo spazio percorso dal mover durante la corsa di ritorno;

2. la funzione $e_{cum,k}(x)$ viene divisa per lo spazio percorso dal mover fino al punto x, al fine di calcolare la forza media agente sul mover durante il ritorno fino al punto x.

$$F_{m,k}(x) = \frac{e_{cum,k}(x)}{x};$$

3. La configurazione viene assegnata ad una classe, secondo il seguente criterio:

$$\text{Classe} = \begin{cases} \text{Negativa} & \text{se } \exists\, x, k \quad \text{t.c. } F_{m,k}(x) \leq -2.5 \text{ N} \\ \text{Borderline} & \text{se } \forall\, x, k \quad \text{t.c. } F_{m,k}(x) > -2.5 \text{ N et } \exists\, x, k \text{ t.c } F_{m,k}(x) \leq 2.5 \text{ N} \\ \text{Positiva} & \text{se } \forall\, x, k \quad \text{t.c. } F_{m,k}(x) > 2.5 \text{ N} \end{cases}$$

Dopo aver testato vari modelli tramite il "Classification Learner App" fornito da MatLab, e la creazione di reti neurali da codice, ci si è orientati verso quest'ultima scelta. Il modello che ha presentato le prestazioni migliori è stata una rete neurale con uno strato nascosto. Tale rete è composta da 11 neuroni nel livello di input (le 11 variabili), 88 neuroni nel livello nascosto, con una sigmoide come funzione di attivazione e 3 neuroni nel livello di output corrispondenti alle 3 classi, con funzione di attivazione softmax. L'algoritmo di addestramento utilizzato è lo "Scaled conjugate gradient backpropagation" e sono necessarie 20 epoche di addestramento. Le prestazioni di tale rete sul set di training e di test sono riportate nelle matrici di confusione in figura 6.4. La precisione della rete è dell'86% sul set di training e del 71% su quello di test. Tali prestazioni sono state ritenute accettabili anche alla luce delle poche configurazioni che soddisfano il vincolo sull'energia ma che vengono classificate come negative. Nel fronte di Pareto che si andrà a costruire rientreranno inevitabilmente delle configurazioni che saranno da scartare, in quanto non è possibile creare un classificatore perfetto. Per questo motivo si decide di aumentare la popolazione dell'algoritmo genetico a 200 individui. Di questi 200 punti, 80 saranno forniti come output da tale algoritmo e, insieme ad altri 40 punti selezionati in modalità random all'interno di A, saranno utilizzati come punti di partenza per l'algoritmo di ricerca diretta del fronte di Pareto. Il fronte finale sarà costituito da 120 punti. L'algoritmo viene impostato per accettare configurazioni classificate come positive o borderline e scartare quelle classificate come negative.

Training set

Classe reale	positiva	borderline	negativa
positiva	83	2	1
borderline	6	50	13
negativa	2	7	63

Classe predetta

Test set

Classe reale	positiva	borderline	negativa
positiva	19	6	1
borderline	4	14	5
negativa	4	3	22

Classe predetta

Figura 6.4: Matrice di confusione sul set di training e di test della seconda rete neurale utilizzata per la predizione della classe di energia delle configurazioni.

Una volta trovato il fronte e validati i risultati con Ansys, vengono calcolate le prestazioni dei modelli predittivi sul fronte di Pareto. Le prestazioni dei modelli di regressione sono riportate in tabella 6.1, in cui vengono riportati $\mathrm{RMSE}_{\mathrm{std}}$ e RMSE. Come già visto, si utilizzano entrambe le metriche metrica perché, per indici quali $F_{start,RV}$, $F^{bck}_{start,k}$, il valore di $\mathrm{RMSE}_{\mathrm{std}}$ viene corrotto da punti il cui valore reale è molto prossimo allo zero. Il classificatore ora ha prestazioni migliori, riportate nella matrice di confusione visibile in figura 6.5: 26 configurazioni su 120 (il 21.7 %) non dovrebbero appartenere al fronte in quanto dovrebbero essere classificate come negative. Non preoccupa più di tanto l'errore di classificazione tra le configurazioni positive e borderline in quanto l'appartenenza all'una o all'altra classe non pregiudica la possibilità di entrare nel fronte e molte delle configurazioni borderline hanno energia minima cumulativa pari a 0, soddisfacendo il vincolo sull'energia. Le prestazioni dei modelli sono ritenute adeguate.

Matrice di confusione sul fronte di Pareto

Classe reale	positiva	borderline	negativa
positiva	47	6	0
borderline	24	17	0
negativa	6	20	0

Classe predetta

Figura 6.5: Matrice di confusione del secondo classificatore sul fronte di Pareto trovato con esso.

Tabella 6.1: Prestazioni dei modelli predittivi sul fronte di Pareto per quel che riguarda lo scenario a 5 W.

INDICE	$\mathbf{RMSE_{std}}$	RMSE
$F^{for}_{m,LV}$	0.84 %	0.48 N
$F^{for}_{m,RV}$	2.96 %	0.57 N
$F^{bck}_{m,LV}$	3.81 %	1.51 N
$F^{bck}_{m,RV}$	10.2 %	0.49 N
r^{for}_{LV}	2.35 %	0.0108 p.u.
r^{for}_{RV}	8.83%	0.090 p.u.
$F_{start,LV}$	3.55%	2.25 N
$F_{start,RV}$	7.87%	3.19 N
$F^{++}_{LV}(0)$	3.61%	2.11 N
$F^{++}_{RV}(0)$	7.85%	2.29 N
$F^{bck}_{start,LV}$	18.2%	5.7 N
$F^{bck}_{start,RV}$	15.4%	4.4 N

6.3.2 Scenario 10 W

Per questo scenario si decide di calcolare un fronte di Pareto con 120 individui in maniera del tutto analoga a quanto fatto prima, con gli stessi algoritmi, vincoli e funzioni obiettivo, utilizzando i modelli predittivi corrispondenti allo scenario $P_{Loss} = 10$ W e costruiti nella sezione 5.5. Il fronte così individuato viene poi validato con l'analisi agli elementi finiti. Le prestazioni dei modelli su tale fronte sono riportate in termini di $\mathrm{RMSE_{std}}$ e RMSE in tabella 6.2, mentre le prestazioni del classificatore sono riportate in figura 6.6. Tali prestazioni sono ritenute buone, anche per quel che riguarda il classificatore, quindi si accetta il fronte di Pareto ottenuto.

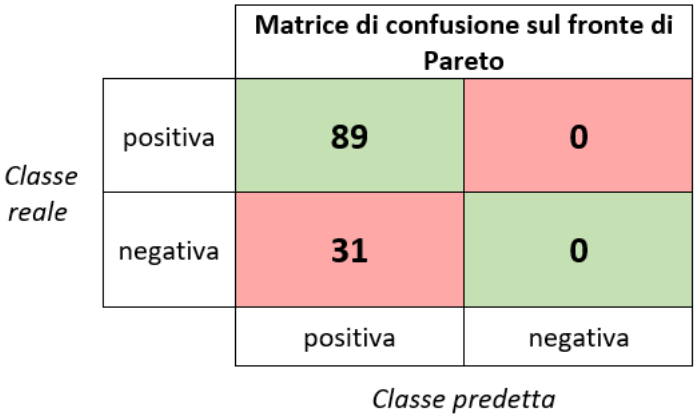


Figura 6.6: Matrice di confusione del classificatore sul fronte di Pareto trovato nello scenario $P_{Loss} = 10$ W.

Tabella 6.2: Prestazioni dei modelli predittivi sul fronte di Pareto per lo scenario 10 W.

INDICE	$\mathbf{RMSE_{std}}$	RMSE
$F^{for}_{m,LV}$	1.9%	0.76 N
$F^{for}_{m,RV}$	2.3 %	0.63 N
$F^{bck}_{m,LV}$	9.4 %	2.3 N
$F^{bck}_{m,RV}$	8.1 %	1.3 N
r^{for}_{LV}	2.7 %	0.014 p.u.
r^{for}_{RV}	3.7 %	0.033 p.u.
$F_{start,LV}$	8.2 %	4.3 N
$F_{start,RV}$	7.3 %	3.9 N
$F^{++}_{LV}(0)$	13.1 %	3.2 N
$F^{++}_{RV}(0)$	10.2 %	2.5 N
$F^{bck}_{start,LV}$	4.7 %	2.6 N
$F^{bck}_{start,RV}$	3.15 %	1.3 N

6.4 Analisi del Fronte di Pareto

6.4.1 Fronte scenario 5 W

Per la scelta della configurazione da adottare viene studiato il fronte di Pareto validato con l'analisi FEM in entrambi gli scenari. Si comincia con lo scenario $P_{Loss} = 5$ W. Per il calcolo del fronte sono state utilizzate 6 funzioni obiettivo ed è quindi impossibile rappresentare queste 6 grandezze contemporaneamente in un grafico. Le funzioni obiettivo più importanti sono senza dubbio le forze medie nelle corse di andata con i relativi ripple, mentre le forze medie nelle corse di ritorno non risultano così importanti al fine delle prestazioni dell'attuatore: è sufficiente che l'attuatore sia in grado di riportare i magneti in posizione iniziale.

Tra le due forze medie, è più sfidante quella sul mover di sinistra, in quanto, come già illustrato nel capitolo 1, è richiesta una forza più elevata per svolgere il lavoro cardiaco associato al rispettivo ventricolo, mentre quella esplicata sul mover di destra soddisfa facilmente il rispettivo lavoro.

In figura 6.7(a) è rappresentato il fronte di Pareto utilizzando come assi principali le forze medie di andata sui due mover; le configurazioni che soddisfano il vincolo sull'energia minima sono colorate in verde, mentre quelle che non lo soddisfano sono in gradazioni di rosso a seconda del valore di E_{min}. Quello rappresentato non è un fronte di Pareto come ce lo si potrebbe aspettare: bisogna infatti ricordare che sono state impostate 6 funzioni obiettivo, quindi è possibile che andando a disegnare il fronte su due dimensioni si trovino delle configurazioni dominate, ma analizzando i punti sulle 6 dimensioni, si trovano solamente punti non dominati.

Le configurazioni che non rispettano i vincoli vengono scartate dal fronte.

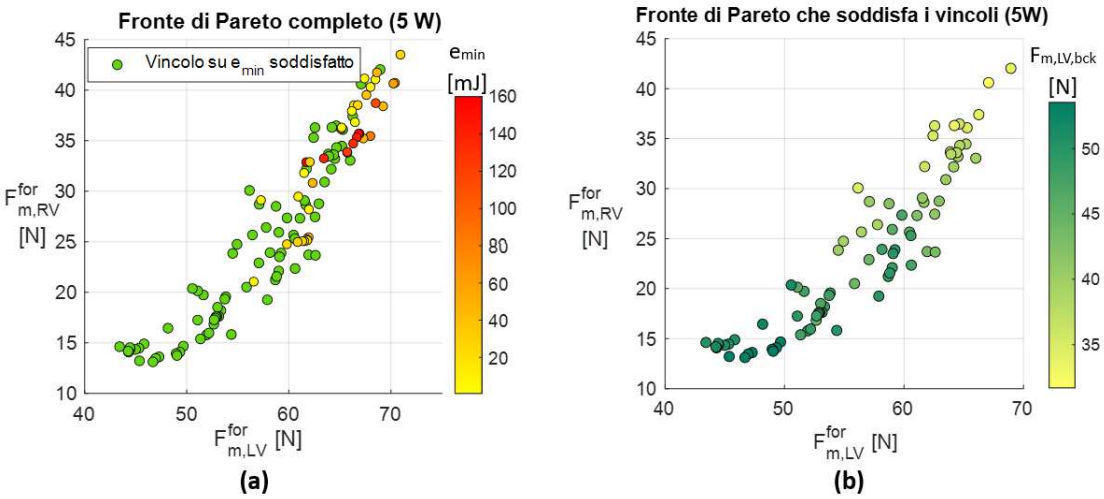


Figura 6.7: Fronte di Pareto per lo scenario $P_{Loss} = 5$ W utilizzando come assi principali le forze medie di andata dei mover. In (a) sono raffigurate anche le geometrie che non soddisfano i vincoli, mentre in (b) solo quelle che li soddisfano. In (b) si può vedere anche l'andamento di $F^{bck}_{m,LV}$

Rimangono così 85 configurazioni mostrate in 6.7(b) tra cui si sceglierà quella da adottare. Da tale immagine si evince come all'aumentare della forza media di andata sul mover di sinistra aumenti la forza media sul mover destro con una proporzionalità lineare. Si può notare inoltre come, all'aumentare delle forze medie della corsa di andata, diminuiscano le forze nella corsa di ritorno, come ci si aspettava. Come detto si cercherà una configurazione che possa massimizzare le forze medie delle corse di andata, senza impedire ai mover di ritornare alla loro posizione iniziale.

In figura 6.8 è riportato il fronte di Pareto proiettato sugli assi forza media e ripple per il mover sinistro. Viene anche riportata la rispettiva forza media di ritorno tramite il colore dei punti, in modo da dare un'idea di come questa vari in relazione alle altre grandezze. Si inizia dall'analisi di queste grandezze in quanto, come già detto, è il profilo di forza del mover sinistro quello più sfidante. In rosso vengono riportate le soluzioni non dominate, tra cui si sceglierà quella da adottare per questo scenario[2].

Analizzando i profili di forza delle configurazioni selezionate si decide di scartare le numero 1, 6 e 9. Le prime perché non presentano miglioramenti sulla forza sul mover sinistro se confrontate rispettivamente con la 2 e la 5; l'ultima viene scartata perché considerata troppo irregolare.

In figura 6.9 è visualizzato il fronte di Pareto utilizzando gli stessi indici usati in figura 6.8, ma riferiti al mover destro, mettendo in evidenza le configurazioni selezionate poc'anzi.

[2]per maggiore chiarezza vengono numerate in ordine crescente.

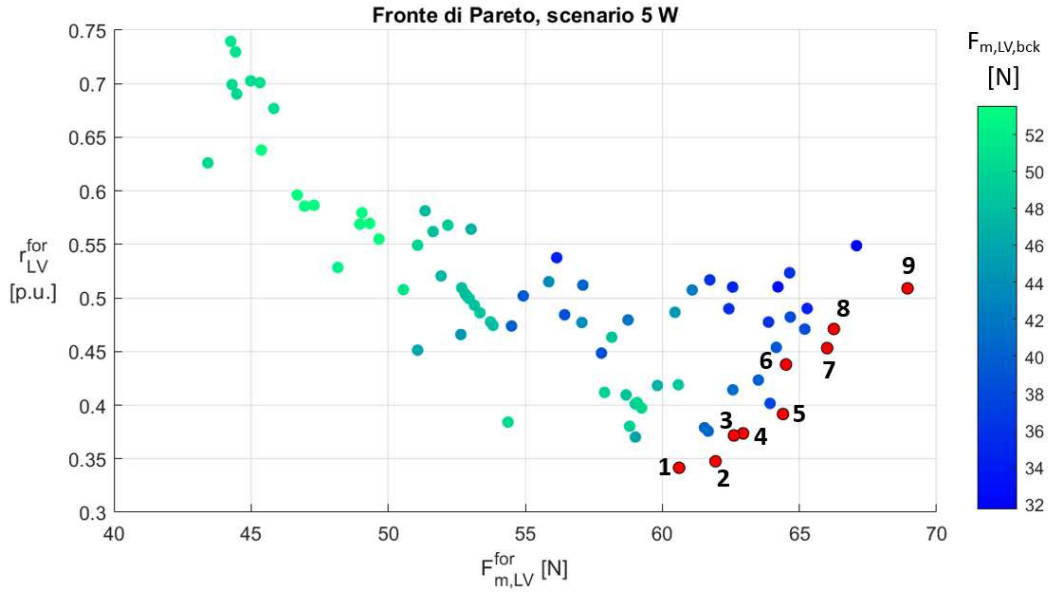


Figura 6.8: Fronte di Pareto nello scenario $P_{Loss} = 5$ W. In rosso sono riportate le soluzioni non dominate, tra cui si andrà a scegliere la configurazione da adottare.

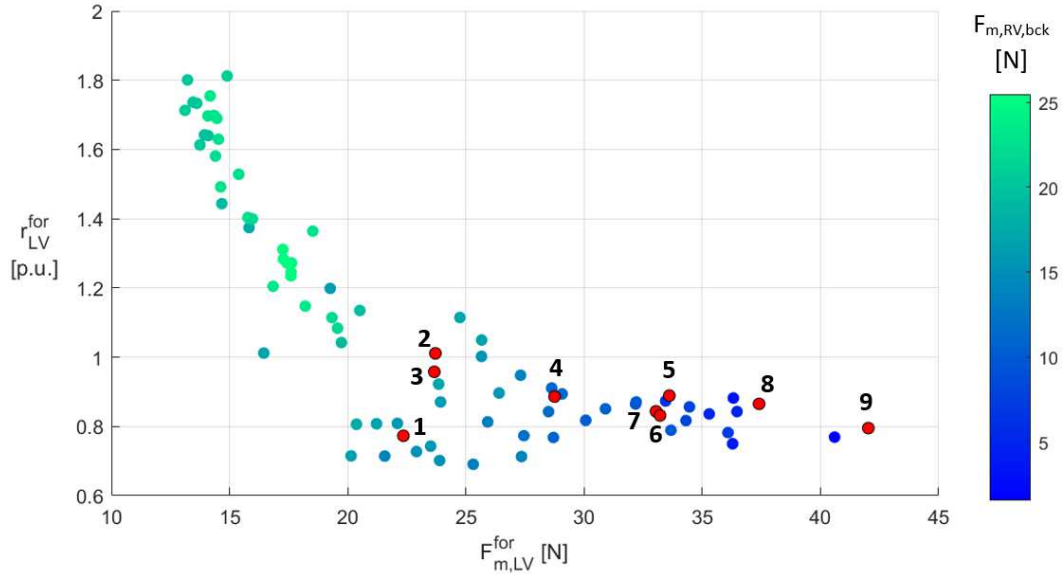


Figura 6.9: Fronte di Pareto nello scenario $P_{Loss} = 5$ W; sono riportate in rosso le configurazioni selezionate in figura 6.8.

6.4.2 Fronte scenario 10 W

Si passa ora all'analisi del fronte di Pareto nello scenario $P_{Loss} = 10$ W, rappresentato in figura 6.10 in termini di forza medie nella fase sistolica dei due mover. Le configurazioni che non rispettano i vincoli (in gradazioni calde in base al valore di e_{min}) vengono scartate. Si osservi come all'aumentare delle forze medie di andata, si trovino più configurazioni che non rispettano il vincolo sull'energia minima, analogamente allo scenario precedente; è individuabile anche una corre-

lazione tra le forze medie di andate dei due mover, simile a quella individuata nello scenario a potenza inferiore.

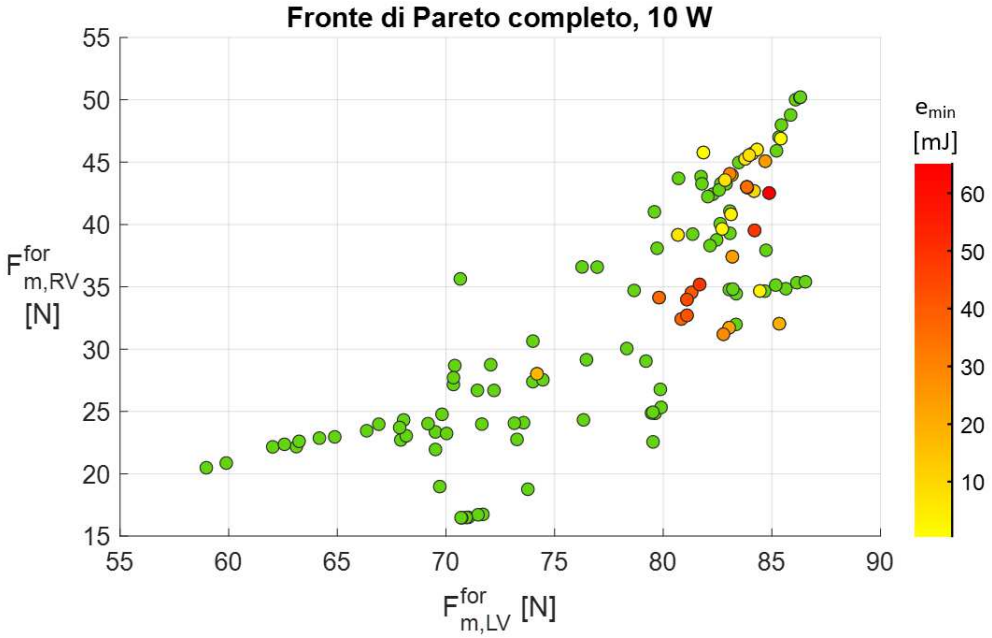


Figura 6.10: Fronte di Pareto completo dello scenario $P_{Loss} = 10$ W. Vengono messe in evidenza le configurazioni che non rispettano il vincolo sull'energia minima, che verranno poi eliminate. Queste si concentrano dove le forze medie d'andata sono più elevate.

Dopo l'eliminazione delle configurazioni che non rispettano i vincoli, il fronte si riduce da 120 a 89 punti e viene mostrato in figura 6.11 utilizzando come assi principali $F^{for}_{m,LV}$ e il relativo ripple; si riporta anche $F^{bck}_{m,LV}$ tramite il colore dei punti.

Dopo l'analisi dei profili di forza di tali configurazioni si decide di tenere le configurazioni 12, 13, 14 e 15; la configurazione 10 viene eliminata perché presenta un profilo irregolare, in particolare i profili iniziano con forze troppo basse in relazione allo scenario (47 N sul mover sinistro e 27 N sul destro) per poi attestarsi a valori più elevati; la 11 viene eliminata in quanto, è quasi uguale alla configurazione 12.

Le 4 configurazioni così selezionate vengono riportate sul fronte utilizzando come assi principali gli indici delle forze del mover di destra (figura 6.12). In tale raffigurazione si nota come si concentrino in una zona ottimale del fronte.

6.4.3 Confronto tra i due scenari

Per confrontare le configurazioni così ottenute si effettuano ulteriori analisi FEM: gli individui selezionati nel primo fronte vengono analizzati aumentando la potenza e riportati sul fronte del secondo scenario e viceversa.

Per prima cosa le configurazioni 12, 13, 14 e 15 vengono analizzate impostando $P_{Loss} = 5$ W e riportate sul rispettivo fronte come mostrato in figura 6.13. Da tale immagine sembra che queste configurazioni (in giallo) siano migliori rispetto

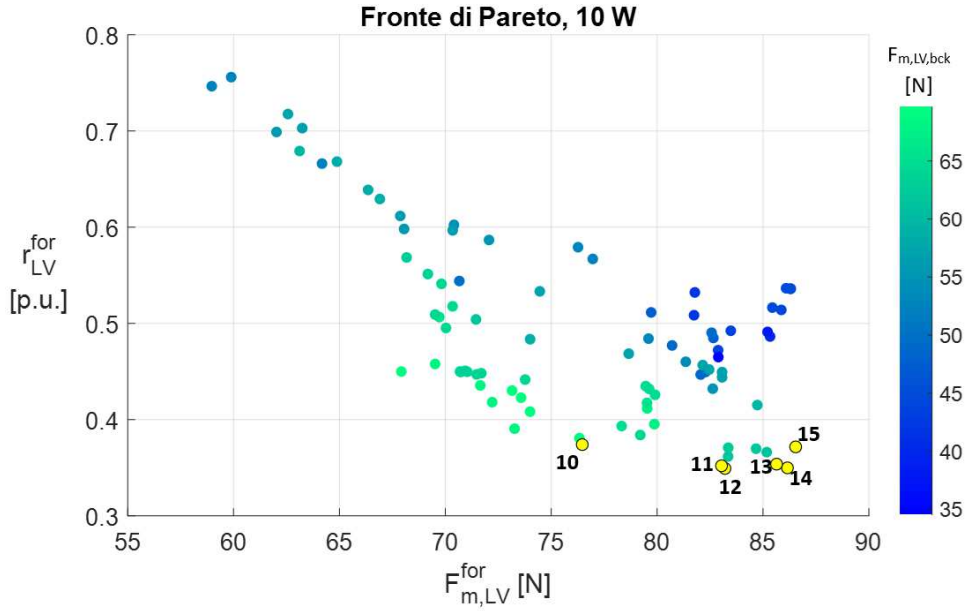


Figura 6.11: Fronte di Pareto dello scenario $P_{Loss} = 10$ W. In giallo sono raffigurate le configurazioni che si considereranno per la successiva ottimizzazione.

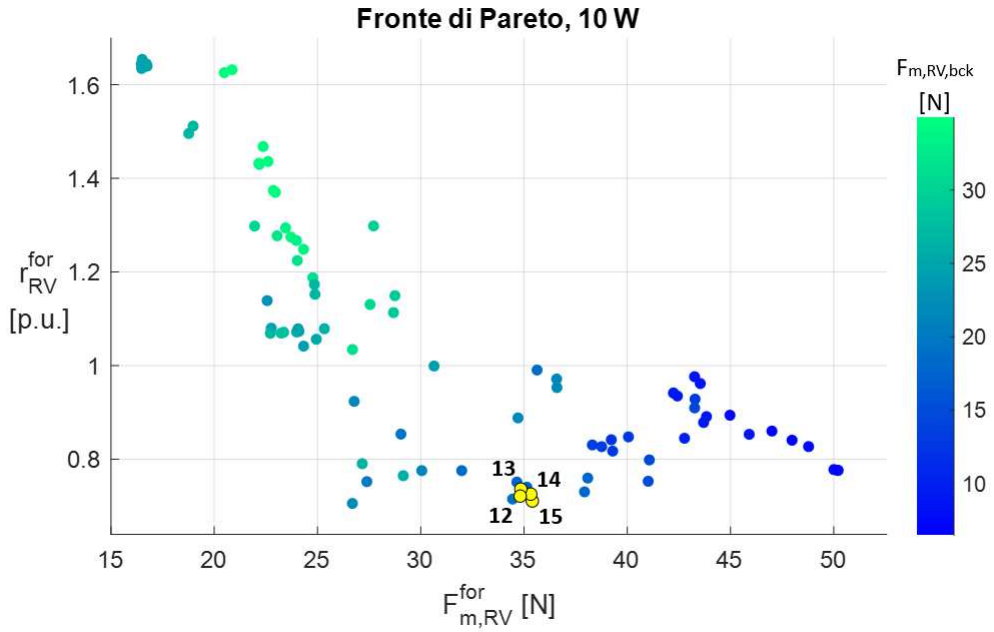


Figura 6.12: Fronte di Pareto nello scenario $P_{Loss} = 10$ W; sono riportate in giallo le configurazioni che erano state selezionate in precedenza.

a quelle selezionate in precedenza (in rosso), in quanto si trovano in posizione dominante. Dall'analisi dei profili di forza emerge però come nessuno di questi individui soddisfi il vincolo sull'energia minima cumulativa, seppur di poco.

Si procede ora al calcolo dei profili di forza delle configurazioni 2, 3, 4, 5, 7 e 8 impostando $P_{Loss} = 10$ W e riportando i risultati nel rispettivo fronte (punti in rosso in figura 6.14). Tali punti, nonostante siano dominati, si collocano in una

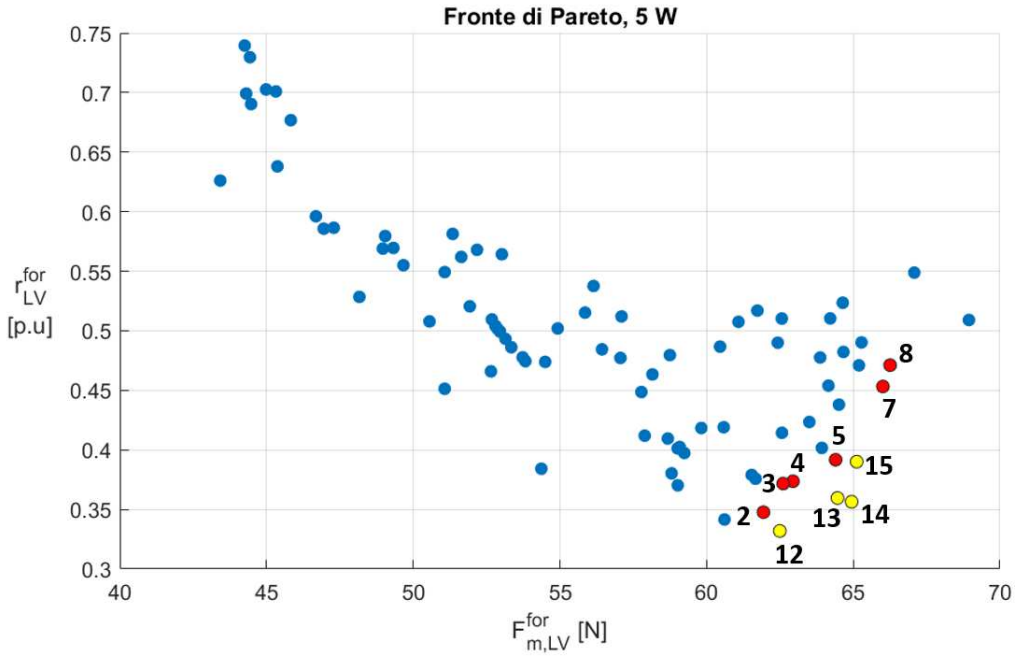


Figura 6.13: Fronte di Pareto dello scenario $P_{Loss} = 5$ W in cui vengono riportate le configurazioni selezionate con $P_{Loss} = 10$ W, riducendo le perdite (in giallo).

buona posizione all'interno del fronte.

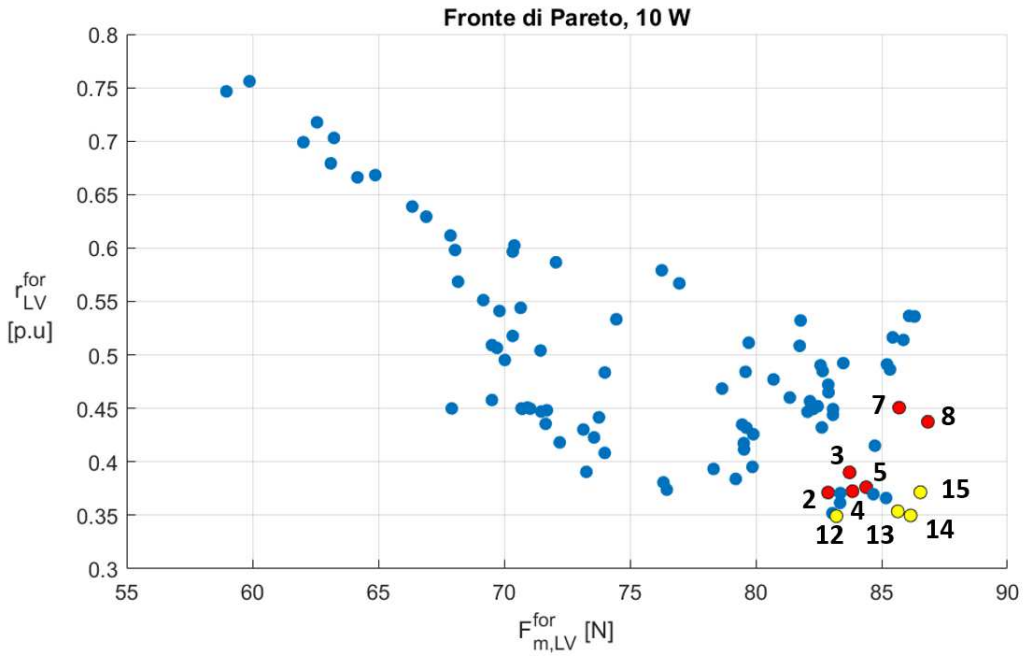


Figura 6.14: Fronte di Pareto dello scenario $P_{Loss} = 10$ W in cui vengono riportate le configurazioni selezionate con $P_{Loss} = 5$ W, aumentando le perdite (in rosso).

Dall'analisi dei profili di forza si è deciso di orientarsi sulla configurazione numero 5 in quanto presenta buone prestazioni in entrambi gli scenari e soddisfa i vincoli. Le prestazioni di tale configurazione, insieme ai valori delle 11 variabili di

progetto sono riportate in tabella 6.3. In figura 6.15 viene mostrata l'architettura dello statore ottenuta a partire da queste variabili. L'altezza di cava inferiore viene aumentata a 3 mm, altrimenti risulta difficilmente sfruttabile.

Tabella 6.3: Le 11 variabili di partenza e gli indici di prestazione della configurazione numero 5, a cui si è giunti alla fine del processo di ottimizzazione multi-obiettivo.

Variabile	Valore
w_{bk1}	12.2 mm
h_{s2}	2.8 mm
h_{sd1}	4.9 mm
h_m	14.3 mm
w_m	3.9 mm
z_{s1}	11.8 mm
z_{s2}	11.8 mm
Δr_{c1}	0.7 mm
Δr_{c2}	3.0 mm
α_{c1}	0.8
α_{c2}	0.72

Indice	5 W	10 W
$F^{for}_{m,LV}$	64.43 N	84.43 N
r^{for}_{LV}	0.391	0.376
$F^{for}_{m,RV}$	33.60 N	40.69 N
r^{for}_{RV}	0.888	0.854
$F^{bck}_{m,LV}$	37.551 N	55.85 N
$F^{bck}_{m,RV}$	5.904 N	14.71 N
$F_{start,LV}$	78.89 N	109.62 N
$F_{start,RV}$	36.13 N	48.44 N

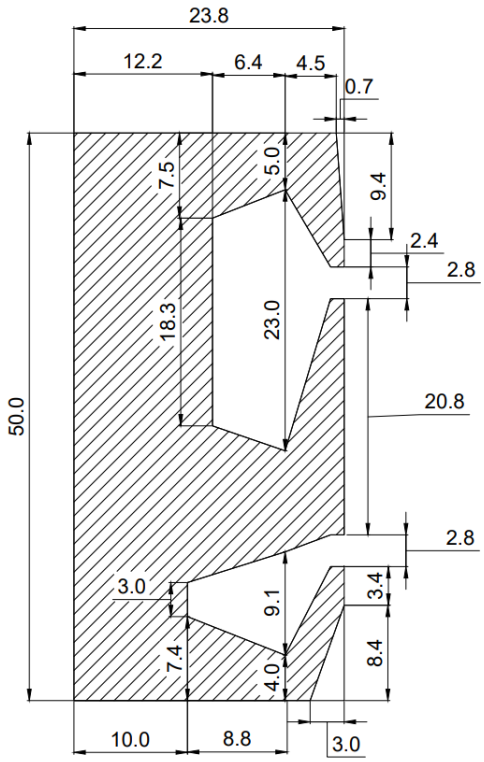


Figura 6.15: Configurazione geometrica dello statore ottenuta mediante la prima ottimizzazione.

6.5 Seconda ottimizzazione

Si vuole ora proseguire nell'ottimizzazione nell'intorno della configurazione selezionata, in modo da ottenere un ulteriore miglioramento delle prestazioni .

Per tale scopo, come prima cosa vengono ristretti i range per le variabili di progetto nell'intorno di quella selezionata. Successivamente si crea un dataset con configurazioni appartenenti a tale spazio con cui vengono addestrati nuovi modelli predittivi. A questo punto si esegue nuovamente l'algoritmo di ricerca del fronte di Pareto, validando successivamente i risultati con la FEA e selezionando la configurazione opportuna. Il tutto viene fatto per lo scenario a 5 W.

6.5.1 Restrizione dei range

Indicando con $x_{i,c}$ il generico indice di prestazione i della configurazione c, appartenente al fronte a 5 W, e con $\bar{x}_i$ l'indice i della configurazione 5, per ogni individuo c viene calcolata la distanza normalizzata d_c dalla numero 5, tenendo conto solamente delle forze medie di andata sui due mover e i rispettivi ripple, secondo la

$$d_c = \sqrt{\sum_{i=1}^{4} \left(\frac{x_{i,c} - \bar{x}_i}{\bar{x}_i} \right)^2} \tag{6.3}$$

Vengono quindi selezionati gli otto individui appartenenti al fronte a 5 W con indice d_c più basso. Così facendo si trovano le configurazioni più vicine a quella selezionata in termini delle funzioni obiettivo ritenute più importanti. L'escursione di ogni variabile di progetto tra queste otto configurazione viene utilizzata come nuovo range ristretto.

I valori delle variabili di progetto della configurazione 5 e i nuovi range di variazione sono riportati in tabella 6.4. Sono riportati anche i range utilizzati in precedenza in modo da avere un rapido confronto.

Tabella 6.4: Valori delle variabili di progetto per la configurazione numero 5, selezionata in precedenza, range ristretti con l'approccio indicato in questo paragrafo e range iniziale.

Variabile	Valore n°5	Range ristretto	Range iniziale
w_{bk1}	12.22 mm	[11.6 - 12.7] mm	[9.0 - 13.5] mm
h_{s2}	2.85 mm	[1.5 - 3.9] mm	[1.5 - 4.0] mm
h_{sd1}	4.92 mm	[4.5 - 5.2] mm	[4.0 - 6.0] mm
h_m	14.85 mm	[14.0 - 14.9] mm	[12.0 - 16.0] mm
w_m	3.92 mm	[3.5 - 4.0] mm	[2.4 - 4] mm
z_{s1}	11.85 mm	[11.5 - 14.0] mm	[9.0 - 14.0] mm
z_{s2}	11.88 mm	[11.0 - 13.0] mm	[9.0 - 13.0] mm
Δr_{c1}	0.69 mm	[0.6 - 2.1] mm	[0.5 - 2.5] mm
Δr_{c2}	3.06 mm	[2.8 - 3.2] mm	[1.2 - 3.2] mm
α_{c1}	0.80	[0.3 - 0.8]	[0.3 - 0.8]
α_{c2}	0.72	[0.5 - 0.8]	[0.3 - 0.8]

6.5.2 Costruzione dei modelli predittivi

Per addestrare i modelli predittivi è necessario prima di tutto creare il dataset di addestramento. Vengono quindi create e analizzate tramite FEA 115 configurazioni, ognuna delle quali costruita con le 11 variabili di progetto impostate ad un valore random all'interno del relativo range di variazione. Questo dataset è più ridotto rispetto a quello usato in precedenza in quanto i range sono molto ristretti.

Con il data set realizzato vengono addestrati i nuovi modelli predittivi. Sulla base dell'esperienza pregressa ci si orienta verso un modello di GPR per i modelli regressivi, mentre si opta per una rete neurale con uno strato nascosto per quanto riguarda la classificazione in base all'energia minima.

I vari modelli regressivi vengono ottimizzati con un algoritmo Bayesiano tramite MatLab analogamente a quanto già fatto in precedenza; per la protezione contro l'overfitting si utilizza la validazione crociata dividendo il data set in 5 fold e si riserva il 30% del data set come test set. Non vengono costruiti nuovi modelli predittivi di $F_{LV}^{++}(0)$, $F_{start,RV}^{bck}$ e $F_{start,LV}^{bck}$ in quanto i valori di questi indici sono abbondantemente dentro i vincoli in tutto il data set. Anche i modelli predittivi delle forze medie durante la corsa di ritorno in sono costruiti, in quanto, in questa seconda fase, non verranno impostate come funzioni obiettivo. Ora è importante solamente il vincolo sull'energia minima cumulativa durante la fase di diastole.

Le prestazioni e gli iperparametri dei nuovi modelli sono riportati in tabella 6.5

Tabella 6.5: Modelli GPR che garantiscono le prestazioni migliori per quanto riguarda la predizione sul test set per il nuovo data set in termini di RMSE . Sigle utilizzate: NM = nonisotropic matern, NI SE = non isotropic squared exponential, NI RQ= nonisotropiq rational quadrtic

INDICE	**BF**	**KERNEL**	$\boldsymbol{\sigma}$	**SD**	**RMSE** (test set)
$F_{m,LV}^{for}$	Zero	NM 5/2	$3.88{\cdot}10^{-4}$	Si	0.0889 N
$F_{m,RV}^{for}$	Costante	NM 3/2	0.0168	No	0.137 N
r_{LV}^{for}	Costante	NM 5/2	0.0262	Si	0.00403 p.u.
r_{RV}^{for}	Zero	NI RQ	0.0421	Si	0.0060 p.u.
$F_{start,LV}$	Costante	NM 5/2	2.896	No	0.287 N
$F_{start,RV}$	Lineare	NI SE	1.87	No	0.624 N
$F_{RV}^{++}(0)$	Zero	NI RQ	6.336	Si	0.675 N

La rete neurale ha lo scopo di prevedere se la configurazione esaminata presenti energia minima cumulativa pari a zero (classificata come positiva) o maggiore a zero (classificiata come negativa). Nel primo caso i mover riescono a ritornare nella loro posizione iniziale senza forze esterne, nel secondo no. Viene quindi eliminata la classe borderline al fine di una maggiore precisione. Il data set viene diviso come segue: 55% training set, 15% validation set e 30% test set. Vengono testate diverse architetture, modificando il numero di neuroni dello strato

nascosto e la relativa funzione di attivazione, fino al raggiungimento di buone prestazioni della rete. Il modello ottenuto si compone di uno strato nascosto formato da 16 neuroni con la sigmoide come funzione di attivazione. Sono inoltre presenti gli 11 neuroni di input e i due neuroni di output. Le prestazioni della rete sono mostrate nelle matrici di confusione di figura in cui la classe 1 corrisponde alla classe positiva e la 2 a quella negativa 6.16

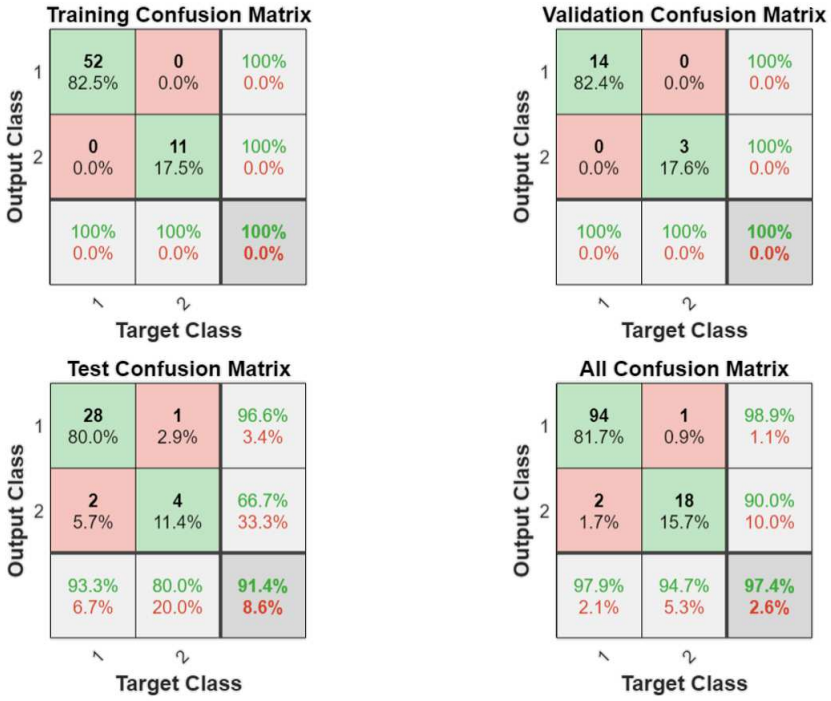


Figura 6.16: Matrici di confusione della rete per la classificazione delle configurazioni in base al vincolo sull'energia minima cumulativa. La classe 1 contiene le architetture per cui è soddisfatto tale vincolo, la classe 2 raggruppa quelle che non lo soddisfano.

6.5.3 Calcolo del fronte di Pareto

Si esegue l'algoritmo di ricerca del fronte di Pareto introdotto in precedenza utilizzando questi nuovi range e modelli predittivi. Vengono impostate ora solo quattro funzioni obiettivo: le due forze medie nella fase sistolica e i relativi ripple. Si impostano inoltre i parametri degli algoritmi in modo che il fronte sia composto da 50 individui. La modalità di gestione dei vincoli è la stessa utilizzata in precedenza.

Trovato il fronte si procede alla sua validazione tramite FEA, valutando poi le prestazioni dei modelli predittivi in termini di RMSE, riportato in tabella 6.6. Le prestazioni di tali modelli sono ottime.

La rete neurale per la classificazione basata sull'energia minima ha un'accuratezza del 25%, un valore non accettabile. Questo succede perché, come già accadeva prima, un incremento delle forze nella fase sistolica può essere ottenuto tramite l'aumento dell'effetto di cogging, che si oppone durante il ritorno degli

Tabella 6.6: Prestazioni dei modelli regressivi sul fronte di Pareto individuato

Indice predetto	RMSE
$F_{m,LV}^{for}$	0.265 N
$F_{m,RV}^{for}$	0.373 N
r_{LV}^{for}	0.0070 p.u.
r_{RV}^{for}	0.0175 p.u.
$F_{start,LV}$	0.245 N
$F_{start,RV}$	1.48 N
$F_{RV}^{++}(0)$	0.619 N

stessi nella fase diastolica. L'algoritmo di ricerca del fronte tende ad aumentare le forze medie di andata, andando a trovare le falle del classificatore.

Per superare questo problema si riaddestra il classificatore aggiungendo al dataset iniziale i 50 individui appena trovati e modificandone l'architettura. Si ricerca nuovamente il fronte di Pareto con l'algoritmo e lo si valida tramite FEA. Così facendo le prestazioni del classificatore sugli individui appartenenti al nuovo fronte migliorano, arrivando ad un'accuratezza del 54%. Tale valore non è ritenuto accettabile, quindi si esegue nuovamente questa procedura. Con questa seconda iterazione, l'accuratezza del classificatore migliora al 76%, un valore accettabile, quindi il processo viene fermato. Questa rete neurale si compone di uno strato nascosto con 75 neuroni e la sigmoide come funzione di attivazione. Ci sono poi gli undici neuroni di input e i due di output.

Questo processo iterativo è visibile in figura 6.17. È interessante notare il processo adattivo del classificatore: se il modello precedente sbagliava la predizione, il modello successivo si terrà alla larga da quella zona; viceversa se la predizione era corretta, la nuova rete selezionerà altri punti lì vicino.

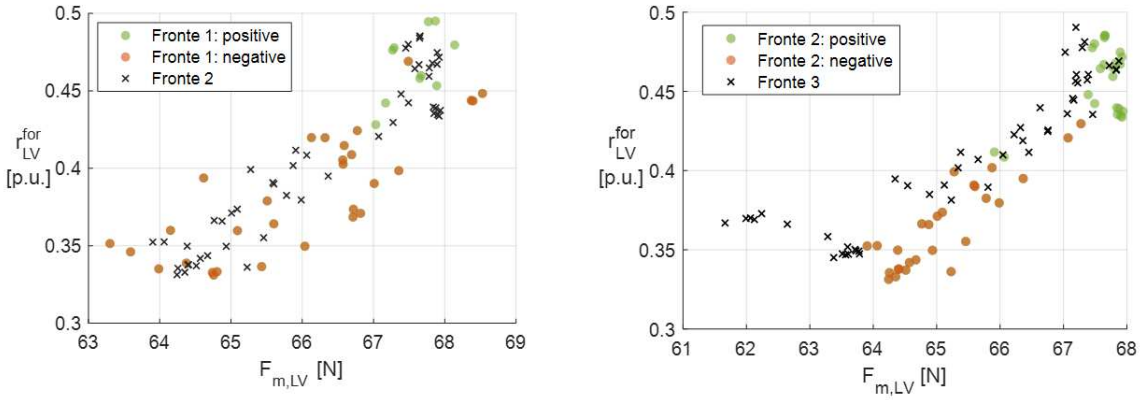


Figura 6.17: Fronte di Pareto in termini di forza media sistolica e ripple sul mover sinistro durante le varie iterazioni per la creazione del classificatore sull'energia minima. Si può notare come il nuovo si adatti in base agli errori, andando a selezionare individui lontani dai punti sbagliati dalla rete precedente.

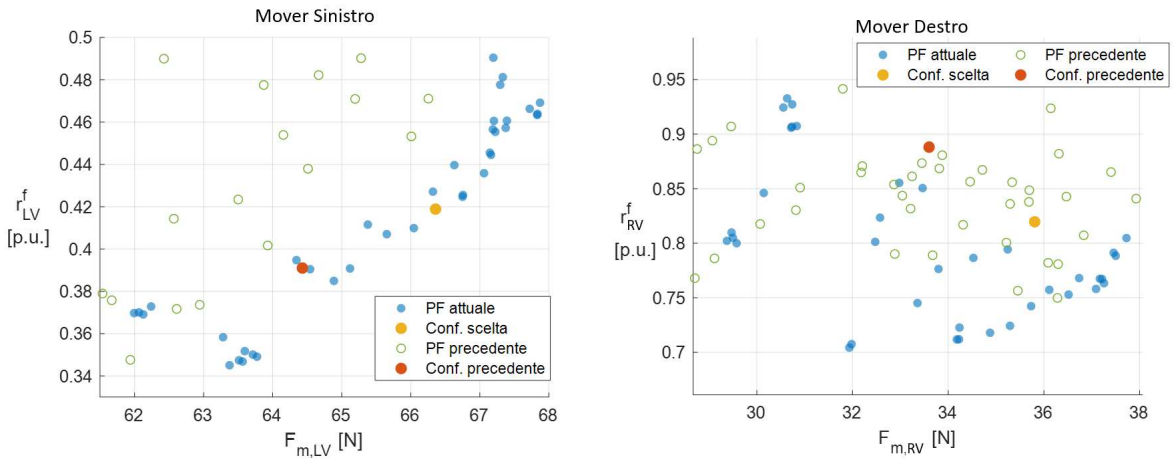


Figura 6.18: Configurazione scelta prima e dopo la rifinizione. Sono visualizzabili anche i due fronti di Pareto ottenuti nelle due fasi per fare un confronto.

Si seleziona quindi la configurazione più opportuna, sulla base della posizione sul fronte di Pareto, del rispetto dei vincoli e della facilità costruttiva. In figura 6.18 si riporta la posizione di tale architettura e di quella a cui si era giunti prima di questa rifinizione. Sono riportati inoltre i fronti di Pareto trovati prima e dopo questo passaggio. Si può osservare come si ottenga un fronte più definito e granulare rispetto a prima. Inoltre questo secondo fronte è dominante rispetto al primo.

Capitolo 7

Analisi dell'attuatore

7.1 Architettura

7.1.1 Variabili di progetto

Alla fine del processo di ottimizzazione si è giunti ad una configurazione ben precisa, le cui 11 variabili di partenza e gli indici di prestazioni sono riportati in tabella 7.2.

Tabella 7.1: Le 11 variabili di partenza e gli indici di prestazione della configurazione ottimizzata.

Variabile	Valore
w_{bk1}	12.40 mm
h_{s2}	2.75 mm
h_{sd1}	4.91 mm
h_m	14.98 mm
w_m	3.89 mm
z_{s1}	12.65 mm
z_{s2}	12.75 mm
Δr_{c1}	0.55 mm
Δr_{c2}	2.99 mm
α_{c1}	0.532
α_{c2}	0.794

Indice	5 W	10 W
$F^{for}_{m,LV}$	66.4 N	86.5 N
r^{for}_{LV}	0.418	0.393
$F^{for}_{m,RV}$	35.8 N	44.5 N
r^{for}_{RV}	0.820	0.769
$F^{bck}_{m,LV}$	38.0 N	56.6 N
$F^{bck}_{m,RV}$	5.61 N	115.7 N
$F_{start,LV}$	75.8 N	105.7 N
$F_{start,RV}$	37.6 N	51.7 N

A partire da queste undici variabili di progetto, come descritto nel paragrafo 4.2.1, è possibile costruire l'attuatore. Facendo riferimento alla figura 7.1, i valori delle varie quote, arrotondati al decimo di millimetro per vincoli costruttivi, sono riportati in tabella 7.2.

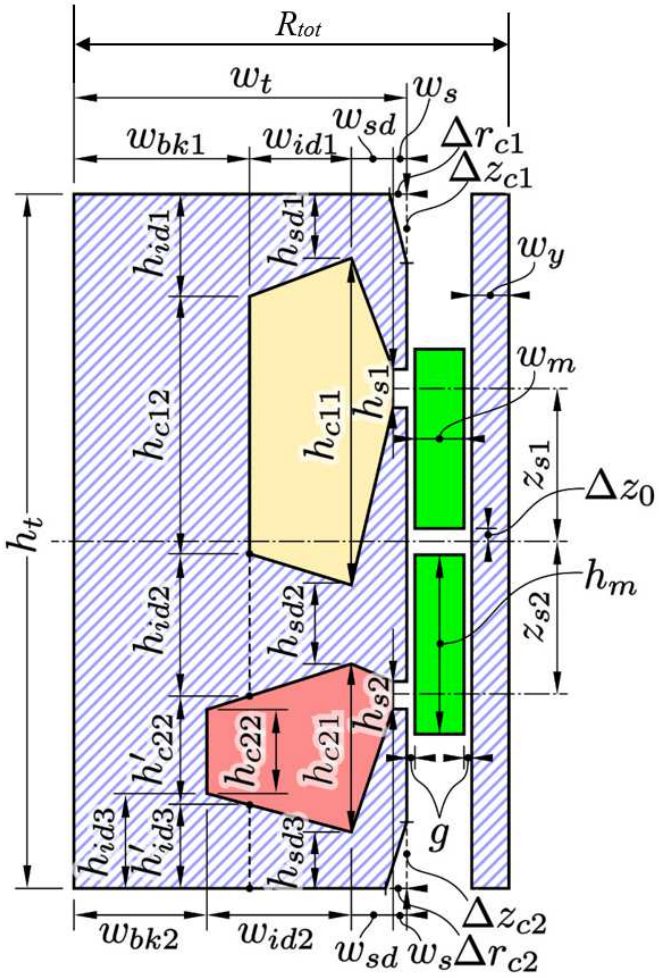


Figura 7.1: Schema dell'architettura dell'attuatore quotata

Tabella 7.2: Valori delle quote dell'attuatore ottimizzato arrotondati per vincoli costruttivi al decimo di millimetro.

Variabile	**Valore** [mm]	**Variabile**	**Valore** [mm]	**Variabile**	**Valore** [mm]
w_{bk1}	12.4	w_{id1}	6.2	w_{sd}	4.0
w_s	1.2	w_t	23.8	w_m	3.9
g	0.5	w_y	2.5	w_{bk2}	10.0
w_{id2}	8.6	R_{tot}	31.2	Δr_{c1}	0.5
Δz_{c1}	5.8	Δr_{c2}	3.0	Δz_{c2}	8.6
h_{id1}	7.4	h_{c12}	18.6	h_{id2}	12.9
h_{c22}	3.0	h_{id3}	7.4	h_{sd1}	4.9
h_{c11}	23.0	h_{sd2}	8.9	h_{c21}	9.2
h_{sd3}	4.0	h_{s1}	2.8	h_{s2}	2.8
z_{s1}	12.6	z_{s2}	12.7	h_m	15.0
h_t	50.0	Δ_{z0}	1.25		

7.1.2 Costruzione

L'attuatore risulta composto di tre parti:

Statore

Statore massiccio cilindrico in lega di ferro cobalto (Vacoflux® 48), con due cave in cui alloggiare i due avvolgimenti. La geometria quotata dello statore è visibile in figura 7.2

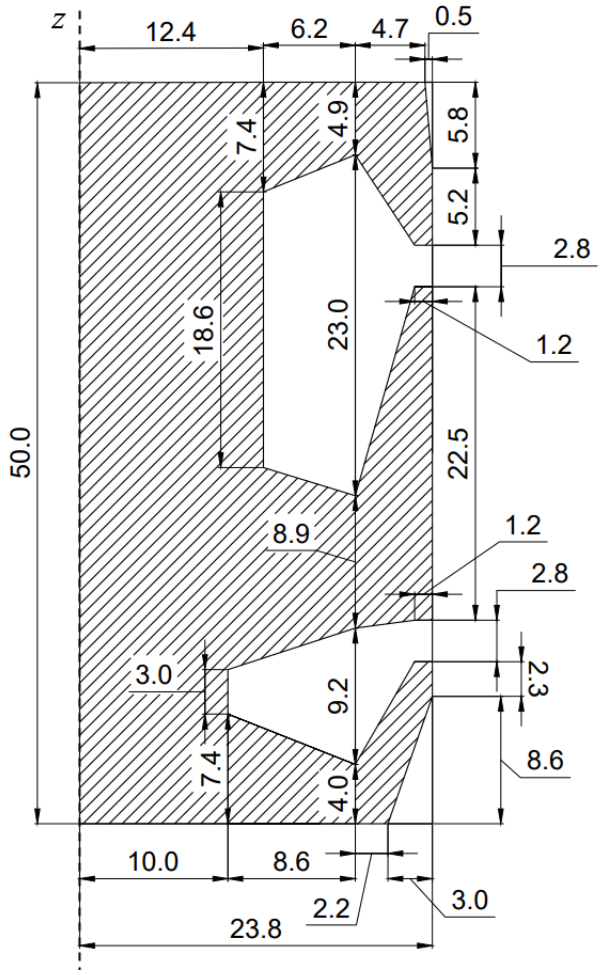


Figura 7.2: Configurazione quotata dello statore ottenuta alla fine del processo di ottimizzazione. Viene raffigurata una sezione trasversale dell'attuatore, il quale presenta una simmetria radiale lungo l'asse z. Le quote sono espresse in mm.

Magneti

Due magneti in NdFeB-48 a forma di anello con raggio interno di 24.3 mm, raggio esterno di 28.2 mm e altezza di 15.0 mm, dotati di magnetizzazione radiale, con il polo nord sul lato interno e il polo sud su quello esterno, come visibile in 7.3

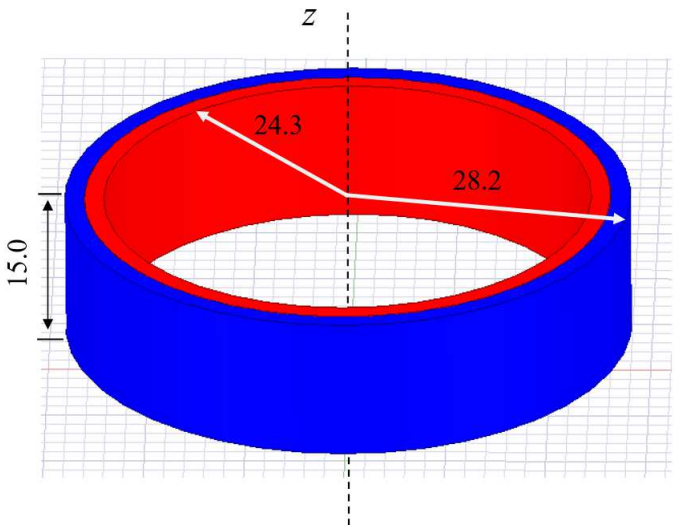


Figura 7.3: Configurazione quotata di uno dei due magneti ad anello. Il polo nord interno è colorato in rosso, mentre il polo sud esterno in blu. Le quote sono espresse in mm.

Giogo esterno

Giogo esterno a forma di anello in lega di ferro cobalto (Vacoflux® 48) con raggio interno di 28.7 mm, raggio esterno di 31.2 mm e altezza di 50 mm. Il giogo esterno è visibile in 7.4

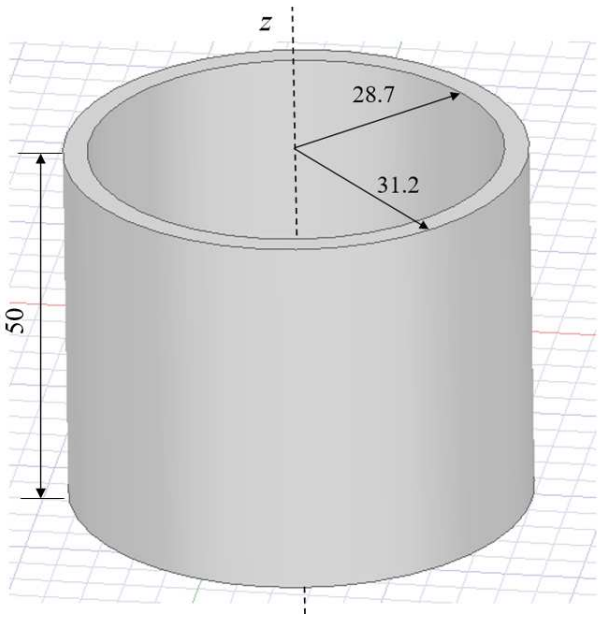


Figura 7.4: Configurazione quotata del giogo esterno. Quote espresse in mm.

Assieme

L'assieme dei vari componenti è mostrato in figura 7.5. La massa totale dell'attuatore risulta di 934 g, inferiore rispetto ai 982 g del modello precedente [5] ed

entro il limite di 1 kg fissato come vincolo. Ogni magnete possiede una massa di 72 g, il nucleo di statore di 478 g, le due bobine insieme hanno massa stimata di 120 g e il giogo esterno di 191 g.

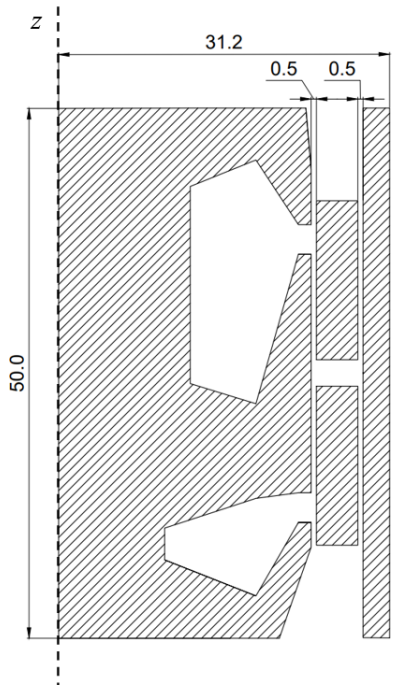


Figura 7.5: Assieme dei componenti dell'attuatore, il quale è ottenuto tramite rotazione di 360° attorno all'asse z. Quote espresse in mm.

Con questa configurazione le sezioni trasversali della bobina superiore e inferiore risultano essere rispettivamente di 181 mm^2 e di 76 mm^2, in rapporto di 2.5:1; i volumi di tali bobine risultano rispettivamente di 19.99 cm^2 e di 7.812 cm^2, anch'essi in rapporto di circa 2.5:1. Con tali valori, è possibile calcolare la densità di corrente, una volta impostate le perdite per effetto Joule, utilizzando la formula

$$\delta = \sqrt{\frac{P_{Loss}}{k_f \cdot \rho_{cu} \cdot (V_{LC} + V_{RC})}} \tag{7.1}$$

In tale formula k_f è il coefficiente di riempimento o fill factor (stimato in 0.5), ρ_{cu} la resistività del rame a 20°C e V_{LC} e V_{RC} i volumi rispettivamente della bobina sinistra e destra. Nello scenario $P_{Loss} = 5$ W si ottiene $\delta = 4.68$ A/mm^2, nello scenario $P_{Loss} = 10$ W si ottiene invece $\delta = 6.62$ A/mm^2. Moltiplicando la densità di corrente per l'area trasversale delle parti attive nelle bobine si ottengono le relative amperespire; la formula utilizzata è quindi

$$N_k I_k = \delta \cdot k_f \cdot A_{bob,k} \tag{7.2}$$

Per la bobina sinistra si ottiene un valore di $N_{LV} I_{LV}$ pari a 423.4 A per lo scenario a 5 W e di 598.8 A per quello a 10 W; per la bobina destra si ottengono valori di $N_{RV} I_{RV}$ pari a 178.4 A e 252.3 A rispettivamente per lo scenario a potenza inferiore e superiore.

Ipotizzando di utilizzare per l'avvolgimento un cavo di rame di sezione di 1 mm^2, nella bobina di sinistra si possono inserire 90 spire, 38 in quella di destra.

7.2 Analisi delle prestazioni

Le forze esplicate da tale configurazione nello scenario 5W sono comparate con quelle ottenute prima del processo di rifinitura, seguito nel paragrafo 6.5, in figura 7.6. Si può notare un aumento della forza nella fase sistolica su entrambi i mover dovuto all'effetto di cogging. Questo aumento non è peggiorativo per quel che riguarda le forze in fase diastolica. In figura 7.7 è riportato lo stesso confronto tra i profili di forza riferito allo scenario a potenza superiore. Si può notare anche qui un miglioramento delle forze medie esplicate. Secondo tali dati, l'attuatore è in grado di svolgere il lavoro cardiaco con $P_{Loss} = 5$ W.

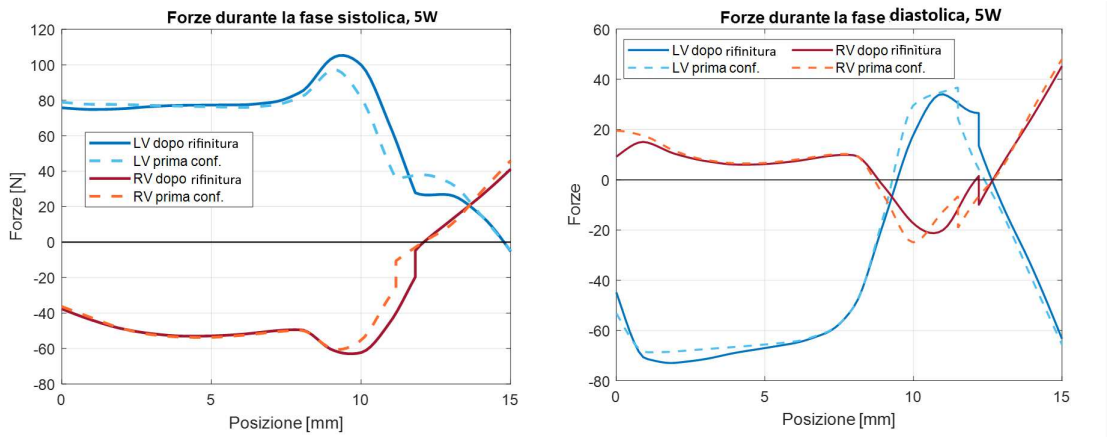


Figura 7.6: Profili di forza sui mover nello scenario 5W della configurazione selezionata prima e dopo il processo di rifinitura presentato nel paragrafo 6.5.

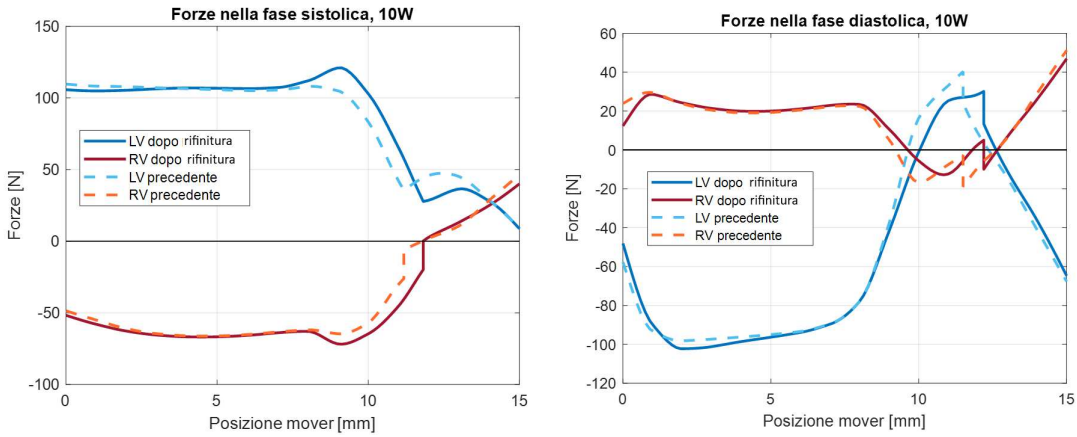


Figura 7.7: Profili di forza sui mover nello scenario 10W della configurazione selezionata prima e dopo il processo di rifinitura presentato nel paragrafo 6.5.

Per la strategia di controllo è prevista l'inversione della polarità di corrente quando i mover si trovano a 11.8 mm di distanza rispetto alla loro posizione

iniziale per la fase sistolica. Per la fase diastolica è possibile utilizzare lo stesso punto di inversione di polarità, oppure adottare l'altra strategia che prevede lo switch nella posizione $x^*_{sw,bck} = 12.2$ mm. Si prosegue nell'analisi adottando quest'ultima strategia.

In figura 7.8 viene riportata la forza di cogging agente sui due mover al variare della loro posizione, ricavata nello stesso modo in cui sono stati ricavati gli altri profili di forza, impostando un valore nullo per le ampere spire dei due avvolgimenti. Si può notare come questa sia molto elevata in prossimità delle aperture di cava mentre, quando i magneti sono a fine corsa diventa negativa, tendendo a richiamarli verso l'interno. Questa forza non è nulla quando i magneti si trovano nella posizione iniziale, in quanto è presente un effetto repulsivo tra i due mover. Ciò significa che sarà necessaria l'applicazione di una corrente per mantenere i magneti fermi nella posizione iniziale.

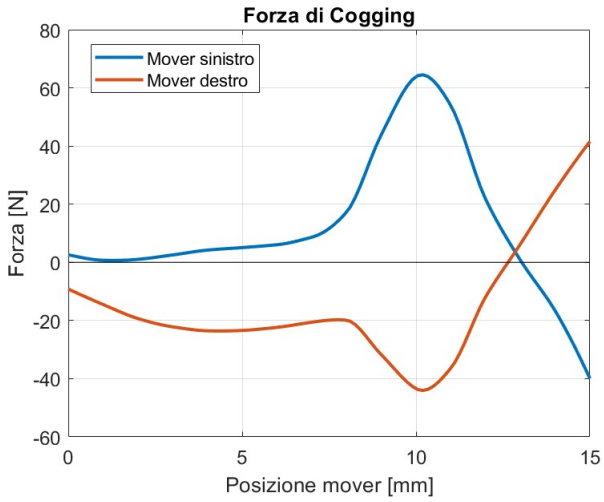


Figura 7.8: Forza di cogging agente sui due mover.

Tramite il software utilizzato è possibile anche calcolare le autoinduttanze (L_{LV} e L_{RV}) e la mutua induttanza dei due avvolgimenti (M) al variare della posizione dei mover. Tali valori sono riportati in figura 7.9 per lo scenario a 5 W. Si possono notare delle variazioni in tali valori dovuti alla saturazione del ferro in determinati punti. Le cose sono simili nello scenario a 10 W, solo che si ha una diminuzione più ampia e marcata in quanto il ferro va incontro maggior saturazione a causa della maggiore corrente circolante negli avvolgimenti. Si può osservare come la saturazione del ferro si presenti in due momenti diversi per le due corse, in particolare nella corsa di andata l'inversione della corrente aiuta a eliminare tale saturazione. I valori di autoinduttanza restano costanti per buona parte della corsa, fattore che può aiutare nel controllo. Si rileva una mutua induttanza significativa tra i due avvolgimenti, di valore negativo.

Un altro parametro che vale la pena considerare è il flusso magnetico a vuoto concatenato con una spira dei due avvolgimenti, al variare della posizione dei magneti, ricavabile dall'analisi FEM. Tale andamento è riportato in figura 7.10(a). La derivata di tale grandezza rispetto a x è proporzionale alla tensione indotta a

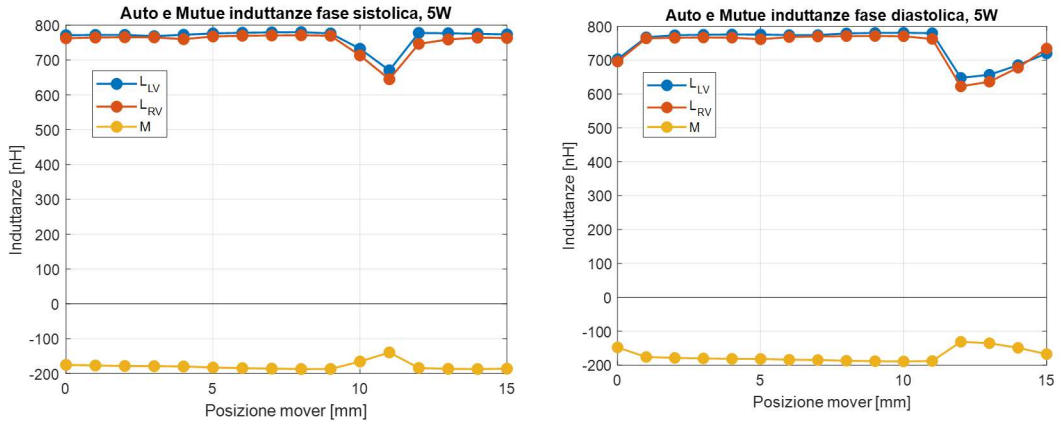


Figura 7.9: Autoinduttanze (L_{LV} e L_{RV}) e mutua induttanza (M) tra i due avvolgimenti al variare della posizione dei mover per la fase sistolica e diastolica nello scenario a 5W.

vuoto su una spira dell'avvolgimento secondo la formula

$$e_k^0 = \frac{d\phi}{dx} \cdot v \tag{7.3}$$

dove v è la velocità del mover. La derivata del flusso concatenato con una spira è visibile in figura 7.10(b). Da questo grafico si evince come si riesca trasmettere una potenza meccanica più elevata ai mover nelle prime fasi della corsa, dove il modulo di $\frac{d\phi}{dx}$ è più elevato. Dove questo valore si abbassa si riuscirà a trasmettere una forza ridotta. Tuttavia la forza risultante non diminuirà troppo perché in quella zona la forza di cogging è più elevata.

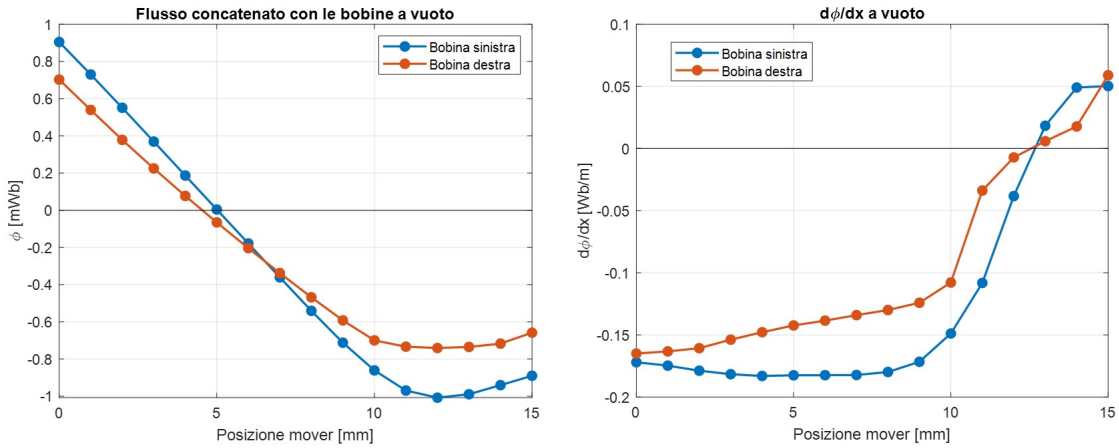


Figura 7.10: Flusso a vuoto concatenato con una spira dei due avvolgimenti statorici in funzione della posizione dei mover e sua derivata rispetto alla posizione del mover.

Conclusioni

Risultati

Con il presente lavoro si consegue un grande miglioramento sulle prestazioni dell'attuatore. Confrontando il design ottenuto con quello di partenza, presentato in [5], si riscontra un miglioramento delle forze su entrambi i mover e una diminuzione delle perdite sugli avvolgimenti. In particolare le forze medie nelle fase sistolica aumentano da 58.1N a 66.4N per il mover sinistro e da 25 N a 36 N per quello destro. Il tutto senza peggiorare la regolarità delle corse e garantendo il soddisfacimento dei vincoli imposti. Inoltre le perdite per effetto Joule risultano diminuite del 54%, passando dai 13W iniziali agli attuali 5W.

Volendo analizzare più nel dettaglio, si può notare che solo una parte di questo miglioramento è attribuibile al processo di ottimizzazione seguito. L'altra parte è dovuta all'inversione della posizione relativa di mover e avvolgimenti. Questo ha infatti permesso di ridurre il volume delle parti attive, diminuendo le perdite per effetto Joule. Per distinguere i due contributi, è utile confrontare la configurazione ottimizzata con quella presentata nel capitolo 3, in cui si erano invertite le posizioni di bobine e mover.

Confrontando tali configurazioni, fissando le perdite del dispositivo a 5 W, si nota un miglioramento della forza media sul mover sinistro nella fase sistolica, del 34%, che aumenta da 49.3 N a 66.4 N. La forza media sul mover destro subisce un incremento del 36% passando da 26.2 N a 35.8 N. Con tali forze medie è possibile svolgere il lavoro richiesto da un cuore sano a riposo di una persona adulta in salute. Aumentando le Ampere spire sugli avvolgimenti ad un valore tale per cui le perdite per effetto Joule risultino di 10 W, si ottiene una forza media sul mover sinistro durante la fase sistolica di 86.5 N e su quello destro di 44.5 N. I ripple delle forze rimangono bassi, garantendo una profilo di forza sufficientemente costante per entrambi i mover.

Bisogna considerare che il calcolo dei profili di forza è stato fatto considerando commutazioni di corrente ideali. Ciò non sarà vero nella pratica a causa delle induttanze presenti nel circuito. Come conseguenza diretta le forze medie avranno un valore leggermente inferiore.

Con la configurazione ottenuta tutti i vincoli sono soddisfatti, e in particolare le perdite ohmiche sono molto ridotte in relazione alla forza esplicata.

Sviluppi futuri

Ora che l'architettura dell'attuatore è definita, è possibile realizzare un prototipo. Servirà poi determinare la strategia di controllo adatta, tenendo conto della

dinamica del dispositivo.

Si procederà dunque al progetto di un controllore che possa adattare le prestazioni del dispositivo alle effettive esigenze del paziente e al test pratico, utilizzando un secondo apparecchio che riproduca la forza resistente prodotta dalla pressione delle arterie.

Processo di ottimizzazione

Scopo secondario di questo progetto è quello di mettere a punto un processo di ottimizzazione generalizzabile in situazioni anche molto diverse rispetto a quella in esame. Tale processo, che è quello seguito nell'arco di questo lavoro, può essere così riassunto:

1. Individuazione delle variabili di progetto da poter variare al fine di ottimizzare degli indici di prestazione. Questo passaggio è fondamentale, in quanto tali indici devono essere scelti molto bene, perché è su di essi che viene tarata l'ottimizzazione. Per poter impostare degli indici significativi è necessario conoscere molto bene il progetto e capire quali sono gli output più importanti.

2. Definizione del range di variazione di ogni variabile di progetto.

3. Utilizzo del metodo di Taguchi e dell'analisi della varianza per determinare le variabili più influenti sulle prestazioni. Tale metodo, mediante la conduzione di esperimenti secondo una matrice ortogonale permette di conoscere l'andamento degli indici di prestazione in funzione delle variabili di progetto.

4. Eliminazione delle variabili di progetto meno influenti, impostandole al valore più conveniente in termini di indici di prestazione secondo l'analisi di Taguchi.

5. Eventuale restringimento dei range per le variabili rimaste.

6. Conduzione degli esperimenti con il nuovo range di variazione secondo una matrice ortogonale. Questo costituirà il training set per la costruzione di un modello predittivo degli indici di prestazione, riducendo notevolmente il tempo di calcolo del singolo esperimento. Per avere un training set significativo, questo deve essere abbastanza esteso. Per aumentarne la numerosità è possibile condurre una seconda volta gli esperimenti secondo la stessa matrice ortogonale, assegnando ad ogni variabile una colonna diversa rispetto alla prima, in modo da non ripetere esperimenti già fatti.

7. Conduzione di un numero adeguato di esperimenti (almeno pari al 30% di quelli effettuati nel punto 6), impostando le variabili di progetto a valori random all'interno dei loro range di variazione. Questi dati verranno utilizzati come test set per i modelli predittivi.

8. Costruzione dei modelli predittivi utilizzando tecniche di machine learning.

9. Calcolo del fronte di Pareto mediante un algoritmo ibrido stocastico-diretto, utilizzando i modelli predittivi costruiti al punto 8.

10. Validazione dei risultati sul fronte di Pareto tramite la conduzione di esperimenti. Se i risultati dei modelli predittivi non sono ritenuti soddisfacenti, utilizzare questi nuovi dati come dati di training per il riaddestramento dei modelli.

11. Scelta della combinazione più conveniente dei parametri sulla base delle informazioni ricavate dal fronte di Pareto.

12. Restringimento dei range delle variabili nell'intorno della configurazione scelta.

13. Costruzione di modelli predittivi ottimizzati per la predizione degli indici in questo nuovo range.

14. Calcolo del Fronte di Pareto in questa zona, validazione dei risultati ed eventuale riaddestramento dei modelli predittivi utilizzando questi nuovi dati. Ripetere fino a che non si ottengono prestazioni adeguate dei modelli.

15. Scelta della configurazione più conveniente su questo nuovo fronte ristretto.

Il calcolo di un nuovo fronte di Pareto più ristretto consente di aumentare la precisione dei modelli. In quella zona,infatti, la variabilità degli indici di prestazione è più bassa, il che porta a modelli più accurati e ad un fronte più definito e preciso.

Questo processo è generalizzabile e può essere applicato a qualsiasi problema di ottimizzazione multi obiettivo. In particolare è consigliato quando le variabili di progetto sono molte e i tempi di calcolo di un singolo esperimento sono relativamente elevati.

Una delle criticità del processo consiste nella creazione di modelli predittivi accurati, per i quali potrebbero essere richieste diverse prove, soprattutto perché è spesso necessario un training test ampio. Per questo motivo si consiglia l'utilizzo del GPR, in quanto fornisce buone prestazioni con un set di dati più limitato rispetto a quello richiesto per l'addestramento di una rete neurale affidabile.

Bibliografia

[1] Michael A Laflamme and Charles E Murry. Heart regeneration. *Nature*, 473(7347):326–335, 2011.

[2] Robert ES Bowen, Thomas J Graetz, Daniel A Emmert, and Michael S Avidan. Statistics of heart failure and mechanical circulatory support in 2020. *Annals of translational medicine*, 8(13), 2020.

[3] Philipp Angleitner, Alexandra Kaider, Jacqueline M Smits, Arezu Z Aliabadi-Zuckermann, Emilio Osorio-Jaramillo, Günther Laufer, and Andreas O Zuckermann. The adapted heart donor score. *Transplant International*, 34(3):546–560, 2021.

[4] Mauro Andriollo, Enrico Fanton, and Andrea Tortella. A review of innovative electromagnetic technologies for a totally artificial heart. *Applied Sciences*, 13(3):1870, 2023.

[5] Mauro Andriollo, Enrico Fanton, Michele Forzan, and Andrea Tortella. Design and analysis of a dual mover linear oscillating actuator for a totally artificial heart. In *2022 International Conference on Electrical Machines (ICEM)*, pages 408–414. IEEE, 2022.

[6] G. Anastasi, S. Capitani, M.L. Carnazza, S. Cinti, et al. *Trattato di anatomia umana - Volume 1*. Cap. 4: Apparato circolatorio sanguigno. EdiErmes, Milano, IV ed., 2006.

[7] H. Pape, A. Kurtz, and S. Silberngal. *Physiologie*. Georg Thieme Verlag KG - Stuttgart - Deutschland, VII ed., 2014.

[8] Hidekatsu Fukuta and William C Little. The cardiac cycle and the physiologic basis of left ventricular contraction, ejection, relaxation, and filling. *Heart failure clinics*, 4(1):1–11, 2008.

[9] Wikipedia. Diagramma di wiggers - wikipedia, l'enciclopedia libera. `http://it.wikipedia.org/w/index.php?title=Diagramma_di_Wiggers&oldid=101920483`, 2019. [Online; in data 8-agosto-2023].

[10] ANDREW N Redington, HUON H Gray, MARGARET E Hodson, ML Rigby, and PJ Oldershaw. Characterisation of the normal right ventricular pressure-volume relation by biplane angiography and simultaneous micromanometer pressure measurements. *Heart*, 59(1):23–30, 1988.

[11] Grassi S. Meccanica cardiaca, 2014. Università degli studi di perugia.

[12] Theresa A McDonagh, Marco Metra, Marianna Adamo, Roy S Gardner, Andreas Baumbach, Michael Böhm, Haran Burri, Javed Butler, Jelena Čelutkienė, Ovidiu Chioncel, et al. 2021 esc guidelines for the diagnosis and treatment of acute and chronic heart failure: Developed by the task force for the diagnosis and treatment of acute and chronic heart failure of the european society of cardiology (esc) with the special contribution of the heart failure association (hfa) of the esc. *European heart journal*, 42(36):3599–3726, 2021.

[13] Vinay Kumar, Abul K Abbas, Nelson Fausto, and Jon C Aster. *Robbins and Cotran pathologic basis of disease*. Elsevier health sciences, 2014.

[14] Michael R Bristow and Brian D Lowes. Trattamento dell'insufficienza cardiaca.

[15] Ministero della salute. Report trapianti e donazione di organi, tessuti e cellule. l'attività della rete nazionale trapianti 2021. `https://www.trapianti.salute.gov.it/imgs/C_17_cntPubblicazioni_506_allegato.pdf`, 2022. [Online; accessed 11-Aug-2023].

[16] Piotr Ponikowski, Stefan D Anker, Khalid F AlHabib, Martin R Cowie, Thomas L Force, Shengshou Hu, Tiny Jaarsma, Henry Krum, Vishal Rastogi, Luis E Rohde, et al. Heart failure: preventing disease and death worldwide. *ESC heart failure*, 1(1):4–25, 2014.

[17] Donald M Lloyd-Jones, Martin G Larson, Eric P Leip, Alexa Beiser, Ralph B D'agostino, William B Kannel, Joanne M Murabito, Ramachandran S Vasan, Emelia J Benjamin, and Daniel Levy. Lifetime risk for developing congestive heart failure: the framingham heart study. *Circulation*, 106(24):3068–3072, 2002.

[18] Salim S Virani, Alvaro Alonso, Hugo J Aparicio, Emelia J Benjamin, Marcio S Bittencourt, Clifton W Callaway, April P Carson, Alanna M Chamberlain, Susan Cheng, Francesca N Delling, et al. Heart disease and stroke statistics-2021 update: a report from the american heart association. *Circulation*, 143(8):CIR0000000000000950, 2021.

[19] Jaimin R Trivedi, Allen Cheng, Ramesh Singh, Matthew L Williams, and Mark S Slaughter. Survival on the heart transplant waiting list: impact of continuous flow left ventricular assist device as bridge to transplant. *The Annals of thoracic surgery*, 98(3):830–834, 2014.

[20] Daniel Timms. A review of clinical ventricular assist devices. *Medical engineering & physics*, 33(9):1041–1047, 2011.

[21] Martina Todesco, Carlo Zardin, Laura Iop, Tiziana Palmosi, Pietro Capaldo, Filippo Romanato, Gino Gerosa, and Andrea Bagno. Hybrid membranes for

the production of blood contacting surfaces: physicochemical, structural and biomechanical characterization. *Biomaterials Research*, 25(1):26, 2021.

[22] Hiroshi Akima. A new method of interpolation and smooth curve fitting based on local procedures. *Journal of the ACM (JACM)*, 17(4):589–602, 1970.

[23] Hiroshi Akima. A method of bivariate interpolation and smooth surface fitting for irregularly distributed data points. *ACM Transactions on Mathematical Software (TOMS)*, 4(2):148–159, 1978.

[24] André Pohlmann, Marc Leßmann, Andreas Fritschi, Thomas Finocchiaro, Ulrich Steinseifer, and Kay Hameyer. Experimental validation of the linear drive train for a total artificial heart system. *Mechatronics*, 23(2):222–226, 2013.

[25] André Pohlmann, Marc Lessmann, Thomas Finocchiaro, Thomas Schmitz-Rode, and Kay Hameyer. Numerical computation can save life: Fem simulations for the development of artificial hearts. *IEEE transactions on magnetics*, 47(5):1166–1169, 2010.

[26] Gabor Kovacs, Andrea Berghold, Stefan Scheidl, and HJERJ Olschewski. Pulmonary arterial pressure during rest and exercise in healthy subjects: a systematic review. *European Respiratory Journal*, 34(4):888–894, 2009.

[27] Thomas Finocchiaro, Thomas Butschen, Paul Kwant, Ulrich Steinseifer, Thomas Schmitz-Rode, Kay Hameyer, and Marc Leßmann. New linear motor concepts for artificial hearts. *IEEE Transactions on Magnetics*, 44(6):678–681, 2008.

[28] Ranjit K Roy. *A primer on the Taguchi method*. Society of Manufacturing Engineers, 2010.

[29] TB Asafa, Naji Tabet, and SAM Said. Taguchi method–ann integration for predictive model of intrinsic stress in hydrogenated amorphous silicon film deposited by plasma enhanced chemical vapour deposition. *Neurocomputing*, 106:86–94, 2013.

[30] Dali Vilma Francis, T Aiswarya, and Trupti Gokhale. Optimization of the incubation parameters for biogenic synthesis of wo3 nanoparticles using taguchi method. *Heliyon*, 8(9), 2022.

[31] Bappa Acherjee. Laser transmission welding of dissimilar plastics: analyses of parametric effects and process optimization using grey-based taguchi method. In *Modern manufacturing processes*, pages 131–144. Elsevier, 2020.

[32] Bahadir Erman Yuce, Peter Vilhelm Nielsen, and Pawel Wargocki. The use of taguchi, anova, and gra methods to optimize cfd analyses of ventilation performance in buildings. *Building and Environment*, 225:109587, 2022.

[33] TB Asafa, Naji Tabet, and SAM Said. Taguchi method–ann integration for predictive model of intrinsic stress in hydrogenated amorphous silicon film deposited by plasma enhanced chemical vapour deposition. *Neurocomputing*, 106:86–94, 2013.

[34] Sunita Parinam, Mukesh Kumar, Neelam Kumari, Vinod Karar, and Amit L Sharma. An improved optical parameter optimisation approach using taguchi and genetic algorithm for high transmission optical filter design. *Optik*, 182:382–392, 2019.

[35] M. Andriollo, G. Martinelli, A. Morini, and A. Tortella. *I trasformatori, teoria ed esercizi*. Società editrice Esculapio, 2016.

[36] Nicola Bianchi, Silverio Bolognani, et al. *Metodologie di progettazione delle macchine elettriche*, volume 1. CLEUP Editrice, 2001.

[37] Chin-Ping Fung. Manufacturing process optimization for wear property of fiber-reinforced polybutylene terephthalate composites with grey relational analysis. *wear*, 254(3-4):298–306, 2003.

[38] Soniya Lalwani, Sorabh Singhal, Rajesh Kumar, and Nilama Gupta. A comprehensive survey: Applications of multi-objective particle swarm optimization (mopso) algorithm. *Transactions on combinatorics*, 2(1):39–101, 2013.

[39] Phil Kim. Matlab deep learning, with machine learning, neural networks and artificial intelligence. 2017.

[40] Collins Ayuya. Regularization to prevent overfitting. `https://www.section.io/engineering-education/regularization-to-prevent-overfitting/` [Consultato il 30/08/2023].

[41] Christopher KI Williams and Carl Edward Rasmussen. *Gaussian processes for machine learning*, volume 2. MIT press Cambridge, MA, 2006.

[42] Thomas Beckers. An introduction to gaussian process models. *arXiv preprint arXiv:2102.05497*, 2021.

[43] Mathworks: Select data for regression or open saved app session. `https://it.mathworks.com/help/stats/select-data-and-validation-for-regression-problem.html` [Consultato il 30/08/2023].

[44] Mathworks: rand, uniformly distributed random numbers genartor. `https://it.mathworks.com/help/matlab/ref/rand.html` [Consultato il 30/08/2023].

[45] Mathworks: Bayesian optimization workflow. `https://it.mathworks.com/help/stats/bayesian-optimization-workflow.html` [Consultato il 31/08/2023].

[46] D Kalyanmoy. Multi-objective optimization using evolutionary algorithms john wiley & sons, 2001.

[47] Mathworks: paretosearch algorithm. https://it.mathworks.com/help/gads/paretosearch-algorithm.html?searchHighlight=paretosearch%20algorithm&s_tid=srchtitle_support_results_1_paretosearch%2520algorithm [Consultato il 02/09/2023].