



UNIVERSITA' DI PADOVA
FACOLTA' DI INGEGNERIA

SCUOLA DI DOTTORATO DI RICERCA IN
INGEGNERIA INDUSTRIALE

INDIRIZZO ENERGETICA

CICLO: XXI

**CRITERI DI PROGETTAZIONE
E PROCEDURE DI
OTTIMIZZAZIONE DEI
CONVOGLIATORI PALETTATI
DELLE POMPE PLURISTADIO.**

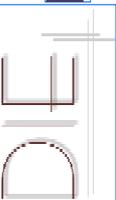
DIRETTORE DELLA SCUOLA:
Chiar.mo Prof. Paolo Bariani

SUPERVISORE:
Chiar.mo Prof. Guido Ardizzone

DOTTORANDO: Francesco De Zan

31 gennaio 2009

DIPARTIMENTO DI INGEGNERIA ELETTRICA



UNIVERSITA' DEGLI STUDI DI PADOVA

Dipartimento di Ingegneria Elettrica

Scuola di Dottorato di Ricerca in Ingegneria Industriale

Indirizzo Energetica

Ciclo XXI

**CRITERI DI PROGETTAZIONE E PROCEDURE
DI OTTIMIZZAZIONE DEI CONVOGLIATORI
PALETTATI DELLE POMPE PLURISTADIO.**

Direttore della Scuola: Chiar.mo Prof. Paolo Francesco Bariani

Supervisore: Chiar.mo Prof. Guido Ardizzon

Dottorando: Francesco De Zan

31 gennaio 2009

Sommario

Le normative in fase di approvazione in sede europea introdurranno a breve nuovi e più restrittivi vincoli all'efficienza delle apparecchiature che fanno uso di energia elettrica. Tali normative imporranno alle aziende produttrici la riprogettazione di una buona parte delle pompe attualmente in produzione.

In questa tesi si propone una procedura parametrica di progettazione delle pompe sommerse multistadio per integrare le tecniche di progettazione attualmente a disposizione. La procedura proposta consente un'interfaccia agevole con i codici di calcolo CFD e con le tecniche di ottimizzazione. Inoltre si propone una procedura per la creazione in sede industriale di banche dati di simulazioni CFD per caratterizzare il funzionamento e le prestazioni di componenti o particolari che devono essere progettati. Contestualmente alla creazione della banca dati si propone una metodologia di design of experiments per la corretta pianificazione di tali banche dati.

Si è proceduto alla progettazione di uno stadio di pompa verticale sommersa a flusso misto facendo uso delle procedure proposte. Inoltre si è pianificata e realizzata una banca dati per caratterizzare le prestazioni dei diffusori-convogliatori delle pompe sommerse a flusso misto; l'accesso a tale banca dati ha consentito la scelta di un ingombro assiale corretto all'interno della progettazione.

La soluzione proposta in sede di progettazione è stata ingegnerizzata e verificata tramite codici CFD nella sua configurazione finale. La verifica numerica ha consentito di indagare l'effetto dei particolari introdotti in sede di ingegnerizzazione. Si è riscontrato che alcuni particolari che vengono solitamente introdotti in fase di progettazione e sono ritenuti solitamente ininfluenti, modificano sensibilmente il campo di moto e pregiudicano il corretto accoppiamento di girante e diffusore.

Un prototipo della soluzione proposta è stato realizzato in materiale plastico e sono state effettuate delle prove di collaudo su una pompa sommersa ad uno e a due stadi. I dati sperimentali si discostano sensibilmente dai dati ricavati dalle simulazioni numeriche. Non è stato possibile effettuare tutte le verifiche necessarie per giustificare la discrepanza tra i dati numerici e quelli sperimentali. Alcuni risultati parziali hanno comunque fornito importanti indicazioni su come deve essere impostata correttamente una simulazione numerica di una pompa sommersa.

Abstract

European Community will soon introduce new and more restrictive regulations on the efficiency of products using electric energy. These regulations will impose to pump manufacturers the redesign of several pumps already in production.

In this thesis a parametric procedure for the design of multistage submersible pumps aimed at integrating the present design criteria is presented. The proposed procedure allows an easy interaction with the CFD codes and with the optimization techniques. Furthermore, a procedure for the creation of databases of numerical analysis results was developed to describe the performance of parts or details of the machine that have to be designed. Together with this procedure, to properly generate the inputs for the database creation, a DOE technique was also developed and applied.

A stage of a multistage submersible mixed-flow pump was designed by means of the proposed procedure and a database for the characterization of the performance of the meridional channels of mixed-flow interstage interstage was built. The database allowed the proper choice of the axial extension of the interstage channel.

The obtained solution was added with all the manufacturing devices, and the final configuration was numerically analysed by the CFD code. Some manufacturing details that are usually considered negligible in the design practice, showed to have a great influence on the flow field and to avoid the correct interaction between the stages.

A plastic prototype was realized by rapid prototyping technique and experimental tests were carried out on a single stage and a two stage submersible pump. The experimental results were found to be very different from the numerical results. However, there was not enough time for further analyses aimed at understanding the reason of such a difference, although some partial results gave interesting indication about how to properly set the parameters of a numerical simulation on a submersible pump to fully describe the flow field inside it.

Dottorato a tema di ricerca vincolato finanziato da
CALPEDA S.p.a.

Indice

Introduzione	4
Contenuti della tesi	6
1 Stato dell'arte	9
1.1 Classificazione e layout dei canali interstadio della macchine multistadio	9
1.2 Criteri di progettazione-1	10
1.2.1 Stepanoff	11
1.2.2 Lazarkiewicz e Trokolansky	12
1.2.3 Meng e Jackson	13
1.2.4 Ellis	15
1.3 Criteri di progettazione-2	17
1.3.1 Tecniche di progettazione inversa	18
1.3.2 Tecniche di ottimizzazione	20
1.4 Esigenze industriali	22
1.4.1 Interazione con i codici di calcolo fluidodinamici	23
1.4.2 Raccolta di informazioni	24
1.4.3 Progettazione di serie di pompe	25
2 Progettazione parametrica	27
2.1 Strumenti matematici per la parametrizzazione-Curve di Bezier	28
2.1.1 Curve di Bezier	29
2.2 Parametrizzazione della palettatura della girante e del canale interstadio	31
2.2.1 Diverse scelte per la parametrizzazione	31
2.2.2 Parametrizzazione della palettatura della girante-Giranti a basso numero tipico	32
2.2.3 Parametrizzazione della palettatura della girante	37
2.2.4 Palettatura del diffusore	40
2.3 Parametrizzazione della forma del condotto meridiano del diffusore	46
3 Algoritmi di ottimizzazione	51
3.1 Differenti approcci all' ottimizzazione	52
3.1.1 Ottimizzazione pura	52
3.1.2 Ottimizzazione assistita da metamodelli	52

3.2	Ottimizzazione vincolata	53
3.3	Simulated Annealing	53
3.4	Algoritmi genetici-GA	55
3.4.1	Elementi di base di un GA	56
3.4.2	Strategie di selezione	58
3.4.3	Accoppiamento e mutazione	59
3.4.4	Codifica reale dei GA	60
3.5	Particle Swarm Optimization (PSO)	61
3.5.1	PSO modificati	62
3.5.2	PSO vincolato	64
3.5.3	Condizione di arresto	64
4	DOE e modelli di regressione	69
4.1	Design of Experiments	69
4.1.1	Curse of dimensionality	71
4.2	Modelli di regressione	71
4.2.1	Addestramento	73
4.2.2	Stima dell'accuratezza dell'approssimazione e confronto tra modelli di regressione	74
4.2.3	Overfitting	77
4.2.4	Superfici di risposta	79
4.2.5	Reti Neurali (Artificial Neural Networks-ANN)	80
4.2.6	Reti di funzioni radiali	85
4.2.7	Kriging	90
5	Utilizzo industriale delle tecniche di ottimizzazione	95
5.1	Banca dati	96
5.2	DOE-Ipercubi latini stratificati	98
5.3	Accesso alla banca dati	101
6	Progettazione di uno stadio di pompa sommersa	103
6.1	Dati di progetto e vincoli produttivi	103
6.1.1	La realizzazione di pompe centrifughe tramite stampaggio di materie plastiche	104
6.2	Analisi fluidodinamiche	108
6.2.1	Meshatura	108
6.2.2	Pre-Processing	109
6.2.3	Solving	110
6.2.4	Post-Processing	110
6.3	Esecuzione della progettazione	111
6.3.1	Fasi della progettazione	111
6.4	Progettazione della girante	112

6.5	Scelta del canale meridiano del diffusore	127
6.5.1	Schema del protocollo	129
6.5.2	Linguaggio di gestione	129
6.5.3	Aspetti peculiari:meshatura.	132
6.5.4	Aspetti peculiari:analisi dei risultati	138
6.5.5	Creazione della banca dati	141
6.5.6	Accettazione della banca dati	149
6.5.7	Utilizzo dei dati: scelta del canale meridiano D	152
6.6	Progettazione delle pale del diffusore-convogliatore	156
6.6.1	Tipologia di diffusore scelta	158
6.6.2	Progettazione del diffusore	159
7	Simulazioni numeriche complete e collaudo del progetto proposto	165
7.1	Funzionamento in sede di fuori progetto	165
7.2	Prototipo di pompa sommersa e impianto di collaudo.	168
7.3	Discretizzazione della macchina in configurazione di collaudo e impostazioni delle simulazioni	172
7.4	Risultati numerici	175
7.4.1	Osservazioni sui risultati numerici	178
7.5	Dati sperimentali di collaudo e confronto con i risultati numerici	183
	Conclusioni	189
	Simbologia	191
	Elenco delle figure	192
	Bibliografia	197

Introduzione

La borsa di studio è stata finanziata dall'azienda Calpeda S.p.a. Il tema di ricerca nasce dall'esigenza di avere a disposizione dei criteri di progettazione per i diffusori-convogliatori delle macchine multistadio che offra un approccio razionale alla progettazione di tali organi utilizzando le moderne tecniche CFD e di ricerca dell'ottimo. Tuttavia, dato che la progettazione integrata di girante e canale interstadio è un requisito fondamentale per ottenere elevate prestazioni dalle pompe multistadio, risulta che la corretta progettazione del diffusore non può prescindere dalla progettazione della girante. Nella presente tesi quindi viene proposta una procedura completa per la progettazione di tali macchine comprensiva anche della progettazione della girante e non esclusivamente rivolta ai canali interstadio.

I livelli di efficienza raggiunti dalle pompe progettate con i criteri di progettazione classici, frutto di decine di anni di esperienza, sono tuttora soddisfacenti. Tuttavia in sede europea sono in fase di approvazione nuove normative molto restrittive per quello che riguarda gli *energy using products* (EUP). Queste normative introdurranno nuovi vincoli più restrittivi sull'efficienza minima delle apparecchiature che fanno uso di energia. E' necessario quindi affiancare ai criteri di progettazione standard anche le tecniche proposte nello stato dell'arte della letteratura tecnico-scientifica. Negli ultimi anni un contributo fondamentale è stato dato dall'introduzione dei software di analisi CFD, che hanno consentito di indagare la fenomenologia fluidodinamica all'interno delle macchine. Recentemente a questi sono stati associati algoritmi di ricerca dell'ottimo come gli algoritmi genetici o l'ottimizzazione per sciami di particelle (EA,PSO). Questi consentono di ricercare, grazie ad un uso massivo di simulazioni CFD, le geometrie che massimizzano le prestazioni. Questi inoltre sono spesso assistiti da metamodelli matematici o statistici (Reti Neurali Artificiali, Reti di Funzioni a Base Radiale, Kriging) e da metodologie di *design of experiments*(DOE) per ridurre lo sforzo computazionale.

In questa tesi si propone un metodo parametrico per la definizione delle palettatura della girante e per la palettatura dei diffusori-convogliatori delle macchine multistadio. Per le giranti si distingue tra geometria delle pale a semplice curvatura, per le quali si propone una parametrizzazione basata sul controllo del momento angolare della quantità di moto del fluido, e geometria a doppia curvatura, per le quali si parametrizzano le forme delle sezioni di riferimento sul piano conforme tramite curve di Bezier. Per il diffusore-convogliatore si propongono due diverse parametrizzazioni della forma delle pale; la prima prevede pale a spessore costante con forma assegnata in modo analogo a quanto fatto per le pale delle giranti a semplice curvatura, la seconda prevede pale a spessore variabile la cui forma è ricavata da

un'assegnato andamento delle velocità di attraversamento nel condotto interpalare.

Si propone un metodo per la pianificazione e la realizzazione di banche dati di simulazioni CFD sviluppato a partire dall'osservazione delle esigenze industriali. I risultati di tali banche dati vengono generalizzati tramite metodi di regressione nei quali accedere con algoritmi di ricerca dell'ottimo, al fine di immagazzinare informazioni da utilizzare nella progettazione di più macchine simili. Per consentire un'adeguata pianificazione delle simulazioni i cui risultati andranno a comporre la banca dati si propone una nuova tecnica DOE, che consente di infittire progressivamente il campione di dati su cui addestrare i metamodelli massimizzando l'efficienza dello sforzo di campionamento.

La procedura proposta nella tesi è stata infine applicata ad un caso pratico di progettazione di uno stadio di pompa sommersa multistadio a flusso misto proposto da Calpeda S.p.a. Preliminarmente alla progettazione esecutiva della girante e del diffusore-riconvogliatore è stata pianificata e realizzata una banca dati di risultati di simulazioni di dei canali meridiani dei diffusori delle pompe multistadio a flusso misto. L'accesso ai dati raccolti ha consentito una scelta ponderata dell'ingombro assiale in relazione alle perdite di natura fluidodinamica interne al condotto.

La soluzione proposta è stata prototipata ed è stato effettuato il collaudo. Inoltre il modello della pompa sommersa nella configurazione di collaudo è stato discretizzato in tutti i particolari e verificato grazie ad una complessa simulazione CFD. Il confronto tra il collaudo e la simulazione numerica suggerisce numerosi studi che non è stato possibile affrontare visti i tempi e che ci si propone di affrontare nel futuro.

Contenuti della tesi

Il lavoro raccolto nella tesi si sviluppa per capitoli. I contenuti sono di seguito raccolti schematicamente nel seguente modo:

- **Capitolo 1 Stato dell'arte:** il primo capitolo consiste nell'analisi dello stato dell'arte per quello che riguarda le tecniche di progettazione delle turbomacchine multistadio e le tecniche di ottimizzazione applicate alle simulazioni al computer; una sezione è dedicata all'analisi delle esigenze della produzione industriale;
- **Capitolo 2 Progettazione parametrica:** in questo capitolo vengono presentate le parametrizzazioni per la palettatura delle giranti e del diffusore; viene inoltre presentata una parametrizzazione della forma dei condotti meridiani dei canali di ritorno utilizzati nelle pompe sommerse multistadio a flusso misto;
- **Capitolo 3 Algoritmi di ottimizzazione:** questa parte è dedicata alla descrizione dei principali algoritmi per la ricerca dell'ottimo utilizzabili con le simulazioni al computer (Simulate Annealing, Algoritmi Genetici, Particle Swarm Optimization);
- **Capitolo 4 *Design of Experiments* e modelli di regressione:** in questo capitolo sono raccolte le principali tecniche di regressione di dati utilizzati per ap-

prossimare i risultati delle simulazioni al computer; inoltre una sezione è dedicata al design of experiments utilizzato in questo campo;

- **Capitolo 5 Utilizzo industriale delle tecniche di ottimizzazione:** descrizione dell'approccio proposto in questa tesi per l'utilizzo delle tecniche di ottimizzazione e dei modelli di regressione; tale approccio prevede la creazione di banche dati per la caratterizzazione delle prestazioni dei componenti delle turbomacchine tramite simulazioni che consentono di immagazzinare dati da utilizzare in più situazioni;
- **Capitolo 6 Progettazione di uno stadio di pompa sommersa:** vengono riportati tutti i passi che sono stati necessari per la progettazione di uno stadio di pompa verticale sommersa, con particolare attenzione all'interazione tra la girante e la palettatura del diffusore; inoltre viene realizzata una banca dati di simulazioni al computer per caratterizzare le prestazioni dei canali meridionali dei diffusori da accoppiarsi con giranti a flusso misto;
- **Capitolo 7 Simulazioni numeriche complete e collaudo del progetto proposto:** in questo capitolo viene presentato il prototipo realizzato a partire dalla soluzione presentata nel capitolo precedente e i dati ricavati dal collaudo di questa macchina; il modello della pompa, nella sua configurazione di collaudo, è stato discretizzato e sono state realizzate delle simulazioni per indagare la fenomenologia interna della macchina; i risultati delle simulazioni hanno messo in luce l'influenza dei dettagli costruttivi sulle prestazioni e il confronto con i dati sperimentali hanno fornito numerose indicazioni su come devono essere impostate nel modo più corretto queste simulazioni.

Capitolo 1

Stato dell'arte

1.1 Classificazione e layout dei canali interstadio della macchine multistadio

Le turbomacchine multistadio possono essere divise sostanzialmente in due gruppi in base al flusso interno alla girante. Infatti si riscontrano differenze sostanziali tra macchine con girante a flusso radiale e macchine con girante a flusso misto.

La classificazione secondo il numero tipico di macchina tipica delle turbomacchine prevede macchine a flusso radiale per numeri tipici $K < 1.5$. Oltre 1.5 fino a 2.5 si fa generalmente uso di macchine con condotto meridiano con un flusso inclinato rispetto alla direzione assiale e perciò si parla di macchine a flusso misto. Ancora oltre si entra nel campo delle macchine assiali. Nella pratica industriale che riguarda le turbomacchine multistadio sommerse la classificazione proposta prima non è sempre rispettata. Per contenere gli ingombri radiali in alcuni casi vengono utilizzate giranti a flusso misto là dove il numero tipico di macchina suggerirebbe giranti a flusso radiale, il tutto a scapito dell'efficienza.

Nella pratica industriale i canali interstadio delle turbomacchine multistadio assumono configurazioni diverse. Questi sono progettati con configurazioni diverse a seconda del flusso in uscita alla girante; questo per l'ovvia ragione che questi organi devono accogliere il flusso in uscita dalla girante e quindi devono possedere una forma che fisiologicamente si adatti a quella della giranti.

Per le giranti con flusso puramente radiale generalmente si fa uso di canali di interstadio a sviluppo radiale, quindi composti da un tratto radiale centrifugo a valle della girante che ne costituisce la naturale prosecuzione e che può essere occupato o meno da una schiera di pale diffusive. Di seguito un canale a gomito di 180° costringe la corrente verso un canale con flusso radiale centripeto occupato da una schiera di pale che completa la rimozione della componente tangenziale della velocità assoluta. La figura 1.1 mostra un disegno esplicativo della schematizzazione appena proposta.

La palettatura alloggiata all'interno dei canali interstadio ha configurazioni di tipo diverso. Esistono canali interstadio con palettatura all'interno del solo canale di ritorno, oppure con due diverse schiere nel diffusore e nel canale di ritorno, oppure con un'unica schiera di pale

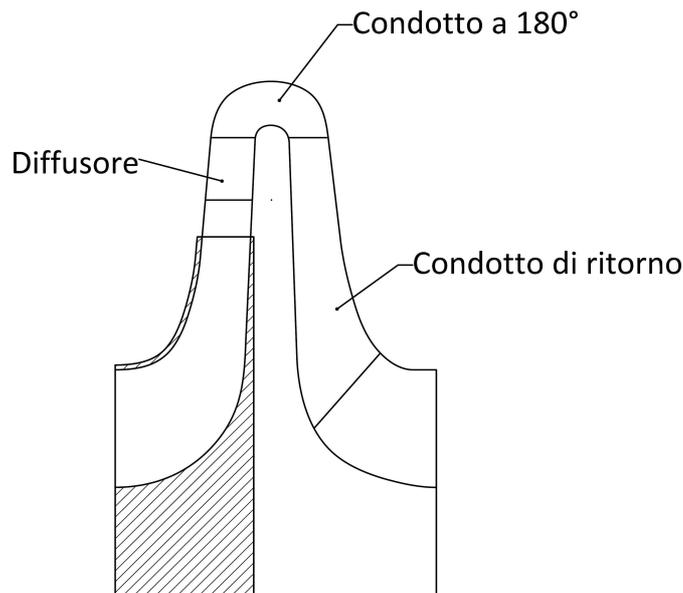


Figura 1.1: Schematizzazione delle varie parti di un canale interstadio a flusso radiale.

continua che parte dal diffusore e occupa anche il gomito a 180° e il canale radiale centripeto. In figura 1.2 sono rappresentati questi tre diversi tipi di configurazioni.

Per le macchine a flusso misto generalmente si fa uso di diffusori con un'unica schiera di pale che accoglie il flusso in uscita dalla girante in direzione mista se non addirittura assiale e lo convoglia verso la girante successiva. Questi diffusori vengono a volte chiamati *bowl diffuser*, letteralmente 'diffusori a boccia', e sono caratterizzati da una forma esterna del condotto meridiano molto raccordata dove in pratica non è possibile distinguere tra diffusore, gomito a 180° e canale di ritorno come in figura 1.3 (vedi [42]).

Esistono in letteratura anche esempi di tipi di accoppiamento diversi di girante e diffusore per soddisfare esigenze diverse. In particolare sono stati utilizzati con successo i diffusori a boccia in accoppiamento con delle giranti a basso numero tipico, quindi con flusso all'uscita radiale (vedi [14]). Questo tipo di diffusore accoppiato con una girante a flusso radiale garantisce ingombri radiali molto più contenuti rispetto ai diffusori con flusso radiale, a scapito però di un maggior ingombro assiale. L'idea che sta dietro a questa soluzione è quella che per deviare la corrente tratti di canale a flusso radiale possono essere sostituiti con dei tratti a sviluppo assiale.

Si veda [13] per una panoramica di alcune soluzioni che riguardano i canali interstadio delle pompe multistadio accoppiati a giranti a flusso radiale.

1.2 Criteri di progettazione-1

I criteri di progettazione che sono stati proposti in letteratura fino agli anni ottanta sono raccolti in [55]. Si riportano in tale sede i risultati riportati dall'autore che risultano utili per il prosieguo della discussione.

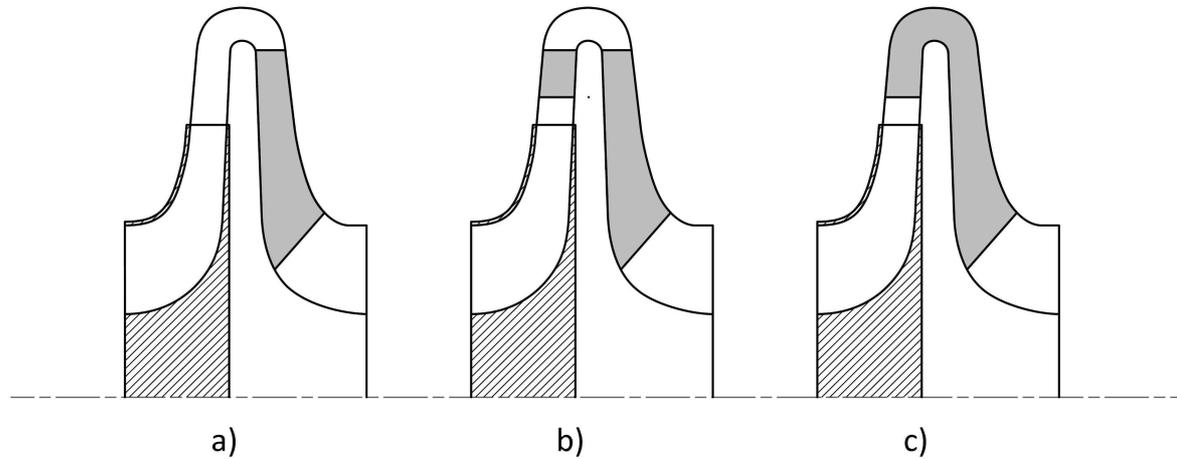


Figura 1.2: Diverse tipologie di palettature nei canali interstadio: a) schiera palare nel solo canale di ritorno; b) doppia schiera palare nel diffusore e nel canale di ritorno; c) schiera palare continua.

1.2.1 Stepanoff

Stepanoff nel suo libro, vedi [43], distingue tra progettazione del solo diffusore radiale e progettazione della schiera del diffusore accoppiata con condotti di ritorno palettati.

Con riferimento alla figura 1.4, basandosi su un'estesa sperimentazione, Stepanoff suggerisce le seguenti direttive per un disegno razionale del diffusore radiale:

- sezione trasversale del diffusore a raggio idraulico massimo, quindi sono preferibili le sezioni quadrate;
- canale di forma rettilinea;
- angolo di divergenza delle superfici palari minori o uguali a quelli tipici dei diffusori

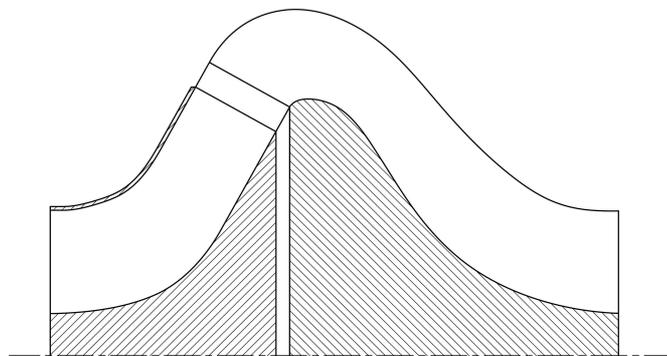


Figura 1.3: Diffusore a boccia per macchine con girante a flusso misto.

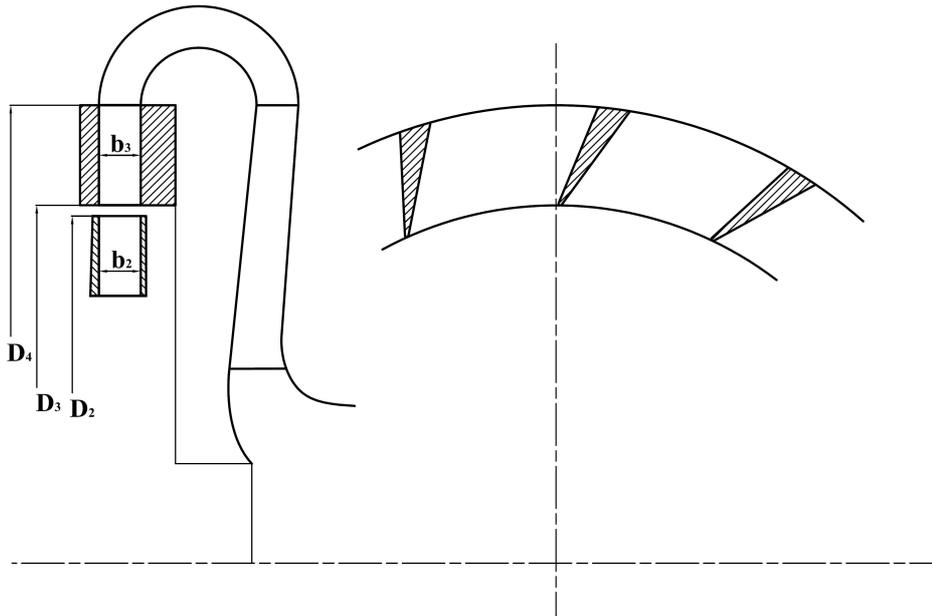


Figura 1.4: Grandezze interessate nella definizione dei canali interstadio secondo Stepanoff.

rettilinei piani: orientativamente 6° , per sezioni rettangolari si può arrivare anche a 11° ;

- la lunghezza della parte conica minore di quattro volte la lunghezza della gola;
- numero di pale minimo per realizzare efficacemente la zona divergente;
- larghezza assiale $b_3 \cong 1.1b_2$;
- rapporto tra i diametri $\frac{D_4}{D_3} = 1.3 \div 1.6$;
- l'area della gola può essere assunta pari all'area della voluta nelle medesime condizioni.

Per quanto riguarda la profilatura della pala del diffusore radiale accoppiato con una pala di canale di ritorno palettato, l'autore consiglia di assumere lo sviluppo di una pala unica, combinando la pala del diffusore con quella del condotto di ritorno, facendo uso del metodo del triangolo degli errori. Il numero delle pale è fissato mediante i criteri che intervengono nella scelta del numero delle pale delle giranti, ed è minore di due o tre unità rispetto a quest'ultimo.

1.2.2 Lazarkiewicz e Trokolansky

La procedura esposta da Stepanoff per il disegno delle pale conduce a perdite per incidenza all'ingresso della palettatura del condotto di ritorno: all'uscita della prima schiera statica si verifica una leggera deviazione a causa del numero finito di pale; all'ingresso della schiera successiva l'ostruzione palare e le disuniformità del moto tendono ad aumentare l'angolo di

flusso. Per tener conto di questo Lazarkiewicz e Troskolansky [44] riportano una procedura originariamente proposta da Pfeleiderer.

Il calcolo delle velocità C_U all'uscita del diffusore viene eseguita con la seguente espressione:

$$C_{U5} = \frac{1}{1 + C_{fd}} \left(C_{U4} + C_{fd} \frac{r_2}{r_4} C_{U3} \right)$$

$$tg\alpha_5 = \frac{C_{M4}}{C_{U5}}$$

in cui:

$$C_{fd} = \frac{r_4^2 \psi_d}{Z_d M_{st}}$$

e ψ_d è un coefficiente che assume valori (0.6 ÷ 0.8); Z_d è il numero di pale del diffusore; M_{st} è il momento statico della linea di corrente rispetto all'asse di rotazione:

$$M_{st} = \int_{r_1}^{r_4} r ds.$$

L'angolo di ingresso nel condotto di ritorno viene aumentato:

$$tg\alpha_6 = tg\alpha_5 k \frac{1}{\xi}$$

in cui ξ è l'ostruzione palare; $k \cong 1.2$ è un coefficiente che tiene conto della disuniformità del moto.

Gli stessi autori consigliano di assegnare alle pale, all'uscita del condotto di ritorno, un'ulteriore deviazione di $3^\circ \cong 5^\circ$, onde evitare residui tangenziali.

1.2.3 Meng e Jackson

Quello presentato in [31] non si tratta di un metodo di progettazione completo ma piuttosto di una serie di punti base utili allo studio razionale di un sistema continuo di diffusione. E' stato presentato ed applicato dalla Rockwell per la progettazione dei sistemi diffusivi delle pompe principali dell'idrogeno per i razzi e per lo space-shuttle. Il diffusore per il quale è stata predisposta la procedura è costituito da due tratti rettilinei collegati da un gomito a sezione costante; è quindi diverso dai diffusori delle macchine reversibili. Tuttavia si riporta ugualmente questo procedimento perchè presenta alcuni aspetti interessanti.

Condizioni d'ingresso

L'esperienza della Rockwell consiglia un ampio traferro tra rotore e statore allo scopo di smorzare le irregolarità provocate dal moto impulsivo della girante sul fluido; il traferro è crescente con l'angolo secondo una legge dedotta sperimentalmente. Per un angolo della corrente di 15° risulta per esempio:

$$\frac{traf}{D_2} = 0.12$$

Ostruzione causata dallo strato limite

L'ostruzione dello strato limite nella gola del diffusore riduce l'area del passaggio e la successiva decelerazione: secondo gli autori il suo calcolo è molto importante.

Gli autori utilizzano il metodo di Truckenbrodt per il calcolo dello sviluppo dello strato limite, a partire dalla girante fino alla gola del diffusore lungo tutte le pareti lambite dal fluido. Il calcolo è iterativo in quanto strato limite e nucleo centrale a potenziale si influenzano reciprocamente.

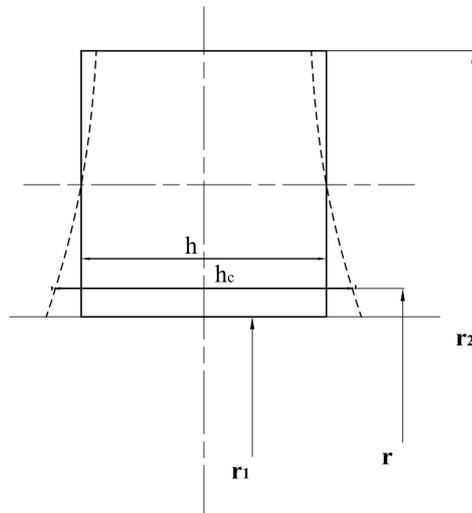


Figura 1.5: Correzione della sezione secondo Meng e Jackson

Gomito

Secondo l'esperienza Rockwell è preferibile che il gomito presenti sezione costante. La sezione di flusso può essere corretta allo scopo di ridurre le perdite e limitare l'insorgere dei moti secondari. Meng e Jackson fanno riferimento ad una procedura di correzione dedotta sperimentalmente alla NASA, secondo la quale una sezione qualsiasi può essere modificata nella sua larghezza trasversale in funzione del raggio locale di curvatura:

$$h_c(r) = h(r) * C(r)$$

dove

$$C(r) = \left[\frac{r_1 + r_2}{2r} \right]^{1.5}$$

Come si può vedere facilmente dall'ultima relazione il fattore correttivo tende ad allargare le zone a raggio più stretto ed a restringere quelle a raggio ampio.

Nei gomiti a sezione circolare questa correzione ha provocato una diminuzione delle perdite di circa il 50%.

Modello di previsione dello stallo all'ingresso

La schiera delle pale in ingresso viene approssimata da una schiera rettilinea; quando l'angolo del fluido è minore di quello di inclinazione della pala il comportamento della schiera è riconducibile a quello di un semidiffusore a pareti rettilinee. Sono allora utilizzabili i dati sperimentali noti per i diffusori ad asse rettilineo, al fine di stabilire l'angolo di incidenza della corrente per cui avviene lo stallo all'ingresso ed il valore della relativa portata; è possibile anche risalire alle perdite di energia dovute allo stallo. Gli autori dimostrano che il metodo di previsione, nonostante la sua approssimazione, fornisce risultati sufficientemente precisi.

1.2.4 Ellis

Ellis in [22] propone un metodo monodimensionale per i canali di ritorno dei compressori che può essere facilmente esteso alle pompe multistadio.

Il criterio è basato su due punti fondamentali: l'orientazione palare e la minima deviazione angolare del fluido.

Da uno studio sulle perdite nei gomiti a sezione rettangolare è risultato che le perdite si minimizzano ponendo due delle superfici del canale ortogonalmente al gradiente di pressione: si instaurerebbero, in tal modo, moti secondari di minor intensità. Per costringere il fluido alla minima deviazione angolare, la rimozione della C_U deve essere eseguita nel gomito a 180° . Il metodo si prefigge di determinare dapprima le coordinate cilindriche della linea media della pala, poi la superficie palare media ed infine lo spessore della pala.

Linea media palare

Si suppone che il moto sia monodimensionale e che i parametri del moto, con le loro derivate, siano adeguatamente rappresentati dal valore valutato lungo la linea meridiana media. Quest'ultima è l'intersezione della superficie media della pala con la superficie di rivoluzione media tra i due dischi.

Si stabilisce una sezione meridiana adottando ampi raggi di curvatura compatibilmente con gli ingombri: restano così fissate già due delle tre coordinate cercate. Si adotta, quindi, una legge di variazione del momento della quantità di moto rC_U ed una legge di decelerazione, lungo la linea meridiana:

$$\lambda = rC_U = \lambda(m)$$

$$c = c(r)$$

dove con m si intende l'ascissa lungo la linea meridiana media a partire dall'ingresso del gomito.

L'autore non suggerisce alcun criterio per la scelta delle due leggi, rimandando ciò ai futuri sviluppi sperimentali e adotta momentaneamente variazioni lineari.

Resta così stabilito l'angolo della linea di corrente media che, per le ipotesi fatte, coincide con la linea palare media:

$$\beta = \arcsin \frac{\lambda(m)}{r(m)c(m)}$$

L'anomalia circonferenziale della linea media della pala è quindi allora:

$$\theta = \int_0^m \operatorname{tg} \beta \frac{dm}{r}$$

Superficie palare media

Imponendo che la superficie palare media venga orientata ortogonalmente al gradiente di pressione, l'angolo δ , che rappresenta l'inclinazione locale della pala rispetto al piano meridiano, è espresso dalla seguente relazione:

$$\operatorname{tg} \delta = \frac{\left(\frac{\lambda}{r}\right)^2 \cos \gamma + \frac{r}{r_c} c_M^2}{c_M \frac{\partial \lambda}{\partial m}}$$

in cui γ è l'angolo di inclinazione della linea di corrente nel piano meridiano ed r_c è il locale raggio di curvatura della linea meridiana media. La superficie palare media è ora descritta dal movimento, lungo la linea media, di un segmento di retta la cui proiezione meridiana è, in ogni punto, ortogonale alla linea meridiana media. Esso sarà anche inclinato, rispetto al piano meridiano, dell'angolo δ calcolato con l'equazione precedente.

Spessore palare

Lo spessore palare è implicitamente definito dalle assunzioni fatte. Calcolato l'angolo palare costruttivo e assegnata la velocità di attraversamento:

$$\frac{Q}{z} = c_M(m) \cdot 2\pi \cdot r(m) \cdot h(m) \cdot (1 - \xi(m))$$

dove ξ è l'ostruzione palare. Quindi:

$$\xi(m) = 1 - \frac{Q/z}{c(m) \cdot \cos \beta(m) \cdot 2\pi \cdot r(m) h(m)}$$

Lo spessore tangenziale t risulta quindi:

$$t(m) = \xi(m) \cdot r(m) \cdot \frac{2\pi}{z}$$

L'autore raccomanda di verificare alla fine che lo sviluppo della pala sia dolce, privo di discontinuità, e che gli angoli all'ingresso e all'uscita abbiano i valori desiderati.

La procedura di progettazione è stata verificata sperimentalmente da Ellis, con esiti apparentemente buoni; è doveroso precisare, comunque, che quelle misure non possono essere considerate completamente affidabili poichè nell'apparato di prova vi erano forti fenomeni di scia provocati dalla palettatura induttrice stazionaria.

La caratteristica delle geometrie calcolate con questa procedura è la forma schiacciata, con due angoli fortemente acuti, delle sezioni di flusso appartenenti alla zona di gomito. Infatti il

gradiente di pressione causato dalla curvatura meridiana è, in quella zona, maggiore di quello tangenziale. I raggi idraulici di quelle sezioni risulteranno, quindi, generalmente bassi.

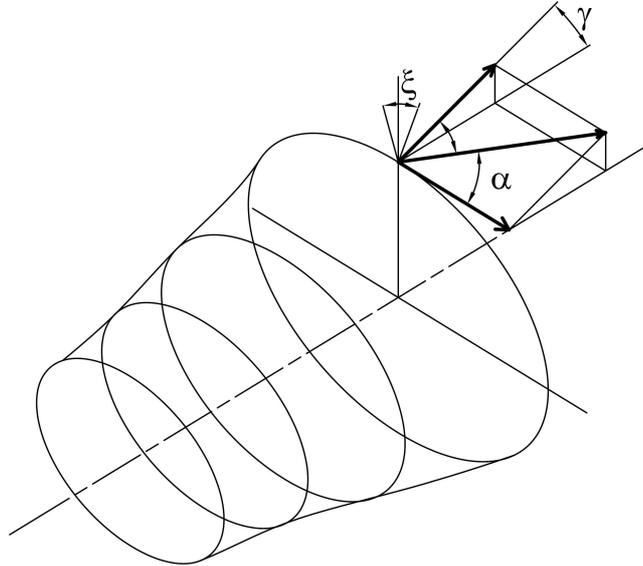


Figura 1.6: Formulazione degli angoli in Ellis.

1.3 Criteri di progettazione-2

Successivamente all'introduzione delle tecniche di calcolo per l'analisi fluido dinamica nella progettazione delle turbomacchine sono stati proposti altri metodi di progettazione delle turbomacchine nella loro complesso e quindi anche dei diffusori palettati.

E' bene porre l'accento sul fatto che è l'introduzione stessa dei codici di calcolo che ha consentito enormi sviluppi alla progettazione delle turbomacchine, sia che se ne faccia un uso diretto, cioè per analizzare il comportamento fluidodinamico di un componente con una data geometria, sia che se ne faccia un uso inverso, cioè per ricavare la geometria di un determinato condotto che garantisce alla corrente un determinato campo di moto.

Con i metodi tradizionali presentati in precedenza la progettazione avveniva per tentativi successivi. Una prima geometria veniva prototipata e si modificava la geometria cercando di capire come correggere tale geometria a partire dai risultati del collaudo del prototipo. Un nuovo prototipo veniva realizzato a partire dalla nuova geometria e veniva effettuato il collaudo. Di fatto si giungeva alla geometria definitiva attraverso tentativi successivi, ciascuno dei quali richiedeva un prototipo. Questo metodo è molto dispendioso perchè le sole informazioni fornite dal collaudo della macchina non garantiscono una completa comprensione dei fenomeni che caratterizzano il campo di moto all'interno di una geometria, per cui spesso era necessaria la realizzazione di numerosi prototipi, prima di giungere ad un risultato definitivo.

L'introduzione dei codici di calcolo CFD ha reso possibile uno studio accurato della fluidodinamica interna delle turbomacchine. Tali codici possono essere utilizzati per indagare il funzionamento di una turbomacchina e per prevederne le prestazioni. Possono quindi essere

utilizzati per realizzare un collaudo virtuale di una geometria proposta in sede di progettazione e per individuare la causa di un eventuale mal funzionamento, andando così ad integrare le tecniche tradizionali di progettazione.

L'utilizzo dei codici di calcolo non sostituisce comunque lo studio della fluidodinamica interna, ma la integra. L'analisi e l'interpretazione dei risultati numerici richiede conoscenze fluidodinamiche approfondite. Allo stesso modo la modifica della geometria per ottenere i miglioramenti desiderati del campo di moto richiede di stabilire una relazione di causa effetto tra la geometria dei condotti e la fluidodinamica interna che può essere realizzata solo da persone esperte.

1.3.1 Tecniche di progettazione inversa

Parallelamente all'utilizzo diretto dei codici di calcolo sono state proposte nuove tecniche di progettazione.

Molte di queste possono essere raggruppate nella famiglia delle tecniche inverse di progettazione; queste raggruppano tutte quelle tecniche che prevedono di definire la geometria della pala a partire dallo scambio di energia e dal campo di moto desiderato della girante. Numerosi sono i metodi inversi che sono stati presentati. Quello che distingue tra di loro i metodi che sono stati proposti è essenzialmente la formulazione utilizzata delle equazioni del moto e la loro integrazione.

In [11] Zangeneh riassume e integra vari metodi inversi proposti fino a quel momento in un'unica formulazione. Viene proposto un metodo inverso tridimensionale per fluidi comprimibili e per macchine a sviluppo sia radiale che assiale. Tale metodo prevede l'assegnazione di una distribuzione della componente rC_U circonferenziale media della velocità nel condotto interpolare sul piano meridiano della girante. Integrando la distribuzione assegnata si ricava la forma tridimensionale della pala che realizza il campo di moto assegnato con la forma:

$$r\bar{C}_U = \frac{Z}{2\pi} \int_0^{\frac{Z}{2\pi}} rV_U d\theta$$

Per ricavare il profilo tridimensionale nel suo lavoro Zangeneh propone l'integrazione delle equazioni della corrente aviscosa all'interno della pala.

In un lavoro successivo dell'autore [17] viene presentata una parametrizzazione dell'andamento assegnato della componente $\partial rC_U / \partial m$, cioè della derivata lungo l'ascissa curvilinea della linea di corrente della componente rC_U circonferenziale media. Tale andamento è prescritto generalmente con un andamento composto da un primo tratto parabolico, un tratto rettilineo e un ultimo tratto parabolico collegati con la continuità della derivata prima. In questo modo con appena tre parametri è possibile definire la distribuzione su una sezione di riferimento. La scelta di parametrizzare la distribuzione di questa grandezza è motivata dal fatto che questa derivata ha una proporzionalità diretta con il carico palare sulla sezione della pala, cioè sulla differenza di pressione tra la faccia aspirante e quella premente secondo la seguente formula:

$$p^+ - p^- = \frac{2\pi}{Z} W_m \frac{\partial r C_u}{\partial m}.$$

In questo modo si può intervenire direttamente sull'andamento della grandezza più importante per controllare lo sviluppo dello strato limite e il campo di moto nel condotto interpalare.

Fa riferimento la figura 1.7 in cui si vede l'andamento che è stato descritto applicato per la sezione di mozzo e quella di corona, tra il bordo di ingresso (LE) e il bordo di uscita (TE). L'andamento di una sezione generica viene ricavato per interpolazione lineare di quelli prescritti su queste sezioni.

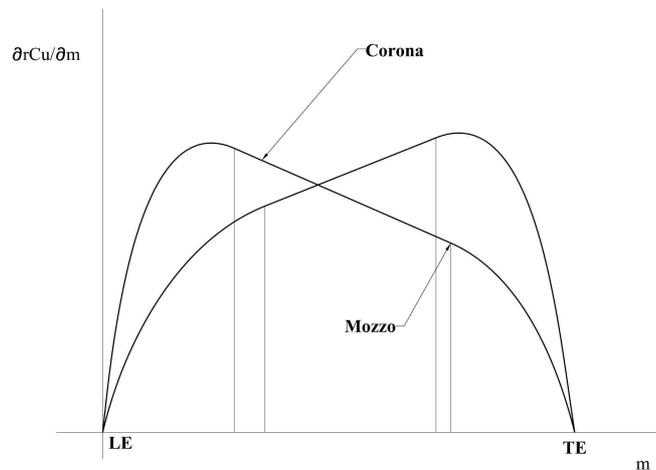


Figura 1.7: Assegnazione per l'andamento del momento angolare della quantità di moto proposta in [17].

Sempre nel suo lavoro [17] l'autore mette in luce come tale tecnica di progettazione può essere utilizzata per controllare l'andamento dei moti secondari all'interno della corrente proprio controllando l'andamento del gradiente di pressione lungo le facce delle pale. Tuttavia vista la natura aviscosa della metodologia inversa questa può essere applicata solo in fase di progettazione per modificare il gradiente palare delle pressioni. Si è reso comunque necessaria una verifica con un codice di calcolo che risolva le equazioni di Navier-Stokes per visualizzare le perdite.

Infine in [16] si mostra un caso di progettazione di un diffusore a boccia di una pompa multistadio.

Un'altra tecnica inversa è quella presentata da Demeulenaere et al. in [20]. Il metodo proposto arriva a definire una geometria della pala a partire da una prefissata distribuzione del numero di Mach lungo la pala, e quindi delle velocità. Per fare ciò gli autori hanno proposto una modifica ad un solutore euleriano e sono arrivati alla soluzione tramite un processo iterativo di piccole modifiche alla geometria della pala. Anche in questo caso però si è fatto uso di un solutore euleriano che non conteggia lo sviluppo dello strato limite. In seguito una modifica a questo metodo è stata proposta da de Vito et al. in [21]. In questo

lavoro ad un codice di calcolo inverso di natura euleriana viene affiancato un codice viscoso per l'analisi della geometria che ad ogni iterazione si ricava.

Il metodo proposto in [20] è stato utilizzato da Veress e Van den Breambussche in [19] per la riprogettazione della palettatura di un canale interstadio di un compressore multistadio. Gli autori in questo lavoro partono dai dati sperimentali presentati in [56] e prendono come geometria di riferimento un canale interstadio palettato solo nel canale di ritorno. Viene poi proposto una palettatura lunga continua che occupa tutto il condotto meridiano a partire dalla zona a flusso radiale subito a valle della girante lungo tutto il gomito e il canale di ritorno. Tale palettatura è stata progettata imponendo un'andamento del carico palare nella sezione media della pala e ricavando l'andamento degli angoli palari su tale sezione. Tale andamento è stato poi riportato su mozzo e corona per ricavare i profili delle due sezioni, applicando delle correzioni per rendere uniforme l'angolo di avvolgimento alle due sezioni. La geometria proposta ha mostrato decisi miglioramenti rispetto alle prestazioni della geometria di riferimento in termini di recupero di pressione statica, di perdita di pressione totale e di capacità di deviazione della schiera. Le analisi numeriche evidenziano come i ricircoli che si notavano nella geometria di riferimento sono quasi del tutto riassorbiti. Oltre a questa nuova geometria gli autori presentano una palettatura ricavata con il metodo inverso proposto in [20] da Demeulenaere. A fronte della complicazione matematica di questo nuovo progetto e della geometria non regolare ottenuta questa tecnica ha prodotto un aumento sensibile della deviazione indotta dalla schiera, ma solo un lieve aumento del recupero di pressione statica.

A giudizio di chi scrive il punto di forza maggiore dei metodi inversi tridimensionali sta nell'aver legato tra di loro la forma della pala e il campo di moto atteso tramite un'efficace parametrizzazione. Tuttavia si distinguono per la complessità matematica che vi sta dietro, e, pur mostrandosi capaci di controllare efficacemente il carico palare e i moti secondari, propongono delle geometrie molto complicate e non assicurano di individuare la geometria con le prestazioni migliori. Inoltre, come fa notare Van den Breambussche in [50], la complessità dei moti secondari è tale che per risolverli è necessario fare uso di codici di calcolo 3D e qualunque inaccuratezza nella determinazione del campo di moto potrebbe portare a grossolani errori; i metodi di progettazione inversa risolvono le equazioni aviscose del moto, o, al più, tengono conto dello strato limite tramite funzioni di parete e quindi potrebbero portare spesso ad errori.

1.3.2 Tecniche di ottimizzazione

A partire dalla seconda metà degli anni novanta si è cominciato ad applicare le tecniche di ottimizzazione a quelli che venivano chiamati computer experiments. Queste metodologie possono essere applicate a qualunque campo che faccia uso di analisi numeriche; i due grandi campi di applicazione sono quello strutturale e quello termofluidodinamico.

I metodi di progettazione inversa trovano la geometria che induce il campo di moto aviscoso prescritto in fase di progettazione. Tale geometria può di fatto essere considerata un ottimo se si considera il campo di moto (o di pressione) desiderato come l'ottimo da raggiungere. Tuttavia questo non può essere considerata la geometria ottima, in quanto sono svariati i

parametri che coinvolgono il funzionamento di una turbomacchina, non solo la regolarità del campo di moto che pure è un indice di buon funzionamento. E' nata così l'esigenza di identificare la geometria che di volta in volta rappresenta l'ottimo secondo i parametri che si ritengono di volta in volta più importanti in fase di progettazione. Le cosiddette tecniche di ottimizzazione per i computer experiments si propongono proprio come obiettivo quello di individuare la geometria che meglio risponde a dei parametri richiesti minimizzando delle funzioni di costo appositamente scelte.

Negli ultimi vent'anni la letteratura prodotta nel campo dell'ottimizzazione è stata particolarmente ricca. Solo una parte riguarda l'applicazione di tali tecniche alla progettazione delle turbomacchine. L'applicazione di tali tecniche alle turbomacchine rientra in quel ramo che prende il nome generico di *shape optimization*; sotto questo nome vanno tutte le tecniche che individuano come risultato una geometria le cui caratteristiche, strutturali o funzionali, rappresentano un ottimo rispetto ad alcuni parametri di interesse. La geometria viene ottimizzata grazie alle simulazioni al computer dalle quali si possono ricavare i valori delle grandezze che si vogliono ottimizzare.

Le tecniche più utilizzate sono le *algoritmi genetici* (GA), la *simulated annealing* (SA) e la tecnica degli sciame di particelle o *particle swarm optimization* (PSO). Tali algoritmi possono essere utilizzati nel contesto delle simulazioni al computer in quanto effettuano la ricerca dell'ottimo senza dover calcolare il gradiente della funzione di costo e possono quindi operare su uno spazio di indagine discreto come quello fornito dalle simulazioni al computer.

Per realizzare un'ottimizzazione è necessario portare a termine numerose simulazioni con un grande dispendio di risorse di calcolo. Per limitare lo sforzo computazionale necessario sono stati proposti approcci che fanno uso di modelli matematici di regressione per estendere a tutto lo spazio di indagine i risultati di poche simulazioni scelte con cura. In questo caso si parla di ottimizzazione assistita da metamodelli (*metamodel assisted optimization, surrogate based optimization*). La scelta delle simulazioni da realizzare viene effettuata grazie alle tecniche statistiche di *emphdesign of experiment* (DOE). In [12] gli autori propongono una dettagliata panoramica di quasi tutte le tecniche che attualmente rappresentano lo stato dell'arte per quello che riguarda le simulazioni al computer.

In [18] Pierret et al. presentano un metodo di progettazione che parte da una parametrizzazione della geometria e ricava la geometria desiderata minimizzando una funzione obiettivo del tipo:

$$OF = P_{Meca} + P_{RP} + P_{\xi} + P_{Mach}.$$

In questa funzione obiettivo il primo termine P_{Meca} è il termine che rappresenta le costanti di resistenza meccanica, perchè si vuole presentare una geometria che sia compatibile con le esigenze produttive. Il secondo termine P_{RP} rappresenta i target di prestazione richiesti; gli ultimi due termini della funzione obiettivo P_{ξ} e P_{Mach} servono rispettivamente a quantificare le perdite di energia nella geometria e il rispetto dell'andamento del numero di Mach lungo le facce della pala. Di fatto quest'ultimo termine consente di assegnare il campo di moto e di pressione a ridosso delle facce delle pale. La ricerca dell'ottimo di questa funzione avviene

tramite un utilizzo combinato di simulazioni CFD, modelli matematici di regressione (nella fattispecie reti neurali) e simulated annealing.

La tecnica proposta dagli autori opera nel modo seguente: un modello di regressione viene addestrato su un database di simulazioni CFD consentendo di ricavare la relazione tra i parametri che definiscono la geometria e le prestazioni della geometria. Su questo metodo di regressione si utilizza una tecnica di ottimizzazione per trovare un ottimo. A partire dai parametri individuati come l'ottimo del modello si effettua una simulazione CFD di verifica su quei dati. Se i risultati della simulazione coincidono con quelli del modello, a meno di uno scarto accettabile, si ritiene conclusa la geometria in questione e individuato l'ottimo; in caso contrario si ripete la procedura arricchendo il database con la simulazione effettuata.

La funzione obiettivo costruita come indicato sopra consente, tramite la tecnica proposta sopra, di rintracciare la geometria che meglio sposa i parametri conteggiati di ciascun termine. Una funzione obiettivo così strutturata consente inoltre di soppesare in modo diverso ciascun termine in modo da privilegiare alcune esigenze a discapito di altre (per esempio prestazioni e ingombri piuttosto che efficienza). Inoltre è possibile utilizzare questo metodo come se fosse un metodo inverso. Infatti, come visto, è possibile specificare come termine della funzione obiettivo l'andamento per punti del numero di mach lungo i lati della pala. In questo modo si individua la forma della pala che ricava un determinato campo di moto e quindi si ottiene un metodo inverso. Rispetto ai metodi inversi presentati nel paragrafo 1.3.1 una pala progettata con un metodo inverso che fa uso di queste tecniche di ottimizzazione sono più precise perchè risolvono il campo di moto con i programmi CFD che risolvono le equazioni di Navier-Stokes invece che con solutori aviscosi.

Tuttavia questi metodi sono molto dispendiosi dal punto di vista computazionale; inoltre, come si fa notare in [18] 'The development of the objective function is a delicate and long procedure. It may contain a lot terms quantifying such design objectives as manufacturing constraints and cost' (La predisposizione della funzione obiettivo è una procedura lunga e delicata, perchè potrebbe contenere molti termini per quantificare tutti gli obiettivi di progetto, vincoli e oneri costruttivi). Anche l'implementazione dei modelli di regressione e degli algoritmi di ottimizzazione rappresentano punti critici per la complessità insita in tali tecniche e nel loro utilizzo con i programmi di simulazione.

1.4 Esigenze industriali

I metodi fin qui presentati rappresentano lo stato dell'arte nella letteratura tecnico-scientifica. Gli articoli scientifici non possono essere letti con lo stesso piglio con cui si legge la letteratura manualistica. Le esigenze a cui rispondono i casi sperimentali trattati nella letteratura scientifica non sono le stesse alle quali si risponde all'interno di un ufficio tecnico quando si tratta di progettare un nuovo pezzo.

Fino agli anni ottanta gli articoli scientifici fornivano delle interpretazioni fisiche ai fenomeni fluidodinamici interni alle turbomacchine e verificavano la teoria con indagini sperimentali. In sede di progettazione da tali articoli si ricavava la comprensione e la genesi dei fenomeni

interni alle turbomacchine, consentendo così di individuare la strategia di intervento migliori per consentire il controllo dei fenomeni in questione. Si capisce come gli articoli scientifici non devono essere letti come dei manuali veri e propri di progettazione, ma bensì come degli strumenti da applicare nel modo più consono alla progettazione.

Allo stesso modo gli articoli che presentano le procedure più moderne di progettazione che fanno uso delle tecniche CFD e degli algoritmi di ricerca dell'ottimo devono essere letti non con l'ottica di trovare ciò che risolve un determinato problema, ma con il piglio di chi vuole un metodo per risolvere i problemi. E' fondamentale quindi applicare tali tecniche tenendo conto delle esigenze dell'industria, non cercare le tecniche che tengono conto delle esigenze dell'industria.

Con il proposito di proporre una procedura di progetto che consente di applicare nel migliore dei modi le tecniche di calcolo numerico e gli algoritmi di ricerca dell'ottimo in campo industriale si è reso necessario individuare quelle che sono le istanze principali della progettazione delle turbomacchine all'interno di un'azienda. Tali istanze sono state individuate anche grazie al contributo del dialogo tra l'azienda finanziatrice della borsa di dottorato e l'università, soprattutto nella collaborazione che ha portato alla realizzazione del prototipo che sarà presentato nei prossimi capitoli.

1.4.1 Interazione con i codici di calcolo fluidodinamici

La prima e più importante istanza che si può individuare per la progettazione nasce dall'osservazione che al giorno d'oggi non si può prescindere dall'utilizzo delle tecniche numeriche di calcolo. Da quando sono state introdotte hanno portato grandi progressi nella progettazione fluidodinamica e ad oggi sono in commercio versioni commerciali dei codici CFD che rendono disponibile alle aziende quelle tecniche che fino a qualche anno fa erano proprietà dei soli enti di ricerca.

La progettazione di un componente di una macchina, utilizzando dei codici di calcolo numerico, solitamente segue lo schema che si va ad illustrare. A partire da una soluzione di primo tentativo (Tentativo '0') (che può essere un componente già esistente che vuol essere migliorato, oppure un componente progettato con le tecniche standard di progettazione) si indaga il comportamento di quest'ultimo grazie ad una simulazione al computer. Dai risultati di questa simulazione un tecnico esperto è in grado di individuare dove sono i difetti della geometria attualmente in analisi e dove e in quale modo è possibile intervenire per apportare miglioramenti.

La fase di analisi della geometria di primo tentativo consente di individuare una nuova geometria che si suppone migliorativa. In taluni casi tale geometria deve essere studiata o controllata al fine di verificarne la realizzabilità grazie alle tecniche produttive che si desiderano adottare. Dopo la fase di verifica di ingegnerizzazione, tale geometria può essere simulata anch'essa con i codici di calcolo. Tale simulazione può essere confrontata con la simulazione precedente (Tentativo '0') che funge così da benchmark¹. Dalla simulazione il progettista può

¹Per confrontare i risultati di due, o più, simulazioni è necessario che i parametri di tali simulazioni siano stati imposti coerentemente affinché i risultati di tali simulazioni siano tra di loro confrontabili

quindi ricavare un controllo se si stanno operando modifiche che possono portare al risultato corretto. Qualora sia necessario è possibile procedere generando altre geometrie e raggiungere, grazie al feedback delle simulazioni effettuate sui tentativi precedenti.

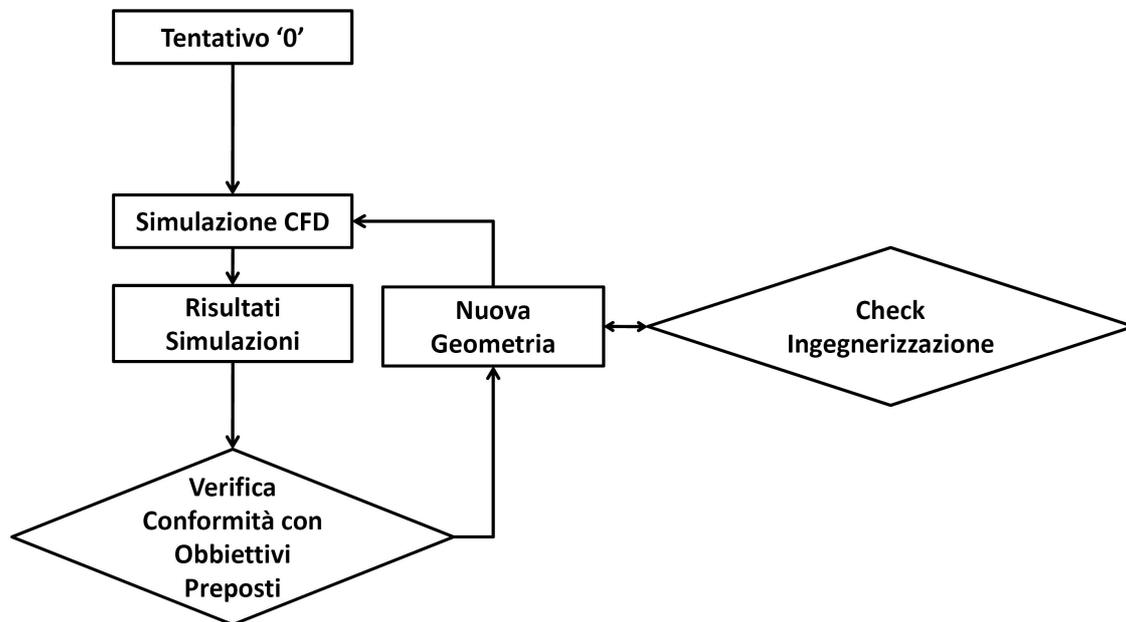


Figura 1.8: Schema dell'utilizzo delle tecniche CFD nella progettazione.

Lo schema proposto in figura 1.8 mostra l'utilizzo delle simulazioni CFD in fase di progettazione. Nel processo sono chiare le interazioni con i codici di calcolo CFD; inoltre la verifica di ingegnerizzazione viene realizzata con software appositi. Un criterio di progettazione per sposarsi con le esigenze di un'azienda che fa uso di questi moderni software deve interfacciarsi agevolmente con questi programmi e consentire di eliminare, il più possibile, i tempi morti nelle varie fasi dello schema senza tuttavia perdere la capacità di adattarsi a quelle che sono le richieste di genericità di una progettazione.

1.4.2 Raccolta di informazioni

Lo schema presentato in figura 1.8 mostra le varie fasi della progettazione industriale che fa uso dei codici di calcolo; si evidenzia altresì un'altra istanza fondamentale per un'azienda. All'interno del percorso di progettazione il passo che porta all'individuazione di nuove geometrie è preceduto da un'analisi dei risultati raccolti per tutti i tentativi precedenti; in altre parole dall'analisi delle informazioni ricavate per i tentativi precedenti e quindi dall'esperienza acquisita in precedenza.

Una delle chiavi per una progettazione razionale consiste nell'opportunità di raccogliere informazioni per migliorare il progetto in corso e da utilizzare in altre sedi. Questa esigenza non è stata formalizzata negli ultimi anni, ma è cosa nota nel campo delle turbomacchine. I numeri tipici di macchina, di flusso e di pressione sono frutto di una riorganizzazione dei

dati storici sul funzionamento e sulle prestazioni delle macchine e quindi rappresentano una rielaborazione delle informazioni raccolte per la progettazione di altre macchine.

Da quanto appena detto risulta che un criterio di progettazione non può considerarsi efficace per un'azienda se non prevede al suo interno una raccolta razionale di informazioni. Tali informazioni devono essere raccolte con dei metodi razionali, in modo da poter essere estese ad altri casi progettuali.

Oltre a ciò queste informazioni devono rispondere a delle richieste di universalità, fatto spesso trascurato in un'azienda dove molto spesso capita che le informazioni che derivano dalla progettazione di un componente siano a disposizione di un singolo individuo, quasi sempre il progettista che ha seguito lo sviluppo del componente. Portando all'estremo questo discorso si nota che in un'azienda spesso una singola persona detiene, come conoscenze personali, buona parte del patrimonio tecnico di un'azienda. In questo modo l'azienda è ricattabile dal progettista nel senso che lui solo ha l'accesso alle conoscenze tecniche dell'azienda e si troverebbe in difficoltà qualora lui venisse a mancare. Questa situazione è frutto di due diverse cause: prima di tutto una mancanza di pianificazione da parte dell'azienda che imponga un programma di archiviazione delle informazioni. In secondo luogo spesso un'azienda non ha a disposizione gli strumenti fondamentali per archiviare informazioni che siano poi accessibili a tutti. Manca in sostanza un linguaggio codificato e universale sotto il quale trascrivere le informazioni.

Quindi affinché sia efficace un criterio di progettazione deve prevedere anche la possibilità di immagazzinare informazioni secondo uno schema razionale (di conseguenza universale) che metta queste informazioni a disposizione di più persone. In questo modo le informazioni sono a disposizione di più individui e diventano patrimonio di tutta l'azienda.

1.4.3 Progettazione di serie di pompe

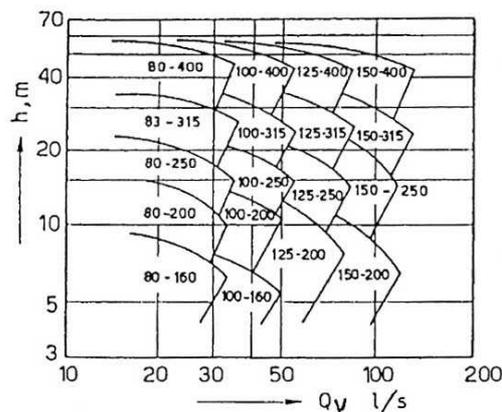


Figura 1.9: Esempio di suddivisione, in termini di portata e prevalenza, del campo di funzionamento, dei vari modelli di una serie di pompe.

Un'ulteriore istanza può essere individuata dalla pratica industriale. Ciascuna azienda presenta le sue pompe a catalogo. Questo significa che ciascuna tipologia di pompa è presen-

tata in una serie. All'interno di ciascuna serie vari modelli si dividono, in termini di portata e prevalenza, il campo complessivo di funzionamento delle varie pompe. Le pompe di ciascuna serie sono accumulate dalla stessa metodologia produttiva e vengono dimensionate seguendo delle procedure di standardizzazione, in modo da pianificare la progettazione di più pompe simili.

Risulta opportuno pianificare delle campagne metodiche di raccolta di dati in sede di progettazione di una pompa che possano essere utilizzate anche in sede di progettazione di pompe della stessa serie. Tale forma di pianificazione può essere intesa come una vera e propria esigenza per una procedura razionale di progettazione di un'azienda dove si deve produrre a catalogo.

Capitolo 2

Progettazione parametrica

A giudizio di chi scrive un approccio parametrico alla progettazione è quello che meglio soddisfa le esigenze industriali analizzate nel paragrafo 1.4.

Una progettazione parametrica prevede la descrizione della geometria con funzioni matematiche di vario tipo che consentono di individuare superfici nello spazio. Ciascuna funzione è specificata da alcuni parametri. Variando i valori dei parametri si ottengono delle geometrie diverse. Esiste quindi una relazione univoca diretta tra i valori dei parametri e la forma geometrica nello spazio.

Con 'parametrizzazione di una geometria' solitamente si intende un processo sistematico con il quale si schematizza una forma geometrica, più o meno complessa, con una serie di entità geometriche meno complesse. Ciascuna entità è definita da una funzione specificata da uno o più parametri. Una geometria è individuata dai valori che assumono i parametri che la specificano; variando uno o più valori di questo insieme si ottiene modifica la geometria.

Il disegno tradizionale della forma geometrica di un pezzo è non parametrico. Nel caso della definizione di una singola geometria la definizione non parametrica è sicuramente più veloce; tuttavia nel caso sia necessario definire più geometrie analoghe o cambiare la forma geometrica sulla quale si sta lavorando, una definizione parametrica è molto più efficiente. Oltre a garantire un controllo preciso delle modifiche che si possono apportare a questi pezzi, viene garantita una modifica delle geometrie molto più rapida. La corrispondenza tra i valori dei parametri e le geometrie consente di individuare con precisione ciascuna ciascuna forma.

Con 'progettazione parametrica', o 'approccio parametrico alla progettazione', si intende quando si definisce parametricamente la geometria di un pezzo che si vuole progettare in modo da consentire rapide modifiche della forma per migliorare determinati aspetti o particolari.

Per quanto riguarda il campo macchinistico è possibile riassumere in alcuni punti quelli che sono i principali vantaggi che si possono trarre dalla progettazione parametrica:

- le forme complesse che assumono le turbomacchine ben si sposano con forme definite da funzioni matematiche;
- la definizione matematica delle superfici consente di esportare facilmente la forma di tali sezioni di questa con metodi non parametrici (ad es. come liste di punti nelle cartesiane) che consentono di ricostruire agevolmente all'interno di ambienti

non parametrici (come i programmi di analisi fluidodinamica o strutturale) le geometrie che si vanno a progettare;

- il controllo sulle modifiche che si possono apportare alla geometria è totale;
- grazie ai codici di calcolo CFD è possibile indagare il comportamento fluidodinamico di un progetto e quindi legare le prestazioni alla geometria sotto analisi. Con una definizione parametrica della geometria è possibile quindi legare direttamente i valori dei parametri che definiscono la geometria alle prestazioni e quindi stabilire una relazione precisa causa-effetto tra variazione della geometria e prestazioni;
- la definizione parametrica della geometria risulta efficace anche in rapporto ad esigenze di standardizzazione; è possibile affrontare la progettazione di più elementi simili utilizzando un'appropriata parametrizzazione.

Parametrizzare una geometria è un'operazione che sta all'inizio di una progettazione; la principale criticità di questa fase è insita proprio nel fatto che è preliminare a tante altre. La scelta della parametrizzazione da utilizzare deve tener conto di tutte le fasi successive, e un parametrizzazione errata potrebbe avere ripercussioni su tutte le fasi successive.

2.1 Strumenti matematici per la parametrizzazione-Curve di Bezier

La matematica mette a disposizione molti strumenti, sotto forma di funzioni o operatori, da utilizzare per la parametrizzazione.

Qualunque funzione, forma geometrica in due o tre dimensioni o operatore matematico può concorrere alla definizione parametrica di una geometria. Per esempio un arco di parabola può rappresentare la sezione di una superficie. Al variare dei coefficienti dell'espressione analitica dell'arco di parabola varia la forma sul piano della parabola e quindi la forma della sezione rappresentata dall'arco di quella parabola. Tali coefficienti possono essere fatti variare liberamente per definire nella massima generalità la forma della parabola che vogliamo, oppure alcuni possono essere vincolati per soddisfare ad esigenze altre esigenze (per esempio di continuità ad un estremo, o di passaggio per un punto).

Le entità analitiche che vanno a comporre la definizione parametrica di una forma geometrica devono essere semplici in rapporto alla complessità della forma da rappresentare. Una parametrizzazione trae giovamento anche dalla flessibilità di queste entità; con questo si intende la possibilità di definire con una stessa parametrizzazione forme geometriche molto diverse tra di loro e la possibilità di aggiungere parametri o di vincolarne altri. Infatti prima di procedere con una parametrizzazione è bene osservare attentamente la forma e come cambia questa forma al variare dei parametri in modo da vincolare quelli che sono superflui.

2.1.1 Curve di Bezier

Le curve e le superfici proposte nella formulazione di Bezier in [23] si sono con il tempo selezionate rispetto alle altre per la formulazione semplice ed elegante e per la disparità di forme che riescono ad assumere.

Tali curve hanno già da tempo mostrato la loro utilità nel campo macchinistico, vista la facilità con la quale consentono di ottenere le forme più disparate mantenendo i vincoli di continuità richiesti alle forme. Il primo ad introdurle con questo scopo fu Casey nel 1983 in [10]. Recentemente sono state utilizzate con successo nelle tecniche di *shape optimization*: Van den Braembussche utilizza le curve di Bezier per parametrizzare il condotto meridiano di giranti di macchine radiali in [50]. In [1] Goel et al. usano le curve di Bezier per parametrizzare la geometria delle pale dei diffusori radiali.

Bezier ha proposto una semplice parametrizzazione della forma di curve a partire dalla posizione di pochi punti, e di superfici a partire dalle curve. Una curva di Bezier è definita dalle coordinate di una successione di punti nello spazio dei quali solo il primo e l'ultimo giacciono sulla curva stessa. I punti estremi su cui passa la curva sono detti *punti di appoggio*, mentre quelli interni della successione sono detti *punti di controllo*; congiungendo con dei segmenti rettilinei i punti in successione della serie si ottiene quello che prende il nome di *poligono* della curva.

L'esempio più semplice è una curva di Bezier di primo grado che può essere scritta come:

$$\mathbf{R} = (1 - u)\mathbf{P}_0 + u\mathbf{P}_1$$

dove \mathbf{R} è il vettore delle coordinate di un punto della curva (x, y, z) e \mathbf{P}_0 e \mathbf{P}_1 sono punti del poligono. u è il parametro che varia tra 0 e 1 lungo la curva. In corrispondenza a $u = 0$ risulta $\mathbf{R} = \mathbf{P}_0$ mentre per $u = 1$ $\mathbf{R} = \mathbf{P}_1$. Facendo variare il parametro u tra 0 e 1 il vettore \mathbf{R} descrive una linea retta come in figura 2.1.

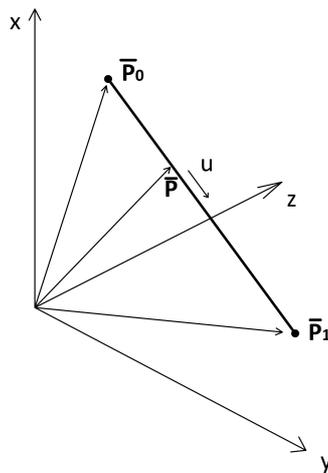


Figura 2.1: Esempio di curva di Bezier di primo grado.

La curva di bezier di secondo grado richiede un poligono a tre punti e può essere definita

da:

$$\mathbf{R} = (1 - u)^2 \mathbf{P}_0 + 2u(1 - u) \mathbf{P}_1 + u^2 \mathbf{P}_2$$

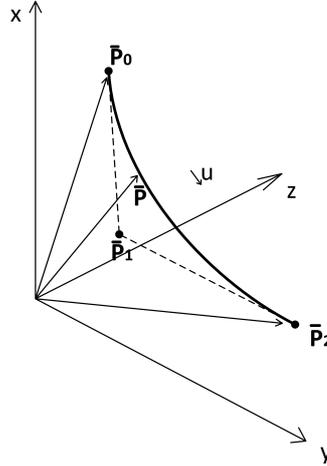


Figura 2.2: Esempio di curva di Bezier di secondo grado.

Come si può vedere dalla figura 2.2 il secondo punto del poligono \mathbf{P}_1 non giace sulla curva ma le tangenti alla curva in corrispondenza ai punti estremi della curva sono in direzione del punto intermedio.

La definizione analitica della curva di Bezier di terzo grado utilizza un poligono a quattro punti:

$$\mathbf{R} = (1 - u)^3 \mathbf{P}_0 + 3u(1 - u)^2 \mathbf{P}_1 + 3u^2(1 - u) \mathbf{P}_2 + u^3 \mathbf{P}_3$$

Di fatto la formulazione analitica delle curve di Bezier è la media ponderata delle coordinate dei punti che compongono il poligono di definizione della curva. Le funzioni ponderali sono i polinomi di Bernstein:

$$B_k^n(u) = \binom{n}{k} u^k (1 - u)^{n-k}$$

e la formulazione analitica di una curva di Bezier generica di grado n è data da:

$$\begin{aligned} \mathbf{R} &= \sum_{k=0}^n \mathbf{P}_k B_k^n(u) = \\ &(1 - u)^n \mathbf{P}_0 + n(1 - u)^{n-1} u \mathbf{P}_1 + \frac{n(n-1)}{2} (1 - u)^{n-2} u^2 \mathbf{P}_2 + \\ &\quad + \cdots + \\ &+ \binom{n}{k} u^k (1 - u)^{n-k} \mathbf{P}_k + \cdots + u^n \mathbf{P}_n \end{aligned}$$

Alcune importanti proprietà delle curve di Bezier sono:

- una curva di grado n è definita da $n + 1$ punti nello spazio;
- la tangente alla curva in corrispondenza al primo punto del poligono è in direzione del secondo e, analogamente, la tangente in corrispondenza all'ultimo punto è in direzione del penultimo punto;
- è possibile incrementare il grado del polinomio per ottenere più gradi di libertà per la definizione della curva qualora ce ne fosse bisogno; inizialmente si può aumentare il grado del polinomio senza alterare la forma della curva, semplicemente aggiungendo un punto di definizione nel poligono che già sta all'interno del poligono.

A partire dalle curve è possibile definire in modo analogo delle parametrizzazioni anche delle superfici. Per dettagli su questo rimandiamo a [10].

Nella progettazione delle turbomacchine le curve di Bezier sono state utilizzate per affrontare i problemi più disparati. Per la loro particolare flessibilità possono essere utilizzate indistintamente per parametrizzare la forma geometrica di una curva o l'andamento monodimensionale di una particolare grandezza.

2.2 Parametrizzazione della palettatura della girante e del canale interstadio

Il lavoro di questa tesi si concentra sulla progettazione delle turbomacchine multistadio. In questo paragrafo si fa innanzi tutto una premessa sui modi in cui sono state scelte queste parametrizzazioni. Poi si esplicano le parametrizzazioni che sono state scelte per designare la forma della palettatura delle giranti, distinguendo tra macchine a basso numero tipico e ad alto numero tipico, e per la palettatura dei diffusori-riconvogliatori.

2.2.1 Diverse scelte per la parametrizzazione

Analizzando la letteratura più recente che riguarda i metodi di progettazione delle turbomacchine che fanno uso di tecniche di parametrizzazione, ci si rende conto che le parametrizzazioni di una forma geometrica possono essere generalmente suddivise in due categorie: le parametrizzazioni delle entità geometriche stesse che compongono la forma in questione e la parametrizzazione di grandezze geometriche o funzionali legate alle forme geometriche.

Nel primo caso una forma geometrica, sia essa una superficie piuttosto che una linea, viene definita grazie ad una formulazione analitica, polinomiale o di altra natura, e la parametrizzazione è quindi direttamente legata alla forma che si sta analizzando. È questo quello che è stato utilizzato in [50] da Van den Braembussche per definire la forma geometrica delle generatrici delle superfici di rivoluzione di mozzo e di corona di una girante centrifuga.

Alla variazione dei valori dei parametri quindi corrisponde una variazione diretta della forma della geometria. Il vantaggio di questa formulazione è il controllo totale sulla geometria, quindi delle variazioni delle forme e dei limiti di variazione e di estensione spaziali delle forme.

Questo tipo di approccio si sposa bene con le tecniche di ottimizzazione. Lasciando la totale libertà di forma si consente ad un algoritmo di ricerca dell'ottimo di andare ad indagare anche geometrie e che il buonsenso non indicherebbe come valide, ma che, al computo delle prestazioni risultano invece efficaci.

Nel secondo caso la variazione delle forme non è controllata direttamente. E' il caso per esempio di quando si parametrizzano gli andamenti di alcune grandezze funzionali e a partire da queste si ricava la forma tridimensionale. In [14] viene parametrizzata per esempio l'andamento degli angoli della pala di un diffusore a boccia lungo la linea di riferimento della generatrice sul piano meridiano del mozzo e analogamente per la corona. A partire da tale andamento si ricava la forma nello spazio integrando l'anomalia angolare. Analogamente per i metodi inversi così come proposti da Goto e Zangeneh in [16] si assegna il campo di moto desiderato con una parametrizzazione dell'andamento della variazione del momento angolare della quantità di moto del fluido nel condotto interpolare. Il vantaggio di questo tipo di parametrizzazione è la possibilità di avere un controllo diretto di alcune grandezze fondamentali del moto del fluido. Lo svantaggio è un non diretto controllo della forma che può assumere la geometria. Spesso in questi casi è necessario generare e visualizzare la geometria stessa per controllare l'adeguatezza di tali geometrie nei confronti di eventuali vincoli geometrici.

Tali metodi di parametrizzazione risultano particolarmente efficaci per stabilire una relazione di causa effetto tra le variazioni della geometria e le variazioni delle prestazioni della geometria che si va ad analizzare proprio perchè i parametri scelti sono spesso parametri direttamente legati al campo di moto.

In questo lavoro per la parametrizzazione si è scelto un approccio che, dove possibile, consentisse di parametrizzare delle grandezze legate al moto del fluido e solo indirettamente legate alla forma geometrica da assegnare alle superfici, avendo però controllato che le penalizzazioni sulla possibilità di intervenire sulla geometria non fossero troppo pesanti.

2.2.2 Parametrizzazione della palettatura della girante-Giranti a basso numero tipico

Le tipologie di girante si differenziano notevolmente in base al numero tipico di macchina. Per giranti a basso numero tipico di macchina ($K < 0.5$) lo sviluppo meridiano della girante è molto allungato in direzione radiale. Al crescere del numero tipico di macchina la forma del condotto meridiano della girante diventa sempre più sviluppata in direzione assiale, fino a diventare del tutto assiale.

Nella conformazione a basso numero tipico di macchina lo sviluppo allungato del condotto meridiano consente di progettare pale a semplice curvatura. Le pale a semplice curvatura sono definite grazie alla forma di una sola sezione di riferimento della pala. Le altre sezioni vengono ricavate da quella di riferimento che viene estrusa in direzione assiale. Ad una vista circonferenziale le sezioni di una pala sono sovrapposte come si può notare in figura 2.3.

Per parametrizzare la forma di una pala a semplice curvatura è sufficiente parametrizzare la forma di una sezione di riferimento alla quale vengono assegnati gli angoli di ingresso e di uscita e viene assegnato l'andamento tra i due.

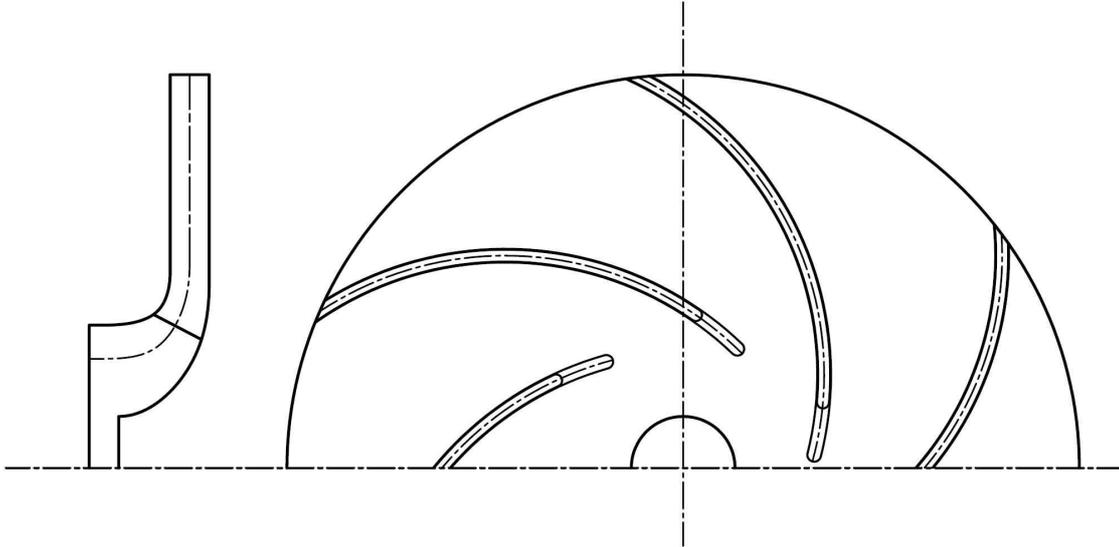


Figura 2.3: Schematizzazione della parte idraulica di una pompa centrifuga a basso numero tipico.

In questa tesi si suggerisce di parametrizzare la forma della linea media della sezione della pala in corrispondenza della superficie del mozzo; la forma tridimensionale della pala e la forma delle linee medie delle altre sezioni della pala si ricavano a partire da quella del mozzo per estrusione in direzione dell'asse della macchina.

Per quello che riguarda la scelta della parametrizzazione della forma della linea media della sezione del mozzo si è scelto di parametrizzare l'andamento della derivata del momento angolare della quantità di moto del fluido su questa sezione; quindi si l'andamento della componente $\partial r C_u \setminus \partial m$, della velocità, dove m è l'ascissa curvilinea della linea media della sezione di riferimento sul piano meridiano come si vede in figura 2.3. L'andamento scelto per tale grandezza è quello proposto in figura 2.5 ed è composto da un primo tratto parabolico, un secondo tratto lineare e un terzo tratto ancora parabolico. A tali tratti viene inoltre imposta la continuità della derivata prima nei punti di giunzione.

In corrispondenza del bordo di ingresso e del bordo di uscita a tale andamento di $\partial r C_u \setminus \partial m$ viene assegnato un valore nullo; in questo modo si assegna un carico palare nullo alla pala in corrispondenza dei bordi di ingresso e di uscita perchè tale grandezza è linearmente legata al carico palare dalla formula:

$$p^+ - p^- = \frac{2\pi}{Z} C_m \frac{\partial r C_u}{\partial m}. \quad (2.1)$$

Imporre un valore nullo del carico palare in corrispondenza del bordo di ingresso ha lo scopo di migliorare il comportamento della palettatura ai carichi parziali, mentre imporre il

2.2. PARAMETRIZZAZIONE DELLA PALETTATURA DELLA GIRANTE E DEL CANALE INTERSTADIO

carico palare nullo in corrispondenza al bordo di uscita è necessario per contenere gli effetti di scia e la deviazione della corrente relativa a valle della pala.

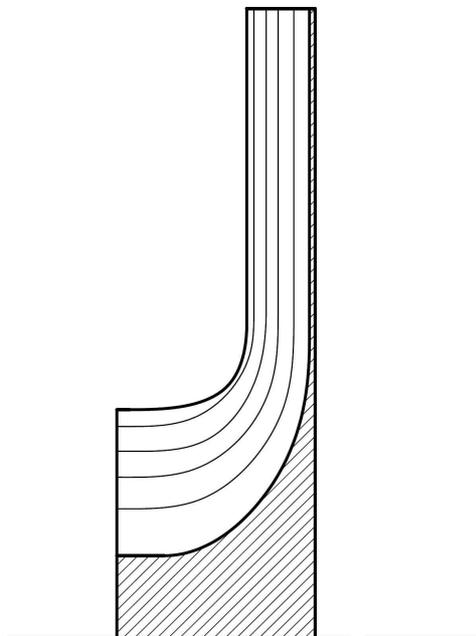


Figura 2.4: Analisi delle linee di corrente nel condotto meridiano di una girante.

In tal modo per individuare una forma della pala sono necessari solo tre parametri. Si sceglie di assegnare la posizione dei punti di giunzione tra i tratti NC e ND e la pendenza α del tratto lineare, similmente a quello suggerito da Zangeneh, vedi [11].

Integrando tale andamento è possibile ricavare il valore della componente tangenziale C_U della velocità lungo tutta la linea di riferimento.

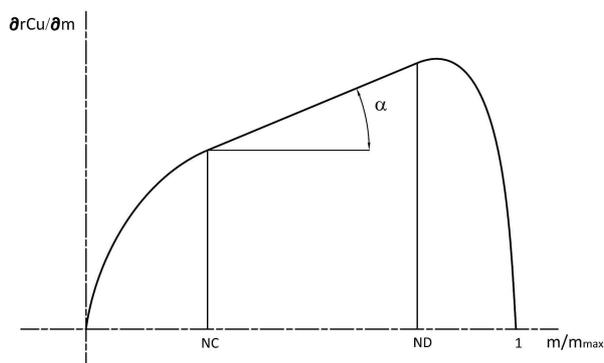


Figura 2.5: Andamento della derivata del momento angolare della quantità di moto lungo la linea media della sezione di riferimento di una girante a semplice curvatura.

I valori degli angoli costruttivi della pala si ricavano dall'andamento appena trovato, note

che siano le velocità meridiane della corrente lungo la linea di riferimento¹ grazie alla relazione:

$$\tan(\beta) = \frac{C_M}{U - C_U}$$

Integrando infine tale andamento è possibile individuare trovare l'anomalia circonferenziale θ della pala secondo la relazione:

$$\tan\beta = \frac{r d\theta}{dm} \rightarrow \theta = \int \tan\beta \frac{dm}{r}$$

Il vantaggio di questo tipo di parametrizzazione è che si ha un controllo diretto su una grandezza, fondamentale per caratterizzare le prestazioni di una schiera di pale, come il carico palare. I tre parametri scelti consentono di caratterizzare agevolmente il carico palare della pala che si sta progettando, infatti risulta abbastanza semplice modificare il carico di una pala e spostando i valori massimi di $\partial r C_u \setminus \partial m$ verso il bordo di ingresso o verso il bordo di uscita.

Principale handicap di questa parametrizzazione è la mancanza di un controllo diretto sulla forma della pala, quindi la mancanza di controllo diretto su una grandezza fondamentale come l'angolo di avvolgimento della pala.

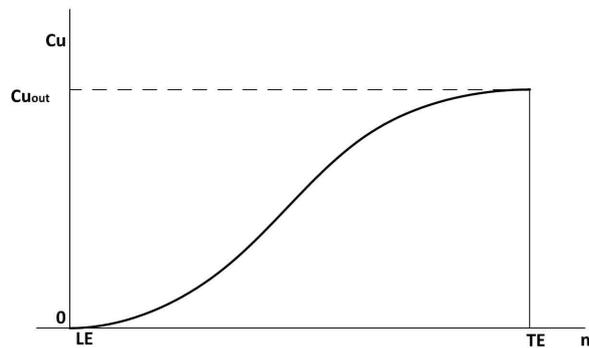


Figura 2.6: Andamento della velocità circonferenziale che risulta da un'assegnazione come in figura 2.5.

Lo schema proposto in figura 2.7 riassume l'applicazione della parametrizzazione proposta per le palettature delle giranti a basso numero tipico di macchina. Si parte quindi dalla definizione di un canale meridiano che rispetti i vincoli di ingombro e le cui dimensioni siano adatte a garantire lo scambio di energia. Si effettua un'analisi della corrente meridiana con metodi a potenziale dal quale è possibile, scelto il numero di pale e la posizione ricavare l'andamento della velocità meridiana sulla sezione di riferimento. Si assegna l'angolo costruttivo di ingresso della pala (ricavato tramite considerazioni sull'incidenza della corrente in ingresso o tramite indicazioni che derivano da simulazioni CFD di altri casi) e l'angolo costruttivo al bordo di uscita (ricavato con considerazioni sullo scambio di energia e grazie al feed back di altre simulazioni CFD). Si assegnano i parametri ND,ND e ALFA che consentono di definire l'andamento della componente $\partial r C_U \setminus \partial m$ lungo la sezione di riferimento e, per integrazione,

¹La velocità meridiana lungo la linea di riferimento si ricava tramite un'analisi a potenziale della corrente nel condotto meridiano come in figura, oppure indirettamente può essere stimata dalle aree di attraversamento del condotto meridiano secondo metodi ormai consueti nella pratica della progettazione macchinistica.

2.2. PARAMETRIZZAZIONE DELLA PALETTATURA DELLA GIRANTE E DEL CANALE INTERSTADIO

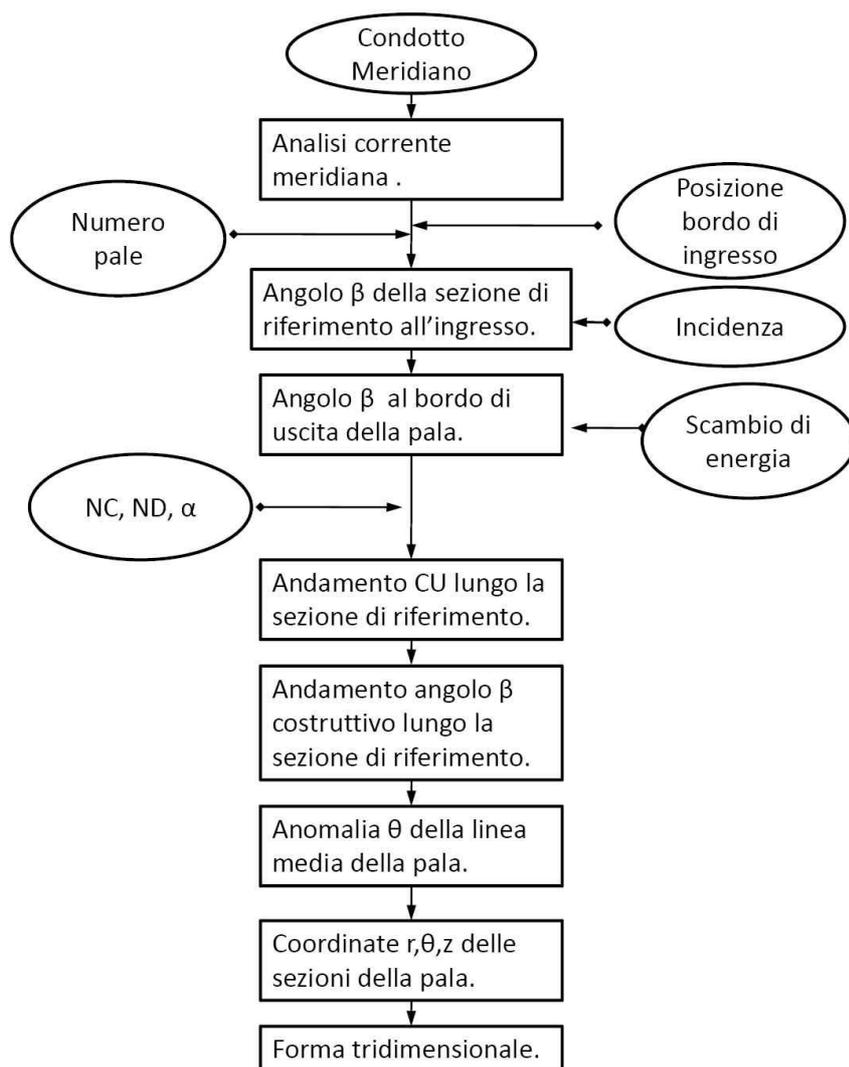


Figura 2.7: Flowchart del dimensionamento della pala di una girante a basso numero tipico di macchina.

si ricava l'andamento della componente C_U della velocità. Noto questo di ricavano i valori dell'angolo β_c lungo la pala e dell'anomalia circonferenziale θ' della sezione di riferimento della pala. Nota questa, dalla condizione di 'semplice curvatura' della pala è possibile ricavare le coordinate (r, θ, z) di tutte le sezioni e quindi la forma tridimensionale della pala.

2.2.3 Parametrizzazione della palettatura della girante

Le giranti di pompe con numero tipico di macchina superiore a 0.5 non sono sagomate con pale a semplice curvatura. La dimensione del condotto meridiano in corrispondenza del bordo di ingresso è tale che il gradiente della velocità in direzione *hub to shroud* della velocità meridiana non è trascurabile. L'angolo del bordo di ingresso di una pala di questo tipo di giranti varia da mozzo a corona e la forma della palettatura di queste giranti è fortemente svergolata.

Nel caso di giranti con flusso della corrente meridiana in uscita inclinato rispetto alla direzione radiale (è il caso delle giranti miste con numero tipico di macchina alto) anche l'angolo costruttivo della pala in corrispondenza del bordo di uscita varia tra le varie sezioni dal mozzo alla corona. Tale angolo varia perchè si mantiene costante lo scambio di energia tra le varie sezioni e quindi cambiando il raggio il bilancio dell'energia impone che cresca la deviazione imposta alla corrente interpalare per compensare la differente componente centrifuga.

La forma della pala che ne risulta è particolarmente complessa e richiede dei metodi di produzione costosi.

E' stata esclusa in questa sede una progettazione inversa della palettatura. In questo tipo di giranti una progettazione inversa prevede di assegnare l'andamento della derivata $\partial r C_u / \partial m$ su tutta la superficie media della pala. Solitamente si assegnano due diversi andamenti su mozzo e su corona, ciascuno come visto nel paragrafo 2.2.2, per un totale di sei parametri. Per le sezioni intermedie si ricava l'andamento corrispondente come interpolazione lineare tra i due. Oltre al carico è necessario assegnare la condizione di impilamento tra i vari profili che aggiunge un ulteriore parametro. A partire da questi parametri è possibile ricavare la forma della pala integrando le equazioni del moto in forma aviscosa. Tale metodologia introduce delle complicazioni matematiche non trascurabili nella fase di progettazione. Inoltre la pala che risulta ha una forma estremamente complicata. Questo, per quello che vedremo nei capitoli successivi, non si sposa con molte tecniche di produzione che si vanno ad applicare per la produzione di queste pompe.

Si è scelto quindi di adottare una parametrizzazione della geometria delle pale che consenta un controllo più diretto sulla forma geometrica della pala: la pala viene costruita parametrizzando la forma della superficie media sul piano conforme $m' - \theta$ (dove la coordinata conforme m' è definita come $m' = \int dm/r$). La forma delle sezioni della pala sono disegnate su questo piano per mezzo di curve di Bezier di terzo grado; quindi per definire ciascuna curva è necessario assegnare quattro punti, due di passaggio agli estremi (che corrispondono al punto del profilo sul naso e al punto sulla coda) e due interni che servono a definire la forma della curva. La forma dei bordi di ingresso può essere assegnata rettilinea o curva secondo le esigenze.

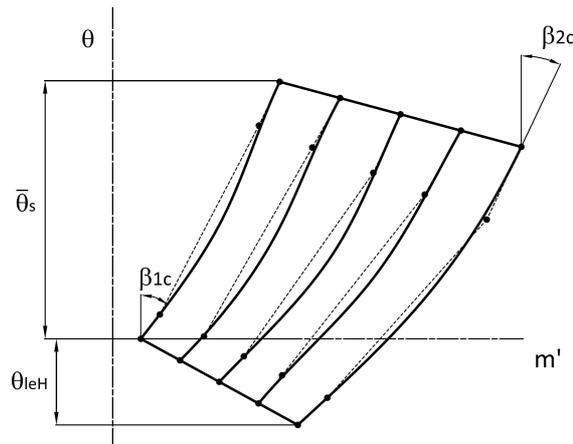


Figura 2.8: Schema della superficie media della pala sul piano conforme $m' - \theta$

I parametri necessari per la definizione completa della superficie media della pala, rappresentati nella figura 2.10, sono quindi:

- l'angolo di avvolgimento del profilo della pala in corrispondenza della corona $\bar{\theta}_S$;
- l'anomalia circonferenziale θ_{leH} ;
- l'inclinazione ξ della proiezione del bordo di uscita sul piano circonferenziale rispetto alla direzione radiale passante per il punto finale della linea media della sezione della corona sul piano circonferenziale .

Oltre a questi parametri è necessario definire la forma delle sezioni della pala sul piano conforme, cioè assegnare la posizione dei punti interni delle curve di Bezier che definiscono tali sezioni. Per ciascuna sezione sono noti gli angoli di ingresso e di uscita e quindi il secondo e il terzo punto della curva di Bezier di ciascuna sezione giacciono rispettivamente sulle rette che partono dai punti della sezione corrispondenti a bordo di ingresso e di uscita e con un'inclinazione pari all'angolo di ingresso e di uscita della pala su quella sezione, come mostrato in figura 2.9.

Rimangono, per ciascuna sezione, due gradi di libertà che corrispondono alla posizione dei punti lungo queste rette. Si è scelto di definire tali posizioni con le coordinate adimensionizzate dell'ascissa curvilinea di ciascun profilo; tali parametri corrispondono per ciascuna sezione ai valori $\bar{m}'_2 = \frac{m'_2}{m'_{tot}}$ e $\bar{m}'_3 = \frac{m'_3}{m'_{tot}}$ secondo la simbologia introdotta in figura 2.9. Inoltre per diminuire il numero di parametri da assegnare si è deciso di imporre tali valori uguali per tutti i profili che si sceglie di utilizzare per la definizione della pala; quindi sono necessari solo due ulteriori parametri per definire la forma completa della pala.

La parametrizzazione che è stata definita in questa sede consente con cinque parametri di gestire la forma della pala in modo completo, mantenendo un controllo diretto sulla forma geometrica. Inoltre è possibile stabilire anche l'avvolgimento della pala, che è una grandezza fondamentale per il controllo delle perdite nella girante, e la condizione di impilamento dei

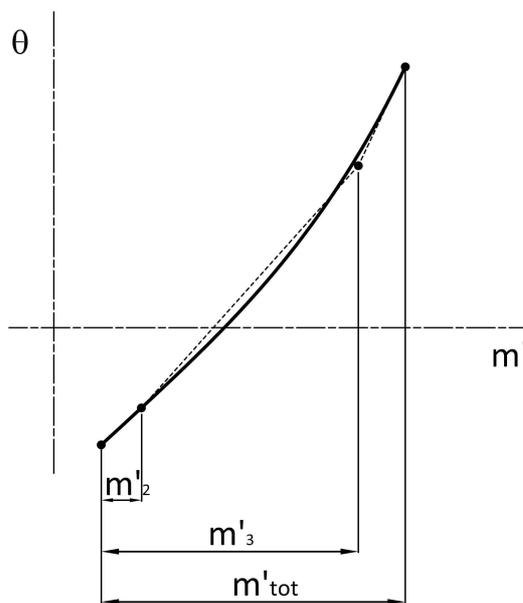


Figura 2.9: Parametrizzazione di una sezione della pala sul piano conforme.

profili della pala. L'effetto fluidodinamico di quest'ultima è più intuitivo se inteso come inclinazione della pala rispetto alla superficie di rivoluzione del mozzo. Modificare la condizione di impilamento significa modificare lo scambio di energia quando si mantengono invariati i profili della pala e quindi gli angoli in uscita. Infatti l'inclinazione della pala è tale da indurre sul fluido una forza con componente assiale verso la superficie del mozzo, si aumenta la pressione media in corrispondenza di quest'ultimo. Localmente aumentare la pressione media comporta, per la conservazione della rotalpia totale all'interno della girante, una diminuzione delle velocità di attraversamento nel condotto interpalare. Secondo la relazione 2.1 vista in precedenza che lega le componenti della velocità alla differenza di pressione tra le facce, si vede che mantenere costante il profilo delle sezioni delle pale comporta mantenere costante la quantità $\partial r C_u \backslash \partial m$. Risulta che modificando localmente la velocità meridiana modifico la differenza di pressione tra le facce della pala e all'uscita questo si traduce in una minore o maggiore deviazione della corrente reale rispetto alla corrente euleriana e quindi in un diverso scambio di energia.

Tale forma di parametrizzazione si discosta, come filosofia, da quella presentata nel paragrafo 2.2.2; non si gestiscono direttamente i parametri di flusso della macchina, ma si controlla direttamente la forma della pala. Il principale svantaggio è che è più complicato per un operatore individuare una relazione di causa-effetto tra il campo di moto della girante e la variazione dei parametri della geometria della pala. Il principale vantaggio di questa formulazione è la possibilità di controllare direttamente la forma della superficie media e la possibilità quindi di vincolare direttamente la forma, in risposta a determinate esigenze di forma.

La procedura che porta quindi a dimensionare una palettatura è riassunta in figura 2.11. Questa parte da un condotto meridiano dimensionato grazie all'esperienza². Stabilito il

²La sagomatura del condotto meridiano della girante può essere fatta con i criteri standard che sono

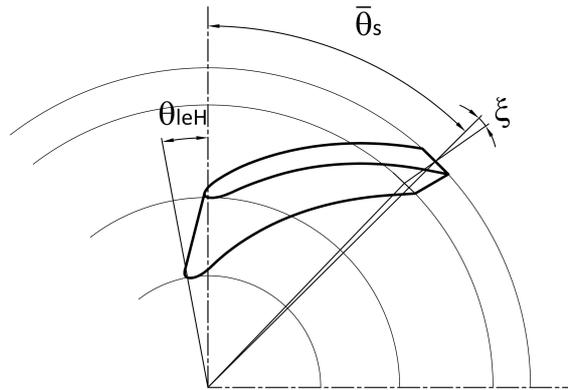


Figura 2.10: Rappresentazione sul piano circolare di una pala corrispondente a quella di figura 2.8.

numero di pale, si ricava la dislocazione delle linee di corrente sul piano meridiano. Assegnata la posizione del bordo di ingresso della palettatura si assegna la distribuzione degli angoli di ingresso della palettatura lungo tutte le sezioni della palettatura, che risponde a considerazioni di natura fluidodinamica. Si assegna inoltre l'andamento dell'angolo costruttivo lungo il bordo di uscita della pala in modo da garantire il corretto scambio di energia. Si sceglie l'angolo di avvolgimento della sezione alla corona della palettatura e l'anomalia circonferenziale θ_{leH} del bordo di ingresso in corrispondenza alla sezione del mozzo e l'angolo ξ dell'inclinazione della superficie media della pala in uscita. Scelti i parametri \bar{m}'_2 e \bar{m}'_3 per le sezioni delle pale è possibile disegnare sul piano conforme $m' - \theta$ la superficie media della pala come in figura 2.8. Dal piano conforme è possibile ricavare le coordinate r, θ, z per tutti i punti della pala e quindi ricavare la forma della pala nello spazio cartesiano.

2.2.4 Palettatura del diffusore

In questa sezione si presenta la parametrizzazione della geometria dei diffusori-convogliatori delle turbomacchine multistadio palettate. Si presentano due diverse parametrizzazioni della forma dei diffusori. La prima prevede pale a spessore costante, mentre la seconda prevede pale con spessore variabile lungo la linea media come si può vedere in figura 2.12. Entrambe le parametrizzazioni fanno riferimento alla definizione di una linea media della palettatura sulla sezione di riferimento. Per ciascun punto della linea media della sezione di riferimento la traccia della superficie media della pala su un piano perpendicolare alla linea media della pala sul piano meridiano è rappresentata da un segmento inclinato rispetto ad una direzione ortogonale alla linea media sul piano meridiano.

Entrambe le parametrizzazioni sono applicabili per la palettatura sia dei canali di ritorno di canali interstadio da accoppiare a giranti a flusso radiale che di diffusori-convogliatori *a boccia*. In particolare nel primo caso, qualora fosse richiesto di progettare delle palettatura a

presentati sui manuali a partire dai diagrammi statistici che consentono una scelta del diametro esterno della girante, e sagomatura per mezzo della distribuzione delle aree di attraversamento.

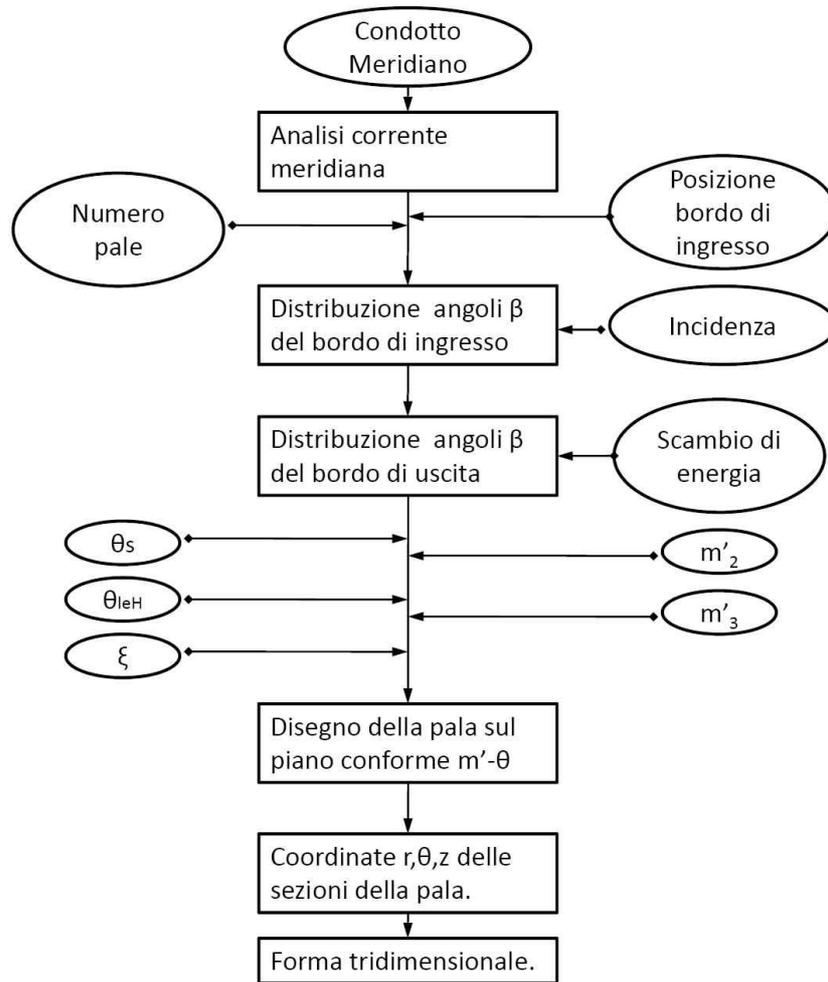
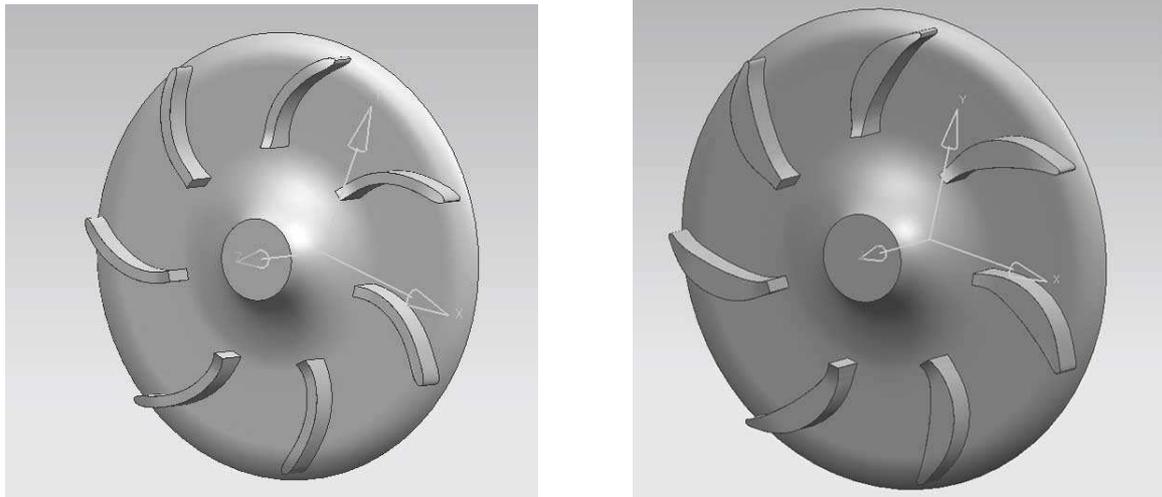


Figura 2.11: Flowchart del dimensionamento della pala di una girante a numero tipico di macchina $K > 0.5$.

semplice curvatura (quindi nella quale le varie sezioni delle pale del canale di ritorno sono tra loro sovrapposte in una vista circonferenziale), tali parametrizzazioni prevedono la possibilità di ottenere queste geometrie, semplicemente impostando che la traccia delle pale sia disposta sempre in direzione assiale.

Le due forme di parametrizzazione sono per molti versi analoghe tra di loro, ma la forma delle palettature che ne risulta è profondamente diversa. Nel primo caso si presenta la parametrizzazione di pale a spessore costante, dove lo spessore viene assegnato in base alle esigenze produttive; la seconda forma di parametrizzazione prevede la parametrizzazione di pale di diffusore a spessore variabile dove lo spessore viene ricavato a partire da un'assegnata distribuzione di velocità. Nella figura 2.12 si può vedere un esempio delle forme che ne risultano, per ciascuna delle parametrizzazioni, per le palettature di un canale di ritorno radiale.



(a) Convogliatore con pale a spessore costante.

(b) Convogliatore con pale a spessore variabile.

Figura 2.12: Differenti tipi di palettature che possono essere alloggiati in un canale di ritorno di un condotto multistadio.

Palettatura a spessore costante

La prima forma di parametrizzazione della pala dei diffusori-convogliatori, ha come punto di partenza la geometrie del canale meridiano dove verranno alloggiati le pale in questione e dal progetto della girante. Il dato dello spessore della pala è un vincolo di progetto, e quindi spesso si ricava da vincoli di fabbricazione.

A partire dal canale meridiano del diffusore si analizza l'andamento della velocità di attraversamento all'interno del canale con metodi a potenziale, così come si è fatto per le giranti. Assegnata la posizione del bordo di ingresso e del bordo di uscita è noto quindi lo sviluppo della linea media della sezione di riferimento sul piano meridiano, e quindi anche i valori della velocità di attraversamento. Si assegna quindi l'andamento della rimozione del momento angolare della quantità di moto così come si è visto per le pale delle girante a semplice curvatura nel paragrafo 2.2.2. Si assegna cioè l'andamento della derivata $\partial r C_u \backslash \partial m$ lungo la linea media. Anche qui risulta conveniente assegnare tale andamento sempre con la parametrizzazione vista nel paragrafo 2.2.2. Quindi sono necessari tre parametri per individuare tale andamento che sono, come in figura 2.5, le posizioni dei punti di congiungimento tra i tratti **NC** e **ND** e la pendenza α del tratto lineare.

L'angolo costruttivo della pala in corrispondenza del bordo di ingresso è individuato grazie dall'analisi CFD della corrente a valle della girante; posizionando correttamente una superficie di riferimento in corrispondenza della posizione del bordo di ingresso della pala del diffusore è possibile calcolare le componenti della velocità assoluta e quindi l'angolo che deve essere assegnato alla pala del diffusore. L'angolo della pala in corrispondenza del bordo di uscita deve garantire un corretto raddrizzamento della corrente assoluta a valle del diffusore-convogliatore in modo garantire un corretto funzionamento della girante dello stadio successivo.

Note queste grandezze, integrando l'andamento della derivata $\partial r C_u \backslash \partial m$ lungo tutta la pala si ricava il valore dell'anomalia θ della pala e quindi le coordinate dei punti della pala.

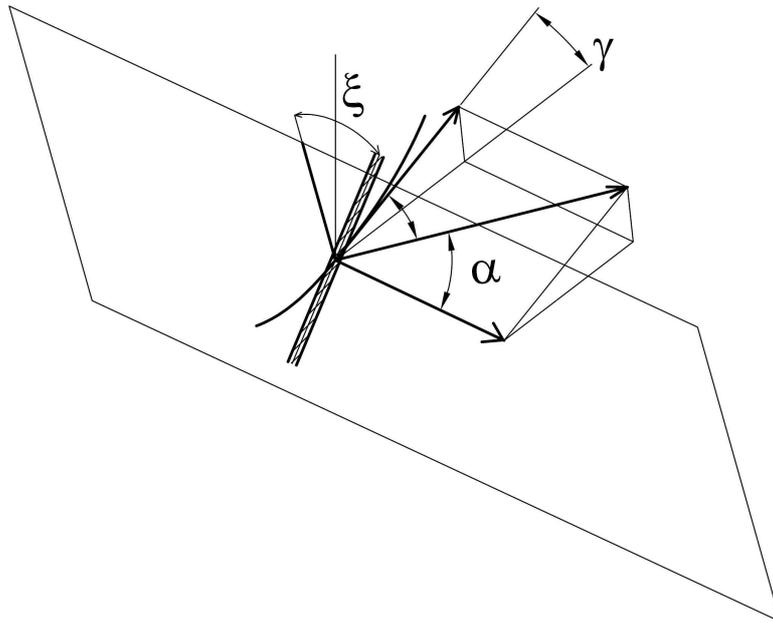


Figura 2.13: Inclinazione della pala rispetto alla linea media della pala.

È necessario definire anche l'inclinazione della traccia della sezione della pala. Questo significa imporre l'andamento dell'angolo ξ definito in figura 2.14. Tale angolo è l'inclinazione della traccia della superficie media della pala rispetto alla perpendicolare alla linea media sul piano meridiano. L'andamento di tale angolo lungo la linea media della superficie di riferimento della pala può essere parametrizzato in vari modi. La soluzione che è stata scelta, a vantaggio della semplicità, prevede di assegnare un valore per questo angolo in corrispondenza al bordo di ingresso della pala e uno in corrispondenza del bordo di uscita. Si assume, per tutti i punti interni della pala un andamento lineare tra questi due valori; in questo modo sono necessari soltanto due parametri per definire l'andamento di questo angolo lungo tutta la pala che sono i valori di tale inclinazione in corrispondenza del bordo di ingresso e del bordo di uscita.

Palettatura a spessore variabile

Si propone anche un'altro tipo di parametrizzazione della palettatura dei canali interstadio. Lo schema di parametrizzazione ricalca in molti punti quello proposto nel paragrafo precedente. Tuttavia si introduce anche una parametrizzazione dello spessore della pala che richiede due ulteriori parametri per definire in modo completo la forma della palettatura. La geometria che ne risulta da una parametrizzazione di questo tipo è estremamente differente da quella vista nella sezione precedente per i diffusori a spessore costante.

L'idea per questa forma di palettatura nasce dall'osservazione di alcuni tipi di canali interstadio di pompe multistadio in cui lo spessore della pala raddrizzatrice non è costante, ma varia lungo lo sviluppo della pala stessa. In questo modo, aumentando lo spessore palare

2.2. PARAMETRIZZAZIONE DELLA PALETTATURA DELLA GIRANTE E DEL CANALE INTERSTADIO

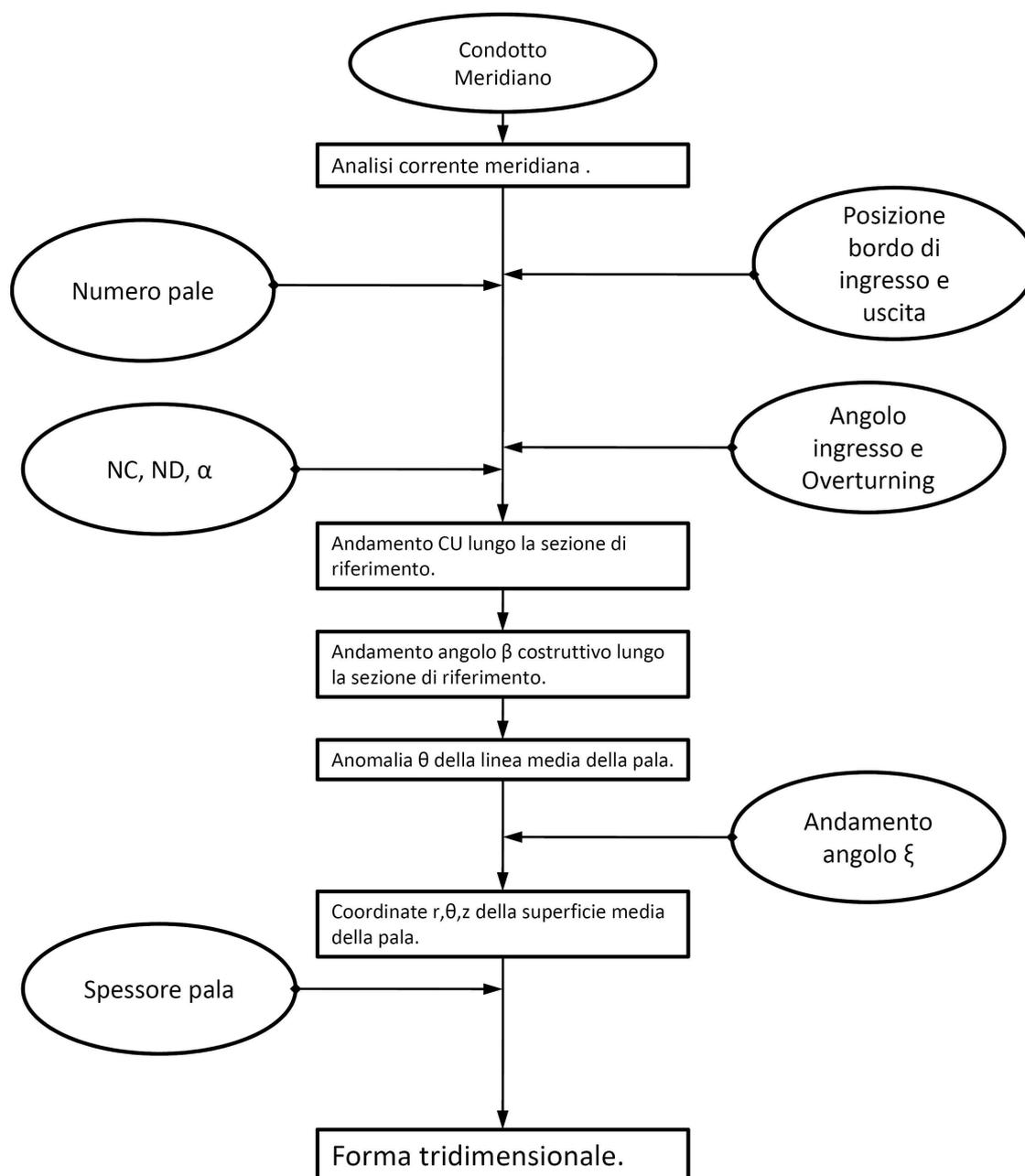


Figura 2.14: Flowchart del dimensionamento della pala del diffusore a spessore costante.

si aumenta la velocità di attraversamento nel canale interpolare in modo tale da prevenire eventuali stalli o distacchi di vena.

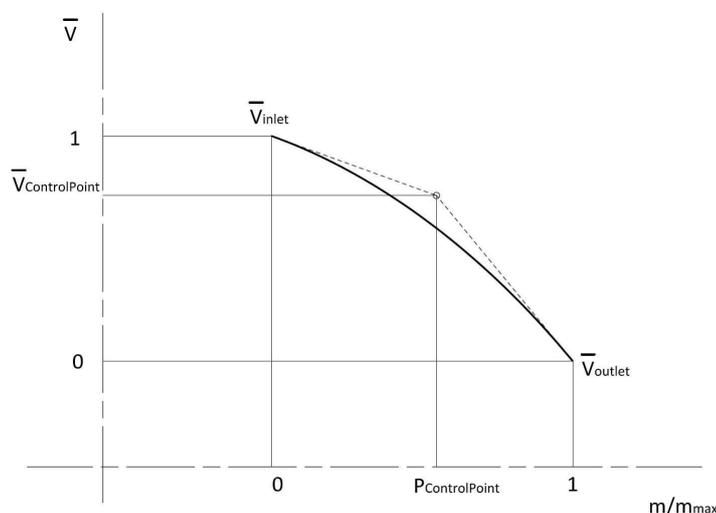


Figura 2.15: Parametrizzazione dell'andamento delle velocità nel condotto interpolare del diffusore.

La pala in questione è definita anche qui a partire dal condotto meridiano e dalla posizione dei bordi di ingresso e di uscita. Noti questi con un'analisi a potenziale si ricava l'andamento delle velocità di attraversamento lungo la linea media del canale meridiano. Si assegna quindi l'andamento della derivata $\partial r C_u \backslash \partial m$ lungo la linea media, utilizzando anche in questa occasione la parametrizzazione proposta che consente di definire in modo completo questo andamento con i tre parametri **NC**, **ND**, e α , dal quale, per integrazione si ricava l'andamento della velocità C_U lungo la linea media.

Si assegna inoltre l'andamento della velocità assoluta C lungo la linea media della pala. E' possibile quindi ricavare punto per punto l'angolo costruttivo della pala che è imposto pari a quello di flusso:

$$\alpha_C = \alpha = \arccos\left(\frac{C_U}{C}\right).$$

Si ricava inoltre punto per punto la velocità di attraversamento $C'_m = \sqrt{C^2 - C_U^2}$ all'interno del condotto interpolare. Da tale andamento si ricava l'andamento dell'ostruzione palare ζ come rapporto tra questa velocità ricavata e l'andamento della velocità meridiana nel canale meridiano in assenza di pale C_m che si ricava dall'analisi a potenziale:

$$\zeta = \frac{C'_m}{C_m}$$

da questo è possibile quindi ricavare l'andamento dello spessore palare normale alla pala.

L'andamento degli spessori quindi è ricavato a partire dalla velocità di attraversamento del fluido nel condotto interpolare, che è proprio la grandezza che si vuole controllare per limitare i distacchi. Visto che gli stalli sono dovuti alle decelerazioni nel condotto interpolare, si scegli

di assegnare l'andamento della velocità di attraversamento di tale velocità come una funzione monotona tra la velocità in corrispondenza del bordo di ingresso e quella in corrispondenza del bordo di uscita che sono note. Per assegnare questa velocità si utilizza una curva di Bezier di secondo grado, come in figura 2.15. Per imporre tale andamento monotono, è sufficiente posizionare il punto di controllo della curva all'interno del rettangolo che ha per diagonale il segmento che congiunge i punti che rappresentano le velocità in corrispondenza del bordo di ingresso e del bordo di uscita. Quindi sono necessari due ulteriori parametri per definire tale andamento che sono le coordinate del punto appoggio. Tali parametri normalizzati vengono quindi definiti come la posizione normalizzata del punto di controllo (P_{CP}) e la velocità corrispondente al punto di controllo (\bar{V}_{CP}), normalizzata tra gli estremi che sono la velocità in ingresso e quella in uscita, .

Resta da definire l'inclinazione della traccia della palettatura rispetto linea media nel piano meridiano, che è analoga a quanto visto per la parametrizzazione della palettatura a spessore costante.

Sono necessari quindi un totale di sette parametri per definire la palettatura contro i cinque sufficienti per la palettatura a spessore costante. Lo schema proposto per la parametrizzazione è riassunto in figura 2.16.

2.3 Parametrizzazione della forma del condotto meridiano del diffusore

Anche lo sviluppo meridiano del diffusore-convogliatore della turbomacchine pluristadio varia al variare dei dati di progetto della macchina e segue in un certo modo la forma della girante. Come visto nel primo capitolo per stadi con giranti a basso numero tipico di macchina il canale meridiano prevede una netta distinzione tra la zona del diffusore a valle della girante che accoglie un flusso radiale in direzione centrifuga e la zona del canale di ritorno in cui il flusso è centripeto e convogliato verso l'aspirazione della girante successiva. Per le macchine con giranti con numero tipico più alto la geometria del convogliatore si fa più compatta e le due zone tendono a confondersi quando si utilizzano i cosiddetti diffusori a boccia (*bowl diffusers*), arrivando ad utilizzare a volte, per giranti a flusso misto, diffusori a boccia con il bordo di ingresso disposto in un tratto di canale meridiano a flusso assiale.

Le pompe sommerse multistadio hanno peculiarità leggermente differenti dalle pompe comuni. Queste macchine devono essere calate in pozzi, quindi hanno dei vincoli di ingombro radiali molto restrittivi. Le taglie delle pompe sommerse multistadio sono definite in base al diametro del pozzo in cui possono essere calate. Le più comuni sono 3-4-5-6-8-10-12 pollici. Per poter mantenere questi ingombri anche con macchine che elaborano portate elevate spesso il campo di utilizzo delle macchine a flusso misto invade quello delle macchine a flusso radiale. I diffusori-convogliatori con imbocco assiale e flusso risultano spesso usati in questo tipo di macchine.

Per quello che riguarda gli organi statorici di macchine pluristadio a basso numero tipico di macchina è rimandiamo a [32] per un esempio di parametrizzazione del condotto completo.

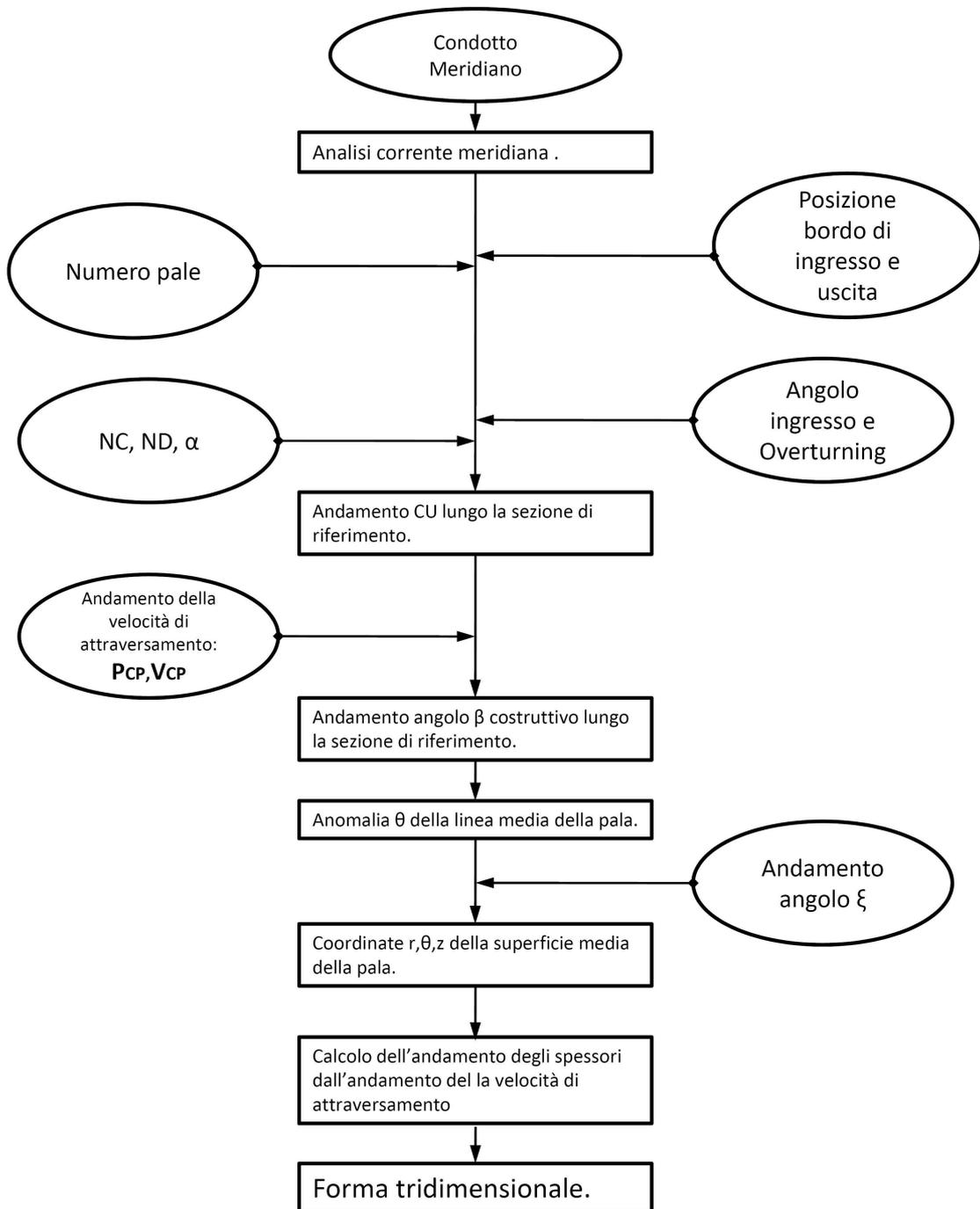


Figura 2.16: Flowchart del dimensionamento della pala del diffusore a spessore variabile.

2.3. PARAMETRIZZAZIONE DELLA FORMA DEL CONDOTTO MERIDIANO DEL DIFFUSORE

Presentiamo in questo paragrafo, invece, una parametrizzazione della geometria del canale meridiano dei diffusori da accoppiare nelle macchine multistadio sommerse con giranti a flusso misto. Si è scelto di approssimare il canale in questione con un breve tratto assiale che accoglie la corrente in uscita dalla girante, e con un tratto a flusso misto radiale-assiale a pareti parallele rettilinee che segue un tratto di raccordo. Per generare il tratto di raccordo si è fatto uso di una curve di Bezier di terzo grado, una come raccordo del profilo del mozzo e una come raccordo del profilo della corona.

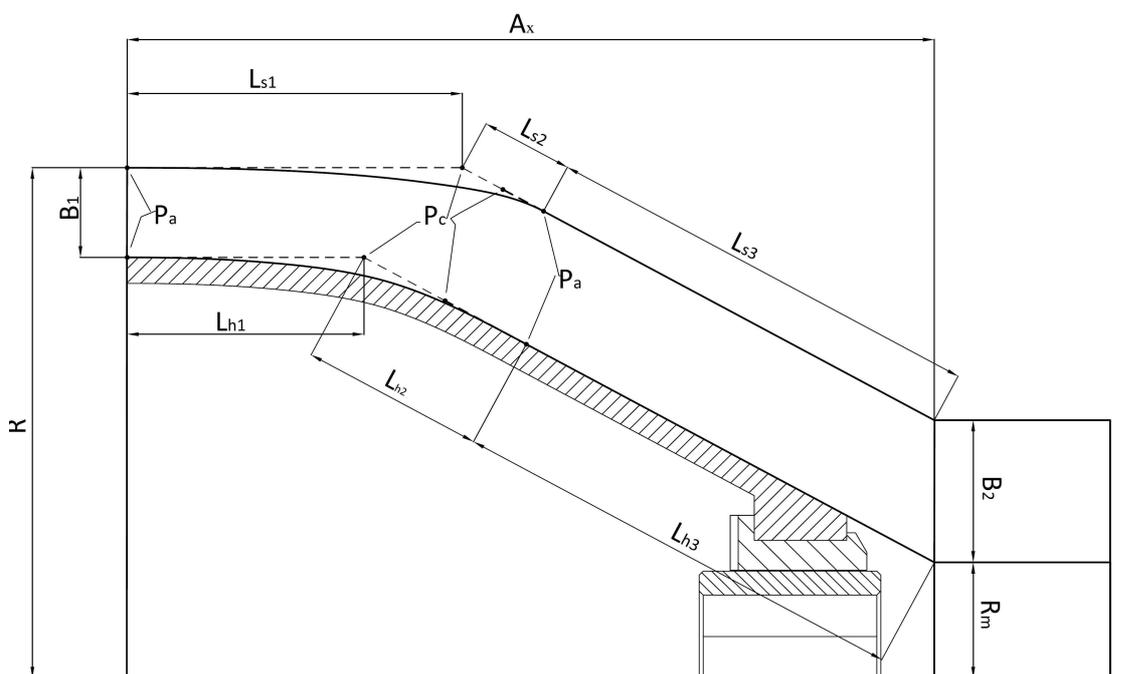


Figura 2.17: Grandezze coinvolte nella parametrizzazione del condotto meridiano.

La figura 2.17 mostra le varie grandezze coinvolte nella definizione del condotto come lo si era definito; l'analisi attenta di questa geometria ha portato a questa parametrizzazione che non esaurisce tutte le geometrie possibili per questo organo e non conteggia tutti particolari di forma che sono presenti nel pezzo che verrà realizzato (boccole, bronzine, anelli di rasamento...ecc), ma consente di avere delle informazioni valide sul profilo ideale che si può realizzare e dal quale è bene scostarsi il meno possibile in fase di ingegnerizzazione.

Il profilo realizzato è composto da un tratto curvo che accoglie la corrente in direzione assiale e la piega verso il secondo tratto a flusso misto, costituito da un tratto a pareti parallele. Per generare il tratto curvo si è fatto uso di due curve di Bezier di terzo grado, come dire con due punti di controllo. Si è scelto di posizionare il secondo punto del poligono di definizione di ciascuna curva sul punto di intersezione tra la retta parallela all'asse della macchina passante per il primo punto di ciascun profilo e il prolungamento del segmento che genera la parte rettilinea del profilo. Inoltre si è scelto di posizionare il terzo punto del poligono (il secondo

che non giace sulla curva) in corrispondenza del punto medio del segmento che congiunge il precedente punto di controllo con l'ultimo punto del poligono, cioè con il punto dove termina la curva di Bezier e incomincia il tratto rettilineo. Tale scelta di posizionamento è stata necessaria innanzi tutto perché garantisce che la tangente delle curve in ingresso al canale sia orizzontale, quindi in asse con la macchina; allo stesso modo così si garantisce la continuità della derivata prima del profilo in corrispondenza congiungimento tra il tratto curvo e il tratto rettilineo. Un risultato di questo tipo si sarebbe ottenuto anche con una curva di Bezier con un solo punto di controllo, posizionato sull'intersezione tra la retta parallela all'asse passante per il primo punto di appoggio e il prolungamento del tratto rettilineo del canale; tuttavia una scelta del genere portava a generatrici del tratto di raccordo con cambi di concavità sicuramente dannosi a livello fluidodinamico. E' stato necessario introdurre il secondo punto di appoggio e vincolarlo in quella posizione per garantire la regolarità della curva.

Nella figura 2.17 si può vedere il significato geometrico dei parametri che sono stati utilizzati per definire i profili di mozzo e corona. Sono stati inoltre indicati i punti del poligono del di definizione delle curve di Bezier; sia quelli interni che i due quelli estremi.

Si è scelto di fare uso di grandezze adimensionali rapportando tutti i parametri ad una grandezza di riferimento in modo da creare una parametrizzazione adatta a definire organi con dimensioni diverse tra di loro. La grandezza che è stata scelta a questo scopo è la dimensione di ingombro radiale, scelta per l'importanza che ha in questa macchina. Per definire una geometria sono necessari sette parametri adimensionali. Alcuni parametri sono stati definiti rapportando una lunghezza ad altre lunghezze a loro volta rapportate con quella radiale di riferimento in modo da rendere più agevole e intuitiva la definizione di tali parametri.

Lunghezze coinvolte della definizione di queste geometrie sono innanzitutto:

$$Lung_h = L_{h2} + L_{h3}$$

$$Lung_s = L_{s2} + L_{s3}$$

La lista dei sette parametri utilizzati per questa definizione diventa quindi:

- $\frac{A_x}{R}$
- $\frac{B_2}{R}$
- $\frac{R_m}{R}$
- $\frac{B_1}{B_2}$
- $\frac{L_{h3}}{Lung_h}$
- $\frac{L_{s3}}{Lung_s}$
- $\frac{L_{h1}}{A_x}$.

Ciascuna geometria, quindi, è definita assegnando un valore per ciascun parametro e un valore per l'ingombro radiale massimo della parte idraulica di questi condotti.

2.3. PARAMETRIZZAZIONE DELLA FORMA DEL CONDOTTO MERIDIANO DEL DIFFUSORE

Capitolo 3

Algoritmi di ottimizzazione

La letteratura che riguarda il campo dell'ottimizzazione ormai è molto ampia. In realtà le molteplici tecniche di ottimizzazione che sono state proposte nel corso degli anni offrono i risultati migliori per campi di applicazione diversi. Infatti spesso le tecniche sono state pensate per essere applicate ad una determinata casistica, oppure hanno subito migliorie nel tempo in relazione ad un preciso campo di applicazione.

Esistono quindi le più svariate tecniche di ottimizzazione. Per esempio in presenza di funzioni semplici, intendendo funzioni regolari e continue, risultano ancora efficaci i metodi di ricerca dei minimi utilizzando le derivate parziali. Nell'ambito della programmazione lineare un algoritmo che si è dimostrato nel passato molto efficace è l'algoritmo del semplice. I problemi di programmazione lineare sono problemi di ricerca di massimi o minimi di funzioni lineari su uno spazio vincolato. Tale algoritmo è computazionalmente molto efficace perchè non richiede il calcolo delle derivate, tuttavia mostra i suoi limiti quando cresce notevolmente il numero delle variabili in gioco.

Risulta chiaro che ogni qualvolta che si approccia un problema di ottimizzazione è necessario scegliere la tecnica che meglio si adatta al caso in analisi.

Per risolvere problemi di ottimizzazione che coinvolgono simulazioni al computer le tecniche che si sono mostrate più efficaci sono tecniche che nascono dall'osservazione di fenomeni naturali o biologici nel tentativo di descriverli tramite un formalismo matematico; di fatto la genesi culturale di questi metodi è simile a quella delle reti neurali, che sono nate dall'osservazione dei sistemi nervosi biologici. I metodi di ottimizzazione in questione sono i metodi della Simulated Annealing (ricottura simulata) che imita il processo fisico di ricottura dei solidi, gli algoritmi genetici (GA-genetic algorithms) che si ispirano al processo di evoluzione e selezione naturale, infine la teoria degli sciame di particelle (PSO-particle swarm optimizer) che prende le mosse dal volo degli stormi di uccelli alla ricerca del cibo.

Tutti questi metodi condividono il tratto fondamentale che non richiedono il calcolo di gradienti, caratteristica indispensabile con casi sperimentali, come sono i computer experiments, dove si hanno a disposizione solo valori puntuali della funzione obiettivo che si vuole minimizzare.

3.1 Differenti approcci all' ottimizzazione

L'approccio alle tecniche di ottimizzazione nella pratica ingegneristica si è con il tempo evoluto verso due strade diverse. Nella progettazione e ottimizzazione ingegneristica, come già visto, si fa largo uso della sperimentazione tramite utilizzo di prototipi. Recentemente l'introduzione dei software di simulazione al computer ha consentito di sostituire, almeno in parte la sperimentazione sul campo con una sperimentazione virtuale.

Nell'evoluzione della letteratura tecnico-scientifica che propone l'uso di tecniche di ottimizzazione si sono selezionati due approcci diversi per l'utilizzo di tali tecniche. Un primo approccio prevede l'impiego della tecnica di ottimizzazione nella sua forma più pura, mentre il secondo approccio ne prevede l'utilizzo assistito da metamodelli.

3.1.1 Ottimizzazione pura

Si può parlare di ottimizzazione pura quando le tecniche di ottimizzazione vengono applicate direttamente sullo spazio di indagine delle variabili; quindi a partire da una soluzione di primo tentativo l'algoritmo di ricerca dell'ottimo per via iterativa aggiorna di volta in volta la soluzione trovata.

Per ciascuna iterazione le tecniche di ottimizzazione prevedono il calcolo di un insieme di possibili soluzioni che si tramuta in un insieme di dati sperimentali da raccogliere. Per questo lo sforzo di raccolta dati può essere particolarmente intenso, implicando il calcolo di parecchie soluzioni per ciascuna iterazione. Per incrementare la convergenza dell'algoritmo è necessario tarare correttamente i parametri del metodo di ottimizzazione e partire da una posizione di primo tentativo sufficientemente vicina all'ottimo cercato, in modo da restringere il campo di indagine ed evitare che il metodo rimanga intrappolato in un minimo locale lontano dalla soluzione reale.

Il vantaggio principale di questo approccio è che viene ricercata e individuata la soluzione ottima tra tutte quelle possibili nello spazio di indagine. Lo svantaggio è che il tempo necessario a completare la raccolta dati potrebbe essere notevole e non stimabile a priori.

3.1.2 Ottimizzazione assistita da metamodelli

Un altro approccio all'utilizzo delle tecniche di ottimizzazione prevede l'utilizzo di tali tecniche assistite da metamodelli.

Questo approccio è stato nel tempo sviluppato con metodi di volta in volta diversi; tuttavia tutti i metodi prevedono di partire da un insieme di dati sperimentali con i quali viene addestrato un modello matematico che consente di prevedere la relazione tra i valori dei parametri e i risultati della sperimentazione anche per valori dei dati di cui non sono disponibili i valori. Tale modello prende il nome di *metamodelli* (in inglese in bibliografia si possono rintracciare con i nomi di *metamodel* o *surrogates model*) perchè si sostituiscono alla fase di sperimentazione; accedendo al modello matematico si possono ricavare i valori che caratterizzano ciascuna soluzione per estrapolazione dai risultati già noti.

L'algoritmo di ottimizzazione quindi accede al metamodello e non direttamente alla sperimentazione (reale o virtuale), garantendo risparmio di tempo e di risorse.

Il principale difetto di questo approccio è dovuto alla fisiologica imprecisione del modello matematico di regressione rispetto alla sperimentazione. Per diminuire questa mancanza di precisione è necessario raccogliere i dati dell'insieme di dati sperimentali nel modo migliore, utilizzando tecniche DOE per massimizzare la precisione e confrontare diversi metodi di regressione al fine di costruire il metamodello più preciso possibile.

3.2 Ottimizzazione vincolata

Buona parte dei problemi di ottimizzazione che devono essere affrontati nella pratica ingegneristica sono problemi di ottimizzazione vincolata, cioè lo spazio di variazione di \mathbf{x} è limitato e la ricerca dell'ottimo deve essere vincolata all'interno di questo spazio; soluzioni esterne non sono accettabili perchè non realizzabili o non aderenti ai problemi in analisi. Generalmente per vincolare lo spazio di indagine si modificano la funzione obiettivo che si intende minimizzare, andando ad incrementare la convessità della funzione da minimizzare intorno allo spazio di indagine.

Un sistema è quello di applicare una funzione di penalizzazione: in tale modo alla funzione che viene minimizzata viene sommato un termine di penalizzazione. Tale termine cresce quanto sono violate le condizioni di vincolo del problema ed è nulla quando la soluzione che viene valutata è all'interno del dominio di indagine. Il termine di penalizzante è, solitamente il prodotto di un coefficiente e di una funzione.

Supponiamo di dover minimizzare una funzione $f(x)$ sotto i vincoli di $c_i(x) \geq 0$. Allora il problema può essere risolto con una serie di problemi di minimizzazione $\min \Phi(x)$ con la funzione $\Phi(x) = f(x) + \sigma g(c(x))$ dove $g(c(x)) = (\min(0, c(x)))^2$. In questo caso σ rappresenta il coefficiente di penalizzazione mentre g è la funzione di penalizzazione. Ad ogni iterazione è possibile incrementare σ per tener conto della convergenza del metodo.

Una metodologia simile è quella che consiste nel creare una barriera incrementando la funzione obiettivo di una funzione che rimane nulla all'interno dello spazio di indagine, ma aumenta in modo proporzionale alla distanza dai bordi quando si esce vincoli dello spazio di indagine.

3.3 Simulated Annealing

La tecnica di ottimizzazione che prende il nome di *simulated annealing* (ricottura simulata) trae ispirazione dal processo fisico di ricottura dei solidi. E' il processo nel quale un solido viene riscaldato fino ad una temperatura massima alla quale tutte le particelle del solido sono libere di muoversi liberamente, e successivamente viene raffreddato lentamente. In questo modo tutte le particelle si assestano allo stato solido in una posizione di bassa energia, a patto di condurre il raffreddamento sufficientemente lentamente.

Esistono molte varianti della tecnica del simulated annealing, ma tutte seguono uno schema

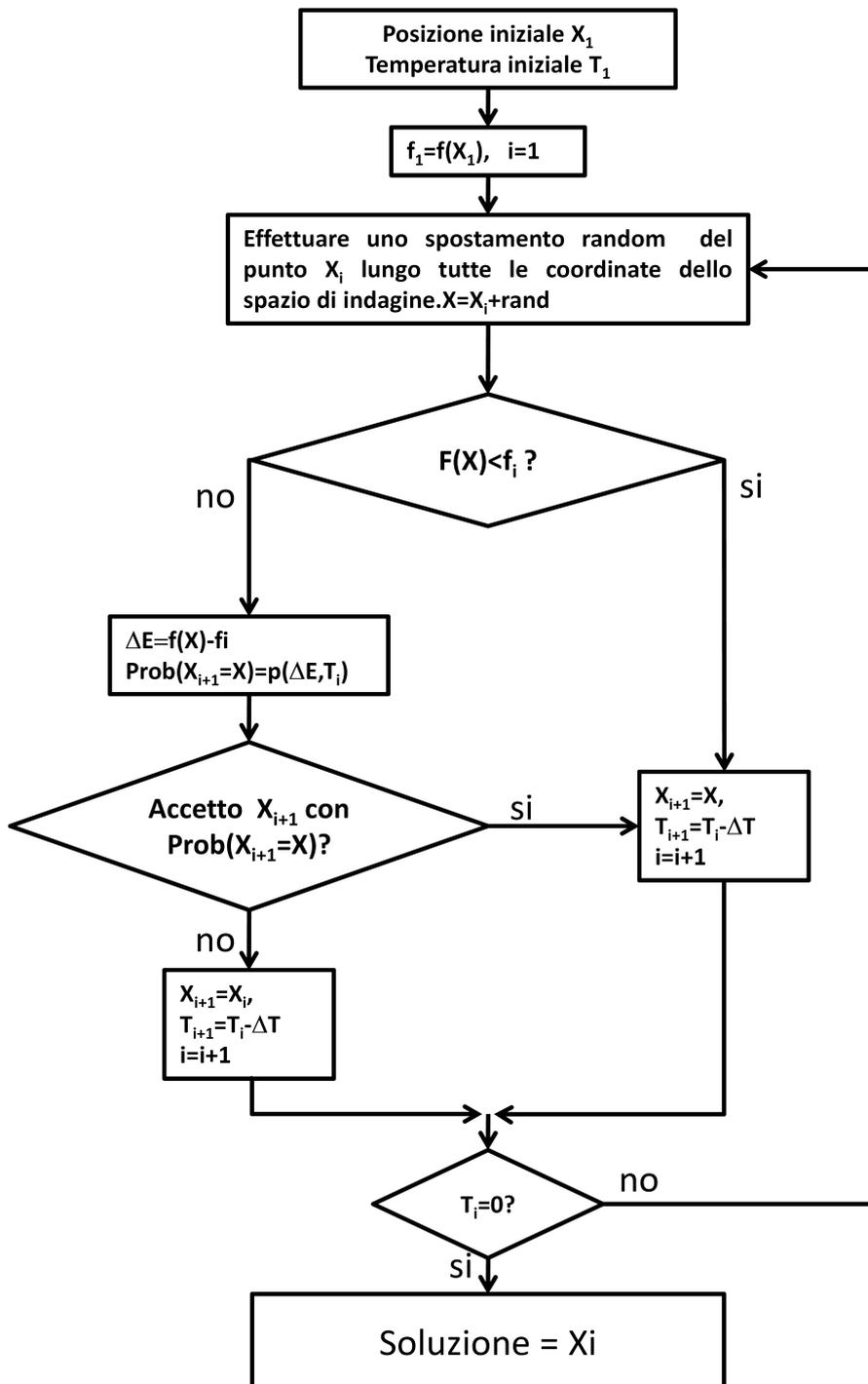


Figura 3.1: Schema dell'ottimizzazione con la Simulated Annealing.

comune. La soluzione di primo tentativo \mathbf{x} viene perturbata in modo casuale arrivando alla soluzione $\bar{\mathbf{x}}$. Se alla nuova posizione corrisponde un valore della funzione obiettivo minore allora $\bar{\mathbf{x}}$ viene accettata come nuova condizione di partenza. In caso contrario la nuova posizione, anche se peggiorativa, può lo stesso essere accettata. La probabilità che la soluzione peggiorativa $\bar{\mathbf{x}}$ possa rimpiazzare \mathbf{x} è calcolata su una temperatura virtuale e su un'energia virtuale. Quest'ultima è pari alla differenza tra i valori della funzione obiettivo calcolati per le due diverse scelte di parametri. La temperatura virtuale invece è un numero che decresce con il passare delle iterazioni del processo evolutivo e riduce la probabilità di accettare come soluzione una soluzione peggiorativa.

Anche se ad un primo sguardo la scelta di accettare una soluzione peggiorativa può non essere sensata, questa acquista logica nel contesto di un'ottimizzazione globale in quanto ci allontaniamo da un soluzione per cercarne una migliore. Questo passaggio è in analogia con il reale processo della ricottura nei solidi, dove ci si muove da una configurazione verso una configurazione più stabile e tale spostamento è basato sulla differenza di energia tra le due posizioni e sulla temperatura. Inoltre distingue la ricottura simulata da una ricerca puramente casuale e consente a tale metodo di aumentare velocità di convergenza, anche se è necessario comunque valutare numerosi casi prima di trovare l'ottimo.

La differenza tra le varie forme di simulated annealing riguarda il numero di variabili che vengono di volta in volta modificate ad ogni iterazione e il modo in cui vengono modificate. Per una riesamina più completa sulla tecnica del simulated annealing si rimanda a [41].

3.4 Algoritmi genetici-GA

Gli algoritmi genetici riproducono il processo evolutivo della specie umana. Ciascun individuo è rappresentato da un cromosoma e rappresenta una delle possibili soluzioni per un problema. La *qualità* di un individuo è misurata mediante una *funzione di fitness*, che è la funzione da minimizzare. In termini biologici la funzione di fitness indica l'adattabilità all'ambiente degli individui: quello che meglio si adattano all'ambiente hanno più probabilità di riprodursi e di trasmettere i propri geni alle generazioni future. Quindi nell'algoritmo genetico la soluzione migliore di un'altra perchè ha un valore inferiore della funzione obiettivo è l'equivalente genetico di un individuo con più chance di trasmettere i propri geni. Per qualunque riferimento agli algoritmi genetici si rimanda a [46].

Sono molteplici i valori da impostare che influiscono sulla convergenza del sistema in termini di efficacia e velocità nel trovare la soluzione. Non esistono valori che funzionano universalmente meglio degli altri e ci si deve affidare all'esperienza per individuare quelli che sono più adatti ad un problema piuttosto che ad un'altro.

Partendo da una popolazione iniziale, un algoritmo genetico produce nuove generazioni che contengono (di solito) individui migliori delle precedenti e quindi l'algoritmo evolve verso l'ottimo globale della funzione di fitness. Come per ciascun algoritmo di quelli che presentiamo in questo capitolo non esistono dimostrazioni che un GA trovi una soluzione globalmente ottima, ma è in grado di trovare una soluzione buona in tempi ragionevoli.

Nel modello tradizionale, i cromosomi sono rappresentati con stringhe di numeri binari, e tutte le generazioni hanno popolazioni eguale numerosità. Ogni cromosoma rappresenta un punto dello spazio di ricerca; quindi è necessario rappresentare le coordinate in base binaria, tutte con la stessa lunghezza. In altre parole ciascun individuo è rappresentato con una stringa di bit sempre della stessa lunghezza. In caso di variabili discrete la conversione è semplice. La rappresentazione di variabili continue è più complicata e deve essere effettuata con accuratezza. Di fatto è necessario troncare le cifre decimali dei numeri, ma l'operazione di troncatura deve essere tale da garantire un adeguato livello di precisione.

Un esempio di come è possibile rappresentare un individuo corrispondente ad un punto di coordinate cartesiane $(x_1, x_2, x_3, x_4) = (18, 3, 1, 4)$ si può vedere in tabella 3.4. Ciascuna coordinata è stata riscritta in base binaria con numeri composti sempre di cinque cifre, e disposti in successione in una stringa più lunga. Il risultato è un cromosoma individuato da una stringa binaria di 20 cifre.

Tabella 3.1: Esempio di cromosoma binario a rappresentazione di un individuo.

← Stringa di lunghezza 20 →			
10010	00011	00001	00100
← x_1 →	← x_2 →	← x_3 →	← x_4 →

Gli operatori utilizzati ricostruiscono in forma matematica le principali caratteristiche del processo di selezione naturale, gli operatori più importati sono la ricombinazione (*crossover*) e la mutazione. La ricombinazione combina i geni tipicamente di due individui per produrre individui figli che ereditano caratteristiche da entrambi i genitori. La mutazione reintroduce nella popolazione materiale genetico perduto.

Questo tipo di ricerca realizza un compromesso tra ricerca della soluzione (*'exploitation'* e esplorazione dello spazio di ricerca (*'exploration'*). Questo corrisponde ad un compromesso tra ricerca globale e ricerca locale; *'exploitation'* eccessiva può portare l'algoritmo a convergere ad una soluzione non accettabile (la ricerca rimane intrappolata in un ottimo locale), *'exploration'* eccessiva può non sfruttare appropriatamente la conoscenza già disponibile rendendo il processo di ricerca molto lento. In altre parole ancora il compromesso è tra ricerca del minimo basandosi sulle informazioni a disposizione e ricerca casuale. In questo punto giace la differenza tra GA e simulated annealing che invece perturba la popolazione con metodi puramente casuali.

3.4.1 Elementi di base di un GA

Una nuova generazione $P(t+1)$ è ottenuta dalla popolazione $P(t)$ per mezzo dei seguenti passi, schematizzati anche nel flow-chart di figura 3.2:

- *valutazione*: si valuta la qualità di ogni individuo tramite la funzione di fitness;

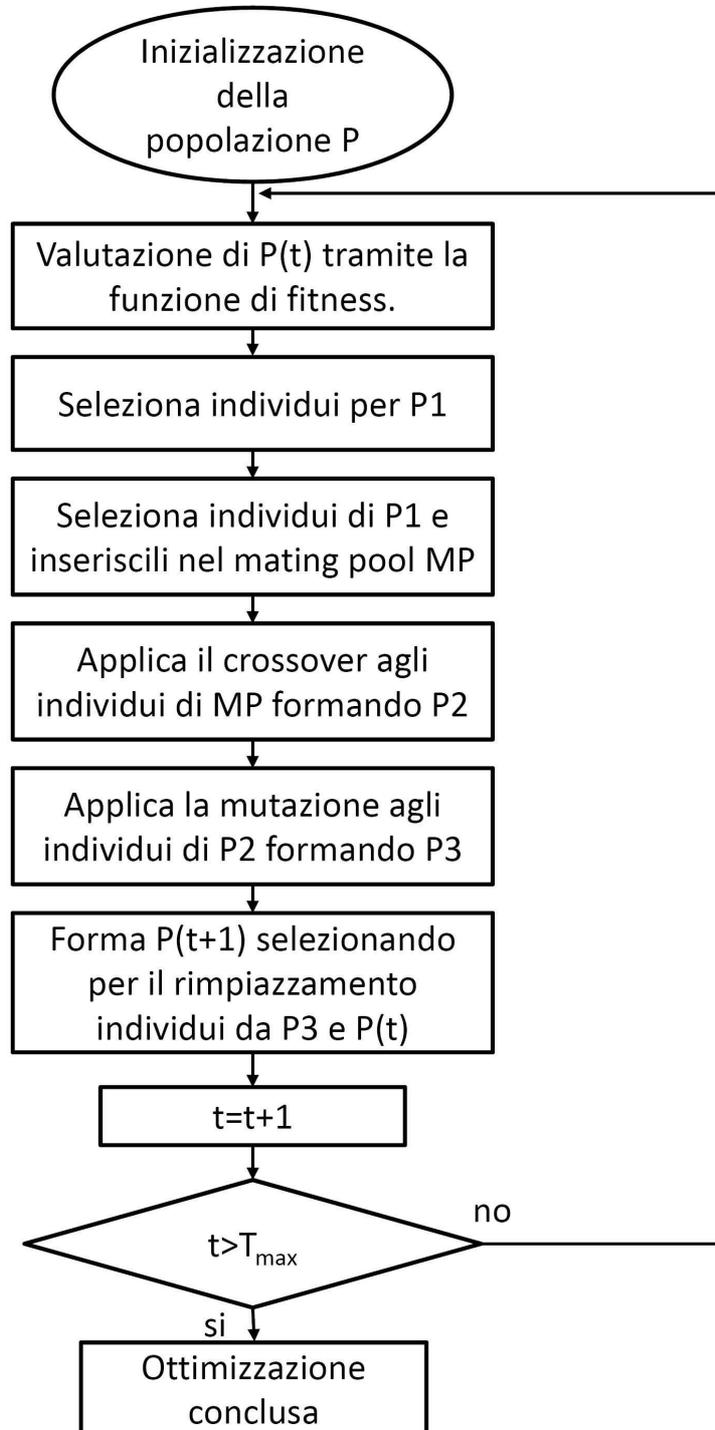


Figura 3.2: Flow-chart dell'evoluzione di una popolazione con un algoritmo genetico.

- *selezione per riproduzione*: gli individui migliori sono selezionati per la riproduzione, tali individui, che sono quelli ritenuti idonei alla riproduzione sono inseriti in una popolazione intermedia. Tra gli individui della popolazione P1 vengono selezionati quelli che compongono il *mating pool*, cioè l'insieme degli individui ai quali sarà applicato il crossover, secondo una determinata probabilità (probabilità di crossover);
- *crossover*: si applica l'operatore di crossover agli individui nel mating poll. Si ottiene una nuova popolazione intermedia P2;
- *mutazione*: l'operatore di mutazione è applicato con una certa probabilità agli individui di P2. Viene prodotta una popolazione intermedia P3;
- *selezione per rimpiazzamento e sopravvivenza*: la nuova generazione P(t+1) contiene gli individui della popolazione P3 ma può includere anche altri individui. Sono possibili più algoritmi di selezione. Ad esempio, un sottoinsieme di individui di P(t) che non sono stati selezionati per la riproduzione, oppure gli individui migliori di P(t).

L'algoritmo innescato su una popolazione di partenza si arresta dopo il passaggio di un numero prefissato n di generazioni; la soluzione al problema è il miglior individuo dell'ultima generazione. Tuttavia il meccanismo di selezione e di generazione di una popolazione tra una generazione e la successiva non tiene memoria del miglior individuo. Quindi è bene confrontare il miglior individuo dell'ultima generazione con il miglior individuo tra tutte le generazioni.

3.4.2 Strategie di selezione

Strategie di selezione per la riproduzione

La selezione per la riproduzione ha un duplice scopo: favorire la riproduzione degli individui con fitness alta e preservare la diversità della popolazione in modo da esplorare tutte le regioni dello spazio di ricerca. La strategia più usata è quella comunemente denominata con selezione proporzionale (*proportional selection*). Supponiamo che la popolazione attuale contenga n cromosomi:

$$P(t) = \{x_1, x_2, \dots, x_n\}$$

data la funzione di obiettivo f , si definisce la fitness totale F della popolazione come

$$F = \sum_{i=1}^n f(x_i).$$

La probabilità di selezione di un cromosoma quindi è quantificabile con:

$$p_i = \frac{f(x_i)}{F}.$$

Per visualizzare il meccanismo di selezione degli individui è possibile pensare all'analogia della ruota: una ruota girevole è divisa in n settori ciascuno corrispondente ad un cromosoma. La dimensione di ciascun settore è proporzionale alla probabilità di selezione p_i del cromosoma x_i . La ruota viene fatta girare n volte. Ogni volta si sceglie un cromosoma che viene copiato nella popolazione P1.

Strategie di selezione per il rimpiazzamento

La selezione per il rimpiazzamento ha lo scopo di scegliere quali, fra gli individui padri e gli individui figli, costituiranno la nuova popolazione. Negli algoritmi genetici tradizionali si assume che la generazione $P(t+1)$ sia costituita dai figli (creati per crossover e mutazione) e da quegli individui nella popolazione intermedia che non sono stati scelti per l'accoppiamento (mating) in modo da mantenere costante il numero degli individui di ciascuna generazione. Un'altra possibilità è copiare un numero fissato di individui (i migliori) da $P(t)$ a $P(t+1)$ senza modificarli. Ovviamente questi individui parteciperanno anche alla selezione per la riproduzione. Questo ultimo modo privilegia l'istanza di ricerca del minimo, a scapito dell'esplorazione dello spazio di ricerca.

3.4.3 Accoppiamento e mutazione

La procedura di accoppiamento parte, generalmente, da due individui scelti a caso nell'insieme del mating. Con l'accoppiamento due cromosomi 'genitori' saranno sostituiti da due cromosomi 'figli' i cui cromosomi sono una combinazione dei due 'genitori'. La metodologia per il crossover tra i cromosomi genitori prevede lo scambio reciproco di sottostringhe di eguale lunghezza. I cromosomi risultanti sono quelli che vanno ad individuare i figli. Per esempio si parla di *one-point crossover* quando si genera un numero casuale l compreso tra 1 e la lunghezza L delle stringhe dei cromosomi e per individuare i figli vengono scambiate le l cifre finali dei padri, come mostrato nella figura 3.3.



Figura 3.3: Esempio di one-point crossover.

Un'accoppiamento con una procedura di *two-point crossover* prevede la divisione casuale di ciascun cromosoma in tre stringhe e lo scambio secondo lo schema della figura 3.4. Questo meccanismo può essere generalizzato nella *n-point crossover*.

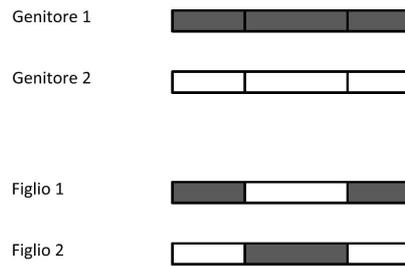


Figura 3.4: Esempio di two-point crossover.

Il processo di mutazione invece agisce su un singolo gene del cromosoma, inteso come cifra della stringa. Grazie alla mutazione possono essere considerati tutti i valori nel processo di ricerca. La mutazione avviene per cambiamento di un gene; essendo ciascun gene una cifra binaria, una mutazione opera modificando $0 \rightarrow 1$ o $1 \rightarrow 0$. Per ciascun gene di un cromosoma la scelta se operare la mutazione avviene generando un numero q con distribuzione uniforme in $[0;1]$. Se $q < P_m$ si opera la mutazione, dove P_m è una probabilità di mutazione selezionata (solitamente $P_m \in [0.001; 0.01]$).

3.4.4 Codifica reale dei GA

Per variabili continue è possibile anche una codifica reale dei cromosomi; ciascun cromosoma diventa una vettore a componenti reali. In questo caso sono possibili anche qui diversi tipi di incrocio tra i cromosomi. Si parla di *crossover-discreto* quando si fa uso di una procedura analoga a quella vista sopra per la codifica binaria. Per ogni cifra i del primo discendente si sceglie con probabilità p fissata il genitore da cui proviene quel gene; il gene corrispondente dell'altro figlio sarà fornito dall'altro genitore. Si fa anche uso del *crossover-convesso*; in questo i discendenti sono espressi come combinazione convessa di due o più genitori. Ad esempio, i cromosomi x e y possono produrre i discendenti u e v i cui geni sono:

$$\begin{cases} u_i = \alpha x_i + (1 - \alpha)y_i \\ v_i = \alpha y_i + (1 - \alpha)x_i \\ \alpha \in [0; 1] \end{cases}$$

Il crossover-convesso può essere visualizzato in uno spazio cartesiano: i figli sono due punti che giacciono sul segmento che congiunge i due punti dei genitori. I due punti figli giacciono ad una distanza α normalizzata dai due estremi del segmento.

La mutazione per la codifica reale opera come per la codifica discreta. Ciascun gene ha una probabilità di mutare e, qualora si verifica la mutazione, il gene in questione viene sostituito con un altro gene pescato casualmente nel range di variazione di quel gene.

3.5 Particle Swarm Optimization (PSO)

L'algoritmo di ottimizzazione tramite sciami di particelle è un'altro algoritmo basato sul concetto di popolazione di individui. Tale algoritmo è stato presentato per la prima volta nel 1995 da Kennedy e Eberhart in [33]. E' un algoritmo che risulta efficace nei casi in cui si applicano gli algoritmi genetici, quindi per la minimizzazione di funzioni dove non è disponibile la conoscenza del gradiente di tali funzioni. Oltre ai vantaggi propri di questi algoritmi che non richiedono la conoscenza del gradiente di queste funzioni, il PSO fa sua una notevole semplicità di implementazione. Nel 2001 van den Bergh ha pubblicato [57] dove è possibile trovare una riesamina completa, almeno fino al momento della pubblicazione del lavoro, dello stato dell'arte per quello che riguarda il PSO.

Questo algoritmo rientra nella più ampia categoria degli algoritmi evolutivi e viene spesso classificato come algoritmo sociale. Questo perchè la genesi di questo algoritmo deriva dall'osservazione del comportamento degli stormi di uccelli in volo alla ricerca di cibo.

L'algoritmo originale, così come è stato presentato in [33], parte da una popolazione, che in questo caso assume il nome di sciame, di particelle, nella quale ciascun elemento rappresenta una potenziale soluzione del problema di ottimizzazione. Sia s il numero di elementi della popolazione. Ciascun elemento è identificato dalla posizione della particella nello spazio di ricerca \mathbf{x}_i . Il movimento della particella nello spazio di ricerca è caratterizzato dalla velocità di spostamento \mathbf{v}_i ; inoltre ogni particella ha un *personal best position* \mathbf{y}_i che è la posizione in cui ha fatto registrare il valore migliore della funzione obiettivo tra tutte quelle esplorate nella ricerca del minimo. Ad ogni iterazione il personal best per ciascuna particella viene aggiornata secondo la seguente espressione:

$$\mathbf{y}_i(t+1) = \begin{cases} \mathbf{y}_i(t) & \text{se } f(\mathbf{x}_i(t+1)) \geq f(\mathbf{y}_i(t)) \\ \mathbf{x}_i(t+1) & \text{se } f(\mathbf{x}_i(t+1)) < f(\mathbf{y}_i(t)) \end{cases}$$

dove f è la funzione obiettivo.

Quindi la posizione del personal best di ciascuna particella viene continuamente aggiornate durante il moto. Viene inoltre definito, ad ogni iterazione la *global best position*, cioè la posizione tra tutti i personal best in cui è stata registrato il valore migliore fino a quel momento della funzione obiettivo. Perciò tale posizione $\hat{\mathbf{y}}$, istante per istante è definita da:

$$\hat{\mathbf{y}}(t) \in \{\mathbf{y}_0(t), \mathbf{y}_1(t), \dots, \mathbf{y}_s(t)\} | f(\hat{\mathbf{y}}(t)) = \min\{f(\mathbf{y}_0(t)), f(\mathbf{y}_1(t)), \dots, f(\mathbf{y}_s(t))\}.$$

Ad ogni iterazione la posizione di ciascuna particella viene aggiornata tramite la velocità secondo la formula

$$\mathbf{x}_i(t+1) = \mathbf{x}_i(t) + \mathbf{v}_i(t+1) \tag{3.1}$$

quindi la direzione di ricerca di ciascuna particella viene rappresentata dalla propria velocità. L'osservazione del movimento in volo degli uccelli ha portato Kennedy e Eberhart

a notare che la direzione in volo di ciascun uccello in uno stormo alla ricerca di cibo è influenzata dalla posizione dove ciascun uccello ha trovato la maggior quantità di cibo nella sua ricerca e dalla posizione dove tutto lo stormo ha individuato la quantità maggiore di cibo in assoluto. Identificando il cibo come il valore della funzione obiettivo è possibile quindi rappresentare la velocità degli uccelli come un vettore \mathbf{v} la cui accelerazione è composta di due termini, uno proporzionale alla distanza di ciascuna particella dal proprio personal best e l'altro proporzionale alla distanza di ciascuna particella dal global best. Di fatto, specificando con $j \in 1 \dots n$ ciascuna dimensione, in modo tale che $v_{i,j}$ denota la j -esima componente della velocità associata alla i -esima particella, l'espressione che aggiorna la velocità ad ogni iterazione dell'algoritmo di ottimizzazione è :

$$v_{i,j}(t+1) = v_{i,j}(t) + c_1 r_{1,j}(t)[y_{i,j}(t) - x_{i,j}(t)] + c_2 r_{2,j}(t)[\hat{y}_j(t) - x_{i,j}(t)].$$

In questa espressione si fa uso di $r_1 \in [0, 1]$ e $r_2 \in [0, 1]$, che sono due vettori casuali scalati delle costanti $0 \leq c_1, c_2 \leq 2$. Tali costanti sono detti anche *coefficienti di accelerazione*. In sostanza, come detto prima la velocità di ciascuna particella all'istante $(t+1)$ è funzione della velocità all'istante precedente e di due termini il primo proporzionale alla distanza tra la posizione attuale della particella e il personal best, il secondo proporzionale alla distanza tra la posizione attuale e il global best. Ciascun termine è soppesato con i coefficienti c_1, c_2 che consentono all'utilizzatore di far prevalere o l'uno o l'altro termine per favorire la convergenza a seconda del caso in analisi. I coefficienti r_1, r_2 sono di natura casuale e sono volti a favorire l'istanza di *exploration* nell'algoritmo.

Generalmente il valore di \mathbf{v} è limitato all'interno di un intervallo $[-\mathbf{v}_{\max}, \mathbf{v}_{\max}]$ per ridurre la probabilità che la particella sfugga dallo spazio di ricerca e o per impedire alle particelle di convergere troppo velocemente rischiando di rimanere intrappolate in un minimo locale. Nel caso lo spazio di ricerca sia definito entro dei limiti quindi dentro a $[-\mathbf{x}_{\max}, \mathbf{x}_{\max}]$ allora si può porre $\mathbf{v}_{\max} = k \times \mathbf{x}_{\max}$ con $0.1 < k < 1$ (vedi [45]).

3.5.1 PSO modificati

Sono state apportate delle modifiche successive al metodo così come è stato proposto inizialmente dagli autori al fine di migliorare le prestazioni in fase di ricerca del minimo. Una delle modifiche che ha portato a migliorare notevolmente la convergenza del metodo è l'introduzione dell'inerzia ω dovuta a Shi e Eberhart in [34]. Questo è un fattore associato al termine della velocità nell'iterazione precedente nell'espressione di aggiornamento della velocità ad ogni iterazione:

$$v_{i,j}(t+1) = \omega v_{i,j}(t) + c_1 r_{1,j}(t)[y_{i,j}(t) - x_{i,j}(t)] + c_2 r_{2,j}(t)[\hat{y}_j(t) - x_{i,j}(t)]. \quad (3.2)$$

Il valore che massimizza le prestazioni dell'algoritmo nella ricerca del minimo dipende dal caso in analisi. Spesse volte è necessario provare varie combinazioni di coefficienti ω, c_1, c_2 prima di individuare la combinazione corretta. Qualora questo non fosse possibile, studi

empirici mostrano comunque che un valore che si colloca nell'intervallo $[0.8,1.2]$ garantisce buoni risultati.

In [57] l'autore propone un'ulteriore modifica al PSO modificato con l'inerzia che prende il nome di *Guaranteed Convergence Particle Swarm Optimizer (GCPSO)*. L'autore parte dall'osservazione che tutti i metodi proposti fino a quel momento hanno una proprietà rischiosa per la convergenza dell'algoritmo: se $\mathbf{x}_i = \mathbf{y}_i = \hat{\mathbf{y}}$, allora l'espressione che aggiorna la velocità del metodo dipende solo dal valore di $\omega v_{i,j}(t)$. Quindi se la posizione di una particella coincide con il proprio personal best e con il global best allora la particella si sposterà da quella posizione solo se sia ω che $v_{i,j}(t)$ sono diversi da zero. Questo implica che se la velocità delle varie particelle che compongono lo sciame è vicina a zero, allora le particelle si attesteranno una volta che la loro posizione coinciderà con il global best; un arresto di questo tipo sarebbe prematuro, perchè non garantisce che il best trovato fino a quell'iterazione sia effettivamente il minimo globale della funzione e non un minimo locale in cui è rimasto intrappolato lo sciame.

Per ovviare a questo difetto viene introdotto un nuovo parametro nello schema dell'algoritmo. Sia τ l'indice della particella che individua il global best dello sciame, cioè per la quale vale l'uguaglianza $\mathbf{y}_\tau = \hat{\mathbf{y}}$. Per questa particella, e solo per questa, viene proposta una nuova equazione per l'aggiornamento della velocità che diventa:

$$v_{i,j}(t+1) = -x_{\tau,j}(t) + \hat{y}_j(t) + \omega v_{\tau,j}(t) + \rho(t)(1 - 2r_{2,j}(t)) \quad (3.3)$$

dove ρ è un fattore definito in base al numero di successi e al numero di insuccessi nella ricerca di un nuovo global best da parte dell'algoritmo; tale fattore è definito come:

$$\rho(t+1) = \begin{cases} 2\rho(t) & \text{se } N_{succ} > s_c \\ 0.5\rho(t) & \text{se } N_{insucc} > f_c \\ \rho(t) & \text{altrimenti} \end{cases}$$

dove N_{succ} e N_{insucc} denota il numero di successi e di insuccessi consecutivi nella ricerca del minimo, con un insuccesso definito come $f(\hat{\mathbf{y}}(t)) = f(\hat{\mathbf{y}}(t))$, ponendo di default $\rho(0) = 1$. Quindi nell'algoritmo devono essere implementati anche i seguenti check per rendere effettivo il metodo:

$$\begin{cases} N_{succ}(t+1) > N_{succ}(t) \implies N_{insucc}(t+1) = 0 \\ N_{insucc}(t+1) > N_{insucc}(t) \implies N_{succ}(t+1) = 0 \end{cases}$$

in modo tale che al verificarsi di un successo il contatore degli insuccessi viene azzerato e viceversa.

Con questa nuova implementazione del metodo la particelle dello sciame continuano ad utilizzare l'espressione di aggiornamento della velocità nell'espressione 3.2, a parte la particella che individua il global best per la quale si fa uso della nuova espressione in 3.3. Sostituendo tale espressione nella 3.1 il termine $-x_{\tau,j}(t)$ annulla la posizione della particella a $\hat{\mathbf{y}}$. A questa posizione un vettore che rappresenta la direzione di ricerca corrente è aggiunto ($\omega v_{\tau,j}(t)$), e il termine $\rho(t)(1 - 2r_{2,j}(t))$ rappresenta uno spostamento casuale nello spazio di indagine con lunghezza massima $2\rho(t)$; in pratica tale termine produce uno spostamento casuale dello

sciame in un'area intorno alla posizione del global best, il cui diametro è controllato dal parametro ρ . La nuova posizione della particella identificata con τ è pari a:

$$x_{\tau,j}(t+1) = \hat{y}_j(t) + \omega v_{\tau,j}(t) + \rho(t)(1 - 2r_2(t)).$$

La scelta ottimale dei valori per i parametri di soglia f_c e s_c dipende dalla funzione obiettivo. L'autore suggerisce per spazi multidimensionali $f_c = 5$, $s_c = 15$. Queste impostazioni implicano che l'algoritmo incrementa velocemente un valore basso di ρ piuttosto che diminuirne un valore alto; questa strategia, che empiricamente ha dato buoni frutti, sembra adatta a favorire l'aspetto esplorativo dell'algoritmo.

3.5.2 PSO vincolato

Nelle PSO è stato possibile applicare anche un'altro tipo di costrizione oltre a quello già visto in all'interno dello spazio di indagine definito dai limiti degli intervalli di variazione delle variabili $[\mathbf{x}_{\min}, \mathbf{x}_{\max}]$. Tale metodologia consiste nell'applicare le leggi della riflessione al movimento della particella. Questa metodologia prende spunto dall'idea delle particelle come di corpi in movimento lungo una direzione nello spazio di indagine. Qualora l'aggiornamento della posizione secondo l'espressione 3.1 porti ad una posizione della particella, esterna ai limiti è possibile applicare le leggi della riflessione al moto di tale particella; una rappresentazione grafica come quella di figura 3.6 rende intuitiva la riflessione: immaginando la particella come una pallina e i limiti imposti dello spazio di indagine come una parete, la pallina non supera la parete ma vi rimbalza.

L'applicazione del metodo di riflessione nell'algoritmo della PSO si configura come un check da effettuare per ciascuna particella dello sciame dopo la fase di aggiornamento della posizione. Se la posizione della particella j risulta dentro i limiti dello spazio di ricerca allora la posizione e la velocità della particella rimangono inalterati. Se una o più componenti della posizione della particella risultano fuori dai rispettivi limiti vanno applicate le leggi della riflessione a tali particelle per la posizione e la velocità:

$$x_{i,j}(t+1) = \begin{cases} x_{i,j}(t+1) & \text{se } x_{i,\min} \leq x_{i,j}(t+1) \leq x_{i,\max} \\ 2x_{i,\max} - x_{i,j}(t+1) & \text{se } x_{i,j}(t+1) > x_{i,\max} \\ 2x_{i,\min} - x_{i,j}(t+1) & \text{se } x_{i,j}(t+1) < x_{i,\min} \end{cases}$$

$$v_{i,j}(t+1) = \begin{cases} v_{i,j}(t+1) & \text{se } x_{i,\min} \leq x_{i,j}(t+1) \leq x_{i,\max} \\ v_{i,j}(t+1) & \text{altrimenti} \end{cases}$$

3.5.3 Condizione di arresto

Nella simulated annealing la condizione di arresto del metodo iterativo è data dal raggiungimento di una determinata temperatura; con gli algoritmi genetici è stabilito a priori il numero di generazioni da far nascere prima di arrestare il metodo. In un certo qual modo si può dire che entrambi i metodi sono esplorativi, cioè si lascia che l'algoritmo esplori lo spazio di

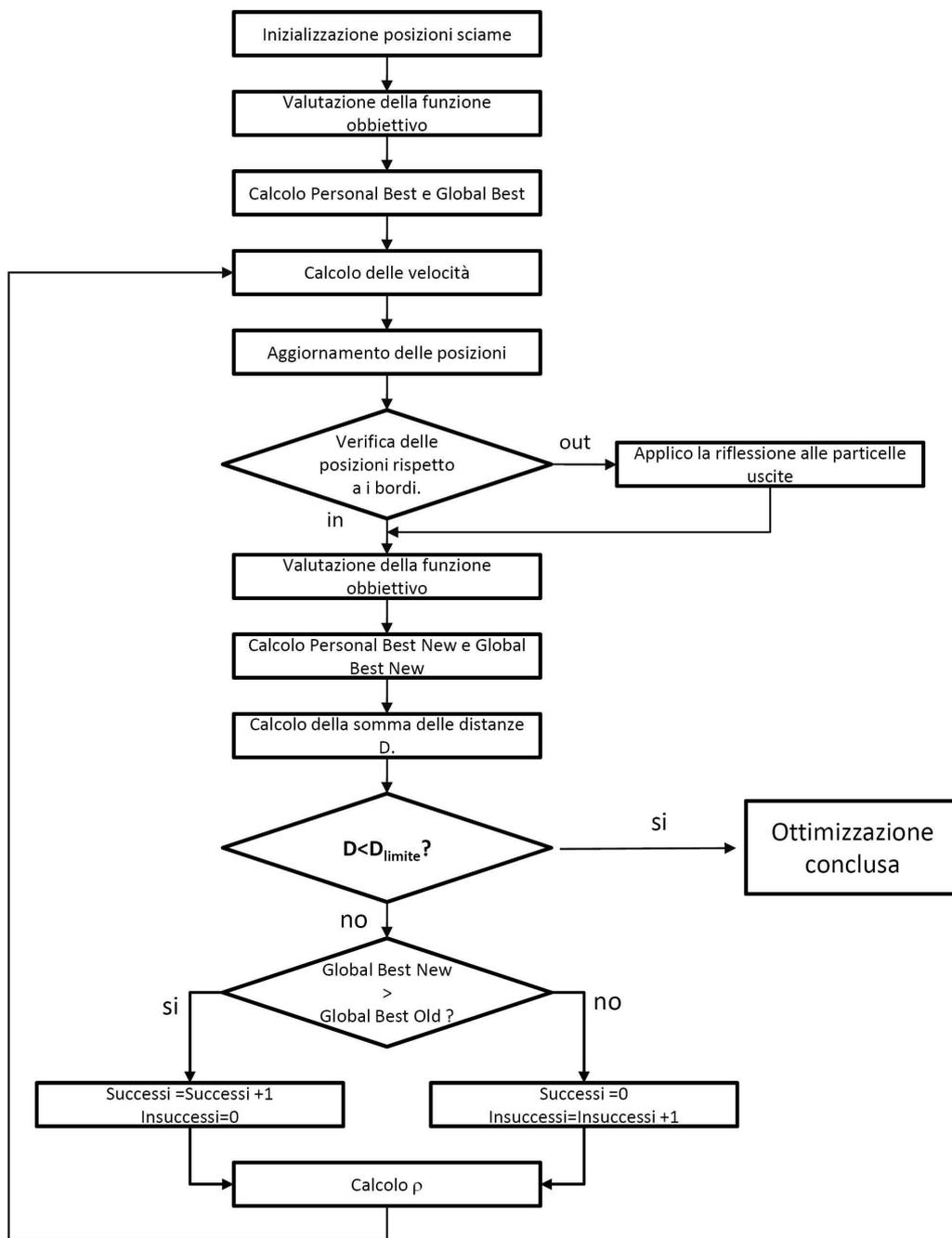


Figura 3.5: Flow-chart della ricerca dell'ottimo tramite PSO.

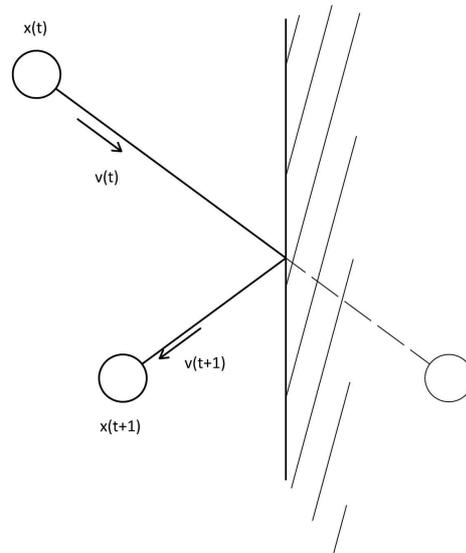


Figura 3.6: Applicazione del vincolo nell'ottimizzazione per mezzo della riflessione: esempio 2D.

indagine, si arresta il metodo dopo alcune iterazioni e si individua l'ottimo come la miglior soluzione finora rintracciata.

Nell'ottimizzazione per sciami di particelle la convergenza è raggiunta quando le particelle hanno individuato il minimo e si sono accumulate intorno ad esso. Quindi è necessario definire una condizione di arresto, che indichi l'avvenuta convergenza del metodo all'intorno del minimo della funzione obiettivo. Tale criterio deve essere in grado di individuare quando lo sciame di particelle ha raggiunto tale minimo quindi quando le particelle si sono concentrate sopra un certo punto.

Il metodo scelto è quello di imporre un limite sulla somma delle distanze tra le varie particelle dello sciame. Quindi dato un sciame di s particelle e sia d_{ij} la distanza tra due generiche particelle i e j , la somma delle distanze a cui fare riferimento per definire la densità di uno sciame di particelle è pari a:

$$D = \sum_{i=1}^s \sum_{j=i+1}^s d_{ij}.$$

In figura 3.7 si può vedere una schematizzazione bidimensionale dello sciame di particelle, e come, convergendo intorno alla posizione in cui si registra il minimo la somma tra le distanze delle varie particelle cali.

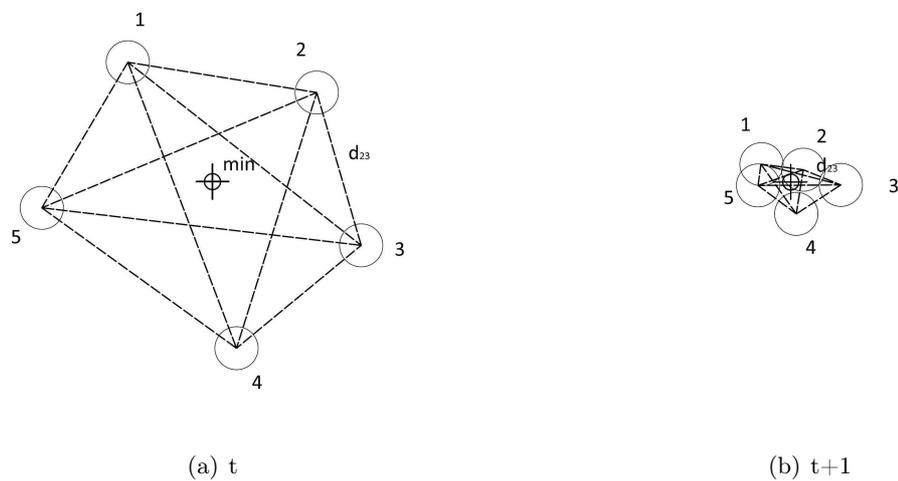


Figura 3.7: Sciame bidimensionale di particelle intorno ad un minimo per iterazioni successive: diminuzione delle distanze reciproche tra le particelle.

Capitolo 4

DOE e modelli di regressione

Come visto nel capitolo 3 gli algoritmi di ottimizzazione nella pratica ingegneristica spesso sono affiancati da tecniche di Design of Experiments e da metamodelli al fine di ridurre lo sforzo computazionale da affrontare. In questo capitolo si presentano le principali tecniche di DOE e i principali metamodelli che sono solitamente utilizzati insieme alle simulazioni al computer.

4.1 Design of Experiments

Con *Design of Experiments* si intendono delle tecniche per la selezione delle prove da fare, nell'indagine sperimentale per massimizzare le informazioni che vengono raccolte mantenendo costante il numero di prove che devono essere realizzate, si veda [38]. Tali tecniche sono applicate anche alle simulazioni al computer quando è necessario pianificare più simulazioni da effettuare per indagare l'influenza di determinati parametri. Le simulazioni al computer vengono considerate come prove sperimentali e infatti non a caso si parla di *computer experiments*; le tecniche di DOE applicate ai computer experiments si differenziano dalle tecniche applicate agli esperimenti di laboratorio veri e propri. Infatti le tecniche di DOE si basano sulla natura degli esperimenti che si vanno a pianificare, quindi per gli esperimenti di laboratorio si basano sull'assunzione che si va ad indagare fenomeni stocastici, mentre le simulazioni al computer sono essenzialmente deterministiche¹.

La scelta delle simulazioni da eseguire è alla base della realizzazione di un metamodello che approssima l'influenza di determinati parametri su determinate grandezze. Nel caso della progettazione delle turbomacchine i metamodelli servono per approssimare l'influenza dei parametri che definiscono la geometria sulle prestazioni. A questo scopo l'algoritmo che prende il nome di Ipercubo Latino (Latin Hypercube Sampling-LHS), citato per la prima volta da McKay et al. in [7], e gli algoritmi da questo derivato, si sono dimostrati più efficaci di un campionamento casuale. Sono facili da generare anche con un numero considerevole di variabili di input e garantiscono che ciascuna porzione in cui è divisa ciascuna variabile di input è rappresentata in modo pari alle altre (nella fattispecie una e una sola volta).

¹Con *deterministiche* si intende che una simulazione al computer, impostata con dei parametri precisi fornirà sempre lo stesso risultato tutte le volte che viene svolta

Dal punto di vista matematico la generazione dei dati di un insieme di campionamento grazie al LHS è piuttosto semplice. Dato uno spazio di scelta di p variabili, ciascuna con il proprio intervallo di variazione I_p , si scelga la numerosità n del campione che si vuole andare ad analizzare. A questo punto è necessario dividere ciascun intervallo I_p in n sottointervalli ciascuno di probabilità costante $1/n$ e si contrassegna ciascun sottointervallo con il corrispondente numero progressivo da 1 a n . Un ipercubo latino non è altro che una matrice $n \times p$, costituita da p colonne ciascuna contenente una permutazione casuale dell'insieme dei numeri naturali $\{1 \dots n\}$. Ciascuna delle n righe contiene l'indicazione degli intervalli delle varie variabili che competono a ciascuno degli n punti dell'insieme di campionamento. Ciascun punto è scelto all'interno di questi intervalli, solitamente posizionato al centro di un intervallo.

In presenza di variabili continue a probabilità uniforme, è possibile ragionare in termini discreti: un ipercubo latino discreto è una matrice $n \times m$ di n punti di campionamento per m variabili. Ciascuna colonna di questa matrice è una permutazione dell'insieme $1 \dots n$. Ogni riga rappresenta, anche qui, la posizione di ciascun punto dell'insieme di campionamento; semplicemente se un punto per l' i -esima variabile ha indicato il numero p , con $1 \leq p \leq n$, allora si intende per quel punto che l' i -esima variabile assume come valore il p -esimo tra n valori equidistanti in cui è diviso l'intervallo I_i . Ciascuna variabile è così discretizzata in n valori equidistanziati.

In figura 4.1 è presentato un esempio di LHS, su uno spazio bidimensionale per un campione di numerosità 10.

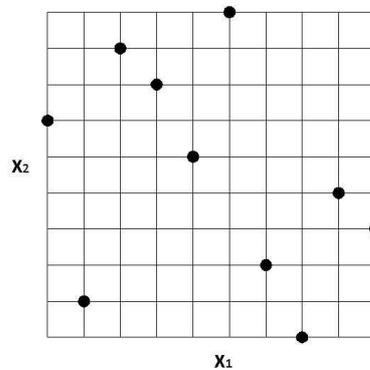


Figura 4.1: Esempio di ipercubo latino.

Le caratteristiche principali per cui l'ipercubo latino risulta efficace nella pianificazione delle simulazioni al computer, o più in generale degli esperimenti deterministici sono che la generazione del campione è casuale, ogni porzione in cui è divisa ciascuna variabile è rappresentata in modo pari alle altre e non ci sono punti punti dell'insieme di campionamento ripetuti.

Tuttavia un campionamento effettuato con l'ipercubo latino classico non è in grado di garantire che ciascuna porzione dello spazio di indagine sia coperta in modo pari alle altre porzioni; a parità di numerosità del campione l'algoritmo è in grado di generare campionamenti molto differenti in quanto a uniformità del campione. Per questo nel corso degli anni sono stati presentati nuovi tipi di ipercubo ottimi che cercano, massimizzando la distanza minime

tra i punti dell'ipercubo, di incrementare, a parità di punti di campionamento, l'efficacia della capacità di rilevazione dell'andamento della grandezza risposta sullo spazio di indagine (vedi [8]).

Parimenti sono state presentate e verificate altre tecniche di campionamento che affiancano, alla generazione classica di un LHS, la tecnica delle matrici ortogonali (orthogonal array), proposta da Leary, Bashkar e Keane, vedi [9]. Questa garantisce la generazione di un ipercubo molto più regolare che si distribuisce meglio sullo spazio di indagine. Lo svantaggio introdotto da questa tecnica è la complicazione che si aggiunge all'algoritmo per generare l'insieme di campionamento.

4.1.1 Curse of dimensionality

In questa sezione dove si parla del campionamento dei punti dell'insieme di dati noti sui quali addestrare una tecnica di regressione, è opportuno citare anche di quello che prende il nome di *curse of dimensionality*, letteralmente la 'condanna delle dimensioni'.

Tale espressione fu introdotta per la prima volta da Bellman in [39] per riferirsi alla crescita esponenziale in volume di uno spazio di indagine associata all'incremento delle dimensioni. Questo, in altre parole, significa che, considerando la distanza euclidea tra due punti di input la grandezza che misura la densità di punti in un insieme e quindi indirettamente la precisione con cui è descritto lo spazio di input, al crescere del numero di variabili dello spazio di input, il numero dei punti necessari per mantenere la stessa densità dei punti cresce esponenzialmente.

La curse of dimensionality è un'ostacolo rilevante nelle operazioni di ottimizzazione o di regressione, soprattutto quanto si opera su range molto ampi per un elevato numero di variabili. In caso di aumento del numero di variabili dello spazio di input il numero di simulazioni necessario per garantire una precisione costante al modello di regressione cresce esponenzialmente con il numero di variabili.

In [40] Friedman sintetizza gli effetti del curse of dimensionality sulla complessità delle funzioni di regressione. Egli afferma che Una funzione definita in uno spazio caratterizzato da un alto numero di dimensioni è molto più complessa di una definita in uno spazio con un numero scarso di dimensioni e la complessità è di difficile previsione.

In questi casi, qualora fosse disponibile, una conoscenza a priori della funzione da rappresentare consente di limitare gli svantaggi; in caso contrario un DOE efficace diventata fondamentale.

4.2 Modelli di regressione

Sempre di più negli ultimi anni i modelli di regressione hanno messo in luce la loro efficacia e la loro duttilità nei problemi di ottimizzazione, soprattutto congiuntamente con i computer experiments. Di seguito si cercherà di argomentare perchè è necessario in molte circostanze ricorrere a modelli di regressione. Le motivazioni che portano a fare uso delle tecniche di regressione molte volte all'interno dei problemi di ottimizzazione che coinvolgono

la progettazione delle turbomacchine valgono per problemi analoghi anche in altri campi di applicazione.

La progettazione di un componente meccanico prevede l'assegnazione di molteplici parametri che ne definiscono la geometria e le caratteristiche. Al variare di questi parametri vengono controllate una o più grandezze che definiscono o rappresentano il funzionamento di tali componenti. Quasi sempre la relazione che intercorre tra i parametri assegnati e le grandezze di controllo non sono semplici; molto raramente seguono andamenti lineari e per range di variazione molto limitati. Inoltre il numero di parametri che devono essere assegnati contemporaneamente è alto e questo si riflette ulteriormente sulla complessità della funzione.

Creare una funzione con gli strumenti matematici dell'analisi, che metta in relazione i parametri alle grandezze di controllo risulta quindi complicato. Spesso si ricorre alla sperimentazione per ricavare tali relazioni. Più recentemente i software di simulazione nei vari campi consentono una sorta di sperimentazione virtuale che consente di individuare la relazione che intercorre tra un componente meccanico e il suo funzionamento.

I dati ricavati dalla sperimentazione, fisica o virtuale, spesso devono essere immagazzinati dentro un modello matematico, perchè sotto tale forma sono più semplici da maneggiare ed è possibile estendere tali risultati anche per altre zone. La complessità dei fenomeni fisici in gioco, il numero di parametri che di volta in volta vengono utilizzati per definire il componente, la non linearità del funzionamento del componente al variare dei parametri di definizione rendono impossibile specificare tramite ragionamento il modello matematico che seguono tali relazioni. Non è quindi possibile ricavare le relazioni che intercorrono tra i parametri tarando i coefficienti di modelli matematici già noti.

I metodi di regressione di cui si deve fare uso per rappresentare il funzionamento di questo componente, ma in generale per rappresentare il funzionamento di qualunque componente industriale, devono essere assolutamente generici. Questi metodi possono essere utilizzati indipendentemente dalla conoscenza del modello matematico che segue il fenomeno o la funzione che si sta approssimando. In pratica, l'utilizzo di questi metodi deve essere possibile come *black-box*: dato un insieme di casi sperimentali che sono stati analizzati per indagare un determinato fenomeno, è possibile utilizzare questi dati per istruire uno o più di questi modelli e rappresentare la funzione che approssima questi dati, qualunque essa sia.

In altre parole questo significa anche poter approssimare una o più grandezze y_i dipendenti da una serie di parametri x_j senza conoscere la funzione matematica che intercorre tra x e y ; non è quindi necessario fare alcuna ipotesi preliminare sulle leggi che possono seguire i fenomeni che si stanno indagando, come per esempio succede con l'approssimazione ai minimi quadrati.

I vantaggi di modelli di questo tipo sono evidenti. A prescindere dal fenomeno dal quale derivano i dati che sono stati raccolti, sia esso un fenomeno fisico isolato, oppure il comportamento di un complesso componente meccanico, oppure una raccolta di dati storici di collaudo, è possibile utilizzare uno di questi metodi di regressione addestrati con questo insieme di dati, e ricavare una approssimazione di questi dati.

Tra le più utilizzate ci sono delle funzioni polinomiali di secondo grado approssimate grazie

al metodo dei minimi quadrati sull'insieme dei punti di addestramento, vedi [36], che prendono anche il nome di Response Surfaces (RS). Oltre a questo sono state utilizzate estesamente le Reti Neurali Artificiali e le Reti Funzioni a Base Radiale Neurali Radiali, ANN e RBFN, rispettivamente. Altro metodo di regressione è quello che prende il nome di Kriging. Si tratta di una tecnica di regressione su base statistica, presentata per la prima volta insieme ai computer experiments da Sacks e altri nel 1989, (vedi [2]).

Tuttavia queste tecniche presentano anche svantaggi che è bene tenere presente in fase di utilizzo. Innanzitutto l'estrema complessità matematica di questi modelli, a parte le RS. Tale complessità rende difficile l'implementazione di questi metodi ad un utente inesperto. È necessario, quindi, se si vuole implementare uno di questi metodi, un approfondito studio della letteratura concernente ciascun metodo. In alternativa è possibile fare uso dei numerosi pacchetti, anche free, che sono disponibili oggi su internet.

Tali metodi comunque richiedono esperienza e studio per essere utilizzati. Le strutture con cui possono essere utilizzati tali metodi di volta in volta sono molteplici e le prestazioni cambiano sensibilmente in base alla struttura scelta. Non esistono regole per la scelta della struttura più adatta, ma tale scelta va fatta caso per caso in base all'esperienza dell'utilizzatore e spesso è necessario testare più strutture differenti per individuare quella che meglio approssima un determinato insieme di dati.

4.2.1 Addestramento

Generalmente il processo attraverso il quale vengono modificati i parametri interni di un modello di regressione, affinché i valori predetti da questi modelli siano più aderenti possibile ai dati che sono stati analizzati, prende il nome generico di *addestramento*. In altri casi si parla di *istruire* un modello, una rete neurale o un quant'altro, con una serie di dati noti in modo che tale modello possa immagazzinare le informazioni contenute in questo insieme di dati e possa usarle per estrapolare i valori della funzione.

Ciascuna tecnica di regressione ha i propri algoritmi di addestramento. Alcune tecniche come le Response Surface fanno uso di metodi ormai collaudati come la tecnica dei minimi quadrati. Altre tecniche come le reti neurali hanno a disposizione algoritmi molto diversi tra loro. Sta all'esperienza di chi che utilizza queste tecniche di regressione la scelta della tecnica che meglio funziona per un determinato caso in analisi. Nella maggior parte dei casi l'addestramento è un processo iterativo che si arresta quando lo scarto sui valori dei pesi ad iterazioni successive o l'errore di regressione compiuto dal modello è inferiore ad un valore di soglia.

Gli algoritmi di addestramento costituiscono sempre la parte matematicamente più complessa delle tecniche di regressione e richiedono, per essere implementati conoscenze specifiche in materia e uno studio approfondito. Tuttavia esistono oggi pacchetti software a disposizione di chi vuole utilizzare queste tecniche che consentono anche all'utente la scelta tra vari algoritmi di addestramento. I vari algoritmi possono essere tra loro classificati in base a due criteri prestazionali che consentono quindi di effettuare una scelta ragionata di quello più indicato per il caso che si sceglie di esaminare. Le prestazioni che devono essere tenute in

considerazione in fase di scelta dell'algoritmo di addestramento sono il tempo necessario per effettuare l'addestramento e la precisione che si riesce a raggiungere con l'approssimazione, sia in termini di errore medio sull'approssimazione degli input noti, sia in termini di regolarità della funzione.

4.2.2 Stima dell'accuratezza dell'approssimazione e confronto tra modelli di regressione

Un'esigenza fondamentale quando si utilizzano le tecniche di regressione è la possibilità di misurare l'accuratezza dell'approssimazione di un insieme di dati, operata con un determinato modello. Oltre alla fondamentale esigenza di avere una stima delle errore che si commette in sede di estrapolazione, una misura del genera potrebbe fornire indicazioni sull'opportunità di infittire, qualora ce ne fosse la possibilità, l'insieme di dati noti per ottenere una regressione migliore. Alternativamente potrebbe essere necessario confrontare l'efficienza di modelli differenti: quindi confrontare tecniche diverse, ma anche confrontare modelli diversi proposti con la stessa tecnica, come potrebbe essere per esempio confrontare l'efficacia di un modello polinomiale di terzo grado con uno polinomiale quadratico, o di una rete neurale con una struttura, piuttosto che con un'altra. Infatti, nonostante la creazione di modelli di regressione sia un'operazione dispendiosa, infittire ulteriormente l'insieme di dati noti con un ulteriore campionamento è molto più oneroso. Risulta quindi più conveniente testare e confrontare diversi metodi di regressione in modo da individuare quello che meglio si adatta all'insieme di dati in questione. Oppure, ancora, un'ulteriore esigenza è poter individuare errore e varianza di dati estrapolati con un metodo di regressione per valutare l'attendibilità dei dati ricavati con i metodi di regressione e calcolare intervalli di confidenza.

L'errore dell'approssimazione commesso da un modello di regressione è definito come:

$$E_{gen} = \lim_{M \rightarrow \infty} \frac{\sum_{i=1}^M (F(x_i) - f(x_i))^2}{M}$$

dove F è il modello di regressione addestrato, f è la funzione che si va ad approssimare, \mathbf{x} è il vettore di input. Risulta quindi che $f(x_i)$ è l'output atteso per l'input x_i (il valore vero), mentre $F(x_i)$ è il valore stimato per quello stesso input dal modello di regressione. In accordo con la definizione appena presentata l'errore di approssimazione altro non è che l'errore quadratico medio su un insieme infinito di punti di cui si conosce il valore della funzione da approssimare, come dire una insieme infinito di dati noti. Ovviamente tale insieme non è disponibile. Quindi per valutare l'errore che si commette in fase di regressione è necessario ricorrere a delle stime su degli insiemi di numerosità finita.

In questo contesto, uno strumento standard utilizzato in statistica è la divisione dell'insieme di dati noti (*split-sampling* o *holdout*) in due sottoinsiemi: il *training set* e il *validation set*, generalmente il primo circa i 2/3 dell'insieme completo e il secondo l'1/3 restante. Il primo, il training set, serve per il calcolo dei parametri del modello. L'altro prende il nome di validation set cioè insieme di controllo. I dati appartenenti all'insieme di controllo non vengono utilizzati nella creazione del modello, quindi le informazione di questo insieme di dati

non sono trasmesse al modello. Calcolando l'errore medio dei dati stimati dalla regressione rispetto ai dati noti dell'insieme di controllo, e la varianza dell'errore, si ottiene una stima dell'errore dell'approssimazione.

Un altro metodo statistico in uso nella stima dell'accuratezza dell'approssimazione di un modello è quello che prende il nome di *cross-validation*. Tale metodo prevede la divisione dell'intero insieme di dati noti in k sottoinsiemi di uguale numerosità. Nella *cross-validation* completa il modello viene addestrato k volte, utilizzando a rotazione ciascuno dei k sottoinsiemi come *validation set* e gli altri $k-1$ come *training set*. La stima dell'errore e della varianza dell'errore dell'approssimazione viene calcolata come media degli errori ricavati dai vari addestramenti.

L'osservazione dei due metodi fin qui presentati mette in luce le analogie tra i due. Infatti è chiaro come il sistema *holdout* può essere visto come un caso particolare di *cross-validation* nel quale la stima dell'errore è fatta su un solo sottoinsieme, senza la ripetizione degli addestramenti del modello e la verifica degli stessi. Infatti a volte per il primo metodo si parla anche di *hold-out cross-validation* mentre la seconda viene specificata con *multi-fold cross-validation* (vedi [37]). Con il metodo *cross-validation* ciascun elemento dell'insieme di dati noti viene utilizzato in un insieme di controllo garantendo così una stima dell'accuratezza della precisione pressoché senza errore e una minore varianza. Lo svantaggio di questa tecnica è che prevede l'addestramento di k modelli di regressione.

Viceversa il metodo *holdout* fornisce una stima dell'errore con una forte varianza; infatti l'errore dipende fortemente dalla scelta di quali elementi compongono l'insieme di addestramento e quali compongono l'insieme di controllo. Se, per assurdo, tutti gli elementi che contengono una determinata informazione finissero nell'insieme di controllo il metodo di regressione, addestrato sugli altri punti, non conserverebbe quell'informazione e quindi non sarebbe in grado di prevederla in fase di estrapolazione della funzione sui punti dell'insieme di test. Quello che si può affermare è che in fase di regressione di insiemi sufficientemente numerosi di punti, il metodo *holdout* garantisce una stima abbastanza accurata dell'errore (anche se sempre più variabile della *cross-validation*); viceversa in caso di regressione di pool di dati scarsamente numerosi, è giocoforza utilizzare il metodo della *cross-validation*.

Portare all'estremo la *cross-validation* su un insieme di dati noti di numerosità n , vuol dire stimare l'errore con una *n-fold cross-validation*. Questo caso particolare, che prende il nome di *leave-one-out cross-validation*, prevede di addestrare il modello di regressione n volte, tenendo di volta in volta un singolo elemento come insieme di controllo e utilizzare gli altri $n-1$ per l'addestramento. Tale tecnica è indicata soprattutto per insiemi con numerosità particolarmente scarsa. Efron in [3] ha dimostrato come l'approccio *leave-one-out* garantisce delle stime senza errore dell'accuratezza della regressione, ma spesso con varianze inaccettabili. Inoltre Breiman e Spector in [4], hanno effettuato una serie di esperimenti di scelta di modelli di regressione e hanno mostrato come per questa disciplina l'approccio *leave-one-out* sia inadeguato.

Altro metodo statistico per stimare l'errore commesso in fase di generalizzazione è quello che prende il nome di *bootstrap*, letteralmente 'laccio per stivali', presentato per la prima volta

in [3]. Con il bootstrap l'errore di approssimazione non è stimato direttamente. Tale metodo parte dalla stima della differenza tra l'errore di generalizzazione e l'errore di addestramento, quest'ultimo anche detto errore apparente, inteso come l'errore commesso da un modello utilizzando lo stesso insieme come training set e validation set contemporaneamente. Tale differenza prende il nome di *ottimismo*. L'errore di generalizzazione quindi può essere stimato come la somma tra l'errore di addestramento calcolato utilizzando tutto l'insieme di dati noti e la stima dell'ottimismo. L'ottimismo è stimato, invece, a partire da una serie di campioni selezionati con ripetizione² dall'insieme di addestramento.

Si indichi con $E_j^{A_l A_v}$ l'errore commesso dal modello addestrato con l'insieme A_l e calcolato sull'insieme A_v di verifica. Se i due insiemi sono uguali allora stiamo indicando l'errore di addestramento, in altro caso è semplicemente l'errore sull'insieme di validazione. Il pedice j invece indica il numero progressivo della stima; infatti dato che il bootstrap è una tecnica statistica, sono richieste numerose stime diverse. È bene osservare che, quando si va a valutare l'errore di un determinato modello, il calcolo di errori diversi al variare di A_l prevede un nuovo addestramento; viceversa il calcolo dell'errore al variare di A_v non prevede un nuovo processo di addestramento, ma semplicemente il calcolo dei valori previsti dal modello per l'insieme di verifica.

Per la stima dell'errore effettuata con il bootstrap, su un modello F , con un insieme di punti I di numerosità N , si parte da diversi insiemi di addestramento A_j di numerosità N pari a quella dell'insieme di punti noti. Gli elementi di tale insieme sono selezionati con ripetizione dall'insieme I . A partire dal modello addestrato su questo insieme A_j è possibile calcolare l'ottimismo opt_j come:

$$opt_j = E_j^{A_j I} - E_j^{A_j A_j}$$

dove $E_j^{A_j I}$ è l'errore commesso sull'insieme di verifica, scelto quest'ultimo come l'intero insieme di punti noti e $E_j^{A_j A_j}$ è l'errore di addestramento. La stima dell'ottimismo \overline{opt} sul modello in analisi, altro non è che la media di J valutazioni dell'ottimismo calcolate su J insiemi di addestramento A_j :

$$\overline{opt} = \frac{\sum_{j=1}^J opt_j}{J}$$

A questo punto, per ottenere una stima dell'errore di generalizzazione con il bootstrap resta solo da calcolare l'errore apparente E^{II} e la stima dell'errore di generalizzazione diventa:

$$\hat{E}_{gen} = \overline{opt} - E^{II}$$

La stima dell'errore di generalizzazione effettuata con il bootstrap offre errori piccoli e

²In statistica *selezionare con ripetizione* significa che nella generazione di un sottoinsieme ogni elemento che viene selezionato dall'insieme per andare a far parte del sottoinsieme viene sostituito nell'insieme; quindi in ciascuna selezione la probabilità che un elemento venga selezionato è costante, inoltre c'è la possibilità che lo stesso insieme venga selezionato più volte.

varianza limitata, tuttavia per arrivare a questa stima è necessario un numero elevato di addestramenti, e quindi tempi lunghi.

Recentemente nell'ambito dello studio della misura dell'accuratezza dei modelli sono stati proposti nuovi metodi per tale stima basati sulla tecnica del bootstrap che velocizzano i tempi. È il caso di Lendasse et al. in [5], dove viene presentata una metodologia che prende il nome di fast bootstrap che si basa sull'osservazione sperimentale che spesso l'ottimismo segue funzioni piane. È possibile effettuare un test di regressione ipotizzando questo andamento, tuttavia è necessario eseguire un test delle ipotesi a posteriori che garantisca questo andamento.

4.2.3 Overfitting

Prima di procedere con la riesamina dei vari metodi di approssimazione che possono essere utilizzati nell'analisi e regressione dei dati, si è ritenuto opportuno fare una premessa sul fenomeno dell'*overfitting*, da tenere sempre ben in considerazione in fase di regressione dei dati.

Si parla di overfitting quando un modello di regressione si adatta ai dati osservati con un numero eccessivo di parametri. In questi casi il risultato dell'approssimazione è un modello che si adatta perfettamente all'insieme di dati noti, ma che non ha nessuna corrispondenza con il fenomeno che si sta cercando di approssimare in tutti gli altri tratti. Spesso questo fenomeno viene considerata una violazione del principio del rasoio di Occam, per il quale la soluzione migliore è la soluzione più semplice.

Nel concreto il problema dell'overfitting si presenta in due situazioni differenti: quando il modello con il quale si sceglie di approssimare il problema non è scelto in maniera appropriata per il problema in analisi. Tale caso si verifica, per esempio, quando si sceglie di approssimare un insieme di dati con il metodo dei minimi quadrati e si sceglie per esempio una funzione polinomiale di terzo grado quando il fenomeno in analisi richiederebbe una polinomiale di primo grado. Se, oltre a questo, il numero di punti dell'insieme di addestramento è scarso, la funzione approssimante si adatterà con precisione ai punti dell'insieme, ma per i tratti meno prossimi ai punti dell'insieme di addestramento si discosterà dalla funzione che più correttamente approssima, introducendo delle informazioni in eccesso. Una possibile soluzione al problema dell'overfitting quando si presenta con queste premesse potrebbe essere la scelta di un modello congiuntamente alla conoscenza del modello matematico che segue il fenomeno sotto osservazione, in modo da scegliere quello che meglio si adatta. Tuttavia questo è un caso raro: il più delle volte la funzione che si cerca di approssimare e da cui provengono i dati dell'insieme di addestramento è completamente ignota o addirittura non è neanche noto quali singoli fattori concorrano nel fenomeno in osservazione.

La seconda situazione si presenta invece nei metodi di approssimazione che utilizzano un procedimento di addestramento di natura iterativa. Capita che si imponga come criterio di arresto sull'addestramento un residuo medio molto piccolo che l'algoritmo non riesce a raggiungere. Durante il processo di addestramento i residui inizialmente decrescono, ma poi, con il procedere delle iterazioni, il residuo si stabilizza e non riesce a raggiungere il residuo medio richiesto come criterio di arresto. Il procedimento si arresta perché si è raggiunto il

numero massimo di iterazioni. Durante il processo di apprendimento il modello non solo non ha raggiunto il residuo minimo scarto, ma sono anche state aggiunte informazioni in più che non hanno alcuna aderenza con il fenomeno in osservazione e anzi portano la funzione approssimante ad essere meno precisa.

Per evitare l'overfitting, e aumentare la precisione della regressione, è necessario attuare particolari tecniche di controllo in fase di addestramento del modello. Una delle più utilizzate in statistica è quella che prende il nome di arresto anticipato, o *early stopping*. Tale metodo prevede la divisione dell'insieme di addestramento in due sottoinsiemi. Il primo, l'insieme di addestramento vero e proprio o *training set*, servirà per il calcolo dei parametri del modello. L'altro prende il nome di insieme di controllo, o *validation set*. Ad ogni iterazione è possibile calcolare l'errore che commette il modello utilizzando i dati dell'insieme di controllo.

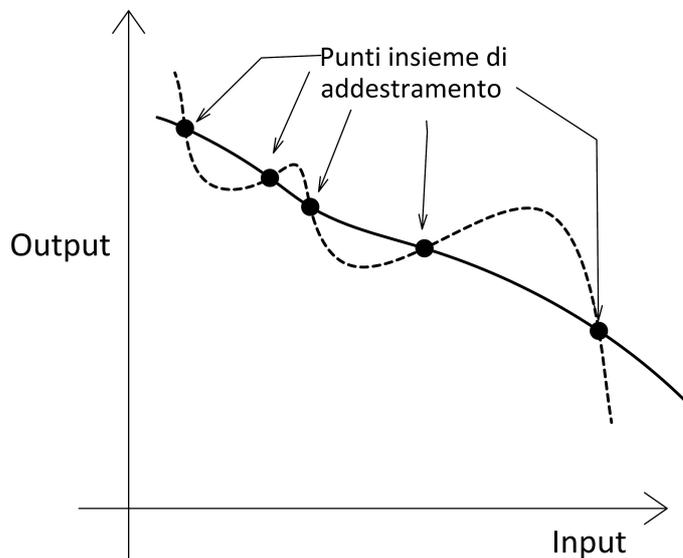


Figura 4.2: Errore dovuto all'overfitting tra la funzione approssimante (linea tratteggiata) e quella da approssimare (linea continua).

Visualizzando l'andamento degli errori di regressione calcolati sul training set e sul validation set al susseguirsi delle iterazioni dell'addestramento, si vede come l'errore sull'insieme di addestramento scende costantemente, questo proprio per la natura degli algoritmi di addestramento che aggiustano i parametri del modello nella direzione della minimizzazione dell'errore. L'andamento dell'errore sull'insieme di controllo sarà decrescente nel primo tratto, successivamente comincerà ad oscillare, se non proprio a crescere. La comparsa dell'andamento crescente degli errori commessi dal modello sull'insieme di controllo è il segnale che indica che l'approssimazione nelle zone distanti dall'insieme di addestramento non è più adeguata. Alla comparsa di questo segnale è bene arrestare anticipatamente il processo di addestramento anche se non si è raggiunto il limite imposto dal criterio di arresto. L'andamento non più decrescente dell'errore di regressione sull'insieme di controllo consente anche di visualizzare il fenomeno dell'overfitting; In corrispondenza del cambio di pendenza della curva degli errori sull'insieme di controllo il metodo di regressione comincia ad immagazzinare informazioni che

non appartengono al fenomeno a cui si riferiscono i dati che si sta cercando di approssimare.

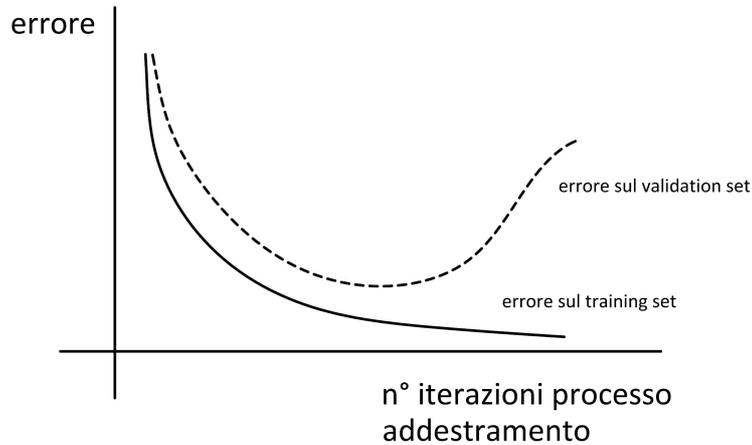


Figura 4.3: Differenza tra l'errore sul **training set** e l'errore sul **validation set** nel processo di addestramento, e comparsa dell'overfitting

4.2.4 Superfici di risposta

Le superfici di risposta polinomiali di primo e secondo grado sono un metodo semplice ed efficace per l'approssimazione dei dati sperimentali. L'utilizzo delle superfici di risposta per l'approssimazione di dati sperimentali è stato ampiamente dibattuto da Montgomery e Meyers in [38] e in [36], dove vengono proposte anche all'interno di un metodo di ricerca degli ottimi grazie a prove sperimentali che prende il nome di *Response Surface Methodology* (RSM). Una superficie di risposta è una funzione che approssima il comportamento di una grandezza al variare degli input con una funzione polinomiale del tipo:

$$y = f(x_1, \dots, x_p) + \epsilon$$

Dove p è il numero di parametri che si suppone influenzino il problema e che sono stati indagati in una serie di esperimenti e dove ϵ è l'errore di approssimazione. Qualora si supponga che l'influenza di ciascun parametro sulla risposta sia di tipo lineare, la superficie di risposta assume la forma di

$$y = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p + \epsilon$$

Dove i β_i sono i parametri di approssimazione. Se si suppone che la relazione tra i parametri e la grandezza risposta sia di secondo ordine allora la funzione diventa del tipo:

$$y = \beta_0 + \sum_{i=1}^p \beta_i x_i + \sum_{i=1}^p \beta_{ii} x_i^2 + \sum_{i < j} \beta_{ij} x_i x_j + \epsilon$$

Nella sua forma di secondo grado le superfici di risposta contengono non solo i termini

di secondo grado di ciascuna parametri x_i^2 , ma anche i termini combinati $x_i x_j$ che danno l'influenza dei vari parametri combinati. Generalmente non si va oltre con il grado delle superfici di risposta. La stima dei valori dei coefficienti di approssimazione si fa utilizzando il metodo ai minimi quadrati che è stato dimostrato forniscono una stima senza errori di questi coefficienti. Vantaggi di questa tecnica di regressione sono indubbiamente la semplicità con cui può essere applicata, il poco tempo richiesto per ottenere i coefficienti e il fatto che si utilizzano polinomi, cioè funzioni semplici il cui comportamento è noto e facilmente investigabile anche in punti distanti dai punti di addestramento. Sicuramente anche il fatto che l'utilizzo di questa tecnica sia stato ampiamente dibattuto costituisce un motivo per farne uso.

4.2.5 Reti Neurali (Artificial Neural Networks-ANN)

Nel corso degli ultimi trent'anni si è fatto largo uso delle reti neurali sia come modello di regressione che come modello di identificazione.

Le reti neurali artificiali fanno propri tutti i vantaggi e gli svantaggi dei modelli di regressione che sono già stati messi in luce, in termini di capacità di approssimare funzioni fortemente non lineari. L'osservazione del funzionamento del sistema nervoso umano ha portato all'ideazione delle reti neurali artificiali. I sistemi nervosi biologici sono organizzati in reti fortemente interconnesse di unità elementari, i neuroni, che si scambiano e rielaborano informazioni. Le unità elementari sono estremamente semplici e possono portare a termine solo un numero limitato di operazioni di base. Le capacità di un sistema nervoso sono il risultato dell'interconnessione tra le cellule di base e crescono tanto più quanto più cresce il numero di unità di base e il numero di interconnessioni. In modo analogo le reti neurali artificiali cercano di riprodurre il funzionamento dei sistemi nervosi biologici e sono per questo composte da molteplici elementi che operano congiuntamente. Proprio come nei sistemi nervosi naturali le unità base delle ANN sono in grado di effettuare solo poche e semplici operazioni e le potenzialità della rete crescono all'aumentare del numero di unità di base e delle interconnessioni. Le unità di base prendono il nome, per analogia, di *neuroni*.

Di contro la complessità analitica di una rete è notevole. Per addestrare una rete neurale, si fa uso di algoritmi molto complessi e difficilmente intellegibili. Inoltre l'addestramento è su base aleatoria: per uno stesso insieme di dati e utilizzando lo stesso algoritmo di addestramento, due addestramenti diversi portano a reti neurali diverse con precisioni diverse.

Neuroni

Le operazioni compiute da un singolo neurone sono estremamente semplici.

Un neurone elabora un input scalare i per ottenere un output sempre scalare o . L'input viene moltiplicato per un peso w e successivamente sommato ad uno scostamento b . Il risultato viene poi processato con una funzione di trasferimento f , secondo lo schema di figura 4.4. Pertanto la relazione che intercorre tra l'input e l'output di un singolo neurone è del tipo:

$$o = f(\mathbf{wi} + \mathbf{b})$$

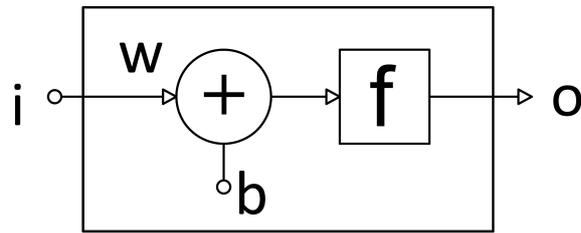


Figura 4.4: Schematizzazione di un neurone di rete neurale.

I neuroni si classificano in base al tipo di funzione di trasferimento che utilizzano. Queste sono generalmente delle funzioni monotone crescenti, lineari o non lineari, che mappano un argomento scalare in un altro scalare. Il risultato della funzione di trasferimento può appartenere ad un insieme illimitato, come quando per esempio si utilizza una funzione lineare, oppure limitato, è il caso per esempio di una funzione come l'arcotangente. Come già detto sono svariate le funzioni di trasferimento che possono essere utilizzate nei neuroni delle ANN. Le più comuni sono quella puramente lineare (che ha come immagine l'intero insieme dei numeri reali), oppure quella logaritmico sigmoideale (con immagine $]0, 1[$).

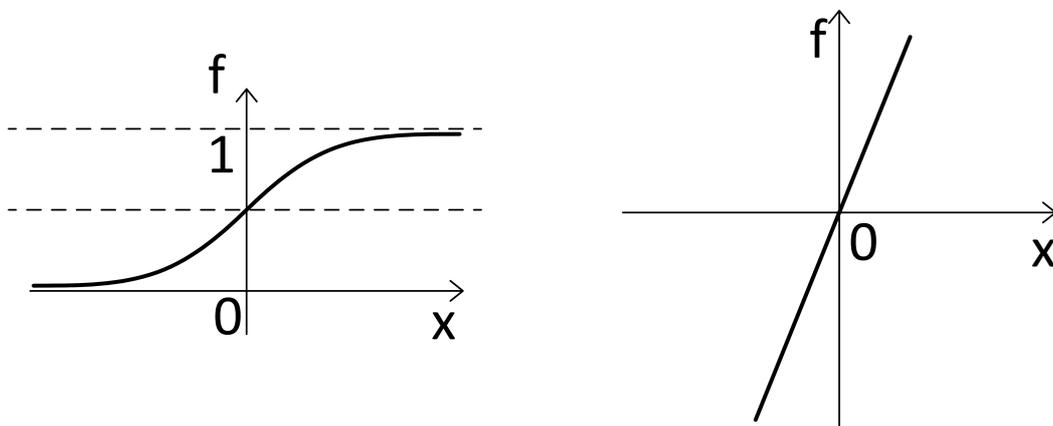


Figura 4.5: Funzione di trasferimento tangenziale e lineare dei neuroni delle reti neurali

Strutture

Una rete neurale è formata da molteplici neuroni organizzati secondo determinate strutture. Più neuroni che elaborano in parallelo gli stessi input formano un *layer*. Una rete neurale è composta di n layer stratificati. I neuroni di uno stesso layer hanno tutti per input gli output

del layer precedente, e gli output diventano gli input per i neuroni del layer successivo. Il primo layer riceve come input direttamente i valori delle variabili indipendenti della funzione da approssimare. L'ultimo layer, ha come output i valori della stessa funzione. Tutti i layer eccetto il layer di output, vengono denominati layer nascosti (*hidden layer*).

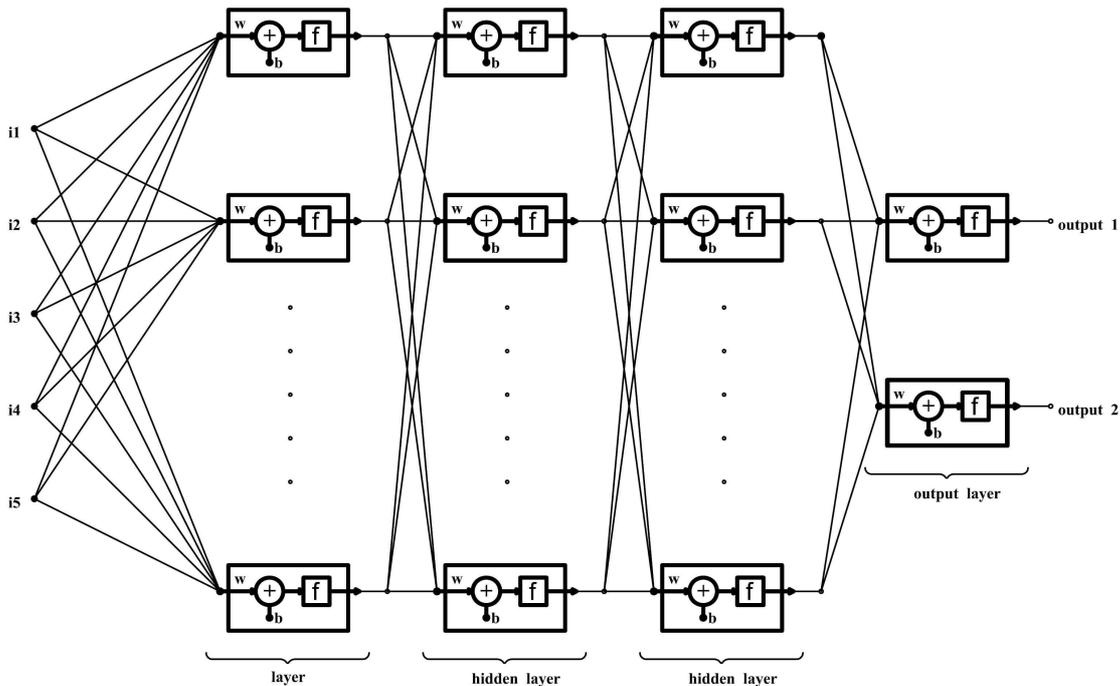


Figura 4.6: Esempio della struttura di una rete neurale a più layer

Generalmente i neuroni di uno stesso layer sono tutti dello stesso tipo, cioè sono caratterizzati dalla stessa funzione di trasferimento. Ciascun neurone ha il suo peso e il suo scostamento e i propri parametri della funzione di trasferimento.

La rete neurale più semplice che si possa pensare è composta da un solo neurone. Viceversa la complessità di una rete neurale è potenzialmente illimitata, potendo aggiungere alla rete più e più strati (layers). Ovviamente reti neurali con strutture differenti, in termini di numero di layer nascosti, di numero di neuroni per ciascun layer e di tipo di neuroni per layer, consentono di approssimare un insieme di dati con precisioni diverse. Per esempio una rete neurale con un numero scarso di neuroni che viene addestrata su un insieme di dati numeroso può non essere sufficiente per garantire un'approssimazione corretta. Viceversa in altre situazioni, con un numero eccessivo di neuroni può comportare problemi di overfitting.

La rete neurale 'migliore' per approssimare una funzione o un insieme di dati, può essere definita come la rete neurale che garantisce il miglior compromesso tra efficacia della regressione sull'insieme di addestramento e semplicità della struttura della rete, in modo da evitare problemi di overfitting, garantendo così un'approssimazione efficace anche nelle zone dello spazio di indagine distanti dai punti di addestramento.

Non esistono regole che consentono di stabilire in anticipo l'ordine di grandezza della precisione che si può ottenere in fase di regressione utilizzando una rete neurale. Per questo non è possibile scegliere a priori quella che è la struttura ottimale della rete per approssimare un determinato insieme di punti. Si può dimostrare (vedi [37]) che è possibile approssimare con una rete neurale con un unico layer nascosto, qualunque funzione continua. Questo teorema garantisce quindi, che qualora si voglia approssimare un insieme di dati è possibile fare uso di una rete neurale con un unico layer nascosto. Non garantisce però che la struttura ad un solo layer nascosto sia la più efficiente o la più precisa.

Ugualmente non è possibile stimare a priori il numero di neuroni che deve avere il layer nascosto, per garantire la miglior precisione. Esistono tuttavia delle formule che forniscono una stima del numero di neuroni da utilizzare per la struttura di una rete neurale al variare del numero di input da immettere e della numerosità dell'insieme di addestramento. Tuttavia tali formule forniscono solo un ordine di grandezza del numero di neuroni, ma non forniscono una stima precisa, e risultano quindi poco utili ai fini pratici; la tendenza comunque è che al crescere del numero di input e della numerosità dell'insieme di addestramento cresce l'ampiezza della rete da addestrare per ottenere una precisione corretta, si veda [37].

Addestramento e Backpropagation (Retropropagazione)

L'addestramento delle reti neurali è di natura iterativa, e tramite successivi aggiustamenti arriva a scegliere i parametri corretti della rete. Ogni iterazione del processo di addestramento prende il nome di *epoca*. Il processo di addestramento è un processo aleatorio. Infatti addestramenti diversi, sempre sullo stesso insieme di output noti, della stessa rete neurale portano a parametri differenti della rete neurale, e quindi a reti neurali di precisione differente. È importante quindi, effettuare più addestramenti della stessa rete neurale e scegliere quello che garantisce la precisione migliore.

A partire da un insieme di dati che riguardano il fenomeno sotto osservazione, la rete neurale è addestrata mediante un opportuno algoritmo che modifica i pesi interni della rete neurale in modo da minimizzare lo scarto tra i valori attesi degli output noti e i valori calcolati della rete per i corrispondenti input.

Esistono in letteratura numerosi algoritmi di addestramento delle reti neurali. Quasi tutti si basano sulla metodologia che prende il nome di *backpropagation*, presentata in modo completo nel 1986 in [47]. Questa prevede un passo in avanti (*forward*) in cui viene predetto il valore dell'output corrispondente all'input noto tramite la rete neurale, e successivamente un passo all'indietro (*back*) nel quale l'errore commesso tra il valore esatto e quello predetto viene propagato all'interno della rete neurale e vengono aggiustati i valori dei pesi dei vari neuroni. Grazie a questo procedimento è possibile aggiustare iterativamente i parametri della rete in modo che possa 'imparare' dai dati noti per minimizzare l'errore della rete neurale. Le epoche si ripetono fino a quando l'errore che la rete neurale commette in fase di approssimazione sull'insieme di dati noti non scende sotto una certa soglia, oppure quando si raggiunge un numero massimo di epoche stabilito.

Gli algoritmi che si basano sul concetto di *backpropagation* sono numerosi. La matemat-

ica che sta dietro a questi algoritmi è molto complicata e non è intellegibile senza uno studio approfondito della materia. I vari algoritmi possono essere classificati secondo il tempo che richiedono per effettuare l'addestramento e la precisione che consentono di raggiungere alla rete, in termini di stima dell'errore di generalizzazione. Ovviamente queste due istanze sono tra di loro opposte quindi gli algoritmi che consentono di raggiungere la precisione migliore sono quelli che impongono i tempi di addestramento maggiori. Tra i vari si può citare l'algoritmo di Levenberg-Marquardt che applicato è stato applicato alle reti neurali in [35] e garantisce ottimi risulti in termini di velocità di addestramento.

Regularization

In fase di addestramento di un modello di regressione si può fare uso di una tecnica che prende il nome di *regularization* e che consente di incrementare la capacità di generalizzazione del metodo su cui è applicata. Presentiamo tale metodo a riguardo dell'addestramento delle reti neurali artificiali. La regolarizzazione comporta la modifica della funzione che misura l'errore di generalizzazione del modello che generalmente è calcolato come la somma dei quadrati degli errori che commette la rete sui dati noti dell'insieme di addestramento. Tipicamente gli algoritmi di addestramento in retropropagazione delle reti neurali minimizzano una funzione di costo è la media della somma dei quadrati degli errori che la rete commette sull'insieme di addestramento. Cioè, un in insieme di N dati noti:

$$F = mse = \frac{1}{N} \sum_{i=1}^N e_i == \frac{1}{N} \sum_{i=1}^N (t_i - a_i)^2$$

dove t_i sono gli output noti dell'insieme di addestramento e a_i sono i valori predetti dalla rete neurale.

E' possibile creare una nuova funzione di costo che viene minimizzata in fase di addestramento, aggiungendo un termine che consiste del valore medio della somma di tutti i parametri che vengono modificati:

$$msw = \frac{1}{n} \sum_{j=1}^n w_j^2.$$

Per cui la funzione di costo diventa

$$mse = \gamma mse + (1 - \gamma) msw$$

dove γ è un rapporto di pesatura che varia tra $[0,1]$. L'utilizzo di questa funzione di costo impone alla rete di aver dei pesi e degli scostamenti contenuti e questo garantisce una funzione di regressione più regolare e una minore sensibilità all'overfitting.

La difficoltà insita in questo metodo è la scelta corretta del parametro di pesatura tra i due termini. Se tale parametro è troppo largo c'è il rischio che si presenti l'overfitting; se è troppo stretto la rete non approssima i dati dell'insieme di addestramento.

4.2.6 Reti di funzioni radiali

Un'altra tecnica di regressione che spesso viene presentata insieme alle reti neurali è la tecnica delle reti di funzioni radiali. In letteratura questa tecnica viene riferita a nomi diversi tra loro; principalmente radial basis function network (RBFN) e radial basis neural network (RBNN).

Le funzioni a base radiale furono presentate per risolvere problemi di interpolazione multidimensionale nella prima metà degli anni sessanta, ma furono presentate come variante delle reti neurali alla fine degli anni ottanta, da qui l'accostamento con le reti neurali. Per sommi capi infatti, le RBFN sono delle reti neurali con un unico layer nascosto, e la funzione di trasferimento utilizzate nei neuroni del layer nascosto sono delle funzioni a base radiale.

Funzioni di basi radiali

Le funzioni di base radiale sono funzioni definite attorno ad un centro, e il cui valore è calcolato sulla distanza, generalmente euclidea, dal centro della funzione in questione.

Gli esempi più comuni di funzioni di base radiale sono:

- gaussiana

$$\phi(\mathbf{x}) = \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}_c\|^2}{2\sigma^2}\right)$$

con $\sigma > 0$;

- multiquadrica inversa

$$\phi(\mathbf{x}) = \left(\frac{1}{\|\mathbf{x} - \mathbf{x}_c\|^2 + \sigma^2}\right)^{1/2}$$

con $\sigma > 0$;

- multiquadrica

$$\phi(\mathbf{x}) = \left(\|\mathbf{x} - \mathbf{x}_c\|^2 + \sigma^2\right)^{1/2}$$

con $\sigma > 0$;

dove \mathbf{x}_c è il vettore della posizione del centro della funzione radiale, mentre σ è un parametro che gestisce l'ampiezza della funzione, cioè di fatto la pendenza della funzione nei dintorni del centro.

A titolo di esempio una funzione radiale $\phi : R^2 \rightarrow R$ può essere visualizzata come una funzione nello spazio simmetrica rispetto ad un asse perpendicolare al piano di input e passante per il centro della funzione come in figura 4.7.

La funzione gaussiana e la multiquadrica inversa sono funzioni *localizzate* (intendendo con questo $\phi \rightarrow 0$ per $\|\mathbf{x} - \mathbf{x}_c\| \rightarrow \infty$), mentre la funzione multiquadrica è *non locale* ($\phi \rightarrow \infty$ per $\|\mathbf{x} - \mathbf{x}_c\| \rightarrow \infty$). Come tecnica di interpolazione le RBFN vengono utilizzate per rappresentare

una funzione y , della quale è disponibile il valore per P punti diversi, come combinazione lineare di P funzioni di base radiale, ossia come

$$y(x) = \sum_{p=1}^P w_p \phi(\|\mathbf{x} - \mathbf{x}_p\|);$$

dove \mathbf{x}_p sono i valori di input per i quali è nota la funzione, che diventano i centri delle P funzioni radiali, e w_p sono i P pesi ignoti che vengono ricavati risolvendo il sistema che si ricava imponendo le P condizioni di interpolazione:

$$y(\mathbf{x}_p) = y_p.$$

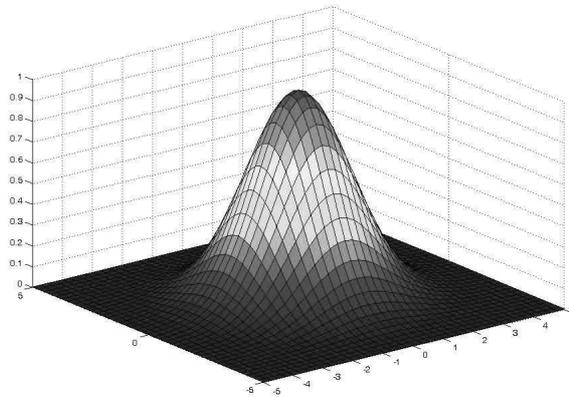


Figura 4.7: Esempio di funzione di trasferimento a base radiale.

Neurone di rete di funzioni a basi radiali

Il neuroni del layer nascosto delle reti di funzioni di base radiale, rappresentato in figura 4.8 sono diversi dai neuroni delle reti neurali artificiali. Per una rete che deve approssimare una funzione con un vettore di input appartenente ad uno spazio r -dimensionale i singoli neuroni hanno ciascuno r pesi interni $w_1 \dots w_r$ che rappresentano di fatto nello spazio r -dimensionale il centro attorno al quale è costruita la funzione a base radiale di quel neurone. A partire da questo centro viene calcolata la distanza dell'input, che successivamente viene moltiplicata per l'altro peso del neurone in questione, b , che fornisce l'ampiezza della funzione. Il tutto viene poi processato con la funzione di trasferimento vera e propria.

La funzione più utilizzata nelle reti di funzioni a base radiale è la funzione gaussiana, già vista prima. Un neurone con una funzione di trasferimento del tipo $\phi(d) = \exp(-d^2) = e^{-d^2}$ e con un peso adimensionale equivalente a $b = 1/\sigma\sqrt{2}$ processa gli input come una funzione gaussiana.

Da questo risulta che un neurone di rete di funzioni a basi radiali che utilizza una funzione gaussiana, elabora un input con una funzione a base radiale centrata in \mathbf{w} con una dispersione σ pari a $1/b\sqrt{2}$. Quindi la funzione che elabora l'input \mathbf{p} può essere scritta come:

$$\phi(p) = \exp(- (b\|\mathbf{p} - \mathbf{w}\|)).$$

Un'ulteriore considerazione va fatta è sul significato di ampiezza della funzione a base radiale. Come si è già detto il parametro che moltiplica la distanza $\|\mathbf{p} - \mathbf{w}\|$ tra l'input e il centro della funzione gaussiana, rappresenta tale ampiezza. Quello che succede è che la funzione gaussiana assume valore unitario quando l'input della funzione coincide con il centro della funzione, e poi andrà decrescendo verso zero man mano che ci si allontana dal centro. Tale funzione inoltre produrrà un output pari a 0.5 quando il punto di input si trova ad una distanza dal centro pari a $0.8326/b$. Quindi aumentando b la funzione, che è monotona decrescente con la distanza dal centro, assume valore 0.5 per punti più vicine al centro e, di fatto, si chiude intorno all'asse di simmetria della funzione. Viceversa abbassando b la funzione si apre intorno all'asse di simmetria, diventando più piatta.

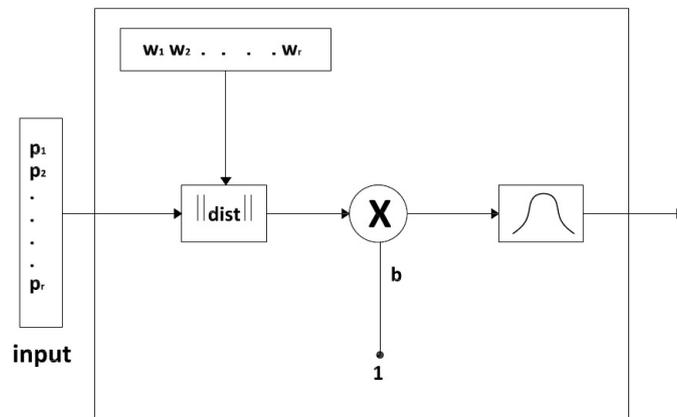


Figura 4.8: Schema di neurone di rete di funzione a base radiale.

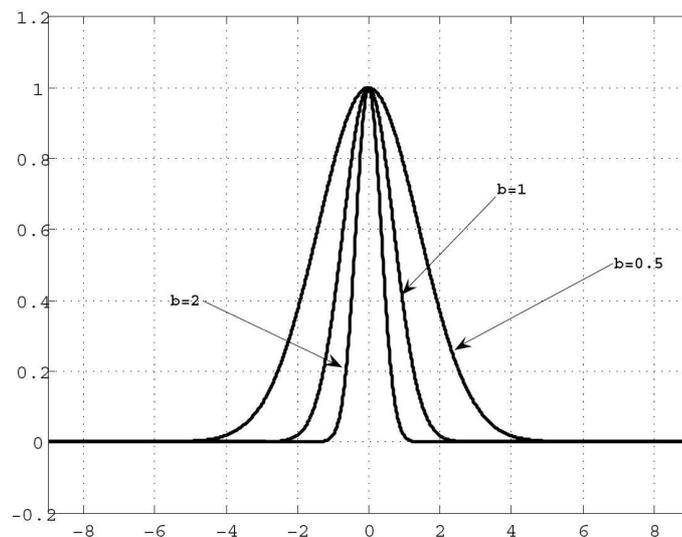


Figura 4.9: Esempio dell'influenza del parametro b sul comportamento della funzione di trasferimento .

Struttura

Come già anticipato, la struttura di una rete di funzioni a base radiale è molto semplice. Essenzialmente è composta da una rete con due layer; il primo layer, quello nascosto, è composto da neuroni caratterizzati da una funzione di trasferimento di base radiale, generalmente di tipo gaussiano. Il layer di output è caratterizzato da una funzione di trasferimento ad output lineare.

In genere le reti di funzioni a base radiali sono funzioni create per interpolare o approssimare funzioni con un output unico. Quindi il layer di output è caratterizzato da un unico neurone, che di fatto pesa i risultati degli output delle funzioni radiali del layer nascosto. Il risultato di una rete di funzioni a basi radiali con output singolo e P neuroni nel layer nascosto, schematicamente è:

$$y(x) = \sum_{p=1}^P w_p \phi(\|\mathbf{x} - \mathbf{x}_p\|).$$

Questa funzione, ipotizzando di utilizzare dei neuroni con funzione di trasferimento di tipo gaussiano diventa:

$$y(x) = \sum_{p=1}^P w_p \exp\left(-b \cdot \|\mathbf{x} - \mathbf{x}_p\|^2\right).$$

Così come è possibile dimostrare che le reti neurali con un solo layer nascosto sono degli approssimatori universali, cioè sono in grado di approssimare qualunque funzione continua, allo stesso modo una rete di funzioni radiali è un approssimatore universale.

Come vedremo, alcuni processi di addestramento delle reti di funzioni a base radiale non richiedono di scegliere la struttura della rete più efficiente per approssimare l'insieme di dati, ma l'algoritmo stesso sceglie la struttura.

Addestramento

Il processo di addestramento delle reti di basi radiali differisce dagli algoritmi di addestramento che si utilizzano per le reti neurali. Di fatto però anche questi ricercano i pesi e gli scostamenti delle reti che minimizzano gli errori di approssimazione.

Una nota particolare va fatta sul carattere della funzione che si vuole ottenere con l'addestramento. È possibile infatti ottenere una rete di funzioni a base radiali che si comporta non come una funzione approssimante, ma bensì come una funzione interpolante. Si parla in questo caso di RBFN esatta. Una rete di funzioni radiali esatta addestrata su un insieme T di dati noti ha errore nullo sui punti dell'insieme T stesso. Per arrivare a questo risultato si fa uso di una rete di funzioni con un numero di neuroni nascosti pari alla numerosità dell'insieme T ; inoltre i centri delle funzioni radiali sono scelti in corrispondenza di punti dell'insieme di dati noti T . In fase preliminare non resta che indicare il parametro che controlla l'ampiezza delle funzioni radiali. L'algoritmo di addestramento calcola i pesi restanti imponendo la condizione di interpolazione. Il principale svantaggio di questa soluzione di addestramento è che una rete

di funzioni radiali esatta ha tanti neuroni nascosti quanti sono i dati noti. Inoltre tale tipo di addestramento fornisce una rete non adeguata in tutti i casi in cui si deve approssimare un insieme di dati numeroso; in questi casi la rete è molto complessa e, oltre al tempo necessario per addestrarla, è quasi sempre soggetta al problema dell'overfitting.

Un addestramento più efficiente è di natura iterativa e si muove nella direzione della minimizzazione dell'errore di approssimazione. Questo parte da una rete con un numero basso di neuroni nascosti e infittisce progressivamente il layer nascosto aggiungendo ad ogni iterazione un neurone in più. Il processo di infittimento del layer nascosto si arresta quando l'errore della rete è sufficientemente basso, oppure quando è stato raggiunto un numero massimo di soglia stabilito per i neuroni oltre il quale l'incremento di precisione della rete ad ogni iterazione dell'algoritmo di addestramento non è più significativo o si rischia di mandare la rete in overfitting. Tuttavia è bene aggiungere che l'informazione riguardante il numero massimo di neuroni che possono essere utilizzati per la rete si acquisisce solo con l'esperienza. Di volta in volta il centro della funzione radiale del neurone nascosto che viene aggiunto ad ogni iterazione, è scelto tra i punti dell'insieme di input noti come quello che minimizza l'errore della rete.

Sono stati proposti in letteratura altri algoritmi di addestramento, che si differenziano tra di loro per complessità e per precisione della rete ottenuta. Quasi tutti questi partono da una rete con struttura assegnata; quindi presentano lo svantaggio di dover scegliere a priori il numero di neuroni nascosti. La scelta di tale numero non è ovviamente supportata da alcuna formula e quindi è di natura sostanzialmente empirica e si rifà all'esperienza del programmatore.

In breve una tecnica molto semplice per addestrare una rete con n neuroni nascosti potrebbe consistere nello scegliere come centri per le funzioni radiali n punti casuali dell'insieme di punti noti, badando che questi siano perlomeno distribuiti uniformemente sullo spazio di indagine. L'algoritmo di addestramento poi si limiterà a calcolare i pesi w_p del layer di output, in modo da minimizzare lo scarto. Tale approccio all'addestramento tuttavia, raggiunge un discreto grado di precisione solo per insiemi noti molto numerosi. Un passo in più viene rappresentato da un aggiustamento delle posizioni dei centri al fine di individuare una disposizione dell'insieme dei centri uniforme rispetto alla posizione dei punti dell'insieme noto che costituiscono il training set. Infine il passo oltre ancora è quello che prevede la scelta dei centri delle funzioni radiali completamente svincolata dai punti dell'insieme di dati noti, ma scelti in modo da minimizzare l'errore di approssimazione della rete.

Scelta dell'ampiezza delle funzioni radiali

Come già visto l'ampiezza della funzione di trasferimento a base radiale dei neuroni modifica la forma della funzione di trasferimento ed è un parametro che deve essere impostato quando in sede di predisposizione della struttura della rete. Di fatto, per come sono costruite le RBFN, l'ampiezza della funzione rappresenta l'influenza che il dato di un punto, cioè il valore della funzione nel suo centro, ha sullo spazio circostante; questo congiuntamente sia in termini di intensità di influenza che di zona di influenza.

Tale dato è una scelta del programmatore da effettuare prima dell'addestramento. Non esistono criteri di scelta per questo parametro e la scelta è dettata dall'esperienza di chi utilizza le reti. I criteri che sono stati presentati sono indicati per un determinato algoritmo di addestramento, e quasi sempre non sono altro che tentativi di formalizzare con una formula la regola generica che dice che l'ampiezza delle funzioni di trasferimento non deve essere né eccessiva né eccessivamente stretta.

Va da sé che la scelta di tale parametro è, per così dire empirica. Il valore corretto dipende da tanti fattori: tipo della funzione di approssimare, numerosità dell'insieme di dati noti, struttura della rete e da tanti altri fattori non altrettanto intellegibili. Se non si hanno a disposizione esperienze precedenti che forniscono indicazioni sull'ampiezza da utilizzare per tipologie di dati analoghe, è bene effettuare sempre uno studio dell'influenza di tale parametro sperimentando come valori diversi influiscano sul risultato in approssimazione.

Confronto tra reti neurali e reti di funzione a base radiale

Sia le reti neurali che le reti di funzioni a base radiale sono esempi di rete di funzioni con struttura a layer. Quindi è logico che per ogni rete neurale esiste una particolare rete di funzioni a base radiale capace di imitare il comportamento di tale rete e viceversa(vedi [37]). Tuttavia esistono delle differenze e delle similitudini tra le due tipologie di rete che è bene evidenziare e che è bene tenere presente quando si utilizzano queste due tecniche:

- una RBFN ha un solo strato nascosto di neuroni, mentre le reti neurali possono essere anche più di uno strato nascosto di layer;
- quando sono utilizzate in fase di regressione la migliore scelta per entrambe le tecniche prevede di utilizzare neuroni di output lineari;
- l'argomento della funzione di attivazione di ciascun neurone nascosto delle RBFN sono le distanze euclidee tra il vettore di input e il centro della funzione di attivazione mentre nelle reti neurali è il prodotto interno tra il vettore di input e il vettore peso del neurone;
- una NN costituisce un approssimatore globale di una funzione non lineare, mentre una RBFN costruisce un'approssimazione globale a partire da degli approssimatori locali.

Dalle considerazioni fatte sopra si può affermare che entrambe le tecniche consentono di generare delle approssimazioni globali di funzioni non lineari, tuttavia le reti neurali consentono di raggiungere, su uno stesso insieme di addestramento, una determinata precisione con un numero di parametri inferiore a quello delle RBFN.

4.2.7 Kriging

La tecnica di regressione che prende il nome di Kriging è stata sviluppata negli anni sessanta per stimare la produttività delle vene minerarie, e prende il nome dall'ingegnere minerario

Krige, che fu il primo a proporla. Nel 1989 Sacks e altri proposero in [2] l'utilizzo di questa tecnica statistica per approssimare insiemi di dati derivanti da osservazioni di fenomeni deterministici come i computer experiments. Questa tecnica appartiene alla famiglia degli algoritmi di approssimazione ai minimi quadrati, come le RS; tuttavia ha la capacità ulteriore, rispetto a queste ultime di fornire il valore esatto in corrispondenza dei punti dei dati di addestramento. Dato un insieme di dati noti di numerosità m di una funzione $y : R^n \rightarrow R$, la forma matematica che assume il modello ottenuto con il Kriging è del tipo:

$$\hat{y} = F(\beta, \mathbf{x}) + z(\mathbf{x})$$

In questo modello F è un modello di regressione lineare composto da una combinazione lineare di un numero arbitrario di funzioni polinomiali f_i , e dove β è il vettore dei β_i coefficienti della combinazione lineare. $z(\mathbf{x})$ è una funzione stocastica che rappresenta l'errore che compie la funzione di regressione lineare F rispetto a y e impone il passaggio di \hat{y} per i punti y_i dell'insieme di dati noti.

Come funzione lineare F si assume una combinazione lineare di p funzioni $f_i : R^n \rightarrow R$, con coefficienti di regressione β_i :

$$F(\beta, x) = \beta_1 f_1(x) + \dots + \beta_p f_p(x).$$

Generalmente le funzioni f_i possono essere delle funzioni di regressione polinomiali, solitamente di grado nullo (costanti), di primo o di secondo grado.

Utilizzando una funzione polinomiale di grado nullo di fatto significa utilizzare un modello di regressione lineare costante, che applica, semplicemente una traslazione in senso verticale dei valori della seconda parte del modello. Di contro utilizzare una polinomiale di secondo grado significa utilizzare come funzione di regressione lineare una RS.

Alcuni autori, vedi per esempio [6], utilizzando la prima strada per la scelta del grado della funzione polinomiale, lasciando quindi che la seconda parte del modello interpoli i punti dell'insieme di dati noti e sostanzialmente definisca l'andamento della funzione. Tuttavia questa scelta, almeno ad esperienza di chi scrive non sembra essere efficace. La tecnica del Kriging risulta essere efficace proprio per come è costruita: una funzione di regressione, che rappresenta l'andamento del fenomeno nella sua traccia principale, accoppiato con una funzione di natura stocastica che impone il passaggio per i punti noti. La parte stocastica del modello di fatto rappresenta un aggiustamento del modello di regressione, al fine di interpretare il comportamento effettivo del fenomeno nella sua interezza, compresi sia fattori casuali che gradi di libertà che non erano stati considerati.

Utilizzare quindi una funzione di regressione costante, significa lasciare alla parte stocastica del modello il compito di interpolare nel migliore dei modi la funzione, quando questa parte sarebbe dedicata a rilevare scostamenti più tenui. In questo caso il modello è difficilmente controllabile e il risultato anche se interpola con precisione i punti dell'insieme noto non è in grado di rappresentare con precisione la funzione che si sta cercando di approssimare.

La scelta del modello da utilizzare nel Kriging viene effettuata dal programmatore nel-

la fase precedente di regressione. La scelta del modello influenza in modo significativo la precisione che si ottiene in fase di regressione, e, pertanto, deve essere fatta con cura. Fortunatamente la scelta fatta sulle funzioni polinomiali di grado nullo, primo o secondo sembra offrire alternative sufficienti a coprire tutte le situazioni. In particolare quella di grado nullo sembra efficace solamente in pochi casi, mentre già quella di primo grado sembra affidabile per quasi tutte le situazioni.

Il processo stocastico $z(\mathbf{x})$ del modello è assunto con media nulla e covarianza tra $z(x)$ e $z(\mathbf{w})$ pari a:

$$E(z_l(\mathbf{x}), z_l(\mathbf{w})) = \sigma^2 R(\theta, \mathbf{x}, \mathbf{w})$$

per $l=1, \dots, q$.

Nell'espressione precedente σ_l è la varianza del processo per la l -esima componente, R è il modello di covarianza con parametro θ . In pratica si vuole rappresentare la deviazione del processo dalla funzione di regressione come un processo stocastico, anche se non è sempre così (infatti i processi come i computer experiments sono di natura deterministica). Imponendo il modello della covarianza tra $z(\mathbf{x})$ e $z(\mathbf{w})$ si vuole imporre quella che è correlazione (o, da un altro punto di vista l'interazione) tra i valori di due punti nello spazio. Le variabili della forma di questa correlazione sono il parametro θ e le posizioni dei punti \mathbf{x} e \mathbf{w} nello spazio di input della funzione. La letteratura riguardante questa tecnica di regressione ha concentrato la sua attenzione su funzioni radiali prendendo come riferimento la distanza tra i due punti \mathbf{x} e \mathbf{w} dello spazio di input. Nello specifico ci si è concentrati su funzioni del tipo:

$$R(\theta, \mathbf{x}, \mathbf{w}) = \prod_{j=1}^n R_j(\theta_j, x_j - w_j).$$

Più in specifico la tipologia delle funzioni R_j , alcune delle quali sono rappresentate in figura 4.10, che di solito si usano sono:

Tabella 4.1: Funzioni di correlazione utilizzate per il Kriging.

Nome	$R_j(\theta, dist_j)$
Exp	$exp(-\theta_j \ d_j\)$
Gauss	$exp(-\theta_j \ d_j\)$
Lin	$max\{0, 1 - \theta_j \ d_j\ \}$
Spherical	$1 - 1.5\xi_j + 0.5\xi_j^3, \quad \xi_j = min\{0, 1 - \theta_j \ d_j\ \}$
Cubic	$1 - 3\xi_j^2 + 2\xi_j^3, \quad \xi_j = min\{0, 1 - \theta_j \ d_j\ \}$
Spline	$s(\xi_j), \quad \xi_j = min\{0, 1 - \theta_j \ d_j\ \}$

dove

$$\varsigma(\xi_j) = \begin{cases} 1 - 15\xi_j^2 + 30\xi_j^3 & \text{per } 0 \leq \xi_j \leq 0.2 \\ 1.25(1 - \xi_j)^3 & \text{per } 0.2 < \xi_j < 1 \\ 0 & \text{per } \xi_j \geq 1 \end{cases}$$

Da notare che tutte le funzioni modello sono asintotiche a zero; questo è compatibile con il principio che segue la funzione di correlazione dei dati, cioè che rappresentano la mutua influenza, o il comportamento, delle varie zone: tanto più due zone sono vicine, tanto più sono tra di loro correlate.

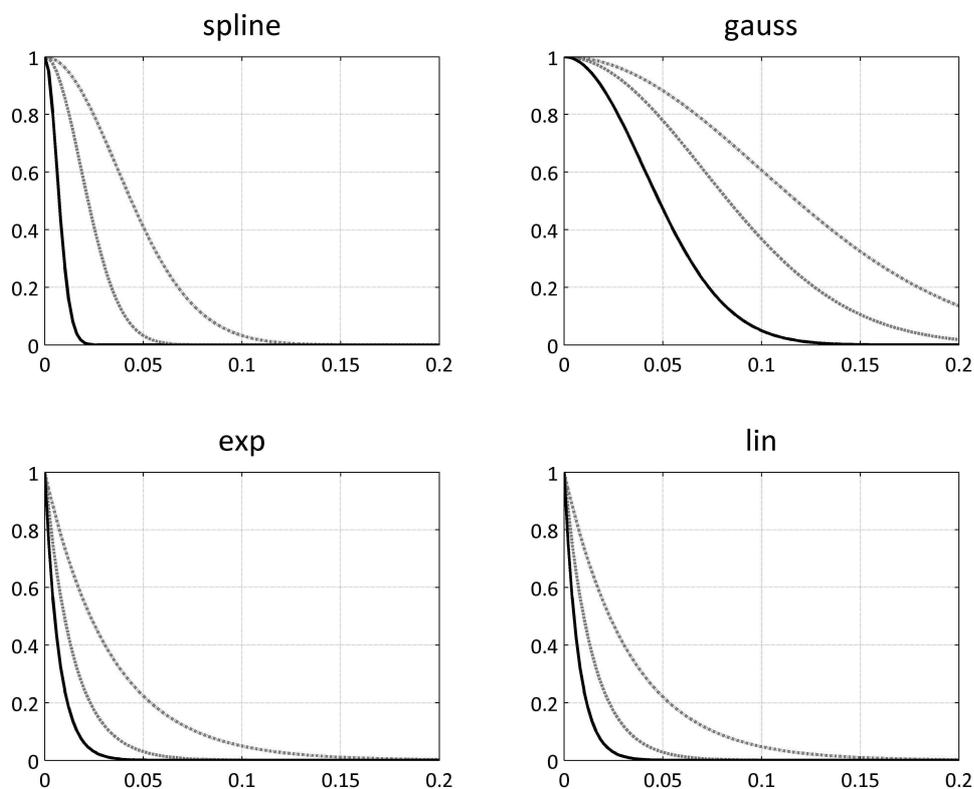


Figura 4.10: Esempi di diverse funzione di correlazione che possono essere utilizzate con il metodo del Kriging. L'intensità delle linee corrisponde a parametri θ più alti.

Come per la funzione di regressione anche in questo caso la scelta della tipologia della funzione stocastica del processo deve essere supportata dall'esperienza del programmatore in modo da poter minimizzare l'errore di approssimazione. Inoltre è bene porre attenzione al fenomeno in osservazione: infatti le tipologie di funzioni che si utilizzano sono sostanzialmente di due tipi. Da come si può notare nella figura, i modelli della funzione di regressione si differenziano sostanzialmente per il comportamento all'origine. Le funzioni gaussiane, cubica, e spline hanno un comportamento parabolico. Quelle esponenziale, lineare e sferica hanno un comportamento lineare in corrispondenza dell'origine. Pertanto se il fenomeno che si va

ad approssimare segue una funzione che è continua e differenziabile su tutto lo spazio le correlazioni che funzionano meglio sono quelle che hanno andamento parabolico all'origine. Quando viceversa si approssima un fenomeno con comportamento lineare all'origine le correlazioni che seguono questo comportamento saranno più efficaci. La figura mostra appunto il comportamento dei vari modelli di correlazione. Da notare come varie il comportamento al variare del parametro θ della funzione di regressione. Al crescere di tale parametro la singola funzione risulta più raccolta intorno all'asse.

Un'ulteriore osservazione è che per come sono strutturate le funzioni di correlazione si possono avere coefficienti di correlazione θ_l differenti per ciascuna direzione di sviluppo della funzione. Questo per rappresentare l'anisotropia della funzione, cioè il diverso comportamento della funzione lungo ciascuna direzione.

Per un insieme di dati noti stabilito, scelte la tipologia della funzione di regressione e il modello della funzione di correlazione, i coefficienti della funzione di regressione e della funzione che individuano il modello di Kriging che meglio approssima la funzione da approssimare sono di difficile calcolo. Come è noto anche da quanto detto per le ANN e le RBFN i principi che sono alla base di queste moderne tecniche di regressione sono comprensibili da qualunque individuo con un background matematico e statistico medio-alto. Per l'implementazione di questi metodi sono necessarie conoscenze matematiche specialistiche ed uno studio approfondito.

Tuttavia anche per il Kriging esistono dei toolbox che consentono di utilizzare questa tecnica in tutte le sue potenzialità (vedi [49] e [48]). Utilizzando questi toolbox è possibile ottenere applicare la tecnica di Kriging ad un insieme di dati noto e ricavare un'approssimazione dovendo indicare solamente la tipologia di funzione di regressione e il modello di funzione di correlazione che l'esperienza dell'operatore indicano come più favorevoli all'interpolazione.

Capitolo 5

Utilizzo industriale delle tecniche di ottimizzazione

Gli algoritmi di ricerca dell'ottimo che sono stati presentati nel capitolo 3, sia che siano utilizzati in forma pura che in associazione con i metamodelli presentati nel capitolo 4.2, possono essere utilizzati in sede industriale. Nell'ambito turbomacchinistico vengono utilizzate in concomitanza con le simulazioni CFD per individuare la geometria che minimizza una funzione di costo. Tale approccio è quello presentato nella letteratura tecnico-scientifica negli ultimi anni; questo si è dimostrato efficace nella progettazione di svariate tipologie di turbomacchine, sia per ottimizzare geometrie complesse come la fluidodinamica di turbomacchine centrifughe, che singoli dettagli come forma di una geometria come in [12].

Gli esempi di ottimizzazione presentati in letteratura si rifanno ad approcci che sono pensati per ottimizzare una soluzione unica. Per applicare in mancanza di conoscenze pregresse si parte da una parametrizzazione assolutamente generica e senza imporre alcun vincolo alle geometrie che possono essere assunte se non quelli minimi richiesti dalla progettazione; l'algoritmo viene lasciato libero di ricercare la soluzione direttamente sullo spazio di indagine oppure su un metamodello che generalizza lo spazio di indagine stesso. L'efficacia di un utilizzo di questo tipo è molto varia: sono molti i parametri che influenzano la rapidità di raggiungimento di una soluzione: ampiezza dello spazio di indagine, complessità del modello da ottimizzare, parametri dell'algoritmo di ricerca dell'ottimo, precisione della regressione effettuata con il metamodello adottato. Non è possibile stimare in anticipo il numero di simulazioni necessarie per raggiungere una soluzione quindi il tempo.

In alternativa è possibile utilizzare delle conoscenze pregresse per circoscrivere il campo di indagine intorno in una zona desiderata che il progettista sa essere efficace. Questo significa utilizzare l'esperienza per circoscrivere la zona in cui applicare l'algoritmo di ricerca dell'ottimo ad un intorno dell'ottimo stesso.

Quindi un'applicazione di tali metodi in questo senso è dedicata principalmente ad una progettazione caso per caso. Infatti non è prevista alcuna pianificazione per raccogliere dati ed estendere i risultati ricavati nella progettazione di una macchina ad altre macchine, oppure di raccolta di informazioni per la progettazione standardizzata di pompe in serie. Tali

informazioni vanno comunque ricavate da progetti precedenti o da ottimizzazione precedenti, lasciando così al progettista il compito di estendere le informazioni ricavate in un caso in altri. In parole comuni, questi metodi non prevedono a chi li utilizza di accumulare un'esperienza universale fruibile da chiunque si ritrovi ad avere a che fare con la progettazione e quindi non consentono di soddisfare del tutto le esigenze di una produzione industriale esposte nel paragrafo 1.4.

Si presenta in questa tesi un differente utilizzo di questi strumenti matematici di ottimizzazione che prevede la creazione di una banca dati di valori in fase preliminare alla progettazione di un pezzo o di un organo o di un particolare fluidodinamico. La corretta pianificazione della creazione della banca dati consente di ottenere un insieme di valori dal quale ricavare dati da utilizzare per la progettazione di più macchine simili.

5.1 Banca dati

Il metodo che si propone per l'utilizzo degli strumenti di ottimizzazione in sede industriale è un metodo per la creazione di banche dati di simulazioni CFD per caratterizzare il comportamento di un particolare da utilizzare in sede di progettazione e l'influenza di variazioni di forma del particolare sulle prestazioni. Un'adeguata pianificazione iniziale di questa banca dati consente anche di prevedere geometrie simili nel layout, ma differenti nelle proporzioni in modo da garantire la fruibilità delle informazioni contenute in questa banca dati anche in casi di progettazione successivi.

La creazione di una banca dati prevede innanzitutto la parametrizzazione della geometria dell'organo, o del particolare che si vuole studiare; lo scopo è quello di stabilire una relazione tra la forma del pezzo in questione e i valori numerici assunti dai parametri che la caratterizzano per stabilire, transitivamente, una relazione tra i risultati numerici delle simulazioni e i dati dei parametri. La parametrizzazione della geometria deve essere concepita con lungimiranza in modo che possa conteggiare la maggior parte delle forme e delle proporzioni che può assumere tale particolare nei vari casi. Non si deve allo stesso tempo scendere nel dettaglio e realizzare una parametrizzazione omnicomprensiva; un numero eccessivo di parametri non è opportuno in alcun caso perchè introduce complicazioni eccessive nel trattamento dei dati per quanto visto nel paragrafo 4.1.1. Allo stesso modo i range di variazione dei parametri devono essere scelti in modo da essere sufficientemente ampi per consentire di coprire le proporzioni più disparate, ma non deve essere troppo ampio da individuare delle geometrie irrealizzabili o rendere troppo complesso il modello.

A partire dalla parametrizzazione individuata si devono svolgere delle simulazioni per caratterizzare il funzionamento delle varie geometrie al variare dei valori dei parametri in questione. Ciascuna geometria deve essere caratterizzata con uno o più valori scalari che consentono il confronto con le altre ricavati dalla fase di postprocessing della simulazione. In fase preliminare quindi è necessario pianificare con attenzione la creazione della banca dati e la fase di post processing; questo in modo tale effettuare simulazioni tra di loro confrontabili e da aver a disposizione i parametri più corretti per il confronto.

I risultati numerici delle simulazioni devono poi essere generalizzati e archiviati in un metamodello. Per ottenere una banca dati che sia efficace ed utilizzabile in circostanze diverse si deve effettuare una regressione con una precisione omogenea su tutto lo spazio di indagine individuato dai range di variazione dei singoli parametri. Non ci sono regole che consentano di stimare il numero di simulazioni necessarie per ottenere un'adeguata precisione in fase di regressione. Perciò fissando un numero di simulazioni non si ha l'assicurazione che i dati ricavati da queste siano sufficienti a garantire una precisione adeguata in fase di regressione. Le tecniche di DOE fissano il numero di simulazioni e individuano quali simulazioni effettuare per massimizzare il numero di informazioni che si possono ricavare.

In questa sede quindi si presenta una nuova tecnica DOE che consente di infittire progressivamente il campione di simulazioni effettuate. E' quindi possibile addestrare un metamodello infittendo iterativamente l'insieme di addestramento fino a raggiungere la precisione desiderata.

Lo schema delle operazioni per la realizzazione della banca dati è quello presentato in figura 5.1. Si utilizza un insieme di dati iniziale per addestrare un modello di regressione e si verifica la stima dell'errore di approssimazione; se l'errore è inferiore ad una soglia che si ritiene accettabile, si può procedere con l'utilizzo dei dati raccolti in sede di progettazione. In caso contrario è possibile, per iterazioni successive, infittire l'insieme di addestramento grazie alla tecnica DOE e addestrare il metamodello fino al raggiungimento di un'errore che si ritiene accettabile.

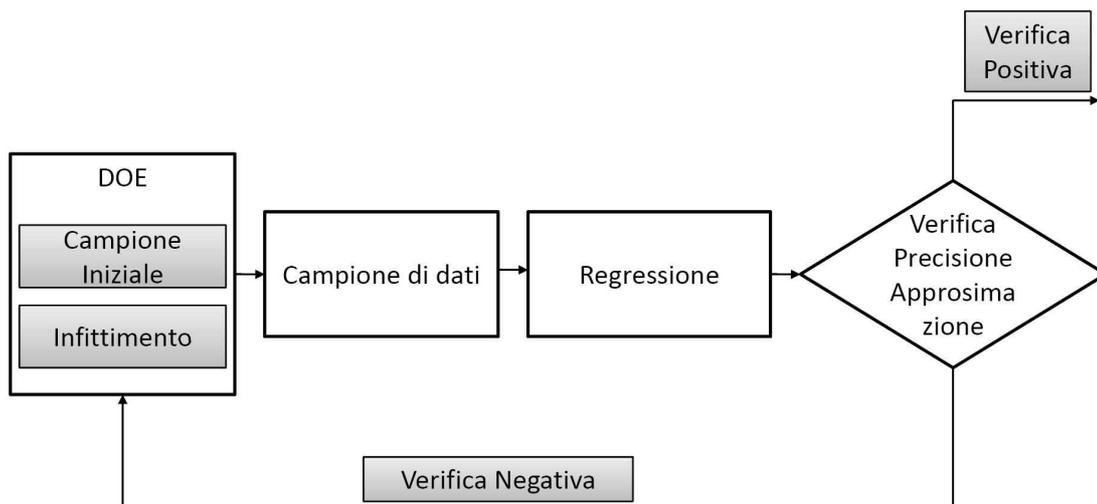


Figura 5.1: Diagramma di flusso che esplica il processo di creazione di una banca dati.

5.2 DOE-Ipercubi latini stratificati

La pianificazione delle simulazioni da svolgere per creare la banca dati da utilizzare in sede di progettazione, consiste nella scelta dei dati che vanno a costituire gli input delle simulazioni da effettuare e, in un secondo momento, gli input dell'insieme utilizzato per l'addestramento dei modelli di regressione.

Non è possibile stabilire a priori la numerosità del campione di dati, e quindi della banca dati, che è necessario adottare per ottenere una regressione con un determinato livello di precisione. Questo per via della complessità e la varietà dei fenomeni implicati nella generalizzazione, della complessità matematica dei diversi metamodelli che si vanno ad utilizzare in sede di regressione, delle diverse tecniche DOE che si possono utilizzare e del numero di parametri di input che vengono utilizzati di volta in volta.

Spesso, nei problemi di ottimizzazione, la dimensione dell'insieme di campionamento viene scelta in base al tempo necessario per una simulazione, al tempo complessivo e alle risorse di calcolo a disposizione; fissato il numero di simulazioni si utilizzano tecniche DOE per individuare i parametri di input delle simulazioni e il regressione che l'esperienza ha indicato essere come più efficace per quel tipo di problema. Questo non garantisce di arrivare ad un determinato livello di precisione.

La creazione di una banca dati di risultati di simulazioni al computer ha delle peculiarità diverse rispetto a quelle di un problema di ottimizzazione tradizionale. Mentre in quest'ultimo si indaga un campo ristretto di variazione dei parametri di input vincolati dallo specifico caso in analisi, nella creazione di una banca dati si indaga uno spazio di variazione ben più ampio. In un problema di ottimizzazione tradizionale ci si può limitare a individuare un trend generico delle variabili per poi andare a raffinare la soluzione, campionando altri punti, nell'intorno dell'ottimo. Viceversa nella creazione di una banca dati bisogna garantire uno standard di precisione generalizzato su tutto lo spazio di indagine.

Quando si va ad affrontare un problema di questo tipo si corre il rischio di sottostimare il campione, cioè di completare le simulazioni e la raccolta dei risultati, effettuare la regressione e ritrovarsi con una soluzione troppo approssimata. Non esistono criteri per infittire ulteriormente l'insieme di punti di campionamento a meno di non inserire punti scelti casualmente o nell'intorno delle zone in cui risultano i maggiori errori di approssimazione. Non è possibile infittire un ipercubo mantenendo le peculiarità dell'ipercubo latino, cioè che i punti all'interno dell'insieme di campionamento non si ripetono e che ciascuna porzione in cui è divisa ciascuna variabile di input sia rappresentata in modo pari alle altre; un ipercubo latino è un design bloccato, in altre parole se aggiungo punti ad un ipercubo non ottengo più un ipercubo. Si corre anche il rischio contrario, cioè che ci troviamo con un numero eccessivo di risultati, quando sarebbero bastati molti meno punti, quindi tempo e risorse di calcolo, per raggiungere quel risultato. Anche questo è un problema non da poco; nella creazione delle banche dati si ha a che fare con centinaia di simulazioni, a volte migliaia, quindi ottimizzare i tempi è estremamente importante.

Risulta chiaro che per questo tipo di problema è necessario individuare il numero preciso

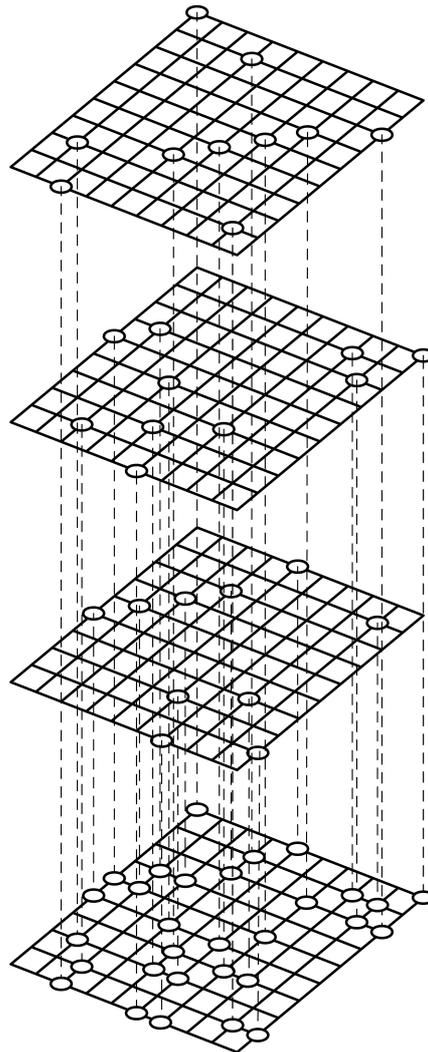


Figura 5.2: Esempio di sovrapposizione di ipercubi.

di punti necessari a garantire la precisione richiesta in fase di regressione.

Quello che proponiamo qui è una nuova tecnica di campionamento che, a partire dal LHS e dalle sue peculiarità, consenta un progressivo infittimento dell'insieme di campionamento mantenendo invariati i principi di selezione dei punti, consentendo in questo modo di incrementare la numerosità dell'insieme e quindi l'accuratezza della regressione dei dati.

Si è cercato di partire dal campionamento con l'ipercubo latino standard e di modificarlo. Riportiamo anche qui le caratteristiche che sono alla base di questa tecnica di campionamento :

- generazione di un campione casuale;
- ogni porzione in cui è divisa ciascuna variabile è rappresentata una sola volta;
- non ci sono punti dell'insieme di campionamento ripetuti, cioè osservazione ripetute.

Si è scelto quindi di optare per un insieme composto da ipercubi latini sovrapposti di eguale

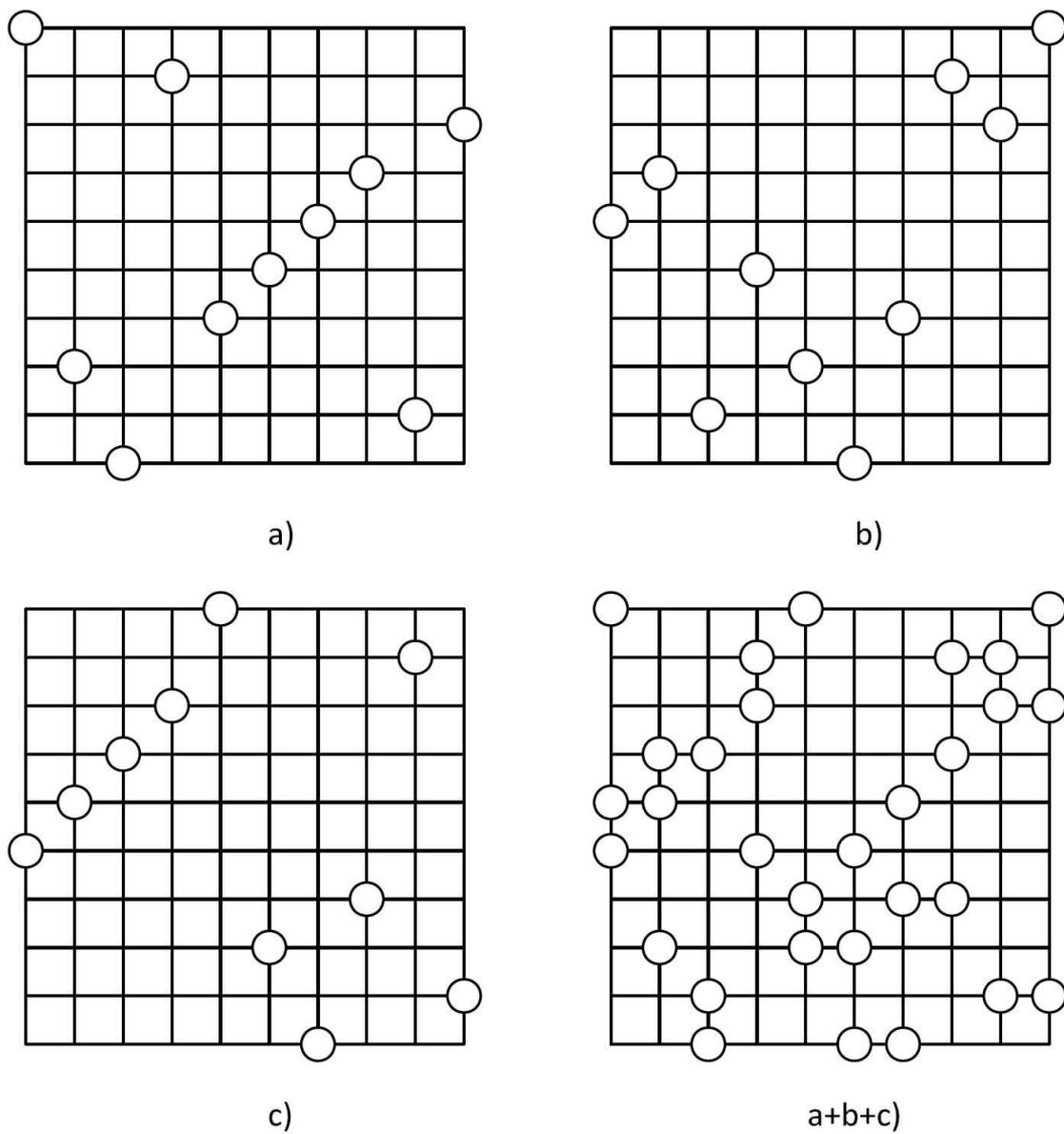


Figura 5.3: Risultato della sovrapposizione di tre ipercubi diversi a), b) e c) sovrapposti in a+b+c)

numerosità. Cioè scelto il valore n di cui si vuole incrementare di volta in volta la numerosità dell'insieme di campionamento, una volta generato un ipercubo latino di numerosità n è possibile sovrapporre a questo un altro ipercubo latino, sempre di numerosità n e generato casualmente sullo stesso spazio di scelta delle variabili, ma controllando che non vi siano tra il primo e il secondo elementi ripetuti; qualora si verificasse questa eventualità si procederebbe nella generazione di un nuovo ipercubo e nel controllo che non vi siano elementi in comune con quello già esistente, proseguendo così finché la verifica non ha dato esito positivo. Il nuovo insieme di campionamento di numerosità $2n$ è composto dal primo e dal secondo ipercubo insieme, e mantiene tutte le caratteristiche fondanti dell'ipercubo latino se si eccettua che ciascuna porzione di ciascuna variabile non è rappresentata una sola volta, bensì lo stesso numero di volte delle altre. E' possibile, per step successivi, incrementare progressivamente l'insieme di campionamento senza dover generare da capo il DOE e conservando i risultati degli esperimenti già fatti. Ad ogni passo si può effettuare nuovamente una regressione e valutare se il grado di precisione raggiunto è sufficiente.

Si osserva che questo sistema di campionamento non è in grado di indicare in anticipo quale è la numerosità del campione necessaria per raggiungere una determinata precisione in fase di regressione e quindi non è possibile fare stime in anticipo del tempo che può richiedere la creazione di una banca dati sufficientemente precisa. Il vantaggio è che è in grado di indicare la precisione raggiunta con il numero di simulazioni fatte fino ad un certo punto evitando di fare più simulazioni di quelle necessarie. Tutto ciò mantenendo le peculiarità del LHS, quindi massimizzando il numero di informazioni che si possono ottenere da ogni strato successivo di punti di campionamento che si sono aggiunti.

Resta da definire tuttavia il valore n dell'incremento dell'insieme di campionamento. Non è stato possibile indagare in questa sede l'influenza del valore dell'incremento n . Si noti però che ciascuno strato aggiunto al campione corrisponde ad un ipercubo latino di numerosità n . Quindi l'ampiezza dell'intervallo di variazione di ciascuna variabile è suddiviso in n parti uguali. Quindi è bene scegliere un numero n tale da suddividere questo intervallo in un numero adeguato di porzioni tale da garantire un'indagine adeguata dei valori dei vari parametri. Allo stesso tempo non deve però essere eccessiva, in modo da poter individuare con sufficiente precisione la numerosità ottima dell'insieme di addestramento, onde evitare simulazioni in più.

Un approccio di questo tipo comunque non sostituisce un'analisi critica del problema che si va ad affrontare. E' sempre necessario in fase preliminare capire il rapporto costi-benefici di un'operazione come la generazione di una banca dati; questo perchè potrebbe rivelarsi comunque un'operazione dispendiosa e non economica per quello che si deve fare.

5.3 Accesso alla banca dati

Una banca dati creata con il sistema proposto in questa tesi, fa uso di metamodelli. In questi metamodelli i risultati delle simulazioni sono raccolti e generalizzati. Sono quindi a disposizione delle *black box* dai quali, per ciascuna geometria individuata da un'insieme

di parametri, è possibile ricavare i valori prestazionali che caratterizzano la geometria in questione e che consentono un confronto con altre.

E' quindi possibile accedere direttamente a questi valori e utilizzarli per confrontare due o più geometrie tra di loro in modo da effettuare una scelta ponderata tra le stesse. L'accesso ai dati in questo caso è direttamente nel metamodello noti che siano i parametri che caratterizzano una geometria.

Avere a disposizione i dati nella forma di cui sopra consente un accesso agli stessi anche tramite gli algoritmi di ottimizzazione. Nel caso di un'ottimizzazione tradizionale gli algoritmi di ricerca dell'ottimo guidano la scelta delle simulazioni stesse e la zona dello spazio di indagine in cui effettuare la ricerca. In questo caso si sfruttano questi algoritmi per accedere alla banca dati e selezionare i risultati migliori. I valori di alcuni parametri possono essere vincolati all'interno di range di variazione, o addirittura a valori precisi, scelti in base a considerazione di diversa natura, come analisi di vincoli di ingombro o di ingegnerizzazione. Stabiliti questi vincoli un'algoritmo di ottimizzazione riesce a ritrovare in poco tempo i valori degli altri parametri a cui corrispondono le prestazioni migliori.

Capitolo 6

Progettazione di uno stadio di pompa sommersa

Le metodologie che sono state presentate nei capitoli precedenti sono state applicate alla progettazione di uno stadio di pompa sommersa multistadio. Per la progettazione dello stadio in questione si fa uso delle parametrizzazioni viste nel capitolo 2 che consentono un rapido interfaccia con i programmi CFD. Inoltre si è creata una banca dati di simulazioni al computer per caratterizzare la forma dei canali meridiani dei diffusori palettati a flusso misto delle pompe sommerse verticali. Per fare questo è stata predisposta una procedura per la gestione in automatico delle simulazioni CFD.

Si è proceduto innanzitutto con il dimensionamento della girante. Nota la geometria della girante, tramite la banca dati che è stata creata si è scelto il condotto meridiano del canale interstadio che risponde meglio alle esigenze di progetto, e infine si è provveduto al dimensionamento delle pale del condotto statorico che garantiscono la miglior interazione tra le due schiere di pale in movimento relativo.

6.1 Dati di progetto e vincoli produttivi

Il caso a cui è stata applicata la metodologia di progetto presentata fin qui, nasce da delle esigenze della ditta finanziatrice. Le normative che sono in fase di approvazione in sede europea impongono di progettare i nuovi modelli secondo standard elevati.

I dati di progetto prevedono che ciascuno stadio abbia una portata di massimo rendimento a $60 \text{ m}^3/\text{h}$ e una prevalenza per stadio alla portata di progetto di almeno 10 m di colonna d'acqua. Secondo le normative che sono in fase di approvazione il rendimento idraulico di tale stadio alla portata di massimo rendimento deve essere pari a 0.80. Inoltre è imposto un vincolo di ingombro radiale: tale stadio, infatti, è previsto per essere alloggiato in una pompa sommersa da 6 pollici, che significa che tali pompe devono poter essere calate in pozzi di diametro interno di 6 pollici. Tenuto conto di tutti i particolari necessari per la costruzione e l'assemblaggio di un prodotto finito l'azienda finanziatrice ha stimato in 118.5 mm il diametro massimo della parte idraulica dello stadio. Un ulteriore vincolo è rappresentato

dalla geometria della bocca di aspirazione della girante; infatti il diametro D_0 è vincolato a 63 mm dalla dimensione dell'anello di rasamento che realizza la tenuta tra la girante e la cassa, mentre il diametro al mozzo è vincolato a 28.3 mm per poter alloggiare una sede esagonale per gli alberi utilizzati.

Oltre ai dati di progetto e ai vincoli dimensionali imposti, ci sono ulteriori restrizioni da applicare alla macchina per quanto riguarda le tecniche di produzione. Infatti si prevede la produzione di tali macchine tramite stampaggio di materie plastiche per contenere i costi di produzione. In particolare i vincoli di forma imposti alla parte idraulica della macchina cambiano notevolmente secondo la forma degli stampi che si intende adottare, e la scelta del tipo di stampo che è possibile adottare impone un'attenta riflessione sui rapporti tra costi e benefici e qualche riga in più di spiegazione.

6.1.1 La realizzazione di pompe centrifughe tramite stampaggio di materie plastiche

Nella pratica industriale l'utilizzo delle materie plastiche sta invadendo sempre di più settori che una volta erano interessati principalmente da materiali di origine ferrosa nonostante le prestazioni siano peggiori, soprattutto in confronto agli acciai inox. Questo perchè nel momento in cui la ricerca è riuscita a compensare il naturale deficit strutturale delle materie plastiche nei confronti delle leghe ferrose, i pezzi prodotti in plastica si rivelano indubbiamente più economici e leggeri.

Tuttavia la realizzazione di componenti in materiale plastico impone delle pesanti restrizioni di forma. Infatti lo stampaggio impone dei vincoli in termini di spessori. Questi non possono essere eccessivamente sottili per non pregiudicare la resistenza meccanica e non ostacolare il flusso del materiale in fase di stampaggio. E d'altro canto non possono neppure eccedere per evitare problemi di ritiro in fase di raffreddamento del pezzo. Inoltre deve essere garantita l'assenza di sottosquadri nella forma nella direzione di estrazione delle varie parti dello stampo e gli angoli tra la superfici piane dello stampo e la direzione di movimento delle varie parti dello stampo in modo da garantire la possibilità di estrazione del pezzo. Tutti questi vincoli impongono notevoli restrizioni alla forma del pezzo che deve essere progettato e limitano notevolmente la complessità della forma che può essere prodotta.

La tipologia più semplice di stampo, detta anche 'apri e chiudi', che può essere adottata è quella composta di una matrice (la parte fissa dello stampo) e di un punzone che si sposta con movimento rettilineo, vedi figura 6.1. Questo tipo di stampo consente di realizzare solo pezzi semplici; nei quali non sono consentiti sottosquadri lungo la direzione di estrazione dello stampo. Per realizzare pezzi con geometria più complessa vengono utilizzati stampi con più parti mobili oltre al punzone. Sono parti più piccole che consentono di realizzare particolari nel pezzo con superfici in sottosquadro rispetto alla direzione di estrazione del punzone; tali parti prendono il nome di carri. I carri hanno generalmente un movimento laterale e devono essere studiati attentamente perchè nel loro movimento non devono intralciarsi a vicenda e con il punzone. I carri possono avere movimento rettilineo oppure, in soluzioni più costose movimento lungo una linea curva.

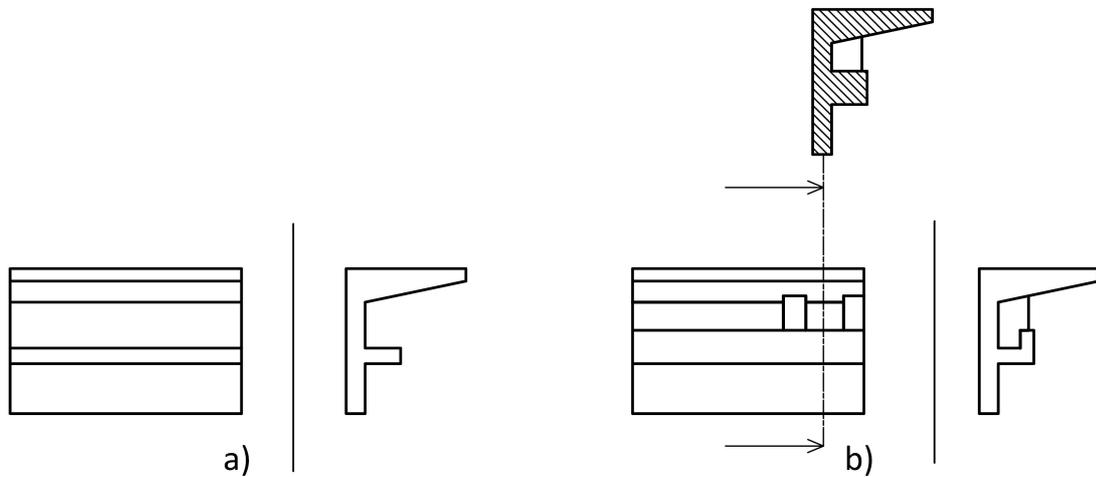


Figura 6.1: Differenti geometrie da produrre per stampaggio.

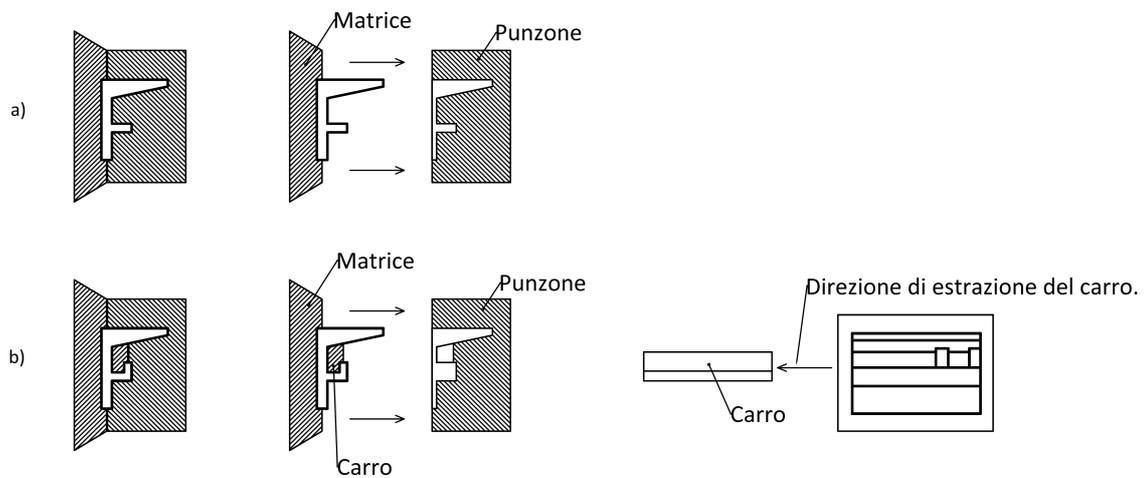


Figura 6.2: Diversi tipi di stampo in risposta alla diversa esigenza di produzione: a) stampo apri e chiudi per stampi senza sottosquadri in una direzione; b) stampo con carro mobile estratto lateralmente per pezzi complessi con sottosquadri.

Ovviamente questa soluzione è più costosa a livello produttivo, ma comunque consente di realizzare giranti a doppia curvatura per stampaggio con materiale plastico.

Per meglio comprendere le diverse forme che devono essere prodotte in modi differenti in figura 6.1 sono proposti due pezzi simili in forma, ma con diversi livelli di complessità. Denominato con **a)** è un pezzo semplice, senza sottosquadri in più direzioni; il pezzo denominato con **b)** è simile nella forma, ma presenta dei particolari che fanno sì che non esistano direzioni senza superfici in sottosquadro per il pezzo in questione. Nella successiva figura 6.2 vengono

presentate le rispettive soluzioni per lo stampaggio dei due diversi pezzi. Il pezzo **a)** può essere stampato con un semplice stampo apri e chiudi, composto da matrice e punzone. Il pezzo, denominato con **b)**, può essere prodotto con uno stampo con un carro estratto lateralmente.

Nella produzione delle giranti delle pompe lo stampo 'apri e chiudi' può essere utilizzato solamente per giranti con geometria a semplice curvatura. In questi casi le pale vengono prodotte di pezzo con il disco che va a realizzare la geometria del mozzo (che prende così il nome di disco portapala); con un altro pezzo viene realizzato il coperchio che va a costituire l'altro disco. Tale tipologia di stampo è molto semplice e necessita di presse poco elaborate; quindi è di fatto il metodo di stampaggio più economico, al punto che spesso per le pompe si preferisce a volte realizzare giranti a semplice curvatura anche per macchine con numero tipico che richiederebbe giranti a doppia curvatura.

Una soluzione di stampaggio con i carri consente invece di realizzare, pur con limitazioni di forma, giranti a doppia curvatura. Per la svergolatura tipica della parte iniziale delle pale a doppia curvatura queste pale presentano un sottosquadro nella direzione assiale e quindi non possono essere realizzate di pezzo con il disco portapala con uno stampo di tipo 'apri e chiudi'. E' possibile realizzare queste pale di pezzo con il disco portapala per mezzo di stampi con dei carri mobili a movimento laterale. In tale modo la faccia della pala che non presenta sottosquadri per il punzone viene definito da quest'ultimo, mentre l'altra faccia viene definita dalla superficie di un carro che viene estratto lateralmente in una direzione che non presenta sottosquadri.

E' interessante fornire un rapporto indicativo del costo delle diverse tipologie di stampo. Fatto 1 il costo di uno stampo semplice di tipo 'apri e chiudi', uno stampo con carri mobili a movimento rettilineo costa 3 e uno con carri mobili con movimento curvo costa 5.

Le pale dei canali di ritorno a flusso centripeto solitamente vengono progettate a semplice curvatura rispetto alla direzione assiale e quindi sono stampabili di pezzo con il disco portapala. Viceversa per le pale dei diffusori a boccia valgono le considerazioni fatte per le giranti a doppia curvatura, solo che la causa di un eventuale sottosquadro in direzione dell'asse della macchina non è più da ricercarsi nella svergolatura della pala in corrispondenza del bordo di ingresso, ma nell'overturning in corrispondenza del bordo di uscita o nel ricoprimento tra le pale in corrispondenza del bordo di ingresso: una diffusore con pale con direzione assiale in corrispondenza del bordo di uscita non presenta sottosquadri in direzione dell'asse della macchina e può essere facilmente realizzato di pezzo con il disco portapala che definisce la superficie del mozzo. Tuttavia in questo tipo di diffusore viene a mancare la totale rimozione della componente tangenziale della velocità assoluta a causa del gradiente di pressione interpalare. Una pala del diffusore con overturning presenta un bordo di uscita piegato oltre la direzione dell'asse della macchina e quindi una porzione di superficie in sottosquadro rispetto alla direzione assiale. Tuttavia, anche in questo caso questi pezzi possono essere prodotti con stampi con carri mobili. In questo caso un carro definisce entrambe le facce ma di pale successive.

Le limitazioni di carattere produttivo devono essere tenute in considerazione già in fase di progettazione: non è economico realizzare geometrie della parte idraulica che poi non possono

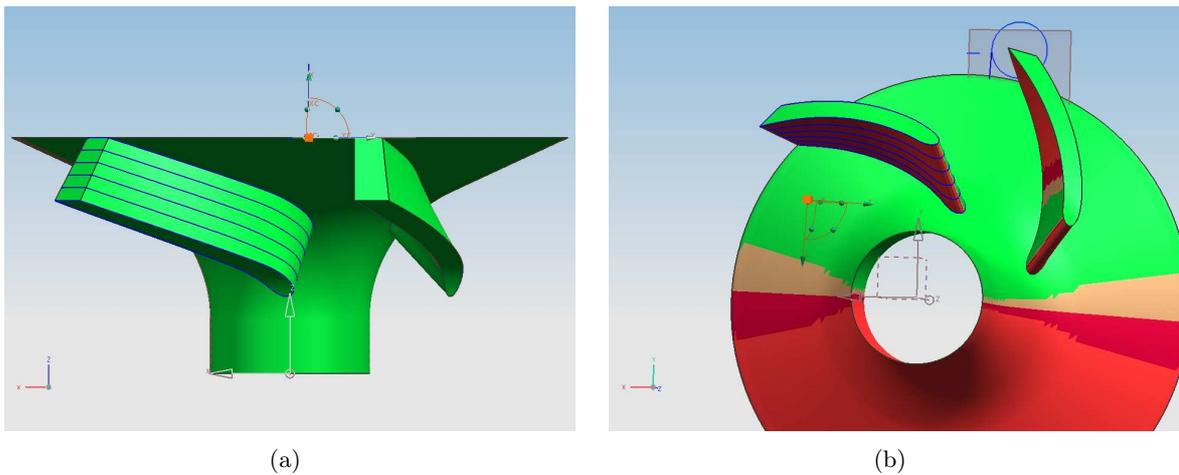


Figura 6.3: Visualizzazione degli angoli di sottosquadro sulle facce di una pala di una girante a semplice curvatura

essere realizzate con le tecniche di produzione disponibili in azienda. Questo rappresenta il principale motivo della difficoltà di applicare tecniche di progettazione inversa nell'ambito aziendale, dove la forma della pala risponde solo alle esigenze fluido dinamiche. Viceversa le parametrizzazioni proposte in questa tesi partono dalla definizione della geometria della pala per consentire un diretto controllo sulla forma e un rapido interfaccia con programmi di modellazione solida che consentono di verificare la realizzabilità.

In figura 6.4 si può notare le tre fasi principali della realizzazione tramite stampaggio di cui si è detto. Nella fase denominata con **a** il punzone si allontana in direzione assiale dal pezzo e lascia libero lo spazio per i movimenti dei carri, che con un movimento laterale (**b+c**) si sganciano dal pezzo e lo liberano completamente.

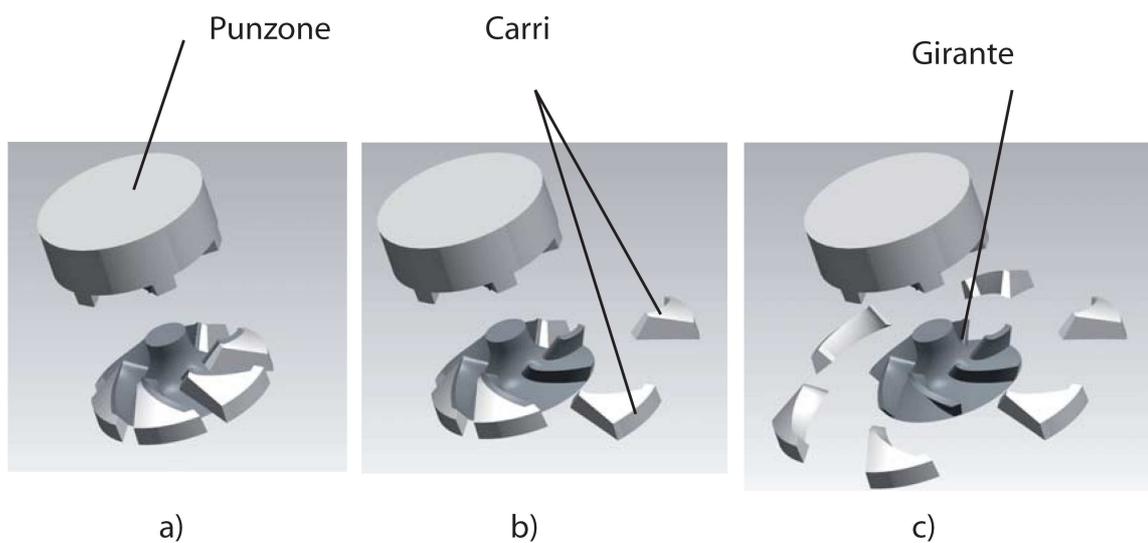


Figura 6.4: Realizzazione tramite stampaggio delle girante a doppia curvatura.

6.2 Analisi fluidodinamiche

Per quanto riguarda le simulazioni fluidodinamiche si è fatto uso del pacchetto commerciale CFD Ansys-CFX11.

La scelta di utilizzare tale pacchetto commerciale è motivata dall'esperienza del gruppo di ricerca. Tale programma infatti garantisce un approccio completo e intuitivo all'analisi e alla sintesi delle turbomacchine, mantenendo comunque la possibilità di personalizzare i parametri sia fluidodinamici che numerici della simulazione. Ciascun programma di questo pacchetto è dedicato ad una delle operazioni necessarie per effettuare una simulazione fluidodinamica. Tali operazioni si organizzano secondo uno schema sequenziale, quindi anche l'utilizzo di tali programmi è sequenziale. L'input di ogni programma è l'output del precedente.

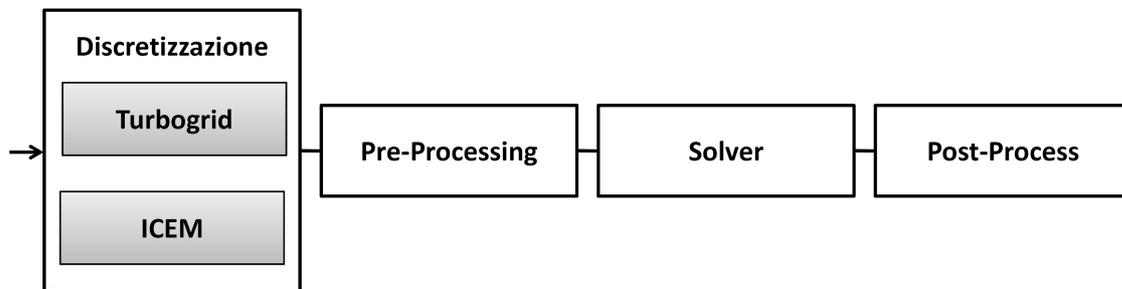


Figura 6.5: Diagramma a blocchi delle operazioni da svolgere per portare a termine una simulazione.

Nella figura 6.5 è riassunta la sequenza delle operazioni necessarie a portare a termine una simulazione CFD e i programmi dedicati.

Tali programmi, oltre ad essere disponibili da interfaccia grafica, sono anche utilizzabili in modalità batch (letteralmente 'in serie'). Con questa modalità è possibile richiamare i vari programmi senza attivare la GUI (graphic user interface), ma avviando il programma tramite un file .bat e indicando al programma la serie di istruzioni da eseguire in un file di script, cioè con un file di testo.

6.2.1 Meshatura

Nel dettaglio, per portare a termine la simulazione si deve partire dalla geometria e dai dati di funzionamento dell'organo che si intende analizzare. Il primo passo consiste nella discretizzazione della geometria; questa fase prende anche il nome di meshatura. I programmi del pacchetto che vengono utilizzati a tal fine sono CFX-Turbogrid e ICEM.

Turbogrid è un programma specifico. È creato a posta per meshare palettature di turbomacchine e legge come input dei file nei quali la geometria dei condotti meridiani e della palettatura delle macchine sono ricostruiti per punti. A partire da questi il programma è in grado di ricostruire la geometria e di discretizzare l'organo in analisi con mesh composte da celle esaedriche. Per le discretizzazioni dei volumi attorno alle pale, il programma utilizza

alcune maschere predefinite che si adattano alle varie geometrie sotto analisi e che prendono il nome di *topology*.

ICEM è un programma più generico di meshatura, che consente di meshare qualunque geometria con la mesh più adatta. La geometria può essere letta come file di punti, oppure importata direttamente sotto forma elettronica da un modellatore solido. Il programma inoltre mette a disposizione anche alcune operazioni per la modifica o la creazione della geometria. Le mesh possono essere create in automatico, facendo uso di alcune opzioni che consentono posizionare i nodi della griglia secondo criteri matematici. In alternativa è possibile definire con precisione le dimensioni della griglia lungo spigoli e superfici del volume solido e utilizzare delle funzioni di densità per i volumi in modo da avere un controllo diretto.

Generalmente sono possibili più strade che per ottenere una mesh di un volume fluido che si differenziano per qualità della discretizzazione e per il tempo necessario per ottenere la discretizzazione. Come in qualunque operazione, una soluzione qualitativamente migliore richiede strumenti più raffinati e un impiego di tempo superiore.

I nodi di una discretizzazione sono disposti ai vertici di celle; la *connettività* di un nodo è definita dal numero delle celle che hanno quel nodo come vertice. Come regola generale nella discretizzazione di volumi per il calcolo fluidodinamico sono possibili due soluzioni: si parla di *mesh strutturata* quando la connettività dei nodi di una mesh è costante, mentre si parla di *mesh non strutturate* quando la connettività di ciascun nodo varia. A livello numerico le mesh strutturate sono da considerarsi preferibili. Le mesh strutturate consentono inoltre un controllo più preciso delle dimensioni delle celle, dei rapporti tra le dimensioni di celle contigue e degli infittimenti necessari nelle zone critiche.

Le mesh si differenziano anche per la forma delle celle. La maggior parte delle mesh di volumi utilizzate in fluidodinamica fanno uso di celle di forma tetraedrica o esaedrica. Una parte minore di casi vengono discretizzati con mesh miste o con elementi di altro tipo. Nelle simulazioni di volumi di fluido in movimento sono da preferire le mesh esaedriche; infatti la forma delle celle consente di orientare le facce dell'esaedro in direzione della corrente, cosa che garantisce una maggiore stabilità numerica. In linea di principio è possibile ottenere una mesh strutturata sia con elementi esaedrici che con elementi tetraedrici, tuttavia le mesh strutturate sono sempre create con celle di tipo esaedrico perché è molto più semplice controllare la connettività dei nodi con questo tipo di celle.

6.2.2 Pre-Processing

I file contenenti mesh e geometria del caso sotto analisi sono poi letti dal programma successivo della sequenza: Pre. Tale programma è adibito al pre-processamento della simulazione, cioè quell'insieme di istruzioni che consentono di definire i parametri necessari per ottenere una simulazione aderente al caso in analisi e portare a convergenza tale simulazione. In questa fase vengono infatti definite le condizioni al contorno, le condizioni di inizializzazione, i modelli del fluido che devono essere utilizzati e i criteri di convergenza. Una volta impostati tutti questi parametri è possibile portare a termine la simulazione.

6.2.3 Solving

Il file prodotto dal Pre viene successivamente utilizzato dal Solver. Questo è in realtà il cuore della simulazione; questo programma risolve per via iterativa in forma conservativa con il metodo dei volumi finiti le equazioni di Navier-Stokes. Anche in questo caso i risultati sono contenuti in un file che diventa l'input per il programma successivo; l'estensione dei file dedicati a memorizzare i risultati delle simulazioni è .res. Tale fase è quella dove il programma richiede minor intervento per l'operatore.

6.2.4 Post-Processing

I risultati ottenuti dalla simulazione vengono infine interpretati in fase di Post-Processing. Il programma dedicato a questo scopo consente di visualizzare il campo di moto all'interno del volume di fluido in analisi. I risultati prodotti dal solutore consistono nei valori di tutte le variabili che servono a rappresentare il campo di moto del fluido. I valori sono immagazzinati nel file di risultati per ciascun nodo della discretizzazione del volume di fluido. Tali programmi non sostituiscono un'approfondita e completa conoscenza della fisica dei fenomeni che governano il complesso funzionamento di queste macchine, ma bensì integrano tale conoscenza, senza la quale non sarebbe possibile interpretare correttamente i risultati di tali simulazioni e non se ne ricaverebbe alcuna informazione utile in fase di progettazione.

A partire da questi risultati è possibile innanzitutto capire se la simulazione è stata impostata nel modo più corretto, cioè se la discretizzazione operata in fase di meshatura è precisa o se le condizioni al contorno sono state impostate bene. Inoltre un'analisi attenta consente di capire in che termini una simulazione è confrontabile con altre, come impostare altre simulazioni in modo che siano confrontabili con quella presente e quali soluzioni è bene mettere in essere per migliorare la soluzione attualmente in analisi. In pratica la fase di post-processing fornisce il necessario feed-back all'operatore per l'utilizzo più consono del programma stesso e lo sviluppo di nuove soluzioni.

Per quello che concerne l'utilizzo delle simulazioni nella progettazione la fase di post processing è la più importante, quella che richiede la maggiore esperienza da parte dell'operatore. L'interpretazione dei risultati deve essere fortemente critica. La corretta interpretazione dei risultati è fondamentale per il prosieguo del lavoro sia che il fine della simulazione sia l'indagine approfondita e la comprensione dei fenomeni che governano il comportamento del fluido, sia che si tratti di uno step intermedio all'interno di un più ampio percorso di progettazione.

Il programma dedicato alla fase di post-processing consente di estrapolare questi valori su tutto il volume. Successivamente con opportune visualizzazioni è possibile rappresentare il campo di moto e tutte le variabili coinvolte. Oltre alle visualizzazioni più intuitive vengono messe a disposizione dell'operatore anche altri tipi di visualizzazioni, più complessi e di difficile interpretazione, ma che consentono di ricavare ulteriori importanti informazioni; c'è altresì la possibilità di calcolare nelle zone più disparate i valori delle grandezze scalari, al fine di quantificare anche numericamente il funzionamento dell'organo in analisi.

Le forme di rappresentazione dei risultati sono una sorta di linguaggio che viene utilizzato

dal programma, e che è necessario comprendere a fondo per interpretare correttamente la fenomenologia della porzione di fluido simulata. Anche se cercano di essere il più intuitivi possibili tali forme di rappresentazione nascondono delle insidie e devono essere utilizzati spesso in modo combinato. Oltre alle rappresentazioni grafiche è necessario fare uso di indagini quantitative sulle grandezze coinvolte nel campo di moto per completare l'analisi delle informazioni fornite dalla simulazione. Grazie a queste analisi quantitative, oltre a integrare le informazioni messe a disposizione dalle visualizzazioni, è possibile calcolare i dati generali del funzionamento dell'organo, cioè quelle grandezze macroscopiche come rendimenti o incrementi di energia che si usano per classificare, analizzare e confrontare il comportamento delle macchine.

6.3 Esecuzione della progettazione

La fase operativa della progettazione di uno stadio di pompa verticale sommersa è complicata a causa della complessità degli organi da progettare e della complessità dell'interazione in funzionamento tra i due.

La normativa europea impone il calcolo dei rendimenti della pompa nella sua interezza, quindi conteggiando anche perdite di natura meccanica e non solo di natura idraulica. I target in sede di progettazione sono rappresentati da un valore di prevalenza e da un rendimento complessivo della macchina.

L'utilizzo dei codici di calcolo CFD, consente di stimare il rendimento idraulico della pompa e lo scambio di energia; nota la geometria della parte idraulica una simulazione CFD dettagliata della pompa in tutti i suoi particolari consente di ricavare lo scambio di energia e il rendimento idraulico della macchina. Questo, moltiplicato per il rendimento meccanico stimato della pompa consente di calcolarne il rendimento complessivo e verificare se questo è pari a quello che si vuole raggiungere. Queste simulazioni richiedono di discretizzare non solo le parti principali delle macchine, come palettature della girante e diffusore, ma anche i canali di trafilamento e tutti i dettagli costruttivi, di fatto andando a simulare la macchina nella configurazione di collaudo. Tali simulazioni richiedono tempi notevoli che male si sposano con la progettazione, dove si deve raggiungere la soluzione che soddisfa gli obiettivi prefissi per tentativi successivi che devono essere portati a termine in tempi contenuti.

Si è preferito distinguere quindi tra simulazioni da utilizzare in sede di progettazione, che analizzano una geometria semplificata senza i canali di trafilamento, e delle simulazioni di verifica, nelle quali tutte le parti della macchina sono discretizzate e possono essere utilizzate per verificare l'entità precisa dello scambio di energia e delle perdite di natura fluidodinamica all'interno della macchina.

6.3.1 Fasi della progettazione

La progettazione esecutiva di questa macchina si è composta essenzialmente di tre fasi principali. Dopo aver scelto il layout di massima della parte idraulica, si è proceduto alla progettazione della girante della macchina. Parallelamente alla progettazione della girante si è

proceduto con la realizzazione di una banca dati di simulazioni CFD che consenta di caratterizzare al meglio la forma del canale meridiano del diffusore secondo la parametrizzazione del canale meridiano vista nel paragrafo 2.3. I risultati ottenuti sono stati utilizzati, come vedremo, per la scelta corretta dell'ingombro assiale del diffusore. La banca dati che è stata realizzata è stata pianificata in modo tale da conglobare anche geometrie di canali diffusori di differenti taglie in modo tale da poter essere utilizzata anche per la progettazione di altre macchine simili.

Oltre alla realizzazione della banca dati per la caratterizzazione dei canali meridiani dei diffusori le fasi che risultano della progettazione della macchina in questione sono schematicamente:

- analisi dei dati progetto e dei vincoli di ingombro;
- scelta del layout della macchina (diametri esterni della girante, larghezza aspirazione del diffusore);
- dimensionamento della girante: verifica fluidodinamica della girante;
- scelta del canale meridiano del diffusore;
- dimensionamento delle pale del diffusore;
- verifica fluidodinamica del comportamento e delle prestazioni dei due organi nel loro funzionamento accoppiato.

6.4 Progettazione della girante

La progettazione delle pompe sommerse multistadio è fortemente vincolata dall'ingombro radiale delle macchine. In questo caso, come detto sopra, l'ingombro radiale massimo previsto per la parte idraulica di queste macchine è di 118.5 mm. E' necessario quindi studiare attentamente in fase preliminare la conformazione meridiana di uno stadio.

I dati di progetto per la macchina sono $H=10$ m, $Q=60$ m³/h. Inoltre queste macchine sono previste per essere azionate da un motore elettrico sommerso a due poli. La velocità di rotazione dell'albero quindi può essere fissata a 2900 giri/min. Sotto queste condizioni si ottiene una macchina con numero tipico di macchina pari a $K=1.26$.

I diagrammi statistici e i manuali di progettazione suggeriscono per questo numero tipico di macchina una geometria per la girante di tipo radiale con pale a doppia curvatura. I diagrammi statistici indicano per questo numero tipico di macchina un numero tipico di pressione $\psi = 0.336$ e un numero tipico di flusso $\phi = 0.155$. Questi valori suggeriscono una girante con flusso all'uscita radiale e diametro esterno pari a 112.5 mm. Questa conformazione è quella che garantisce i rendimenti più elevati. Tuttavia una girante radiale con questo diametro esterno non è compatibile con i vincoli di ingombro radiale imposti.

La figura 6.6 mette in luce come non ci sia spazio sufficiente tra la bocca di uscita della girante a flusso puramente radiale e la parete della camicia esterna dello stadio. Non c'è spazio

Tabella 6.1: Tabella riassuntiva dei dati di progetto e dei vincoli dello stadio di pompa sommersa.

Parametri	Valore
Portata	60 m ³ /h
Regime di rotazione	2900 rpm
Prevalenza	10 m
Diametro massimo dello stadio	118.5 mm
Diametro bocca di aspirazione	63 mm
Diametro mozzo	28.3 mm
Numero tipico di macchina K	1.26
Numero di pressione suggerito ψ	0.336
Numero di flusso suggerito ϕ	0.155

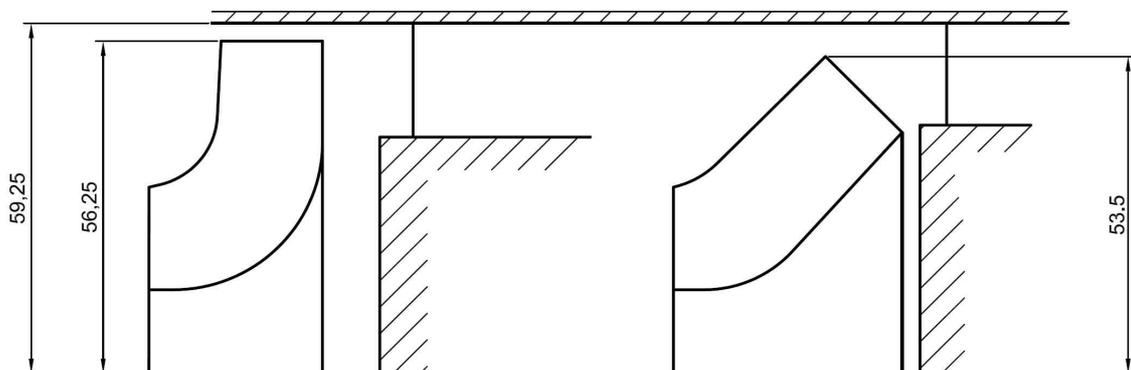


Figura 6.6: A sinistra: dimensioni della geometria di girante radiale alloggiata negli ingombri radiali imposti; a destra: possibile esempio di soluzione a flusso misto per la stessa dimensione radiale.

per inserire un diffusore radiale in uscita alla girante. Non si è ritenuto opportuno progettare una girante con diametro esterno ridotto che garantisca lo scambio energia sufficiente, infatti sarebbe necessario un angolo costruttivo della pala in corrispondenza del bordo di uscita molto alto che porterebbe la macchina ad una forte instabilità funzionale.

In questo caso è necessario optare quindi per una girante con flusso meridiano misto, accoppiata con un diffusore-convogliatore a flusso misto, anche se i dati di progetto sug-

gerirebbero una girante radiale per ottenere i massimi risultati in termini di prestazioni.

Per determinare l'inclinazione rispetto alla direzione assiale del condotto meridiano della girante è necessario effettuare uno studio del layout meridiano della macchina e dell'interazione tra il canale meridiano inclinato della girante e l'ingresso assiale nel diffusore palettato, in modo da stabilirne la corretta inclinazione. Per effettuare tale studio si è scelto di disegnare vari possibili condotti meridiani, accoppiati con un tratto assiale di canale statorico a schematizzare la bocca di ingresso del diffusore.

La geometria del canale meridiano è stata disegnata nel seguente modo: scelta l'inclinazione della linea media del condotto meridiano e imposta un'area di attraversamento netta alla bocca d'uscita della girante si disegna la linea media della girante con un arco di cerchio seguito da un tratto rettilineo inclinato. Si disegnano per punti i profili di mozzo e di corona assegnando un'andamento delle aree di attraversamento normali alla linea media lineare tra quello l'area di ingresso e quella della bocca d'uscita. Tali profili vengono poi approssimati con un arco di circonferenza nella zona di ingresso e un tratto rettilineo nella parte terminale in modo da ottenere una definizione più semplice delle generatrici.

Tali pompe montano giranti flottanti, in grado di muoversi assialmente, quindi la posizione relativa tra girante e diffusore non è costante, tuttavia è possibile stimare un valore medio della distanza minima tra girante e diffusore pari a 3 mm. I condotti sono stati disegnati seguiti da una schematizzazione del bordo di ingresso di un diffusore a valle. Il diffusore a valle è dimensionato in modo che l'area di attraversamento sia pari all'area di attraversamento allo scarico della girante. Questa non è una scelta definitiva, ma consente, all'osservazione dei vari schemi, di capire se per garantire un corretto imbocco del diffusore da parte della corrente in uscita dalla gira si deve allargare o restringere la bocca di aspirazione del diffusore. In altre parole se è necessario accelerare o decelerare la corrente nel traferro tra la girante e il diffusore.

Nella figura 6.7 si vede lo studio che è stato portato avanti per le giranti a flusso misto, con linea media inclinata di 45° rispetto alla direzione assiale. La prima riga di condotti si riferisce a giranti con diametro massimo pari a 108 mm, mentre la seconda riga schematizza tutte giranti con diametro massimo pari a 104 mm. Da sinistra a destra, su ciascuna riga, calano i valori delle aree di attraversamento in uscita. Osservando tali esempi è possibile valutare quali soluzioni sono caratterizzate da una geometria più dolce e quali impongono un raccordo tra la direzione assiale e quella radiale più brusco. Inoltre è possibile individuare quali modifiche devono essere apportate alla bocca di ingresso del diffusore accogliere correttamente il flusso e che tipo di effetto potrebbe esserci e quindi cercare di capire quale potrebbe essere la configurazione che consente l'interazione più corretta tra girante e diffusore. In figura 6.8 è possibile vedere un grafico analogo per le giranti inclinate di 60° rispetto alla direzione assiale.

Tali schemi possono anche essere interpretati per mezzo delle grandezze adimensionali tipiche delle turbomacchine. Tra le righe varia il numero di pressione ψ (in particolare è più alto per la seconda riga dove i diametri sono più bassi). Mentre su ciascuna riga, muovendosi da sinistra verso destra diminuisce l'area di attraversamento in uscita della girante quindi aumenta il numero di flusso ϕ .

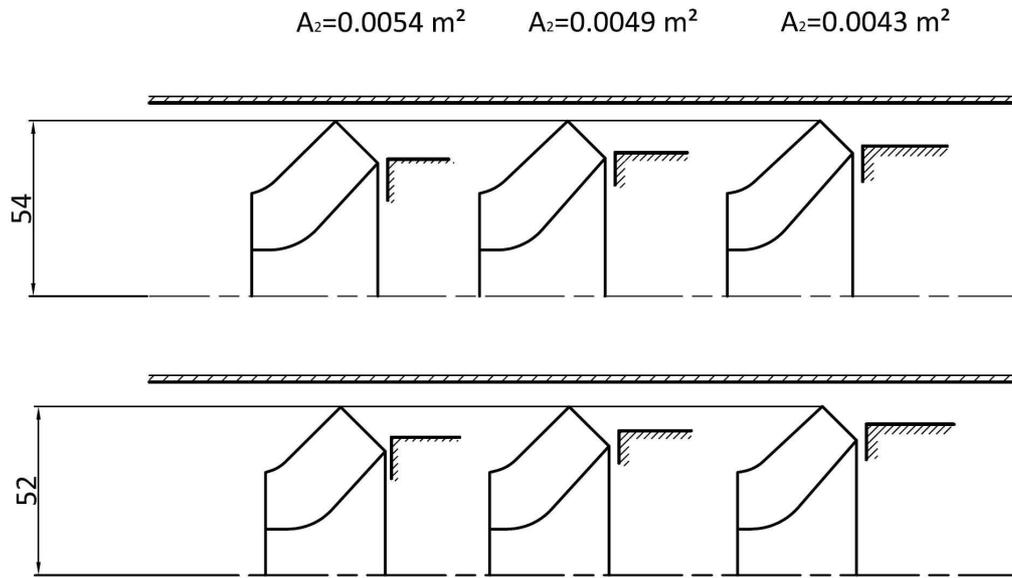


Figura 6.7: Studio di diversi canali meridiani della giranti a flusso misto con linea media del canale meridiano inclinata di 45° rispetto alla direzione assiale.

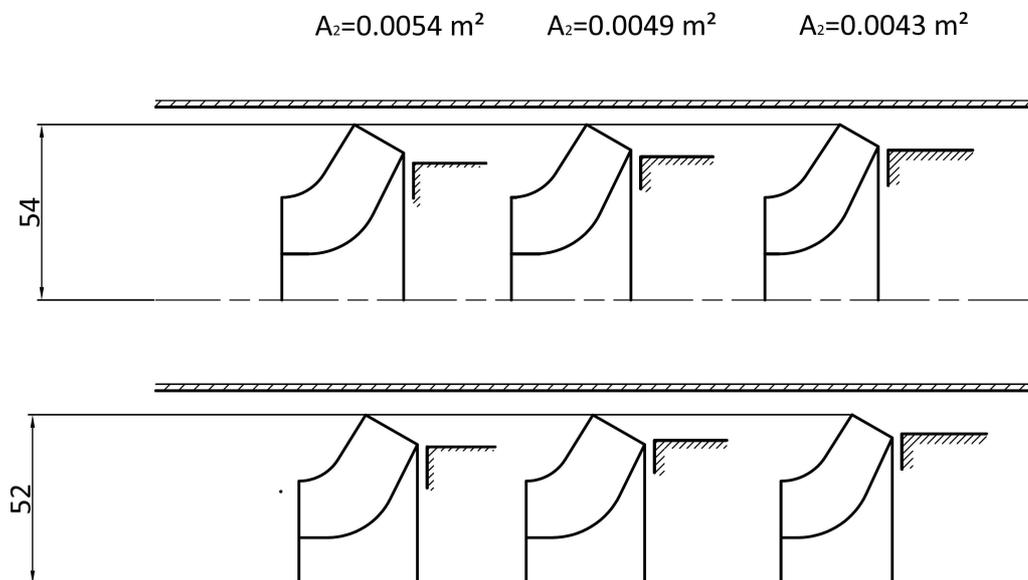


Figura 6.8: Studio di diversi canali meridiani della giranti a flusso misto con linea media del canale meridiano inclinata di 60° rispetto alla direzione assiale.

Scelto il rendimento condotto meridiano della girante si è proceduto con il dimensionamento delle pale della girante. Per effettuare la progettazione della girante grazie ai codici

di calcolo CFD è necessario stimare la prevalenza che deve elaborare la girante al netto delle perdite dovute all'interazione con gli organi statorici, misurata tra due superfici di riferimento subito a monte e subito a valle della girante. Come già detto il rendimento idraulico atteso per questa macchina è pari all'80% e la prevalenza minima da garantire è pari a 10 m di colonna d'acqua. Non si hanno a disposizione dati riguardanti la ripartizione delle perdite tra girante e canale interstadio in turbomacchine analoghe a quella in progettazione. Tuttavia da esperienze precedenti è noto che giranti con numero tipico di macchina intorno all'unità e con flusso regolare, realizzano un rendimento idraulico quantificabile intorno al 92-93%. Da questo è possibile ricavare la quotaparte delle perdite di natura idraulica che si attende negli organi statorici della macchina in questione; sia H l'energia del singolo stadio, e H_t l'energia teorica euleriana della macchina, tale che $\eta_{id} = \frac{H}{H_t} = 0.8$ dove η_{id} è il rendimento idraulico complessivo. Sia quindi ΔH_{gir} l'altezza corrispondente alle perdite nella girante e ΔH_{diff} analogamente per il diffusore. Vale allora la seguente relazione:

$$H = H_t - \Delta H_{gir} - \Delta H_{diff};$$

sia quindi $\eta_{gir} = 1 - \frac{\Delta H_{gir}}{H_t}$ il rendimento idraulico della girante; riarrangiando la formula precedente si ricava:

$$\begin{aligned}\Delta H_{diff} &= H_t - H - \Delta H_{gir} = \\ &= H_t - (1 - \eta_{gir})H_t - H = \\ &= \eta_{gir}H_t - H =\end{aligned}$$

e quindi riassumendo la seguente formula

$$\Delta H_{diff} = H \left(\frac{\eta_{gir}}{\eta_{id}} - 1 \right)$$

consente una stima delle perdite negli organi statorici. Quindi la prevalenza attesa, come indicazione di primo tentativo, da una simulazione CFD della sola girante è pari a 11.5 m.

Per raggiungere questo risultato sono stati necessari numerosi tentativi successivi e quindi numerose analisi CFD. Non è possibile in questa sede ricostruire il percorso che ha portato ad individuare la geometria più consona della girante, ma ci si limita a mostrare le peculiarità della soluzione che è stata proposta e le osservazioni salienti che hanno portato a questo risultato.

Per quello che riguarda le simulazioni della sola girante si è simulata una sola periodicità della macchina quindi un solo condotto interpolare; per effettuare la simulazione nel modo più corretto è stato discretizzato un tratto di adduzione a pareti parallele e un tratto di condotto anulare a sviluppo assiale allo scarico per ricreare nel modo più corretto lo condizione di flusso a valle della girante nella zona in cui si andrà a posizionare il bordo di ingresso del diffusore. La meshatura è stata realizzata con il programma Ansys-Turbogrid. Si sono realizzate delle mesh con un numero complessivo di nodi di circa 250000 unità in modo da garantire una corretta discretizzazione della forma geometrica. E' stato utilizzato l'opzione O-grid per

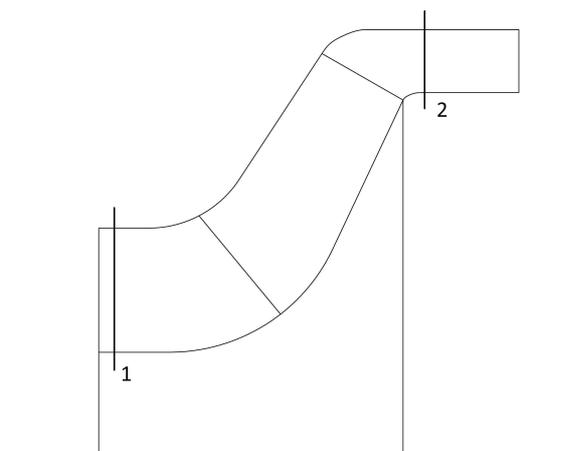


Figura 6.9: Schema del volume discretizzato e superfici di riferimento utilizzate per il computo dello scambio di energia delle simulazioni CFD.

creare dei layer di nodi più fitti intorno alla superficie della pala. Per ciascun caso che è stato analizzato la geometria era tale da garantire inoltre un angolo minimo tra le facce dei vari elementi di minimo 25° . Il modello di turbolenza che è stato utilizzato per le simulazioni è il modello SST (Shear Stress Transport); il vantaggio di tale modello è che non è necessaria una definizione della mesh estremamente accurata, è sufficiente garantire una y^+ della mesh inferiore a 200 in prossimità delle pareti per avere una corretta definizione dello strato limite, così come suggerito in [52]. L'infittimento utilizzato per questa mesh è tale da consentire di raggiungere sempre questo limite senza controlli a posteriori sui risultati della simulazione. Si è ritenuto quindi tali scelte di modellazione sufficienti a garantire una corretta discretizzazione del problema.

Nella figura 6.9 si può notare uno schema della vista meridiana del volume simulato, con la proiezione della pala e lo schema delle zone ausiliarie all'adduzione e allo scarico. Tali zone sono state modellate per consentire di analizzare nel modo migliore la corrente all'uscita della girante. Sono state eseguite delle prove per capire se fosse necessario modellare uno scarico con un'estensione assiale superiore, ma l'estensione scelta per queste simulazioni si è mostrata sufficiente a visualizzare gli effetti di miscelamento nelle scie delle pale e non pregiudica la stabilità numerica della simulazione.

La forma del raccordo tra il bordo di uscita della girante e il tratto assiale corrispondente alla posizione del bordo di ingresso del diffusore è stato semplificato rispetto alla forma che assume nella realtà, dove sono presenti i canali di trafileamento. Si è scelto di semplificare in tale modo la forma del canale meridiano nella progettazione iniziale della girante perchè si è ritenuto fosse quella più semplice da modellare consentendo comunque di ricavare tutte le informazioni necessarie sullo scambio di energia.

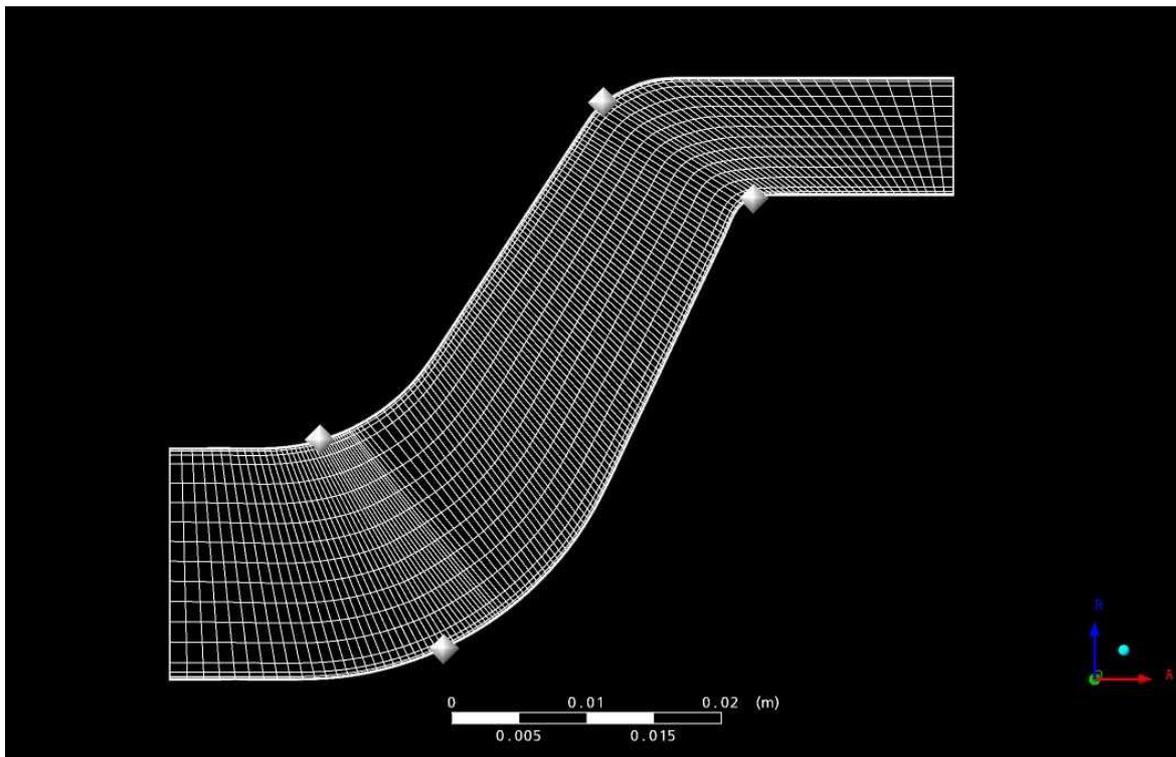


Figura 6.10: Visione, sul piano meridiano, dei nodi della discretizzazione del volume della pala e delle zone ausiliarie di adduzione e di scarico.

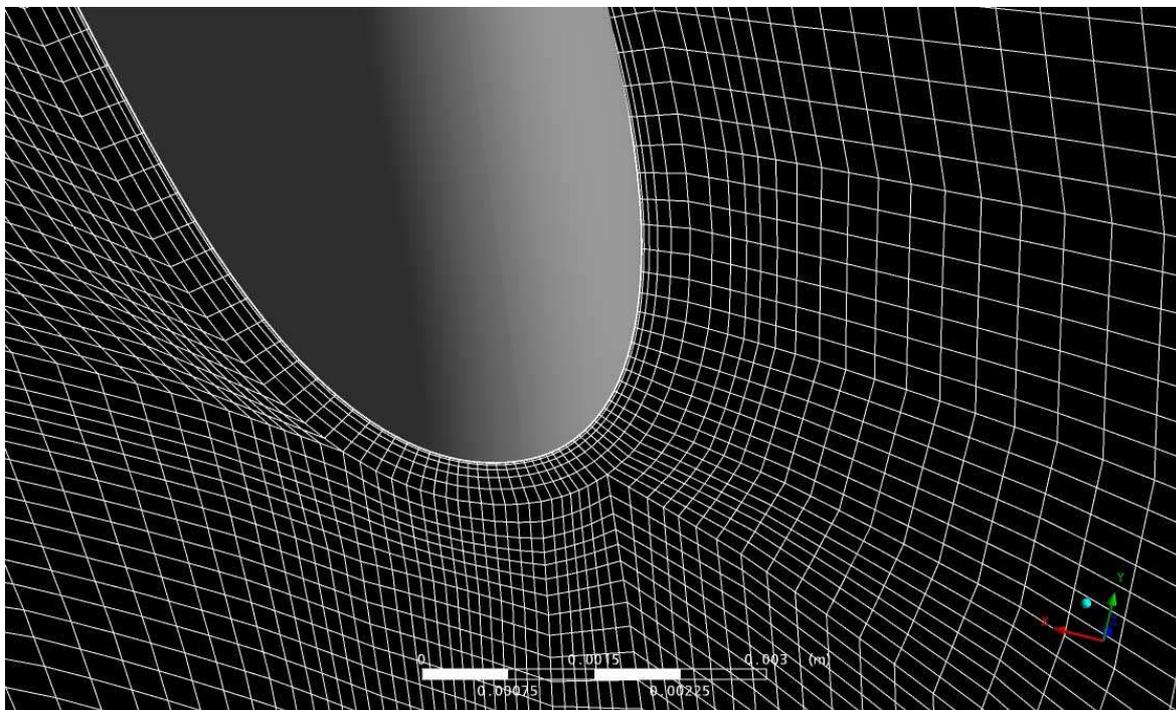


Figura 6.11: Particolare della distribuzione dei nodi della discretizzazione attorno alla pala.

Spessore palare

La prima osservazione che è necessario fare, riguarda lo spessore delle pale. Come già ricordato nel paragrafo 6.1.1 queste macchine sono prodotte materiali plastici tramite stampaggio; gli

spessori massimi e minimi sono critici nello stampaggio e soggetti a restrizioni rigorose. Nella soluzione proposta in questo lavoro lo spessore delle pale è stato imposto costante e pari a 4 mm, assegnato sul piano conforme. Il valore dello spessore di stampaggio non può scendere al di sotto di questo perchè ne risulterebbe pregiudicato il flusso di materiale plastico all'interno dello stampo e anche la resistenza meccanica della pala. Come valore massimo delle spessore non si può superare i 5 mm. Uno spessore superiore a questo impone dei problemi di ritiro in fase di raffreddamento del pezzo stampato. All'interno di questo range di spessori si è scelto l'estremo inferiore e quindi uno spessore pari a 4 mm. Tale scelta è motivata dall'osservazione di mantenere contenute le velocità nel condotto interpale per contenere le perdite e gli effetti di scia.

Numero di pale

Il numero delle pale che è stato scelto per la soluzione proposta è pari a 6. Nelle macchine multistadio in cui le giranti sono accoppiate a diffusori palettati il numero delle pale della schiera delle pale del diffusore e della girante sono dei numeri primi tra di loro. Questo per evitare problemi di rumore dovuti alla risonanza delle pulsazioni delle onde di pressione generate dal passaggio della pale della girante in corrispondenza delle pale del diffusore. Soluzioni simili a quella sviluppata nell'azienda in questione hanno 5 pale per la girante e 6 per il diffusore. Inizialmente si è scelto di procedere con una soluzione con una girante con questo stesso numero di pale. Quello che si è notato è che, mantenendo costante il numero delle pale oltre un certo valore dell'angolo al bordo di uscita la corrente relativa diventa poco sensibile alla forma della pala: non si riesce ad incrementare la prevalenza elaborata dalla girante incrementando l'angolo in uscita delle pale. Le pale di queste giranti hanno dei numeri tipici di pressione molto alti rispetto ad altre macchine con lo stesso numero tipico di macchina perchè hanno un diametro medio della girante ridotto. Gli angoli costruttivi della pala in corrispondenza del bordo di uscita della pala sono molto alti e quindi la deviazione della corrente relativa dalla direzione costruttiva della pala è notevole. Quindi la soluzione a 5 pale non si è rivelata opportuna per raggiungere la prevalenza obiettivo in quanto la deviazione della corrente non lo consente, qualunque forma della pala si scelga. Per riuscire ad aumentare la capacità di guida della schiera di pale è stato necessario diminuire il passo pale della schiera e portare a 6 il numero di pale della girante, nonostante questo possa penalizzare il rendimento della macchina in quanto aumenta il perimetro bagnato all'interno della girante e il coefficiente di ostruzione.

Posizione del bordo di ingresso

La posizione del bordo di ingresso non si è rivelata molto influente per le prestazioni di queste macchine. La sagomatura a doppia curvatura della pala della macchina consente di ottenere buone incidenze anche per bordi di ingresso posizionati molto a ridosso del tratto assiale di adduzione. Inoltre queste macchine, per la loro peculiarità di lavorare in immersione e quindi sotto battente non sono sensibili al fenomeno della cavitazione. Quindi l'unica considerazione che guida il posizionamento del bordo di ingresso è di avere una pala sufficientemente lunga

per consentire una guida efficace della corrente (quindi di fatto che ci sia sufficiente spazio per consentire alla pala di guidare il fluido evitando distacchi di vena) senza prevedere un allungamento eccessivo della pala stessa che potrebbe penalizzare la girante in termini di perdite.

Inclinazione del canale meridiano

Il raggiungimento di uno scambio di energia sufficiente è la questione chiave nella progettazione della girante. Inoltre lo scambio di energia deve essere quello più corretto per questo tipo di macchine. In altre parole non bisogna raggiungere uno scambio di energia a scapito della stabilità di funzionamento della girante. Quindi la progettazione deve essere accompagnata da una verifica fluidodinamica della geometria proposta anche per le portate di fuori progetto al fine di controllare che la forma della curva caratteristica sia la più consona possibile.

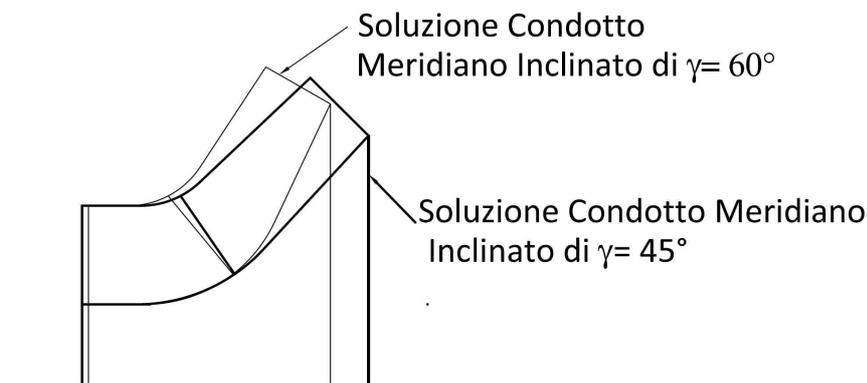


Figura 6.12: Confronto tra due condotti meridiani con diversa inclinazione rispetto all'asse della macchina.

L'esigenza di raggiungere il corretto scambio di energia impone la modifica congiunta di molte caratteristiche della girante: primi tra tutti l'inclinazione del canale meridiano e l'angolo di uscita della pala. Come già detto, quanto meno il condotto meridiano in uscita dalla girante è inclinato rispetto alla direzione assiale tanto più basso è il diametro esterno medio della bocca di uscita della girante. Questo, come si vede dalla figura 6.12, per un duplice motivo:

- a parità di dimensione della bocca di aspirazione e di estensione della pala lungo il piano meridiano un condotto meno inclinato consente diametri esterni della girante più contenuti;
- l'inclinazione della bocca di uscita è tale che, a parità di diametro esterno della girante e di larghezza del condotto meridiano al bordo di uscita, il diametro della sezione al mozzo è inferiore rispetto ad una soluzione con condotto più inclinato; in generale quindi il diametro medio della bocca di uscita è maggiore nel caso del condotto con direzione radiale.

Riassumendo, una soluzione con condotto meridiano in uscita più inclinato rispetto all'asse della macchine consente di ottenere macchine con diametri della bocca di uscita superiori, quindi con numeri di pressione ψ , a parità di scambio di energia inferiori. Questo significa anche che nello scambio complessivo di energia all'interno della girante pari a:

$$\Delta H = \frac{U_2^2 - U_1^2}{2} + \frac{W_1^2 - W_2^2}{2} + \frac{C_2^2 - C_1^2}{2}$$

a parità di energia scambiata, in una girante con il condotto più inclinato lo scambio di energia si ripartisce in maggior parte nella differenza tra le componenti centrifughe rispetto ad una soluzione con condotto meno inclinato rispetto alla direzione assiale dove lo scambio di energia fa affidamento sulla variazione corrente relativa.

Questo si traduce in giranti che, per elaborare lo scambio di energia richiesto, hanno pale con angolo in uscita tanto più alto, quanto minore è l'inclinazione del condotto meridiano rispetto alla direzione assiale. In particolare per le sezioni al mozzo, dove l'effetto centrifugo si fa sentire meno, con condotti meridiani inclinati di un angolo $\gamma = 45^\circ$ l'angolo costruttivo della pala si avvicina ai 90° . Angoli costruttivi tanto grandi non sono consueti per le pompe e non offrono buoni risultati dal punto di vista del comportamento fluidodinamico per giranti a poche pale. Per questo motivo ci si è scelto un condotto inclinato di un angolo $\gamma = 60^\circ$ rispetto alla direzione assiale e con un raggio massimo in uscita pari a 112 mm.

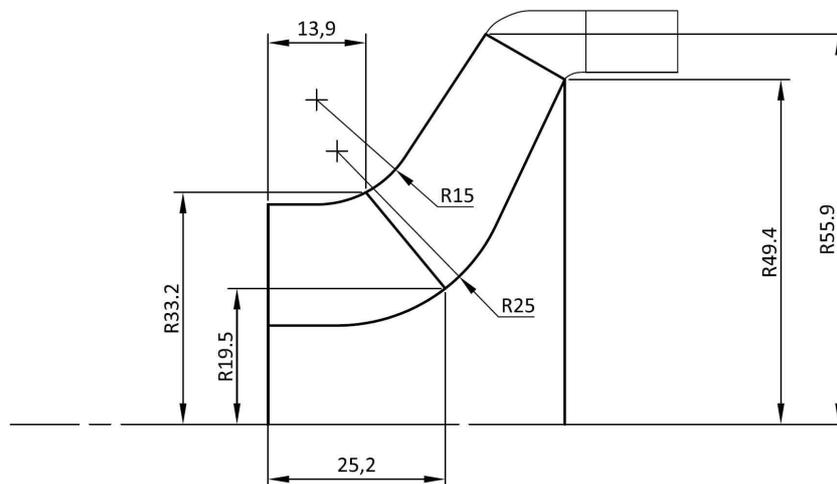


Figura 6.13: Caratteristiche principali del canale meridiano della girante progettata con particolare dei bordi di ingresso.

Tale condotto corrisponde ad una scelta del numero tipico di pressione $\psi = 0.339$ mentre un numero di flusso $\phi = 0.228$ e rappresenta, a giudizio di chi scrive il miglior compromesso che è stato possibile ottenere tra la necessità di utilizzare una girante a flusso misto per contenere l'ingombro radiale e il corretto scambio di energia.

Forma della pala

La progettazione della palettatura è vincolata da più istanze. Si deve:

- consentire lo scambio di energia richiesto; quindi l'angolo in uscita della pala deve garantire di raggiungere questo risultato; inoltre, per quanto visto nel paragrafo 2.2.3 anche il carico palare e la condizione di impilamento dei profili della pala imponendo una diversa deviazione della corrente relativa, modificano lo scambio di energia;
- minimizzare le perdite nella girante; per minimizzare le perdite di energia i parametri principali sono, in assenza di ricircoli all'interno del condotto interpalare, l'angolo della pala in corrispondenza del bordo di ingresso e l'estensione della pala, quindi l'avvolgimento θ dei profili;
- la forma della palettatura deve rispettare i vincoli di realizzabilità come visto nel paragrafo 6.1.1.

Sono stati necessari numerosi tentativi per individuare una soluzione adeguata. La scelta iniziale della prevalenza attesa per la girante è sembrata adeguata. Sono stati fatti dei tentativi per dimensionare giranti con prevalenze superiori imponendo degli angoli al bordo di uscita della pala superiori, ma le giranti che risultano sono caratterizzate da una forte instabilità funzionale e la prevalenza cala al di sotto della portata di progetto.

La girante che si propone come soluzione al problema di progettazione (denominata Girante **P**) è stata verificata tramite simulazioni CFD ed è caratterizzata da una curva di rendimento con massimo posizionato correttamente in corrispondenza della portata di progetto. La prevalenza alla portata di progetto è pari a 11.45 m, quindi in linea con le stime attese e la curva caratteristica non mostra instabilità funzionali di alcun tipo (come è possibile vedere in figura 6.14, dove le superfici di riferimento utilizzate per il computo dello scambio di energia sono quelle schematizzate in figura 6.9). In tabella 6.2 sono riportati i valori dei parametri di progetto utilizzati nella progettazione.

Il rendimento alla portata di progetto è, sempre da simulazione CFD, pari a circa $\eta_{Bep} = 0.95$ e si mantiene alto per un buon tratto intorno alla portata di massimo rendimento come è possibile notare in figura 6.14.

Le immagini di figura 6.15 mostrano la velocità e l'andamento dell'incremento di pressione nel condotto interpalare per superfici a diverse posizioni *hub to shroud*. Le immagini testimoniano il buon comportamento delle pale per quello che riguarda lo scambio di energia e la regolarità della corrente. Da notare un'incidenza positiva al naso della girante che sposta il punto di ristagno dal naso geometrico della pala verso il lato in pressione. Questo è comunque ritenuto accettabile perchè privilegia il funzionamento alle alte portate che sono quelle più critiche per queste macchine vista la limitatezza delle sezioni di attraversamento che comporta alte perdite alle alte portate. In figura 6.16 si vede inoltre l'andamento della velocità di

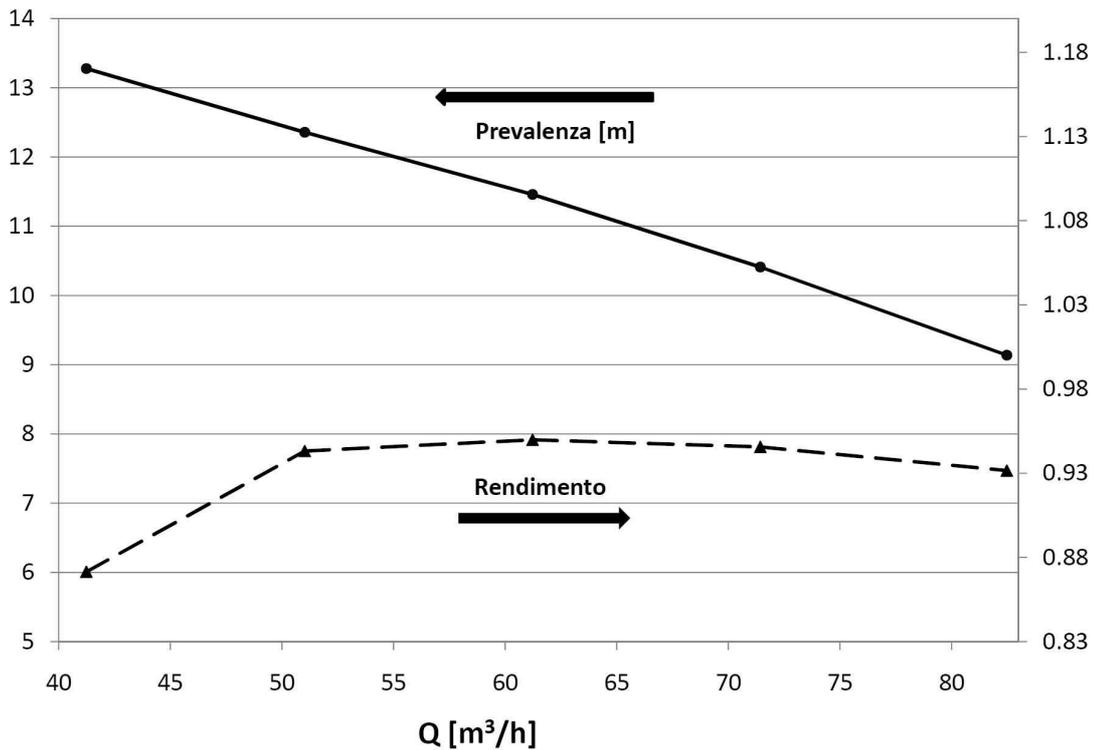


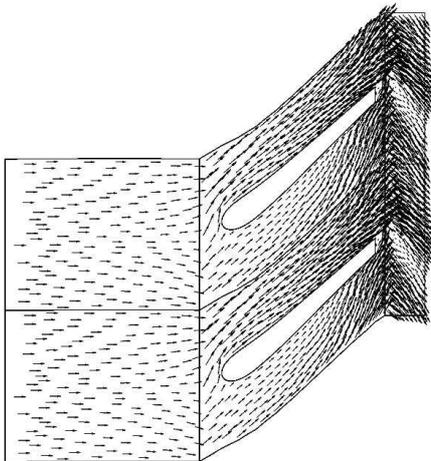
Figura 6.14: Curve caratteristiche ricavate da simulazioni numeriche della girante progettata denominata Girante **P**.

Tabella 6.2: Tabella riassuntiva dei valori dei parametri della Girante **P**.

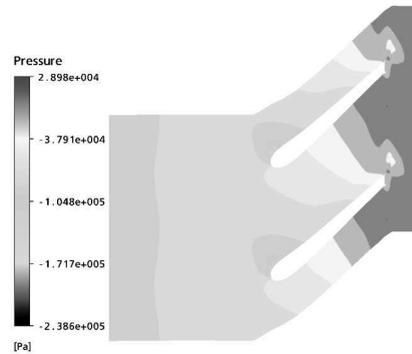
Parametri	Valore
$\bar{\theta}_S$	45
θ_{leH}	-10
ξ	-10
\bar{m}'_2	0.1
\bar{m}'_3	0.95

attraversamento in corrispondenza della bocca di uscita della girante. Come si può notare la disuniformità della corrente in uscita dal condotto interpolare è contenuta e quindi garantisce un buon accoppiamento con la schiera palare a valle.

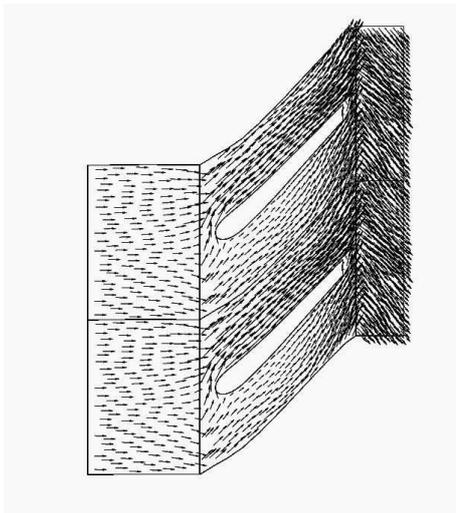
Nei grafici di figura 6.17 si può vedere l'andamento risultante per la girante **P** dell'angolo costruttivo della pala che risulta dalla forma assegnata alla pala. Come si vede tale andamento è costante lungo tutta la pala a parte nel tratto della coda della pala dove l'angolo cresce



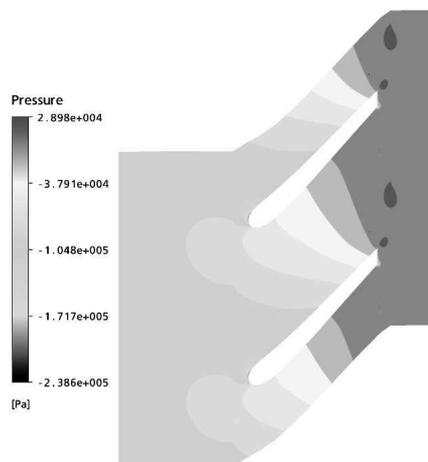
(a) Velocità blade to blade alla sezione al mozzo



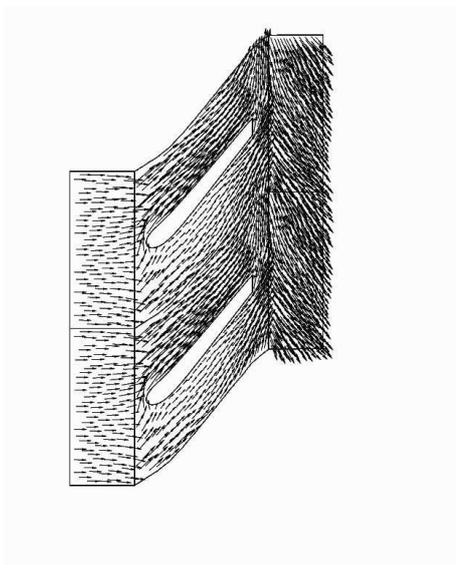
(b) Incremento di pressione statica alla sezione al mozzo



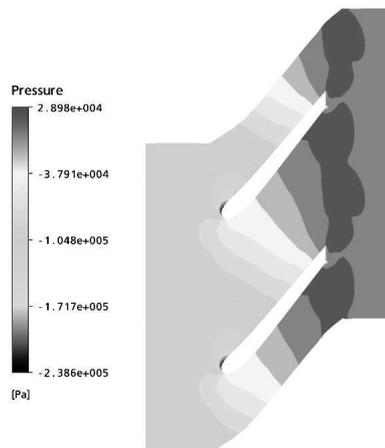
(c) Velocità blade to blade alla sezione media



(d) Incremento di pressione statica alla sezione media



(e) Velocità blade to blade alla sezione alla corona



(f) Incremento di pressione statica alla sezione alla corona

Tabella 6.3: Risultati numerici ricavati da simulazioni numeriche per la Girante **P**.

Portate [m ³ /h]	40	50	60	70	80
H [m]	13.28	12.36	11.46	10.41	9.14
Potenza [W]	1711	1820	2012	2142	2203
Ht [m]	15.23	13.10	12.06	11.01	9.81
η_{id}	0.87	0.94	0.95	0.946	0.93
Velocità meridiana scarico [m/s]	4.41	5.08	5.96	6.87	7.96
Velocità circonferenziale scarico [m/s]	8.56	7.61	7.03	6.43	5.74
Angolo corrente assoluta [deg]	26.5	33.6	40.4	47	54.4

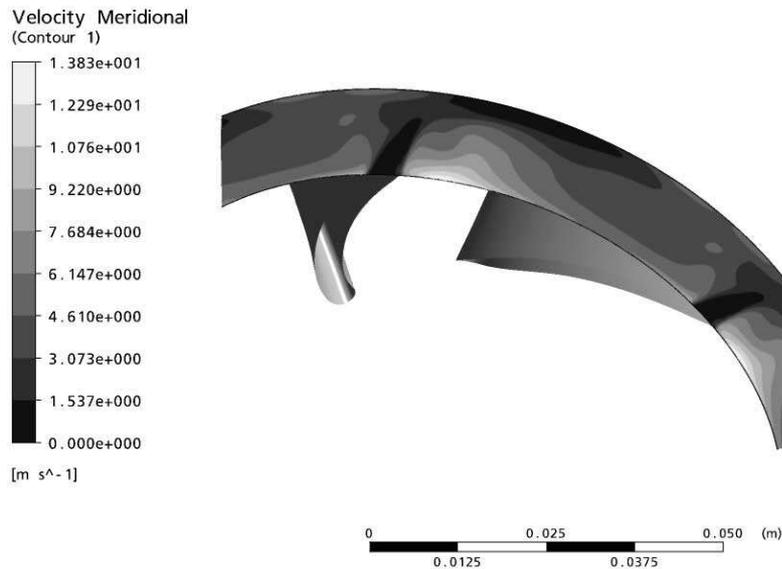
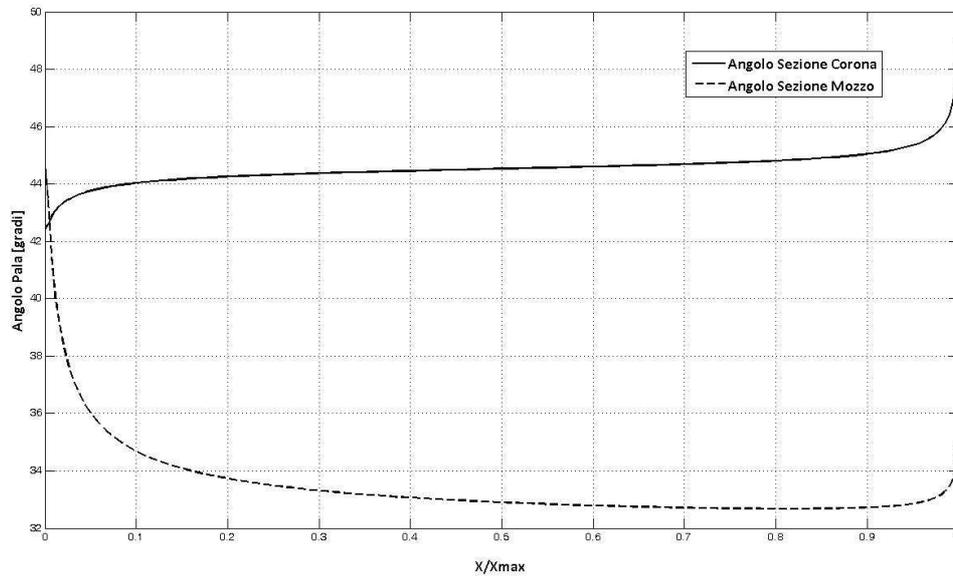
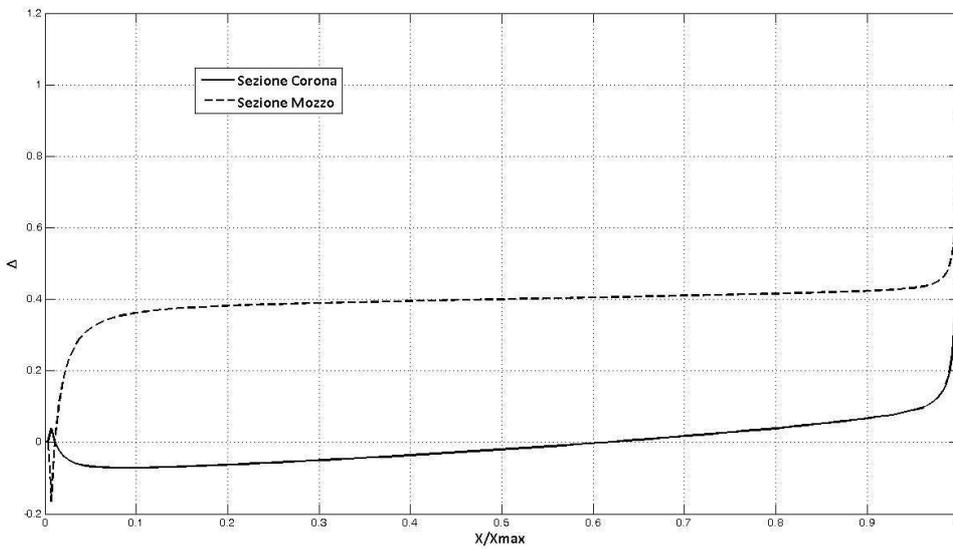


Figura 6.16: Modulo della velocità meridiana in corrispondenza del bordo di uscita della pala.

intensamente sia per la sezione al mozzo che per la sezione alla corona. Nel secondo grafico è diagrammato l'andamento della derivata della componente RC_U lungo la linea media



(a) Andamento dell'angolo β costruttivo della pala



(b) Andamento del momento angolare della quantità di moto.

Figura 6.17: Andamento di alcune grandezze rappresentative della forma della pala della girante **P**

delle sezioni di mozzo e corona che deriva dalla forma assegnata della pala nel piano conforme; tale andamento deriva da calcoli monodimensionali e avviscosi quindi è da considerarsi puramente indicativo del comportamento atteso della pala. Sotto queste premesse il comportamento atteso per la corrente nel condotto interpolare è regolare e confrontabile con gli andamenti solitamente assegnati per le sezioni di riferimento delle pale a semplice curvatura. La grandezza Δ diagrammata ha un'espressione pari a:

$$\Delta = \frac{RC_U - RC_{U1}}{RC_{U2} - RC_{U1}}$$

dove 1 corrisponde al bordo di ingresso della pala e 2 al bordo di uscita.

La forma ricavata dalla pala è realizzabile inoltre con uno stampo con carri mobili così come illustrato nel paragrafo 6.1.1. La figura 6.3 mostra in verde gli angoli di sforno corretti per il bordo di uscita e l'estradosso della pala nella direzione di estrazione tangenziale. Questo garantisce la possibilità che tali superfici possono essere definite da un carro che si muove nella direzione di analisi. La figura 6.3 mostra invece come la superficie dell'intradosso della pala, evidenziata in rosso possa essere realizzata dal punzone dello stampo e come la linea di chiusura tra punzone e carro utilizzato per costruire la pala stia in corrispondenza del naso.

6.5 Scelta del canale meridiano del diffusore

La progettazione del diffusore-riconvogliatore della macchina in questione secondo le metodologie presentate nel paragrafo 2.2.4 parte dalla forma del canale meridiano di questo organo. L'ingombro radiale di questa macchina è vincolato dall'ingombro massimo della parte idraulica della macchina in questione; rimane quindi, oltre alla scelta della forma delle generatrici del condotto, da stabilire l'ingombro assiale di questi organi. La scelta corretta è quella che garantisce le perdite minori. Tale scelta deriva dal giusto compromesso tra la ricerca di velocità più contenute e la riduzione della superficie bagnata; istanze che in questo caso sono in contrasto tra di loro: aumentando l'estensione assiale del condotto assiale è possibile disegnare dei condotti più dolci che garantiscono velocità più contenute, ma viceversa all'aumentare dell'estensione assiale aumenta anche la superficie bagnata dal fluido.

La forma del canale meridiano di questi organi è comune per molte altre macchine di taglie differenti; si è scelto quindi di applicare la procedura presentata nel capitolo 5 e di creare una banca dati di simulazioni per caratterizzare il comportamento fluidodinamico di questi canali, al variare dei valori dei parametri che ne definiscono la geometria secondo la parametrizzazione presentata nel paragrafo 2.3.

In questa parametrizzazione ciascuna geometria è identificata da sette parametri adimensionali, in quanto tutte le dimensioni che caratterizzano la geometria e le proporzioni del canale in questione fanno riferimento, direttamente o indirettamente, all'ingombro radiale massimo di ciascun pezzo. La scelta di rappresentare le geometrie con dei parametri adimensionali consente di confrontare geometrie differenti tra di loro senza curarsi della scala del pezzo. Una volta creata la banca dati è possibile associare a ciascuna geometria uno o più valori che individuano le prestazioni della geometria in analisi. In sede di progettazione sarà possi-

bile confrontare tra loro geometrie con forme incluse nella parametrizzazione utilizzata per la banca dati e con dimensioni comparabili.

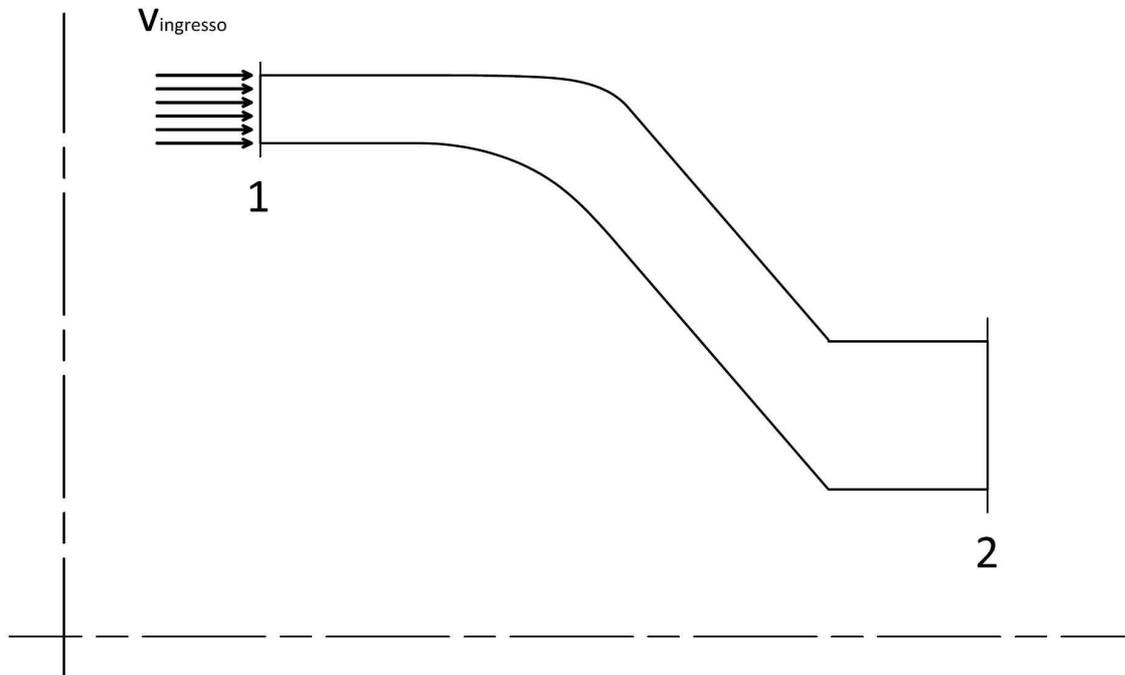


Figura 6.18: Esempificazione delle condizioni al contorno utilizzate per le simulazioni dei canali meridiani dei diffusori.

Per realizzare le simulazioni CFD che compongono la banca dati è stato necessario assegnare delle dimensioni ai condotti in analisi e dei valori ai parametri di flusso. Si è scelto di assegnare a tutte le geometrie che sono state simulate lo stesso ingombro assiale e la stessa velocità del fluido nella sezione di ingresso. In questo modo in tutte le simulazioni il flusso è caratterizzato dallo stesso numero di Reynolds, perchè per tutte la dimensione di riferimento e il valore di riferimento della velocità è costante. La velocità che è stata imposta è tale da garantire in ingresso un numero di Reynolds tipico di correnti in moto turbolento pienamente sviluppato. Con queste premesse una raccolta dati di questo tipo non consente di prevedere le prestazioni delle macchine in termini assoluti in tutte le condizioni di moto; tuttavia si ritiene che quanto proposto sia sufficiente a garantire la confrontabilità tra le varie soluzioni, cioè che assegnate due o più geometrie sia possibile individuare il rapporto tra le prestazioni di queste in modo da scegliere quella che si comporta nel modo migliore rispetto alle nostre esigenze. D'altra parte non è di alcuna utilità sapere quali che sono le perdite in un condotto meridiano dove mancano le palette; risulta invece utile capire quale condotto meridiano e in quale misura garantisce le minori perdite.

La dimensione radiale di riferimento che è stata scelta per portare avanti le simulazioni è di 59.25 mm. Tale dimensione è quella di massimo ingombro radiale del caso di progettazione che è stato sviluppato per questa tesi, come detto nel paragrafo 6.1. Si è assunto questa estensione

per contestualizzare la creazione della banca dati al primo caso a cui viene applicata. La velocità imposta in ingresso è pari a 5 m/s.

Per la pianificazione e la realizzazione di questa banca dati si è utilizzato il metodo presentato nel capitolo 5 che ha consentito la scelta ottimale dell'ingombro assiale del diffusore e della forma del canale meridiano.

Per ottenere una banca dati di precisione soddisfacente è necessario portare a termine un numero elevato di simulazioni CFD; la procedura proposta è stata pensata per essere applicata con un protocollo che consente di gestire in automatico tutte le fasi della creazione di tale banca dati, dalla fase di design of experiments, alla fase di regressione passando per la discretizzazione dei singoli condotti e la realizzazione delle varie simulazioni. È infatti indispensabile disporre di un protocollo automatico che gestisca tutte le fasi del procedimento; il numero di simulazioni da portare a termine non è noto, ma è sicuramente tale che la quantità di lavoro e di tempo necessari sono insostenibili per un operatore.

Nei paragrafi successivi sono riportate le caratteristiche principali di tale protocollo, come è stata realizzata la regressione dei risultati ottenuti dalle simulazioni e come sono stati utilizzati questi dati per ricavare informazioni utili per la scelta del canale meridiano del diffusore.

6.5.1 Schema del protocollo

Tutto il metodo che porta alla realizzazione banca dati di simulazioni e alla funzione di regressione che approssima con precisione il comportamento di questi componenti, può essere schematizzato con un flow-chart abbastanza semplice. La creazione di un diagramma a blocchi suggerisce lo schema logico del procedimento e anche lo schema che può seguire un algoritmo per creare in automatico la banca dati di simulazioni.

Nello schema proposto in figura 6.19 si evidenziano le parti principali della metodologia, e come ciascuna parte può essere visualizzata e gestita come sottosezioni tra di loro separate o collegate da passaggi logici.

Le parti principali in cui si può dividere lo schema sono essenzialmente tre: in una prima parte viene gestito il DOE del metodo che comporta la creazione dei dati di input del campione iniziale e i successivi infittimenti del campione che saranno richiesti dalla procedura. La seconda parte è quella che gestisce, a partire dai dati di input delle simulazioni da svolgere le simulazioni al computer, cioè la campagna di raccolta dei dati forniti dai computer experiments. Infine, nella terza parte si procede alla regressione dei risultati ricavati dalla banca dati, e alla verifica dell'interpolazione.

6.5.2 Linguaggio di gestione

Per rendere operativo tale protocollo è stato necessario fare uso di un linguaggio di programmazione che consente l'interconnessione tra le varie parti e sottoparti del protocollo. Lo strumento che si è scelto di adottare a questo scopo è il linguaggio Matlab.

Tale linguaggio consente innanzitutto di implementare gli algoritmi per gestire il design of experiments, la regressione dei dati e il controllo della precisione raggiunta in fase di re-

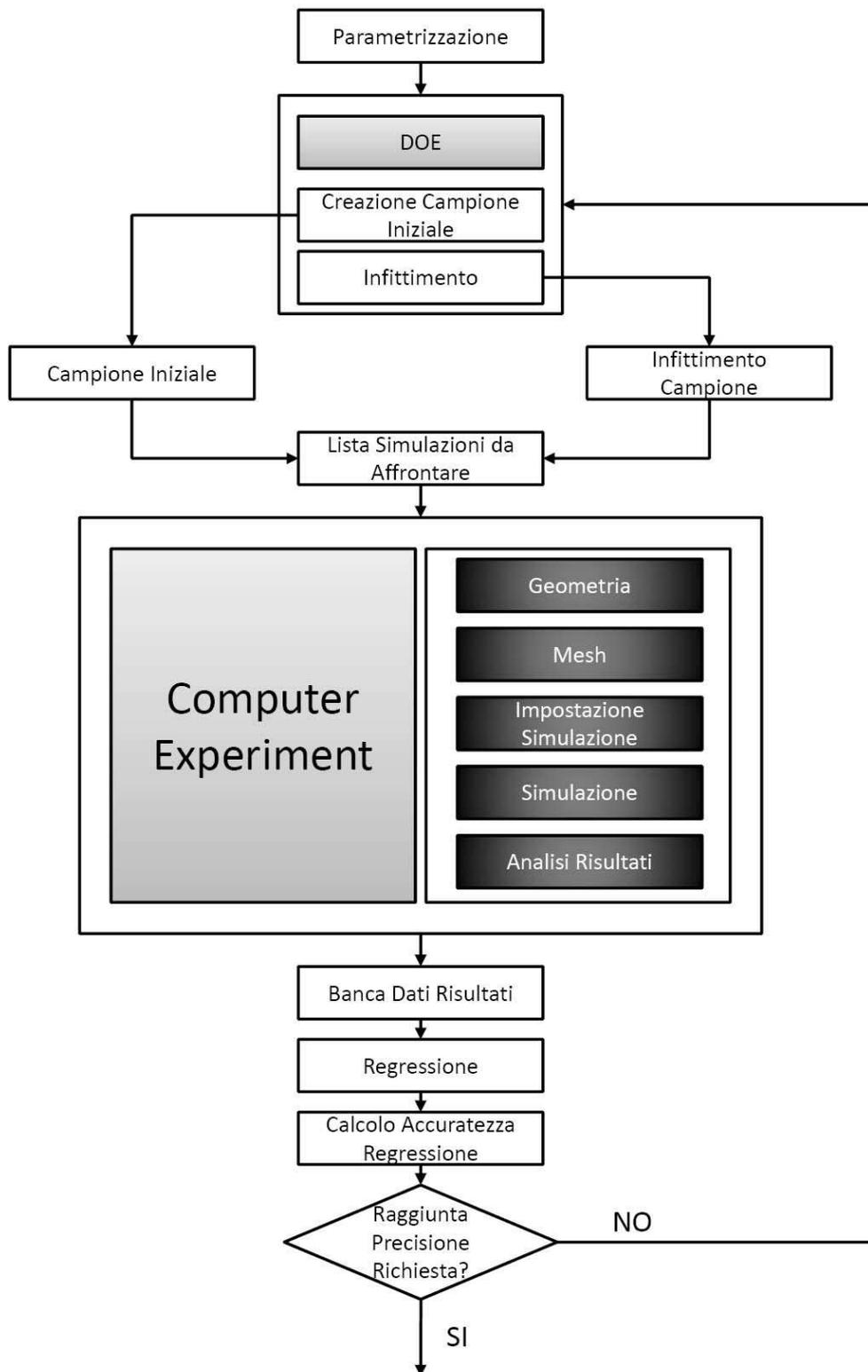


Figura 6.19: Diagramma a blocchi della protocollo per la gestione automatica della banca dati di simulazioni.

gressione secondo i metodi che si scelgono, tra quelli presentati nel capitolo 4 o secondo altre tecniche non prese in considerazione in questa tesi ma che si rivelano adatte al problema che si sta analizzando. Inoltre permette di implementare dei programmi di calcolo numerico per la realizzazione delle geometrie secondo le procedure presentate nel capitolo 2.

Altre fondamentali caratteristiche che rendono Matlab adatto a gestire il protocollo che si vuole creare possono essere schematizzate come segue:

- Matlab è in grado di interfacciare i vari programmi commerciali coinvolti nella realizzazione delle simulazioni dalle quali si ricavano i risultati della banca dati;
- consente di gestire file e cartelle del sistema operativo, con tutte le operazioni tipiche (creare, cancellare, spostare, rinominare);
- consente di immagazzinare i dati ricavati e gestite gli output delle simulazioni, creando sia dei report delle singole simulazioni che dei file di tabelle in cui vengono salvate e catalogate tutte le informazioni che si possono ricavare dalle simulazioni.

Utilizzo di Matlab

In sostanza tutta la procedura che gestisce la creazione della banca dati e la regressione dei risultati delle simulazioni può essere gestita tramite Matlab. Il protocollo che è stato schematizzato nel diagramma a blocchi in figura 6.19 può essere realizzato nella sua interezza con una funzione che all'interno è strutturata in ulteriori sottofunzioni che ricalcano le caselle dello schema. Quindi una prima funzione si preoccupa di generare il DOE iniziale della banca dati, un'altra di richiamare in batch i programmi del pacchetto di simulazione e quindi di effettuare la simulazione inserendo di volta in volta gli input corretti ecc. ecc.

Per quello che riguarda la fase di regressione dei risultati ricavate dalla banca dati di simulazioni Matlab mette a disposizione dei toolbox interni che consentono di addestrare delle reti neurali o delle reti di funzioni a base radiale impostate secondo la struttura che si preferisce (Neural Network Toolbox e Radial Basis Network Toolbox). Il toolbox di statistica consente inoltre di addestrare velocemente delle response surface su insiemi di dati noti senza dover implementare alcuna funzione (vedi [51]).

Sul web, inoltre, è possibile rintracciare numerosi toolbox che possono essere inseriti nella libreria di Matlab e richiamati al momento opportuno. Alcuni di questi toolbox risultano molto utili e mettono a disposizione tecniche statistiche e matematiche altrimenti non accessibili e di non facile implementazione. È a disposizione un toolbox gratuito, sviluppato da un gruppo di ricercatori della DTU-Danmarks Tekniske Universitet che consente di utilizzare la tecnica di regressione del Kriging sempre tramite Matlab (vedi [48]).

Matlab inoltre è stato utilizzato anche come programma di calcolo per la gestione del design of experiments della banca dati per la scelta dei casi da simulare e i successivi infittimenti.

Tale porzione del protocollo è matematicamente molto semplice e richiede solo poche righe di codice per essere implementata.

Più complesso da sviluppare è la parte che riguarda la gestione della geometria parametrica, che richiede la realizzazione di funzioni più articolate in modo da individuare per punti le geometrie dei canali da realizzare secondo la parametrizzazione illustrata nel paragrafo 2.3.

Matlab è stato utilizzato anche per gestire tutte le parti intermedie del protocollo. In pratica tutta la struttura del protocollo automatico è stata creata grazie ad una serie di funzioni Matlab annidate. L'insieme di queste funzioni è in grado di spostare, rinominare e cancellare cartelle, aprire e modificare file di testo e di creare ed eseguire file .bat che richiamano i programmi del pacchetto Ansys-CFX per eseguire le simulazioni fluidodinamiche. Di fatto Matlab è stato utilizzato per mettere in essere tutte le operazioni di immagazzinamento e gestione proprie di una database.

Per attuare tutto ciò è stato anche necessario prevedere un sistema di catalogazione dei dati e delle simulazioni. Infatti di volta in volta, per consentire alla funzione base di interfacciare le varie sottofunzioni e i programmi fluidodinamici, si deve rintracciare un determinato file piuttosto che un altro ed eseguire una determinata operazione al posto di un'altra. Un sistema di catalogazione e archiviazione delle simulazioni e delle informazioni ad esse collegate consente anche di spezzettare le operazioni da svolgere e di gestire al meglio le varie fasi della creazione della banca dati. Infatti la logica a blocchi proposta come traccia non impone al suo interno che sia necessario svolgere la sequenza delle operazioni che compongono una simulazione, dalla generazione della geometria alla interpretazione dei risultati, in una sequenza unica per ciascuna simulazione. Il risparmio in termini di tempo, o la convenienza dell'operatore, potrebbe richiedere di raggruppare tutte le operazioni di un certo tipo che riguardano un determinato insieme di simulazioni e solo successivamente completare la sequenza di operazioni per quelle simulazioni.

6.5.3 Aspetti peculiari: meshatura.

La sequenza proposta sopra è del tutto ideale. Nella realtà le possibilità di automatizzare la sequenza sono limitate dal caso in analisi. In altre parole, non sempre i programmi a disposizione hanno delle procedure interne in grado di gestire completamente in automatico le varie fasi. In molte situazioni, per rendere automatica la creazione della banca dati, è necessario prevedere delle soluzioni ad hoc che nel concreto significa cercare delle soluzioni personalizzate per la geometria che si vuole analizzare o ottimizzare. Le fasi che si sono rivelate critiche da automatizzare sono, come ci si poteva aspettare, quelle che nella normale esecuzione di una simulazione richiedono il maggior intervento umano. Tali fasi sono, quindi, quelle di meshatura e di post-processing.

Come si è visto in precedenza il pacchetto di programmi a disposizione Ansys-CFX mette a disposizione due diversi programmi per la meshatura:

- CFX-Turbogrid, programma specifico dedicato alla discretizzazione con mesh strut-

turate di organi palettati di turbomacchine, radiali o assiali;

- ANSYS-ICEM, programma per la meshatura general purpose che consente la discretizzazione sia con mesh strutturate che non strutturate di geometrie generiche.

Ciascuno di questi programmi, dopo aver caricato la geometria e realizzato la mesh, memorizza il risultato su dei file che poi sono importati nel programma di pre-processing in fase di impostazione delle condizioni al contorno della simulazione. Generalmente questa fase è molto delicata, e l'operatore impiega a volte parecchio tempo per realizzare la mesh voluta.

Il volume geometrico che è stato discretizzato e simulato per indagare ciascuna geometria è quello mostrato in figura 6.20. Oltre alla parte che rappresenta proprio il volume del canale meridiano del diffusore individuato dalla rivoluzione attorno all'asse delle due generatrici nel piano meridiano, si sono aggiunte due zone ausiliarie a sviluppo assiale a monte e valle del canale, per simulare correttamente le condizioni del moto della corrente. L'estensione assiale delle pareti parallele della zona ausiliaria a monte del diffusore è stata scelta pari a 1 cm tale e quale all'estensione assegnata alle pareti parallele della zona ausiliaria a valle.

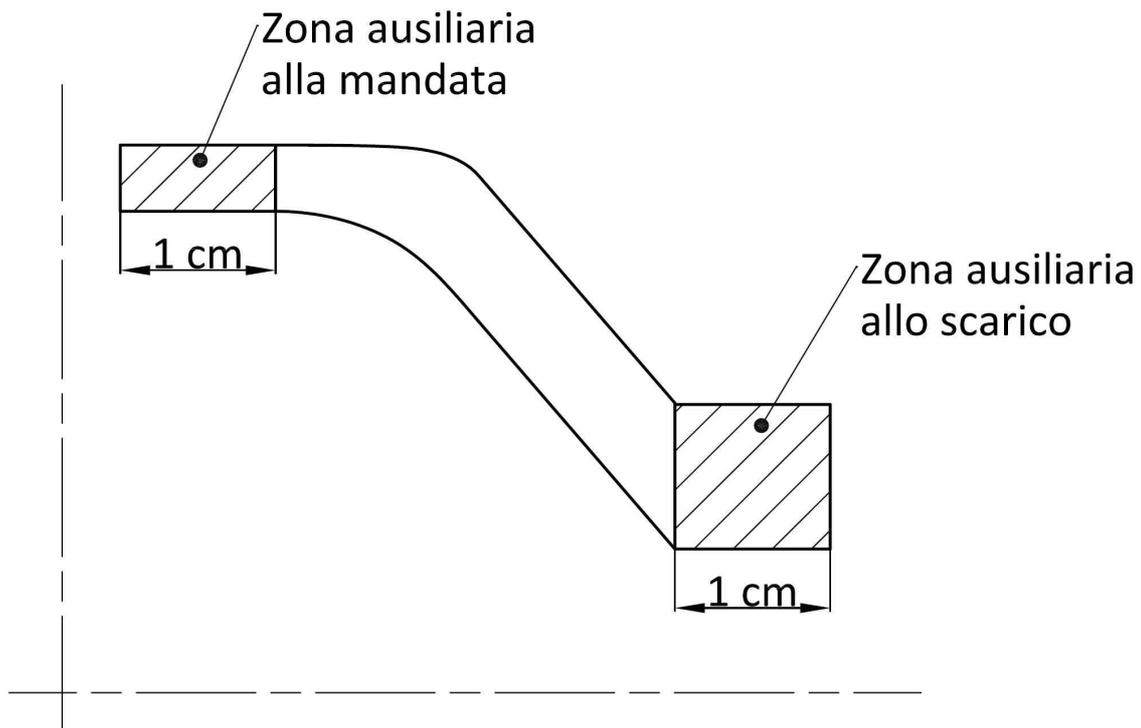


Figura 6.20: Zone ausiliarie utilizzate per la corretta modellazione del flusso all'interno del condotto meridiano del diffusore.

Tali condotti sono assialsimmetrici per qualunque angolo di periodicità. Questi condotti sono dei canali meridiani senza palettatura, ottenuti per rivoluzione attorno all'asse della macchina delle curve generatrici. Quindi l'assialsimmetria è completa e qualunque piano di

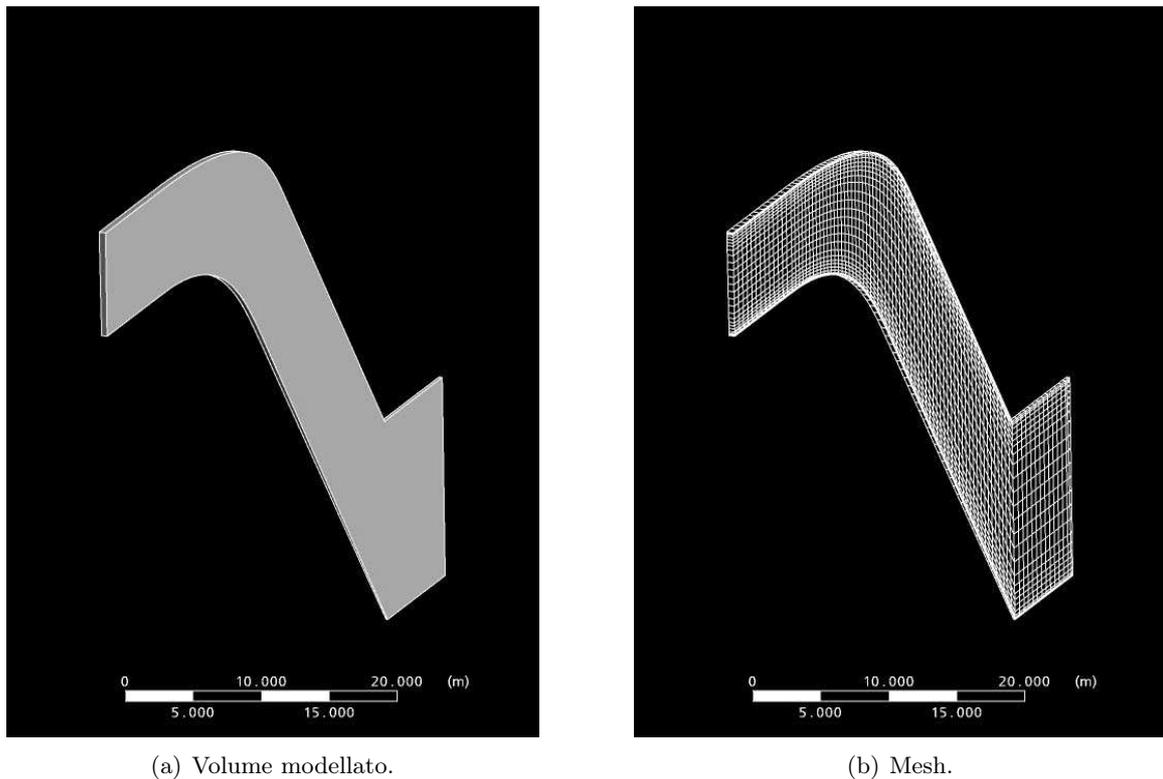
sezione passante per l'asse della macchina rappresenta una periodicità. Tale particolarità geometrica deve essere sfruttata per semplificare la complessità delle simulazioni da svolgere. Il flusso della corrente è sviluppato completamente nella direzione del piano meridiano quindi la meshatura deve svilupparsi prevalentemente sul piano meridiano e può essere più semplice in direzione circonferenziale. L'ideale sarebbe quello di effettuare una simulazione bidimensionale del canale meridiano, ma questo non è possibile. La discretizzazione realizzata con le celle ha uno sviluppo tridimensionale e quindi devono esistere nodi disposti tra di loro con distanza non nulla in tutte e tre le direzioni dello spazio.

Si è cercato di raggiungere il massimo livello di semplificazione possibile: per contenere i tempi e le risorse di calcolo necessarie, ma senza perdere informazioni utili al confronto di varie geometrie. Quindi ciascuna soluzione scelta dal DOE è stata discretizzata solo in un piccolo settore assialsimmetrico di 1° . In questo modo i nodi della mesh sono stati disposti sul piano meridiano che corrisponde ad una delle due facce periodiche del volume grazie ad una mesh strutturata, come quella mostrata in figura 6.21. Successivamente la posizione di questi è stata copiata, per rotazione sull'altra faccia del volume simulato che corrisponde all'altra periodicità, distante, in direzione circonferenziale 1° . In questo modo il volume è discretizzato con due soli layer di nodi in direzione circonferenziale, ciascuno su una periodicità. Questa scelta è quella che consente di limitare il più possibile i nodi in direzione circonferenziale, consentendo di concentrarli il più possibile nella direzione del flusso. Inoltre i nodi sulle facce periodiche sono coincidenti in direzione circonferenziale e questo garantisce minor utilizzo di risorse di calcolo al solutore.

La parte più complicata da sviluppare nella discretizzazione di questi condotti non è tanto la meshatura di questi condotti intesa come scelta dei parametri, quanto il rendere automatico queste operazioni.

Purtroppo i programmi commerciali utilizzati nel pacchetto Ansys-CFX11 sono inadeguati a questo scopo. Questi mettono a disposizione la possibilità di registrare delle macro con la sequenza di istruzioni da svolgere per giungere alla discretizzazione del volume in questione. Tali macro sono dei file di testo e possono essere modificati con semplicità grazie a dei codici di calcolo tipo Matlab per sostituire i dati di input per creare la geometria e i parametri della mesh. In tal modo è possibile applicare la sequenza di istruzioni utilizzata nella discretizzazione di una geometria ad un'altra geometria, anche in modalità batch. Molto spesso, però, questo non è sufficiente per la realizzazione di una mesh strutturata corretta. Il programma dedicato a questa operazione del pacchetto Ansys-CFX11, ICEM¹, risulta molto instabile e la sequenza di operazioni spesso si arresta in corrispondenza della creazione della geometria. L'output del programma è molto spesso un messaggio di errore oppure un file con una discretizzazione assolutamente sbagliata, che necessitano di un intervento manuale da parte di un operatore. Ovviamente questo non si sposa con una procedura automatica dove è necessario avere a disposizione una serie di operazioni che consentano di giungere ad un risultato senza l'intervento di un operatore.

¹Il programma dedicato nello specifico alla realizzazione della meshatura delle turbomacchine Turbo-grid è predisposto solo per la creazione di mesh di palettaggi completi, e non è un supporto valido per la discretizzazione di soli canali meridiani senza palettatura.



(a) Volume modellato.

(b) Mesh.

Figura 6.21: Esempio di settore assialsimmetrico di 1° utilizzato per modellare il canale meridiano di un diffusore e relativa mesh.

Per riuscire a giungere ad una procedura completamente automatica è stato necessario sviluppare un algoritmo diretto per la discretizzazione dei canali meridiani. L'algoritmo che è stato sviluppato individua le coordinate precise dei nodi della discretizzazione e consente di organizzare la discretizzazione in file di formato *Patran Neutral File*. Questi sono dei file di testo con estensione *.out* dove le coordinate dei nodi sono organizzate secondo una particolare formattazione. Pertanto è possibile, tramite una funzione di Matlab individuare le coordinate tridimensionali dei nodi di una discretizzazione e stampare un file di testo secondo la formattazione dei file neutrali di Patran di estensione *.out* che vanno a costituire un file di mesh che può essere letto dal programma di impostazioni delle simulazioni.

La discretizzazione è stata individuata grazie ad una griglia sul piano meridiano che utilizza le due direzioni principali di sviluppo del pezzo: la direzione del fluido, individuata dalle generatrici di mozzo e corona, e la direzione di alcune pseudo-ortogonali a queste. Lo schema utilizzato per la definizione di questa mesh prevede di dividere in n parti di lunghezza uguale le generatrici di mozzo e corona, e di individuare così i nodi che giacciono lungo le generatrici. Congiungendo con segmenti rettilinei i nodi di pari cardinalità sulle due generatrici si individuano le pseudo ortogonali al moto della corrente. Per individuare i nodi interni queste linee vengono divise in m segmenti, questa volta garantendo un opportuno infittimento in corrispondenza dei nodi della parete; l'infittimento in prossimità delle pareti è stato realizzato con una legge geometrica. La griglia di punti che risulta da questo tipo di discretizzazione è quello mostrato in figura 6.22.

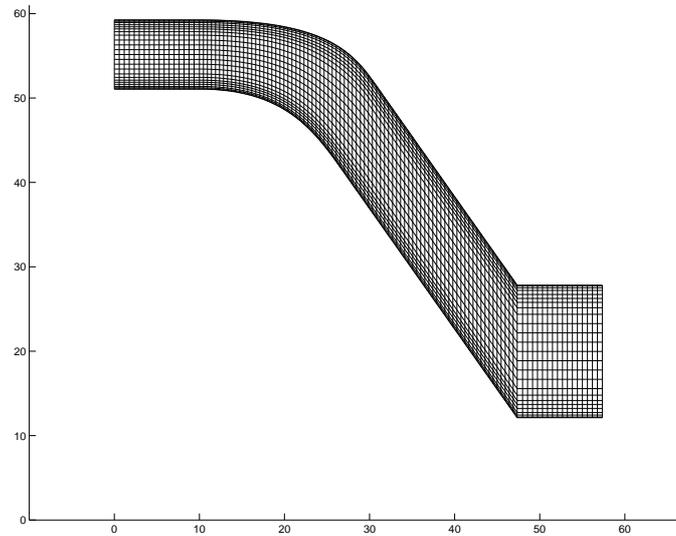


Figura 6.22: Esempio della discretizzazione che si ottiene con l’algoritmo proposto sul piano meridiano.

La densità della griglia in direzione *inlet to outlet* è stata definita per mezzo della distanza **Dist** in mm tra i nodi. In direzione *hub to shroud* sono stati posti 24 nodi, con un infittimento verso le pareti definito sulla lunghezza normalizzata delle pseudo-ortogonali. In corrispondenza delle pareti la distanza normalizzata tra i nodi contigui segue la distribuzione: 0.01, 0.01, 0.02, 0.03.

Sono stati eseguiti degli studi di sensibilità su queste mesh al fine di verificare le capacità predittive di questo tipo di discretizzazione e la qualità della mesh stessa. Infatti le scelte fatte portano a mesh anomale, senza nodi *interni*, ma solo con nodi che giacciono sulle superfici delle condizioni al contorno. Per verificare che questo tipo di modellizzazione non comporti differenze rispetto ad una modellizzazione più completa sono stati svolti degli studi di confronto per poter affermare che i risultati di una simulazione con una mesh senza nodi interni, abbia risultati analoghi ad una mesh con nodi interni. Sono state confrontate diverse geometrie di canale meridiano, ciascuna delle quali è stata modellata sia con un settore circolare di 1° discretizzato con due strati di nodi in direzione circonferenziale, che con un settore circolare di 9° discretizzato con 10 strati di nodi in direzione circonferenziale, spaziatati ciascuno di 1° . Sono state scelte due diverse geometrie ed è stato fatto variare un parametro per ciascuna geometria. I dati riportati rispettivamente nelle tabelle 6.4 e 6.5 e nei grafici di figure 6.23 e 6.24 mostrano come non ci siano sostanziali differenze tra una simulazione senza nodi interni e una con nodi interni. Questo non solo in assoluto, tra simulazioni corrispondenti, ma anche nell’andamento che si riscontra tra i risultati delle simulazioni al variare dei parametri, con scarti tra le simulazioni con un due soli layer di nodi e le simulazioni con dieci layer inferiori al 1%.

Ulteriori simulazioni su modelli di settore più ampi in senso circonferenziale e discretizzati

con un numero superiore di layer superiore a dieci non hanno mostrato differenze, con il caso analizzato. Di questi non si riportati i risultati numerici perchè non sono stati raccolti in modo sistematico e completo.

Tabella 6.4: Confronto delle perdite di energia totale al variare del parametro L_{H3} e del numero di layer utilizzati per la discretizzazione in direzione circonferenziale.

L_{H3}	Numero di layer	Perdita di energia [Pa]	Numero di layer	Perdita di energia [Pa]
0.4	2	6927	10	6926
0.5	2	6899	10	6904
0.6	2	6890	10	6894
0.7	2	6878	10	6885
0.8	2	6860	10	6867

Tabella 6.5: Confronto delle perdite di energia totale al variare del parametro A_x/R e del numero di layer utilizzati per la discretizzazione in direzione circonferenziale.

A_x/R	Numero di layer	Perdita di energia [Pa]	Numero di layer	Perdita di energia [Pa]
0.4	2	8273	10	8467
0.5	2	7030	10	7113
0.6	2	6959	10	6958
0.7	2	7305	10	7307
0.8	2	7865	10	7868
0.8	2	8630	10	8617

E' stato inoltre realizzato uno studio di sensibilità della mesh per identificare il numero di nodi necessario per garantire un corretto valore dei risultati in termini di perdite di energia. La densità della griglia è stata progressivamente aumentata diminuendo la distanza Dist e sono state misurate le perdite di energia tra ingresso e uscita del condotto sia per mesh a 2 layer di nodi che a 10 layer di nodi in direzione circonferenziale. Questo ha reso possibile identificare il valore della dimensione della griglia che garantisce di affrancare i risultati delle simulazioni dall'infittimento della griglia e scegliere l'infittimento più opportuno della griglia. In tabella 6.6 sono riportati i risultati in termini di perdite di energia al variare del numero di layer in senso circonferenziale e di Dist per una geometria di condotto con flusso non stallato. Tali risultati sono riportati sotto forma di grafico anche in figura 6.25.

Come si può osservare dalla tabella 6.6 e della figura 6.25 per valori di $Dist < 0.9$ l'influenza della densità della griglia è trascurabile. Quindi tale valore è stato scelto per tutte le simulazioni effettuate. Quello che si può affermare è che sono sufficienti griglie che portano la numerosità complessiva dei nodi sopra le 4000 unità. Gli infittimenti impostati intorno alle

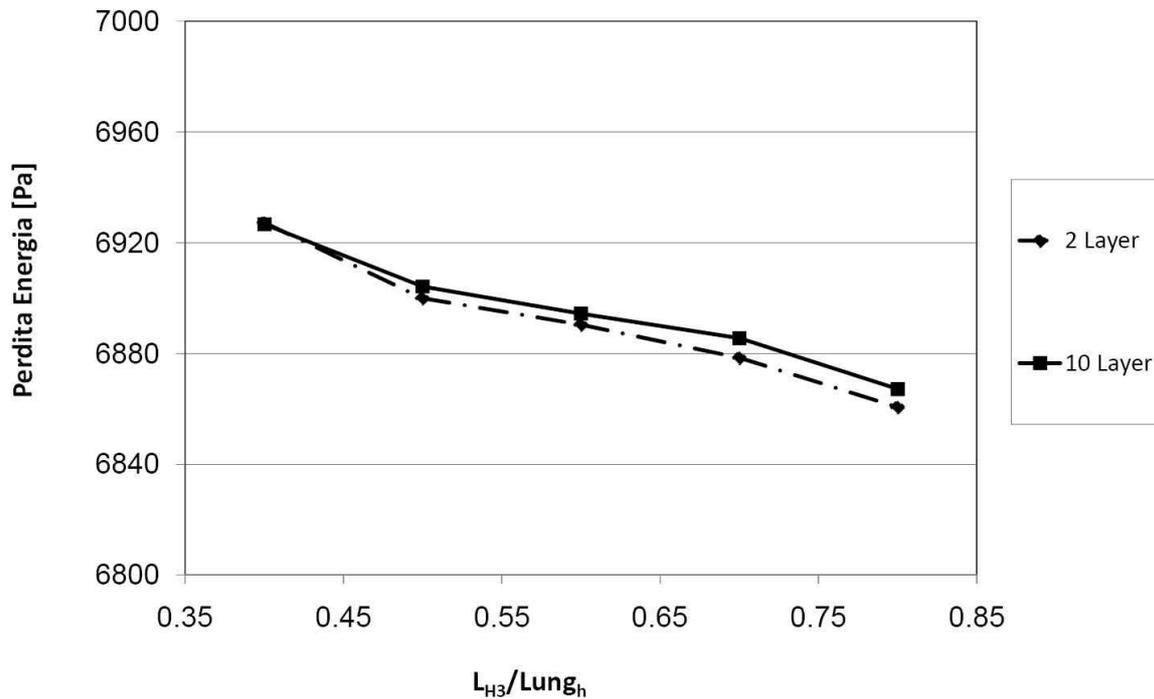


Figura 6.23: Andamento delle perdite al variare del parametro L_{H3} e del numero di layer di nodi in direzione circonferenziale, con riferimento alla tabella 6.23

Tabella 6.6: Risultati dello studio di sensibilità della mesh.

Dist	Numero di layer	Perdita di energia [Pa]	Numero di layer	Perdita di energia [Pa]
2	2	11000	10	11000
1.7	2	10200	10	10200
1.4	2	9600	10	9600
1.1	2	9260	10	9260
0.8	2	9060	10	9060
0.7	2	9000	10	9020
0.6	2	9020	10	9030

pareti comunque garantisce valori di y^+ della mesh inferiori a 200 sia al mozzo che alla corona. Questo rispetta quelli che sono i consigli di modellazione forniti da [52] per simulazioni che fanno uso del modello di turbolenza SST.

6.5.4 Aspetti peculiari: analisi dei risultati

Come già detto nel paragrafo 6.2.4 la fase di analisi dei risultati, o post-processing, è quella che generalmente richiede maggior intervento umano, in quanto l'esperienza di quest'ultimo è

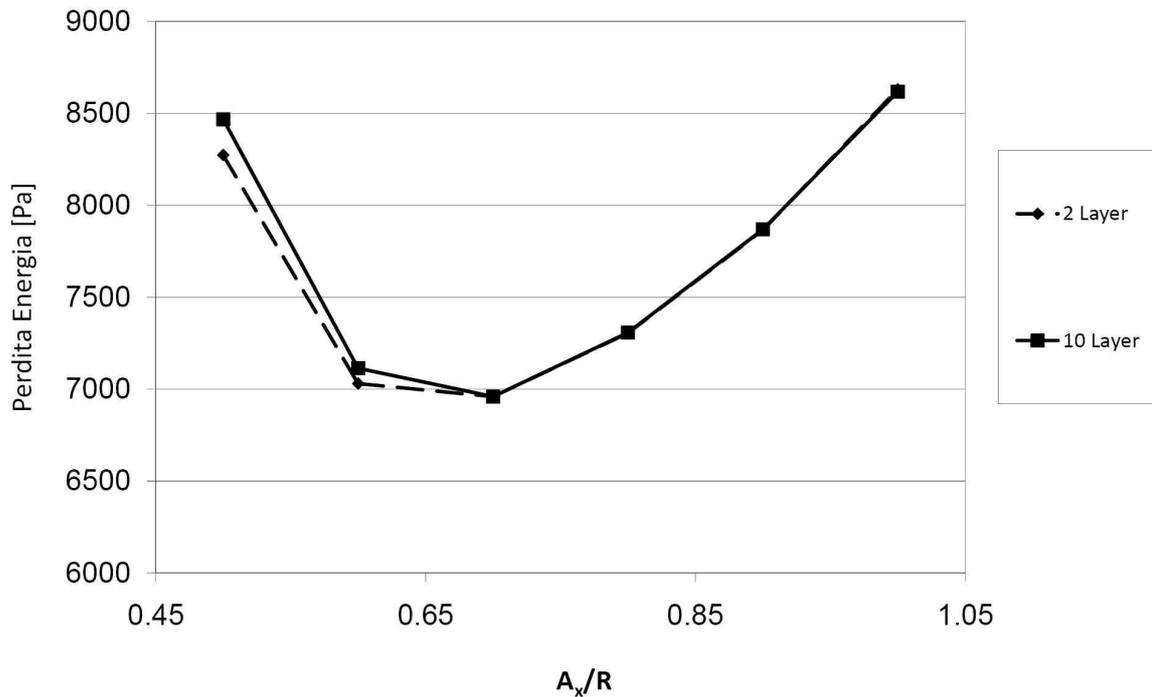


Figura 6.24: Andamento delle perdite al variare del parametro L_{H3} e del numero di layer di nodi in direzione circonferenziale, con riferimento alla tabella 6.24

fondamentale per la corretta interpretazione dei dati. Molti fenomeni, infatti, possono essere compresi solo a partire dall'osservazione diretta del campo di moto nel suo completo. D'altro canto però quanto è maggiore l'intervento umano in una determinata operazione tanto più diventa difficile automatizzare tale operazione. Un post-processing completo si sposa bene con la progettazione caso per caso, ma non con le tecniche di ottimizzazione, che richiedono infatti la messa in opera di protocolli automatici.

Quando si utilizzano le tecniche di ottimizzazione nella progettazione fluidodinamica si deve ricondurre il funzionamento fluidodinamico delle turbomacchine a parametri numerici, in modo da poter ricostruire una funzione matematica da minimizzare. In [50] o in [1] è possibile trovare degli esempi su come sono costruite queste funzioni.

Nei casi più complessi, dove i parametri che devono essere ottimizzati sono molteplici, la misura delle prestazioni delle macchine e quindi le funzioni da minimizzare sono composte da molti termini come si è visto nel paragrafo 1.3.2 parlando di [50]. In questi casi si cerca di ricostruire anche andamenti di pressioni o di velocità; sono quindi necessari molti termini per poter costruire delle funzioni obiettivo del genere, a cui corrispondono quindi molti dati da ricavare nelle fasi di post-processing delle simulazioni, che quindi diventano molto pesanti.

Anche per lo scopo che si prefigge questa tesi è necessario caratterizzare il comportamento di ciascuna geometria con delle stime numeriche delle prestazioni. Quello che si deve costruire in altre parole è una misura delle prestazioni e della funzionalità fluidodinamica delle varie geometrie al fine di poter confrontare varie soluzioni.

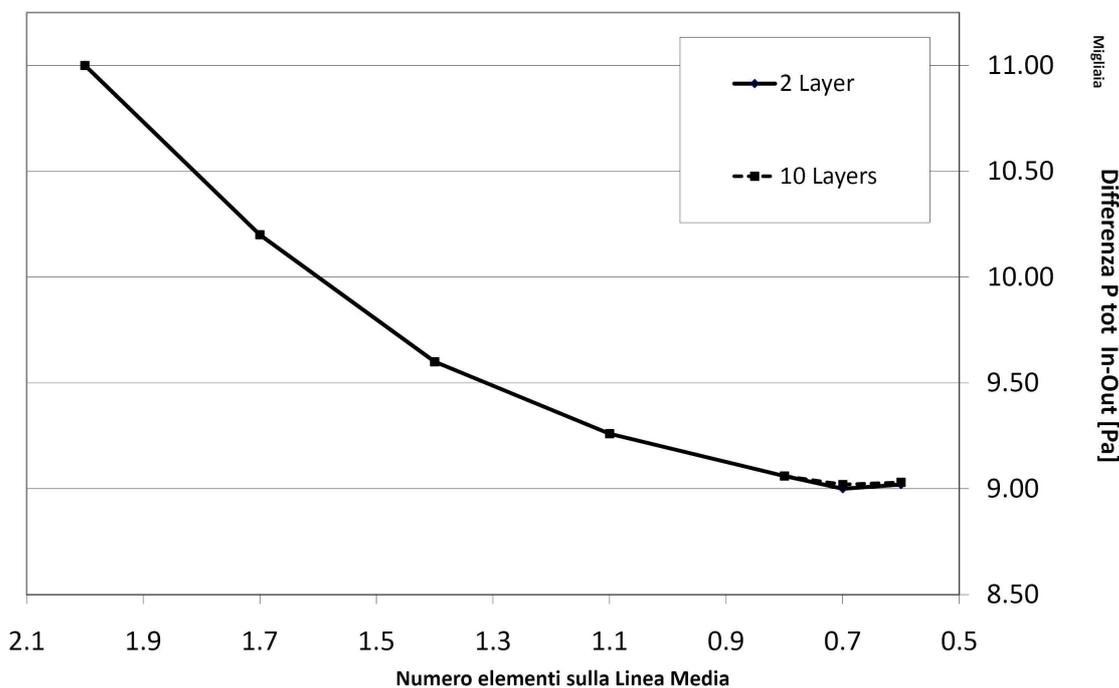


Figura 6.25: Influenza dell'infiltramento della discretizzazione sui risultati della simulazioni; dati di tabella 6.6

Per la creazione della banca dati in questione è necessario tenere presente meno parametri. Di fatto in questo caso il parametro più importante da considerare, che è anche l'unico, è la perdita di pressione totale tra ingresso e uscita del canale meridiano, proprio perché questo è il parametro che si può investigare in questo tipo di analisi e che può dare informazioni utili in sede di progettazione. Infatti le perdite di energia tra ingresso e uscita consentono di fare una graduatoria di merito tra più soluzioni e di capire quale ha una forma più dissipativa.

Quindi la fase di post-processing di questi canali risulta limitata alle perdite di energia tra ingresso e uscita del canale che è stato analizzato. Quindi, la differenza di pressione totale del fluido tra ingresso e uscita del canale è quella che deve essere rilevata per ciascuna geometria nella fase di post-processing. Le superfici di riferimento utilizzate per il calcolo della differenza di pressione sono le superfici denominate con 1 per l'ingresso e 2 per l'uscita in figura 6.18. Di fatto la scelta ricade sul calcolo delle perdite di energia totale per tutto il condotto modellato, inglobando anche le zone ausiliarie per conteggiare anche le perturbazioni indotte in queste zone a monte e valle; si è scelto di conteggiare anche le perdite in queste zone perché si ritiene siano parte essenziale del funzionamento del condotto anche se queste non fanno parte di fatto del canale.

Il calcolo di queste perdite è semplice dal punto di vista del post-processing per queste simulazioni. Le funzioni che sono messe a disposizione dal programma CFXPost consentono di calcolare la pressione totale su varie tipologie di superficie di riferimento. Il calcolo dell'energia totale è fatto come riferimento sulla pressione totale. CFXPost mette a disposizione la

possibilità di mediare il valore della pressione totale su tutta la superficie di riferimento; la media in questione è una media integrale pesata sulla portata che attraversa le superfici di riferimento (vedi [53]). In termini matematici l'energia media sulla superficie di riferimento può essere espressa come:

$$\bar{E} = \frac{\int mE}{\int m}$$

dove E è l'energia specifica (la pressione totale), m è la massa che attraversa un'unità di superficie nell'unità di tempo.

In tale modo si è caratterizzata ciascuna geometria analizzata con le perdite di energia nel condotto in questione.

La fase di post-processing, che si concretizza con il calcolo delle perdita di energia, può essere svolta con una procedura batch dal programma CFXPost; in questa procedura la sovrastruttura in Matlab compila un file di testo *.cse* che contiene la serie di istruzioni che CFXPost necessita per calcolare la differenza di pressione totale tra l'ingresso e l'uscita del canale. Il programma viene richiamato e la procedura contenuta nel file *.cse* viene avviata avviando un file eseguibile di dos *.bat* da Matlab. Come output di questa serie di istruzioni il programma CFXPost rende un file di testo dove è contenuto il valore della perdita di pressione in questione, che può essere letta e archiviata.

6.5.5 Creazione della banca dati

La realizzazione effettiva della banca dati rappresenta la messa in opera di tutte le parti della procedura proposte, con le peculiarità presentate nei paragrafi precedenti per la realizzazione della discretizzazione e per la fase di post-processing. L'impianto seguito dall'algoritmo è quello schematizzato in figura 6.19.

Per la creazione della banca dati, i parametri che sono stati fatti variare sono i sette riportati nel paragrafo 2.3 utilizzati per definire la forma del condotto meridiano del diffusore. Il campo di variazione di ciascun parametro è riportato in tabella 6.7

I valori riportati in tabella sono stati scelti con un'analisi delle geometrie delle pompe multistadio sommerse di piccola e media taglia attualmente in produzione nell'azienda finanziatrice e nella concorrenza. Gli estremi di variazione sono stati scelti in modo tale da poter prevedere l'accoppiamento con giranti a flusso misto con proporzioni tipiche per queste macchine, ricavate sia dai cataloghi dei principali costruttori di pompe e da dimensionamenti di massima suggeriti dai diagrammi statistici. Quindi i valori dei diametri di mozzo, bocca di aspirazione e bocca di mandata sono in linea con i valori che la pratica suggerisce per queste macchine. Inoltre con la scelta di tali estremi di variazione si è cercato di inglobare solo geometrie plausibili: per questo, per esempio, la larghezza del condotto all'ingresso è naturalmente inferiore a quella all'uscita, in quanto vanno riferite all'andamento delle aree di attraversamento delle giranti a cui si interfaccia il diffusore a monte e a valle.

Si è scelto un incremento della numerosità del campione di dati pari a 24 unità. Questo

Tabella 6.7: Intervalli di variazione utilizzati nella creazione della banca dati dei parametri di definizione della geometria del canale meridiano del diffusore-convogliatore.

Parametro	Estremo inferiore dell'intervallo di variazione	Estremo superiore dell'intervallo di variazione
$\frac{A_x}{R}$	0.5	1.5
$\frac{B_2}{R}$	0.2	0.35
$\frac{R_m}{R}$	0.2	0.3
$\frac{B_1}{B_2}$	0.4	0.75
$\frac{L_{H3}}{L_{unGH}}$	0.4	0.9
$\frac{L_{S3}}{L_{unGS}}$	0.4	0.9
$\frac{L_{H1}}{A_x}$	0.2	0.6

corrisponde, per quanto visto nel paragrafo 5.2, alla dimensione degli ipercubi latini che rappresentano l'unità base del DOE utilizzato per infittire di volta in volta la banca dati di simulazioni dopo la verifica della precisione raggiunta con i metodi di regressione. Tale dimensione è stata scelta perchè garantisce, a giudizio di chi scrive, un'adeguata suddivisione degli intervalli di variazione dei parametri di input e quindi consente di esplorare con sufficiente precisione gli andamenti delle variabili.

I dati ricavati dalle simulazioni sono stati utilizzati per addestrare dei modelli di regressione e immagazzinare i dati in questi modelli; prima dell'addestramento delle tecniche di regressione, è stata applicata ai dati la funzione di Matlab *mapstd*. Lo scopo di tale funzione (vedi [51]) è quello di preprocessare i valori dell'insieme di dati noto. Tale funzione scala i valori dell'insieme a cui è applicato in un nuovo insieme caratterizzato da valore medio 0 e deviazione standard 1. Tale funzione viene applicata sia all'insieme degli input, che all'insieme dei corrispondenti output noti. I valori degli input e degli output preprocessati hanno lo stesso ordine di grandezza, inoltre hanno lo stesso valore medio e la stessa deviazione standard. In questo modo risulta più efficiente il processo di addestramento dei modelli di regressione che spesso sono di natura stocastica ed effettuano troncamenti delle cifre decimali.

Tra i vari metodi proposti nel paragrafo 2.3 per la stima dell'accuratezza della regressione, si è scelto di utilizzare un *split sample* dividendo in 4 parti l'insieme di dati noti ed utilizzando 3/4 come insieme di addestramento e 1/4 come insieme di verifica. La scelta è caduta su questo metodo che è meno elaborato, perchè gli altri metodi, come la *cross validation* o il *bootstrapping* prevedono di effettuare più addestramenti per ogni insieme di dati noti, e questo comporta una quantità di tempo sensibilmente maggiore.

Si è scelto di addestrare contemporaneamente una rete neurale, un kriging e una rete

di funzioni a base radiale per avere un confronto del comportamento di varie tecniche di regressione. Per poter addestrare più modelli di regressione diversi da confrontare tra di loro è stato necessario portare a termine un numero considerevole di simulazione per un totale di 3600 condotti meridiani simulati. Questo numero corrisponde a 150 ipercubi latini sovrapposti che forniscono il totale degli input su cui realizzare le simulazioni. A ciascuno ipercubo aggiunto ai precedenti corrisponde un ciclo del metodo iterativo proposto durante il quale si eseguono le simulazioni corrispondenti all'input individuato (con le quali si intendono tutte le fasi oltre a quelle di soluzione, quindi meshatura, pre e post processing).

Per portare a termine queste operazioni è stato necessario impegnare una workstation con quattro processori per circa 8 giorni.

In figura 6.26 si riportano alcuni esempi delle geometrie che sono state generate dall'algoritmo e di come sono state discretizzate.

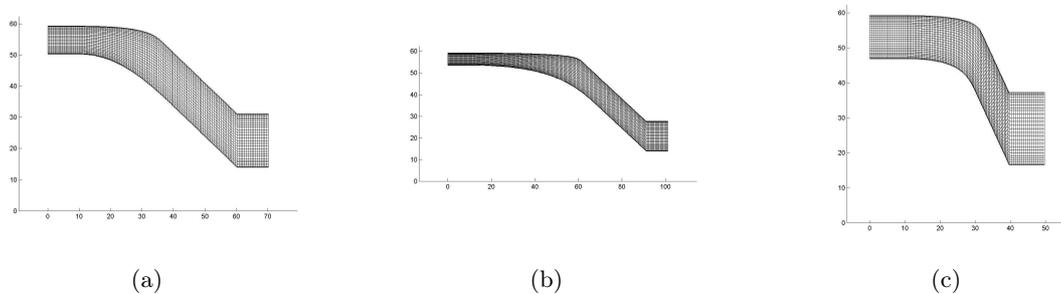


Figura 6.26: Esempi di diverse geometrie che sono state discretizzate per la realizzazione della banca dati.

Risultati addestramento reti neurali

Si è scelto di addestrare delle reti neurali ad un solo layer nascosto di neuroni con funzioni di trasferimento tangenziale e lineare rispettivamente per il layer nascosto e per quello di output. Per ciascun ciclo del metodo di creazione della banca dati sono state addestrate più reti neurali con diverse strutture in modo da ricercare la struttura più corretta per approssimare l'insieme di dati di addestramento. Sono state addestrate reti neurali con le seguenti strutture:

- per i cicli dall'1 all'86, corrispondenti a insiemi con numerosità da 24 a 2064 sono stati addestrate reti neurali con 10, 20, 30, 40 e 50 neuroni nel layer nascosto;
- per i cicli dall'87 al 150, corrispondenti a numerosità dal 2088 al 3600, sono state addestrate reti neurali con layer nascosti con 50 e 60 neuroni.

Per ciascuna struttura di rete neurale per ciascun ciclo, poi, sono state addestrate quattro diverse reti neurali. E' consigliato infatti, posta la natura stocastica degli algoritmi di

addestramento delle reti neurali, addestrare più reti e scegliere quella con l'errore minore. Tuttavia questo non sempre è possibile, perchè l'operazione di addestramento delle reti neurali, quando cresce il numero dei neuroni e il numero di elementi dell'insieme di addestramento è un'operazione lunga.

Tabella 6.8: Stime in percentuale dell'errore di approssimazione per diverse reti neurali addestrate fino al 54° ciclo di infittimento dell'insieme di dati noti.

Numero di neuroni				
10	3126.9	2284.5	2519.6	1587.6
20	2284.9	2100.1	11012	1183.7
30	2176.4	1886.7	1943.7	1898.1
40	2004.4	1732.4	1023	2060.6
50	1291.3	2555.2	2154	2349.7

Tabella 6.9: Stime in percentuale dell'errore di approssimazione per diverse reti neurali addestrate al 105° ciclo di infittimento dell'insieme di dati noti.

Numero di neuroni				
50	296.07	573.03	378.41	358.03
60	393.93	373.72	512.65	384.6

Nelle tabelle 6.8 e 6.9 sono riportati, a titolo di esempio, i valori delle stime dell'errore di approssimazione rilevate per le reti neurali addestrate rispettivamente per il 54° e il 105° ciclo del metodo per la creazione della banca dati. Il 54° ciclo corrisponde ad un insieme di 1296 valori noti che si dividono in un insieme di addestramento di 972 unità e un insieme di verifica di 324 unità. Il 105° ciclo corrisponde ad una banca dati di 2520 simulazioni; di queste i dati ricavati da 1890 sono utilizzati per l'insieme di addestramento, i restanti 630 per l'insieme di verifica.

I risultati dell'addestramento delle varie reti neurali sono presentati nella figura 6.27, dove al crescere del numero di simulazioni che vanno a comporre la banca dati, quindi del numero di cicli, è riportato, la stima dell'errore di che compie la rete neurale più precisa. Si può vedere come già dopo 42 cicli, corrispondenti a 1008 simulazioni, la stima dell'errore scende a valori considerevolmente inferiori a quelli che si avevano all'inizio. Lasciamo al paragrafo successivo i commenti per quello che riguarda l'accettazione di tali modelli come definitivi per la banca dati.

Nel grafico di figura 6.28 abbiamo invece un confronto dell'addestramento di reti neurali con strutture diverse. Sono diagrammate le linee di tendenza dell'andamento della stima dell'errore per diverse strutture delle reti neurali, cioè per diversi numeri di neuroni nel layer nascosto. Non sono stati diagrammati gli errore effettivi al procedere dei cicli per l'irregolarità

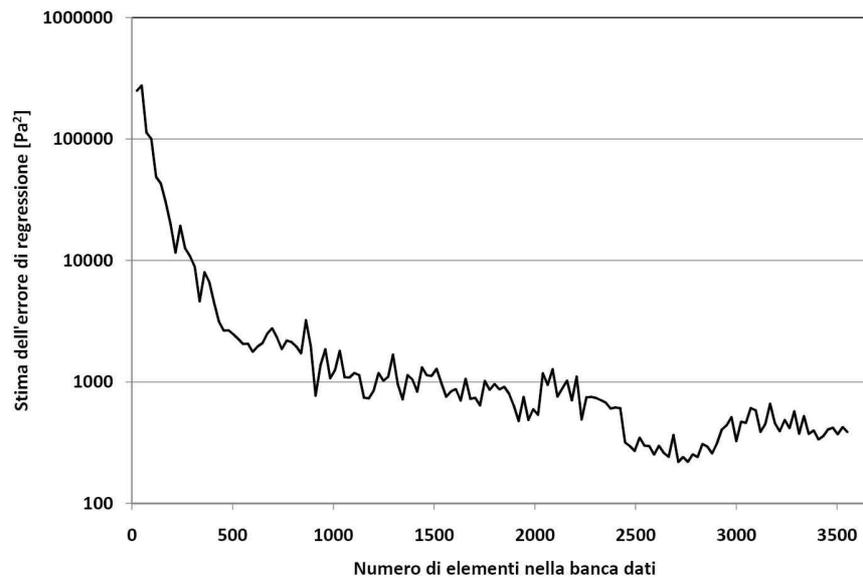


Figura 6.27: Andamento dell'errore di regressione nell'addestramento di reti neurali.

dell'andamento di questi che avrebbe reso difficile la comprensione del grafico. Le linee di tendenza, invece, consentono di intuire chiaramente l'influenza della struttura della rete sulla precisione che questa può raggiungere. Come già fatto notare nel paragrafo 4.2.5 al crescere della numerosità dell'insieme di addestramento è necessario un numero maggiore di neuroni nel layer nascosto; quindi una rete con un numero basso di neuroni risulta inadeguata se si ha come target degli errori molto bassi in fase di regressione. Viceversa se si ritiene sufficiente un errore medio più alto o si hanno a disposizione un numero limitato di simulazioni sono da preferire reti con strutture più semplici. In tale caso il tempo necessario all'algorithmo di addestramento è inferiore e anche la precisione che si può raggiungere è superiore a quella che si può raggiungere con reti di con strutture più complesse.

Risultati addestramento reti di funzioni a basi radiali

Per quello che riguarda l'addestramento della rete di funzioni a base radiale, la funzione di addestramento di Matlab *newrb* è in grado di individuare la dimensione più corretta del layer nascosto aumentando il numero di neuroni ad ogni epoca dell'addestramento fintantochè l'errore cala, e arrestando l'addestramento quando ricomincia a crescere, in modo da prevenire l'overfitting.

Si è scelto quindi di addestrare delle reti di funzioni a base radiale di tipo gaussiano. Per individuare la rete che meglio generalizza i dati dell'insieme di addestramento sono state portate avanti più prove al variare dell'ampiezza della funzione di trasferimento per individuare il valore che fornisce i risultati migliori in termini di errore di generalizzazione. I valori scelti per il parametro *Spread* che definisce l'ampiezza della funzione di trasferimento intorno al suo centro sono riportati in tabella 6.10. In particolare la grandezza *Spread* indica la distanza dal centro a cui la funzione radiale assume valore 0.5. In pratica ogni qualvolta un input sta

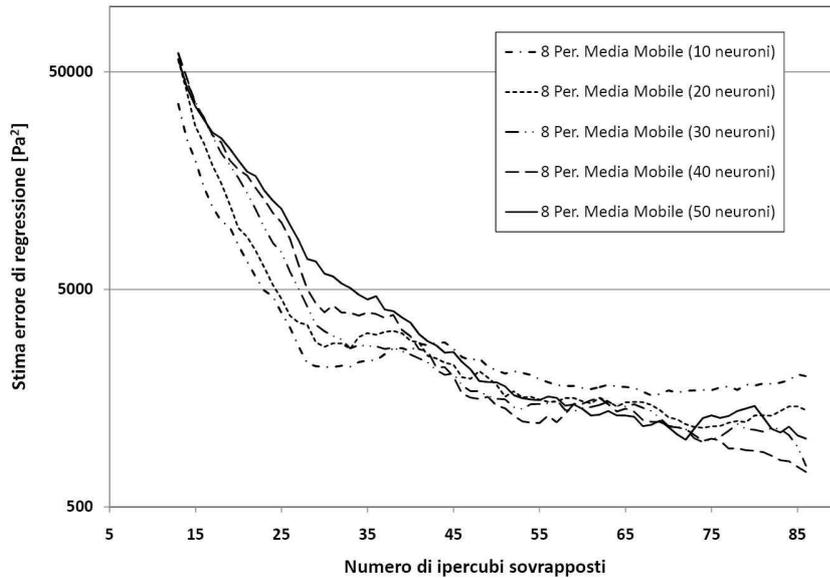


Figura 6.28: Influenza della struttura delle reti neurali sull'errore di regressione.

ad una distanza pari o inferiore a *Spread* dal vettore che rappresenta il centro di una delle funzioni che compongono la rete allora la funzione corrispondente avrà un valore superiore a 0.5.

Tabella 6.10: Valori del parametro di ampiezza testati.

Spread	2.5	5	7.5	10
---------------	-----	---	-----	----

I risultati sono presentati in figura 6.29, dove in ascissa è presente la numerosità della banca dati di simulazioni, mentre in ordinata il valore dell'errore di approssimazione. La scala è necessariamente logaritmica per permettere un'adeguata visualizzazione. Come si può notare la linea che identifica il valore 2.5 di *Spread* è sensibilmente più imprecisa degli altri valori che invece mostrano precisioni sostanzialmente comparabili; tra queste non esiste una scelta preferenziale in quanto la variazione di questo parametro non comporta variazioni nel tempo necessario all'addestramento della rete.

Risultati addestramento Kriging

La tecnica del Kriging è di origine statistica. Il Kriging prevede un modello matematico dato dalla somma di due termini: un termine che è un modello di regressione lineare dove i vari termini sono funzioni polinomiali, e un termine di natura stocastica che impone l'interpolazione dei dati noti ed è calcolato imponendo la forma della covarianza tra i punti.

Il Kriging non ha una struttura interconnessa come le reti neurali e le reti di funzioni a base radiale, ma è un modello matematico basato sulla somma di due termini di diversa natura. Per utilizzare nel modo più corretto la funzione di Kriging bisogna indicare i tipi di

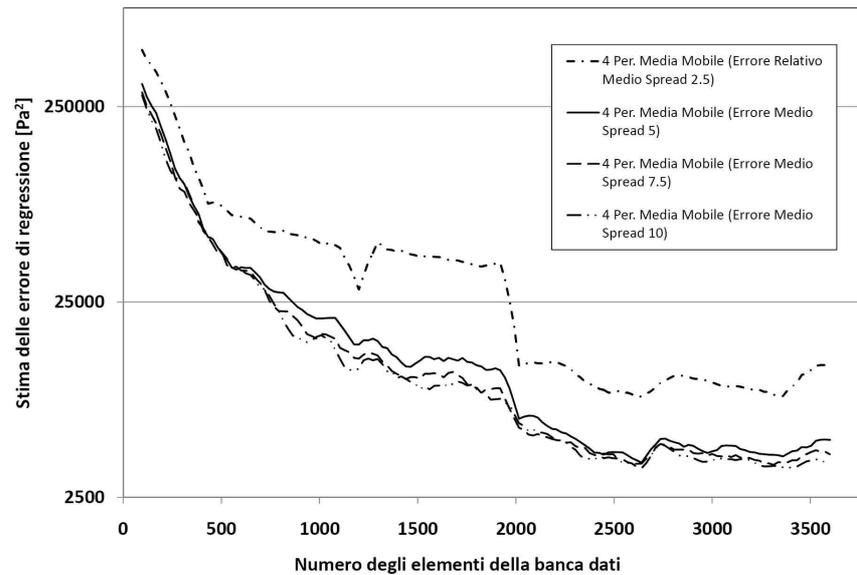


Figura 6.29: Influenza dell'ampiezza delle funzioni a base radiale sulla precisione delle RBFN.

funzioni che compongono il modello, più adatti per l'insieme di dati che si sta approssimando. E' stato quindi necessario individuare il grado della funzione polinomiale di regressione e il tipo di funzione che deve seguire la covarianza del termine stocastico tra quelli citati nella tabella 4.1.

Sono state provate funzioni di regressione di grado nullo, di primo e di secondo grado, cioè tutte quelle messe a disposizione dal toolbox Dace utilizzato (vedi il paragrafo 4.2.7 o [2]). I tre gradi della funzione polinomiale sono stati provati, in tutte le combinazioni possibili, con funzioni di covarianza gaussiane, cubiche e spline. La scelta di queste tre è motivata dal comportamento parabolico di queste in corrispondenza all'origine perchè si ritiene questo più adatto al caso in analisi.

Nelle due sottotabelle di tabella 6.5.5 sono riportati i risultati della regressione con i modelli del Kriging. Nella prima sottotabella sono indicati i minimi errori di regressione stimati per ciascun modello inteso come combinazione di grado di funzione di regressione polinomiale e di tipo di funzione di correlazione. Nella seconda sottotabella è indicata la numerosità della banca dati, quindi il livello di infittimento, necessario per raggiungere l'errore indicato nella prima sottotabella. I dati riportati mostrano come i modelli di regressione con funzione di covarianza di tipo cubico siano assolutamente inadeguati per approssimare i dati a disposizione. I modelli addestrati con funzione gaussiana e spline invece mostrano comportamenti e ordine di grandezza della stima dell'errore confrontabili; tuttavia il modello di correlazione gaussiano è quello che garantisce la miglior approssimazione e come la funzione polinomiale di secondo grado sia, tra le varie, quella che fornisce errori minori.

E' interessante notare che le maggiori precisioni sono state ottenute per una numerosità della banca dati pari a 1752 per tutte e tre i gradi delle funzioni di regressione, sia per le funzioni di correlazione gaussiane che spline. Una possibile spiegazione può essere cercata in una particolare combinazione di dati tra insieme di addestramento e insieme di verifica che

Tabella 6.11: Minimo errore raggiunto con ciascun modello di Kriging raggiunto (a), e numerosità per cui è stato raggiunto tale errore.

(a) Minimo errore stimato.

		Grado del polinomio di regressione		
		Nulla	Primo	Secondo
Funzione di correlazione	Cubica	12741	25377	26144
	Spline	4805	4092	3855
	Gaussiana	2971	3103	2619

(b) Numerosità della banca dati.

		Grado del polinomio di regressione		
		Nulla	Primo	Secondo
Funzione di correlazione	Cubica	120	120	936
	Spline	1752	1752	1752
	Gaussiana	1752	1752	1752

rende l'approssimazione tramite Kriging insensibile al modello scelto.

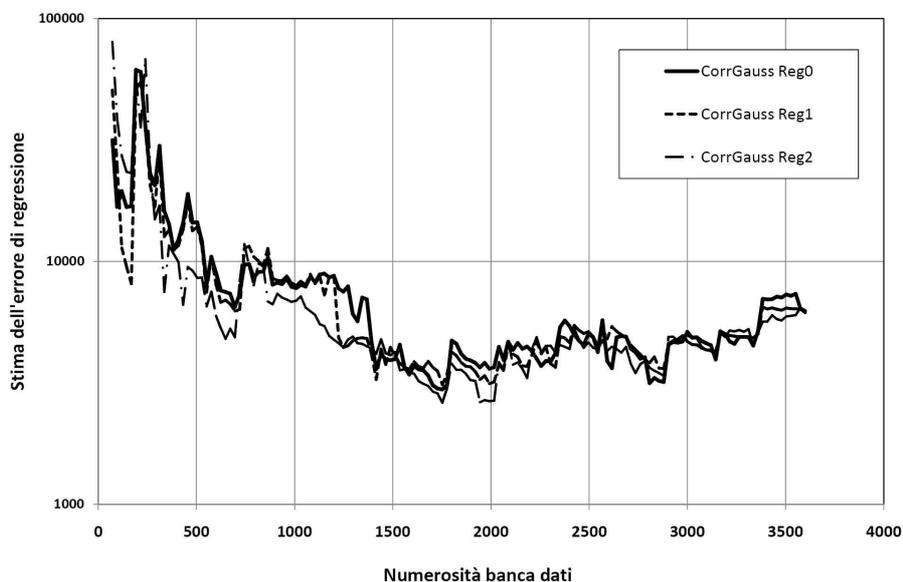


Figura 6.30: Risultati dell'addestramento di diversi modelli di Kriging.

Nella figura 6.30 sono diagrammate le storie di convergenza di tre modelli di Kriging, tutti e tre con funzione di correlazione che definisce il termine stocastico del modello di tipo gaussiano. Come è possibile vedere il grado del polinomio di regressione non ha un'influenza sostanziale sulla precisione che può essere raggiunta dal modello in sede di approssimazione.

6.5.6 Accettazione della banca dati

Nella figura 6.31 vediamo un confronto tra i risultati in termini di stime dell'errore di regressione dei vari metodi dove si può vedere un confronto tra l'andamento dell'errore di regressione sull'approssimazione dei risultati delle simulazioni della banca dati dei tre metodi proposti. Come si può notare le reti neurali sono sensibilmente più precise delle reti di funzioni a base radiale o del Kriging.

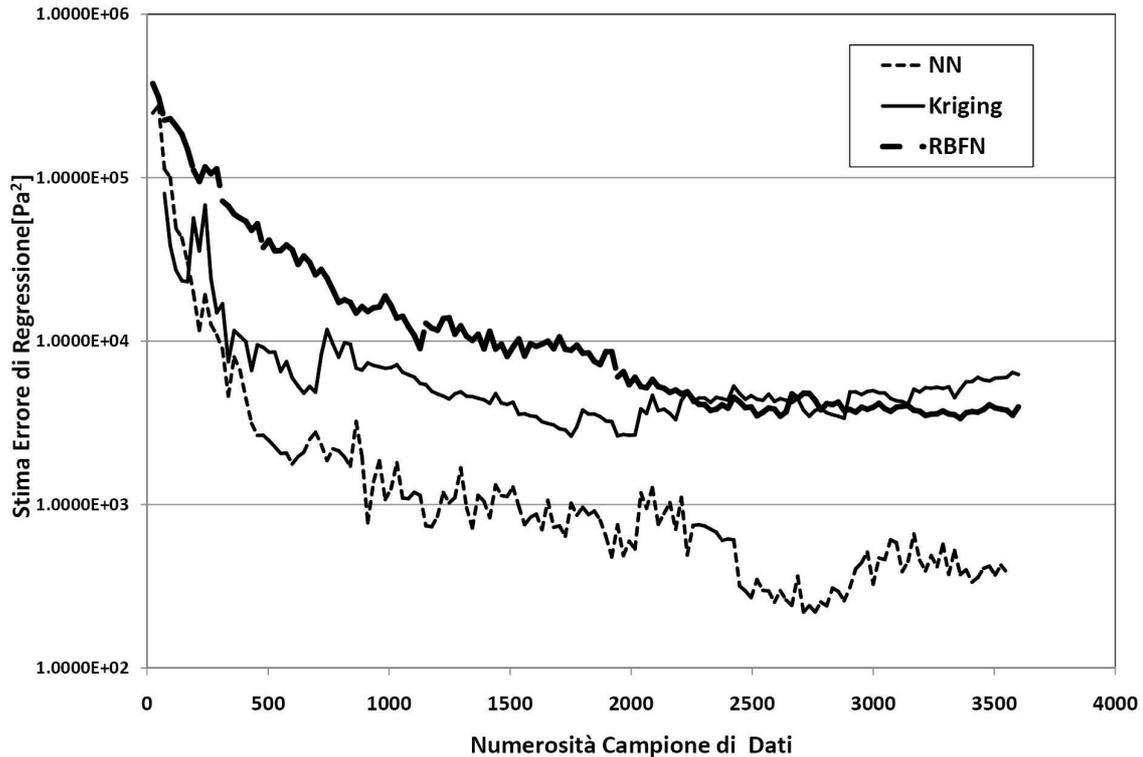


Figura 6.31: Confronto tra i vari modelli di regressione.

I risultati sono qui proposti anche per confrontare le diverse efficienze in fase di regressione dei tre metodi utilizzati; per cui il numero di simulazioni che è stato portato avanti è decisamente elevato e il tempo dedicato all'addestramento di modelli di regressione diversi è notevole. In ambito industriale, dove le esigenze in termini di tempo sono stringenti l'approccio è sicuramente differente e si orienta sull'utilizzo di una sola tecnica di approssimazione, la cui scelta è dettata dall'esperienza. Da quanto visto nella figura 6.31 dove è possibile confrontare questi metodi le reti neurali sembrano dare i risultati migliori.

Il numero di cicli necessari per accettare la banca dati è sicuramente inferiore a quello utilizzato in questa sede: sarà sufficiente arrestare le simulazioni prima di quanto fatto in questa sede. Per arrestare la creazione della banca dati è necessario individuare un criterio di accettazione. Sicuramente l'osservazione dei diagrammi come quelli delle figure precedenti, che mostrano la precisione raggiunta in sede di generalizzazione dai dati raccolti, fornisce delle indicazioni sull'opportunità di proseguire o meno con la raccolta dei dati.

Prendendo a riferimento il grafico dell'andamento dell'errore in fase di addestramento delle

reti neurali ad un singolo layer nascosto con 20 neuroni in quest'ultimo che si riporta in figura 6.32 si può notare come l'andamento dell'errore di regressione scenda bruscamente con l'aggiunta dei primi *strati* di punti all'insieme che compone la banca dati, e si è stabilizzato su un errore, più o meno alto. Tale andamento indica che non è possibile incrementare sensibilmente la precisione del modello aggiungendo altri dati. Raggiungere questo livello di stabilizzazione dell'errore è sicuramente un'indicazione di non proseguire oltre con il campionamento. Volendo arrestare la creazione del modello anche prima è necessario avere cognizione fisica del livello di errore che si è raggiunto.

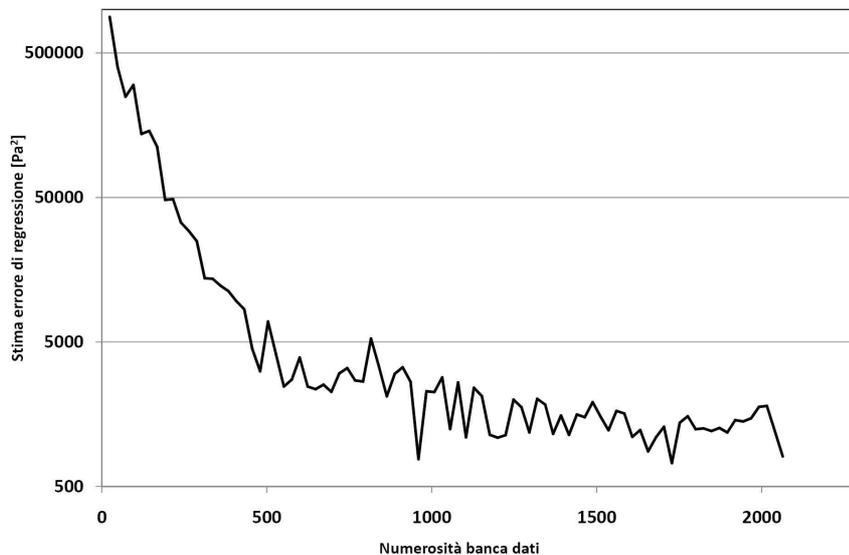


Figura 6.32: Andamento della stima dell'errore di regressione al variare della numerosità della banca dati per una rete neurale con un layer nascosto con 20 neuroni.

La stima dell'errore che è stata utilizzata è l'errore quadratico medio. La radice quadrata di questo consente di trovare un valore fisico dell'errore commesso che fornisce un'ordine di grandezza dell'errore che il modello commette in fase di predizione. Tale valore può essere utile per giudicare il livello di precisione raggiunto, stabilire se tale errore è accettabile o meno e se è necessario effettuare altre simulazioni. Nella 6.12 sono riportate varie colonne che servono ad intuire come è possibile fare uso di questi dati per avere una stima della precisione che ci si può attendere in fase di regressione durante la raccolta dei dati.

Nella prima colonna è indicato il numero di simulazioni che vanno a formare la banca dati; nella seconda colonna sono indicati i cicli che sono stati realizzati e quindi il numero di ipercubi latini che sono stati sovrapposti per individuare gli input della simulazioni della banca dati. Il numero di neuroni sta ad indicare la struttura delle reti neurali che sono state addestrate con quei dati. La colonna successiva è la stima dell'errore che il modello commette in fase di regressione e quindi lo scarto quadratico medio che si registra sull'insieme di verifica, l'unità di misura di questo è Pa^2 . La colonna successiva, denominata con *Errore sulla misura*,

Tabella 6.12: Valori dell'errore di regressione per una rete neurale ad un layer nascosto di 20 neuroni.

Npunti	Cicli	Numero Neuroni	Stima Errore [Pa ²]	Errore sulla misura [Pa]	Errore sulla media [%]	Errore sul minimo [%]
24	1	20	888740	942.730078	63.0167164	157.1216797
144	6	20	144440	380.0526279	25.4045874	63.34210466
264	11	20	29283	171.122763	11.4386874	28.52046049
384	16	20	11241	106.0235823	7.08713785	17.67059705
504	21	20	6925.3	83.21838739	5.56272643	13.86973123
624	26	20	2457.2	49.57015231	3.31351285	8.261692052
744	31	20	3299.3	57.43953342	3.839541	9.573255571
864	36	20	2102.5	45.85302607	3.06504185	7.642171012
984	41	20	2280.9	47.75876883	3.19243107	7.959794805
1104	46	20	1094.4	33.08171701	2.21134472	5.513619501
1224	51	20	1133.5	33.66749174	2.25050078	5.611248623
1344	56	20	1844.1	42.94298546	2.87052042	7.157164243
1464	61	20	1506.6	38.81494558	2.59458192	6.469157596
1584	66	20	1601.4	40.01749617	2.67496632	6.669582696
1704	71	20	1294.1	35.97360143	2.4046525	5.995600239
1824	76	20	1260.9	35.50915375	2.37360653	5.918192292
1944	81	20	1409.3	37.54064464	2.50940138	6.256774107
2064	86	20	807.6	28.41830396	1.89961925	4.736383993

rappresenta la radice quadrata della stima dell'errore, e quindi il valore dell'errore così come ce lo si potrebbe attendere su una possibile previsione fatta dal modello su un punto qualsiasi dello spazio di indagine, e per questo l'unità di misura di questo valore non è altro che Pa. Per capire l'impatto che un errore di questa entità potrebbe avere su una valore estrapolato con la rete neurale, è possibile ricavare il valore percentuale dell'errore rispetto al valore medio dei risultati che sono stati utilizzati per addestrare la rete neurale (quindi di fatto l'errore percentuale sul valore medio di perdita di energia) e il valore percentuale dell'errore che si commette sulla minima misura che è stata rilevata della banca dati. Quest'ultimo di fatto rappresenta il peso di un errore come quello stimato nel caso più penalizzante.

Anche nelle colonne che concernono gli errori relativi si può vedere come procede l'andamento dell'influenza che un errore può avere sulla previsione del modello al variare delle iterazioni e fornisce informazioni preziose per capire se la precisione raggiunta con quel metodo può essere ulteriormente aumentata o se è sufficiente per quello che ci si propone.

Tuttavia questi ragionamenti sono fatti su i valori medi di tutte le misure effettuate e non tengono conto quindi della varianza e della varianza dell'errore. E' necessario infatti indagare anche la varianza dell'errore in fase di regressione e controllare che questa non sia eccessiva. In questa tesi la dispersione degli errori è stata controllata grazie alla funzione di Matlab *postreg* che consente di diagrammare i valori attesi degli output dell'insieme di verifica contro i valori stimati dal modello. La regressione è corretta se i dati si dispongono lungo bisettrice del primo quadrante; qualora i dati siano dispersi attorno alla retta allora la varianza dell'errore

di regressione è sicuramente alta.

6.5.7 Utilizzo dei dati: scelta del canale meridiano D

I dati a disposizione possono essere utilizzati secondo le procedure proposte nel paragrafo 5.3, quindi possono essere utilizzate per confrontare due o più geometrie differenti, oppure per effettuare un'ottimizzazione vincolata e ricavare la forma del condotto che minimizza le perdite e può essere accoppiato con la girante progettata.

Con quest'ultimo scopo è stato possibile ricavare la forma del condotto meridiano che minimizza le perdite distribuite e può essere accoppiato con la girante che è stata proposta nel paragrafo 6.4 (Girante **P**). Si è utilizzato quindi un algoritmo di ricerca agli sciami di particelle modificato (*GCPSO*) come riportato nel paragrafo 3.5.1 per individuare i valori dei parametri che caratterizzano tale condotto. È stato utilizzato uno sciame di 100 particelle, con coefficienti di accelerazione c_1, c_2 entrambi pari a 1.2; peso dell'inerzia sarà pari a $\omega = 0.7298$. I valori di soglia utilizzati sono $f_c = 5, s_c = 20$, per evitare una convergenza troppo precipitosa del metodo verso un minimo. Per ogni ricerca dell'ottimo l'algoritmo è stato applicato 5 volte per verificare che l'algoritmo non fosse rimasto intrappolato in un minimo locale.

La ricerca dell'ottimo effettuata è stata vincolata in tre parametri in modo da selezionare il condotto meridiano di un diffusore in grado di accogliere la girante **P**.

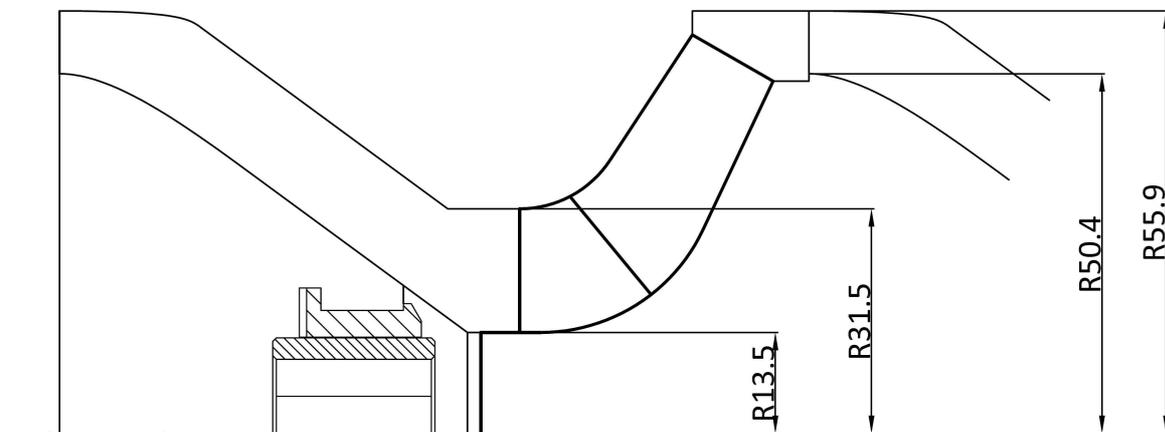


Figura 6.33: Schematizzazione dell'accostamento di girante e diffusore.

Nella tabella 6.13 sono riportati i valori degli estremi di variazione dei valori dei parametri utilizzati nella ricerca del condotto meridiano ottimo. Sono evidenziati i parametri che sono stati fissati ad un valore costante per consentire l'accoppiamento con la girante **P** secondo il disegno di figura 6.33. Sono stati bloccati i parametri che riguardano le dimensioni vincolate, quindi diametro del mozzo, diametro della bocca di aspirazione e di scarico del diffusore. Nella figura 6.33 si può notare come il diametro esterno della bocca di aspirazione del diffusore è stato posto pari al diametro massimo che poteva essere utilizzato per il diffusore, mentre il diametro interno è stato scelto in modo da realizzare un corretto accoppiamento con lo scarico

Tabella 6.13: Vincoli utilizzati nella ricerca del condotto ottimo.

Parametro	Estremo Inferiore	Estremo Superiore
$\frac{A_x}{R}$	0.5	1.5
$\frac{B_2}{R}$	0.2928	
$\frac{R_m}{R}$	0.2388	
$\frac{B_1}{B_2}$	0.5076	
$\frac{L_{H3}}{L_{ungH}}$	0.4	0.9
$\frac{L_{S3}}{L_{ungS}}$	0.4	0.9
$\frac{L_{H1}}{A_x}$	0.2	0.6

della girante. In questo caso le dimensioni sono tali che il rapporto tra l'area netta dell'ingresso del condotto del diffusore e l'area netta del condotto di scarico della girante è pari a circa 0.7, quindi il flusso nel traferro tra girante e diffusore-riconvogliatore risulta accelerato.

Tabella 6.14: Valori dei parametri che individuano la geometria ottima

Parametro	$\frac{A_x}{R}$	$\frac{B_2}{R}$	$\frac{R_m}{R}$	$\frac{B_1}{B_2}$	$\frac{L_{h3}}{L_{ung_h}}$	$\frac{L_{s3}}{L_{ung_s}}$	$\frac{L_{h1}}{A_x}$
Valore	1.0924	0.2	0.4	0.9	0.2928	0.2388	0.5076

L'ottimizzazione secondo i vincoli di tabella 6.13 ha portato ad individuare un condotto caratterizzato dai valori riportati in tabella 6.14. La geometria individuata è stata simulata ed è stata ricavata una perdita di pressione totale pari a 750 Pa contro una perdita di 690 Pa circa predetta dalla rete neurale. Osservando i valori ricavati con l'ottimizzazione si può notare come la geometria che minimizza le perdite con i vincoli presentati sopra, utilizzi per tre parametri dei valori che stanno all'estremo dell'intervallo di variazione; tale geometria è quella che, per un rapporto tra ingombro assiale e ingombro radiale pari a 1.093, mantiene la sezione di attraversamento più ampia; fatto questo comprensibile, in quanto queste proporzioni di condotto non introducono distacchi e quindi le perdite sono comandate dalla velocità di attraversamento.

La geometria ricavata dall'ottimizzazione, fornisce un condotto che può essere utilizzato in sede di progettazione per sagomare le generatrici del diffusore-riconvogliatore; la scelta di questo condotto, a parte i vincoli imposti sopra, è guidata solo dall'aspetto delle perdite di energia nel convogliatore.

Risulta più interessante accedere alla banca dati per vedere la relazione che sussiste tra diverse esigenze che vengono considerate in sede di progettazione. In prima istanza a guidare la progettazione è la necessità di raggiungere lo scambio di energia minimo. Oltre a questa prima istanza, durante la progettazione di questo tipo di macchine le altre richieste che sono tenute in considerazione sono le perdite e i vincoli di ingombro. Le perdite devono essere

contenute per raggiungere gli standard di efficienza più elevati. Questo però va contro la necessità di limitare il più possibile gli ingombri delle macchine; infatti la diminuzione degli ingombri richiede spesso di imporre al fluido percorsi più tortuosi e quindi velocità e perdite maggiori.

Per quello che riguarda il target di efficienza che si vuole raggiungere con questo tipo di macchine si è già detto quando si è parlato dei vincoli introdotti in sede di progettazione. Gli ingombri radiali sono un vincolo di progetto che non può essere aumentato; si è già utilizzato l'ingombro massimo e quindi non è possibile utilizzare tale parametro per aumentare le prestazioni. Rimane libero come parametro dimensionale l'ingombro assiale del canale meridiano del diffusore-riconvogliatore.

Le istanze da valutare in sede di progettazione, tra di loro contrastanti, sono l'ingombro assiale e le perdite di energia. Al fine di investigare il mutuo andamento di queste due istanze è stato possibile ricavare agevolmente, grazie alla banca dati, la geometria e le corrispondenti perdite, per vari valori dell'ingombro assiale. Sono state realizzate diverse ottimizzazioni del condotto meridiano, vincolando il parametro $\frac{A_x}{R}$ per vari valori, oltre ai tre parametri $\frac{B_2}{R}$, $\frac{R_m}{R}$ e $\frac{B_1}{B_2}$ che garantiscono l'accoppiamento con la girante **P**.

In presenza di un flusso non stallato, quello che ci si aspetta è che al crescere dell'ingombro assiale le perdite calino perchè la geometria del condotto diventa meno acuta e quindi calano le velocità e le perdite di energia, fino ad un minimo (che corrisponde a quello individuato nell'ottimizzazione precedente, per poi ricominciare a salire per il crescere eccessivo della superficie bagnata.

Nella tabella 6.15 sono riportati i valori di perdita per la geometria ottima che corrisponde al che ottimizza tali perdite in corrispondenza dell'ingombro assiale segnato. Il grafico di figura 6.34 riporta tali dati e consente di fare alcune considerazioni utili per la scelta dell'ingombro assiale in sede di progettazione.

L'osservazione del grafico di figura 6.34 che riporta il valore delle perdite per la geometria ottima con un determinato ingombro assiale, evidenzia un andamento come quello atteso. La geometria che minimizza le perdite in assoluto ha un'ingombro assiale di 1.09 volte l'ingombro radiale come trovato nell'ottimizzazione di tabella 6.14. L'osservazione di tale grafico è molto importante perchè consente di fare una scelta ragionata dell'ingombro assiale del condotto della macchina che si vuole utilizzare. Le geometrie con ingombro assiale superiore a quello ottimo, quindi con $\frac{A_x}{R} > 1.1$ sono da ritenersi inadeguate, perchè sono penalizzanti sia per le perdite che per l'ingombro assiale. Per valori del parametro inferiori a 1.1, l'andamento le due esigenze hanno degli andamenti opposti: al calare dell'ingombro crescono le perdite. Così nell'intento di contenere gli ingombri si penalizzano le perdite, e viceversa. Non avendo indicazioni su quale di queste due istanze, e in che misura, privilegiare rispetto all'altra, si può osservare che per valori del parametro di ingombro assiale compresi tra 0.95 e 1.1 le perdite non subiscono sostanziali variazioni. In quel tratto si può affermare che l'influenza dell'ingombro assiale sulle perdite di energia è contenuta e anzi si può dire che sono indipendenti. Per valori inferiori dell'ingombro assiale le perdite salgono considerevolmente e quindi le due istanze non possono più essere considerate indipendenti.

Tabella 6.15: Risultati ricavati per le perdite all'interno del condotto meridiano al variare dell'ingombro assiale.

Valore perdita [Pa]	$\frac{A_x}{R}$	Valore perdita [Pa]	$\frac{A_x}{R}$
973.35	0.5	691.35	1.0172
895.2	0.53448	689.56	1.0517
843.64	0.56897	688.85	1.0862
810.92	0.60345	689.08	1.1207
796.64	0.63793	690.18	1.1552
789.87	0.67241	692.08	1.1897
782.02	0.7069	694.68	1.2241
769.78	0.74138	697.97	1.2586
754.24	0.77586	701.89	1.2931
738.67	0.81034	706.39	1.3276
724.93	0.84483	711.44	1.3621
713.8	0.87931	717.03	1.3966
705.24	0.91379	723.11	1.431
698.9	0.94828	729.72	1.4655
694.38	0.98276	736.84	1.5

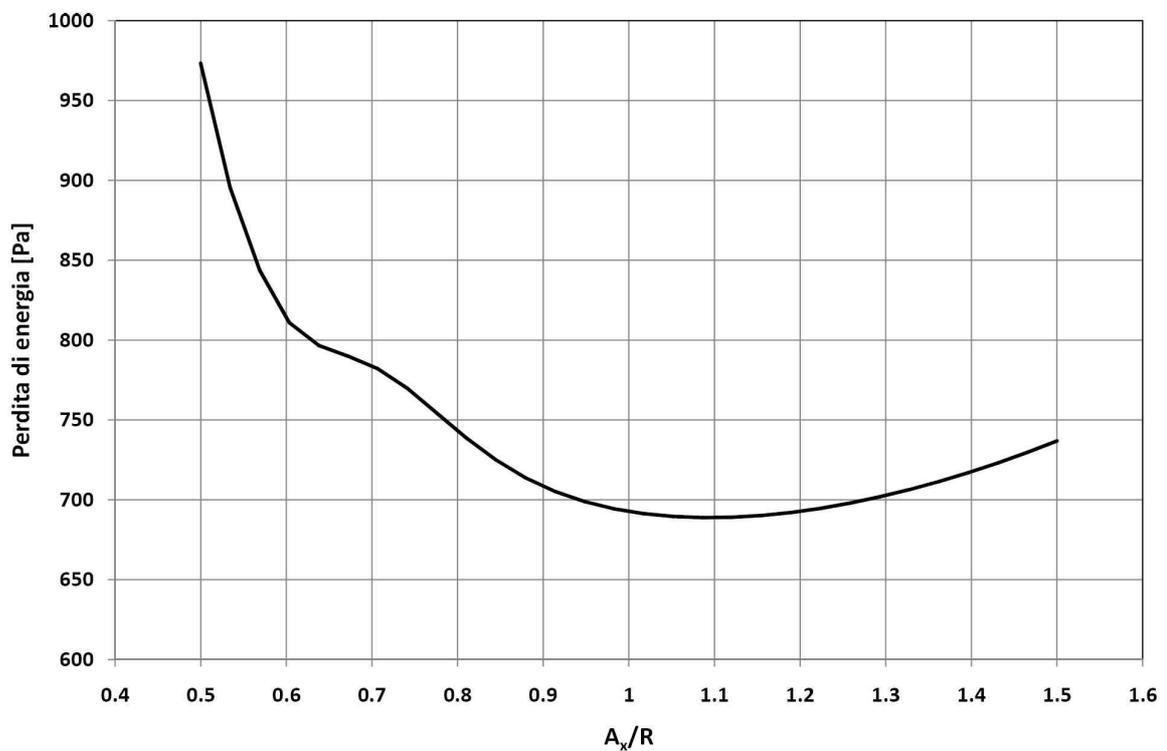


Figura 6.34: Influenza dell'ingombro assiale sulle perdite nel condotto meridiano.

Considerando che un ingombro assiale elevato non è da considerarsi favorevole(nelle macchine composte da un numero elevato di stadi la lunghezza eccessiva dell'albero potrebbe

risentire termini rigidità torsionale e velocità critiche quindi è da contenere) si è scelto quindi un condotto meridiano con un parametro $\frac{A_x}{R} = 0.95$. l'insieme dei valori del condotto sono riportati in tabella 6.16 che è stato identificato come condotto **D**; per quello che è stato osservato poche righe fa la scelta è motivata dal fatto che la riduzione dell'ingombro assiale a questo valore, rispetto al valore 1.1, garantisce perdite minime contro una riduzione di oltre il 10% dell'ingombro assiale e quindi rappresenta il giusto compromesso tra le due esigenze.

Tabella 6.16: Parametri scelti per il condotto meridiano **D** del diffusore.

Parametro	$\frac{A_x}{R}$	$\frac{B_2}{R}$	$\frac{R_m}{R}$	$\frac{B_1}{B_2}$	$\frac{L_{h3}}{L_{ung_h}}$	$\frac{L_{s3}}{L_{ung_s}}$	$\frac{L_{h1}}{A_x}$
Valore	0.94828	0.2	0.4	0.9	0.2928	0.2388	0.5076

6.6 Progettazione delle pale del diffusore-convogliatore

La progettazione delle pale del diffusore parte, necessariamente, dalla conoscenza dei dati geometrici e di funzionamento della girante **P**. Effettuata la scelta del condotto meridiano del diffusore-convogliatore è possibile procedere con il dimensionamento delle pale del diffusore secondo le procedure presentate in 2.2.4.

Così come per la girante, la progettazione è avvenuta per tentativi successivi fino al raggiungimento di una soluzione che rispettasse gli obiettivi che erano stati posti, anche il progetto del diffusore è avvenuta per step successivi. Dopo ciascun progetto effettuato si è fatto uso di simulazioni CFD per verificare il comportamento fluidodinamico.

La progettazione del diffusore è stata realizzata tenendo conto dell'interazione tra organo statorico e organo rotante. Le simulazioni CFD che sono state utilizzate per la progettazione sono molto complesse e richiedono risorse di calcolo e tempi non trascurabili. Infatti per tenere conto dei fenomeni di interazione tra organo rotante e organo statorico nella simulazione di uno stadio di macchina ogni diffusore progettato deve essere validato ai codici di calcolo nella configurazione accoppiata con la girante **P** con la quale è previsto funzionare. E' stato necessario effettuare simulazioni dell'insieme girante diffusore-convogliatore per investigare il funzionamento del convogliatore progettato ogni volta che è stato realizzato un nuovo progetto. Per tener conto a pieno dei fenomeni di interazione tra le due schiere si sono effettuate delle simulazioni in modalità Transient; è stato necessario discretizzare girante e diffusore nella loro interezza. Infatti le due schiere palari che devono essere accoppiate hanno numeri di pale differenti e primi tra loro. Quindi non è possibile simulare l'accoppiamento degli organi sfruttando le periodicità che derivano dall'assialsimmetria dei componenti.

Anche la geometria dei diffusori stessi rappresenta un vicolo che limita l'utilizzo di simulazioni semplificate per affrontare il problema. Infatti i diffusori a spessore variabile sfruttano il fattore di ostruzione per controllare lo sviluppo della corrente nel condotto interpalare. Semplificare il modello da analizzare simulando un diffusore con un numero di pale pari a

quello della girante e diverso da quello che nella realtà viene applicato cambia in modo inaccettabile il coefficiente di ostruzione palare e quindi i risultati di una simulazione sarebbero da considerarsi non attinenti al modello che si vuole progettare. Una simulazione di questo tipo sarebbe altresì accettabile per i diffusori a spessore costante dove la semplificazione introdotta è meno importante, e quindi, interpretando con senso critico i risultati delle simulazioni, è possibile estrapolare da simulazioni semplificate risultati per la simulazione dello stadio completo. Tuttavia non è assolutamente applicabile per stadi di pompa sommersa che fanno uso di diffusori-convogliatori a spessore variabile.

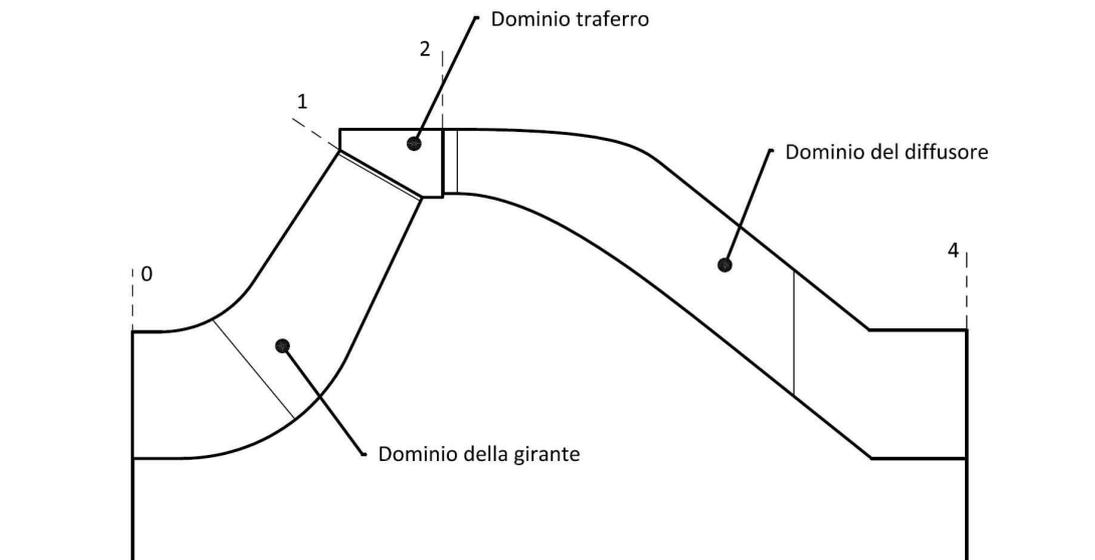


Figura 6.35: Schematizzazione del volume discretizzato per la modellazione dello stadio completo di pompa verticale sommersa: divisione in domini e superfici di riferimento.

Ciascuna simulazione è composta da tre mesh accostate come in figura 6.35; da tale figura si intuisce anche come è stato modellato il condotto libero tra la bocca di efflusso della girante e il bordo di ingresso del diffusore e quali sono le superfici di riferimento utilizzate per individuare le varie parti della mesh. Non sono state simulate tutte le parti tra girante e diffusore e girante e cassa, perchè ritenute poco influenti per la fluidodinamica del diffusore. Tutto l'insieme della macchina è stato simulato, come è già stato detto con un'interfaccia tra organi rotanti e organi statorici di tipo *Transient*; il passo temporale delle simulazioni è stato fissato in $1.72e-04$ s, che corrisponde nella rotazione della girante ad un passo angolare di 3° , tale da consentire un numero di Courant medio inferiore a 5. Ciascun passo angolare è stato risolto in 6 iterazioni (*coefficient loop*, vedi [54]), e la simulazione è stata portata avanti per 240 *timestep*, che corrispondono a 2 giri completi della girante intorno all'asse. Il modello di turbolenza utilizzato è SST, coerentemente con quanto fatto per le simulazioni della sola giranti. Tutta la discretizzazione che è stata svolta ha richiesto 2050000 nodi con una mesh di tipo esaedrico strutturata, ripartiti rispettivamente in 600000 per la mesh della girante,

1300000 per il diffusore e 150000 per il traferro di collegamento tra girante e diffusore.

Questi impostazioni consentono di raggiungere una soluzione con risultati stabili con circa 100 ore di calcolo utilizzando 3 processori e 4 Gb di memoria Ram su una workstation; i risultati sono stati considerati stabili quando i residui massimi dell'equazione di continuità scende sotto il limite di $1e-4$ e l'andamento della prevalenza elaborata dalla macchina al passare dei timestep manifesta un andamento oscillante periodico.

Raggiunti questi obiettivi, sono stati salvati 5 risultati ad intervallati di 5 timestep ciascuno; i risultati sono stati calcolati come i valori medi dei 5 risultati.

6.6.1 Tipologia di diffusore scelta

Prima di presentare la geometria che è stata individuata e suggerita all'azienda finanziatrice, è bene fare una precisazione riguardo alla tipologia di parametrizzazione che si è scelto di adottare per la progettazione tra le due presentate nel paragrafo 2.2.4.

Il diffusore che si è scelto di progettare è a spessore variabile perchè tale tipologia ha dato i risultati migliori per quello che riguarda l'accoppiamento con la girante **P**. Riportiamo in questo paragrafo le considerazioni, a livello qualitativo, che hanno portato alla scelta del diffusore a spessore variabile a scapito di quello a spessore costante.

Per discriminare tra i vari tipi di diffusore sono state portate avanti delle simulazioni per investigare il comportamento fluidodinamico di diffusori isolati. Tali simulazioni sono state effettuate su un dominio della sola pala al quale è stato aggiunto un tratto di condotto assiale fittizio a monte del bordo di ingresso. Il dominio fluido che è stato discretizzato corrisponde a quello di figura 6.20. In ingresso è stata imposta una velocità con direzione pari alla direzione media che è stata registrata in uscita dalla girante in corrispondenza della sezione di rilevamento a valle della girante in sezione 6.9.

Sotto queste premesse le perdite fatte registrare dal diffusore a spessore costante si sono mostrate sempre inferiori a quelle fatte registrare dal diffusore a spessore variabile se simulati nel comportamento isolato. Inoltre, a parità di angolo della corrente assoluta a valle della schiera del diffusore, l'angolo costruttivo richiesto della pala a spessore variabile è superiore rispetto a quello della pala spessore costante. In altre parole per raddrizzare completamente la corrente a valle del diffusore è necessario imporre alla pala con spessore variabile un overturning superiore di 10° - 13° rispetto a quello della pala a spessore costante.

In entrambe i casi la corrente è stabile e regolare lungo tutte le sezioni del condotto interpalare. Le perdite superiori che si registrano nel condotto a spessore variabile sono quindi da attribuirsi alle velocità di attraversamento maggiori. Infatti ulteriori prove sono state condotte con diffusori progettati con profili a SV dove si sono imposte delle velocità di attraversamento più contenute, ottenendo quindi dei profili di pala più sottili, e le perdite che si sono registrate sono state più basse.

Al contrario simulazioni di diffusori preceduti dalla girante, quindi secondo la configurazione utilizzata in sede di progettazione, evidenziano risultati opposti: i diffusori a spessore a variabile garantiscono allo stadio della pompa sommersa un rendimento complessivo mediamente più alto di 5 punti percentuali rispetto ad uno stadio che fa uso di un diffusore

a spessore costante. Le simulazioni del diffusore a spessore costante accoppiate con la girante, evidenziano degli intensi ricircoli che interessano quasi tutto il condotto interpalare del diffusore in corrispondenza del mozzo e sono caratterizzati da perdite intense di energia totale. I diffusori a spessore variabile invece mantengono più stabile la corrente all'interno del canale interpalare in virtù delle velocità di attraversamento più alte e controllate grazie all'andamento degli spessori palari.

Una graduatoria di merito del funzionamento di questi diffusori basata sui dati ricavati dalle simulazioni dei diffusori isolati risulterebbe controproducente in sede di progettazione, perchè non tiene conto dell'interazione completa tra girante e diffusore. La presenza della girante a monte, infatti, con un flusso più disturbato e inclinato sul piano meridiano rispetto alla direzione assiale, destabilizza la corrente nella sezione al mozzo del diffusore e fanno registrare perdite maggiori.

6.6.2 Progettazione del diffusore

In questa sezione si presentano le scelte fatte in sede di progettazione e i risultati che hanno portato a proporre il progetto del diffusore-convogliatore. Ogni nuovo tentativo che è stato realizzato è stato validato ai codici di calcolo per la portata di $60 \text{ m}^3/\text{h}$ in modo da ricevere dalla simulazione CFD il feedback per la progettazione; diversamente da quanto fatto in sede di progettazione della girante, in cui venivano simulate più portate di funzionamento per ottenere indicazioni anche nel comportamento fuori progetto. Il tempo e le risorse necessarie per effettuare una simulazione completa non consentono di realizzare in tempi accettabili simulazioni che consentano di investigare anche il comportamento di fuori progetto della macchina completa. Si è ritenuto sufficiente comunque progettare il diffusore nell'interazione con il diffusore solo alla portata di progetto di $60 \text{ m}^3/\text{h}$ in quanto questa è la portata per la quale sono stati fissati i target in termini di prestazioni da raggiungere.

La ristrettezza di tempi e la necessità di ottenere risultati in tempi contenuti, non ha altresì consentito di effettuare delle simulazioni comparative per investigare l'influenza dei singoli parametri che definiscono la geometria del diffusore. La scelta di molti parametri è motivata da considerazioni e valutazioni non validate da un'effettiva indagine numerica del comportamento della pala al variare della geometria identificata con i valori dei parametri.

La geometria che si va a presentare, denominata con Diffusore **SV**, risponde ai requisiti prestazionali imposti come target in sede di progettazione, ma è una geometria che non rappresenta l'ottimo in assoluto.

Si è scelto di progettare diffusori con 7 pale per rispettare la regola che suggerisce numeri primi tra di loro nelle schiere palari in moto relativo tra loro. Nella figura 6.36 sono riportate le dimensioni di massima del canale meridiano del diffusore, la posizione dei bordi di ingresso e di uscita della pala e come questi si collocano rispetto al volume che è stato modellato per le simulazioni CFD. Sempre nella stessa figura 6.36 sono visualizzate anche come si collocano, a ridosso del mozzo, i particolari delle tenute intorno all'albero che è necessario prevedere. Come si vede esiste spazio sufficiente per alloggiare tali particolari.

I parametri **NC**, **ND** e α , che designano l'andamento della derivata $\partial r C_u \setminus \partial m$ lungo la

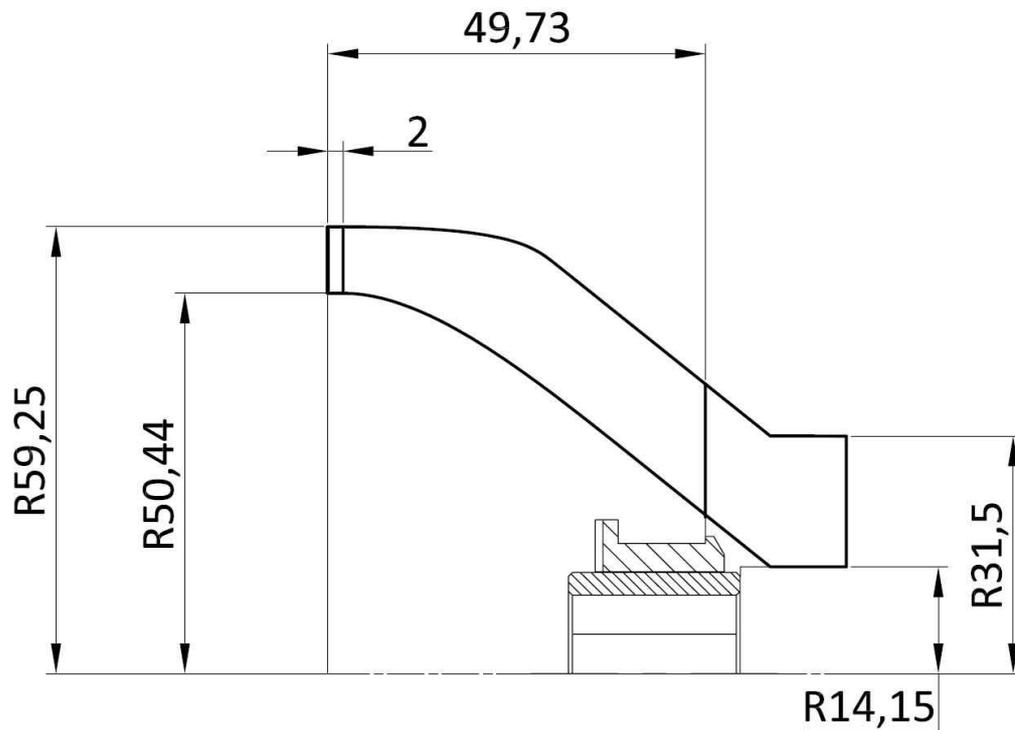


Figura 6.36: Schizzo con le indicazioni delle dimensioni principali della palettatura.

linea media della sezione di riferimento, sono stati scelti in modo da garantire un'andamento costante di tale derivata. La forma della pala nel piano meridiano suggerisce che la pala può lavorare uniformemente lungo tutto il suo sviluppo e quindi sembra appropriato imporre un valore costante di questa derivata. I valori scelti sono quindi riportati in tabella 6.17.

Tabella 6.17: Parametri utilizzati per la definizione dell'andamento del momento angolare della quantità di moto nel diffusore.

NC	ND	ALFA
0.2	0.75	0

Per quello che concerne l'andamento dell'angolo ξ lungo la linea media, che definisce l'inclinazione della traccia della pala rispetto al piano meridiano, è stato posto costante e pari a 20° , con verso tale da orientare la pala in modo da indurre una forza sulla corrente diretta verso l'asse della macchina. In tale modo si cerca di orientare la pala in modo da mantenere aderente la corrente alla superficie del mozzo che, per quanto osservato anche nel paragrafo 6.6.1, è la sezione critica.

Per quello che riguarda l'angolo di ingresso della pala del diffusore deve essere scelto in modo consentire la miglior interazione rotore-statore possibile nel funzionamento. L'angolo

della corrente assoluta in corrispondenza della sezione a valle della girante come in figura 6.9, è stato valutato pari a 40° ed è stato utilizzato come input per il dimensionamento della pala del diffusore. Le simulazioni numeriche hanno mostrato che tale valore dell'angolo costruttivo della pala garantisce l'incidenza migliore alla portata di progetto.

L'angolo di uscita della pala del diffusore, quindi l'overturning, deve garantire un completo raddrizzamento della corrente assoluta a valle del diffusore, in modo da consentire una corretta interazione tra uno stadio e il successivo. Come si è già notato, il diffusore con pale a spessore variabile consente richiede degli overturning maggiori a causa della deviazione maggiore della corrente rispetto alla corrente euleriana. E' stato necessario imporre un'overturning pari a 21.5° rispetto alla direzione assiale per ottenere un completo raddrizzamento della corrente nelle simulazioni CFD. L'angolo della corrente è stato misurato a valle del diffusore in corrispondenza della superficie di riferimento 4 di figura 6.35.

L'andamento della velocità assoluta nel condotto interpalare è assegnato con i valori dei parametri come definiti in 2.2.4 pari a $P_{CP} = 0.3$ e $\bar{V}_{CP} = 0$, quindi con la velocità che individua il punto di controllo è pari alla velocità assoluta della corrente in ingresso. Tale andamento assegnato è stato diagrammato nella prima immagine di figura 6.38. Nella seconda immagine della stessa figura, invece, è stato diagrammato l'angolo costruttivo della linea media della pala tra ingresso e uscita del diffusore derivante dalle assunzioni fatte per la progettazione.

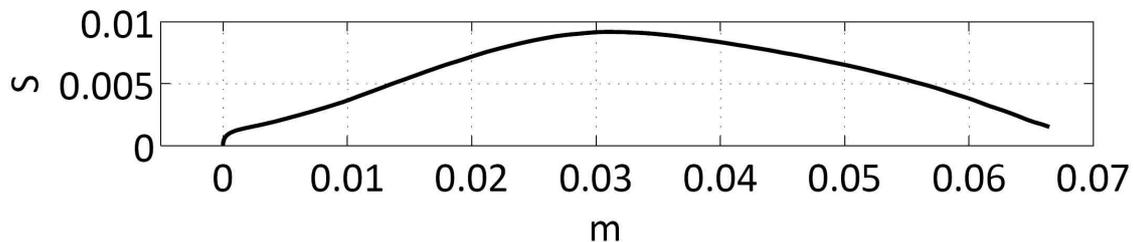
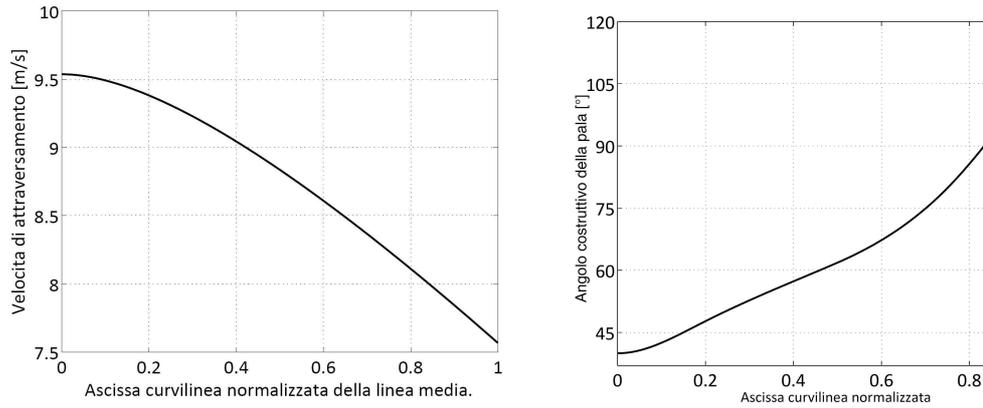


Figura 6.37: Distribuzione degli spessori normali alla linea media ricavati dall'assegnazione della velocità di attraversamento per il diffusore **SV**.

Questo andamento delle velocità di attraversamento impone la distribuzione degli spessori di figura 6.37, dove in ascissa troviamo il valore dell'ascissa curvilinea sulla linea media del piano conforme, e in ordinata il corrispondente semispessore della pala. Il valore massimo del semispessore che risultata è pari a 9.2 mm e si registra per un punto della pala collocato ad un'ascissa curvilinea pari al 46% della lunghezza totale della linea media.

Il canale interpalare del diffusore che ne risulta è riportata in figura 6.39 dove si vede la particolare forma della pala.

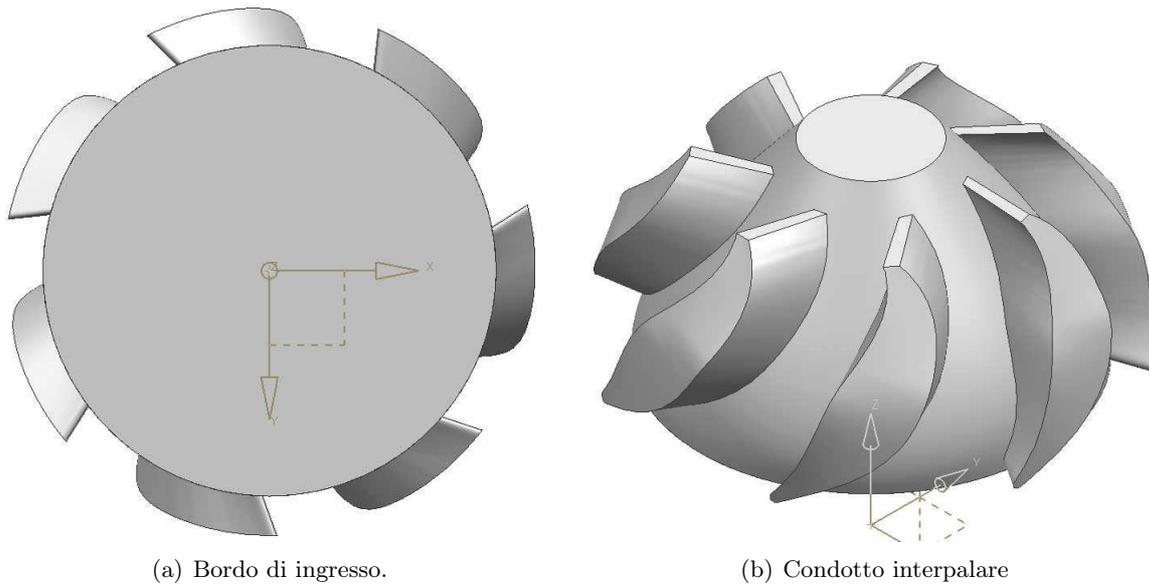
La simulazione dello stadio composto dalla girante **P** e dal diffusore **SV** alla portata di progetto di $60 \text{ m}^3/\text{h}$, ha messo in luce il buon comportamento del diffusore insieme con tale girante. Si visualizzano solo lievi ricircoli in corrispondenza della sezione del mozzo del diffusore. Il flusso della corrente in ingresso al diffusore è regolare; il punto di ristagno sulle pale del diffusore si colloca in corrispondenza del naso della pala come è testimoniato nella figura 6.40.



(a) Andamento assegnato della velocità nel condotto interpalare.

(b) Andamento dell'angolo costruttivo della pala.

Figura 6.38: Andamento assegnato della velocità di attraversamento nel condotto interpalare per il diffusore **SV** e corrispondente andamento degli angoli costruttivi della pala.



(a) Bordo di ingresso.

(b) Condotto interpalare

Figura 6.39: Vista frontale e laterale del condotto interpalare del diffusore con particolare del bordo di ingresso (a) e del condotto interpalare (b).

Nella figura 6.41 è visualizzato il campo di moto e il campo di pressione per le sezioni in corrispondenza al mozzo, alla corona e mediana. In queste si vede come il campo di moto della girante non risulta influenzato dall'interazione con il diffusore. Il campo di moto e lo scambio di energia sono regolari in tutto il canale meridiano. Si registra un lieve stallo della corrente in corrispondenza della sezione al mozzo nel diffusore che però si riassorbe e interessa solo le sezioni fino al 20 % in direzione *hub to shroud*.

Infine, come si può vedere sempre dalla figura 6.41 gli effetti di scia a valle del diffusore sono contenuti e l'entità dell'overturning imposta sembra essere sufficiente a raddrizzare completamente la corrente. I dati numerici ricavati dalla progettazione dei risultati delle sim-

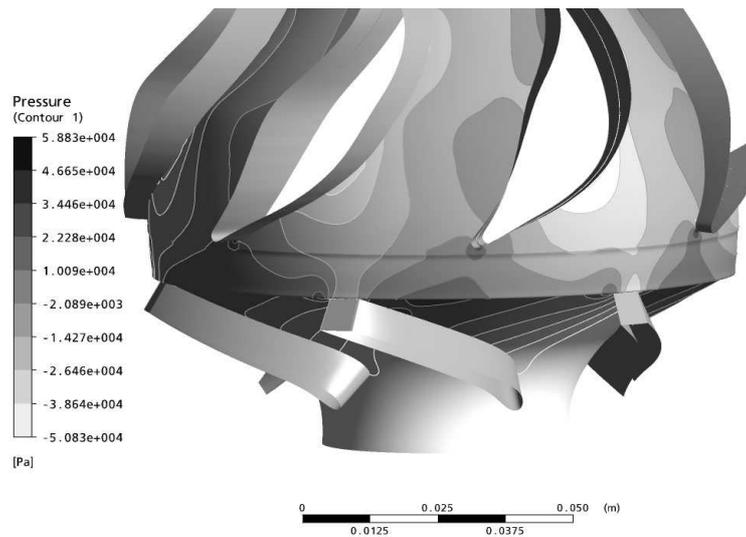


Figura 6.40: Punti di ristagno in corrispondenza del naso della palettatura del diffusore.

ulazioni sono riportati in tabella 6.18. Si può notare come siano sostanzialmente rispettate le previsioni fatte in sede preliminare alla progettazione per lo scambio di energia e nella verifica ai codici di calcolo CFD della sola girante: una girante funzionante correttamente elabora energia per circa 11.5 metri di colonna d’acqua. Di questi circa 1.5 sono persi nell’interazione con il diffusore e nel diffusore stesso. Il rendimento complessivo dello stadio, alla portata di progetto è quindi pari all’83 %, qualche punto superiore alle richieste.

Tabella 6.18: Risultati numerici ricavati per la simulazione della girante **P** accoppiata con il diffusore **SV** alla portata di progetto di 60 m³/h.

Q [m ³ /h]	H ₀₁ [m]	H ₁₂ [m]	H ₂₄ [m]	H _{tot} [m]	Ht [m]	η_G	η_{id}	α_{OUT} [deg]
60	11.63	-0.59	-0.91	10.13	12.20	0.954	0.830	89:90

In conclusione si è proposta all’azienda finanziatrice come progetto esecutivo i modelli della girante **P** e del diffusore **SV** per i dati di progetto e i target prestazionali che erano stati presentati.

In tabella 6.18 la potenza è stata misurata grazie alla coppia lungo la direzione dell’asse di rotazione per le superfici in movimento. Questa potenza è stata scalata con la portata di massa e la costante gravitazionale per ricavare la grandezza indicata come Ht. Che quindi diventa $Ht = \frac{P}{Wg}$ e rappresenta l’energia specifica trasmessa al fluido in metri di colonna d’acqua.

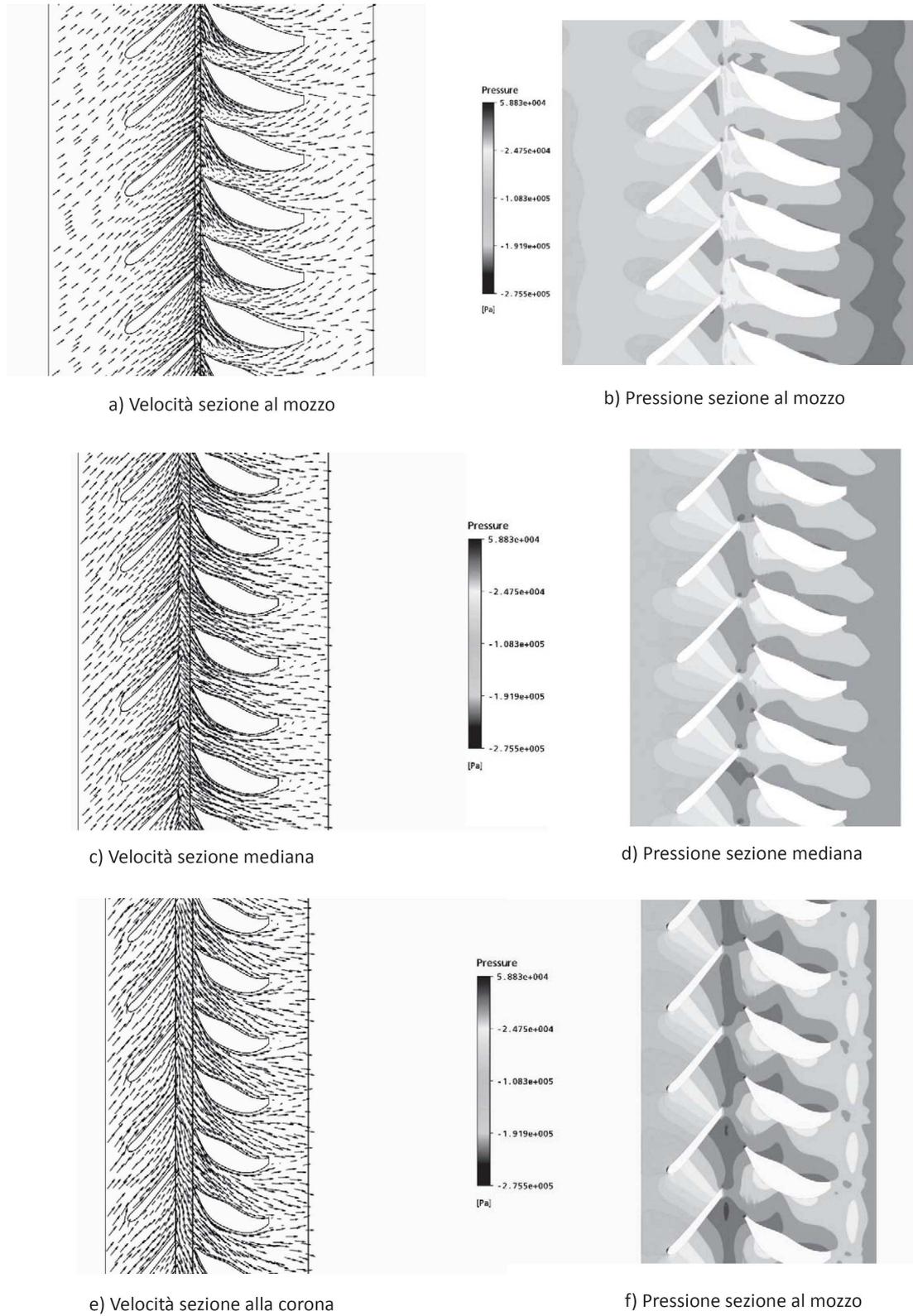


Figura 6.41: Campo di moto e campo di pressione nell'interazione tra girante e diffusore alla portata di progetto per varie sezioni blade to blade.

Capitolo 7

Simulazioni numeriche complete e collaudo del progetto proposto

Il modello proposto all'aziende finanziatrice, composto dalla girante **P** e dal diffusore **SV** sono stati prototipati ed è stata effettuata una prova di collaudo su una pompa sommersa da 6 pollici a due stadi. In questo capitolo si presentano i risultati di un'analisi numerica completa sul comportamento della stadio progettato nella configurazione utilizzata per la progettazione nel capitolo precedente, al fine di investigare il comportamento di fuori progetto. Oltre a questo si presentano i dati derivanti da un'analisi numerica completa della macchina nella configurazione delle pompa che è stata collaudata. La configurazione che è stata simulata quindi è composta di corpo di aspirazione, due stadi e corpo della valvola di scarico, in tutti i particolari che sono interessati dal fluido. Vengono presentati infine i risultati delle prove di collaudo effettuati con la macchina prototipata.

Nella figura 7.1 si può vedere il modello 3D del prototipo di pompa che utilizza l'idraulica progettata nel capitolo precedente.

Anticipiamo che i risultati del collaudo si discostano sensibilmente da quelli attesi dalle simulazioni sia per la configurazione utilizzata in sede di progettazione che per la configurazione utilizzata in sede di collaudo che di fatto tiene conto di tutti i particolari che interessano l'idraulica della pompa.

7.1 Funzionamento in sede di fuori progetto

L'analisi numerica utilizzata in sede di progettazione e presentata nel paragrafo 6.6.2 prevede la simulazione del modello di stadio con girante e canale interstadio accoppiati alla sola portata di progetto, cioè quella per la quale erano fissati gli obiettivi in termini di prestazioni della macchina. In sede di progettazione non è stato possibile verificare il funzionamento della macchina per le portate di fuori progetto e quindi non è stato possibile tracciare la curva di funzionamento della macchina e investigare quindi in modo completo l'interazione della girante **P** e le varie soluzioni di diffusore proposte.

Come criterio discriminante tra le varie soluzioni di diffusore vagliate si è utilizzato il

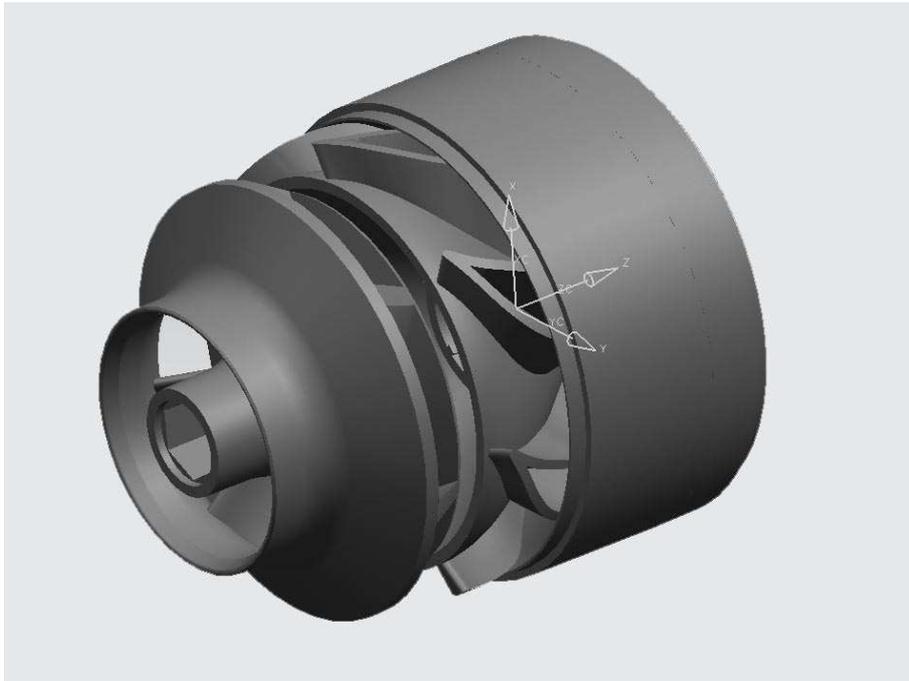


Figura 7.1: Modello 3D del prototipo.

funzionamento della macchina alla sola portata di progetto della macchina; questo perchè gli obiettivi in termini di prestazioni sono stati fissati per la sola portata di $60 \text{ m}^3/\text{h}$ e perchè indagare, per ciascuna soluzione, anche le portate di fuori progetto rappresentava un lusso non ammissibile.

Per la soluzione definitiva rappresentata dalla girante **P** e dal diffusore **SV** è stata effettuata un'indagine sperimentale più approfondita. Si riporta di seguito le curve di funzionamento della soluzione di progetto ricavata dalle simulazioni CFD effettuate per la configurazione utilizzata in sede di progetto alle portate di 40, 50, 60 e $75 \text{ m}^3/\text{h}$.

Come si può notare dalla figura 7.2 l'accostamento del diffusore alla girante sposta il punto di massima efficienza a portate inferiori a quella di progetto, quando quella della girante nel funzionamento isolato era perfettamente centrata sulla portata di progetto; per la portata di $50 \text{ m}^3/\text{h}$ si registra un rendimento idraulico pari a 84.7%. La curva della prevalenza non manifesta flessi, e quindi l'interazione funzionale tra girante e diffusore non produce instabilità funzionali.

La potenza è stata misurata grazie alla coppia lungo la direzione dell'asse di rotazione per le superfici in movimento. Questa potenza è stata scalata con la portata di massa e la costante gravitazionale per ricavare la grandezza indicata come Ht . Che quindi diventa $Ht = \frac{P}{Wg}$ e rappresenta l'energia specifica trasmessa al fluido in metri di colonna d'acqua

Come già puntualizzato nel capitolo precedente le prestazioni attese alla portata di progetto indicano che il diffusore **SV** soddisfa gli obiettivi che erano stati imposti. Tuttavia l'interazione sposta di una quantità non trascurabile il punto di massima efficienza verso le portate minori. Anche se il rendimento cresce ulteriormente, tale variazione non è trascurabile. Infatti le normative classificano le pompe in base alla portata di massimo rendimento

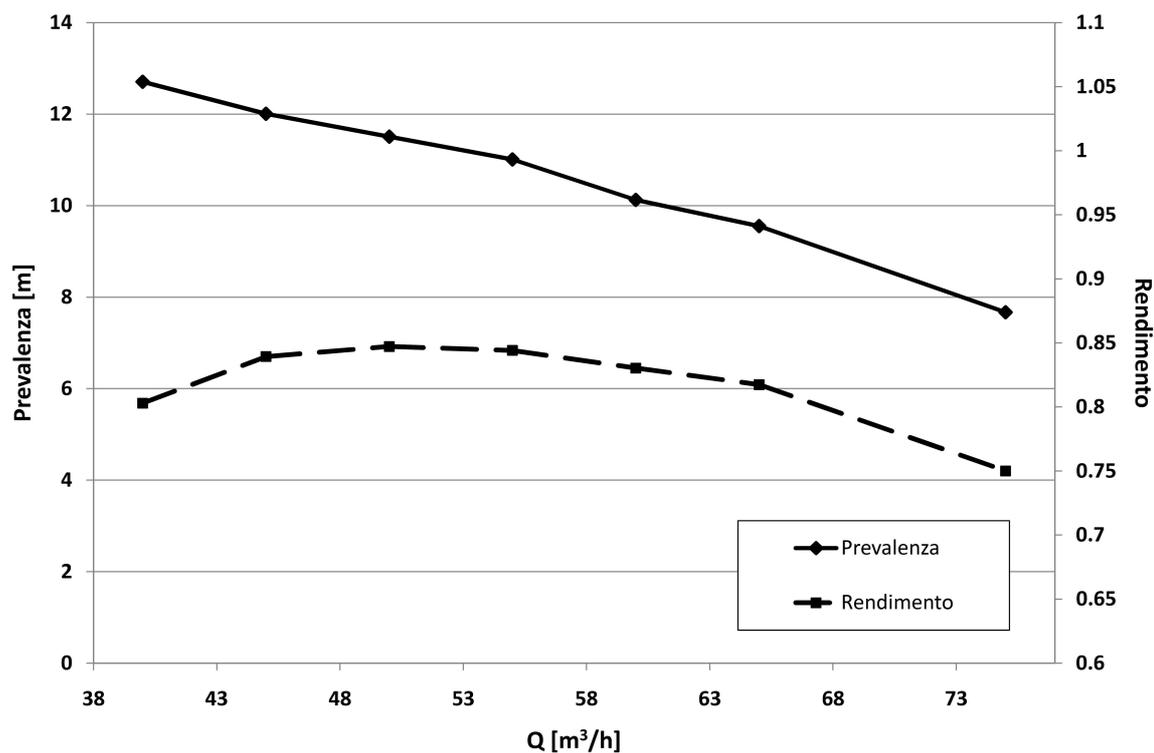


Figura 7.2: Curve di funzionamento per il prototipo ricavate da simulazione numerica.

e su queste viene valutata l'efficienza della pompa. Il valore di rendimento dell'80% utilizzato come obiettivo in sedi di progettazione è stato ricavato per una pompa con portata di massimo rendimento a 60 m³/h. Quindi tale macchina non risponde in modo corretto agli obiettivi di progetto. E' quindi necessario rivedere il progetto del diffusore-convogliatore per ottenere un risultato in linea con le attese.

Tabella 7.1: Risultati delle simulazioni numeriche per la configurazione di progetto.

Q [m ³ /h]	40	45	50	55	60	65	75
H ₀₁ [m]	14.52	13.54	12.93	12.44	11.63	11.13	9.66
H ₁₂ [m]	-0.66	-0.60	-0.57	-0.57	-0.59	-0.61	-0.74
H ₂₄ [m]	-1.16	-0.93	-0.86	-0.86	-0.91	-0.96	-1.25
H _{tot} [m]	12.71	12.01	11.50	11.01	10.13	9.55	7.67
H _t [m]	15.82	14.31	13.58	13.04	12.20	11.69	10.23
η _G	0.912	0.946	0.952	0.954	0.954	0.952	0.945
η _{tot}	0.803	0.839	0.847	0.844	0.830	0.817	0.750
α _{OUT} [deg]	89.5:96	90:92	90:91	89:90	89:90	90:91	92.5:93

7.2 Prototipo di pompa sommersa e impianto di collaudo.

Il modello consegnato all'azienda finanziatrice rappresenta solo la parte idraulica principale della macchina. Nel prototipo effettivo sono presenti tutte le zone ausiliarie bagnate dal fluido. La parte idraulica del modello che è stata effettivamente realizzata differisce in alcuni particolari dal modello che è stato discretizzato e verificato ai codici di calcolo nella fase di progettazione i cui risultati estesi sono presentati nel paragrafo 7.1.

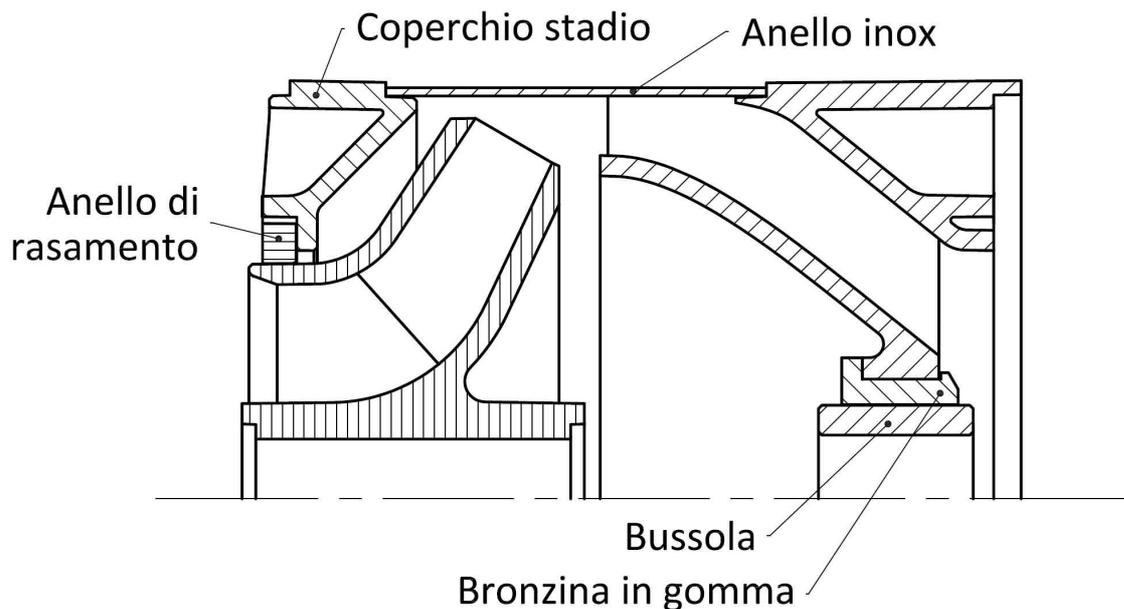


Figura 7.3: Sezione 2D del prototipo realizzato con tutti i particolari costruttivi.

Il prototipo è stato realizzato con una configurazione già ingegnerizzata, nel senso che la geometria del pezzo prototipato prevede tutti i particolari indispensabili per la produzione in serie tramite stampaggio. La configurazione realizzata quindi prevede gli spessori della giusta dimensione e gli angoli di sforno delle superfici corrette. Inoltre il prototipo è un assemblato composto di più parti. Nel disegno di figura 7.3 è rappresentato il modello prototipato dello stadio nella sua interezza. Oltre alla girante e il diffusore sono stati riportati i componenti più importanti che vanno a costituire la superficie bagnata dal fluido. In particolare l'anello inox che costituisce la parete che contiene il fluido in uscita dalla girante, il coperchio stadio che chiude la zona posteriore del diffusore dello stadio precedente e realizza la sede dove è alloggiato l'anello di rasamento radiale che realizza la tenuta tra l'anello esterno della girante e il diffusore. In questo caso il rasamento è realizzato per contenere al minimo le perdite per trafilamento; il gioco che si realizza tra l'anello e la girante è di 0.1 mm sul raggio. Tra il diffusore e l'albero a sede esagonale (non riportato in figura) il sostegno viene realizzato tramite una bussola in acciaio inox e una bronzina in gomma.

Le zone interessate dal flusso principale della corrente dove sono presenti le maggiori differenze tra il modello utilizzato in sede di progettazione e quello realizzato come prototipo sono quelle in corrispondenza del traferro tra girante e diffusore e dello scarico del diffusore.

In corrispondenza del traferro non sono presenti pareti come in figura 6.35, bensì il condotto meridiano si apre verso le zone di ristagno oltre il disco della corona e quello del mozzo. Allo scarico del diffusore la parete fisica che costituisce il mozzo a valle della schiera del diffusore è rappresentata dalla superficie della bronzina e della boccola. Anche la zona di condotto che definisce la corona a valle del diffusore è differente da quella di figura 6.35 utilizzata per le simulazioni del capitolo 6 e nel paragrafo 7.1. Nella parete è ricavato uno scanso per consentire l'alloggiamento dell'anello di rasamento dello stadio successivo. Di fatto nel modello che è stato prototipato le pareti del canale meridiano del diffusore subito a valle del bordo di uscita della schiera di pale del diffusore sono sensibilmente differenti da quelle del modello discretizzato in sede di progettazione. In quelle zone la corrente tende a staccare e si manifestano dei ricircoli.

Il prototipo del modello di figura 7.1 è stato realizzato in materiale FullCore, tramite stampaggio a getto. La tecnica di prototipazione utilizzata realizza il pezzo per strati sovrapposti di tale materiale sagomati con la forma delle sezioni del pezzo. Ogni strato ha una profondità di 16 μm ; la rugosità delle superfici del prototipo è dettata dalla profondità di tali strati, e quindi assolutamente migliore rispetto a prototipi realizzati per sinterizzazione o stereolitografia. La tecnica di prototipazione rapida per stampaggio a getto di polimeri, inoltre non ha limiti di forma sui pezzi che è possibile realizzare, infatti l'impiego di un particolare gel di supporto consente di realizzare superfici in sottosquadro rispetto alla direzione del getto. Tale gel viene poi rimosso in un secondo momento grazie con un semplice lavaggio con una soluzione con soda acqua e soda caustica.

In figura 7.5 sono riportate due immagini del prototipo in questione.

Il collaudo è realizzato con una pompa sommersa a due stadi. La configurazione da collaudo quindi è quella di una pompa sommersa commerciale da 6. Oltre agli stadi sono presenti altre parti che completano la pompa. La figura 7.4 mostra l'assemblaggio della pompa completa. Da basso la macchina è composta delle seguenti parti:

- **motore sommerso:** il motore sommerso di azionamento è a due poli in bagno d'olio; in sede di collaudo della parte idraulica della macchina si fa uso di motori tarati;
- **corpo di aspirazione:** sopra il motore è presente il corpo di aspirazione della macchina, questo è realizzato in ghisa di fusione e presenta quattro feritoie attraverso le quali viene aspirato il fluido dalla vasca; in sede di utilizzo, le feritoie sono ricoperte da una grata che impedisce alle impurità più voluminose di attraversare la girante; il corpo di aspirazione alloggia al proprio interno il giunto tra l'estremità dell'albero al motore l'estremità dell'albero alla girante;
- **corpi stadio:** dopo il corpo di aspirazione sono presenti i corpi dei due stadi in serie, calettati sull'albero;
- **corpo della valvola:** realizzato in ottone serve ad alloggiare la valvola di non ritorno posizionata allo scarico dell'ultimo stadio e la filettatura per l'innesto della tubatura di mandata;

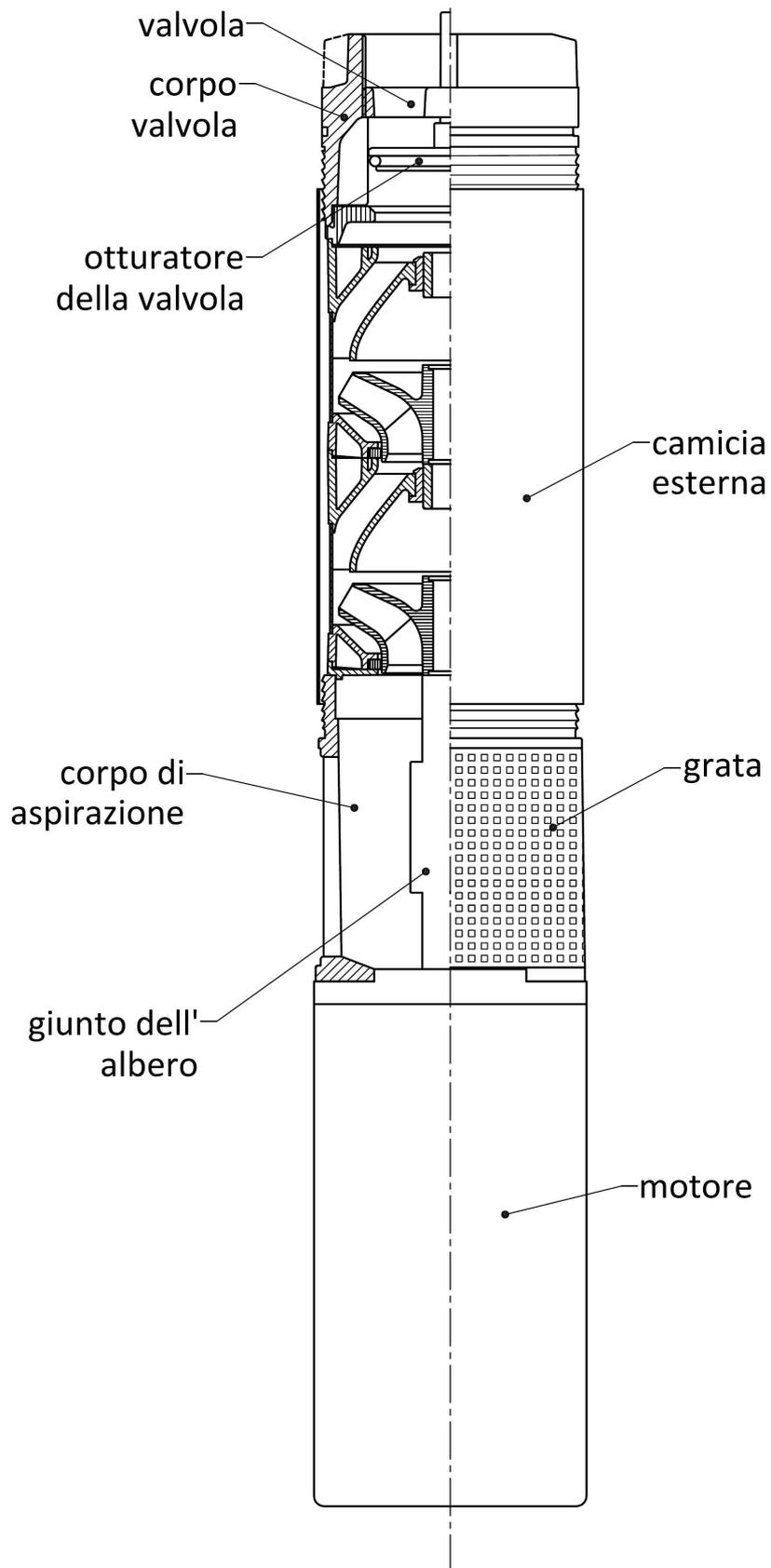
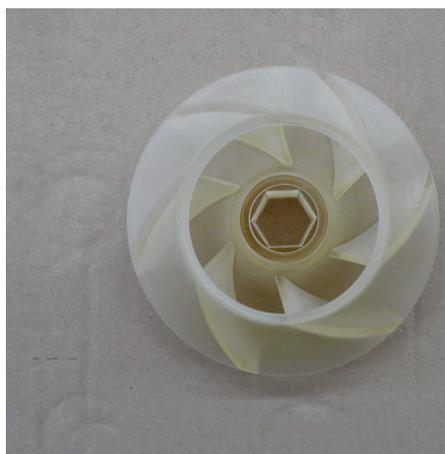


Figura 7.4: Disegno della pompa sommersa completa di tutte le parti che la compongono.



(a) Girante



(b) Particolare del bordo di ingresso del diffusore e del condotto interpale.

Figura 7.5: Immagini dei prototipi realizzati.

- **camicia esterna:** la camicia in alluminio copre e protegge i corpi stadio della macchina ed è agganciata in basso al corpo di aspirazione e in alto al corpo della valvola di mandata.

La normativa europea UNI EN ISO 9906 (vedi [58]) prevede il collaudo delle pompe comprendendo tutti gli accessori che la compongono e ne costituiscono parte integrante. I collaudi che poi vengono riportata nei catalogo dei produttori misurano le prestazioni di pompe in funzionamento sia della valvola che della grata. In questa tesi chiameremo *configurazione di collaudo* la configurazione della macchina senza gli accessori della grata e della valvola nel corpo di scarico che introducono le perdite di carico; il collaudo viene effettuato nella configurazione senza questi accessori che introducono forti perdite di carico per avere un riscontro confrontabile con le simulazioni numeriche dove le geometrie di questi componenti sono state trascurate in quanto di difficile modellizzazione e per aver un'indicazione sulla prevalenza del funzionamento della parte idraulica principale della macchina con meno perdite parassite possibile.

7.3. DISCRETIZZAZIONE DELLA MACCHINA IN CONFIGURAZIONE DI COLLAUDO E IMPOSTAZIONI DELLE SIMULAZIONI

L'impianto di prova utilizzato risponde alle caratteristiche indicate dalla normativa UNI EN ISO 9906 nella sezione dedicata alle prove sulle pompe verticali sommerse. In figura 7.6 è riportato lo schema dell'impianto di prova utilizzato per effettuare il collaudo secondo la normativa citata. Il corpo completo della pompa viene immerso nella vasca. La pompa scarica in un impianto a valle dove viene effettuata la misurazione della pressione statica di mandata e della portata. La misura della prevalenza quindi risulta pari a

$$H = z'_2 - z'_1 \frac{p_{M2}}{\rho g} + \frac{\rho_{f1}}{\rho} (z_{M2} - z_{2'}) + \frac{U_2^2}{2g}.$$

I vari simboli fanno riferimento allo schema di figura 7.6 ; ρ_{f1} è la densità del fluido nel tubo di raccordo tra la tubazione e il trasduttore di pressione. e il termine $\frac{\rho_{f1}}{\rho} (z_{M2} - z_{2'})$ rappresenta la correzione del valore della lettura della pressione p_{M2} secondo la differenza di quota.

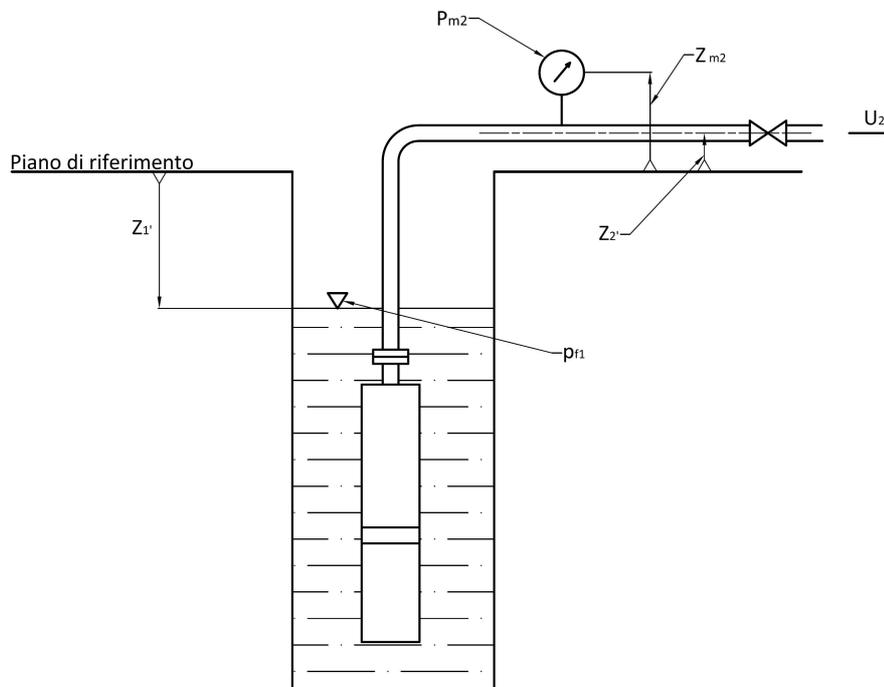


Figura 7.6: Schema dell'impianto di prova per pompe verticali sommerse secondo la normativa UNI 9906.

7.3 Discretizzazione della macchina in configurazione di collaudo e impostazioni delle simulazioni

La configurazione della macchina utilizzata in sede di collaudo è stata discretizzata in tutti i suoi particolari per effettuare un simulazione completa che fosse comparabile con le prove di collaudo effettuate. Il volume che è stato discretizzato è rappresentato in forma schematica nella suo profilo meridiano nella figura 7.7 .

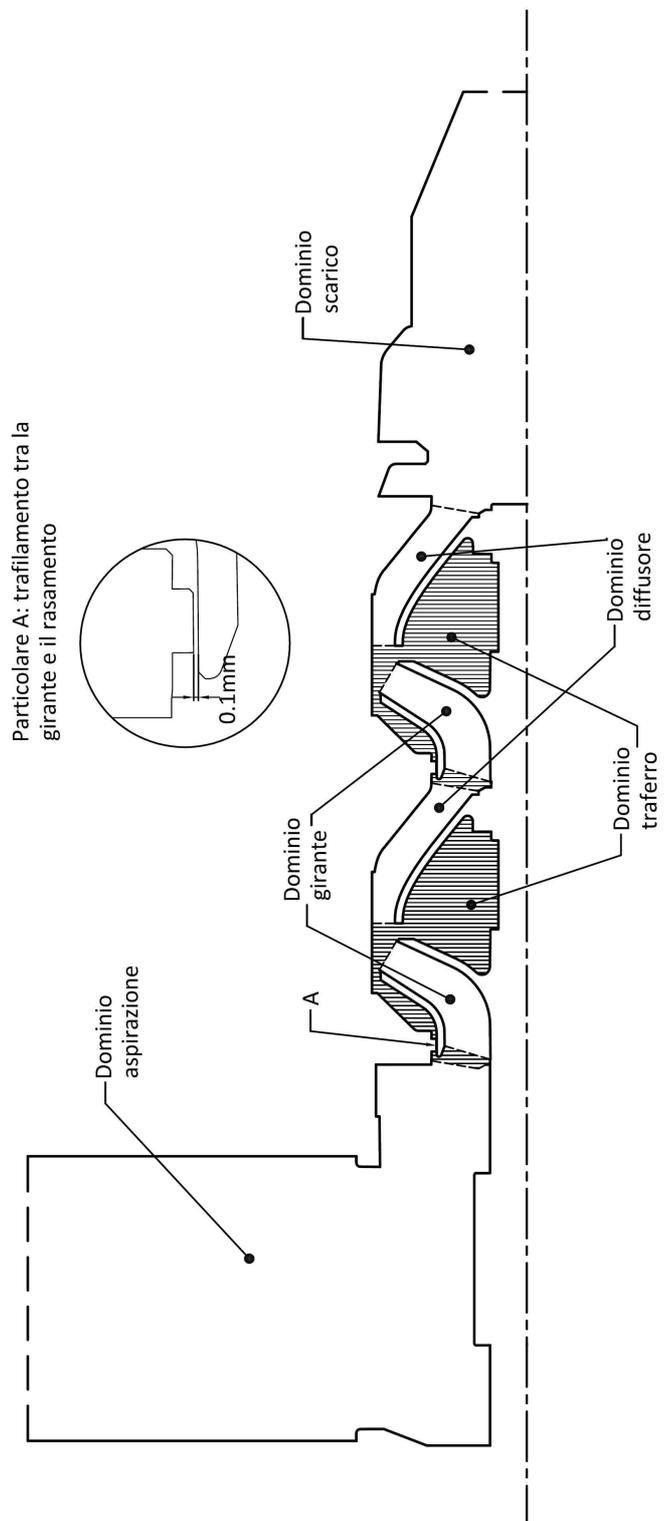


Figura 7.7: Schema dei domini in cui è stato suddiviso il modello per poter essere discretizzato.

Per contenere il numero di nodi necessario per effettuare la simulazione, già considerevole, si è scelto di non schematizzare tutto il volume della vasca in cui è immersa la pompa. È stato discretizzata una porzione anulare del volume della vasca in cui è immersa la pompa, intorno alle feritoie dell'aspirazione, ritenendo questa sufficiente per ricavare nel modo più completo le condizioni di moto in ingresso alla girante; è stata imposta una condizione di parete liscia per le due superfici dell'anello ortogonali all'asse della macchina, mentre in corrispondenza della superficie cilindrica laterale è stata imposta una condizione al contorno di *inflow* (vedi [54]) dove viene imposta la portata di massa e la direzione del fluido ortogonale alla superficie. Anche il corpo della valvola di scarico è stato discretizzato, comprese le razze che fungono da guida per il movimento dell'otturatore. Il modello che è stato discretizzato non prevede la grata all'aspirazione e la valvola nel corpo di scarico. Allo scarico del corpo di valvola è stato necessario aggiungere un tratto di condotto fittizio a pareti coniche per facilitare la convergenza della simulazione. Alla superficie della discretizzazione in corrispondenza dello sbocco nel tubo di mandata è stata imposta una condizione al contorno di *outflow*, stabilendo il livello di pressione statica media. Nella figura 7.7 sono rappresentati i domini in cui è stata suddivisa la parte idraulica bagnata dal fluido. Si è scelto di imporre questa coppia di condizioni al contorno per la simulazione in quanto è la combinazione più stabile numericamente tra quelle possibili per una simulazione CFD.

È stato necessario avvalersi di software differenti per discretizzare i diversi domini e l'operazione di discretizzazione è stata particolarmente complessa. I domini che interessano lo stadio sono stati discretizzati con mesh esaedriche strutturate, se pur con differenti infittimenti. I domini di aspirazione e scarico sono stati discretizzati con mesh non strutturate tetraedriche meno fitte di quelle che interessano i condotti interni. In particolare per la meshatura è stato utilizzato il programma del pacchetto Ansys ICEM; la mesh è stata realizzata con gli algoritmi automatici ai quali sono stati aggiunti dei layer di nodi attorno alle pareti a sviluppo perpendicolare alle pareti.

Le mesh della discretizzazione delle palettature di girante e diffusore sono state realizzate con Turbogrid e sono stati utilizzati infittimenti comparabili con quelli utilizzati per le simulazioni utilizzate in sede di progettazione del diffusore per discretizzare lo stadio completo. Sono stati utilizzati due domini ausiliari (denominati in figura 7.7 come 'Dominio traferro') che comprendono il volume del traferro tra la girante e il diffusore, il volume di acqua situato tra il disco del mozzo della girante e l'albero della macchina, il volume tra il disco della corona della girante e il fondo del diffusore dello stadio precedente, il canale di trafilamento tra la girante e l'anello di rasamento e il traferro tra il bordo di uscita del diffusore e il bordo di ingresso della girante. Questi sono stati discretizzati con mesh strutturate esaedriche grazie al programma ICEM e in figura 7.7 sono rappresentati evidenziati; in figura inoltre è evidenziato il particolare della canale di trafilamento in corrispondenza dell'anello di rasamento. Di fatto il dominio della girante si interfaccia solo con il dominio del traferro.

Il numero di nodi utilizzato si ripartisce come è stato riportato in tabella 7.2. Per il dettaglio del trafilamento in figura 7.8 è riportato il numero di nodi utilizzato per discretizzare il canale di trafilamento tra il disco della girante e l'anello di rasamento. Come si vede sono stati

utilizzati 7 nodi equispaziati in direzione radiale per discretizzare un canale di 0.1 mm.

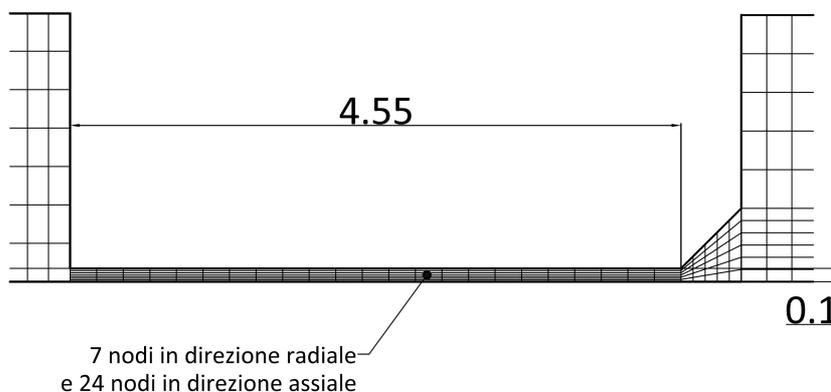


Figura 7.8: Discretizzazione del trafilemento tra girante e anello di rasamento.

Tabella 7.2: Numero di nodi utilizzati per la discretizzazione di ciascun dominio in cui è stato suddiviso il modello della pompa utilizzata in sede di collaudo

Numero nodi per dominio				
aspirazione	traferro	girante	diffusore	corpo valvola
452497	952920	507414	639597	151781

Tenendo presente che i domini della girante, del diffusore e del traferro sono ripetuti due volte, il totale dei nodi utilizzati è pari a 4804140.

Le simulazioni alle varie portate sono state eseguite con modello di turbolenza SST con interfaccia Transient tra domini con sistemi di riferimento in movimento reciproco. Il valore del time step utilizzato è tale da garantire un rotazione di 1° ad ogni timestep. Per ogni timestep si è fatto uso di 7 iterazioni di convergenza. La dimensione della mesh non consente di effettuare degli studi di sensibilità sulla mesh, tuttavia i risultati hanno mostrato dei valori di y^+ in buona parte inferiori al valore di 200, e comunque mai superiori al valore di 300 per i domini della girante e del diffusore; il numero di courant medio è inferiore a 1.5 e quello massimo si mantiene inferiore a 30. Tali valori sono stati ritenuti accettabili per la simulazioni e i risultati delle simulazioni sono stati ritenuti validi.

Con queste impostazioni una simulazione impiega, per avere risultati stabili, circa 2 settimane di calcolo su una workstation utilizzando 4 processori e 10.5 Gb di memoria Ram.

7.4 Risultati numerici

A causa della quantità di risorse richieste per una simulazione, è stato possibile effettuare simulazioni solo ad alcune portate per verificare il funzionamento e le prestazioni della macchina a

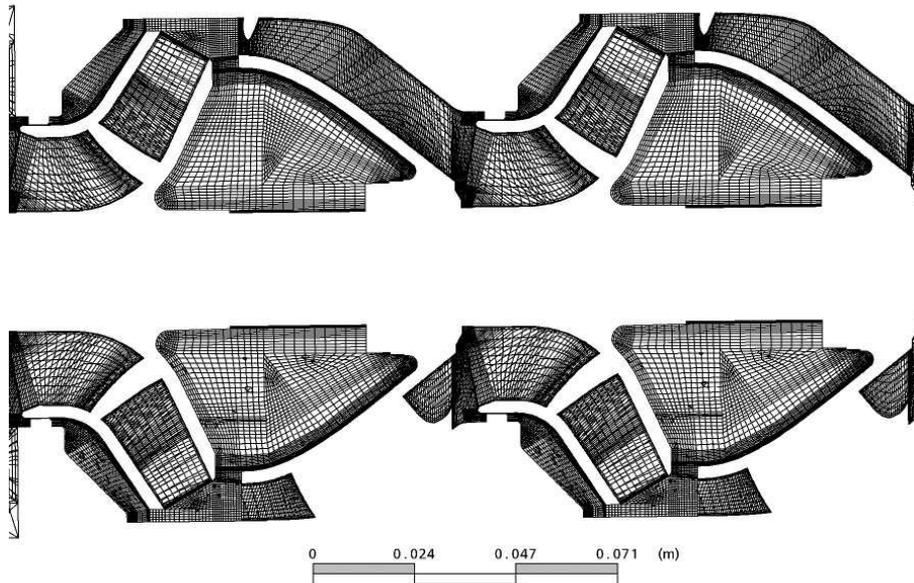


Figura 7.9: Discretizzazione del modello dello stadio.

varie portate. I risultati hanno comunque messo in luce aspetti interessanti che devono essere tenuti in considerazione.

I risultati delle simulazioni hanno reso possibile innanzi tutto valutare come influiscono sulle prestazioni le modifiche introdotte in sede di ingegnerizzazione.

I risultati numerici riportati in questa tesi sono stati ottenuti come media dei risultati ottenuti per numerosi risultati intermedi delle simulazioni. Tali risultati sono stati ottenuti, dopo che la simulazione ha raggiunto risultati stabili dello scambio di energia, per una serie di timestep equidistanti in modo da coprire un angolo di rotazione sufficientemente ampio. In tabella 7.3 sono riportati i dati numerici delle simulazioni effettuate per le portate di 43, 54, 60 m^3/h . Tali portate sono state scelte per coprire uniformemente il range di portate attorno a quella di massimo rendimento rilevata nelle simulazioni in configurazione di progetto. Oltre ai dati che riassumono il funzionamento della macchina nella sua interezza sono riportati anche numerosi dati che consentono un'indagine approfondita dei risultati delle simulazioni.

In tabella 7.3 **1S** indica zone o grandezze che si riferiscono al primo stadio, mentre **2S** che si riferiscono al secondo stadio. Il primo gruppo di risultati, indicato con **Scambio di energia**, indica la differenza di energia, misurata in metri di colonna d'acqua, tra l'inizio e l'uscita di ciascuna zona considerata. Le superfici di controllo che sono state utilizzate per rilevare tali valori corrispondono alle interfacce dei domini, come indicati in figura 7.7. Per cui per aspirazione si intende la differenza di energia totale del fluido tra le superficie di ingresso e di uscita del dominio dell'aspirazione. Va precisato che per *adduzione* si intende la zona di condotto che sta a monte delle giranti, quindi lo scambio di energia indicato in corrispondenza di Adduzione 1S indica la differenza di energia del fluido tra la superficie in corrispondenza del termine del dominio dell'aspirazione e la superficie in corrispondenza dell'imbocco del dominio

della girante, e per Adduzione 2S si intende la zona delimitata a monte dal dominio del primo canale interstadio e a valle dalla girante del secondo stadio. Per *traferro* si intende la zona di fluido tra la superficie di sbocco del dominio della girante e l'ingresso del diffusore; quindi, quando in tabella 7.3 si indica lo scambio riferito al traferro si intende la differenza di energia totale del fluido tra superficie di uscita dal dominio della girante e superficie di ingresso del dominio del diffusore, quindi di fatto si identificano le perdite per mescolamento a valle della girante.

Lo scambio di energia complessivo di ciascuno stadio è stato conteggiato tra la superficie di ingresso della zona di adduzione e lo la superficie di uscita dal dominio del diffusore. Lo scambio complessivo è stato invece conteggiato tra la superficie di ingresso nel dominio di aspirazione e una superficie piana nel dominio dello scarico subito a monte del tratto a pareti coniche.

Anche qui come già fatto per i risultati delle simulazioni in configurazione di progetto, la potenza è stata misurata grazie alla coppia lungo la direzione dell'asse di rotazione per le superfici in movimento. Questa potenza è stata scalata con la portata di massa e la costante gravitazionale per ricavare la grandezza indicata come Ht . Che quindi diventa $Ht = \frac{P}{Wg}$ e rappresenta l'energia specifica trasmessa al fluido in metri di colonna d'acqua. Tale grandezza quindi può essere confrontata direttamente con lo scambio di energia (\mathbf{H}) che rappresenta l'energia trasmessa al fluido al netto delle perdite. Inoltre è stata specificata sia la potenza complessiva trasmessa dalle pareti di ciascuno stadio che la potenza trasmessa per ventilazione, che rappresenta quella trasmessa al fluido dalle pareti esterne al condotto interpalare, cioè quelle che definiscono le zone di fluido tra il disco del mozzo e l'albero e tra il disco della corona e la superficie posteriore dello stadio precedente.

La potenza riferita all'albero è quella calcolata a partire dal momento torcente registrato alle pareti dell'albero della macchina nel corpo di aspirazione.

Il calcolo dei rendimenti complessivi del singolo stadio è stato realizzato il rapporto tra l'energia specifica al netto delle perdite tra la superficie di ingresso e di uscita di ciascuno stadio e la corrispondente energia trasmessa al fluido Ht . Per quello che riguarda il rendimento interno della girante è stato calcolato tenendo conto della portata che effettivamente attraversa la girante che per effetto dei trafiletti è superiore a quella che viene elaborata dalla macchina.

Quindi il calcolo dei rendimenti idraulici è stato effettuato secondo la formulazione

$$\eta = \frac{H}{Ht}$$

mentre per i rendimenti interni delle due giranti si è utilizzata la seguente formulazione:

$$\eta_{Gir} = \frac{WgH_{Girante}}{\eta_{Vol}P_{Gir}}$$

dove W è la portata di massa in ingresso alla macchina.

Il calcolo degli angoli di flusso è stato effettuato a partire dalle componenti meridiana e circonferenziale della velocità in ingresso del diffusore (*Traferro*), in corrispondenza dell'uscita

del diffusore (*Scarico Diff*) e in corrispondenza della superficie piana al termine del corpo della valvola di scarico, cioè subito a monte dell'inizio del tratto conico ausiliario.

7.4.1 Osservazioni sui risultati numerici

La girante del primo stadio mostra un buon comportamento, testimoniato dal rendimento interno che è solo di poco inferiore a quello realizzato nella simulazione nella configurazione di progetto. Anche il diffusore convogliatore del primo stadio mostra un buon comportamento in termini di perdite di energia e nel complesso il rendimento e lo scambio di energia del primo stadio risultano in linea con i risultati delle simulazioni presentate in 7.1.

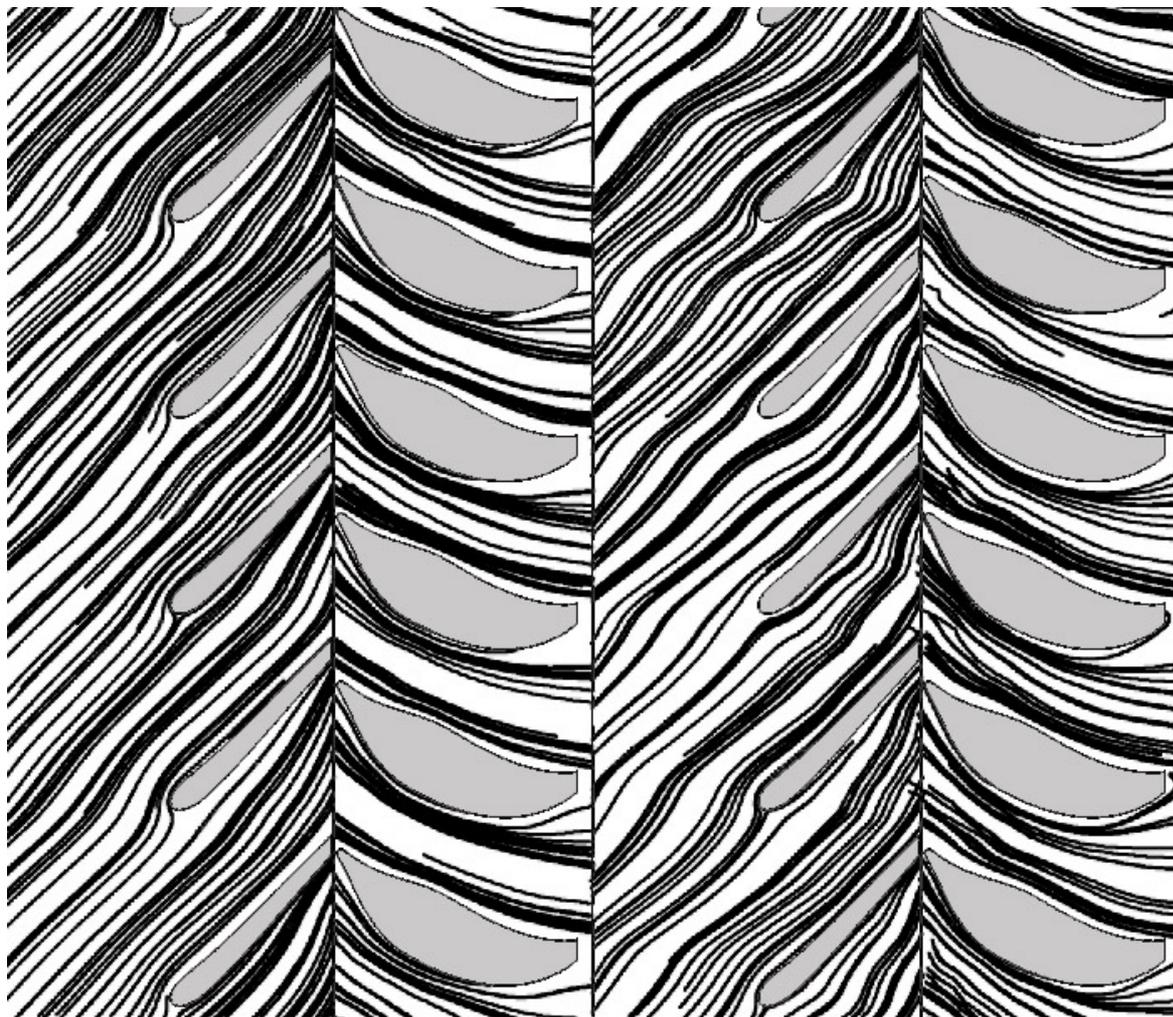


Figura 7.10: Campo di moto sulla superficie interpale.

I risultati numerici del secondo stadio risultano sensibilmente differenti da quelli del primo stadio e da quanto atteso in sede di progettazione. In particolare sia lo scambio di energia che il rendimento risultano inferiori rispetto al primo stadio. Questo è da attribuirsi ad una non corretta interazione tra i due stadi che compongono la pompa. In figura 7.10 si possono osservare le linee di corrente tracciate in corrispondenza della sezione media per i condotti

Tabella 7.3: Dettaglio dei risultati ottenuti per le simulazioni numeriche nella configurazione di collaudo.

Portata [m³/h]	43	54	60
	Scambio di energia (H) [m]		
Aspirazione	-0.08	-0.13	-0.15
Adduzione 1S	0.06	0.09	0.11
Girante 1S	13.31	11.98	11.23
Traferro 1S	-0.66	-0.52	-0.46
Diffusore 1S	-1.10	-0.64	-0.73
Adduzione 2S	-0.22	-0.30	-0.33
Girante 2S	12.88	11.16	9.83
Traferro 2S	-0.81	-0.56	-0.53
Diffusore 2S	-1.07	-0.66	-0.74
Valvola	-1.20	-1.28	-1.10
Scambio di energia 1S	11.50	10.73	9.91
Scambio di energia 2S	10.78	9.63	8.23
Scambio di energia Totale	21.00	18.95	16.89
	Potenza [W]		
Potenza 1S	1780	1984	2068
Potenza 1S Interna	1711	1915	2007
Ventilazione 1S	69	69	61
Potenza 2S	1805	1996	1973
Potenza 2S Interna	1735	1927	1906
Ventilazione 2S	70	70	68
Potenza Albero	5	5	4
	Ht [m]		
Primo Stadio	15.19	13.48	12.65
Primo Stadio Interna	14.60	13.01	12.28
Ventilazione 1S	0.59	0.47	0.37
Secondo Stadio	15.40	13.57	12.07
Secondo Stadio Interna	14.80	13.09	11.66
Ventilazione 2S	0.60	0.47	0.41
Albero	0.04	0.03	0.03
	Angoli Flusso [gradi]		
Traferro 1	33.8	41.2	45.3
Traferro 2	34.0	41.6	45.6
Scarico Diff 1	91.8	89.3	84.7
Scarico Diff 2	87.7	87.3	86.2
Valvola	90.0	95.0	89.4
	Rendimento [%]		
Girante 1S	92.7	93.3	92.5
Girante 2S	88.5	86.2	85.2
Primo Stadio	75.7	79.6	78.4
Secondo Stadio	70.0	71.0	68.2
Totale	68.5	70.0	68.3
Volumetrico 1S	98.3	98.6	98.9
Volumetrico 2S	98.4	98.8	99.0

interpalari dei due stadi. Da tale figura è possibile osservare come il transito della corrente nel primo stadio sia molto regolare, confermando il buon funzionamento dello stadio già evidenziato dai risultati quantitativi. La corrente nel secondo stadio non presenta problemi grossolani, tuttavia il moto visualizzato è chiaramente disturbato. Andando ad analizzare i risultati numerici si nota come a valle del primo diffusore la corrente non risulti completamente raddrizzata e sia presente una componente residua di prerotazione. La non corretta rimozione della componente tangenziale della velocità del primo stadio comporta uno scadimento delle prestazioni della girante del secondo stadio testimoniato anche dal rendimento interno della seconda girante che è sensibilmente inferiore a quella della prima girante. Inoltre la prerotazione residua limita lo scambio di energia che la stessa girante può realizzare. La figura 7.11 mostra una visualizzazione tramite vettori del campo di moto a valle nel corpo dello scarico, a valle quindi del secondo diffusore, dove si possono notare come sia presente anche in questo caso una componente tangenziale residua della velocità.

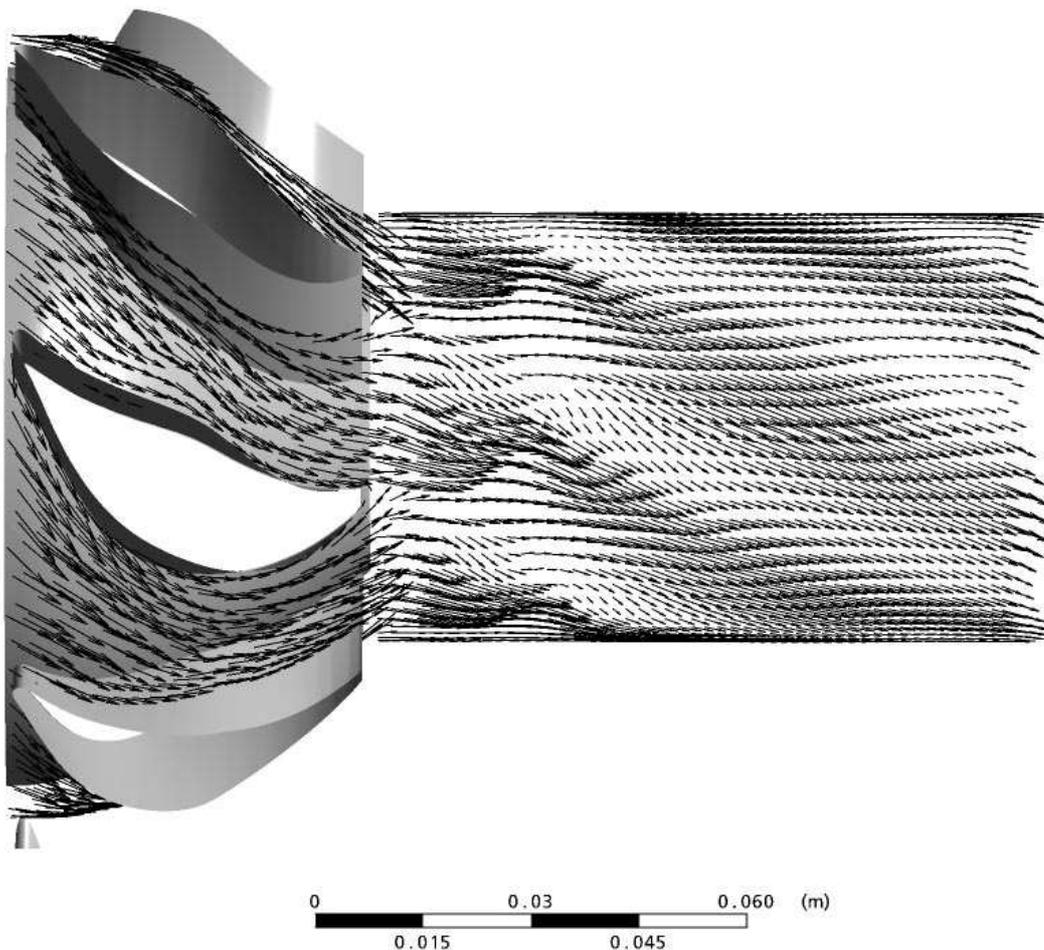


Figura 7.11: Campo di moto della corrente a valle del secondo diffusore.

Dai risultati presentati si può osservare che il diffusore non convoglia alla girante successiva una flusso perfettamente raddrizzato. Alla portata di progetto di $60 \text{ m}^3/\text{h}$ l'angolo della

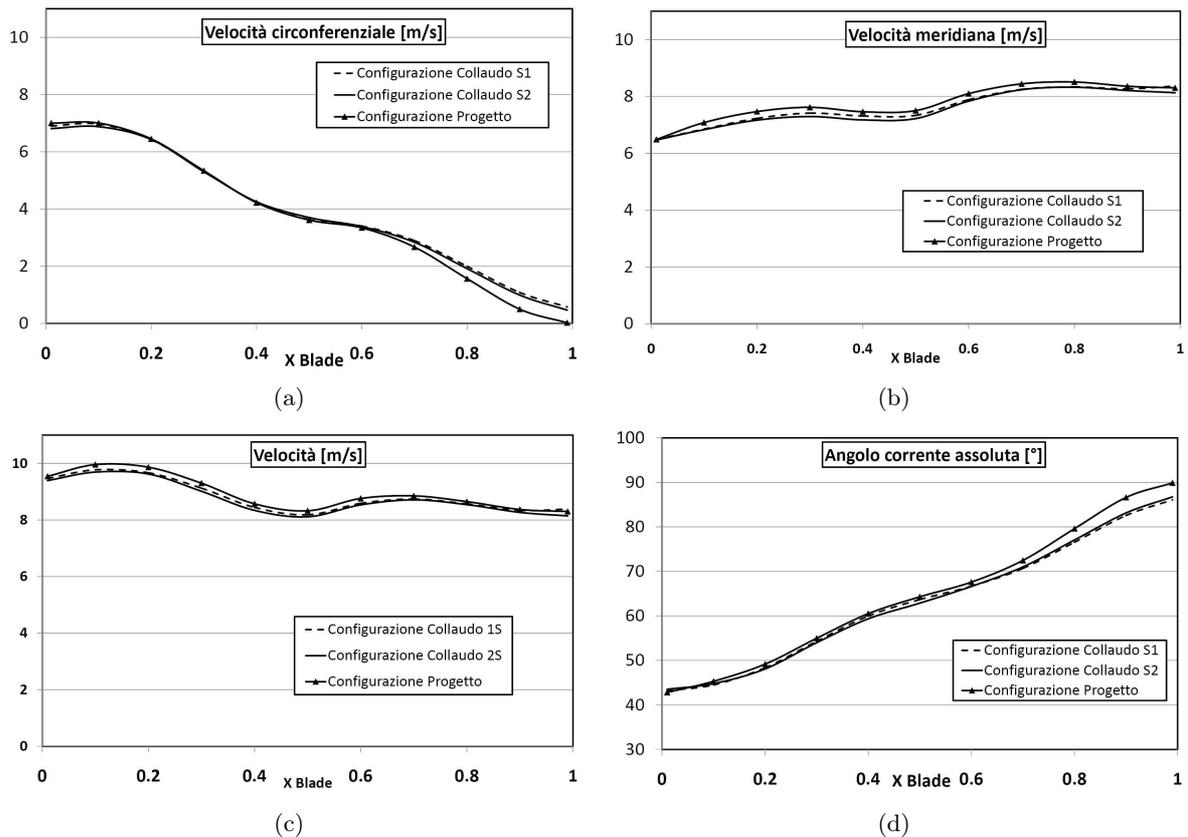


Figura 7.12: Andamenti delle componenti della velocità nei condotti interpalari.

corrente in uscita dal diffusore risulta pari a circa 85° , inferiore a quello atteso, e misurato nelle simulazioni nella configurazione di progetto.

In figura 7.12 si riportano i valori dell'andamento della velocità assoluta, circonferenziale e meridiana e l'angolo della corrente assoluta, rispetto all'ascissa curvilinea normalizzata della linea media della pala per le pale dei diffusori nella configurazione da collaudo e per la pala nella configurazione di progetto. Da tale confronto si può notare come la velocità meridiana, imposta dalla continuità, non mostri differenze tra la configurazione di progetto e la configurazione di collaudo. Viceversa la velocità circonferenziale è uguale nelle due configurazioni per la prima metà della pala, mentre a partire dal 50% dell'ascissa curvilinea cala di meno nella configurazione di collaudo, lasciando una componente residua a valle della pala. Anche l'andamento dell'angolo della velocità assoluta della corrente tra bordo di ingresso e bordo di uscita mostra come da un'ascissa curvilinea normalizzata pari al 50% circa in poi la pala nella configurazione di collaudo non sia più in grado di raddrizzare la corrente come ci si aspettava.

La pala analizzata nella configurazione di progetto non differisce da quella analizzata nella configurazione di collaudo; questo porta a ritenere che il mancato raddrizzamento sia da ricercarsi nella diversa sagomatura del canale meridiano in coda alla pala tra la configurazione di progetto e quella in configurazioni di collaudo. Questa influenza era stata trascurata in sede di progettazione, ma invece queste analisi mostrano come abbiano un'influenza tutt'altro che trascurabile e che non possano essere tralasciate.

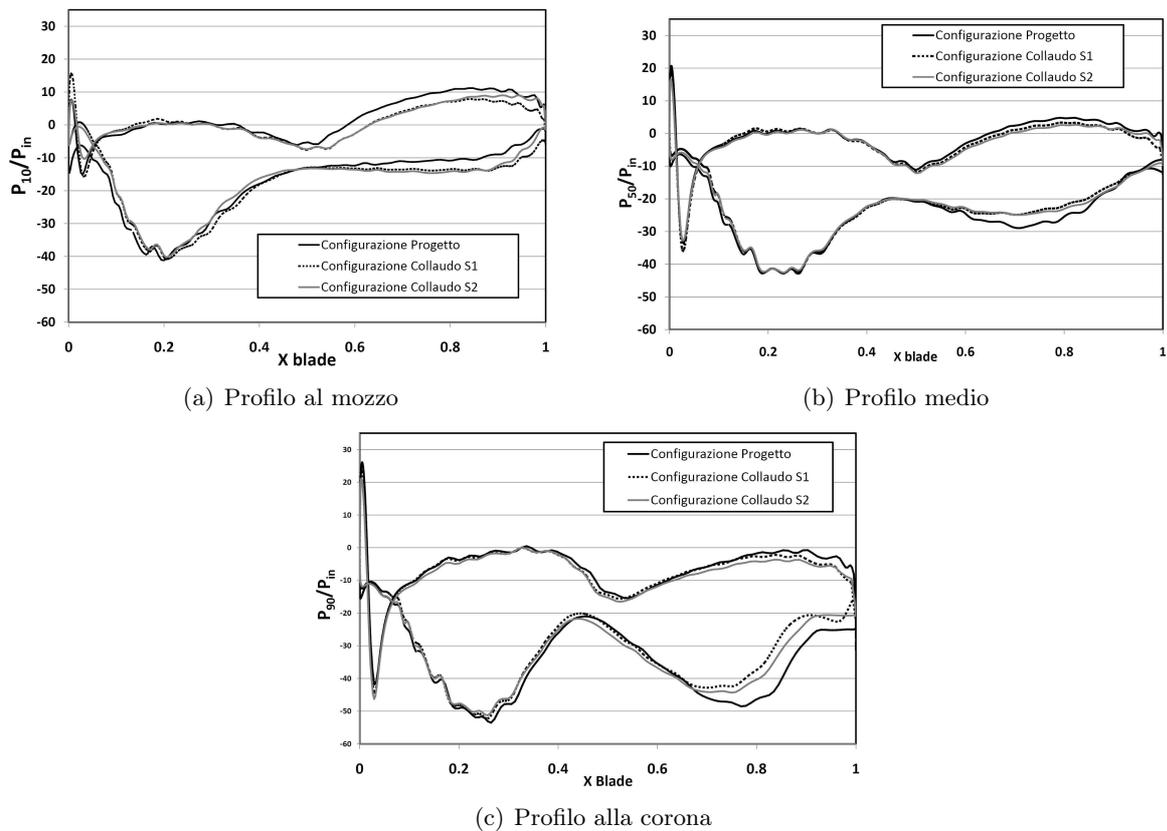


Figura 7.13: Andamento della pressione su estradosso e intradosso: confronto tra le pale del diffusore nella configurazione di progetto e nella configurazione di collaudo.

Anche il confronto degli andamenti della pressione sulle facce delle pale per le due diverse configurazioni, riportato in figura 7.13, mostra come la differenza di pressione tra le facce nella configurazione di progetto sia superiore, a testimonianza di una maggiore capacità di guida.

Si riportano in figura 7.14 i campi di moto in ingresso delle giranti del primo e del secondo stadio, a testimonianza dell'influenza sul campo di moto della forma dei condotti meridiani. Si riporta inoltre in figura 7.15 l'andamento complessivo dell'incremento di pressione statica all'interno dei condotti dei due stadi.

Le simulazioni realizzate quindi per la configurazione di collaudo mostrano come l'interazione tra girante e diffusore dello stesso stadio sia positiva e si riscontrano risultati analoghi a quelli della configurazione di progetto.

Tuttavia i cambiamenti imposti al modello del diffusore in sede di ingegnerizzazione, e motivati dalla necessità di rendere la macchina facilmente assemblabile introducono delle irregolarità non trascurabili nel campo di moto. Tali irregolarità sono, a giudizio di chi scrive la prima causa della non perfetta interazione tra gli stadi successivi.

La simulazione completa quindi fornisce un importante feedback alla progettazione: per consentire una corretta interazione tra gli stadi è necessario rivedere il progetto del diffusore in modo da eliminare tale prerotazione e consegnare alla girante successiva una corrente completamente raddrizzata. Per fare questo la strada che l'osservazione indica è quella di ricondurre,

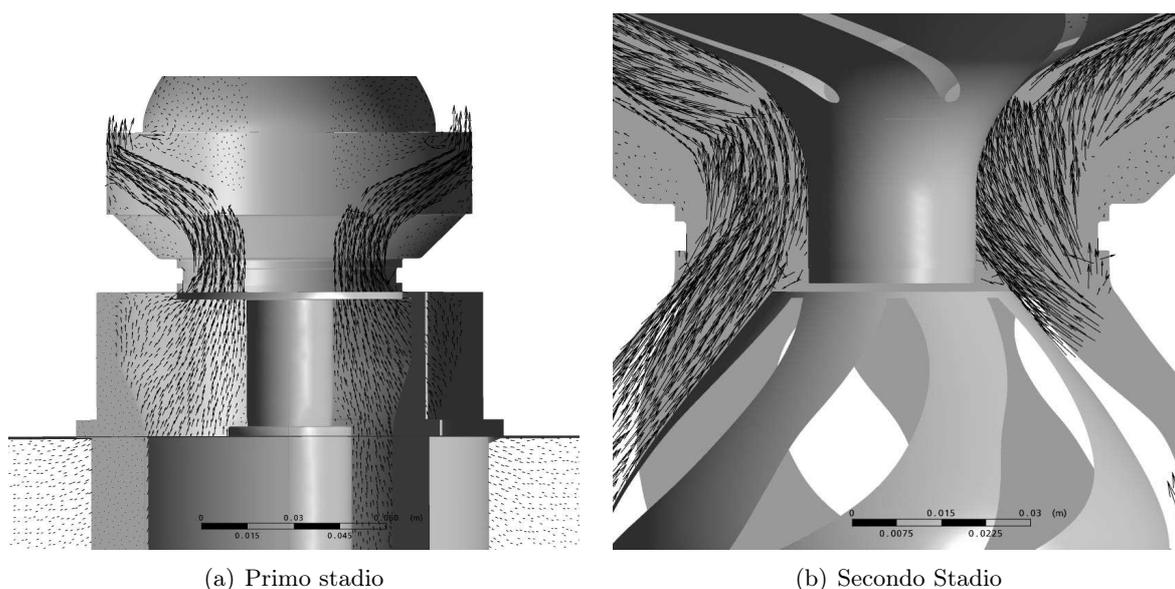


Figura 7.14: Visualizzazione trasversale del campo di moto in ingresso alle giranti.

almeno in parte se possibile, la geometria del prototipo verso la geometria utilizzata in sede di progetto, e di aumentare l'angolo di overturning della pala del diffusore.

E' da notare quindi come le semplificazioni che possono essere apportate al modello in fase di progettazione, o viceversa le modifiche che vengono introdotte in sede di ingegnerizzazione al progetto, e che si considerano ininfluenti nella maggior parte dei casi, invece hanno un'influenza fondamentale e non possono essere trascurate in sede di progettazione.

Inoltre le simulazioni mostrano come i risultati siano promettenti, ma è ancora possibile intervenire per incrementare l'efficienza della pompa, che infatti, stando alle previsioni da queste simulazioni non raggiunge gli obiettivi che si erano prefissati; in primis è necessario individuare una soluzione che raggiunga il massimo del rendimento in corrispondenza della portata di progetto, in modo da rispettare a pieno quelli che sono i vincoli della normativa, che prevedono il calcolo dei parametri prestazionali in corrispondenza della portata di massimo rendimento effettiva.

7.5 Dati sperimentali di collaudo e confronto con i risultati numerici

Sono stati effettuati due diversi collaudi della macchina nella configurazione a due stadi che è stata anche simulata, e in una configurazione con uno stadio solo, per poter ricavare, per confronto le prestazioni del primo e del secondo stadio.

I collaudi sono stati effettuati su una pompa sommersa da 6 pollici come quella illustrata in figura 7.4 in un impianto di prova come quello riportato in 7.2, senza la presenza della grata esternamente all'aspirazione e senza l'otturatore nel corpo valvola. Questo perchè la soluzione che è stata discretizzata nelle simulazioni che sono state indagate nel paragrafo precedente non prevedevano tali parti. Inoltre, anche se la norma prevede di effettuare collaudi con le

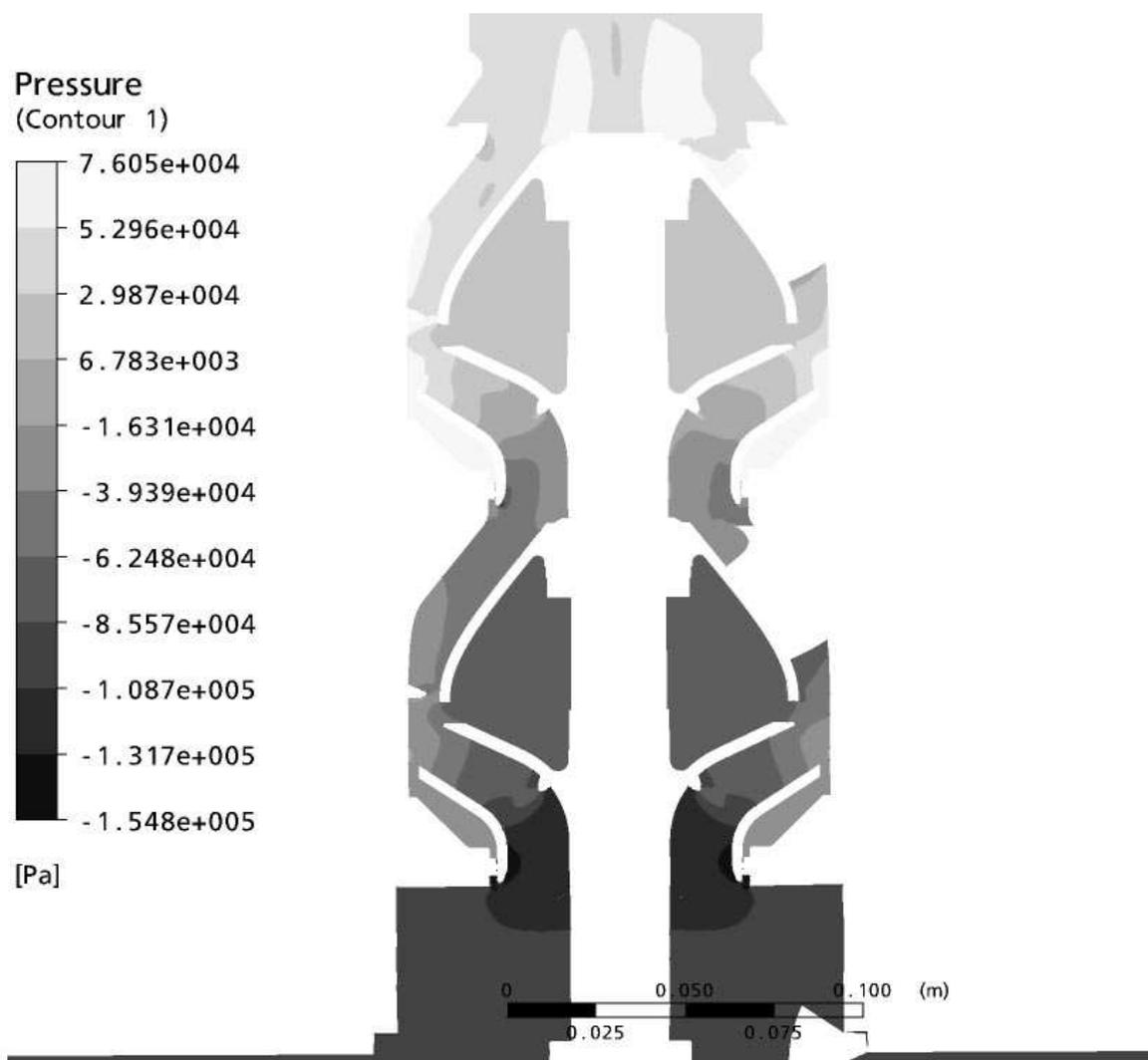


Figura 7.15: Incremento di pressione statica complessivo all'interno della macchina nella configurazione di progetto.

pompe complete di tutti i componenti, e quindi anche di questi due, si è ritenuto opportuno effettuare delle prove delle pompe in queste condizioni per avere un dato più diretto sulle prestazioni. I due componenti che sono stati rimossi infatti introducono delle intense perdite di carico, difficilmente stimabili, che quindi non consentono di isolare con precisione le prestazioni fluidodinamiche dello stadio della macchina.

Nelle figure 7.16 e 7.17 sono riportate le curve di prevalenza e di rendimento della pompa con uno stadio e con due stadi. Purtroppo i dati di collaudo sono stati resi disponibili solo negli ultimi giorni della stesura della tesi e non sono state possibili delle indagini sperimentali più complete.

Le prestazioni che si sono registrate sono state inferiori alle attese in termini di prestazioni, cosa che già le simulazioni della configurazione di collaudo avevano evidenziato nel paragrafo 7.4. Tuttavia i risultati del collaudo si discostano in modo rilevante anche dalle previsioni numeriche. Infatti, se il dato del rendimento è viziato dalle perdite meccaniche che non possono

essere stimate, il dato sperimentale della prevalenza numerico dovrebbe rispecchiare quello sperimentale perchè nella simulazione è stato tenuto conto di tutte le parti idrauliche che concorrono allo scambio di energia nella pompa, ad eccezione delle perdite nel tratto di tubo a valle della pompa tra il corpo valvola e la presa di rilevamento della pressione a valle della macchina. Nella figura 7.16 sono riportati a confronto i dati sperimentali della prevalenza con i dati delle simulazioni nella configurazione di collaudo. Come è possibile vedere, confrontando la curva di collaudo della pompa a due stadi e la curva ottenuta grazie alle simulazioni in configurazione di collaudo si può notare come, mentre la curva sperimentale raggiunge un valore di 13.5 m di prevalenza ad una portata di 60.4 m³/h, la simulazione CFD calcola alla portata di 60 m³/h una prevalenza di 16.9 m. Inoltre lo scostamento tra i dati numerici della prevalenza e i dati sperimentali è consistente anche agli altri valori di portata.

Anche le curve del rendimento sono sensibilmente differenti: le perdite meccaniche non sono note quindi non è possibile confrontare il valore del massimo rendimento dalla curva di collaudo e i dati del rendimento dalle simulazioni numeriche. Tuttavia le perdite meccaniche sono costanti, quindi la portata di massimo rendimento rilevata a collaudo dovrebbe coincidere con quella dei risultati numerici. Invece, in sede di collaudo la pompa ha fatto registrare una portata di 43.5 m³/h in corrispondenza del punto di massimo rendimento, contro il valore di 54 m³/h fatto registrare dalle simulazioni in configurazione di collaudo.

Tuttavia l'osservazione delle curve di collaudo della pompa con uno stadio solo hanno reso possibili ulteriori considerazioni. Infatti come è possibile notare il rendimento del primo stadio ha un massimo per 40 m³/h, ancora inferiore a quella registrata per la pompa a due stadi e ben inferiore a quella ricavata dai risultati numerici sia in configurazione di collaudo che in configurazione di progetto (vedi tabella 7.1).

I risultati del collaudo non sono arrivati in tempo per consentire tutte le verifiche del caso. Tuttavia possono essere avanzate delle ipotesi che spiegano la discrepanza tra questi risultati. Infatti sembra improbabile che una discrepanza tanto grande possa essere attribuita ad errori del codice di calcolo nella stima delle perdite della macchina. La discretizzazione effettuata segue le indicazioni di meshatura indicate in [52] per quello che riguarda le perdite distribuite, inoltre precedenti esperienze mostrano che i codici di calcolo utilizzati consentono una buona stima delle prestazioni delle turbomacchine.

L'osservazione dei dati sperimentali di collaudo sembra indicare la presenza di una prerotazione della corrente a monte del primo stadio. Infatti il primo stadio ha un rendimento complessivo soddisfacente tenendo conto che in sede di collaudo sono presenti tutte le perdite di carico dovute alla presenza del corpo di aspirazione e alla coppia della valvola di scarico. Tuttavia la portata di massimo rendimento risulta spostata a portate sensibilmente inferiori a quelle registrate per le simulazioni di progetto. L'inserimento del secondo stadio, torna a spostare verso le portate più alte il punto di massimo rendimento, indicando come il secondo stadio abbia un comportamento più prossimo a quello atteso.

La presenza di una prerotazione a monte del primo stadio fornisce una possibile spiegazione di come la prevalenza complessiva della macchina sia inferiore alle attese e di come il punto di massimo rendimento del primo stadio sia in corrispondenza di portate basse. Il fatto che

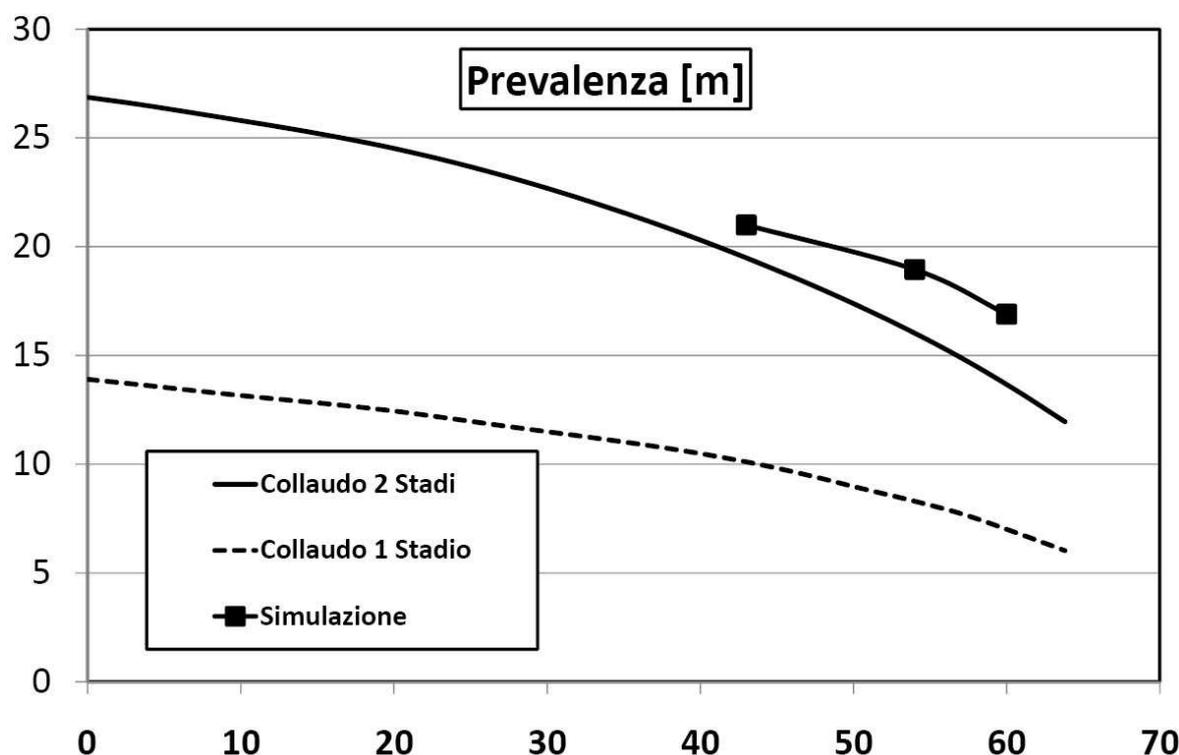


Figura 7.16: Curve della prevalenza di collaudo e confronto con i dati numerici.

il punto di massimo rendimento del secondo stadio torna a spostarsi verso portate superiori è spiegabile con il fatto che il funzionamento di tale stadio è influenzato solo dal campo di moto allo scarico del primo stadio e quindi per il secondo stadio il punto di massimo rendimento si colloca a portate superiori a quello del primo.

Le simulazioni CFD non hanno evidenziato la presenza di una prerotazione a monte della prima girante. In realtà questo sembra dovuto al fatto che il modello e le condizioni al contorno utilizzate non sono appropriate per simulare una pompa sommersa nella sua interezza. Infatti il modello utilizzato per simulare le condizioni al contorno di ingresso nella girante prevede, come illustrato nel paragrafo 7.3, un tratto anulare intorno all'aspirazione con pareti lisce ai lati e una condizione di portata di massa in ingresso con direzione del fluido perpendicolare alla superficie in corrispondenza della superficie cilindrica. In sede di modellazione si è ritenuto tale soluzione sufficiente per consentire alla simulazione numerica di individuare in modo completo le condizioni di flusso in ingresso alla girante del primo stadio. Invece non sembra essere così: è stata effettuata una simulazione del corpo di aspirazione immerso in un volume di fluido maggiore tutto intorno alle feritoie dell'aspirazione molto più esteso di quello utilizzato nella figura della simulazione di collaudo; in figura 7.18 si può vedere la sezione con le dimensioni principali del volume che è stato simulato e il confronto con il volume utilizzato per la simulazione in configurazione di collaudo. E' stata imposta la portata di massa in uscita dal volume di fluido, in corrispondenza dell'inizio del dominio della girante, senza indicare una direzione preferenziale del fluido. Tutta la superficie del volume attorno alla camicia della pompa e del corpo di aspirazione è stato impostata come superficie di *opening*

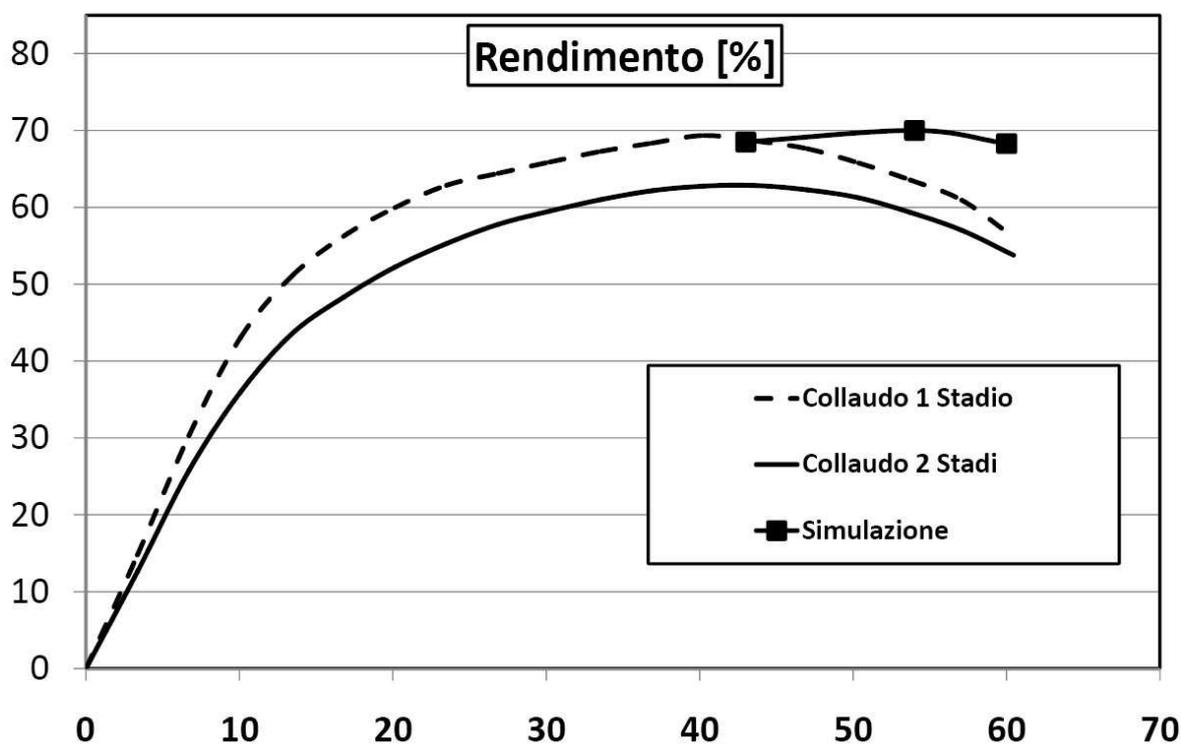


Figura 7.17: Curve del rendimento di collaudo e confronto con i dati numerici.

(nella quale il fluido è libero di scorrere sia in ingresso che in uscita) impostando la pressione totale costante su tutta la superficie; in tale modo la simulazione è libera di identificare la direzione più corretta del fluido in ingresso. Inoltre è stata imposta la condizione di albero rotante. La simulazione mostra che si sviluppa una prerotazione attorno all'albero, dovuta al trascinarsi di questo e che quindi l'angolo in ingresso alla girante è di parecchi gradi differente da quello della direzione assiale. Sembra quindi che le condizioni al contorno imposte nel caso precedente non consentano di calcolare in modo completo il moto della corrente e che impediscano al codice CFD di visualizzare in modo completo la prerotazione.

Non è stato possibile effettuare tutte le indagini del caso sufficienti a validare questi risultati che meritano un'analisi più completa in quanto i risultati sperimentali, la cui osservazione ha fornito lo spunto per queste deduzioni, sono arrivati in fase di stesura della tesi. Tali risultati forniscono comunque importanti indicazioni su come devono essere impostate correttamente le simulazioni delle macchine multistadio e su quella che è l'influenza delle condizioni al contorno sul campo di moto calcolato nelle simulazioni CFD.

7.5. DATI SPERIMENTALI DI COLLAUDO E CONFRONTO CON I RISULTATI NUMERICI

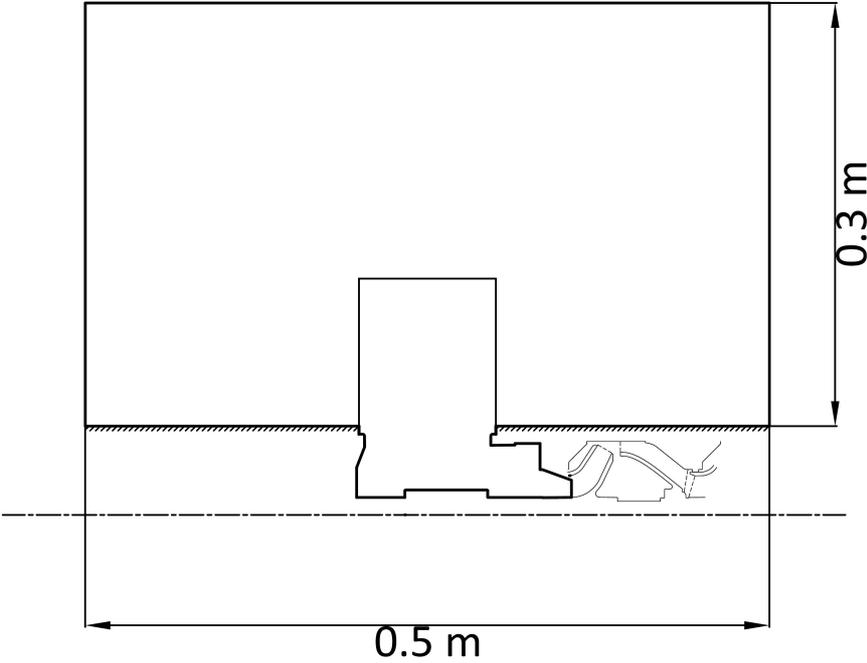


Figura 7.18: Schema del volume discretizzato per analizzare la prerotazione.

Conclusioni

Con la presente tesi di dottorato si è cercato di rispondere alle esigenze della ditta finanziatrice Calpeda S.p.a. che necessitava di una procedura razionale di progettazione per i canali di ritorno delle pompe sommerse multistadio.

In questo lavoro si propone una procedura per la progettazione completa per una macchina multistadio, in quanto la progettazione dei convogliatori palettati delle pompe multistadio non può prescindere dalla progettazione della girante e dall'analisi dell'interazione tra i due organi.

Si ritiene che la procedura proposta consenta una progettazione delle turbomacchine radiali multistadio in tutte le configurazioni che queste possono assumere. Quanto proposto consente inoltre un'interfaccia agevole con i moderni codici di calcolo fluidodinamico, dai quali non si può prescindere per una progettazione avanzata che rispetti parametri prestazionali severi.

Inoltre all'interno della procedura è predisposta una tecnica di raccolta dati che consente di raccogliere dati utili all'azienda in sede di progettazione. Tale tecnica fa uso di modelli di regressione e di algoritmi di ricerca dell'ottimo ai quali è stato affiancato una metodologia di design of experiments sviluppata ad hoc.

Le metodologie proposte sono state applicate alla progettazione di uno stadio di pompa sommersa a flusso misto. La progettazione è stata realizzata grazie ai codici di calcolo CFD che hanno consentito di valutare nel modo più appropriato l'interazione rotore-statore e realizzare una macchina che da dati numerici realizza un rendimento idraulico superiore all'80% per il singolo stadio. A fianco della progettazione è stata realizzata una banca dati di simulazioni CFD per caratterizzare il comportamento dei condotti meridiani dei canali di ritorno. Per realizzare tale banca dati è stato necessario implementare una procedura per la gestione automatica delle simulazioni al computer che ha mostrato una buona flessibilità d'uso e si ritiene potrà essere utilizzata anche in altri casi. L'accesso ai risultati della banca dati ha consentito la scelta più opportuna delle dimensioni di ingombro assiale del diffusore della macchina che è stata progettata.

E' stata inoltre realizzata una simulazione della macchina completa di due stadi, di corpo di aspirazione e valvola di scarico. La simulazione è stata realizzata tenendo conto di tutti i particolari costruttivi e ha permesso di indagare l'effetto non trascurabile dei dettagli introdotti in sede di ingegnerizzazione sui profili ricavati in sede di progettazione; in particolare è stato possibile evidenziare come l'allargamento del condotto meridiano a valle delle pale del diffusore per consentire l'alloggiamento dell'anello di rasamento e della bronzina in gomma, modifichi le capacità di guida della corrente delle pale.

7.5. DATI SPERIMENTALI DI COLLAUDO E CONFRONTO CON I RISULTATI NUMERICI

Il progetto realizzato è stato inoltre prototipato e i risultati delle prove di collaudo sono stati confrontati con i risultati attesi dalle simulazioni numeriche. I risultati sperimentali si sono mostrati sensibilmente differenti dai risultati attesi, in particolare la prevalenza è sensibilmente più bassa e il punto di massimo rendimento risulta spostato verso portate inferiori a quelle attese. Purtroppo i risultati delle prove di collaudo sono arrivati in fase di stesura della tesi e non è stato possibile effettuare tutte le verifiche volte a giustificare la discrepanza tra i dati sperimentali e numerici. Tuttavia i primi risultati che è stato possibile ottenere mostrano la presenza di una prerotazione all'imbocco della prima girante causata dalla rotazione dell'albero all'interno del corpo di aspirazione. Tale prerotazione, che non era stata conteggiata in sede di progetto, rappresenta una spiegazione plausibile alla forma della curva caratteristica ottenuta in sede di collaudo.

Tale prerotazione non è stata visualizzata dalle analisi CFD; in tali simulazioni l'aspirazione dalla vasca è stata modellizzata con un tratto anulare di fluido, con pareti laterali e ingresso di portata di massa con velocità perpendicolare alla superficie in corrispondenza della superficie cilindrica. Simulando una porzione più ampia della vasca intorno alla camicia della pompa e impostando una superficie con condizione al contorno di fluido a pressione totale costante invece si nota l'insorgenza della prerotazione causata dalla rotazione dell'albero. Si ritiene, quindi, che le impostazioni delle condizioni al contorno utilizzate nella simulazione completa non siano indicate per la simulazione delle pompe sommerse.

In conclusione si ritiene che la procedura proposta in questa tesi sia agevole e consenta una progettazione efficace delle pompe multistadio. Rimangono aperti, e si rimandano a studi futuri, gli interrogati proposti dalle simulazioni più complesse riguardo all'influenza dei particolari costruttivi inseriti in sede di ingegnerizzazione sulle prestazioni della macchina e riguardo alle condizioni al contorno che devono essere assegnate per simulare nel modo più corretto una pompa sommersa multistadio.

Simbologia

Dove non specificato diversamente:

α angolo tra la corrente assoluta e la direzione circonferenziale

β angolo costruttivo della pala

ξ inclinazione della traccia della pala

W velocità relativa

C velocità assoluta

U velocità periferica

C_M componente meridiana della velocità

C_U componente circonferenziale della velocità

$m' = \int dm/r$

θ coordinata circonferenziale

ζ coefficiente di ostruzione

B larghezza del canale meridiano

γ inclinazione rispetto all'asse della macchina su un piano meridiano

K numero tipico di macchina

ψ numero tipico di pressione

ϕ numero tipico di flusso

7.5. DATI SPERIMENTALI DI COLLAUDO E CONFRONTO CON I RISULTATI NUMERICI

Elenco delle figure

1.1	Schematizzazione delle varie parti di un canale interstadio a flusso radiale. . .	10
1.2	Diverse tipologie di palettature nei canali interstadio:a) schiera palare nel solo canale di ritorno; b) doppia schiera palare nel diffusore e nel canale di ritorno; c) schiera palare continua.	11
1.3	Diffusore a boccia per macchine con girante a flusso misto.	11
1.4	Grandezze interessate nella definizione dei canali interstadio secondo Stepanoff.	12
1.5	Correzione della sezione secondo Meng e Jackson	14
1.6	Formulazione degli angoli in Ellis.	17
1.7	Assegnazione per l'andamento del momento angolare della quantità di moto proposta in [17].	19
1.8	Schema dell'utilizzo delle tecniche CFD nella progettazione.	24
1.9	Esempio di suddivisione, in termini di portata e prevalenza, del campo di funzionamento, dei vari modelli di una serie di pompe.	25
2.1	Esempio di curva di Bezier di primo grado.	29
2.2	Esempio di curva di Bezier di secondo grado.	30
2.3	Schematizzazione della parte idraulica di una pompa centrifuga a basso numero tipico.	33
2.4	Analisi delle linee di corrente nel condotto meridiano di una girante.	34
2.5	Andamento della derivata del momento angolare della quantità di moto lungo la linea media della sezione di riferimento di una girante a semplice curvatura.	34
2.6	Andamento della velocità circonferenziale che risulta da un'assegnazione come in figura 2.5.	35
2.7	Flowchart del dimensionamento della pala di una girante a basso numero tipico di macchina.	36
2.8	Schema della superficie media della pala sul piano conforme $m' - \theta$	38
2.9	Parametrizzazione di una sezione della pala sul piano conforme.	39
2.10	Rappresentazione sul piano circonferenziale di una pala corrispondente a quella di figura 2.8.	40
2.11	Flowchart del dimensionamento della pala di una girante a numero tipico di macchina $K > 0.5$	41

2.12	Differenti tipi di palettature che possono essere alloggiare in un canale di ritorno di un condotto multistadio.	42
2.13	Inclinazione della pala rispetto alla linea media della pala.	43
2.14	Flowchart del dimensionamento della pala del diffusore a spessore costante. . .	44
2.15	Parametrizzazione dell'andamento delle velocità nel condotto interpalare del diffusore.	45
2.16	Flowchart del dimensionamento della pala del diffusore a spessore variabile. . .	47
2.17	Grandezze coinvolte nella parametrizzazione del condotto meridiano.	48
3.1	Schema dell'ottimizzazione con la Simulated Annealing.	54
3.2	Flow-chart dell'evoluzione di una popolazione con un algoritmo genetico. . . .	57
3.3	Esempio di one-point crossover.	59
3.4	Esempio di two-point crossover.	60
3.5	Flow-chart della ricerca dell'ottimo tramite PSO.	65
3.6	Applicazione del vincolo nell'ottimizzazione per mezzo della riflessione: esempio 2D.	66
3.7	Sciame bidimensionale di particelle intorno ad un minimo per iterazioni successive: diminuzione delle distanze reciproche tra le particelle.	67
4.1	Esempio di ipercubo latino.	70
4.2	Errore dovuto all'overfitting tra la funzione approssimante (linea tratteggiata) e quella da approssimare (linea continua).	78
4.3	Differenza tra l'errore sul training set e l'errore sul validation set nel processo di addestramento, e comparsa dell'overfitting	79
4.4	Schematizzazione di un neurone di rete neurale.	81
4.5	Funzione di trasferimento tangenziale e lineare dei neuroni delle reti neurali	81
4.6	Esempio della struttura di una rete neurale a più layer	82
4.7	Esempio di funzione di trasferimento a base radiale.	86
4.8	Schema di neurone di rete di funzione a base radiale.	87
4.9	Esempio dell'influenza del parametro b sul comportamento della funzione di trasferimento	87
4.10	Esempi di diverse funzione di correlazione che possono essere utilizzate con il metodo del Kriging. L'intensità delle linea corrisponde a parametri θ più alti.	93
5.1	Diagramma di flusso che esplica il processo di creazione di una banca dati. . .	97
5.2	Esempio di sovrapposizione di ipercubi.	99
5.3	Risultato della sovrapposizione di tre ipercubi diversi a), b) e c) sovrapposti in $a+b+c$)	100
6.1	Differenti geometrie da produrre per stampaggio.	105
6.2	Diversi tipi di stampo in risposta alla diversa esigenza di produzione: a) stampo aperti e chiusi per stampi senza sottosquadri in una direzione; b) stampo con carro mobile estratto lateralmente per pezzi complessi con sottosquadri. . . .	105

6.3	Visualizzazione degli angoli di sottosquadro sulle facce di una pala di una girante a semplice curvatura	107
6.4	Realizzazione tramite stampaggio delle girante a doppia curvatura.	107
6.5	Diagramma a blocchi delle operazioni da svolgere per portare a termine una simulazione.	108
6.6	A sinistra:dimensioni della geometria di girante radiale alloggiata negli ingombri radiali imposti;a destra: possibile esempio di soluzione a flusso misto per la stessa dimensione radiale.	113
6.7	Studio di diversi canali meridiani della giranti a flusso misto con linea media del canale meridiano inclinata di 45° rispetto alla direzione assiale.	115
6.8	Studio di diversi canali meridiani della giranti a flusso misto con linea media del canale meridiano inclinata di 60° rispetto alla direzione assiale.	115
6.9	Schema del volume discretizzato e superfici di riferimento utilizzate per il computo dello scambio di energia delle simulazioni CFD.	117
6.10	Visione, sul piano meridiano, dei nodi della discretizzazione del volume della pala e delle zone ausiliarie di adduzione e di scarico.	118
6.11	Particolare della distribuzione dei nodi della discretizzazione attorno alla pala.	118
6.12	Confronto tra due condotti meridiani con diversa inclinazione rispetto all'asse della macchina.	120
6.13	Caratteristiche principali del canale meridiano della girante progettata con particolare dei bordi di ingresso.	121
6.14	Curve caratteristiche ricavate da simulazioni numeriche della girante progettata denominata Girante P	123
6.15	Comportamento fluidodinamico della palettatura della Girante P	124
6.16	Modulo della velocità meridiana in corrispondenza del bordo di uscita della pala.	125
6.17	Andamento di alcune grandezze rappresentative della forma della pala della girante P	126
6.18	Esemplificazione delle condizioni al contorno utilizzate per le simulazioni dei canali meridiani dei diffusori.	128
6.19	Diagramma a blocchi della protocollo per la gestione automatica della banca dati di simulazioni.	130
6.20	Zone ausiliarie utilizzate per la corretta modellazione del flusso all'interno del condotto meridiano del diffusore.	133
6.21	Esempio di settore assialsimmetrico di 1° utilizzato per modellare il canale meridiano di un diffusore e relativa mesh.	135
6.22	Esempio della discretizzazione che si ottiene con l'algoritmo proposto sul piano meridiano.	136
6.23	Andamento delle perdite al variare del parametro L_{H3} e del numero di layer di nodi in direzione circonferenziale, con riferimento alla tabella 6.23	138
6.24	Andamento delle perdite al variare del parametro L_{H3} e del numero di layer di nodi in direzione circonferenziale, con riferimento alla tabella 6.24	139

6.25	Influenza dell'infittimento della discretizzazione sui risultati della simulazioni; dati di tabella 6.6	140
6.26	Esempi di diverse geometrie che sono state discretizzate per la realizzazione della banca dati.	143
6.27	Andamento dell'errore di regressione nell'addestramento di reti neurali.	145
6.28	Influenza della struttura delle reti neurali sull'errore di regressione.	146
6.29	Influenza dell'ampiezza delle funzioni a base radiale sulla precisione delle RBFN.	147
6.30	Risultati dell'addestramento di diversi modelli di Kriging.	148
6.31	Confronto tra i vari modelli di regressione.	149
6.32	Andamento della stima dell'errore di regressione al variare della numerosità della banca dati per una rete neurale con un layer nascosto con 20 neuroni.	150
6.33	Schematizzazione dell'accostamento di girante e diffusore.	152
6.34	Influenza dell'ingombro assiale sulle perdite nel condotto meridiano.	155
6.35	Schematizzazione del volume discretizzato per la modellizzazione dello stadio completo di pompa verticale sommersa: divisione in domini e superfici di riferimento.	157
6.36	Schizzo con le indicazioni delle dimensioni principali della palettatura.	160
6.37	Distribuzione degli spessori normali alla linea media ricavati dall'assegnazione della velocità di attraversamento per il diffusore SV	161
6.38	Andamento assegnato della velocità di attraversamento nel condotto interpalare per il diffusore SV e corrispondente andamento degli angoli costruttivi della pala.	162
6.39	Vista frontale e laterale del condotto interpalare del diffusore con particolare del bordo di ingresso (a) e del condotto interpalare (b).	162
6.40	Punti di ristagno in corrispondenza del naso della palettatura del diffusore.	163
6.41	Campo di moto e campo di pressione nell'interazione tra girante e diffusore alla portata di progetto per varie sezioni blade to blade.	164
7.1	Modello 3D del prototipo.	166
7.2	Curve di funzionamento per il prototipo ricavate da simulazione numerica.	167
7.3	Sezione 2D del prototipo realizzato con tutti i particolari costruttivi.	168
7.4	Disegno della pompa sommersa completa di tutte le parti che la compongono.	170
7.5	Immagine dei prototipi realizzati.	171
7.6	Schema dell'impianto di prova per pompe verticali sommerse secondo la normativa UNI 9906.	172
7.7	Schema dei domini in cui è stato suddiviso il modello per poter essere discretizzato.	173
7.8	Discretizzazione del trafileamento tra girante e anello di rasamento.	175
7.9	Discretizzazione del modello dello stadio.	176
7.10	Campo di moto sulla superficie interpalare.	178
7.11	Campo di moto della corrente a valle del secondo diffusore.	180
7.12	Andamenti delle componenti della velocità nei condotti interpalari.	181

7.13 Andamento della pressione su estradosso e intradosso: confronto tra le pale del diffusore nella configurazione di progetto e nella configurazione di collaudo.	182
7.14 Visualizzazione trasversale del campo di moto in ingresso alle giranti.	183
7.15 Incremento di pressione statica complessivo all'interno della macchina nella configurazione di progetto.	184
7.16 Curve della prevalenza di collaudo e confronto con i dati numerici.	186
7.17 Curve del rendimento di collaudo e confronto con i dati numerici.	187
7.18 Schema del volume discretizzato per analizzare la prerotazione.	188

Bibliografia

- [1] Goel T., Dorney D.J., Haftka R.T., Shyy W.. Improving the Hydrodynamic Performance of Diffuser Vanes via Shape Optimization. *Computer and Fluids*, 2007.
- [2] Sacks J., Wlech W.J., Mitchell T.J., Wynn H.P.. Design and Analysis of Computer Experiments. *Statistical Science*, 1989.
- [3] Efron B.. Estimatinig the Error Rate of a Prediction Rule: Improvement on Cross-validation. *Journal of the American Statistical Association*, 1983.
- [4] Breiman L., Spector P.. Submodel Selection and Evaluation in Regression. *International Statistical Review*, 1992.
- [5] Lendasse A., Simon G., Wertz V., Verleysn M.. Fast Bootstrap Methodology for Regression Model Selection. *Neurocomputing*, 2004.
- [6] Jeong S., Murayama M., Yamamoto K.. Efficient Opitimization Design Method Using Kriging Model. *AIAA Journal of Aircraft*, 2005.
- [7] McKay M.D., Bechman R.J., Conover W.J.. A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code. *Tecnometrics*, 1979.
- [8] Johnson M., Moore L., Ylvisaker D.. Minimax and Maxmin Distance Designs. *Journal of Statistical Planning and Inference*, 1990.
- [9] Simpson T.W., Lin D. K. J.,Chen W.. Sampling Strategies for Computer Experiments:Design and Analysis. *International Journal of Reliability and Application*, 2000.
- [10] Casey M.V.. A Computational Geometry for the Blades and Internal Flow Channels of Centrifugal Compressors. *Transactions of the ASME*, 1983.
- [11] Zangeneh M.. A Compressible Three-Dimensional Design Method for Radial and Mixed Flow Turbomachinery Blades. *International Journal of Numerical Methods in Fluids*, 1991.
- [12] Queipo N.,Haftka R.,Shyy W.. Surrogate-Based Analysis and Optimization.. *Progress in Aerospace Sciences*, 2005.

- [13] Kovats A.. Diffusers of Multistage Centrifugal Pumps. *ASME Mechanics Conference-Houston* , 1983.
- [14] Boncinelli P.,Biagi R.,Focacci A.,Corradini U.,Arnone A.,Bernacca M.,Borghetti M.. Bowl-Type Diffusers for Low Specific-Speed Pumps: An Industrial Application. *ASME Journal of Turbomachinery* , 2008.
- [15] Borges J. E.. A Three-Dimensional Inverse Method for Turbomachinery: Part I-Theory. *Transaction of the ASME*, 1990.
- [16] Goto A., Zangeneh M.. Hydrodynamic Design of Pump Diffuser Using Inverse Design Method and CFD. *ASME Journal of Fluids Engineering*, 2002.
- [17] Zangeneh M., Goto A., Takemura T.. Suppression of Secondary Flows in a Mixed-Flow Pump Impeller by Application of Three-Dimensional Inverse Design Method: Part 1-Design and Numerical Validation. *Transaction of the ASME*, 1996.
- [18] Pierret S.,Van den Braembussche R.A.. Turbomachinery Blade Design Using a Navier-Stokes Solver and Artificial Neural Network. 1999.
- [19] Veress A.,Van den Braembussche R.A.. Inverse Design and Optimization of a Return Channel for a Multistage Centrifugal Compressor. 2004.
- [20] Demeulenaere A.,Van den Braembussche R.A.. Three-Dimensional Inverse Method for Turbomachinery Blading Design. 1998.
- [21] de Vito L.,Van den Braembussche R.A.,Deconink H.. A Novel Two-Dimensional Viscous Inverse Design Method for Turbomachinery Blading. 2003.
- [22] Ellis G.O.. Crossover System for Centrifugal Compressors. 1960.
- [23] Bezier P.. *emploi des machines a commande numerique*. 1970.
- [24] Zangeneh M.. inverse design of centrifugal compressor vaned diffuser in inler shear flows.. 1996.
- [25] Kergourlay G.,Younsi M.,Bakir F.,Rey R.. influence of splitter blades on the flow field of a centrifugal pump: Test-analysis comparison. 2007.
- [26] Goto A., Nohmi M.,Sakurai T.,Sogawa Y. hydrodynamic design system for pump based on 3-d cad,cfd,and inverse design method.. 2002.
- [27] Goel T.,Haftka R.T.,Shyy W.,Queipo N.V.. ensemble of surrogates. 2006.
- [28] Rai M.M. Single and multiple objective optimization with differential evolution and neural networks. In *Lecture Series-Introduction to Optimization Methods and Tools for Multidisciplinary Design in Aeronautics and Turbomachinery*. von Karman Institute for Fluid Dynamics, 2008.

-
- [29] Verstraete T., Van den Braembussche R.A. Multidisciplinary optimization of turbomachinery components including heat transfer and stress prediction. In *Lecture Series-Introduction to Optimization Methods and Tools for Multidisciplinary Design in Aeronautics and Turbomachinery*. von Karman Institute for Fluid Dynamics, 2008.
- [30] Michaelides K.V., Tournlidakis A., Elder R.L. Use of cfd for the three-dimensionale hydrodynamic design of vertical diffuser pumps. In *Advances of CFD in Fluid Machinery Desing*. Professional Engineering Publishing, 2003.
- [31] Meng, Jackson. A Continuous Crossover System Design. 1983. Huston.
- [32] Aungier R.H.. Aerodynamic Design and Analysis of Vaneless Diffusers and Return Channels. 1993. Presented at International Gas Turbine and Aeroengine Congress and Exposition, Cincinnati Ohio.
- [33] Kennedy J., Eberhart R.. Particle Swarm Optimization. 1995. Proceedings IEEE International Conference on Neural Network, Perth Australia.
- [34] Shi J., Eberhart R.. A Modified Particle Swarm Optimization. 1998. International Conference of Evolutionary Computation, Anchorage Alaska.
- [35] Hagan, M.T., Menhaj M.. Training feed-forward networks with the Marquardt algorithm. 1999. IEEE Transactions on Neural Networks.
- [36] Myers H.R., Montgomery D.C.. *Response Surface Methodology*. Wiley, 2002.
- [37] Haykin S.. *Neural Network, a comprehensive foundation*. Prentice Hall, 1999.
- [38] Montgomery D.C.. *Design and Analysis of Experiments*. Wiley, 2001.
- [39] Bellman R.E.. *Dynamic Programming*. Princeton University Press, 1957.
- [40] Cherkassky V., Friedman J.H., Wechsler H.. *From Statistics to Neural Networks: Theory and Pattern Recognition Applications*. Springer-Verlag, 1995.
- [41] van Laarhoven P.J.M., Aarts E.H.L.. *Simulated Annealing: Theory and Application*. Kluwer Academic Publishers, 1987.
- [42] Japikse D., Marscher W. D., Furst R. B.. *Centrifugal Pump Design and Performance*. Concept ETI, 1997.
- [43] Stepanoff A. J.. *Centrifugal and Axial Flow Pumps*. J. Wiley, 1957.
- [44] Lazarkiewicz S., Troskolanski A.T.. *Impeller Pumps*. Pergamon Press, 1957.
- [45] AA VV. *New Ideas in Optimization*. McGraw Hill, 1999.
- [46] Thomas Weise. *Global Optimization Algorithms - Theory and Application*. 2009. Online available at <http://www.it-weise.de/>.

- [47] Rumelhart D. E., Mc Clelland J.L.. Parallel Distributed Processing: Explorations in the Mirostructure of Cognition. Technical report, M.I.T, 1986.
- [48] Lophaven S.N.,Nielsen H.B., Sondergaard J.. DACE-A Matlab Kringing Toolbox. Technical report, Technical University of Denmark, 2002.
- [49] Lophaven S.N.,Nielsen H.B., Sondergaard J.. Aspects of the Matlab Toolbox dace. Technical report, 2002.
- [50] Van den Braembussche R.A.. Optimization of Radial Impeller Geometry. Technical report, 2005.
- [51] AA VV. *Matlab User's Guide*.
- [52] AA VV. *ANSYS-CFX Solver Modeling Guide*.
- [53] AA VV. *ANSYS-CFX Post User's Guide*.
- [54] AA VV. *ANSYS-CFX Pre User's Guide*.
- [55] Macor A.. *Diffusori-convogliatori di macchine reversibili pluristadio per impianti di accumulazione idrica*. PhD thesis, Università di Padova, 1987.
- [56] Rothstein E.. *Experimentelle und Theoretische Untersuchung der Strömungsvorgänge in Rückfürkanälen von Radialverdichterstufen insbesondere solchen mit geringen Kanalbreite*. PhD thesis, RWTH Aachen, Germany, 1984.
- [57] van den Bergh F.. *An Analysis of Particle Swarm Optimizers*. PhD thesis, University of Pretoria, 2001.
- [58] AA VV. Prove di prestazione idrauliche e criteri di accettazione. Livelli 1 e 2. Technical report, 2002.

Ringraziamenti

Ringrazio innanzi tutto i Professori Ardizzon e Pavesi per l'opportunità che mi hanno concesso con questi tre anni di dottorato e per avermi insegnato tante cose che sicuramente nel futuro mi saranno utili. Al loro fianco non posso non mettere, nella lista dei ringraziamenti, la Calpeda S.p.a. per aver contribuito a questo progetto di ricerca.

Dopo i ringraziamenti ufficiali si vorrebbe fare quelli non ufficiali e si vorrebbe ringraziare tutte le persone che ti hanno accompagnato, direttamente o indirettamente negli anni del dottorato, ma in questo caso si rischia di uscire da quelli che sono i ringraziamenti a tono con l'occasione. Ringrazio comunque tutti, anche se non centrate nulla con l'università, con il dottorato e le turbomacchine.

Un grazie particolare va ai miei colleghi di avventura Giovanna, Albero e Antonio che hanno contribuito a questa tesi dal punto di vista umano e dal punto di vista scientifico.

Ringrazio la mia famiglia.

Non mi resta che ringraziare Carla!