

**POLITECNICO DI MILANO**

Scuola di Ingegneria dei Sistemi

Corso di Laurea Magistrale in  
Ingegneria Gestionale



**Multisensor data fusion for quality inspection  
of free-form surfaces**

Relatore: Ch.ma Prof. Bianca Maria COLOSIMO

Tesi di Laurea Magistrale di:  
Luca PAGANI Matr. 749614

Anno Accademico 2010-2011



# Indice

<b>Sommario</b>	<b>xiii</b>
<b>1 Strumenti di misura monosensore</b>	<b>1</b>
1.1 Misura . . . . .	1
1.2 Strumenti di misura . . . . .	2
1.2.1 Coordinate measuring machine . . . . .	2
1.2.2 Luce strutturata . . . . .	4
1.3 Metrologia di larga scala . . . . .	7
1.4 Sistema MScMS-II . . . . .	9
<b>2 Multisensor data fusion</b>	<b>15</b>
2.1 Sensori multirisoluzione . . . . .	15
2.2 Multisensor Coordinate Metrology . . . . .	18
<b>3 Modello gerarchico</b>	<b>25</b>
3.1 Stato dell'arte . . . . .	25
3.2 Processo gaussiano . . . . .	27
3.3 Processi gaussiani per dati multirisoluzione . . . . .	28
3.3.1 Modello a bassa risoluzione . . . . .	29
3.3.2 Modello ad alta risoluzione . . . . .	29
3.4 Modello bassa risoluzione . . . . .	30
3.4.1 Introduzione . . . . .	30
3.4.2 Stima dei parametri . . . . .	31
3.4.3 Valore atteso e predizione in nuovi punti e loro proprietà	33
3.5 Modello alta risoluzione . . . . .	37
3.5.1 Introduzione . . . . .	37
3.5.2 Stima dei parametri . . . . .	39
3.5.3 Valore atteso e predizione in nuovi punti . . . . .	45
<b>4 Valutazione prestazioni modello gerarchico</b>	<b>51</b>
4.1 Introduzione . . . . .	51
4.2 Modelli a bassa risoluzione . . . . .	53
4.3 Modelli con punti ad alta risoluzione . . . . .	56
4.3.1 Modello alta risoluzione . . . . .	57

4.3.2	Modello addition . . . . .	58
4.3.3	Modello fusion . . . . .	59
4.4	Analisi residui . . . . .	61
4.5	Decimazione mediante metrica quadratica dei punti . . . . .	67
4.5.1	Modello fusion su modello con decimazione quadratica . . . . .	72
4.6	Conclusioni . . . . .	78
<b>5</b>	<b>Campionamento punti ad alta risoluzione</b>	<b>81</b>
5.1	Curvatura gaussiana . . . . .	81
5.2	Piano di campionamento proposto . . . . .	85
5.3	Analisi . . . . .	99
5.3.1	Analisi 1 . . . . .	100
5.3.2	Analisi 2 . . . . .	106
5.3.3	Analisi 3 . . . . .	109
5.3.4	Analisi 4 . . . . .	111
5.4	Conclusioni . . . . .	115
	<b>Conclusioni</b>	<b>117</b>

# Elenco delle figure

1.1	Macchina di misura a coordinate [1]	3
1.2	Esempio scanner a luce strutturata	4
1.3	Pattern di luce	5
1.4	Deformazione patter luce	5
1.5	Tecnica di triangolazione	6
1.6	Sistema di misura a coordinate sferiche	8
1.7	Principio di funzionamento triangolazione	9
1.8	Principio di funzionamento trilaterazione	9
1.9	Sistema MScMS II, la linea tratteggiata rappresenta la connessione virtuale tra i sensori ed i marker del tastatore, indicati con A e B	10
1.10	Tastatore wireless	11
1.11	Rete di sensori	11
1.12	Triangolazione per determinare i punti dei marker	11
2.1	Classificazione metodi di acquisizione [45]	18
2.2	Esempio di sistema ottico e meccanico integrato [2]	19
2.3	Classificazione metodi di acquisizione [11]	19
2.4	Tastatore con range interno di misura [3]	20
2.5	Scanner laser [5]	21
2.6	Configurazione sistema [49]	22
2.7	Pezzo misurato [49]	23
3.1	Esempio kernel gaussiano	28
4.1	Artefatto fisico	52
4.2	Punti a bassa risoluzione	52
4.3	Termini che compongono il modello	54
4.4	Superficie predetta	54
4.5	Residui modello	55
4.6	Residui con intervalli di confidenza (blu) e di predizione (verde)	55
4.7	Semiampiezza intervallo di confidenza e immagine	56
4.8	Punti alta accuratezza campionati	56
4.9	Intervallo di confidenza sulla media	58

4.10	Intervallo di confidenza sulla media . . . . .	58
4.11	Effetto processo gaussiano . . . . .	60
4.12	Semi-ampiezze intervalli di confidenza . . . . .	61
4.13	Residui modello alta risoluzione . . . . .	62
4.14	Residui modello <i>addition</i> . . . . .	62
4.15	Residui modello bassa accuratezza e <i>fusion</i> . . . . .	63
4.16	Boxplot residui . . . . .	63
4.17	Boxplot quadrato residui . . . . .	64
4.18	Residui modello bassa (nero) ed alta (rosso) . . . . .	65
4.19	Intervalli di confidenza mediana residui . . . . .	66
4.20	Intervalli di confidenza mediana quadrato residui . . . . .	66
4.21	Triangolazione di Delaunay . . . . .	68
4.22	Triangolazione di Delaunay . . . . .	70
4.23	Modello con decimazione quadratica . . . . .	70
4.24	Residui modello decimazione quadratica . . . . .	71
4.25	Residui e intervalli di confidenza e predizione modello con decimazione quadratica . . . . .	72
4.26	Semiampiezza intervallo di confidenza sul valore atteso . . . . .	72
4.27	Effetto processo gaussiano con decimazione quadratica . . . . .	73
4.28	Residui modello bassa accuratezza e <i>fusion</i> decimazione quadratica . . . . .	74
4.31	Intervalli di confidenza 99% mediana residui . . . . .	75
4.29	Stima densità residui . . . . .	76
4.30	Stima densità quadrati residui . . . . .	77
4.32	Intervalli di confidenza 99% mediana quadrato residui . . . . .	78
5.1	Superficie <i>peaks</i> con curvatura gaussiana . . . . .	84
5.2	Superficie del modello a bassa risoluzione . . . . .	88
5.3	Densità di probabilità campionamento proposto . . . . .	90
5.4	Campionamenti con 1 classe . . . . .	91
5.5	Campionamenti con 25 classi . . . . .	92
5.6	$\omega_1 = 0, 1, \omega_2 = 0, 9$ . . . . .	94
5.7	$\omega_1 = 0, 3, \omega_2 = 0, 7$ . . . . .	95
5.8	$\omega_1 = \omega_2 = 0, 5$ . . . . .	96
5.9	$\omega_1 = 0, 7, \omega_2 = 0, 3$ . . . . .	97
5.10	$\omega_1 = 0, 9, \omega_2 = 0, 1$ . . . . .	98
5.11	Disegno proposto . . . . .	101
5.12	Quadrato latino . . . . .	101
5.13	Disegno misto . . . . .	102
5.14	Superficie con quadrato errori . . . . .	105
5.15	Errore aggiunto alla superficie . . . . .	106
5.16	Superficie con quadrato errori . . . . .	108
5.17	Superficie con quadrato errori . . . . .	110
5.18	Superficie ed errore sistematico . . . . .	111
5.19	Densità di probabilità campionamento proposto . . . . .	112

5.20	Campionamento con disegno proposto . . . . .	113
5.21	Campionamento con quadrato latino . . . . .	113
5.22	Campionamento con metodo misto . . . . .	113
5.23	Superficie con quadrato errori . . . . .	115





# Elenco delle tabelle

1.1	Caratteristiche sistema MScMS-II . . . . .	12
2.1	Specifiche tecniche [49] . . . . .	22
4.1	Parametri modello bassa risoluzione . . . . .	53
4.2	Parametri modello alta risoluzione . . . . .	57
4.3	Stime modello alta risoluzione . . . . .	59
4.4	Statistiche puntuali . . . . .	65
4.5	Parametri modello bassa risoluzione . . . . .	71
4.6	Stime modello alta risoluzione decimazione quadratica . . . . .	73
4.7	Statistiche puntuali modello con decimazione quadratica . . . . .	74
5.1	<i>P-value</i> test uguaglianza media a zero . . . . .	104
5.2	<i>P-value</i> confronti a coppie medie quadrati errori . . . . .	105
5.3	<i>P-value</i> test uguaglianza media a zero . . . . .	107
5.4	<i>P-value</i> confronti a coppie medie quadrati errori . . . . .	107
5.5	<i>P-value</i> test uguaglianza media a zero . . . . .	109
5.6	<i>P-value</i> confronti a coppie medie quadrati errori . . . . .	109
5.7	<i>P-value</i> test uguaglianza media a zero . . . . .	114
5.8	<i>P-value</i> confronti a coppie medie quadrati errori . . . . .	114







# Sommario

Nel campo della metrologia un settore di interesse è la misura di superfici di forma complessa. Gli strumenti di misura utilizzati in questo campo possono essere di due tipi: monosensore e multisensore.

Il primo tipo di strumento utilizza un unico dispositivo per acquisire la forma dell'oggetto, mentre gli strumenti che appartengono alla seconda tipologia ne utilizzano almeno due.

L'utilizzo di strumenti multisensore permette di integrare le informazioni derivanti da diversi fonti al fine di sfruttare i vantaggi delle varie tecnologie utilizzate.

Qian e Wu in [35] descrivono tre possibili tipi di integrazione di dati da fondere:

1. i data set registrati da ogni dispositivo hanno le stesse caratteristiche, per cui è quasi impossibile distinguere le varie fonti;
2. i dati provenienti da ogni sensore non possiedono pattern simili, hanno diverse grandezze per cui hanno poco in comune tra loro;
3. ogni data set ha caratteristiche diverse, ma tutti hanno andamenti simili.

Nel primo caso le differenze tra ogni data set può essere trascurata e può essere utilizzato un unico modello con in input i dati provenienti da tutti i sensori. Nel secondo caso è ragionevole affermare che non esiste un metodo per integrare i dati. Sarebbe utile effettuare dei controlli sui sensori utilizzati per la misura o sull'esecuzione dell'esperimento.

L'ultimo caso è quello più comune nell'analisi degli esperimenti ed è quello analizzato nel presente lavoro di tesi. L'analisi viene generalmente eseguita considerando i data set provenienti da ogni sensore separatamente. Questo approccio permette di giungere a risultati migliori rispetto all'esecuzione di analisi distinte.

Nel presente lavoro si presenta un'estensione del modello sviluppato da Qian *et al.* [34] per integrare dati con diversa accuratezza. Gli autori trattano il caso in cui due set di dati provengono da sensori con diverse accuratezze. Un sensore

fornisce dati con elevata accuratezza, ma la misura necessita di un procedimento lento e costoso. Questo è il caso di un esperimento fisico o di una simulazione molto dettagliata. Gli altri dati invece possono provenire da un'altra simulazione che è più veloce e meno costosa della precedente, ma fornisce risultati meno accurati.

Diversamente dai dati a bassa accuratezza quelli più precisi sono disponibili in numero limitato a causa del tempo o del costo necessario per ottenerli. Per questo motivi si propone un metodo per determinare quali punti ad elevata accuratezza acquisire. I fattori che vengono tenuti in considerazione per un buon posizionamento dei punti ad alta accuratezza sono la complessità della superficie rilevata e l'incertezza di previsione ricavabile per mezzo del modello sui punti a bassa accuratezza.

# Capitolo 1

## Strumenti di misura monosensore

Lo scopo del presente lavoro di tesi è quello di utilizzare un modello matematico per descrivere il comportamento di una generica superficie 2,5D. Queste corrispondono alla classe di superfici che possono essere espresse nella forma  $z = f(x, y)$ . In questo primo capitolo sono riportate le caratteristiche dell'attività di misurazione nel campo metrologico e vengono presentati alcuni strumenti di misura monosensore presenti sul mercato. Infine si presenta il caso particolare di metrologia che prende il nome di metrologia di larga scala, che si occupa della misurazione di oggetti di grandi dimensioni, ed un sistema innovativo di misura sviluppato dal Laboratorio di Qualità e Metrologia Industriale del Politecnico di Torino.

### 1.1 Misura

La misurazione (o misura) è l'assegnazione di un intervallo di valori ad una particolare proprietà fisica (chiamata anche misurando). L'attività svolta per assegnare una misura si definisce misurazione.

Con termine misurando non ci si riferisce all'oggetto o al fenomeno su cui si sta eseguendo una misurazione, ma ad una specifica grandezza che caratterizza quest'ultimo. Per esempio, quando si rileva la temperatura di un liquido, il misurando non è il liquido, ma la temperatura del medesimo.

Per questioni sperimentali e teoriche (ad esempio il principio di indeterminazione di Heisenberg) il misurando non è, in realtà, descrivibile da un solo valore numerico, anche ipotizzando una precisione di misurazione infinita.

Ogni misura viene così definita come un intervallo di valori entro cui probabilmente essa è compresa.

Lo sviluppo della metrologia ha portato a definizioni in chiave statistica della definizione dei misurandi, e all'introduzione del concetto di incertezza di misura. Quest'ultimo, in prima approssimazione, può definirsi come la larghezza dell'in-

tervallo di valori: più l'intervallo è grande, maggiore è l'incertezza di misura.

Le caratteristiche di una misura nel campo metrologico sono presenti nel VIM (Vocabolario Internazionale di Metrologia) in cui sono definiti i termini comunemente utilizzati in metrologia. Tra queste si cita:

- **accuratezza:** grado di concordanza tra un valore misurato ed il vero valore, a questo concetto non è associato un valore numerico. Una misura è detta accurata se presenta un piccolo errore di misura, che è dato dalla differenza tra il valore misurato ed il valore di riferimento;
- **giustezza:** grado di concordanza tra la media di un numero infinito di valori misurati ripetutamente ed un valore di riferimento, la giustezza non è una quantità e non può essere espressa da un valore numerico. Una misura è giusta se ha un piccolo errore sistematico di misura, che corrisponde alla componente di errore di misura che rimane costante durante le repliche della misura, o che varia in modo prevedibile;
- **precisione:** grado di concordanza tra le indicazioni oppure tra valori di misura ottenuti misurando ripetutamente lo stesso oggetto o oggetti simili in condizioni specifiche. La precisione di misura è usualmente espressa da valori numerici tramite l'imprecisione di misura, misurata ad esempio dalla varianza, dalla deviazione standard o dal coefficiente di variazione con condizioni di misura controllate.

## 1.2 Strumenti di misura

Dopo una breve presentazione del concetto di misura si presentano gli strumenti di misura, definiti come strumenti con la funzione di effettuare una misurazione. In fisica e ingegneria, una misurazione è il processo di rapportare una grandezza fisica ad eventi, fenomeni od oggetti del mondo reale.

Lo strumento può presentarsi come un singolo oggetto, oppure può essere composto da due o più elementi che, opportunamente combinati, permettono di eseguire la misura (sistema di misura).

Gli strumenti presentati in questo capitolo sono la macchina di misura a coordinate (CMM, *Coordinate Measuring Machine*) e lo scanner a luce strutturata.

### 1.2.1 Coordinate measuring machine

La macchina di misura a coordinate è uno strumento che permette di acquisire le coordinate spaziali appartenenti alle superfici di un manufatto (figura 1.1). Questa macchina può essere controllata manualmente da un operatore, o per mezzo di un computer. Il sistema di movimentazione di una CMM è tipicamente costituito da tre assi,  $X$ ,  $Y$  e  $Z$ . Questi assi sono tra loro ortogonali e formano un sistema di coordinate tridimensionale. Le misure sono definite attraverso un



tastatore (*probe*) presente sul terzo asse della macchina. Il processo più comune per l'acquisizione è a contatto (meccanico), ma ne esistono anche di tipo ottico, laser o mediante la luce strutturata.



Figura 1.1: Macchina di misura a coordinate [1]

Una macchina di misura a coordinate è tipicamente formata dai seguenti dispositivi:

- una struttura principale, che tipicamente consiste in una struttura di granito ed un banco di lavoro;
- un sensore per la misura, chiamato *probe*;
- un sistema elettronico per il controllo e l'elaborazione dei dati.

Le tipologie di misura di una macchina di misura a coordinate sono:

- punto-punto: si rileva un punto alla volta, ed il sensore si stacca dalla superficie dopo ogni rilevazione. Lo svantaggio di questo metodo è che può essere un procedimento molto lento per la rilevazione di ogni punto;
- scansione: il sensore è sempre a contatto con la superficie e la misura avviene in continuo, è una procedura che permette di acquisire i punti più velocemente rispetto alla modalità precedente, ma è anche meno accurata. Attraverso questa tipologia di misura si possono anche scansionare superfici con geometria incognita.

L'applicazione principale di questo strumento di misura è l'ispezioni geometrica, ma viene utilizzato anche per ispezioni dimensionali o per identificazioni di geometrie incognite. Esistono diversi settori in cui le CMM sono utilizzate come ad esempio l'industria automobilistica, meccanica, elettronica ed aeronautica e militare.

Per permettere la buona riuscita dell'operazione di acquisizione della nuvola di punti la struttura meccanica di una macchina di misura a coordinate deve possedere alcune proprietà, tra le quali si cita:

- rigidezza statica e dinamica della struttura;
- accessibilità a tutte le parti del pezzo da misurare;
- facilità di manovra;
- facilità nell'esecuzione dell'operazione di carico/scarico del pezzo.

Una volta acquisiti i punti dell'oggetto fisico si processano i dati eseguendo una compensazione software degli errori, ricavati grazie alla taratura della macchina che è stata eseguita mediante l'acquisizione di un campione di geometria nota.

I possibili fattori di incertezza, che determinano una buona misura, come descritto in [46], possono derivare da:

- hardware CMM;
- pezzo misurato;
- strategia di campionamento;
- algoritmi di ricostruzione della superficie;
- fattori estrinseci come ad esempio l'operatore.

### 1.2.2 Luce strutturata

La luce strutturata (LS) è un sistema che permette di acquisire una nuvola di punti attraverso l'utilizzo di sensori di digitalizzazione di tipo ottico senza contatto (figura 1.2). Questa tecnica sfrutta la proiezione di frange di luce su una superficie da rilevare.



Figura 1.2: Esempio scanner a luce strutturata

I pattern di luce utilizzati possono essere di due tipi:

- reticolo di linee parallele, come in figura 1.3a;

- tecnica a luce codificata (figura 1.3b): consiste in proiezioni di pattern luminosi che permettono di codificare i punti delle immagini mediante l'utilizzo di livelli di grigio o di colore, oppure mediante particolari successioni temporali dei frame.

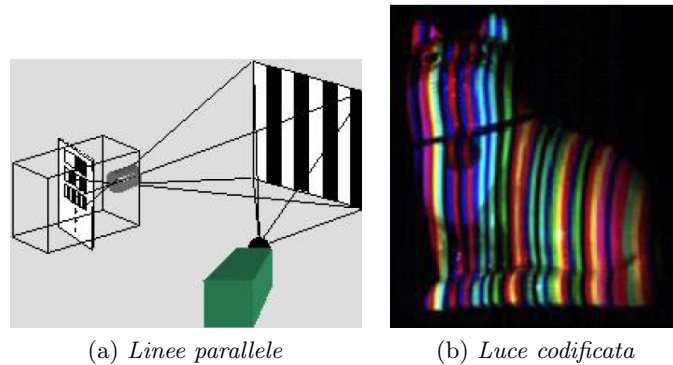


Figura 1.3: Pattern di luce

Prima della digitalizzazione di un oggetto è importante eseguire la calibrazione del sistema di misura. Attraverso questa tecnica è possibile acquisire figure 2D di forma nota per valutare le distorsioni dovute alla distanza tra il sensore e l'oggetto. Lo scopo di questa operazione è quello di stimare gli errori di acquisizione per poi poterli correggere via software.

Il calcolo della distanza del punto dallo scanner è eseguito grazie alla deformazione del pattern di luce proiettato sull'oggetto fisico (figura 1.4); per calcolare l'altezza di un punto si ricorre alla tecnica della triangolazione.

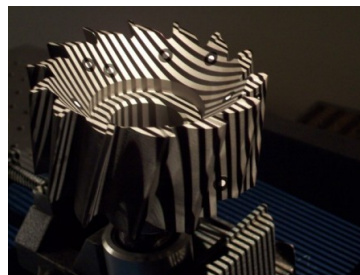


Figura 1.4: Deformazione patter luce

Una breve spiegazione della tecnica di triangolazione può essere fatta osservando la figura 1.5. I calcoli sono eseguiti per determinare le coordinate  $x_h$  e  $z_h$  del punto da rilevare  $H$ . I calcoli da svolgere per ricavare la coordinata  $y_h$  sono simili a quelli riportati di seguito per ricavare la coordinata  $x_h$ .

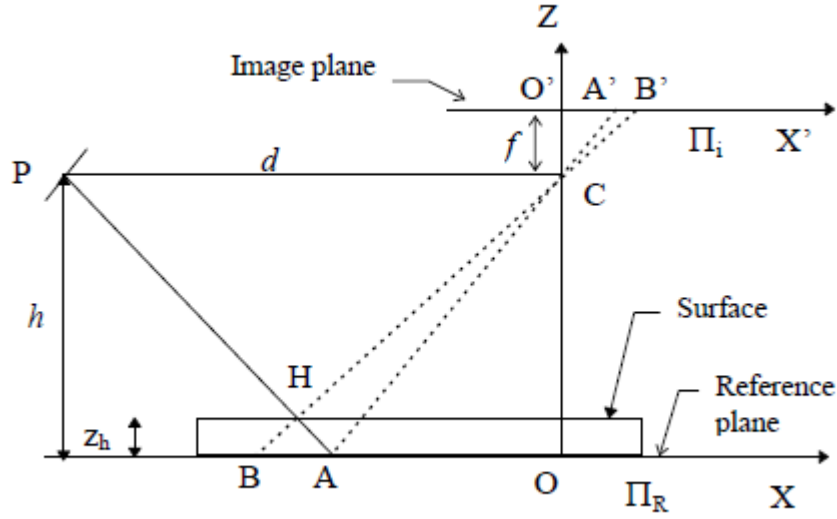


Figura 1.5: Tecnica di triangolazione

Siano  $P$  il punto focale del sistema di proiezione e  $C$  il punto focale del sistema di visione,  $\Pi_R$  un piano di riferimento,  $\Pi_i$  il piano dell'immagine parallelo, in prima approssimazione, a  $\Pi_R$ ,  $O$  e  $O'$  le origini dei sistemi di riferimento che appartengono, rispettivamente, al piano di riferimento ed al piano dell'immagine,  $f$  la distanza focale,  $h$  la distanza tra  $P$  e la sua proiezione ortogonale sul piano  $\Pi_R$  e  $d$  la distanza tra  $P$  e  $C$ . Le quantità sopra definite dipendono dai parametri del sistema di misura e sono quindi note.

Il triangolo  $PHC$  è simile, per costruzione, al triangolo  $AHB$  per cui:

$$\frac{h - z_h}{d} = \frac{z_h}{x_a - x_b} \quad (1.1)$$

per cui

$$z_h = \frac{h(x_a - x_b)}{x_a - x_b + d} \cong \frac{h}{d}(x_a - x_b). \quad (1.2)$$

Considerando invece che il triangolo  $HCH'$  è simile al triangolo  $B'CO'$  si ha che:

$$x_h = \frac{x_{b'}}{d}(h - z_h) = \frac{x_{b'}}{d}h\left(1 - \frac{z_h}{h}\right) \cong \frac{x_{b'}}{d}h. \quad (1.3)$$

La coordinata  $y_h$  è calcolata in modo analogo a  $x_h$  per cui risulta che:

$$y_h = \frac{y_{b'}}{d}(h - z_h) = \frac{y_{b'}}{d}h\left(1 - \frac{z_h}{h}\right) \cong \frac{y_{b'}}{d}h. \quad (1.4)$$

Le principali caratteristiche degli scanner a luce strutturata sono:

- basso costo di acquisto rispetto a strumenti più precisi;
- facile reperibilità sul mercato;
- tecnologia soggetta ad evoluzioni e miglioramenti costanti;
- grande libertà e versatilità nello scegliere il pattern di luce con cui investire l'oggetto.

Rispetto alla macchina di misura a coordinate descritta in precedenza questo sistema di misura effettua delle misure meno accurate, ma il tempo di acquisizione dei punti è molto minore.

### 1.3 Metrologia di larga scala

In molti settori industriali, tra i quali l'automotive, l'aerospaziale, l'industria navale e ferroviaria è emersa la necessità di produrre pezzi con specifiche sempre più stringenti. Come conseguenza vi è la necessità di un sistema di misura per prodotti di grandi dimensioni in molti parti del processo di produzione come l'assemblaggio, l'allineamento ed l'ispezione. La metrologia di larga scala (LSM, *large scale metrology*) è definita in [16] come il campo della metrologia che si occupa del rilevamento della misura di grandi macchine o strutture.

Franceschini, Maisano e Mastrogiacomo in [18] elencano dei requisiti che deve possedere un sistema di LSM che sono:

- portabilità: capacità del sistema di essere facilmente mosso, montato e smontato, per questo sono desiderabili un peso piccolo ed una minima dimensione;
- flessibilità: capacità del sistema di eseguire diversi compiti nel campo della misura (determinazione delle coordinate dei punti, distanze, curve, superfici, ...) ed essere impiegato quando sussistono diverse condizioni ambientali;
- maneggevolezza: facilità di installazione, un rapido set-up e piccoli tempi di calibrazione, prima che il sistema sia operativo; facilità d'uso dell'interfaccia software;
- performance metrologiche: adeguate performance metrologiche in termini di stabilità, ripetibilità, riproducibilità ed accuratezza [ISO 5725 1986];
- scalabilità: la capacità del sistema di misurare diverse forme e volumi con dimensioni lineari fino a 30-60 metri;
- piccolo impatto economico: sono compresi il prezzo di produzione, i costi di installazione, addestramento e manutenzione;

- lavoro *indoor*: il sistema deve poter essere utilizzato in ambienti chiusi, come capannoni o laboratori.

Sono stati sviluppati diversi sistemi di LSM in cui sono state incorporate una moltitudine di tecnologie diverse, come ad esempio ottiche, meccaniche, elettromagnetiche ed acustiche.

Estler *et. al* in [16] affermano che i sensori ottici sono i più efficienti e disponibili nelle applicazioni metrologiche di larga scala, generalmente grazie ai vantaggi nelle ottiche, nella *imagining technology* e nella maggiore velocità di calcolo e precisione.

Galetto, Mastrogiamco e Prali in [19] descrivono diverse classificazioni di ogni sistema di misura in accordo con:

- disposizione dei sensori: centralizzati o distribuiti;
- condizioni di misura: a contatto o non a contatto;
- principi di funzionamento: sistemi che utilizzano due angoli ed una lunghezza, sistemi che utilizzano diversi angoli (triangolazione), sistemi che utilizzano diverse lunghezze (trilaterazione, multilaterazione).

Nei sistemi che utilizzano due angoli ed una lunghezza le coordinate sono espresse in coordinate cilindriche  $(\rho, \varphi, \vartheta)$ . La misura degli angoli è ricavabile grazie ad encoders angolari, mentre il lato è ricavabile grazie ad interferometri.

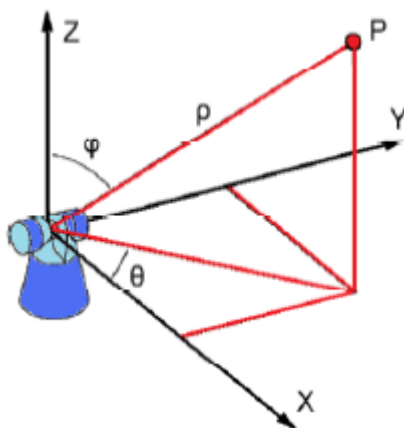


Figura 1.6: Sistema di misura a coordinate sferiche

Al posto di utilizzare due angoli ed un lato è possibile ricavare la posizione di un punto dalle informazioni fornite tramite gli angoli di due o più punti di riferimento. Della triangolazione si riporta solo un'immagine con il principio di funzionamento (figura 1.7) poichè è già stata introdotta in precedenza.

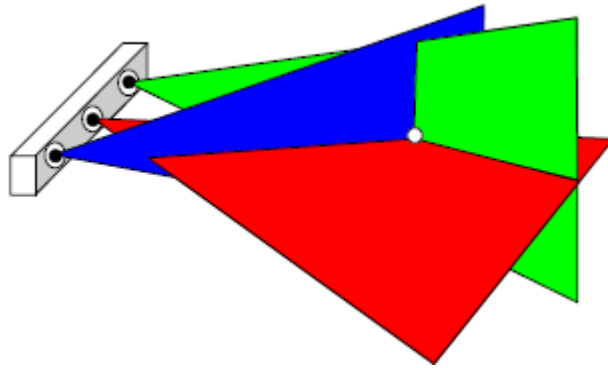


Figura 1.7: Principio di funzionamento triangolazione

La trilaterazione utilizza la locazione nota di tre o più punti di riferimento e misura la distanza da ognuno di essi per determinare la posizione del punto da localizzare (figura 1.8). La localizzazione dei punti può essere eseguita attraverso la soluzione di un sistema non lineare [17]. L'utilizzo di questa procedura provoca dei costi elevati poichè richiede l'utilizzo di almeno tre interferometri, per questo motivo è utilizzata soprattutto nel campo della ricerca.

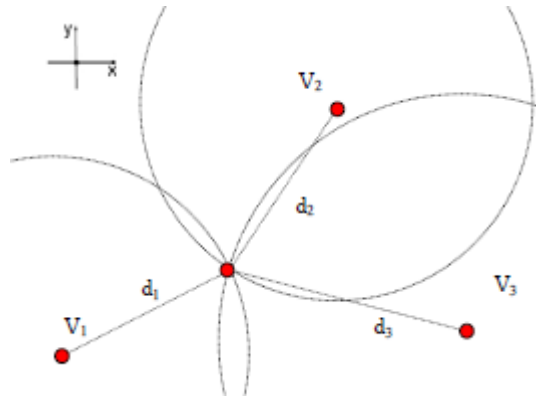


Figura 1.8: Principio di funzionamento trilaterazione

## 1.4 Sistema MScMS-II

Un esempio di sistema di LSM è presente in [19, 20, 21] in cui gli autori descrivono un sistema ad infrarossi (IR, *InfraRed*) con lo scopo di mantenere i costi contenuti, un semplice e rapido rilevamento delle coordinate spaziali di oggetti di grandi dimensioni sfruttando i principi della fotogrammetria. Il sistema è denominato MScMS-II, che è l'acronimo di *Mobile Spatial coordinate Measurement System - II*. Il sistema è stato sviluppato dal Laboratorio di Qualità e Metrologia Industriale del Politecnico di Torino ed è stato presentato per

la prima volta all'ASME *International Manufacturing Science and Engineering Conference*, West Lafayette, IN, nel 2009.

Il sistema è formato da tre elementi principali (figura 1.9): un tastatore wireless, una rete di sensori ed un sistema di elaborazione dei dati che processa i dati derivanti dai sensori. Il sistema è rappresentato in figura 1.9.

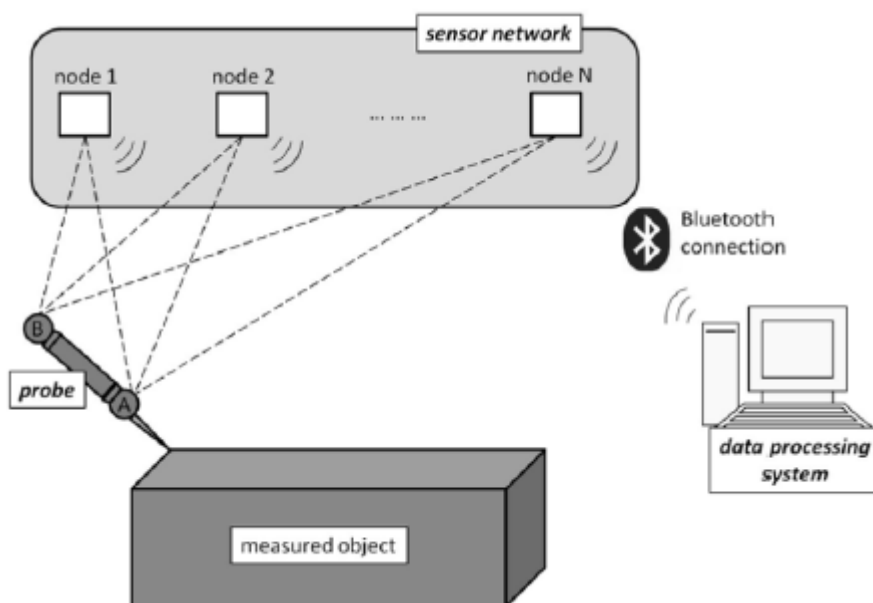


Figura 1.9: Sistema MScMS II, la linea tratteggiata rappresenta la connessione virtuale tra i sensori ed i marker del tastatore, indicati con A e B

Il tastatore mobile wireless (figura 1.10), è dotato di due marker per permettere la corretta identificazione del suo posizionamento nello spazio di misura. La posizione del punto da misurare data dal punto  $\mathbf{x}_V$  ed è possibile ricavarla dalla seguente relazione:

$$\mathbf{x}_V = \mathbf{x}_A + (\mathbf{x}_B - \mathbf{x}_A) \cdot \frac{\|\mathbf{x}_V - \mathbf{x}_A\|}{\|\mathbf{x}_B - \mathbf{x}_A\|} \quad (1.5)$$

dove  $\|\cdot\|$  rappresenta la norma euclidea e  $\mathbf{x}_i = (x_i, y_i, z_i)'$ . La presenza di due punti è necessaria per determinare il punto a contatto con il pezzo  $\mathbf{x}_V$ , poichè nel caso ci fosse un solo marker, con il centro in posizione  $\mathbf{x}_A$ , il punto da misurare potrebbe trovarsi in un qualsiasi punto della sfera centrata in  $\mathbf{x}_A$  con raggio  $\|\mathbf{x}_V - \mathbf{x}_A\|$ .



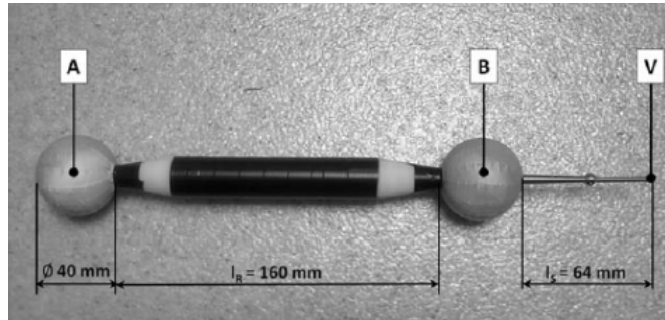


Figura 1.10: Tastatore wireless

La rete, denominata costellazione (figura 1.11), di sensori permette di identificare un corretto posizionamento del tastatore. L'identificazione dei marker sul tastatore è svolta con la tecnica della triangolazione (figura 1.12), dove i punti dei marker sono rappresentati con  $M_1$  e  $M_2$ .

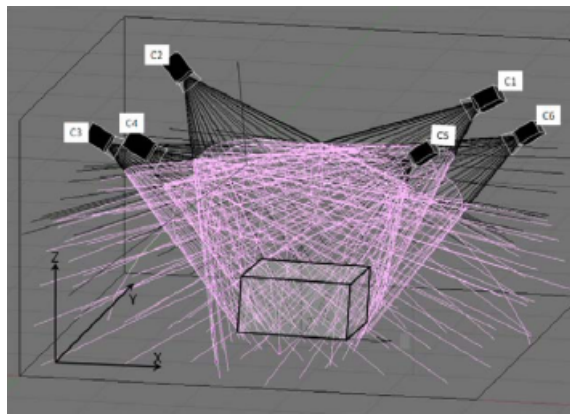


Figura 1.11: Rete di sensori

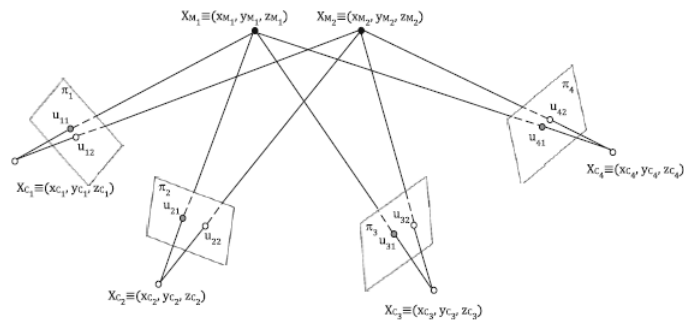


Figura 1.12: Triangolazione per determinare i punti dei marker

I principali fattori che determinano l'incertezza di misura della coordinata 3D rilevata possono essere sintetizzati come segue:

- incertezza delle coordinate 2D, che si riferisce alla proiezione dei punti nelle coordinate 2D dei pixel che formano il piano dell'immagine;
- incertezza della calibrazione della camera, che è associata con i parametri interni ed esterni della camera ottenuti nella fase di calibrazione;
- errore di sincronizzazione della camera, il quale è considerato trascurabile in condizioni statiche, ma va tenuto in considerazione in caso di un tracciamento dei punti dinamico;
- incertezza nelle coordinate 3D dei punti, che può essere ricondotto all'algoritmo di triangolazione per la ricostruzione della nuvola di punti;
- incertezza nelle coordinate della punta del tastatore, che determina l'incertezza dei punti misurati dal sistema MScMS-II

Le caratteristiche di questo sistema di misura sono riassunte in tabella 1.1.

Caratteristica	Descrizione
Scalabilità	Capacità di variare lo spazio misurabile al fine di poter misurare oggetti di elevato volume e di forme complesse.
<i>All around visibility</i>	Capacità del tastatore di raggiungere ogni parte dell'oggetto da qualsiasi direzione.
Connessione wireless	Sensori sono connessi all'unità di elaborazione tramite una connessione Bluetooth.
Ottimizzazione del layout	Capacità del sistema di suggerire automaticamente la posizione ottimale dei sensori in modo da coprire in modo efficiente il volume di lavoro.
Cooperazione tra blocchi di sensori	Permette di ottimizzare l'acquisizione dei punti, il sistema di auto-diagnostica, ed il consumo di energia.
Fusione di sensori	Capacità di integrare il sistema metrologico con altri sensori distribuiti nello spazio per monitorare grandezze diverse (temperatura, umidità, vibrazioni, intensità di luce, ...) al fine di fornire indicazioni delle condizioni ambientali del volume di lavoro e monitorare le condizioni del sistema in fatto di auto-diagnostica ed auto-calibrazione.

Tabella 1.1: Caratteristiche sistema MScMS-II

Il sistema sviluppato è facilmente trasportabile e richiede un tempo di setup limitato grazie alla sua modularità e all'algoritmo di auto-localizzazione della costellazione di sensori che non dipende dall'ambiente in cui si effettua la misura. Grazie alla sua architettura il sistema MScMS-II può essere impiegato per misurare oggetti di grandi dimensioni con forma complessa.



## Capitolo 2

# Multisensor data fusion

In questo capitolo verrà presentata la definizione di fusione di dati con diverse risoluzioni e le sue caratteristiche. Successivamente verranno presentati degli esempi di macchine ibride, cioè dotate di sensori di diversa natura con un sistema di riferimento comune che permettono di ottenere, grazie alla collaborazione, risultati migliori in termini di accuratezza di misura oppure di tempo di processamento, dei singoli strumenti monosensore.

### 2.1 Sensori multirisoluzione

Data la crescente complessità degli oggetti da rilevare e la maggiore accuratezza richiesta, è possibile acquisire superfici *free form* con un'elevata precisione mediante la fusione di sensori con diverse risoluzioni.

La fusione di dati multirisoluzione è definita in [45] come il processo che si occupa della trasformazione di dati provenienti da diverse sorgenti (sensori) in un formato comune al fine di ottenere benefici derivanti dalle informazioni dei vari sensori. L'obiettivo principale della fusione di dati multirisoluzione è il miglioramento della qualità e della sfruttabilità dei risultati provenienti da un processo di misura. Attraverso la sinergia di vari strumenti di misura è possibile ottenere risultati che non sarebbe possibile avere dagli strumenti presi singolarmente, le possibili cause possono essere la limitata accuratezza, i tempi oppure i costi di rilevazione di un numero elevato di dati.

Weckelman *et al.* forniscono dei criteri per la classificazione delle configurazioni basate su più di un sensore. In base alle caratteristiche delle sorgenti di informazioni, i sensori possono essere:

- omogenei: sensori simili, progettati per misurare lo stesso fenomeno fisico, oppure uno comparabile;
- non omogenei: le informazioni acquisite dai sensori non sono direttamente confrontabili, ma richiedono un pre-processamento.

In base ai dati raccolti i sensori possono essere:

- complementari: nel caso in cui i sensori sono tra loro indipendenti, ma i dati provenienti possono essere combinati in modo da avere una rappresentazione migliore dell'oggetto. Un tipico problema che risolvono i sensori complementari è quello dell'incompletezza della misura;
- competitivi: nel caso in cui i sensori sono tra loro indipendenti, ed acquisiscono tutti la stessa immagine. La fusione viene eseguita facendo la media dei vari pixel, in questo caso è simile al caso delle misure ripetute;
- cooperativi: nel caso in cui le misure che provengono da diversi sensori sono utilizzate per derivare informazioni che non sarebbero disponibili dai sensori presi individualmente, come il miglioramento dei dettagli di un oggetto o l'eliminazione delle ambiguità. Un esempio è la macchina di misura a coordinate multisensore che integra diverse tipologie di sensori come a contatto e ottici.

L'ultimo criterio di classificazione è basato sulla metodologia di unione dei dati multirisoluzione:

- fusione per sensori: i sensori misurano la stessa proprietà, ad esempio un sensore a contatto ed uno ottico misurano lo stesso oggetto ottenendo due set di dati distinti che saranno successivamente uniti;
- fusione per attributi: i sensori misurano diverse caratteristiche associate allo stesso ambiente sperimentale;
- fusione per domini: multipli sensori misurano la stessa proprietà su diversi domini del volume di lavoro, un esempio è la fusione di due immagini che raffigurano lo stesso oggetto da due punti di vista diversi e con una certa illuminazione;
- fusione nel tempo: quando nuove misure vengono affiancate a dati storici disponibili, come nel caso della misura dello stesso oggetto eseguita con due calibrazioni diverse.

La fusione dei dati presenta i seguenti passi:

1. pre-processamento;
2. registrazione;
3. ottimizzazione;
4. fusione dei dati;
5. ricostruzione;
6. meshing;

7. conversioni dei dati in altro formato.

I primi quattro passi sono necessari per un'operazione di fusione dei dati, mentre gli ultimi tre sono facoltative.

Il pre-processamento è la prima fase di un processo di fusione dei dati. In questa fase i valori che derivano dal sensore possono subire delle trasformazioni che dipendono dal livello di astrazione ai cui i dati sono combinati. L'integrazione può essere eseguita a livello di segnale, di feature oppure di simbolo.

A livello segnale i dati vengono integrati direttamente.

A livello di feature i dati vengono fusi con lo scopo di avere delle stime migliori di una particolare caratteristica del segnale. I dati possono essere integrati a livello feature quando non c'è coerenza temporale o spaziale dei dati a livello segnale.

A livello di simbolo la classificazione o la rilevazione delle informazioni viene integrata sulla per prendere decisioni basate su funzioni di probabilità.

Un esempio di pre-processamento dei dati del campo delle metrologia è quello della riduzione del dataset quando sono presenti una grande quantità di misurazioni che non determinano un migliore risultato nella ricostruzione della superficie rispetto a un set di dati meno numeroso.

La registrazione è un aspetto critico nella fusione dei dati provenienti da sensori con diverse risoluzioni. Corrisponde all'allineamento e alla trasformazione dei sistemi di coordinate specifici di ogni singolo sensore in uno comune. Tipicamente gli algoritmi di registrazione delle nuvole di punti sono basati sul criterio dei minimi quadrati, cioè la minimizzazione della varianza della distanza tra punti accoppiati. Nei processi di registrazione 3D una pratica comune per individuare queste coppie di punti è l'applicazione di alcuni marcatori sulla superficie.

I metodi per la fusione dei dati servono per decidere quali dati devono essere integrati e come le informazioni ridondanti sono combinate tra loro.

I metodi comunemente utilizzati per la fusione dei dati sono i modelli statistici, i modelli con una logica fuzzy o le reti neurali.

In [45] è presentata una classificazione dei diversi metodi di acquisizione di forme geometriche, riportata in figura 2.3.

Le principali tipologie di sensori sono ottici e a contatto.

I sensori ottici possono acquisire un numero elevato di punti da una superficie con velocità elevata. La possibilità di acquisire i punti senza che ci sia contatto con l'oggetto permette l'acquisizione di superfici flessibili e materiali deformabili. I limiti di questa tecnologia sono la bassa accuratezza, se paragonati ai sensori a contatto, l'occlusione e la regione di acquisizione limitata, così come la grande sensibilità alle condizioni ottiche della superficie. I dati acquisiti attraverso sensori ottici sono spesso affetti da errori ed imperfezioni e possono

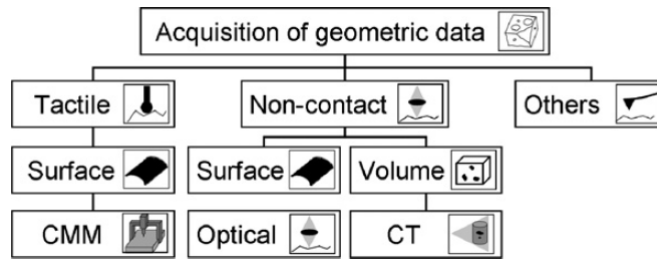


Figura 2.1: Classificazione metodi di acquisizione [45]

mostrare ridondanze in alcune aree e punti mancanti in altre [7]. Nelle misure di grandezze geometriche e dimensionali sono generalmente utilizzati i sensori di distanza, che possono essere classificati in base alla tipologia di acquisizione dei punti, sensori 1D se acquisiscono un punto alla volta, 2D se acquisiscono la superficie tramite delle linee, oppure 3D se acquisiscono un'area [27, 32].

I sensori a contatto sono spesso impiegati nel caso in cui la superficie da misurare permette o richiede l'acquisizione di una nuvola di punti poco densa, come nel caso di pezzi prismatici, superfici di cui è noto il modello CAD o forme di geometria non complessa. Il loro range di misura varia da qualche micrometro a diversi millimetri con un'operazione che può essere eseguita in una, due o tre dimensioni. I sensori a contatto sono generalmente lenti nell'acquisizione dei punti, ma presentano un'elevata accuratezza di misura se paragonati ai sensori ottici. Il principio di funzionamento è basato sull'interazione meccanica tra il sistema di misura e l'oggetto.

## 2.2 Multisensor Coordinate Metrology

La riduzione del *lead time* nel campo del *reverse engineering*, e l'aumento del grado di flessibilità e del livello di automazione del processo di digitalizzazione ha portato allo sviluppo di sistemi che combinano sensori non omogenei come i *probe* meccanici ed i sistemi ottici [31], un esempio è raffigurato in figura 2.2. Il sistema ottico, che può consistere in una videocamera che acquisisce l'intera superficie e fornisce alla CMM le informazioni riguardanti le parti dell'oggetto che dovrà acquisire tramite il *probe* o un altro tipo di sensore [10, 29, 42, 41]. Questi sistemi combinano i vantaggi delle diverse configurazioni, la capacità dei sensori ottici di generare velocemente delle informazioni che riguardano la totalità della superficie, con quella di ottenere delle informazioni molto accurate tipica dei sensori a contatto. I sistemi ibridi sono stati utilizzati inizialmente nel campo del *reverse engineering* e solo ultimamente sono stati utilizzati per scopi di ispezione e controllo.



Una buona integrazione delle informazioni provenienti dai due sensori permette una rapida e precisa ricostruzione di oggetti di forma complessa [8, 39].

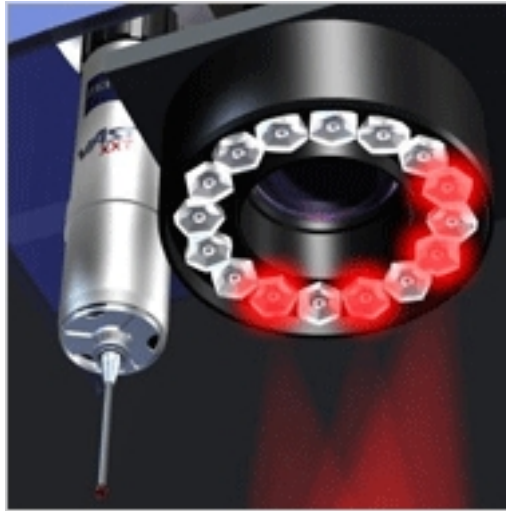


Figura 2.2: Esempio di sistema ottico e meccanico integrato [2]

Una classificazione dei sensori che possono essere montati su un macchina a coordinate in base al principio di funzionamento è presente in figura 2.3.

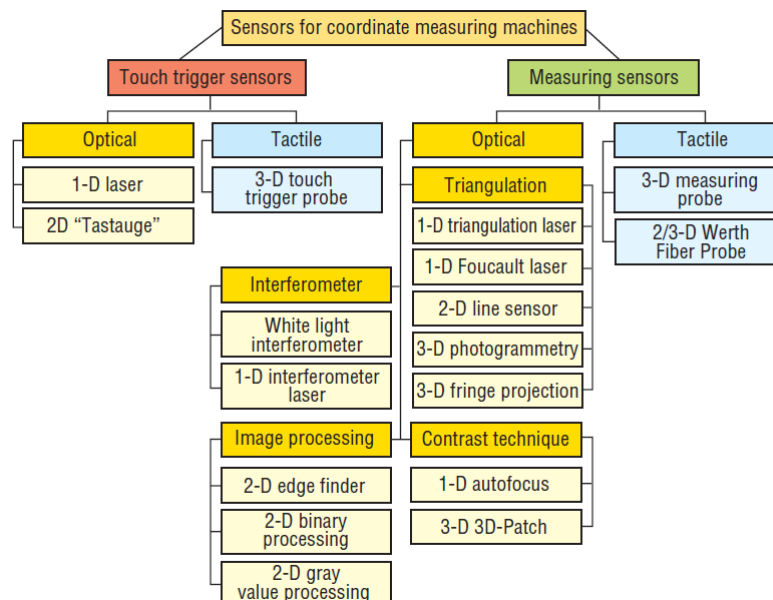


Figura 2.3: Classificazione metodi di acquisizione [11]

I sensori *trigger* producono un segnale dopo che hanno rilevato un punto da misurare, che viene ricevuto dal sistema di misura di ogni asse, i quali permettono di determinare le coordinate di un punto nello spazio. Il movimento degli assi è essenziale per la determinazione delle coordinate del punto rilevato.

I sensori di misura, come quello rappresentato in figura 2.4, hanno un range interno di misura che è dell'ordine di diversi millimetri. A causa di questi ulteriori gradi di libertà un punto è determinato sovrapponendo il valore misurato dal sensore con le coordinate lette dalla macchina di misura. È così possibile determinare un punto anche quando gli assi della macchina di misura a coordinate sono fermi, fino a quando l'oggetto è all'interno del range di misura del sensore.



Figura 2.4: Tastatore con range interno di misura [3]

Un altro importante criterio di differenziazione è il numero di dimensioni che il sensore è in grado di registrare. Questo determina se le informazioni sono riportate in una, due o tre coordinate spaziali.

Nel campo del *reverse engineering*, dove si sono sviluppati questi sistemi multi-sensore, una fusione dei dati provenienti dai sensori può essere svolta partendo dalla nuvola di punti acquisita mediante il sistema di visione, che fornisce delle informazioni sulla forma dell'oggetto. Da questi viene ricostruito il modello CAD del pezzo ed è usato come prima approssimazione del modello finale. La conoscenza, anche se imprecisa, della superficie fornisce una base di partenza per la programmazione della strategia di campionamento dei punti acquisiti con il tastatore meccanico, con una risoluzione del numero di punti ad alte precisione necessari. Queste informazioni vengono poi fuse con quelle del modello CAD preliminare in modo tale da produrre un modello CAD finale accurato. Al posto di sistemi di visione possono anche essere utilizzati degli scanner laser [15, 25, 24]. Lo scanner con sensore laser viene utilizzato per acquisire una nuvola di punti molto densa (figura 2.5), mentre il sensore a contatto della CMM per

definire i contorni dell'oggetto da misurare [5]. Entrambi i sensori sono montati sull'asse  $Z$  della macchina di misura a coordinate, per cui i due sensori hanno un sistema di riferimento comune. Alcune macchine di misura a coordinate hanno la possibilità di un cambio automatico del sensore senza avere la necessità di dover eseguire una calibrazione dello strumento.

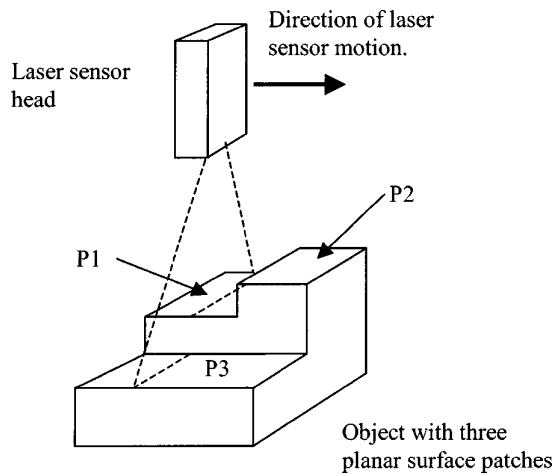


Figura 2.5: Scanner laser [5]

Un esempio è presente in [49] in cui gli autori descrivono il sistema raffigurato in figura 2.6. Nella macchina di misura a coordinate sono presenti due sensori di misura: uno a contatto, mentre il secondo consiste in uno scanner a luce strutturata. Entrambi i sensori sono montati sull'asse  $Z$  della macchina. Il sistema è composto anche da una tavola girevole su cui viene montato il pezzo. Le specifiche tecniche sono riportate nella tabella 2.1.

Il sistema a luce strutturata rappresentato è composto da un proiettore e da due telecamere CCD. Una luce strutturata viene proiettata sull'oggetto da misurare. Le due telecamere ricevono la luce proiettata e l'immagine viene processata per poter determinare la forma 3D dell'oggetto. La scansione dell'intero oggetto da parte della luce strutturata viene eseguita per mezzo della tavola girevole. Per oggetti di forma complessa è necessario ricorrere a viste multiple per poter permettere di scansionare l'oggetto, queste nuvole di punti vengono assemblate per avere la superficie finale.

Il tastatore meccanico viene utilizzato per misurare le posizioni di riferimento, come gli angoli, le feature più importanti e le parti che non è possibile acquisire con lo scanner a luce strutturata.

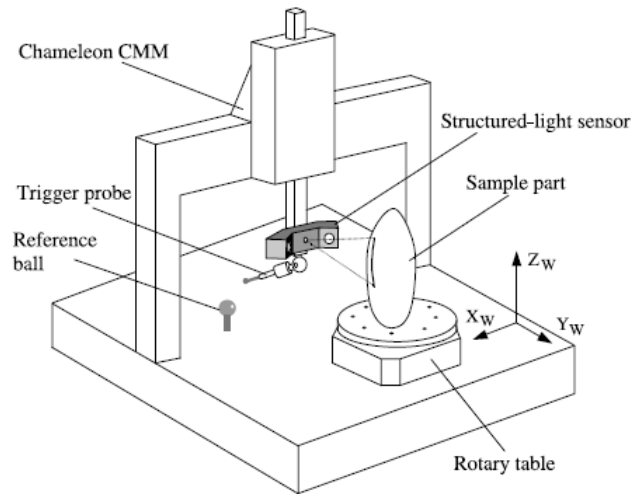


Figura 2.6: Configurazione sistema [49]

Dispositivo	Specifiche
Brown & Sharpe Chameleon CMM	Working volume: $1.500 \times 1.000 \times 800$ mm, MPE: $3 + (3,5L/1000)\mu m$
Renishaw touch trigger probe	PH1 manual probe head, TP6 trigger probe, PEL2 & PEL3 extension bars
Black and white CCD camera	WATEC model WAT-902B, compar 25 mm/F1.8 focus lent
Image acquisition card	EureCard Picolo Pro 2
Rotary table	Model TP21400, positional accuracy: $\pm 10''$
Control card of rotary table	PCL-839 stepper motor card
PC	PIII based AdvanTech Industrial computer, linked to CMM controller

Tabella 2.1: Specifiche tecniche [49]

Nel lavoro gli autori effettuano una fusione dei dati di tipo cooperativo. Viene utilizzato il sensore ottico per acquisire la parte della superficie più regolare, mentre i punti nelle parti interne dei fori e sulle superfici laterali vengono acquisiti con il tastatore meccanico poichè lo scanner a luce strutturata, in quelle zone, commette degli errori che non possono essere trascurati. Il ricorso al tastatore meccanico della CMM è necessario nella parte inferiore delle superfici poichè il pezzo è montato sulla tavola girevole e questo non permette l'acquisizione dei punti per mezzo dello scanner a luce strutturata.

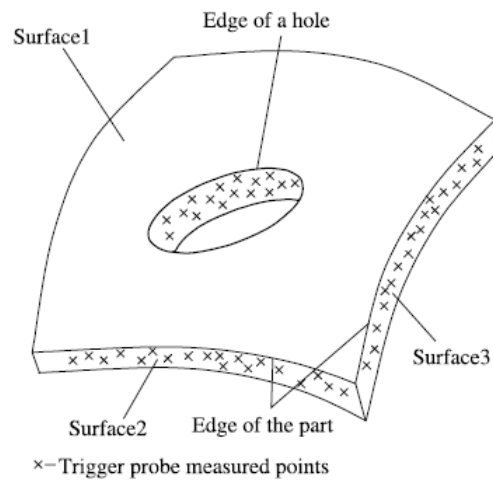


Figura 2.7: Pezzo misurato [49]

Per poter fare una corretta integrazione dei dati i sistemi di riferimento del tastatore meccanico e ottico devono riferirsi ad un sistema di riferimento comune. Per fare questo è necessaria solo un'operazione di traslazione delle nuvole di punti poichè entrambi i sensori sono montati sull'asse  $Z$  della macchina. L'origine del sistema di riferimento è stato scelto dagli autori come il centro della sfera presente in figura 2.6. Grazie a questo punto in comune è possibile eseguire le traslazioni delle due nuvole di punti al fine di integrarle per ottenere la ricostruzione finale dell'oggetto misurato.



## Capitolo 3

# Modello gerarchico

In questo capitolo si presentano i processi gaussiani (PG), utilizzati per la statistica spaziale ([13], [44], [4]), per modellazione di processi stocastici come l'output di simulazioni numeriche ([37], [38]) e per funzioni non definite analiticamente ([33]). Successivamente si presenta l'estensione del modello presente in [34] nel caso in cui la predizione del modello a bassa risoluzione sia affetta da incertezza.

### 3.1 Stato dell'arte

Alcuni autori hanno dimostrato che l'analisi integrata dei dati con diverse scale e risoluzioni porta a risultati migliori rispetto un'analisi separata.

In letteratura sono presenti dei modelli per il trattamento di dati multi-risoluzione che utilizzano i processi gaussiani, la maggior parte di essi hanno un approccio bayesiano.

Xia, Ding, e Mallick, nel loro recente lavoro pubblicato su IIE Transaction [47], forniscono una recensione dettagliata sulle metodologie sviluppate per integrare dati provenienti da simulazione deterministiche con diverse precisioni o la combinazione tra simulazioni e dati provenienti da esperimenti fisici. Si distinguono due principali scuole di pensiero in materia.

Gli autori citano il lavoro da Reese *et al.* [36] in cui vengono presi in considerazione insieme di dati osservati da tre fonti distinte: *computer experiment*, esperimento fisico e opinioni di esperti del settore. Come primo passo, viene costruito un modello appropriato per i diversi dati. Successivamente vengono combinate le informazioni provenienti dalle tre fonti, infine vengono integrati i dati mediante una metodologia flessibile che considera gli errori nei dati delle diverse fonti attraverso un modello ricorsivo gerarchico bayesiano.

Nel secondo approccio identificato da Xia *et al.*, viene inizialmente costruito un modello, per i soli dati a bassa risoluzione, e viene sviluppato, in seguito,

un modello per i dati ad alta risoluzione che utilizzano i valori del modello a risoluzione minore come variabili di input. Quest'ultimo modello è spesso chiamato modello di collegamento poichè collega i dati a bassa ed alta risoluzione eseguendo una trasformazione di scala e un cambiamento di posizione.

Kennedy e O'Hagan [28] implementano un modello di collegamento attraverso un processo gaussiano tra dati provenienti da una simulazione e quelli di un esperimento fisico. Questo processo, noto come calibrazione, è implementato in un contesto bayesiano. Il lavoro di Higdon *et al.* [23] è molto simile al precedente. Anche in questo caso è stato utilizzato un approccio bayesiano unito ad un processo gaussiano con lo scopo di calibrare i parametri di un simulatore utilizzando i dati provenienti da un esperimento fisico. Gli autori sottolineano l'importanza della modellistica che tiene conto della totalità dell'incertezza mediante un modello bayesiano.

Un'ulteriore lavoro in cui viene sviluppato un modello gerarchico è presente nell'articolo di Qian, Wu, Seepersad, Joseph e Allen [34]. In questo caso gli autori non utilizzano approccio bayesiano. Mentre nel lavoro successivo [35], Qian e Wu sviluppano un modello gerarchico che sfrutta i processi gaussiani ed un approccio bayesiano.

Nei due articoli gli autori utilizzano un modello con processo gaussiano per stimare i dati a bassa risoluzione. Poi, nel modello di collegamento sono presenti dei parametri di scala ed uno di shift; questo utilizza informazioni provenienti da un piccolo insieme di dati ad alta risoluzione, che vengono usati per migliorare l'accuratezza della previsione. In [34] il cambiamento di scala è modellato tramite una regressione lineare; in [35] è presente l'utilizzo di un processo gaussiano anche per il parametro di scala permettendo di prendere in considerazione cambiamenti più complessi tra i modelli a bassa ed alta precisione. Nel presente lavoro di tesi si segue l'approccio di [34] e non si sviluppa un approccio bayesiano.

Tornando al lavoro di Xia, Ding e Mallick gli autori forniscono un'applicazione del modello presente in [35] nel campo della metrologia, lo stesso ambito del presente lavoro di tesi. Gli autori utilizzano dei dati ad alta risoluzione provenienti da una macchina di misura a coordinate dotata di tastatore meccanico, mentre i punti a bassa risoluzione sono stati acquisiti con una CMM meno precisa dotata di sensore ottico (OCMM *Optical*CMM). Essi inoltre utilizzano un modello con processo gaussiano che è più adatto a modellare i dati provenienti da un'acquisizione mediante una macchina di misura a coordinate rispetto al kriging universale [48] poichè commette migliori stime degli errori di forma. L'articolo di Xia *et al.* focalizza l'attenzione sulla registrazione delle nuvole di punti provenienti dalle due macchine, nel caso queste non abbiano un sistema di riferimento comune.



### 3.2 Processo gaussiano

Un processo gaussiano è definito come un processo stocastico in cui prendendo un qualsiasi numero finito di variabili aleatorie, appartenenti al processo, hanno una distribuzione congiunta gaussiana.

Per questo motivo un processo gaussiano è completamente definito se sono definiti i suoi primi due momenti, per cui anche se sono definite la media e la varianza. Si definisce la media  $\mu(\mathbf{x}_1)$  e la covarianza  $C(\mathbf{x}_1, \mathbf{x}_2)$  come:

$$\begin{aligned}\mu(\mathbf{x}_1) &= \mathbb{E}[f(\mathbf{x}_1)] \\ C(\mathbf{x}_1, \mathbf{x}_2) &= \Sigma = \mathbb{E}\{[f(\mathbf{x}_1) - \mu(\mathbf{x}_1)][f(\mathbf{x}_2) - \mu(\mathbf{x}_2)]\}\end{aligned}\quad (3.1)$$

Si considera il caso in cui la matrice di varianza covarianza,  $\Sigma$ , sia definita positiva, cioè la variabile aleatoria associata non sia degenerare lungo nessuna direzione. Questa proprietà porta a vantaggi nel calcolo della predizione dati i punti sperimentali [38].

Si focalizza l'attenzione sui processi gaussiani strettamente stazionari poiché solitamente i processi spaziali non presentano derivate descritte dal processo. Inoltre i processi non stazionari sono di difficile utilizzo nel caso di dataset di grandi dimensioni. Sia  $Y(\mathbf{x}_i)$ ,  $\mathbf{x}_i \in \mathbb{R}^d$  un PG, esso è detto strettamente stazionario se, per ogni  $\mathbf{h} \in \mathbb{R}^d$  ed ogni  $\mathbf{x}_1, \dots, \mathbf{x}_n \in \mathbf{X}$ , con  $\mathbf{x}_1 + \mathbf{h}, \dots, \mathbf{x}_n + \mathbf{h} \in \mathbf{X}$ , allora  $(Y(\mathbf{x}_1), \dots, Y(\mathbf{x}_k))'$  e  $(Y(\mathbf{x}_1 + \mathbf{h}), \dots, Y(\mathbf{x}_k + \mathbf{h}))'$  hanno la stessa distribuzione. Questo implica che i vettori aleatori abbiano la stessa media e varianza. Un processo gaussiano stazionario è una rappresentazione realistica della realtà perché la correlazione tra i vari punti non dipende dallo specifico punto dello spazio di misura.

È possibile dimostrare che la covarianza di un processo gaussiano stazionario deve soddisfare la (3.2).

$$Cov(Y(\mathbf{x}_1), Y(\mathbf{x}_2)) = C(\mathbf{x}_1 - \mathbf{x}_2) \quad (3.2)$$

Questo significa che la covarianza dipende dalle posizioni relative dei punti nello spazio, è invariante rispetto alle traslazioni.

Spesso la covarianza di un PG è definita dalla varianza del processo  $\sigma_Y^2$  e dalla funzione di correlazione definita come:

$$R(\mathbf{x}_1, \mathbf{x}_2) = \frac{C(\mathbf{x}_1, \mathbf{x}_2)}{\sigma_Y^2}. \quad (3.3)$$

La funzione di correlazione utilizzata in questo lavoro è il kernel gaussiano anisotropo:

$$R(\mathbf{x}_1, \mathbf{x}_2 | \boldsymbol{\vartheta}) = \exp \left\{ - \sum_{i=1}^d \vartheta_i (x_{1i} - x_{2i})^2 \right\}. \quad (3.4)$$

Il kernel rappresenta la funzione di correlazione che esprime la relazione esistente da due punti nello spazio, cioè la tendenza a variare di un punto quando varia un altro. Questo permette, in fase di predizione, di conoscere la posizione di un punto nello spazio conoscendo quella dei punti in un suo intorno.

Si dimostra che la matrice di correlazione i cui elementi sono quelli descritti nella (3.4) è definita positiva [30] se il kernel gaussiano è applicato in punti diversi. In figura 3.1 sono riportati due esempi di kernel gaussiani, nel grafico di sinistra è rappresentata la funzione con  $d = 2$  e  $\boldsymbol{\vartheta} = (0,05 \ 0,05)'$ , mentre in quello di destra  $\boldsymbol{\vartheta} = (0,05 \ 0,2)'$ . La differenza tra le due è la velocità con cui la correlazione tende a zero lungo le due direzioni, questa è data dai parametri  $\vartheta_1$  e  $\vartheta_2$ .

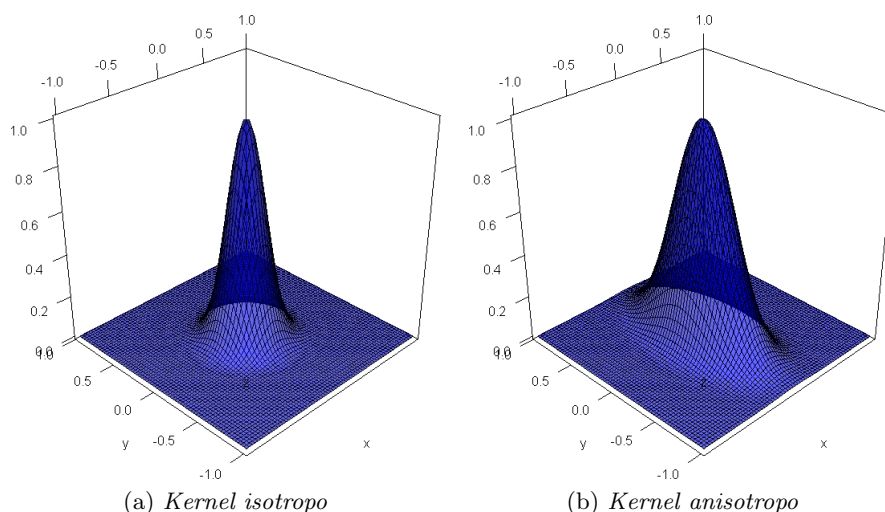


Figura 3.1: Esempio kernel gaussiano

### 3.3 Processi gaussiani per dati multirisoluzione

Si illustrano i due modelli di processo gaussiano sviluppati da Qian e Wu in [34] e [35]. Come precedentemente descritto, gli autori considerano la situazione in cui due tipi esperimenti forniscono dati con diverse risoluzioni. Vi è un esperimento a bassa accuratezza (LE), veloce da eseguire, ma approssimativo, e un esperimento dotato di alta precisione (HE), preciso ma costoso. Non si riporta il modello presente in [47] poichè, anche se il campo di applicazione del modello è lo stesso del presente lavoro, l'attenzione degli autori si è focalizzata sulla registrazione, mediante un approccio bayesiano, delle nuvole di punti derivanti da due sistemi di misura senza un sistema di riferimento comune.

Si ipotizza che LE e HE abbiano in comune un insieme di  $k$  fattori  $\mathbf{x} = (x_1, x_2, \dots, x_k)'$ . Le variabili della sperimentazione dei parametri a bassa accu-

ratezza viene indicato con  $\mathbf{D}_l = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ . I corrispondenti dati a bassa risoluzione sono indicati con  $\mathbf{y}_l = (y_l(\mathbf{x}_1), y_l(\mathbf{x}_2), \dots, y_l(\mathbf{x}_n))$ . Il piano sperimentale per HE viene indicato con  $\mathbf{D}_h = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{n_1}\}$ , mentre le corrispondenti risposte ad alta risoluzione sono indicate con  $\mathbf{y}_h = (y_h(\mathbf{x}_1), y_h(\mathbf{x}_2), \dots, y_h(\mathbf{x}_{n_1}))'$ . Si parte dal presupposto che il numero dati LE disponibili è maggiore del numero di quelli ad alta precisione ( $n > n_1$ ), dal momento che i dati richiedono l'utilizzo di più risorse per essere calcolati. Così gli autori danno per scontato che  $D_h \subset D_L$  senza perdita di generalità. Lo scopo principale di questi modelli è la previsione della risposta ad alta risoluzione in punto dove non è stata eseguita nessuna prova sperimentale.

### 3.3.1 Modello a bassa risoluzione

Gli autori considerano il caso in cui i dati LE provengano da una simulazione deterministica. Così è possibile applicare le tecniche di modellazione descritte nella sezione precedente. Il modello per la risposta LE è:

$$y_l(\mathbf{x}_i) = \mathbf{f}'_l(\mathbf{x}_i)\boldsymbol{\beta}_l + \varepsilon_l(\mathbf{x}_i), \quad \forall i = 1, \dots, n \quad (3.5)$$

dove  $\mathbf{f}_l(\mathbf{x}_i) = (f_0(\mathbf{x}_i), f_1(\mathbf{x}_i), \dots, f_k(\mathbf{x}_i))'$  è un vettore di funzioni note e  $\boldsymbol{\beta}_l = (\beta_{l_0}, \beta_{l_1}, \dots, \beta_{l_k})'$  è un vettore di coefficienti di regressione ignoti.  $\varepsilon_l(\cdot)$  si presume essere un processo gaussiano con media nulla. La sua funzione di covarianza dipende dalla varianza del processo,  $\sigma_l^2$ , e da un vettore di parametri non noti  $\boldsymbol{\vartheta}_l$ . Questi parametri sono i parametri di scala che appaiono nella funzione di correlazione gaussiana.

$$R_l(\mathbf{x}_j, \mathbf{x}_m) = \exp \left\{ - \sum_{i=1}^k \vartheta_{l_i} (x_{ji} - x_{mi})^2 \right\}, \quad (j, m) = 1, \dots, n \quad (3.6)$$

dove  $\vartheta_{l_i} > 0 \quad \forall i = 1, \dots, k$ . Questa particolare classe di funzioni di correlazione produce campioni del corrispondente PG che sono infinitamente differenziabili. Dato che  $R(\cdot, \cdot)$  appartiene alla classe delle funzioni di correlazione gaussiane, il PG  $\varepsilon_l(\cdot)$  è completamente definito dalla sua media,  $\mu$ , che corrisponde a zero, la sua varianza  $\sigma_l^2$  ed i suoi parametri di correlazione  $\boldsymbol{\vartheta}_l$ . Ne consegue che  $y_l(\mathbf{x}_i) \sim \mathcal{PG}(\mathbf{f}'_l(\mathbf{x}_i)\boldsymbol{\beta}_l, \sigma_l^2, \boldsymbol{\vartheta}_l)$ .

In [34] e [35] si introduce l'ipotesi che i fattori considerati nella sperimentazione hanno un effetto lineare, cioè  $\mathbf{f}_l(\mathbf{x}_i) = (1, x_{i1}, x_{i2}, \dots, x_{in})'$ . Inoltre, essi affermano che l'inclusione di un modello semplice nella media del processo gaussiano può portare a vantaggi numerici nella stima dei parametri della correlazione  $\boldsymbol{\vartheta}_l$ . Questa affermazione è confermata dai risultati numerici degli esempi presenti nel loro articolo.

### 3.3.2 Modello ad alta risoluzione

Siccome i punti LE non sono molto accurati e sono disponibili punti HE è conveniente integrare i due dataset per migliorare la qualità di predizione del

modello. Qian e Wu propongono due modelli di collegamento il primo è presentato in [34], mentre il secondo in [35]. Nel seguito si riporta solo il primo modello di collegamento poichè il secondo utilizza un approccio bayesiano che non è presente in questo lavoro di tesi.

### Modello di collegamento

In [34] il modello di connessione tra i dati con diverse precisioni è:

$$y_h(\mathbf{x}_i) = \rho(\mathbf{x}_i)y_l(\mathbf{x}_i) + \delta(\mathbf{x}_i), \quad \forall i = 1, 2, \dots, n_1 \quad (3.7)$$

in cui il parametro di scala  $\rho(\cdot)$  è una funzione lineare:

$$\rho(\mathbf{x}_i) = \rho_0 + \sum_{j=1}^k \rho_k x_{ij}, \quad i = 1, 2, \dots, n_1 \quad (3.8)$$

e  $\delta(\cdot) \sim \mathcal{PG}(\delta_0, \sigma_\delta^2, \boldsymbol{\vartheta}_\delta)$ , cioè un processo gaussiano con media  $\delta_0$ , varianza  $\sigma_\delta^2$  e parametri di correlazione  $\boldsymbol{\vartheta}_\delta$ . Anche in questo modello viene utilizzata la funzione di correlazione gaussiana:

$$R_\delta(\mathbf{x}_j, \mathbf{x}_m) = \exp \left\{ - \sum_{i=1}^k \vartheta_{\delta_i} (x_{ji} - x_{mi})^2 \right\}, \quad (j, m) = 1, \dots, n_1 \quad (3.9)$$

## 3.4 Modello bassa risoluzione

### 3.4.1 Introduzione

In questa prima parte si introduce un modello, (3.10), per descrivere il comportamento degli  $n$  punti acquisiti mediante uno strumento con una bassa accuratezza come in [34], con la differenza che è presente l'errore casuale poichè nel caso preso in considerazione i punti a bassa accuratezza derivano da misure di superfici e sono soggette ad errori. Non si utilizza il modello presente in [35] poichè l'approccio bayesiano è lento se paragonato a quello frequentista, e per un numero molto elevato di punti, come nel caso della metrologia, i tempi richiesti per la stima dei parametri diventano inaccettabili. Sebbene il modello con un approccio bayesiano sia più flessibile è stato dimostrato in [14] che non porta a grandi vantaggi in termini di errore di predizione.

$$\mathbf{y}_l = \mathbf{Z}_l \boldsymbol{\beta} + \boldsymbol{\eta} + \boldsymbol{\varepsilon}_l \quad (3.10)$$

Le componenti sono:

- $\mathbf{y}_l \in \mathbb{R}^n$ : vettore i cui elementi rappresentano i punti a bassa accuratezza;

- $\mathbf{Z}_l \in (n, p)$ :  $\begin{bmatrix} \mathbf{z}'_1 \\ \vdots \\ \mathbf{z}'_n \end{bmatrix}$  è la matrice disegno del modello, quindi  $\mathbf{z}_i = \mathbf{f}(\mathbf{x}_i)$ .
- $\boldsymbol{\beta} \in \mathbb{R}^p$ : vettore dei regressori;
- $\boldsymbol{\eta} \sim \mathcal{PG}_n(\mathbf{0}_n, \sigma_\eta^2 \mathbf{R}_\eta)$ ,  $\boldsymbol{\eta} \in \mathbb{R}^n$ : processo gaussiano di media  $\mathbf{0}_n$ , che rappresenta il vettore nullo di dimensione  $n$ , mentre  $\sigma_\eta^2 \mathbf{R}_\eta$  è la matrice di varianza covarianza, con  $\mathbf{R}_\eta$  matrice correlazione definita successivamente nella (3.11).
- $\boldsymbol{\varepsilon}_l \sim \mathcal{N}_n(\mathbf{0}_n, \sigma_{\varepsilon_l}^2 \mathbf{I}_n)$ ,  $\boldsymbol{\varepsilon}_l \in \mathbb{R}^n$ : errore casuale ed indipendente dalla collocazione del punto nello spazio.

La matrice  $\mathbf{R}_\eta \in (n, n)$  è una matrice di correlazione che rappresenta il kernel gaussiano anisotropo ed è così definita:

$$\mathbf{R}_\eta = [r_{\eta_{ij}}], \quad r_{\eta_{ij}} = \exp \left\{ - \left[ \vartheta_1^l (x_{i1} - x_{j1})^2 + \vartheta_2^l (x_{i2} - x_{j2})^2 \right] \right\} \quad (3.11)$$

Viene considerato solo il caso in cui  $\mathbf{X}_l \in (n, 2)$  poichè i punti oggetto di questa analisi provengono da misure di oggetti fisici. Il passaggio ad una dimensione maggiore è però immediato. Facendo un'ulteriore ipotesi, cioè l'indipendenza tra il processo gaussiano  $\boldsymbol{\eta}(\mathbf{x}_i)$  e l'errore casuale  $\boldsymbol{\varepsilon}_l(\mathbf{x}_i)$  si può giungere alla seguente conclusione:

$$\mathbf{y}_l \sim \mathcal{N}_n(\mathbf{Z}_l \boldsymbol{\beta}, \boldsymbol{\Sigma}_l)$$

dove  $\boldsymbol{\Sigma}_l = \sigma_{\varepsilon_l}^2 \mathbf{I}_n + \sigma_\eta^2 \mathbf{R}_\eta$  rappresenta la matrice di varianza covarianza dei punti a bassa accuratezza.

### 3.4.2 Stima dei parametri

Il metodo utilizzato per la stima dei parametri del modello è la massima verosimiglianza [38] ed, essendo  $\mathbf{y}_l$  un processo gaussiano, la verosimiglianza risulta:

$$L = \frac{1}{(2\pi)^{\frac{n}{2}} |\boldsymbol{\Sigma}_l|^{\frac{1}{2}}} \exp \left\{ - \frac{(\mathbf{y}_l - \mathbf{Z}_l \boldsymbol{\beta})' \boldsymbol{\Sigma}_l^{-1} (\mathbf{y}_l - \mathbf{Z}_l \boldsymbol{\beta})}{2} \right\} \quad (3.12)$$

Si calcola ora il logaritmo di  $L$ , indicato con  $l$ , per permettere una sua più facile massimizzazione.

$$l = -\frac{n}{2} \log 2\pi - \frac{1}{2} \log |\boldsymbol{\Sigma}_l| - \frac{(\mathbf{y}_l - \mathbf{Z}_l \boldsymbol{\beta})' \boldsymbol{\Sigma}_l^{-1} (\mathbf{y}_l - \mathbf{Z}_l \boldsymbol{\beta})}{2} \quad (3.13)$$

Si trova il gradiente di  $l$  per verificare se è possibile trovare analiticamente i punti stazionari.

$$\begin{aligned}
\frac{\partial l}{\partial \boldsymbol{\beta}} &= \mathbf{Z}'_l \boldsymbol{\Sigma}_l^{-1} (\mathbf{y}_l - \mathbf{Z}_l \boldsymbol{\beta}) \\
\frac{\partial l}{\partial \sigma_\eta^2} &= -\frac{1}{2} \text{tr} (\boldsymbol{\Sigma}_l^{-1} \mathbf{R}_\eta) + \frac{1}{2} (\mathbf{y}_l - \mathbf{Z}_l \boldsymbol{\beta})' \boldsymbol{\Sigma}_l^{-1} \mathbf{R}_\eta \boldsymbol{\Sigma}_l^{-1} (\mathbf{y}_l - \mathbf{Z}_l \boldsymbol{\beta}) \\
\frac{\partial l}{\partial \sigma_{\varepsilon_i}^2} &= -\frac{1}{2} \text{tr} (\boldsymbol{\Sigma}_l^{-1}) + \frac{1}{2} (\mathbf{y}_l - \mathbf{Z}_l \boldsymbol{\beta})' \boldsymbol{\Sigma}_l^{-2} (\mathbf{y}_l - \mathbf{Z}_l \boldsymbol{\beta}) \\
\frac{\partial l}{\partial \vartheta_1^l} &= -\frac{\sigma_\eta^2}{2} \text{tr} \left( \boldsymbol{\Sigma}_l^{-1} \frac{\partial \mathbf{R}_\eta}{\partial \vartheta_1^l} \right) + \frac{\sigma_\eta^2}{2} (\mathbf{y}_l - \mathbf{Z}_l \boldsymbol{\beta})' \boldsymbol{\Sigma}_l^{-1} \frac{\partial \mathbf{R}_\eta}{\partial \vartheta_1^l} \boldsymbol{\Sigma}_l^{-1} (\mathbf{y}_l - \mathbf{Z}_l \boldsymbol{\beta}) \\
\frac{\partial l}{\partial \vartheta_2^l} &= -\frac{\sigma_\eta^2}{2} \text{tr} \left( \boldsymbol{\Sigma}_l^{-1} \frac{\partial \mathbf{R}_\eta}{\partial \vartheta_2^l} \right) + \frac{\sigma_\eta^2}{2} (\mathbf{y}_l - \mathbf{Z}_l \boldsymbol{\beta})' \boldsymbol{\Sigma}_l^{-1} \frac{\partial \mathbf{R}_\eta}{\partial \vartheta_2^l} \boldsymbol{\Sigma}_l^{-1} (\mathbf{y}_l - \mathbf{Z}_l \boldsymbol{\beta})
\end{aligned} \tag{3.14}$$

dove

$$\frac{\partial \mathbf{R}_\eta}{\partial \vartheta_i^l} = [r_{\eta_{jk}}^{\vartheta_i^l}], \quad r_{\eta_{jk}}^{\vartheta_i^l} = -(x_{ji} - x_{ki})^2 r_{\eta_{jk}}$$

Uguagliando  $\nabla l$  al vettore nullo è possibile ricavare solo  $\widehat{\boldsymbol{\beta}}$  in forma chiusa.

$$\widehat{\boldsymbol{\beta}} = (\mathbf{Z}'_l \boldsymbol{\Sigma}_l^{-1} \mathbf{Z}_l)^{-1} \mathbf{Z}'_l \boldsymbol{\Sigma}_l^{-1} \mathbf{y}_l \tag{3.15}$$

Lo stimatore, dato  $\boldsymbol{\Sigma}_l$ , è corretto.

*Dimostrazione.*

$$\mathbb{E}(\widehat{\boldsymbol{\beta}} | \boldsymbol{\Sigma}_l) = (\mathbf{Z}'_l \boldsymbol{\Sigma}_l^{-1} \mathbf{Z}_l)^{-1} \mathbf{Z}'_l \boldsymbol{\Sigma}_l^{-1} \mathbb{E}(\mathbf{y}_l) = (\mathbf{Z}'_l \boldsymbol{\Sigma}_l^{-1} \mathbf{Z}_l)^{-1} \mathbf{Z}'_l \boldsymbol{\Sigma}_l^{-1} \mathbf{Z}_l \boldsymbol{\beta} = \boldsymbol{\beta} \tag{3.16}$$

□

Si ricorda che  $\mathbf{y}_l$  è una normale multivariata e  $\widehat{\boldsymbol{\beta}} | \boldsymbol{\Sigma}_l$  è una sua combinazione lineare, per cui hanno la stessa distribuzione di probabilità, si calcola allora la matrice di varianza covarianza dello stimatore.

$$\begin{aligned}
\text{Cov}(\widehat{\boldsymbol{\beta}} | \boldsymbol{\Sigma}_l) &= (\mathbf{Z}'_l \boldsymbol{\Sigma}_l^{-1} \mathbf{Z}_l)^{-1} \mathbf{Z}'_l \boldsymbol{\Sigma}_l^{-1} \text{Cov}(\mathbf{y}_l) \boldsymbol{\Sigma}_l^{-1} \mathbf{Z}_l (\mathbf{Z}'_l \boldsymbol{\Sigma}_l^{-1} \mathbf{Z}_l)^{-1} = \\
&= (\mathbf{Z}'_l \boldsymbol{\Sigma}_l^{-1} \mathbf{Z}_l)^{-1} \mathbf{Z}'_l \boldsymbol{\Sigma}_l^{-1} \boldsymbol{\Sigma}_l \boldsymbol{\Sigma}_l^{-1} \mathbf{Z}_l (\mathbf{Z}'_l \boldsymbol{\Sigma}_l^{-1} \mathbf{Z}_l)^{-1} = \\
&= (\mathbf{Z}'_l \boldsymbol{\Sigma}_l^{-1} \mathbf{Z}_l)^{-1}
\end{aligned} \tag{3.17}$$

Per i risultati precedenti, ed essendo  $\widehat{\boldsymbol{\beta}}$  uno stimatore di massima verosimiglianza, si ha che asintoticamente

$$\widehat{\boldsymbol{\beta}} | \boldsymbol{\Sigma}_l \approx \mathcal{N}_p \left( \boldsymbol{\beta}, (\mathbf{Z}'_l \boldsymbol{\Sigma}_l^{-1} \mathbf{Z}_l)^{-1} \right) \tag{3.18}$$

Avendo ottenuto  $\widehat{\boldsymbol{\beta}}$ , ed essendo unico, se le altre variabili da ottimizzare hanno un solo punto stazionario, si può inserire il suo valore nella funzione  $l$  che

sarà massimizzata per ottenere gli stimatori di massima verosimiglianza degli altri parametri. La nuova funzione  $l_{\hat{\beta}}$  prende il nome di *profile likelihood* [38].

$$\begin{aligned} l_{\hat{\beta}} &= -\frac{n}{2} \log 2\pi - \frac{1}{2} \log |\boldsymbol{\Sigma}_l| - \frac{1}{2} \left[ \mathbf{y}_l - \mathbf{Z}_l (\mathbf{Z}'_l \boldsymbol{\Sigma}_l^{-1} \mathbf{Z}_l)^{-1} \mathbf{Z}'_l \boldsymbol{\Sigma}_l^{-1} \mathbf{y}_l \right]' \\ &\quad \cdot \boldsymbol{\Sigma}_l^{-1} \left[ \mathbf{y}_l - \mathbf{Z}_l (\mathbf{Z}'_l \boldsymbol{\Sigma}_l^{-1} \mathbf{Z}_l)^{-1} \mathbf{Z}'_l \boldsymbol{\Sigma}_l^{-1} \mathbf{y}_l \right] = \\ &= -\frac{n}{2} \log 2\pi - \frac{1}{2} \log |\boldsymbol{\Sigma}_l| - \frac{1}{2} [(\mathbf{I}_n - \mathbf{H}) \mathbf{y}_l]' \boldsymbol{\Sigma}_l^{-1} [(\mathbf{I}_n - \mathbf{H}) \mathbf{y}_l] \end{aligned}$$

in cui si introduce  $\mathbf{H} = \mathbf{Z}_l (\mathbf{Z}'_l \boldsymbol{\Sigma}_l^{-1} \mathbf{Z}_l)^{-1} \mathbf{Z}'_l \boldsymbol{\Sigma}_l^{-1}$ .

Calcoliamo, come fatto in precedenza, il gradiente della funzione  $l_{\hat{\beta}}$

$$\begin{aligned} \frac{\partial l_{\hat{\beta}}}{\partial \sigma_\eta^2} &= -\frac{1}{2} \text{tr} (\boldsymbol{\Sigma}_l^{-1} \mathbf{R}_\eta) + [\mathbf{H} \mathbf{R}_\eta \boldsymbol{\Sigma}_l^{-1} \mathbf{H} \mathbf{y}_l - \mathbf{H} \mathbf{R}_\eta \boldsymbol{\Sigma}_l^{-1} \mathbf{y}_l] \boldsymbol{\Sigma}_l^{-1} (\mathbf{I}_n - \mathbf{H}) \mathbf{y}_l + \\ &\quad + \frac{1}{2} [(\mathbf{I}_n - \mathbf{H}) \mathbf{y}_l]' \boldsymbol{\Sigma}_l^{-1} \mathbf{R}_\eta \boldsymbol{\Sigma}_l^{-1} (\mathbf{I}_n - \mathbf{H}) \mathbf{y}_l \\ \frac{\partial l_{\hat{\beta}}}{\partial \sigma_{\varepsilon_l}^2} &= -\frac{1}{2} \text{tr} (\boldsymbol{\Sigma}_l^{-1}) + [(\mathbf{H} \boldsymbol{\Sigma}_l^{-1} \mathbf{H} - \mathbf{H} \boldsymbol{\Sigma}_l^{-1}) \mathbf{y}_l]' \boldsymbol{\Sigma}_l^{-1} (\mathbf{I}_n - \mathbf{H}) \mathbf{y}_l + \\ &\quad + \frac{1}{2} [(\mathbf{I}_n - \mathbf{H}) \mathbf{y}_l]' \boldsymbol{\Sigma}_l^{-2} (\mathbf{I}_n - \mathbf{H}) \mathbf{y}_l \\ \frac{\partial l_{\hat{\beta}}}{\partial \vartheta_l^i} &= -\frac{\sigma_\eta^2}{2} \text{tr} \left( \boldsymbol{\Sigma}_l^{-1} \frac{\partial \mathbf{R}_\eta}{\partial \vartheta_l^i} \right) + \sigma_\eta^2 \left[ \left( \mathbf{H} \frac{\partial \mathbf{R}_\eta}{\partial \vartheta_l^i} \boldsymbol{\Sigma}_l^{-1} \mathbf{H} - \mathbf{H} \frac{\partial \mathbf{R}_\eta}{\partial \vartheta_l^i} \boldsymbol{\Sigma}_l^{-1} \right) \mathbf{y}_l \right]' \\ &\quad \cdot \boldsymbol{\Sigma}_l^{-1} (\mathbf{I}_n - \mathbf{H}) \mathbf{y}_l + \frac{\sigma_\eta^2}{2} [(\mathbf{I}_n - \mathbf{H}) \mathbf{y}_l]' \boldsymbol{\Sigma}_l^{-1} \frac{\partial \mathbf{R}_\eta}{\partial \vartheta_l^i} \boldsymbol{\Sigma}_l^{-1} (\mathbf{I}_n - \mathbf{H}) \mathbf{y}_l \end{aligned} \tag{3.19}$$

In questo caso non si possono ricavare analiticamente i valori degli stimatori e le variabili da ottimizzare sono presenti in funzioni non lineari, tuttavia queste derivate possono essere usate per essere inserite in un algoritmo di ottimizzazione di tipo *quasi-Newton* ed ottenere una soluzione più precisa rispetto al calcolo del gradiente mediante metodi numerici.

### 3.4.3 Valore atteso e predizione in nuovi punti e loro proprietà

#### Proprietà del predittore lineare

Si verifica ora che lo stimatore della media,  $\hat{y}_l(\mathbf{x}_0)$ , è *BLUP* (*Best Linear Unbiased Predictor*). Si inizia col calcolare la combinazione lineare di  $\mathbf{y}_l$  in cui si esprime  $\hat{\mu}_l(\mathbf{x}_0)$ , definito come  $\hat{\mu}_l(\mathbf{x}_0) \stackrel{def}{=} \hat{\mathbb{E}}(\hat{y}_l(\mathbf{x}_0))$ .

$$\begin{aligned}
\hat{y}_{l_0} &= \mathbf{z}'_{l_0} \hat{\boldsymbol{\beta}} + \sigma_\eta^2 \mathbf{r}'_{\eta_0} \boldsymbol{\Sigma}_l^{-1} (\mathbf{y}_l - \mathbf{Z}_l \hat{\boldsymbol{\beta}}) \\
&= \left[ (\mathbf{z}'_{l_0} - \sigma_\eta^2 \mathbf{r}'_{\eta_0} \boldsymbol{\Sigma}_l^{-1} \mathbf{Z}_l) (\mathbf{Z}'_l \boldsymbol{\Sigma}_l^{-1} \mathbf{Z}_l)^{-1} \mathbf{Z}'_l \boldsymbol{\Sigma}_l^{-1} + \sigma_\eta^2 \mathbf{r}'_{\eta_0} \boldsymbol{\Sigma}_l^{-1} \right] \mathbf{y}_l \\
&= \left\{ \mathbf{z}'_{l_0} (\mathbf{Z}'_l \boldsymbol{\Sigma}_l^{-1} \mathbf{Z}_l)^{-1} \mathbf{Z}'_l \boldsymbol{\Sigma}_l^{-1} + \sigma_\eta^2 \mathbf{r}_{\eta_0} \boldsymbol{\Sigma}_l^{-1} \left[ \mathbf{I}_n - \mathbf{Z}_l (\mathbf{Z}'_l \boldsymbol{\Sigma}_l^{-1} \mathbf{Z}_l)^{-1} \mathbf{Z}'_l \boldsymbol{\Sigma}_l^{-1} \right] \right\} \mathbf{y}_l = \\
&= \mathbf{a}'_* \mathbf{y}_l \tag{3.20}
\end{aligned}$$

dove

$$\mathbf{a}'_* = \mathbf{z}'_{l_0} (\mathbf{Z}'_l \boldsymbol{\Sigma}_l^{-1} \mathbf{Z}_l)^{-1} \mathbf{Z}'_l \boldsymbol{\Sigma}_l^{-1} + \sigma_\eta^2 \mathbf{r}_{\eta_0} \boldsymbol{\Sigma}_l^{-1} \left[ \mathbf{I}_n - \mathbf{Z}_l (\mathbf{Z}'_l \boldsymbol{\Sigma}_l^{-1} \mathbf{Z}_l)^{-1} \mathbf{Z}'_l \boldsymbol{\Sigma}_l^{-1} \right]. \tag{3.21}$$

In cui  $\mathbf{z}_{l_0}$  rappresenta  $f(\mathbf{x}_0)$  ed  $\mathbf{r}_{\eta_0}$  la correlazione tra  $\mathbf{y}_l$  e  $y_l(\mathbf{x}_0)$ .

A questo punto si cerca una proprietà che deve possedere una combinazione lineare di  $\mathbf{y}_l$ , rappresentata da  $\mathbf{a}' \mathbf{y}_l$ , per fare in modo che lo stimatore lineare sia corretto, questa proprietà tornerà utile in seguito:

$$\mathbb{E}(\mathbf{a}' \mathbf{y}_l) = \mathbf{a}' \mathbf{Z}_l \boldsymbol{\beta} = \mathbb{E}[y_l(\mathbf{x}_0)] = \mathbf{z}'_{l_0} \boldsymbol{\beta} \tag{3.22}$$

da cui  $\mathbf{Z}_l \mathbf{a} = \mathbf{z}_{l_0}$ .

Si dimostra che anche lo stimatore trovato in (3.21) soddisfa questa proprietà, infatti:

$$\mathbb{E}[\hat{y}_l(\mathbf{x}_0)] = \mathbf{a}'_* \mathbf{Z}_l \boldsymbol{\beta} = [\mathbf{z}'_{l_0} + \sigma_\eta^2 \mathbf{r}'_{\eta_0} \boldsymbol{\Sigma}_l^{-1} (\mathbf{Z}_l - \mathbf{Z}_l)] \boldsymbol{\beta} = \mathbf{z}'_{l_0} \boldsymbol{\beta} = \mathbb{E}[y_l(\mathbf{x}_0)]$$

Si hanno ora tutti gli strumenti per verificare che lo stimatore è *BLUP*. I passi sono gli stessi utilizzati in [38], con la differenza che in quel caso non è presente l'errore casuale.

Si calcola il *Mean Square Prediction Error* (MSPE) per trovare la combinazione lineare di  $\mathbf{y}_l$  che minimizza questa quantità in un generico punto  $\mathbf{x}_0$ . Il MSPE è definito come  $\mathbb{E} \left\{ \left[ \hat{f}(\mathbf{x}_i) - f(\mathbf{x}_i) \right]^2 \right\}$ .

$$\begin{aligned}
\mathbb{E} \left\{ [\mathbf{a}' \mathbf{y}_l - y_l(\mathbf{x}_0)]^2 \right\} &= \mathbb{E} \left\{ [\mathbf{a}' (\mathbf{Z}_l \boldsymbol{\beta} + \boldsymbol{\eta} + \boldsymbol{\varepsilon}_l) - \mathbf{z}'_{l_0} \boldsymbol{\beta} - \eta(\mathbf{x}_0) - \varepsilon_{l_0}]^2 \right\} = \\
&= \mathbb{E} \left[ \left( \underbrace{\mathbf{a}' \mathbf{Z}_l \boldsymbol{\beta} - \mathbf{z}'_{l_0} \boldsymbol{\beta}}_{\stackrel{(3.22)}{=} 0} + \mathbf{a}' \boldsymbol{\eta} - \eta(\mathbf{x}_0) + \mathbf{a}' \boldsymbol{\varepsilon}_l + \varepsilon_{l_0} \right)^2 \right] = \\
&= \mathbb{E} \left( \mathbf{a}' \boldsymbol{\eta} \boldsymbol{\eta}' \mathbf{a} + \eta(\mathbf{x}_0)^2 + \mathbf{a}' \boldsymbol{\varepsilon}_l \boldsymbol{\varepsilon}'_l \mathbf{a} + \varepsilon_{l_0}^2 - 2 \mathbf{a}' \boldsymbol{\eta} \eta(\mathbf{x}_0) + \right. \\
&\quad \left. 2 \mathbf{a}' \boldsymbol{\eta} \boldsymbol{\varepsilon}'_l \mathbf{a} - 2 \mathbf{a}' \boldsymbol{\eta} \varepsilon_{l_0} - 2 \eta(\mathbf{x}_0) \mathbf{a}' \boldsymbol{\varepsilon}_l + 2 \eta(\mathbf{x}_0) \varepsilon_{l_0} - 2 \mathbf{a}' \boldsymbol{\varepsilon}_l \varepsilon_{l_0} \right) = \\
&= \sigma_\eta^2 \mathbf{a}' \mathbf{R}_\eta \mathbf{a} + \sigma_\eta^2 + \sigma_{\varepsilon_l}^2 \mathbf{a}' \mathbf{a} + \sigma_{\varepsilon_l}^2 - 2 \mathbf{a}' \sigma_\eta^2 \mathbf{r}_{\eta_0} = \\
&= \mathbf{a}' \boldsymbol{\Sigma}_l \mathbf{a} - 2 \sigma_\eta^2 \mathbf{a}' \mathbf{r}_{\eta_0} + \sigma_\eta^2 + \sigma_{\varepsilon_l}^2 \tag{3.23}
\end{aligned}$$



La terza uguaglianza si ha per l'indipendenza tra le varie variabili aleatorie. La funzione da minimizzare è:

$$\begin{aligned} \min_{\mathbf{a}} f(\mathbf{a}) &= \min_{\mathbf{a}} \mathbf{a}' \boldsymbol{\Sigma}_l \mathbf{a} - 2\mathbf{a}' \sigma_\eta^2 \mathbf{r}_{\eta_0} \\ \text{soggetto a} \\ \mathbf{Z}'_l \mathbf{a} &= \mathbf{z}_{l_0} \end{aligned}$$

Dove non compaiono i termini  $\sigma_{\varepsilon_l}^2$  e  $\sigma_\eta^2$  poichè non variano l'ottimizzazione rispetto ad  $\mathbf{a}$ . Il metodo dei moltiplicatori di Lagrange può essere utilizzato per risolvere questo problema di minimizzazione quadratica vincolata, la nuova funzione da massimizzare diventa:

$$\min_{\mathbf{a}, \boldsymbol{\lambda}} L(\mathbf{a}, \boldsymbol{\lambda}) = \min_{\mathbf{a}, \boldsymbol{\lambda}} \mathbf{a}' \boldsymbol{\Sigma}_l \mathbf{a} - 2\mathbf{a}' \sigma_\eta^2 \mathbf{r}_{\eta_0} + 2\boldsymbol{\lambda}' (\mathbf{Z}'_l \mathbf{a} - \mathbf{z}_{l_0})$$

si calcola ora il gradiente della funzione rispetto a  $\mathbf{a}$  e  $\boldsymbol{\lambda}$  e si pone uguale al vettore nullo per cercare eventuali punti stazionari:

$$\begin{aligned} \frac{\partial L(\mathbf{a}, \boldsymbol{\lambda})}{\partial \mathbf{a}} &= 2 \boldsymbol{\Sigma}_l \mathbf{a} - 2\sigma_\eta^2 \mathbf{r}_{\eta_0} + 2 \mathbf{Z}_l \boldsymbol{\lambda} = \mathbf{0} \\ \frac{\partial L(\mathbf{a}, \boldsymbol{\lambda})}{\partial \boldsymbol{\lambda}} &= 2 (\mathbf{Z}'_l \mathbf{a} - \mathbf{z}_{l_0}) = \mathbf{0} \end{aligned}$$

si riscrive ora il sistema lineare in forma matriciale:

$$\begin{bmatrix} \boldsymbol{\Sigma}_l & \mathbf{Z}_l \\ \mathbf{Z}'_l & \mathbf{0} \end{bmatrix} \begin{pmatrix} \mathbf{a} \\ \boldsymbol{\lambda} \end{pmatrix} = \begin{pmatrix} \sigma_\eta^2 \mathbf{r}_{\eta_0} \\ \mathbf{z}_{l_0} \end{pmatrix}$$

si risolve il sistema lineare invertendo la matrice a blocchi:

$$\begin{aligned} \begin{pmatrix} \bar{\mathbf{a}} \\ \bar{\boldsymbol{\lambda}} \end{pmatrix} &= \begin{bmatrix} \boldsymbol{\Sigma}_l & \mathbf{Z}_l \\ \mathbf{Z}'_l & \mathbf{0} \end{bmatrix}^{-1} \begin{pmatrix} \sigma_\eta^2 \mathbf{r}_{\eta_0} \\ \mathbf{z}_{l_0} \end{pmatrix} = \\ &= \begin{bmatrix} \boldsymbol{\Sigma}_l^{-1} - \boldsymbol{\Sigma}_l^{-1} \mathbf{Z}_l (\mathbf{Z}'_l \boldsymbol{\Sigma}_l^{-1} \mathbf{Z}_l)^{-1} \mathbf{Z}_l \boldsymbol{\Sigma}_l^{-1} & \boldsymbol{\Sigma}_l^{-1} \mathbf{Z}_l (\mathbf{Z}'_l \boldsymbol{\Sigma}_l^{-1} \mathbf{Z}_l)^{-1} \\ (\mathbf{Z}'_l \boldsymbol{\Sigma}_l^{-1} \mathbf{Z}_l)^{-1} \mathbf{Z}_l \boldsymbol{\Sigma}_l^{-1} & - (\mathbf{Z}'_l \boldsymbol{\Sigma}_l^{-1} \mathbf{Z}_l)^{-1} \end{bmatrix} \begin{pmatrix} \sigma_\eta^2 \mathbf{r}_{\eta_0} \\ \mathbf{z}_{l_0} \end{pmatrix} = \\ &= \begin{pmatrix} \boldsymbol{\Sigma}_l^{-1} \mathbf{Z}_l (\mathbf{Z}'_l \boldsymbol{\Sigma}_l^{-1} \mathbf{Z}_l)^{-1} \mathbf{z}_{l_0} + \sigma_\eta^2 \left[ \boldsymbol{\Sigma}_l^{-1} - \boldsymbol{\Sigma}_l^{-1} \mathbf{Z}_l (\mathbf{Z}'_l \boldsymbol{\Sigma}_l^{-1} \mathbf{Z}_l)^{-1} \mathbf{Z}'_l \boldsymbol{\Sigma}_l^{-1} \right] \mathbf{r}_{\eta_0} \\ - (\mathbf{Z}'_l \boldsymbol{\Sigma}_l^{-1} \mathbf{Z}_l)^{-1} \mathbf{z}_{l_0} + \sigma_\eta^2 (\mathbf{Z}'_l \boldsymbol{\Sigma}_l^{-1} \mathbf{Z}_l)^{-1} \mathbf{Z}'_l \boldsymbol{\Sigma}_l^{-1} \mathbf{r}_{\eta_0} \end{pmatrix} \end{aligned}$$

si nota che  $\bar{\mathbf{a}}$  coincide con  $\mathbf{a}_*$ , l'hessiana di  $f(\mathbf{a})$  risulta  $2\mathbf{R}_\eta$ , che è definita positiva e questo è sufficiente per affermare che  $\bar{\mathbf{a}}$  è il minimo di  $f(\mathbf{a})$  e quindi anche per  $\mathbb{E} \left\{ [\mathbf{a}' \mathbf{y}_l - y_l(\mathbf{x}_0)]^2 \right\}$ .

**Valore atteso e predizione in nuovi punti: caso media nota**

Per la predizione della risposta in un nuovo punto del piano sperimentale  $\mathbf{x}_0$  si considera prima il caso in cui la media sia nota. Per ipotesi la variabile aleatoria è un processo gaussiano, per cui le risposte nei punti del piano sperimentale e quelle da predire hanno una distribuzione congiunta gaussiana riportata nella (3.24).

$$\begin{pmatrix} \mathbf{y}_l \\ \hat{\mu}_l(\mathbf{x}_0) \end{pmatrix} | \boldsymbol{\Sigma}_l, \boldsymbol{\beta} \sim \mathcal{N}_{n+1} \left( \begin{pmatrix} \mathbf{Z}_l \boldsymbol{\beta} \\ \mathbf{z}'_{l_0} \boldsymbol{\beta} \end{pmatrix}, \begin{bmatrix} \boldsymbol{\Sigma}_l & \sigma_\eta^2 \mathbf{r}_{\eta_{12}} \\ \sigma_\eta^2 \mathbf{r}'_{\eta_{12}} & \sigma_\eta^2 \end{bmatrix} \right) \quad (3.24)$$

Dove  $\mathbf{r}_{\eta_{12}} = \text{Corr}(\mathbf{y}_l, y_l(\mathbf{x}_0))$  e  $\mathbf{z}_{l_0} = f(\mathbf{x}_0)$ .

Si sfruttano ora la proprietà della normale che afferma che la variabile aleatoria condizionata è sempre distribuita come una normale con media e varianza come indicato nella (3.25).

$$\hat{\mu}_l(\mathbf{x}_0) | \mathbf{y}_l, \boldsymbol{\Sigma}_l, \boldsymbol{\beta} \sim \mathcal{N}(\mathbf{z}'_{l_0} \boldsymbol{\beta} + \sigma_\eta^2 \mathbf{r}'_{\eta_{12}} \boldsymbol{\Sigma}_l^{-1} (\mathbf{y}_l - \mathbf{Z}_l \boldsymbol{\beta}), \sigma_\eta^2 - \sigma_\eta^4 \mathbf{r}'_{\eta_{12}} \boldsymbol{\Sigma}_l^{-1} \mathbf{r}_{\eta_{12}}) \quad (3.25)$$

Il valore atteso può essere suddiviso in due parti, la prima  $\mathbf{z}'_{l_0} \boldsymbol{\beta}$  rappresenta l'effetto del termine lineare presente nel modello, il secondo,  $\sigma_\eta^2 \mathbf{r}'_{\eta_{12}} \boldsymbol{\Sigma}_l^{-1} (\mathbf{y}_l - \mathbf{Z}_l \boldsymbol{\beta})$  corrisponde alla parte della predizione dovuta alla correlazione tra punti vicini, rappresenta quindi l'effetto del processo gaussiano. Inserendo i valori degli stimatori al posto dei parametri si può affermare che lo stimatore ottenuto è asintoticamente corretto, non si può affermare di più non sapendo le proprietà degli stimatori degli altri parametri del modello.

*Dimostrazione.*

$$\mathbb{E}(\hat{\mu}_l(\mathbf{x}_0) | \mathbf{y}_l, \boldsymbol{\Sigma}_l) = \mathbf{z}'_{l_0} \mathbb{E}(\hat{\boldsymbol{\beta}}) + \sigma_\eta^2 \mathbf{r}'_{\eta_{12}} \boldsymbol{\Sigma}_l^{-1} (\mathbb{E}(\mathbf{y}_l) - \mathbf{Z}_l \mathbb{E}(\hat{\boldsymbol{\beta}})) = \mathbf{z}'_{l_0} \boldsymbol{\beta}$$

□

Si passa ora alla predizione di nuovi punti, l'unica modifica che va fatta rispetto al caso precedente è l'aggiunta l'errore casuale  $\sigma_{\varepsilon_l}^2$  nella distribuzione congiunta.

$$\begin{pmatrix} \mathbf{y}_l \\ \hat{y}_l(\mathbf{x}_0) \end{pmatrix} | \boldsymbol{\Sigma}_l, \boldsymbol{\beta} \sim \mathcal{N}_{n+1} \left( \begin{pmatrix} \mathbf{Z}_l \boldsymbol{\beta} \\ \mathbf{z}'_{l_0} \boldsymbol{\beta} \end{pmatrix}, \begin{bmatrix} \boldsymbol{\Sigma}_l & \sigma_\eta^2 \mathbf{r}_{\eta_{12}} \\ \sigma_\eta^2 \mathbf{r}'_{\eta_{12}} & \sigma_\eta^2 + \sigma_{\varepsilon_l}^2 \end{bmatrix} \right) \quad (3.26)$$

Quindi, si applica la stessa proprietà del caso precedente per trovare la distribuzione di  $\hat{y}_l(\mathbf{x}_0) | \mathbf{y}_l, \boldsymbol{\Sigma}_l, \boldsymbol{\beta}$ .

$$\hat{y}_l(\mathbf{x}_0) | \mathbf{y}_l, \boldsymbol{\Sigma}_l, \boldsymbol{\beta} \sim \mathcal{N}(\mathbf{z}'_{l_0} \boldsymbol{\beta} + \sigma_\eta^2 \mathbf{r}'_{\eta_{12}} \boldsymbol{\Sigma}_l^{-1} (\mathbf{y}_l - \mathbf{Z}_l \boldsymbol{\beta}), \sigma_\eta^2 + \sigma_{\varepsilon_l}^2 - \sigma_\eta^4 \mathbf{r}'_{\eta_{12}} \boldsymbol{\Sigma}_l^{-1} \mathbf{r}_{\eta_{12}})$$

Anche in questo caso lo stimatore è corretto ed i calcoli sono gli stessi del caso precedente.

### Valore atteso e predizione in nuovi punti: caso media non nota

Si passa ora al caso in cui la media non sia nota, ma stimata dai dati. In questo secondo caso l'incertezza di previsione nel punto predetto è data dal MSPE, [40], con  $\mathbf{a} = \mathbf{a}_*$ , definito nella (3.21). Il risultato finale è presente nella (3.27).

$$\begin{aligned} MSPE [\hat{y}_l(\mathbf{x}_0)|\mathbf{y}_l, \Sigma_l] = & \sigma_{\varepsilon_l}^2 + \sigma_{\eta}^2 - \sigma_{\eta}^4 \mathbf{r}_{\eta_{12}} \Sigma_l^{-1} \mathbf{r}_{\eta_{12}} + \\ & + (\mathbf{z}_{l_0} + \sigma_{\eta}^2 \mathbf{r}'_{\eta_{12}} \Sigma_l^{-1} \mathbf{Z}_l)' \Sigma_{\hat{\beta}} (\mathbf{z}_{l_0} + \sigma_{\eta}^2 \mathbf{r}'_{\eta_{12}} \Sigma_l^{-1} \mathbf{Z}_l) \end{aligned} \quad (3.27)$$

In questo caso la varianza tende, asintoticamente, alla varianza in cui la media è considerata nota. Per il calcolo della varianza della risposta media basta, grazie all'indipendenza tra  $\varepsilon_l(\mathbf{x}_0)$  e  $\hat{\mu}_l(\mathbf{x}_0)$ , rimuovere solo l'errore casuale e si giunge alla (3.28).

$$\begin{aligned} MSPE [\hat{\mu}_l(\mathbf{x}_0)|\mathbf{y}_l, \Sigma_l] = & \sigma_{\eta}^2 - \sigma_{\eta}^4 \mathbf{r}_{\eta_{12}} \Sigma_l^{-1} \mathbf{r}_{\eta_{12}} + \\ & + (\mathbf{z}_{l_0} + \sigma_{\eta}^2 \mathbf{r}'_{\eta_{12}} \Sigma_l^{-1} \mathbf{Z}_l)' \Sigma_{\hat{\beta}} (\mathbf{z}_{l_0} + \sigma_{\eta}^2 \mathbf{r}'_{\eta_{12}} \Sigma_l^{-1} \mathbf{Z}_l) \end{aligned} \quad (3.28)$$

## 3.5 Modello alta risoluzione

### 3.5.1 Introduzione

In questa sezione si estende il modello di secondo stadio presente in [34], nel quale si collegano i punti a bassa accuratezza con quelli con una precisione maggiore non considerando la variabilità nella predizione dei punti di primo stadio. Gli autori non hanno considerato questo caso poichè i punti a bassa accuratezza da loro analizzati provengono da simulatori deterministici, per cui se i punti ad alta risoluzione sono un sottoinsieme di questi, non hanno bisogno di rilassare questa ipotesi. Si osserva che i punti sperimentali di questo modello possono anche non essere un sottoinsieme dei punti che sono stati utilizzati nel piano sperimentale del modello a bassa risoluzione.

Per rappresentare gli  $n_1$  punti che provengono dallo strumento di misura con migliore accuratezza si utilizza il processo gaussiano descritto nella (3.29).

$$\mathbf{y}_h = \mathbf{Z}_h \boldsymbol{\rho} + \delta_0 \mathbf{1}_{n_1} + \boldsymbol{\delta} + \boldsymbol{\varepsilon}_h \quad (3.29)$$

si analizzano singolarmente le varie componenti del modello:

- $\mathbf{y}_h \in \mathbb{R}^{n_1}$ : risposte ad alta accuratezza negli  $n_1$  punti misurati;

- $\mathbf{Z}_h = \begin{bmatrix} \tilde{\mathbf{z}}'_{h_1} y_l(\mathbf{x}_1) \\ \vdots \\ \tilde{\mathbf{z}}'_{h_{n_1}} y_l(\mathbf{x}_{n_1}) \end{bmatrix} \in (n_1, r)$ : matrice disegno del modello ad alta risoluzione.  $y_l(\mathbf{x}_i)$  rappresenta il valore dell' $i$ -esimo punto a bassa risoluzione,

ma non essendo disponibile si utilizza il valore predetto dal modello a bassa risoluzione  $\hat{\boldsymbol{\mu}}_l(\mathbf{x}_i)$ , mentre  $\tilde{\mathbf{z}}'_{h_i} = \mathbf{f}'_h(\mathbf{x}_i)$  sono le righe della matrice disegno convenzionale, nel resto del capitolo sarà omesso il  $\hat{\cdot}$  alla matrice  $\tilde{\mathbf{Z}}_h$  e nel modello ad alta risoluzione comparirà il vettore  $\hat{\boldsymbol{\mu}}_l$  al posto del vettore  $\mathbf{y}_l$ ;

- $\boldsymbol{\rho} \in \mathbb{R}^r$ : vettore dei regressori;
- $\boldsymbol{\delta} \sim \mathcal{PG}_{n_1}(\mathbf{0}_{n_1}, \sigma_{\boldsymbol{\delta}}^2 \mathbf{R}_{\boldsymbol{\delta}})$ ,  $\boldsymbol{\delta} \in \mathbb{R}^{n_1}$ : processo gaussiano utilizzato per determinare effetti non noti di scalibratura dello strumento a bassa risoluzione;
- $\delta_0$ : parametro che rappresenta la traslazione dei punti;
- $\boldsymbol{\varepsilon}_h \sim \mathcal{N}_{n_1}(\mathbf{0}_{n_1}, \sigma_{\boldsymbol{\varepsilon}_h}^2 \mathbf{I}_{n_1})$ ,  $\boldsymbol{\varepsilon}_h \in \mathbb{R}^{n_1}$ : vettore aleatorio che rappresenta l'errore casuale, indipendente dallo specifico punto dello spazio, dello strumento con elevata precisione.

$\mathbf{R}_{\boldsymbol{\delta}}(\vartheta_1^h, \vartheta_2^h)$  rappresenta la matrice di correlazione gaussiana definita in modo analogo a  $\mathbf{R}_{\boldsymbol{\eta}}$  del modello a bassa risoluzione:

$$\mathbf{R}_{\boldsymbol{\delta}} = [r_{\delta_{ij}}], \quad r_{\delta_{ij}} = \exp \left\{ - \left[ \vartheta_1^h (x_{i1} - x_{j1})^2 + \vartheta_2^h (x_{i2} - x_{j2})^2 \right] \right\} \quad (3.30)$$

Anche in questo modello si ipotizza l'indipendenza tra  $\boldsymbol{\delta}$ ,  $\boldsymbol{\varepsilon}_h$  e  $\hat{\boldsymbol{\mu}}_l$ .

Un modo alternativo di scrivere il modello ad alta risoluzione è descritto nella (3.31). Questa notazione è utile per calcolare la matrice di varianza covarianza del vettore  $\mathbf{y}_h$ .

$$\mathbf{y}_h = \mathbf{M}\hat{\boldsymbol{\mu}}_l + \delta_0 \mathbf{1}_{n_1} + \boldsymbol{\delta} + \boldsymbol{\varepsilon}_h \quad (3.31)$$

con

$$\mathbf{M} = \text{diag}(\tilde{\mathbf{Z}}_h \boldsymbol{\rho})$$

Con queste considerazioni si può determinare la distribuzione di  $\mathbf{y}_h$  presente nella (3.32).

$$\mathbf{y}_h \sim \mathcal{N}_{n_1}(\mathbf{Z}_h \boldsymbol{\rho} + \delta_0 \mathbf{1}_{n_1}, \boldsymbol{\Sigma}_h) \quad (3.32)$$

in cui:

$$\boldsymbol{\Sigma}_h = \mathbf{M}\boldsymbol{\Sigma}_0 \mathbf{M}' + \sigma_{\boldsymbol{\delta}}^2 \mathbf{R}_{\boldsymbol{\delta}} + \sigma_{\boldsymbol{\varepsilon}_h}^2 \mathbf{I}_{n_1} \quad (3.33)$$

con  $\boldsymbol{\Sigma}_0 = \text{Cov}(\hat{\boldsymbol{\mu}}_l | \mathbf{y}_l, \boldsymbol{\Sigma}_l)$ , che è facilmente ricavabile generalizzando la (3.25):

$$\text{Cov}(\hat{\boldsymbol{\mu}}_l | \mathbf{y}_l, \boldsymbol{\Sigma}_l) = \sigma_{\boldsymbol{\eta}}^2 \mathbf{R}_{\boldsymbol{\eta}22} - \sigma_{\boldsymbol{\eta}}^4 \mathbf{R}'_{\boldsymbol{\eta}12} \boldsymbol{\Sigma}_l^{-1} \mathbf{R}_{\boldsymbol{\eta}12} \quad (3.34)$$

o la (3.28) se non si vuole utilizzare la varianza asintotica:

$$\begin{aligned}
MSPE(\hat{\boldsymbol{\mu}}_l | \mathbf{y}_l, \boldsymbol{\Sigma}_l) = & \sigma_\eta^2 \mathbf{R}_{\eta_{22}} - \sigma_\eta^4 \mathbf{R}'_{\eta_{12}} \boldsymbol{\Sigma}_l^{-1} \mathbf{R}_{\eta_{12}} + \\
& + (\mathbf{Z}_{l_0} + \sigma_\eta^2 \mathbf{R}'_{\eta_{12}} \boldsymbol{\Sigma}_l^{-1} \mathbf{Z}_l)' \boldsymbol{\Sigma}_{\hat{\boldsymbol{\beta}}} (\mathbf{Z}_{l_0} + \sigma_\eta^2 \mathbf{R}'_{\eta_{12}} \boldsymbol{\Sigma}_l^{-1} \mathbf{Z}_l)
\end{aligned} \quad (3.35)$$

dove  $\mathbf{Z}_{l_0} \in (n_1, p)$  è la matrice disegno del modello a bassa risoluzione nei punti ad elevata accuratezza,  $\mathbf{R}_{\eta_{22}} \in (n_1, n_1)$  è l'applicazione della funzione di correlazione gaussiana associata al processo  $\eta(\cdot)$  nei punti ad alta accuratezza e  $\mathbf{R}_{\eta_{12}} \in (n, n_1)$  è l'applicazione della funzione di correlazione gaussiana associata al processo  $\eta(\cdot)$  tra i punti a bassa ed alta accuratezza.

Si osserva che nel modello di secondo stadio  $\boldsymbol{\rho}$  compare sia nella media che nella matrice di varianza covarianza e che la matrice  $\mathbf{M}$  è simmetrica.

### 3.5.2 Stima dei parametri

Come eseguito in precedenza si trovano le stime dei parametri massimizzando la verosimiglianza presente nella (3.36).

$$L = \frac{1}{(2\pi)^{\frac{n_1}{2}} |\boldsymbol{\Sigma}_h|^{\frac{1}{2}}} \exp \left\{ -\frac{(\mathbf{y}_h - \mathbf{Z}_h \boldsymbol{\rho} - \delta_0 \mathbf{1}_{n_1})' \boldsymbol{\Sigma}_h^{-1} (\mathbf{y}_h - \mathbf{Z}_h \boldsymbol{\rho} - \delta_0 \mathbf{1}_{n_1})}{2} \right\} \quad (3.36)$$

Si procede ora con il calcolo del logaritmo di  $L$ .

$$l = -\frac{n_1}{2} \log 2\pi - \frac{1}{2} \log |\boldsymbol{\Sigma}_h| - \frac{(\mathbf{y}_h - \mathbf{Z}_h \boldsymbol{\rho} - \delta_0 \mathbf{1}_{n_1})' \boldsymbol{\Sigma}_h^{-1} (\mathbf{y}_h - \mathbf{Z}_h \boldsymbol{\rho} - \delta_0 \mathbf{1}_{n_1})}{2}$$

Si distingue ora il caso in cui  $\rho$  sia scalare e  $\mathbf{Z}_h = \hat{\boldsymbol{\mu}}_l$  da quello generale per le differenze nei calcoli delle derivate di  $l$ . Il logaritmo della verosimiglianza in questo caso particolare risulta:

$$\begin{aligned}
l = & -\frac{n_1}{2} \log 2\pi - \frac{1}{2} \log |\rho^2 \boldsymbol{\Sigma}_0 + \sigma_\delta^2 \mathbf{R}_\delta + \sigma_{\varepsilon_h}^2 \mathbf{I}_{n_1}| + \\
& - \frac{(\mathbf{y}_h - \rho \hat{\boldsymbol{\mu}}_l - \delta_0 \mathbf{1}_{n_1})' (\rho^2 \boldsymbol{\Sigma}_0 + \sigma_\delta^2 \mathbf{R}_\delta + \sigma_{\varepsilon_h}^2 \mathbf{I}_{n_1})^{-1} (\mathbf{y}_h - \rho \hat{\boldsymbol{\mu}}_l - \delta_0 \mathbf{1}_{n_1})}{2}
\end{aligned}$$

Data la maggiore semplicità nel calcolo si cercano prima i punti stazionari di  $l$  quando  $\rho$  è uno scalare. Il  $\nabla l$  risulta:

$$\begin{aligned}
\frac{\partial l}{\partial \delta_0} &= \mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} (\mathbf{y}_h - \rho \hat{\boldsymbol{\mu}}_l - \delta_0 \mathbf{1}_{n_1}) \\
\frac{\partial l}{\partial \rho} &= -\rho \operatorname{tr} (\boldsymbol{\Sigma}_h^{-1} \boldsymbol{\Sigma}_0) + \hat{\boldsymbol{\mu}}_l' \boldsymbol{\Sigma}_h^{-1} (\mathbf{y}_h - \rho \hat{\boldsymbol{\mu}}_l - \delta_0 \mathbf{1}_{n_1}) + \\
&\quad + \rho (\mathbf{y}_h - \rho \hat{\boldsymbol{\mu}}_l - \delta_0 \mathbf{1}_{n_1})' \boldsymbol{\Sigma}_h^{-1} \boldsymbol{\Sigma}_0 \boldsymbol{\Sigma}_h^{-1} (\mathbf{y}_h - \rho \hat{\boldsymbol{\mu}}_l - \delta_0 \mathbf{1}_{n_1}) \\
\frac{\partial l}{\partial \sigma_\delta^2} &= -\frac{1}{2} \operatorname{tr} (\boldsymbol{\Sigma}_h^{-1} \mathbf{R}_\delta) + \frac{1}{2} (\mathbf{y}_h - \rho \hat{\boldsymbol{\mu}}_l - \delta_0 \mathbf{1}_{n_1})' \boldsymbol{\Sigma}_h^{-1} \mathbf{R}_\delta \boldsymbol{\Sigma}_h^{-1} (\mathbf{y}_h - \rho \hat{\boldsymbol{\mu}}_l - \delta_0 \mathbf{1}_{n_1}) \\
\frac{\partial l}{\partial \sigma_{\varepsilon_h}^2} &= -\frac{1}{2} \operatorname{tr} (\boldsymbol{\Sigma}_h^{-1}) + \frac{1}{2} (\mathbf{y}_h - \rho \hat{\boldsymbol{\mu}}_l - \delta_0 \mathbf{1}_{n_1})' \boldsymbol{\Sigma}_h^{-2} (\mathbf{y}_h - \rho \hat{\boldsymbol{\mu}}_l - \delta_0 \mathbf{1}_{n_1}) \\
\frac{\partial l}{\partial \vartheta_1^h} &= -\frac{1}{2} \operatorname{tr} \left( \sigma_\delta^2 \boldsymbol{\Sigma}_h^{-1} \frac{\partial \mathbf{R}_\delta}{\partial \vartheta_1^h} \right) + \frac{\sigma_\delta^2}{2} (\mathbf{y}_h - \rho \hat{\boldsymbol{\mu}}_l - \delta_0 \mathbf{1}_{n_1})' \boldsymbol{\Sigma}_h^{-1} \frac{\partial \mathbf{R}_\delta}{\partial \vartheta_1^h} \boldsymbol{\Sigma}_h^{-1} (\mathbf{y}_h - \rho \hat{\boldsymbol{\mu}}_l - \delta_0 \mathbf{1}_{n_1}) \\
\frac{\partial l}{\partial \vartheta_2^h} &= -\frac{1}{2} \operatorname{tr} \left( \sigma_\delta^2 \boldsymbol{\Sigma}_h^{-1} \frac{\partial \mathbf{R}_\delta}{\partial \vartheta_2^h} \right) + \frac{\sigma_\delta^2}{2} (\mathbf{y}_h - \rho \hat{\boldsymbol{\mu}}_l - \delta_0 \mathbf{1}_{n_1})' \boldsymbol{\Sigma}_h^{-1} \frac{\partial \mathbf{R}_\delta}{\partial \vartheta_2^h} \boldsymbol{\Sigma}_h^{-1} (\mathbf{y}_h - \rho \hat{\boldsymbol{\mu}}_l - \delta_0 \mathbf{1}_{n_1})
\end{aligned}$$

con

$$\frac{\partial \mathbf{R}_\delta}{\partial \vartheta_i^h} = [r_{\delta_{jk}^i}], \quad r_{\delta_{jk}^i} = -(x_{ji} - x_{ki})^2 r_{\delta_{jk}}$$

in questo caso si riesce a ricavare solo  $\hat{\delta}_0$  analiticamente:

$$\hat{\delta}_0 = \frac{\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} (\mathbf{y}_h - \rho \hat{\boldsymbol{\mu}}_l)}{\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1}}$$

Avendo trovato l'unico valore, se sono unici anche gli altri, per  $\hat{\delta}_0$ , possiamo inserirlo nella massima verosimiglianza e massimizzare questa funzione rispetto alle altre variabili. Il logaritmo della verosimiglianza diventa:

$$\begin{aligned}
l_{\hat{\delta}_0} &= -\frac{n_1}{2} \log 2\pi - \frac{1}{2} \log |\boldsymbol{\Sigma}_h| - \frac{1}{2} \left[ \mathbf{y}_h - \rho \hat{\boldsymbol{\mu}}_l - \frac{\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} (\mathbf{y}_h - \rho \hat{\boldsymbol{\mu}}_l) \mathbf{1}_{n_1}}{\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1}} \right]' \boldsymbol{\Sigma}_h^{-1} \cdot \\
&\quad \cdot \left[ \mathbf{y}_h - \rho \hat{\boldsymbol{\mu}}_l - \frac{\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} (\mathbf{y}_h - \rho \hat{\boldsymbol{\mu}}_l) \mathbf{1}_{n_1}}{\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1}} \right] = \\
&= -\frac{n_1}{2} \log 2\pi - \frac{1}{2} \log |\boldsymbol{\Sigma}_h| - \frac{1}{2 (\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1})^2} [(\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1} \mathbf{I}_{n_1} - \mathbf{1}_{n_1} \mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1}) \mathbf{y}_h + \\
&\quad + (\mathbf{1}_{n_1} \mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \hat{\boldsymbol{\mu}}_l - \mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1} \hat{\boldsymbol{\mu}}_l) \rho] \boldsymbol{\Sigma}_h^{-1} [(\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1} \mathbf{I}_{n_1} - \mathbf{1}_{n_1} \mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1}) \mathbf{y}_h + \\
&\quad + (\mathbf{1}_{n_1} \mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \hat{\boldsymbol{\mu}}_l - \mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1} \hat{\boldsymbol{\mu}}_l) \rho]
\end{aligned}$$

Per semplicità si definisce:

$$\mathbf{v} = (\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1} \mathbf{I}_{n_1} - \mathbf{1}_{n_1} \mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1}) \mathbf{y}_h + (\mathbf{1}_{n_1} \mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \hat{\boldsymbol{\mu}}_l - \mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1} \hat{\boldsymbol{\mu}}_l) \rho$$

Si passa ora al calcolo del gradiente:

$$\begin{aligned}
\frac{\partial l_{\hat{\delta}_0}}{\partial \sigma_{\delta}^2} &= -\frac{1}{2} \text{tr} (\boldsymbol{\Sigma}_h^{-1} \mathbf{R}_{\delta}) - \frac{1}{2 (\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1})^2} \left\{ \frac{2}{\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1}} \mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{R}_{\delta} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1} \mathbf{v} \boldsymbol{\Sigma}_h^{-1} \mathbf{v} + \right. \\
&\quad \left. - 2 \mathbf{v}' \boldsymbol{\Sigma}_h^{-1} [(\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{R}_{\delta} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1} \mathbf{I}_{n_1} - \mathbf{1}_{n_1} \mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{R}_{\delta} \boldsymbol{\Sigma}_h^{-1}) \mathbf{y}_h + \right. \\
&\quad \left. + (\mathbf{1}_{n_1} \mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{R}_{\delta} \boldsymbol{\Sigma}_h^{-1} \hat{\boldsymbol{\mu}}_l - \mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{R}_{\delta} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1} \hat{\boldsymbol{\mu}}_l) \rho] - \mathbf{v}' \boldsymbol{\Sigma}_h^{-1} \mathbf{R}_{\delta} \boldsymbol{\Sigma}_h^{-1} \mathbf{v} \right\} \\
\frac{\partial l_{\hat{\delta}_0}}{\partial \sigma_{\varepsilon_h}^2} &= -\frac{1}{2} \text{tr} (\boldsymbol{\Sigma}_h^{-2}) - \frac{1}{2 (\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1})^2} \left\{ \frac{2}{\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1}} \mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-2} \mathbf{1}_{n_1} \mathbf{v} \boldsymbol{\Sigma}_h^{-1} \mathbf{v} - 2 \mathbf{v}' \boldsymbol{\Sigma}_h^{-1} \cdot \right. \\
&\quad \cdot [(\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-2} \mathbf{1}_{n_1} \mathbf{I}_{n_1} - \mathbf{1}_{n_1} \mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-2}) \mathbf{y}_h + (\mathbf{1}_{n_1} \mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-2} \hat{\boldsymbol{\mu}}_l - \mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-2} \mathbf{1}_{n_1} \hat{\boldsymbol{\mu}}_l) \rho] + \\
&\quad \left. - \mathbf{v}' \boldsymbol{\Sigma}_h^{-2} \mathbf{v} \right\} \\
\frac{\partial l_{\hat{\delta}_0}}{\partial \vartheta_i^h} &= -\frac{\sigma_{\delta}^2}{2} \text{tr} \left( \boldsymbol{\Sigma}_h^{-1} \frac{\partial \mathbf{R}_{\delta}}{\partial \vartheta_i^h} \right) - \frac{\sigma_{\delta}^2}{2 (\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1})^2} \left\{ \frac{2}{\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1}} \mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \frac{\partial \mathbf{R}_{\delta}}{\partial \vartheta_i^h} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1} \cdot \right. \\
&\quad \cdot \mathbf{v} \boldsymbol{\Sigma}_h^{-1} \mathbf{v} - 2 \mathbf{v}' \boldsymbol{\Sigma}_h^{-1} \left[ \left( \mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \frac{\partial \mathbf{R}_{\delta}}{\partial \vartheta_i^h} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1} \mathbf{I}_{n_1} - \mathbf{1}_{n_1} \mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \frac{\partial \mathbf{R}_{\delta}}{\partial \vartheta_i^h} \boldsymbol{\Sigma}_h^{-1} \right) \cdot \right. \\
&\quad \cdot \mathbf{y}_h + \left. \left( \mathbf{1}_{n_1} \mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \frac{\partial \mathbf{R}_{\delta}}{\partial \vartheta_i^h} \boldsymbol{\Sigma}_h^{-1} \hat{\boldsymbol{\mu}}_l - \mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \frac{\partial \mathbf{R}_{\delta}}{\partial \vartheta_i^h} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1} \hat{\boldsymbol{\mu}}_l \right) \rho \right] + \\
&\quad \left. - \sigma_{\delta}^2 \mathbf{v}' \boldsymbol{\Sigma}_h^{-1} \frac{\partial \mathbf{R}_{\delta}}{\partial \vartheta_i^h} \boldsymbol{\Sigma}_h^{-1} \mathbf{v} \right\} \\
\frac{\partial l_{\hat{\delta}_0}}{\partial \rho} &= -\rho \text{tr} (\boldsymbol{\Sigma}_h^{-1} \boldsymbol{\Sigma}_0) - \frac{1}{2 (\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1})^2} \left\{ \frac{4\rho}{\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1}} \mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \boldsymbol{\Sigma}_0 \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1} \mathbf{v}' \boldsymbol{\Sigma}_h^{-1} \mathbf{v} + \right. \\
&\quad \left. - 2 \mathbf{v}' \boldsymbol{\Sigma}_h^{-1} [2\rho (\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \boldsymbol{\Sigma}_0 \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1} \mathbf{I}_{n_1} - \mathbf{1}_{n_1} \mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \boldsymbol{\Sigma}_0 \boldsymbol{\Sigma}_h^{-1}) \mathbf{y}_h + \right. \\
&\quad \left. + (\mathbf{1}_{n_1} \mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \boldsymbol{\Sigma}_0 \boldsymbol{\Sigma}_h^{-1} \hat{\boldsymbol{\mu}}_l - \mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \boldsymbol{\Sigma}_0 \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1} \hat{\boldsymbol{\mu}}_l) 2\rho^2 + \mathbf{1}_{n_1} \mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \hat{\boldsymbol{\mu}}_l + \right. \\
&\quad \left. - \mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1} \hat{\boldsymbol{\mu}}_l] - 2\rho \mathbf{v}' \boldsymbol{\Sigma}_h^{-1} \boldsymbol{\Sigma}_0 \boldsymbol{\Sigma}_h^{-1} \mathbf{v} \right\}
\end{aligned}$$

Dal gradiente non è possibile ricavare in forma chiusa nessun punto stazionario dei parametri coinvolti nella massimizzazione.

Dopo aver trattato il caso più semplice, si considera la situazione in cui  $\boldsymbol{\rho}$  è un vettore, il gradiente della funzione  $l$  risulta:

$$\begin{aligned}
\frac{\partial l}{\partial \delta_0} &= \mathbf{1}'_{n_1} \Sigma_h^{-1} (\mathbf{y}_h - \mathbf{Z}_h \boldsymbol{\rho} - \mathbf{1}_{n_1} \delta_0) \\
\frac{\partial l}{\partial \boldsymbol{\rho}} &= -\frac{1}{2} \text{tr} \left[ \left( \frac{\partial \mathbf{M} \Sigma_0 \mathbf{M}}{\partial \boldsymbol{\rho}} \right) \Sigma_h^{-1} \right] + \mathbf{Z}'_h \Sigma_h^{-1} (\mathbf{y}_h - \mathbf{Z}_h \boldsymbol{\rho} - \delta_0 \mathbf{1}_{n_1}) + \\
&\quad + \frac{1}{2} \{ \mathbf{I}_{n_1} \otimes [(\mathbf{y}_h - \mathbf{Z}_h \boldsymbol{\rho} - \delta_0 \mathbf{1}_{n_1})' \Sigma_h^{-1}] \} \left( \frac{\partial \mathbf{M} \Sigma_0 \mathbf{M}}{\partial \boldsymbol{\rho}} \Sigma_h^{-1} \right) (\mathbf{y}_h - \mathbf{Z}_h \boldsymbol{\rho} - \delta_0 \mathbf{1}_{n_1}) \\
\frac{\partial l}{\partial \sigma_\delta^2} &= -\frac{1}{2} \text{tr} (\Sigma_h^{-1} \mathbf{R}_\delta) + \frac{1}{2} (\mathbf{y}_h - \mathbf{Z}_h \boldsymbol{\rho} - \delta_0 \mathbf{1}_{n_1})' \Sigma_h^{-1} \mathbf{R}_\delta \Sigma_h^{-1} (\mathbf{y}_h - \mathbf{Z}_h \boldsymbol{\rho} - \delta_0 \mathbf{1}_{n_1}) \\
\frac{\partial l}{\partial \sigma_{\varepsilon_h}^2} &= -\frac{1}{2} \text{tr} (\Sigma_h^{-1}) + \frac{1}{2} (\mathbf{y}_h - \mathbf{Z}_h \boldsymbol{\rho} - \delta_0 \mathbf{1}_{n_1})' \Sigma_h^{-2} (\mathbf{y}_h - \mathbf{Z}_h \boldsymbol{\rho} - \delta_0 \mathbf{1}_{n_1}) \\
\frac{\partial l}{\partial \vartheta_1^h} &= -\frac{1}{2} \text{tr} \left( \sigma_\delta^2 \Sigma_h^{-1} \frac{\partial \mathbf{R}_\delta}{\partial \vartheta_1^h} \right) + \frac{\sigma_\delta^2}{2} (\mathbf{y}_h - \mathbf{Z}_h \boldsymbol{\rho} - \delta_0 \mathbf{1}_{n_1})' \Sigma_h^{-1} \frac{\partial \mathbf{R}_\delta}{\partial \vartheta_1^h} \Sigma_h^{-1} (\mathbf{y}_h - \mathbf{Z}_h \boldsymbol{\rho} + \\
&\quad - \delta_0 \mathbf{1}_{n_1}) \\
\frac{\partial l}{\partial \vartheta_2^h} &= -\frac{1}{2} \text{tr} \left( \sigma_\delta^2 \Sigma_h^{-1} \frac{\partial \mathbf{R}_\delta}{\partial \vartheta_2^h} \right) + \frac{\sigma_\delta^2}{2} (\mathbf{y}_h - \mathbf{Z}_h \boldsymbol{\rho} - \delta_0 \mathbf{1}_{n_1})' \Sigma_h^{-1} \frac{\partial \mathbf{R}_\delta}{\partial \vartheta_2^h} \Sigma_h^{-1} (\mathbf{y}_h - \mathbf{Z}_h \boldsymbol{\rho} + \\
&\quad - \delta_0 \mathbf{1}_{n_1})
\end{aligned}$$

in cui

$$\text{tr}(\mathbf{A}) = \text{tr} \begin{bmatrix} \mathbf{A}_1 \\ \vdots \\ \mathbf{A}_m \end{bmatrix} \stackrel{\text{def}}{=} \begin{pmatrix} \text{tr}(\mathbf{A}_1) \\ \vdots \\ \text{tr}(\mathbf{A}_m) \end{pmatrix}, \quad \text{con } \mathbf{A}_i \in (n, n) \forall i = 1, \dots, m$$

e

$$\frac{\partial \mathbf{M} \Sigma_0 \mathbf{M}}{\partial \boldsymbol{\rho}} = 2 \mathbf{P} \{ (\Sigma_0 \otimes \mathbf{1}_{r \times r}) \bullet [\text{vec}(\mathbf{Z}'_h) \text{vec}(\mathbf{Z}'_h)'] \} (\mathbf{I}_{n_1} \otimes \boldsymbol{\rho}) \in (n_1 r, n_1)$$

dove  $\bullet$  rappresenta il prodotto elemento per elemento e  $\otimes$  il prodotto di Kronecker.

*Dimostrazione.*

$$\begin{aligned}
\mathbf{M} \Sigma_0 \mathbf{M} &= \begin{bmatrix} \mathbf{z}'_{h_1} \boldsymbol{\rho} & 0 & \cdots & 0 \\ 0 & \mathbf{z}'_{h_2} \boldsymbol{\rho} & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \mathbf{z}'_{h_{n_1}} \boldsymbol{\rho} \end{bmatrix} \Sigma_0 \begin{bmatrix} \mathbf{z}'_{h_1} \boldsymbol{\rho} & 0 & \cdots & 0 \\ 0 & \mathbf{z}'_{h_2} \boldsymbol{\rho} & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \mathbf{z}'_{h_{n_1}} \boldsymbol{\rho} \end{bmatrix} = \\
&\begin{bmatrix} \sigma_{11} \boldsymbol{\rho}' \mathbf{z}_{h_1} \mathbf{z}'_{h_1} \boldsymbol{\rho} & \sigma_{12} \boldsymbol{\rho}' \mathbf{z}_{h_1} \mathbf{z}'_{h_2} \boldsymbol{\rho} & \cdots & \sigma_{1n_1} \boldsymbol{\rho}' \mathbf{z}_{h_1} \mathbf{z}'_{h_{n_1}} \boldsymbol{\rho} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n_1 1} \boldsymbol{\rho}' \mathbf{z}_{h_1} \mathbf{z}'_{h_{n_1}} \boldsymbol{\rho} & \sigma_{n_1 2} \boldsymbol{\rho}' \mathbf{z}_{h_{n_1}} \mathbf{z}'_{h_2} \boldsymbol{\rho} & \cdots & \sigma_{n_1 n_1} \boldsymbol{\rho}' \mathbf{z}_{h_{n_1}} \mathbf{z}'_{h_{n_1}} \boldsymbol{\rho} \end{bmatrix}
\end{aligned}$$



la derivata è

$$\begin{aligned} \frac{\partial \mathbf{M} \boldsymbol{\Sigma}_0 \mathbf{M}^*}{\partial \boldsymbol{\rho}} &= 2 \begin{bmatrix} \sigma_{11} \mathbf{z}_{h_1} \mathbf{z}'_{h_1} \boldsymbol{\rho} & \sigma_{12} \mathbf{z}_{h_1} \mathbf{z}'_{h_2} \boldsymbol{\rho} & \cdots & \sigma_{1n_1} \mathbf{z}_{h_1} \mathbf{z}'_{h_{n_1}} \boldsymbol{\rho} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n_1 1} \mathbf{z}_{h_1} \mathbf{z}'_{h_{n_1}} \boldsymbol{\rho} & \sigma_{n_1 2} \mathbf{z}_{h_{n_1}} \mathbf{z}'_{h_2} \boldsymbol{\rho} & \cdots & \sigma_{n_1 n_1} \mathbf{z}_{h_{n_1}} \mathbf{z}'_{h_{n_1}} \boldsymbol{\rho} \end{bmatrix} = \\ &= 2 \{ (\boldsymbol{\Sigma}_0 \otimes \mathbf{1}_{r \times r}) \bullet [\text{vec}(\mathbf{Z}'_h) \text{vec}(\mathbf{Z}'_h)'] \} (\mathbf{I}_{n_1} \otimes \boldsymbol{\rho}) \end{aligned}$$

ma le componenti non sono ordinate, per cui bisogna premoltiplicare il vettore per un'opportuna matrice di permutazione  $\mathbf{P}$ :

$$P = \begin{pmatrix} \mathbf{e}'_1 \\ \mathbf{e}'_{p+1} \\ \vdots \\ \mathbf{e}'_{(n_1-1)p+1} \\ \vdots \\ \mathbf{e}'_{n_1 p} \end{pmatrix}$$

dove  $\mathbf{e}_i \in \mathbb{R}^{n_1}$  rappresenta il vettore la cui  $i$ -esima componente è uno, mentre le altre corrispondono a zero.  $\square$

Anche in questo caso si ricava solo  $\hat{\delta}_0$  analiticamente:

$$\hat{\delta}_0 = \frac{\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} (\mathbf{y}_h - \mathbf{Z}_h \boldsymbol{\rho})}{\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1}}$$

Non potendo ottenere  $\hat{\boldsymbol{\rho}}$  in forma chiusa non si può verificare se è uno stimatore corretto, ma lo è asintoticamente essendo uno stimatore di massima verosimiglianza. Come fatto in precedenza si scrive  $l_{\hat{\delta}_0}$ , e se ne calcola il gradiente.

$$\begin{aligned} l_{\hat{\delta}_0} &= -\frac{n_1}{2} \log 2\pi - \frac{1}{2} \log |\boldsymbol{\Sigma}_h| - \frac{1}{2} \left[ \mathbf{y}_h - \mathbf{Z}_h \boldsymbol{\rho} - \frac{\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} (\mathbf{y}_h - \mathbf{Z}_h \boldsymbol{\rho}) \mathbf{1}_{n_1}}{\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1}} \right]' \boldsymbol{\Sigma}_h^{-1} \cdot \\ &\quad \cdot \left[ \mathbf{y}_h - \mathbf{Z}_h \boldsymbol{\rho} - \frac{\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} (\mathbf{y}_h - \mathbf{Z}_h \boldsymbol{\rho}) \mathbf{1}_{n_1}}{\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1}} \right] = \\ &= -\frac{n_1}{2} \log 2\pi - \frac{1}{2} \log |\boldsymbol{\Sigma}_h| - \frac{1}{2 (\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1})^2} [(\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1} \mathbf{I}_{n_1} - \mathbf{1}_{n_1} \mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1}) \mathbf{y}_h + \\ &\quad + (\mathbf{1}_{n_1} \mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{Z}_h - \mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1} \mathbf{Z}_h) \boldsymbol{\rho}] \boldsymbol{\Sigma}_h^{-1} [(\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1} \mathbf{I}_{n_1} - \mathbf{1}_{n_1} \mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1}) \mathbf{y}_h + \\ &\quad + (\mathbf{1}_{n_1} \mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{Z}_h - \mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1} \mathbf{Z}_h) \boldsymbol{\rho}] \end{aligned}$$

in questo secondo caso:

$$\mathbf{v} = (\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1} \mathbf{I}_{n_1} - \mathbf{1}_{n_1} \mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1}) \mathbf{y}_h + (\mathbf{1}_{n_1} \mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{Z}_h - \mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1} \mathbf{Z}_h) \boldsymbol{\rho}$$

Dopo avere definito, per semplicità nei calcoli,  $\mathbf{v}$  si ricava il gradiente di  $l_{\hat{\delta}_0}$ .

$$\begin{aligned}
\frac{\partial l_{\hat{\delta}_0}}{\partial \sigma_{\delta}^2} &= -\frac{1}{2} \text{tr}(\boldsymbol{\Sigma}_h^{-1} \mathbf{R}_{\delta}) - \frac{1}{2(\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1})^2} \left\{ \frac{2}{\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1}} \mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{R}_{\delta} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1} \mathbf{v} \boldsymbol{\Sigma}_h^{-1} \mathbf{v} + \right. \\
&\quad \left. - 2 \mathbf{v}' \boldsymbol{\Sigma}_h^{-1} [(\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{R}_{\delta} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1} \mathbf{I}_{n_1} - \mathbf{1}_{n_1} \mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{R}_{\delta} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1}) \mathbf{y}_h + \right. \\
&\quad \left. + (\mathbf{1}_{n_1} \mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{R}_{\delta} \boldsymbol{\Sigma}_h^{-1} \mathbf{Z}_h - \mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{R}_{\delta} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1} \mathbf{Z}_h) \boldsymbol{\rho}] - \mathbf{v}' \boldsymbol{\Sigma}_h^{-1} \mathbf{R}_{\delta} \boldsymbol{\Sigma}_h^{-1} \mathbf{v} \right\} \\
\frac{\partial l_{\hat{\delta}_0}}{\partial \sigma_{\varepsilon_h}^2} &= -\frac{1}{2} \text{tr}(\boldsymbol{\Sigma}_h^{-2}) - \frac{1}{2(\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1})^2} \left\{ \frac{2}{\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1}} \mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-2} \mathbf{1}_{n_1} \mathbf{v} \boldsymbol{\Sigma}_h^{-1} \mathbf{v} - 2 \mathbf{v}' \boldsymbol{\Sigma}_h^{-1} \cdot \right. \\
&\quad \cdot [(\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-2} \mathbf{1}_{n_1} \mathbf{I}_{n_1} - \mathbf{1}_{n_1} \mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-2} \mathbf{1}_{n_1}) \mathbf{y}_h + (\mathbf{1}_{n_1} \mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-2} \mathbf{Z}_h - \mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-2} \mathbf{1}_{n_1} \mathbf{Z}_h) \boldsymbol{\rho}] + \\
&\quad \left. - \mathbf{v}' \boldsymbol{\Sigma}_h^{-2} \mathbf{v} \right\} \\
\frac{\partial l_{\hat{\delta}_0}}{\partial \vartheta_i^h} &= -\frac{\sigma_{\delta}^2}{2} \text{tr} \left( \boldsymbol{\Sigma}_h^{-1} \frac{\partial \mathbf{R}_{\delta}}{\partial \vartheta_i^h} \right) - \frac{\sigma_{\delta}^2}{2(\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1})^2} \left\{ \frac{2}{\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1}} \mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \frac{\partial \mathbf{R}_{\delta}}{\partial \vartheta_i^h} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1} \mathbf{v} \boldsymbol{\Sigma}_h^{-1} \mathbf{v} \right. \\
&\quad \left. - 2 \mathbf{v}' \boldsymbol{\Sigma}_h^{-1} \left[ \left( \mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \frac{\partial \mathbf{R}_{\delta}}{\partial \vartheta_i^h} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1} \mathbf{I}_{n_1} - \mathbf{1}_{n_1} \mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \frac{\partial \mathbf{R}_{\delta}}{\partial \vartheta_i^h} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1} \right) \mathbf{y}_h + \right. \right. \\
&\quad \left. \left. + \left( \mathbf{1}_{n_1} \mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \frac{\partial \mathbf{R}_{\delta}}{\partial \vartheta_i^h} \boldsymbol{\Sigma}_h^{-1} \mathbf{Z}_h - \mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \frac{\partial \mathbf{R}_{\delta}}{\partial \vartheta_i^h} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1} \mathbf{Z}_h \right) \boldsymbol{\rho} \right] + \right. \\
&\quad \left. - \sigma_{\delta}^2 \mathbf{v}' \boldsymbol{\Sigma}_h^{-1} \frac{\partial \mathbf{R}_{\delta}}{\partial \vartheta_i^h} \boldsymbol{\Sigma}_h^{-1} \mathbf{v} \right\} \\
\frac{\partial l_{\hat{\delta}_0}}{\partial \boldsymbol{\rho}} &= -\frac{1}{2} \text{tr} \left( \frac{\partial \mathbf{M} \boldsymbol{\Sigma}_0 \mathbf{M}}{\partial \boldsymbol{\rho}} \boldsymbol{\Sigma}_h^{-1} \right) - \frac{1}{2(\mathbf{1}_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1})^2} \left\{ -\frac{2}{\mathbf{1}_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1}} [\mathbf{I}_r \otimes (\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1})] \cdot \right. \\
&\quad \cdot \left( \frac{\partial \mathbf{M} \boldsymbol{\Sigma}_0 \mathbf{M}}{\partial \boldsymbol{\rho}} \boldsymbol{\Sigma}_h^{-1} \right) \mathbf{1}_{n_1} \mathbf{v}' \boldsymbol{\Sigma}_h^{-1} \mathbf{v} - 2 [\mathbf{I}_r \otimes (\mathbf{v}' \boldsymbol{\Sigma}_h^{-1})] \left\{ \left[ \mathbf{I}_r \otimes (\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1}) \cdot \right. \right. \\
&\quad \cdot \left( \frac{\partial \mathbf{M} \boldsymbol{\Sigma}_0 \mathbf{M}}{\partial \boldsymbol{\rho}} \boldsymbol{\Sigma}_h^{-1} \right) \mathbf{1}_{n_1} \otimes \mathbf{I}_{n_1} - \mathbf{I}_r \otimes (\mathbf{1}_{n_1} \mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1}) \left( \frac{\partial \mathbf{M} \boldsymbol{\Sigma}_0 \mathbf{M}}{\partial \boldsymbol{\rho}} \boldsymbol{\Sigma}_h^{-1} \right) \left. \right] \mathbf{y}_h + \\
&\quad \left. - \left[ \mathbf{I}_r \otimes (\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1}) \left( \frac{\partial \mathbf{M} \boldsymbol{\Sigma}_0 \mathbf{M}}{\partial \boldsymbol{\rho}} \boldsymbol{\Sigma}_h^{-1} \mathbf{Z}_h \right) - \mathbf{I}_r \otimes (\mathbf{1}_{n_1} \boldsymbol{\Sigma}_h^{-1}) \left( \frac{\partial \mathbf{M} \boldsymbol{\Sigma}_0 \mathbf{M}}{\partial \boldsymbol{\rho}} \boldsymbol{\Sigma}_h^{-1} \right) \cdot \right. \right. \\
&\quad \left. \left. \cdot \mathbf{1}_{n_1} \otimes \mathbf{Z}_h \right] \boldsymbol{\rho} + \text{vec}(\mathbf{1}_{n_1} \mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{Z}_h - \mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1} \mathbf{Z}_h) \right\} - \mathbf{I}_r \otimes (\mathbf{v}' \boldsymbol{\Sigma}_h^{-1}) \cdot \\
&\quad \left. \cdot \left( \frac{\partial \mathbf{M} \boldsymbol{\Sigma}_0 \mathbf{M}}{\partial \boldsymbol{\rho}} \boldsymbol{\Sigma}_h^{-1} \right) \mathbf{v} \right\}.
\end{aligned}$$

Lo stimatore di massima verisimiglianza è corretto:

$$\begin{aligned}\mathbb{E}(\widehat{\delta}_0 | \mathbf{y}_l, \mathbf{y}_h, \boldsymbol{\Sigma}_l, \boldsymbol{\Sigma}_h) &= \mathbb{E} \left[ \frac{\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} (\mathbf{y}_h - \mathbf{Z}_h \boldsymbol{\rho})}{\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1}} \right] = \\ &= \frac{\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} (\mathbf{Z}_h \boldsymbol{\rho} + \delta_0 \mathbf{1}_{n_1} - \mathbf{Z}_h \boldsymbol{\rho})}{\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1}} = \delta_0\end{aligned}$$

La sua varianza, asintoticamente, risulta:

$$\begin{aligned}\text{Var}(\widehat{\delta}_0 | \mathbf{y}_l, \mathbf{y}_h, \boldsymbol{\Sigma}_l, \boldsymbol{\Sigma}_h) &= \text{Var} \left[ \frac{\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} (\mathbf{y}_h - \mathbf{Z}_h \boldsymbol{\rho})}{\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1}} \right] = \text{Var} \left( \frac{\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{y}_h}{\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1}} \right) = \\ &= \frac{1}{\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1}}\end{aligned}$$

Si conclude che:

$$\widehat{\delta}_0 \approx \mathcal{N} \left( \delta_0, \frac{1}{\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1}} \right)$$

### 3.5.3 Valore atteso e predizione in nuovi punti

Come nel caso a bassa risoluzione precedente si calcola prima il valore atteso in nuovi punti nel caso in cui il parametro  $\delta_0$  è considerato noto. Si definisce  $\widehat{\mu}_h(\mathbf{x}_0) \stackrel{\text{def}}{=} \widehat{\mathbb{E}}(\widehat{\mathbf{y}}_h(\mathbf{x}_0))$ .

$$\begin{aligned}\begin{pmatrix} \mathbf{y}_h \\ \widehat{\mu}_h(\mathbf{x}_0) \end{pmatrix} | \mathbf{y}_l, \mathbf{y}_h, \boldsymbol{\Sigma}_l, \boldsymbol{\Sigma}_h, \delta_0 &\sim \mathcal{N}_{n_1+1} \left( \begin{pmatrix} \mathbf{Z}_h \boldsymbol{\rho} + \delta_0 \mathbf{1}_{n_1} \\ \mathbf{z}'_{h_0} \boldsymbol{\rho} + \delta_0 \end{pmatrix}, \right. \\ &\left. \begin{bmatrix} \boldsymbol{\Sigma}_h & \sigma_\delta^2 \mathbf{r}_{\delta_{12}} + \boldsymbol{\sigma}_0 \\ \sigma_\delta^2 \mathbf{r}'_{\delta_{12}} + \boldsymbol{\sigma}'_0 & \sigma_\delta^2 + \sigma_{00} (\mathbf{z}'_{n_1+1} \boldsymbol{\rho})^2 \end{bmatrix} \right)\end{aligned}$$

Dove  $\mathbf{r}_{\delta_{12}} = \text{Corr}(\mathbf{y}_h, y_h(\mathbf{x}_0))$ ,  $\mathbf{z}'_{h_0}$  rappresenta l'applicazione della funzione  $f_h(\cdot)$  nel generico punto  $\mathbf{x}_0$ ,  $\mathbf{z}'_{h_0} = \mathbf{z}'_{h_0} \widehat{\mu}_l(\mathbf{x}_0)$ ,  $\sigma_{00} = (\mathbf{z}'_{h_0} \boldsymbol{\rho})^2 \cdot \text{Var}[\widehat{\mu}_l(\mathbf{x}_0)]$  ed infine  $\boldsymbol{\sigma}_0 \in \mathbb{R}^{n_1}$  rappresenta il vettore la cui  $i$ -esima componente è  $\sigma_{0i} = (\mathbf{z}'_{h_i} \boldsymbol{\rho})(\mathbf{z}'_{h_0} \boldsymbol{\rho}) \cdot \text{Cov}(\widehat{\mu}_{li}, y_h(\mathbf{x}_0)) \forall i = 1, \dots, n_1$ . Se  $\boldsymbol{\Sigma}_0$  è definito dalla (3.35) basta sostituire alla varianza il valore del MSPE.

Il valore atteso risulta quindi:

$$\begin{aligned}\widehat{\mu}_h(\mathbf{x}_0) | \mathbf{y}_l, \mathbf{y}_h, \boldsymbol{\Sigma}_l, \boldsymbol{\Sigma}_h, \delta_0 &\sim \mathcal{N} \left( \mathbf{z}'_{h_0} \boldsymbol{\rho} + \delta_0 + (\sigma_\delta^2 \mathbf{r}_{\delta_{12}} + \boldsymbol{\sigma}_0)' \boldsymbol{\Sigma}_h^{-1} (\mathbf{y}_h - \mathbf{Z}_h \boldsymbol{\rho} - \delta_0 \mathbf{1}_{n_1}), \right. \\ &\left. \sigma_\delta^2 + \sigma_{00} (\mathbf{z}'_{h_0} \boldsymbol{\rho})^2 - (\sigma_\delta^2 \mathbf{r}_{\delta_{12}} + \boldsymbol{\sigma}_0)' \boldsymbol{\Sigma}_h^{-1} (\sigma_\delta^2 \mathbf{r}_{\delta_{12}} + \boldsymbol{\sigma}_0) \right)\end{aligned}$$

Si aggiunge l'errore casuale del modello ad alta risoluzione per ottenere la predizione in un nuovo punto:

$$\begin{pmatrix} \mathbf{y}_h \\ \hat{y}_h(\mathbf{x}_0) \end{pmatrix} | \mathbf{y}_l, \mathbf{y}_h, \boldsymbol{\Sigma}_l, \boldsymbol{\Sigma}_h, \delta_0 \sim \mathcal{N}_{n_1+1} \left( \begin{pmatrix} \mathbf{Z}\boldsymbol{\rho} + \delta_0 \mathbf{1}_{n_1} \\ \mathbf{z}'_{h_0}\boldsymbol{\rho} + \delta_0 \end{pmatrix}, \begin{bmatrix} \boldsymbol{\Sigma}_h & \sigma_\delta^2 \mathbf{r}_{\delta_{12}} + \boldsymbol{\sigma}_0 \\ \sigma_\delta^2 \mathbf{r}'_{\delta_{12}} + \boldsymbol{\sigma}'_0 & \sigma_\delta^2 + \sigma_{00} (\mathbf{z}'_{h_0}\boldsymbol{\rho})^2 + \sigma_{\varepsilon_h}^2 \end{bmatrix} \right)$$

per cui:

$$\begin{aligned} \hat{y}_h(\mathbf{x}_0) | \mathbf{y}_l, \mathbf{y}_h, \boldsymbol{\Sigma}_l, \boldsymbol{\Sigma}_h, \delta_0 \sim \mathcal{N} \left( \mathbf{z}'_{h_0}\boldsymbol{\rho} + \delta_0 + (\sigma_\delta^2 \mathbf{r}_{\delta_{12}} + \boldsymbol{\sigma}_0)' \boldsymbol{\Sigma}_h^{-1} (\mathbf{y}_h - \mathbf{Z}_h \boldsymbol{\rho} - \delta_0 \mathbf{1}_{n_1}), \right. \\ \left. \sigma_\delta^2 + \sigma_{00} (\mathbf{z}'_{h_0}\boldsymbol{\rho})^2 - (\sigma_\delta^2 \mathbf{r}_{\delta_{12}} + \boldsymbol{\sigma}_0)' \boldsymbol{\Sigma}_h^{-1} (\sigma_\delta^2 \mathbf{r}_{\delta_{12}} + \boldsymbol{\sigma}_0) \right) \end{aligned}$$

Non avendo i veri valori si inserisce la stima di massima verosimiglianza dei parametri, i cui stimatori sono asintoticamente corretti.

Si calcola ora la varianza della media in un punto  $\mathbf{x}_0$  nel caso in cui non sia noto il valore di  $\delta_0$  e  $y_l(\mathbf{x}_0)$ , ma lo sia quello di  $\boldsymbol{\rho}$  poichè compare nella matrice di varianza covarianza di  $\mathbf{y}_h$  e non è possibile avere una sua rappresentazione in forma chiusa.

Si ricava per prima cosa la combinazione lineare di  $\mathbf{y}_h$  in cui è espresso  $\hat{\mu}_h(\mathbf{x}_0)$ :

$$\begin{aligned} \hat{\mu}_h(\mathbf{x}_0) | \mathbf{y}_l, \mathbf{y}_h, \boldsymbol{\Sigma}_l, \boldsymbol{\Sigma}_h &= \hat{\mathbf{z}}'_{h_0}\boldsymbol{\rho} + \hat{\delta}_0 + (\sigma_\delta^2 \mathbf{r}_{\delta_{12}} + \boldsymbol{\sigma}_0)' \boldsymbol{\Sigma}_h^{-1} (\mathbf{y}_h - \hat{\mathbf{Z}}_h \boldsymbol{\rho} - \hat{\delta}_0 \mathbf{1}_{n_1}) = \\ &= \hat{\mathbf{z}}'_{h_0}\boldsymbol{\rho} + \left[ 1 - (\sigma_\delta^2 \mathbf{r}_{\delta_{12}} + \boldsymbol{\sigma}_0)' \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1} \right] \hat{\delta}_0 - (\sigma_\delta^2 \mathbf{r}_{\delta_{12}} + \boldsymbol{\sigma}_0)' \boldsymbol{\Sigma}_h^{-1} \hat{\mathbf{Z}}_h \boldsymbol{\rho} + \\ &\quad + (\sigma_\delta^2 \mathbf{r}_{\delta_{12}} + \boldsymbol{\sigma}_0)' \boldsymbol{\Sigma}_h^{-1} \mathbf{y}_h = \\ &= \left\{ \hat{\mathbf{z}}'_{h_0} - \left[ 1 - (\sigma_\delta^2 \mathbf{r}_{\delta_{12}} + \boldsymbol{\sigma}_0)' \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1} \right] \frac{\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \hat{\mathbf{Z}}_h}{\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1}} - (\sigma_\delta^2 \mathbf{r}_{\delta_{12}} + \boldsymbol{\sigma}_0)' \right. \\ &\quad \cdot \boldsymbol{\Sigma}_h^{-1} \hat{\mathbf{Z}}_h \left. \right\} \boldsymbol{\rho} + \left\{ \left[ 1 - (\sigma_\delta^2 \mathbf{r}_{\delta_{12}} + \boldsymbol{\sigma}_0)' \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1} \right] \frac{\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1}}{\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1}} + \right. \\ &\quad \left. + (\sigma_\delta^2 \mathbf{r}_{\delta_{12}} + \boldsymbol{\sigma}_0)' \boldsymbol{\Sigma}_h^{-1} \right\} \mathbf{y}_h = \\ &= \hat{\mathbf{z}}'_{h_0}\boldsymbol{\rho} - \mathbf{a}'_* \hat{\mathbf{Z}}_h \boldsymbol{\rho} + \mathbf{a}'_* \mathbf{y}_h \end{aligned}$$

con

$$\mathbf{a}'_* = \left[ 1 - (\sigma_\delta^2 \mathbf{r}_{\delta_{12}} + \boldsymbol{\sigma}_0)' \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1} \right] \frac{\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1}}{\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1}} + (\sigma_\delta^2 \mathbf{r}_{\delta_{12}} + \boldsymbol{\sigma}_0)' \boldsymbol{\Sigma}_h^{-1} \quad (3.37)$$

In questo caso lo stimatore è del tipo  $a_0 + \mathbf{a}'_1 \mathbf{y}_h$ , per cui non è BLUP [38] a causa della mancanza di una forma analitica per lo stimatore di  $\boldsymbol{\rho}$ , ma  $\mathbf{a}_*$  è il vettore che minimizza il  $MSEP(\hat{y}_h(\mathbf{x}_0)|\mathbf{y}_l, \mathbf{y}_h, \boldsymbol{\Sigma}_l, \boldsymbol{\Sigma}_h)$ , la dimostrazione è riportata successivamente.

Si è usata la notazione  $\hat{\mathbf{z}}_{h_0}$  e  $\hat{\mathbf{Z}}_h$  per evidenziare che non si ha a disposizione il vero valore di  $y_l(\mathbf{x}_0)$ , e sarà utile nel calcolo del MSPE del predittore nel nuovo punto sperimentale.

Si introduce la seguente proprietà che deve possedere questo stimatore per essere corretto:

$$\begin{aligned} \mathbb{E} \left[ \hat{\mathbf{z}}'_{h_0} \boldsymbol{\rho} - \mathbf{a}' \hat{\mathbf{Z}}_h \boldsymbol{\rho} + \mathbf{a}' \mathbf{y}_h \right] &= \mathbb{E} \left[ \hat{\mathbf{z}}'_{h_0} \boldsymbol{\rho} - \mathbf{a}' \hat{\mathbf{Z}}_h \boldsymbol{\rho} + \mathbf{a}' \hat{\mathbf{Z}}_h \boldsymbol{\rho} + \mathbf{a}' \mathbf{1}_{n_1} \hat{\delta}_0 + \boldsymbol{\delta} + \boldsymbol{\varepsilon}_h \right] = \\ &= \mathbf{z}'_{h_0} \boldsymbol{\rho} + \mathbf{a}' \mathbf{1}_{n_1} \delta_0 = \mathbb{E} [y_h(\mathbf{x}_0)] = \mathbf{z}'_{h_0} \boldsymbol{\rho} + \delta_0 \implies \\ &\implies \mathbf{a}' \mathbf{1}_{n_1} \delta_0 = \delta_0 \implies \mathbf{a}' \mathbf{1}_{n_1} = 1 \end{aligned} \quad (3.38)$$

Si procede al calcolo del MSPE partendo dalla definizione:

$$\begin{aligned} MSPE[\hat{\mu}_h(\mathbf{x}_0)|\mathbf{y}_l, \mathbf{y}_h, \boldsymbol{\Sigma}_l, \boldsymbol{\Sigma}_h] &= \mathbb{E} \left\{ \left[ \hat{\mathbf{z}}'_{h_0} \boldsymbol{\rho} + \mathbf{a}'_* \boldsymbol{\delta} + \underbrace{\mathbf{a}'_* \mathbf{1}_{n_1} \delta_0 - \delta_0}_{\substack{(3.38) \\ =0}} + \right. \right. \\ &\quad \left. \left. - \mathbf{a}'_* \hat{\mathbf{Z}}_h \boldsymbol{\rho} + \mathbf{a}'_* \mathbf{Z}_h \boldsymbol{\rho} + \mathbf{a}'_* \boldsymbol{\varepsilon}_h - \mathbf{z}'_{h_0} \boldsymbol{\rho} - \delta(\mathbf{x}_0) \right]^2 \right\} = \\ &= \mathbb{E} \left\{ \left\{ \hat{\mathbf{z}}'_{h_0} \boldsymbol{\rho} [\hat{\mu}_l(\mathbf{x}_0) - y_l(\mathbf{x}_0)] - \mathbf{a}'_* \mathbf{M} (\hat{\boldsymbol{\mu}}_l - \mathbf{y}_l) + \mathbf{a}'_* \boldsymbol{\delta} + \mathbf{a}'_* \boldsymbol{\varepsilon}_h + \right. \right. \\ &\quad \left. \left. - \delta(\mathbf{x}_0) \right\}^2 \right\} = \\ &= \mathbb{E} \left\{ (\hat{\mathbf{z}}'_{h_0} \boldsymbol{\rho})^2 [\hat{\mu}_l(\mathbf{x}_0) - y_l(\mathbf{x}_0)]^2 + \mathbf{a}'_* \mathbf{M} (\hat{\boldsymbol{\mu}}_l - \mathbf{y}_l) (\hat{\boldsymbol{\mu}}_l - \mathbf{y}_l)' \mathbf{M} \mathbf{a}_* + \mathbf{a}'_* \boldsymbol{\delta} \boldsymbol{\delta}' \mathbf{a}_* + \right. \\ &\quad + \mathbf{a}'_* \boldsymbol{\varepsilon}_h \boldsymbol{\varepsilon}'_h \mathbf{a}_* + \delta^2(\mathbf{x}_0) - 2 \hat{\mathbf{z}}'_{h_0} \boldsymbol{\rho} [\hat{\mu}_l(\mathbf{x}_0) - y_l(\mathbf{x}_0)] (\hat{\boldsymbol{\mu}}_l - \mathbf{y}_l)' \mathbf{M} \mathbf{a}_* \\ &\quad + 2 \hat{\mathbf{z}}'_{h_0} \boldsymbol{\rho} [\hat{\mu}_l(\mathbf{x}_0) - y_l(\mathbf{x}_0)] \mathbf{a}_* \boldsymbol{\delta} + 2 \hat{\mathbf{z}}'_{h_0} \boldsymbol{\rho} [\hat{\mu}_l(\mathbf{x}_0) - y_l(\mathbf{x}_0)] \mathbf{a}'_* \boldsymbol{\varepsilon}_h + \\ &\quad - 2 \hat{\mathbf{z}}'_{h_0} \boldsymbol{\rho} [\hat{\mu}_l(\mathbf{x}_0) - y_l(\mathbf{x}_0)] \delta(\mathbf{x}_0) - 2 \mathbf{a}'_* \mathbf{M} (\hat{\boldsymbol{\mu}}_l - \mathbf{y}_l) \mathbf{a}'_* \boldsymbol{\delta} + \\ &\quad + 2 \mathbf{a}'_* \mathbf{M} (\hat{\boldsymbol{\mu}}_l - \mathbf{y}_l) \delta(\mathbf{x}_0) - 2 \mathbf{a}'_* \mathbf{M} (\hat{\boldsymbol{\mu}}_l - \mathbf{y}_l) \mathbf{a}'_* \boldsymbol{\varepsilon}_h + 2 \mathbf{a}'_* \boldsymbol{\delta} \boldsymbol{\varepsilon}'_h \mathbf{a}_* - 2 \mathbf{a}'_* \boldsymbol{\delta} \delta(\mathbf{x}_0) + \\ &\quad \left. - 2 \mathbf{a}'_* \boldsymbol{\varepsilon}_h \delta(\mathbf{x}_0) \right\} = \\ &= (\hat{\mathbf{z}}'_{h_0} \boldsymbol{\rho})^2 \sigma_{00} + \mathbf{a}'_* \mathbf{M} \boldsymbol{\Sigma}_0 \mathbf{M} \mathbf{a}_* + \sigma_\delta^2 \mathbf{a}'_* \mathbf{R}_\delta \mathbf{a}_* + \sigma_{\varepsilon_h}^2 \mathbf{a}'_* \mathbf{a}_* + \sigma_\delta^2 - 2 \mathbf{a}'_* \boldsymbol{\sigma}_0 - 2 \sigma_\delta^2 \mathbf{a}'_* \mathbf{r}_{\delta_{12}} = \\ &= (\hat{\mathbf{z}}'_{h_0} \boldsymbol{\rho})^2 \sigma_{00} + \sigma_\delta^2 + \mathbf{a}'_* (\mathbf{M} \boldsymbol{\Sigma}_0 \mathbf{M} + \sigma_\delta^2 \mathbf{R}_\delta + \sigma_{\varepsilon_h}^2 \mathbf{I}_{n_1}) \mathbf{a}_* - 2 \mathbf{a}'_* (\sigma_\delta^2 \mathbf{r}_{\delta_{12}} + \boldsymbol{\sigma}_0) = \\ &= (\hat{\mathbf{z}}'_{h_0} \boldsymbol{\rho})^2 \sigma_{00} + \sigma_\delta^2 + \mathbf{a}'_* \boldsymbol{\Sigma}_h \mathbf{a}_* - 2 \mathbf{a}'_* (\sigma_\delta^2 \mathbf{r}_{\delta_{12}} + \boldsymbol{\sigma}_0). \end{aligned}$$

**Osservazione 3.5.1.** Se si è utilizzato  $\Sigma_0$  definito nella (3.34), siccome  $\beta$ , e quindi anche la media, è considerata nota sia ha che  $\mathbb{E} [\hat{\mu}_l(\mathbf{x}_0) - y_l(\mathbf{x}_0)]^2 = \text{Var} [\hat{\mu}_l(\mathbf{x}_0)|\mathbf{y}_l, \Sigma_l]$ , mentre se il  $\Sigma_0$  utilizzato è definito dalla (3.35),  $\mathbb{E} [\hat{\mu}_l(\mathbf{x}_0) - y_l(\mathbf{x}_0)]^2 = \text{MSPE} [\hat{\mu}_l(\mathbf{x}_0)|\mathbf{y}_l, \Sigma_l]$ .

Il MSPE calcolato precedentemente è asintoticamente corretto infatti, facendo tendere  $n_1$  all'infinito,  $\mathbf{a}_*$ , riportato nella (3.37) tende a:

$$\mathbf{a}_* \rightarrow (\sigma_\delta^2 \mathbf{r}_{\delta_{12}} + \boldsymbol{\sigma}_0)' \Sigma_h^{-1} \quad (3.39)$$

poichè  $\frac{1}{\mathbf{1}'_{n_1} \Sigma_h^{-1} \mathbf{1}_{n_1}} = \text{Var}(\hat{\delta}_0) \rightarrow 0$  perché lo stimatore di  $\delta_0$  deriva dalla massimizzazione della verosimiglianza. Per cui:

$$\begin{aligned} \text{MSPE}[\hat{\mu}_h(\mathbf{x}_0)|\mathbf{y}_l, \mathbf{y}_h, \Sigma_l, \Sigma_h] &\rightarrow (\tilde{\mathbf{z}}'_{h_0} \boldsymbol{\rho})^2 \sigma_{00} + \sigma_\delta^2 + (\sigma_\delta^2 \mathbf{r}_{\delta_{12}} + \boldsymbol{\sigma}_0)' \\ &\cdot \Sigma_h^{-1} (\sigma_\delta^2 \mathbf{r}_{\delta_{12}} + \boldsymbol{\sigma}_0) - 2 \sigma_\delta^2 (\sigma_\delta^2 \mathbf{r}_{\delta_{12}} + \boldsymbol{\sigma}_0)' \Sigma_h^{-1} (\sigma_\delta^2 \mathbf{r}_{\delta_{12}} + \boldsymbol{\sigma}_0) \end{aligned} \quad (3.40)$$

che è quello che si voleva dimostrare.

Il MSPE della predizione in un nuovo punto sperimentale  $y_h(\mathbf{x}_0)$  si calcola aggiungendo  $\sigma_{\varepsilon_h}^2$  a  $\text{MSPE}[\hat{\mu}_h(\mathbf{x}_0)|\mathbf{y}_l, \mathbf{y}_h, \Sigma_l, \Sigma_h]$  per l'indipendenza tra  $\varepsilon_h(\mathbf{x}_0)$  e  $\hat{\mu}_h(\mathbf{x}_0)$ :

$$\begin{aligned} \text{MSPE}[\hat{y}_h(\mathbf{x}_0)|\mathbf{y}_l, \mathbf{y}_h, \Sigma_l, \Sigma_h] &= \sigma_{\varepsilon_h}^2 + (\tilde{\mathbf{z}}'_{h_0} \boldsymbol{\rho})^2 \sigma_{00} + \sigma_\delta^2 + \\ &+ \mathbf{a}'_* \Sigma_h \mathbf{a}_* - 2 \mathbf{a}'_* (\sigma_\delta^2 \mathbf{r}_{\delta_{12}} + \boldsymbol{\sigma}_0). \end{aligned} \quad (3.41)$$

Si dimostra ora che  $\mathbf{a}_*$  minimizza il  $\text{MSEP}(\hat{y}_h(\mathbf{x}_0)|\mathbf{y}_l, \mathbf{y}_h, \Sigma_l, \Sigma_h)$ .

La funzione da minimizzare è

$$\begin{aligned} \min_{\mathbf{a}} f(\mathbf{a}) &= \min_{\mathbf{a}} \mathbf{a}' \Sigma_h \mathbf{a} - 2 \mathbf{a}' (\sigma_\delta^2 \mathbf{r}_{\delta_{12}} + \boldsymbol{\sigma}_0) \\ &\text{soggetto a} \\ \mathbf{a}' \mathbf{1}_{n_1} &= 1 \end{aligned}$$

in questa funzione non compaiono i termini  $(\tilde{\mathbf{z}}'_{h_0} \boldsymbol{\rho})^2 \sigma_{00}$ ,  $\sigma_\delta^2$  e  $\sigma_{\varepsilon_h}^2$  poichè non influenzano la minimizzazione non dipendendo da  $\mathbf{a}$ .

Per l'ottimizzazione si utilizza il metodo dei moltiplicatori di Lagrange. La funzione da minimizzare diventa:

$$\min_{\mathbf{a}, \lambda} L(\mathbf{a}, \lambda) = \min_{\mathbf{a}, \lambda} \mathbf{a}' \Sigma_h \mathbf{a} - 2 \mathbf{a}' (\sigma_\delta^2 \mathbf{r}_{\delta_{12}} + \boldsymbol{\sigma}_0) + 2\lambda(\mathbf{a}' \mathbf{1}_{n_1} - 1)$$

il gradiente risulta:

$$\begin{aligned}\frac{\partial L(\mathbf{a}, \lambda)}{\partial \mathbf{a}} &= 2\boldsymbol{\Sigma}_h \mathbf{a} - 2(\sigma_\delta^2 \mathbf{r}_{\delta_{12}} + \boldsymbol{\sigma}_0) + 2\lambda \mathbf{1}_{n_1} \\ \frac{\partial L(\mathbf{a}, \lambda)}{\partial \lambda} &= \mathbf{a}' \mathbf{1}_{n_1} - 1\end{aligned}$$

Uguagliando il gradiente al vettore nullo e risolvendo il sistema lineare si ricava il punto stazionario:

$$\bar{\mathbf{a}} = \left( \boldsymbol{\Sigma}_h^{-1} - \frac{\boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1} \mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1}}{\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1}} \right) (\sigma_\delta^2 \mathbf{r}_{\delta_{12}} + \boldsymbol{\sigma}_0) + \frac{\boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1}}{\mathbf{1}'_{n_1} \boldsymbol{\Sigma}_h^{-1} \mathbf{1}_{n_1}}$$

Il punto stazionario  $\bar{\mathbf{a}}$  corrisponde al vettore  $\mathbf{a}_*$  ed essendo l'hessiana di  $f(\mathbf{a})$  pari a  $2\boldsymbol{\Sigma}_h$ , che è definita positiva, per cui  $\mathbf{a}_*$  è il minimo del  $MSEP(\hat{y}_h(\mathbf{x}_0) | \mathbf{y}_l, \mathbf{y}_h, \boldsymbol{\Sigma}_l, \boldsymbol{\Sigma}_h)$ .





## Capitolo 4

# Valutazione prestazioni modello gerarchico

### 4.1 Introduzione

Dopo aver presentato l'ampliamento del modello di Qian *et al.* si valutano le prestazioni di questo modello gerarchico confrontandole con quelle che si ottengono per mezzo di un modello in cui si considerano i soli punti a bassa risoluzione, un modello in cui si utilizzano i soli punti ad alta risoluzione ed un modello non gerarchico in cui i punti a bassa ed alta risoluzione sono trattati come se provenissero tutti dallo strumento di misura.

Nella seconda parte del capitolo si presenta un altro modello gerarchico, in cui i punti a bassa risoluzione, prima di essere processati, vengono decimati mediante una metrica quadratica. L'utilizzo di una decimazione di punti è necessario quando le nuvole di punti provenienti da uno strumento di misura hanno una dimensione elevata, tale da non rendere possibile il processamento di un numero così elevato di punti. I risultati di questi modelli, a bassa risoluzione e *fusion*, saranno confrontati con quelli dei modelli a bassa risoluzione e *fusion* di riferimento, in cui stime derivano dal set di dati completo.

Per l'analisi di un caso reale si è considerata l'acquisizione della superficie *free form* raffigurata in figura 4.1.

Cavallaro, Moroni e Petró in [9] hanno utilizzato i dati per:

- valutare le prestazioni derivanti da due metodi di calibrazione per lo scanner a luce strutturata;
- valutare le prestazioni di misura di uno scanner a luce strutturata sviluppato dal Dipartimento di Meccanica del Politecnico di Milano.

La superficie è stata acquisita mediante lo scanner a luce strutturata, che rappresentano i punti a bassa risoluzione, mentre i punti ad alta risoluzione sono

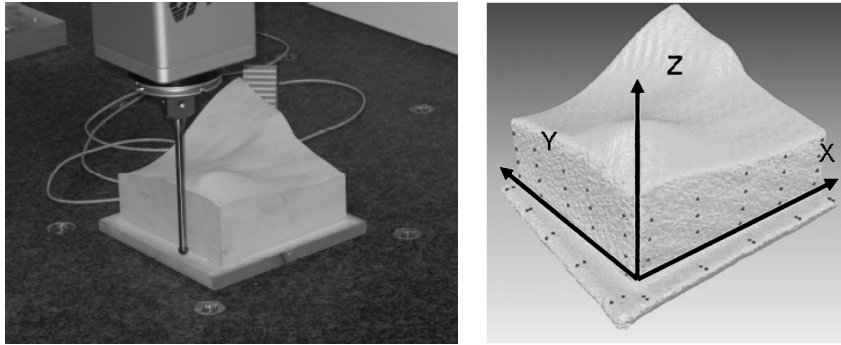


Figura 4.1: Artefatto fisico

stati ottenuti mediante una CMM. Questa macchina di misura ha un *Maximum Permissible Error for Probing* ( $MPE_P$ ), definito nella normativa ISO 10360-2, pari a  $2 \mu m$ .

I punti a bassa risoluzione si riferiscono al sistema di coordinate definito dalle tre superfici ortogonali raffigurate in figura 4.1, in modo da poter misurare gli stessi punti anche con la CMM.

Sebbene ci sia stato un allineamento delle nuvole di punti i due sistemi di misura non hanno un sistema di coordinate comune per cui gli autori hanno allineato le due nuvole di punti mediante una variante dell'algoritmo *Iterative Closest Point* (ICP) presente nel software Rapidform.

Nella figura 4.2 sono presenti i punti a bassa risoluzione acquisiti mediante uno scanner a luce struttura.

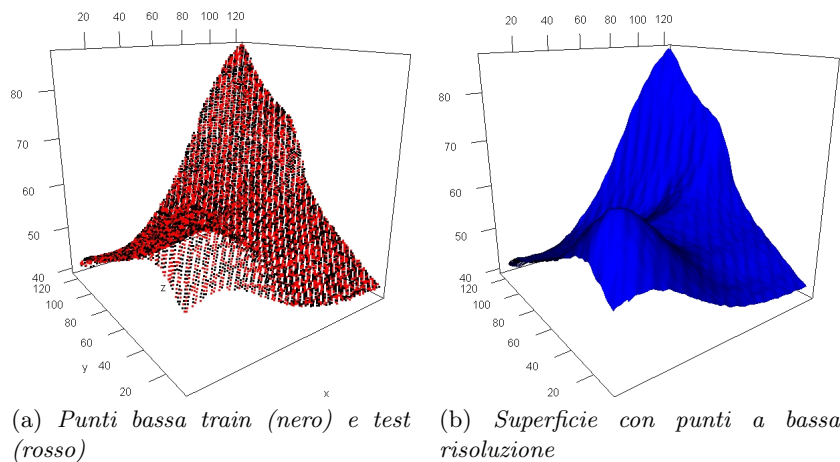


Figura 4.2: Punti a bassa risoluzione

Per poter valutare la validità delle stime che derivano dai modelli analizzati,

la nuvola di punti (composta da 9.635 punti) è stata suddivisa in due insiemi: il primo utilizzato per la stima dei parametri (punti neri), mentre il secondo (punti rossi) come set di dati per testare i risultati. Il primo insieme è formato da un totale di 4.818 punti, da cui sono stati estratti i punti ad elevata accuratezza; mentre il secondo dai rimanenti che corrispondono a 4.817 elementi. Una ricostruzione della superficie ottenuta considerando i punti acquisiti senza la rimozione dell'errore casuale (figura 4.2b), evidenzia come sia molto ondulata.

## 4.2 Modelli a bassa risoluzione

Data la complessità della superficie analizzata è opportuno utilizzare un modello in cui sia presente un processo gaussiano (riportato nella (4.1)); in tale modello è presente un termine lineare poichè la superficie rappresentata dalla nuvola di punti sembra avere un andamento crescente.

$$y_l(\mathbf{x}_i) = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \eta(\mathbf{x}_i) + \varepsilon_l \quad (4.1)$$

I parametri stimati grazie al modello a bassa risoluzione sono riportati in tabella 4.1.

Parametro	Valore stima
$\beta_0$	34,5120
$\beta_1$	0,1377
$\beta_2$	0,1562
$\sigma_\eta^2$	39,4723
$\vartheta_1^l$	0,0006
$\vartheta_2^l$	0,0071
$\sigma_{\varepsilon_l}^2$	0,0183

Tabella 4.1: Parametri modello bassa risoluzione

I punti appartenenti al *training set* vengono ‘spiegati’ per la maggior parte dall’intercetta ( $\beta_0$ ) e dai parametri appartenenti al processo gaussiano (figura 4.3b), dato il grande valore assunto dalla stima di  $\sigma_\eta^2$ . La stima di massima verosimiglianza dei parametri del kernel gaussiano ( $\vartheta_1^l$  e  $\vartheta_2^l$ ) presenta dei valori bassi; questo indica un’alta connessione tra punti vicini e, poichè essi differiscono per un ordine di grandezza, la correlazione tra i punti varia a seconda della direzione (lungo l’asse  $x$  oppure  $y$ ).

Risultano significativi, anche se presentano un valore basso, i coefficienti dei termini lineari (figura 4.3a); questi, infatti, permettono di attenuare l’effetto del processo gaussiano portando ad una minore incertezza attorno al valore predetto. Inserendo il valore degli stimatori nel modello, si ottiene la ricostruzione della superficie presente in figura 4.4; si nota subito la differenza rispetto a quella ottenuta tramite l’utilizzo dei soli punti derivanti dallo scanner a luce strutturata

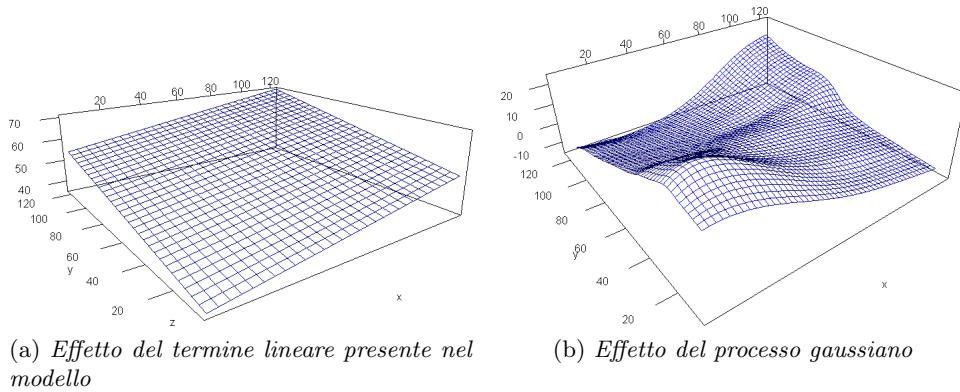


Figura 4.3: Termini che compongono il modello

poichè, eliminando l'errore casuale, la superficie risulta più liscia non dovendo interpolare tutti i punti rumorosi.

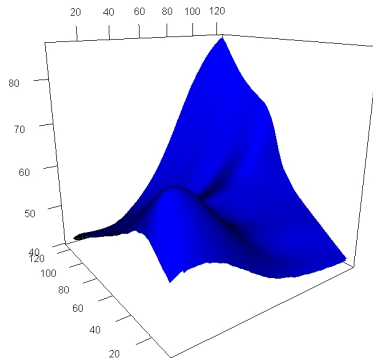


Figura 4.4: Superficie predetta

Per verificare i punti in cui le superfici rappresentate nelle figure 4.2 e 4.4 differiscono, si riporta la superficie fittata dal modello a bassa risoluzione (figura 4.5) in cui ad ogni colore è associato il valore assunto dai residui in un determinato punto.

Dall'analisi della superficie 4.5a e dall'immagine 4.5b si conclude che la differenza tra le due superfici è dovuta all'eliminazione delle ondulazioni derivanti dai punti rilevati. Infatti nel grafico 3D (figura 4.6) si nota che i residui sono casualmente distribuiti nello spazio eccetto in un'unica zona corrispondente alla parte più alta della superficie. In blu sono raffigurate due superfici che rappresentano l'ampiezza degli intervalli di confidenza al 95% sulla media del valore predetto, mentre in verde è disegnata una superficie derivante dall'ampiezza dell'intervallo di predizione con lo stesso valore di errore di primo tipo; i residui compresi in quest'ultimo intervallo sono il 94,56% su un totale di 4.818 unità, quindi vi è una buona stima della variabilità intorno al valore puntuale predetto.

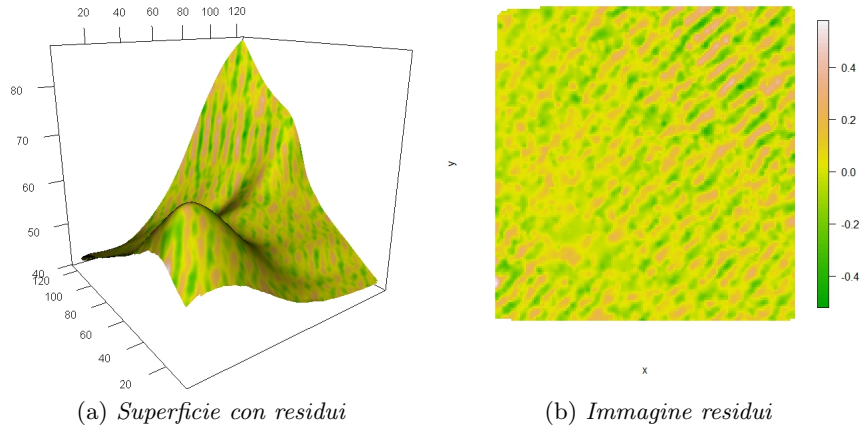


Figura 4.5: Residui modello

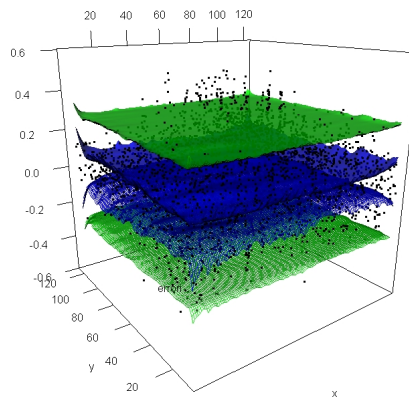


Figura 4.6: Residui con intervalli di confidenza (blu) e di predizione (verde)

Si riporta infine la superficie che deriva dalle differenze tra i valori dell'intervallo di confidenza superiore e il valore stimato dal modello ed una sua immagine (figura 4.7). È possibile evidenziare che per buona parte dell'artefatto la ampiezza del semi-intervallo di confidenza risulta inferiore ai cinque centesimi di millimetro. Da notare un effetto non eliminabile causato dalla mancanza di punti da utilizzare per la stima dei parametri del modello si ha in corrispondenza dei bordi del piano sperimentale, in cui è presente una grande incertezza attorno al valore predetto; tale fenomeno prende il nome di effetto di bordo.

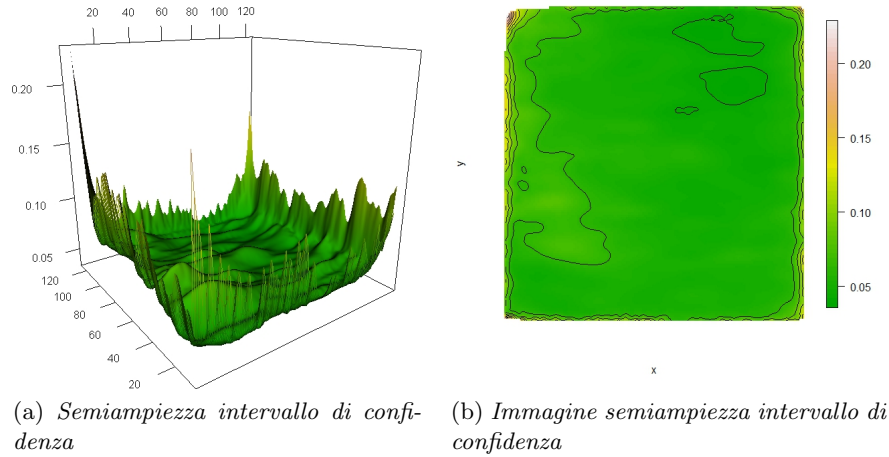


Figura 4.7: Semiampiezza intervallo di confidenza e immagine

### 4.3 Modelli con punti ad alta risoluzione

In questa sezione si presenteranno i modelli ottenuti grazie all'utilizzo dei punti del *training set* che provengono da misure eseguite tramite la macchina di misura a coordinate. Si è deciso di non utilizzare la totalità del data set poiché il tempo necessario per acquisire un numero così elevato di punti sarebbe molto elevato.

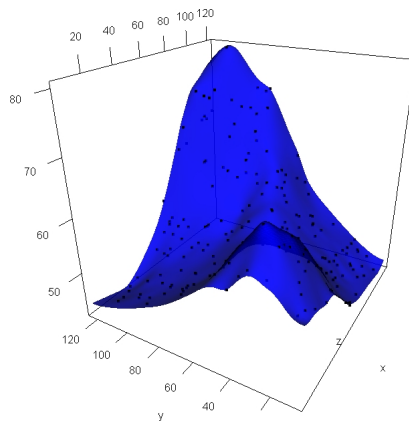


Figura 4.8: Punti alta accuratezza campionati

È stato estratto casualmente un sottoinsieme di 200 punti che serviranno per la stima dei parametri dei modelli analizzati. I punti campionati sono presenti in figura 4.8: in blu è rappresentata la ricostruzione della superficie ottenuta

dai soli punti ad alta risoluzione; il risultato che si ottiene con un numero così limitato di punti è peggiore rispetto all'utilizzo dell'insieme dei punti a bassa accuratezza.

### 4.3.1 Modello alta risoluzione

In questo primo modello non vengono utilizzate le informazioni che derivano dai punti acquisiti con la luce strutturata, ma per la stima dei parametri del modello (4.3) vengono utilizzati solo i punti ad elevata accuratezza disponibili.

$$y_h(\mathbf{x}_i) = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \eta(\mathbf{x}_i) + \varepsilon_h \quad (4.2)$$

Il modello utilizzato è lo stesso della stima dei parametri dei modelli a bassa risoluzione poichè i punti ad alta e bassa risoluzione provenienti dalla stessa superficie, hanno lo stesso andamento. Le stesse considerazioni fatte in precedenza sulle stime dei parametri del modello possono essere fatte anche in questo caso (tabella 4.2), ad eccezione della stima della varianza dell'errore casuale che risulta inferiore di un ordine di grandezza rispetto a quella stimata attraverso i dati provenienti dallo scanner a luce strutturata. Tale differenza è da ricondursi alla differente accuratezza dei due strumenti di misura utilizzati.

Parametro	Valore stima
$\beta_0$	36,6727
$\beta_1$	0,1254
$\beta_2$	0,1358
$\sigma_\eta^2$	41,8094
$\vartheta_1$	0,0005
$\vartheta_2$	0,0050
$\sigma_\varepsilon^2$	0,0016

Tabella 4.2: Parametri modello alta risoluzione

In figura 4.9 si riportano i grafici che riguardano il valore della semiampiezza dell'intervallo di confidenza; l'effetto di bordo risulta essere molto amplificato se paragonato a quello che si ha nei due modelli a bassa risoluzione precedenti. Questo è causato dalla differenza tra il numero di punti ad alta e bassa accuratezza, la quale provoca anche un aumento del valore che l'incertezza di predizione ha all'interno del dominio. L'intervallo di predizione, in questo caso, sovrastima la bontà dei dati poichè all'interno dell'intervallo di predizione, con un  $\alpha$  pari al 5% sono presenti solo il 91,39% delle osservazioni.

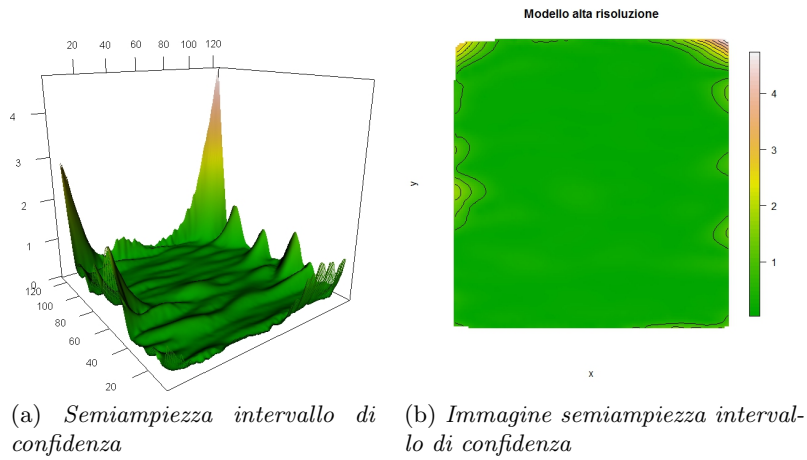


Figura 4.9: Intervallo di confidenza sulla media

### 4.3.2 Modello addition

È stato inoltre testato un modello (denominato *addition*) in cui in cui il *training set* è formato dall'unione dei punti a bassa ed alta risoluzione senza tenere in considerazione la differente risoluzione dei punti. Il modello utilizzato è lo stesso presente nella (4.1).

I risultati ottenuti, sia come stima dei parametri, che in termini di valore dei residui sono molto simili a quelli ottenuti con il modello a bassa risoluzione poichè il set di dati ad alta risoluzione è molto più piccolo di quello a bassa risoluzione, circa il 40%.

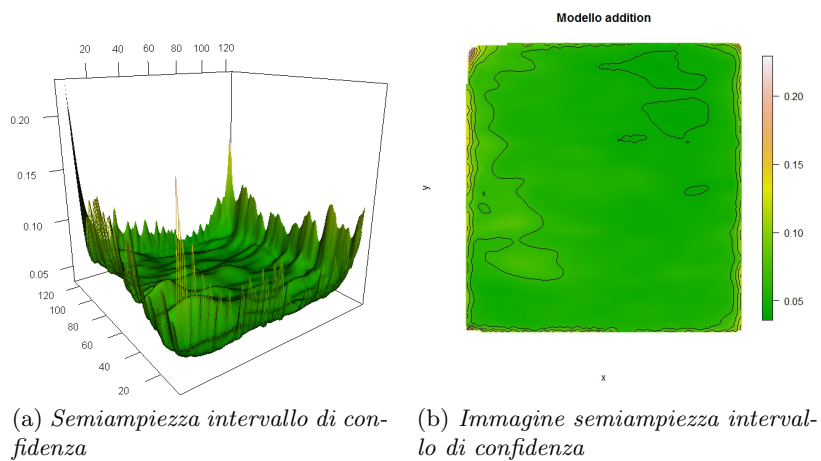


Figura 4.10: Intervallo di confidenza sulla media



Si riportano i solo grafici relativi al valore della semiampiezza dell'intervallo di confidenza sulla media (figura 4.10), in cui si nota l'uguaglianza con quello ottenuto dal modello a bassa risoluzione (figura 4.7). Nel caso in cui non ci siano grandi errori nei punti a bassa risoluzione e fossero presenti delle zone della superficie in cui mancano i punti a bassa risoluzione, questa potrebbe essere una buona tecnica per la fusione dei due set di dati, considerando i pessimi risultati ottenuti nella confidenza sul valor medio stimato ottenuti con un modello in cui si sono utilizzati i soli punti ad alta risoluzione.

### 4.3.3 Modello fusion

I punti ad alta risoluzione campionati vengono utilizzati per stimare i parametri del modello di secondo stadio riportato nella (4.3). Non sono presenti parametri per la correzione spaziale dell'errore dello scanner a luce strutturata, ad eccezione del parametro di scala  $\rho_0$  costante, poichè non sono disponibili informazioni che permettono di aggiungere altri parametri. In questa parte la stima del valore atteso dei punti a bassa risoluzione,  $\widehat{\mu}_l(\mathbf{x}_i)$ , viene fatta tramite le stime ottenute su tutti i punti appartenenti al dataset *train* della nuvola di punti a bassa accuratezza.

$$y_h(\mathbf{x}_i) = \widehat{\mu}_l(\mathbf{x}_i)\rho_0 + \delta_0 + \delta(\mathbf{x}_i) + \varepsilon_h(\mathbf{x}_i) \quad (4.3)$$

Parametro	Valore stima
$\rho_0$	0,9986
$\delta_0$	0,0922
$\sigma_\delta^2$	0,0434
$\vartheta_1^h$	0,0018
$\vartheta_2^h$	0,0242
$\sigma_{\varepsilon_h}^2$	$6,0947 \cdot 10^{-4}$

Tabella 4.3: Stime modello alta risoluzione

La ‘correzione’ dei punti a bassa accuratezza (stime dei parametri del modello fusion, tabella 4.3) viene effettuata per la maggior parte dall'azione congiunta del parametro di shift  $\delta_0$  e dal parametro di scala  $\rho_0$ . Il fattore di scala moltiplica tutti i valori per una quantità minore di uno assegnando ai punti un valore inferiore di quello predetto dal modello a bassa risoluzione, mentre il parametro di shift, essendo positivo, esegue una traslazione parallela all'asse  $z$  riportando i valori più vicini a quelli predetti tramite i punti a luce strutturata. Questi due effetti non si compensano poichè la scala agisce in modo proporzionale al valore assunto dal punto, mentre il processo di traslazione non ne è influenzato. La stima della varianza in questo caso ha un valore inferiore rispetto alle stime di tutti i modelli precedentemente analizzati, è infatti pari circa ad un terzo di

quella stimata tramite i soli punti ad elevata accuratezza grazie alla bontà delle stime dei punti ottenute per mezzo di un modello di primo stadio. Questo valore è maggiore rispetto a quello della CMM, pari a  $4 \cdot 10^{-6}$  come riportato in [9], poichè si ricorda che la stima non è stata eseguita sui soli punti provenienti da misure fatte grazie alla macchina di misura a coordinate, ma viene anche inserita l'incertezza di uno strumento meno accurato come la luce strutturata ed inoltre il numero di punti risulta limitato considerando la complessità della superficie analizzata per effettuare una buona stima dell'errore casuale associato allo strumento.

Il processo gaussiano risulta avere un'incidenza minore rispetto ai casi precedenti (figura 4.11). È possibile notare che questo ha valori assoluti elevati, attorno all'unità, solo in alcune zone della superficie; una possibile spiegazione è la presenza in quelle zone di un errore sistematico dei dati a bassa risoluzione, causato da una lettura non accurata dello scanner a luce strutturata oppure alla penalizzazione di queste zone rispetto ad altre da parte dell'algoritmo di registrazione utilizzato dagli autori per l'allineamento delle nuvole di punti.

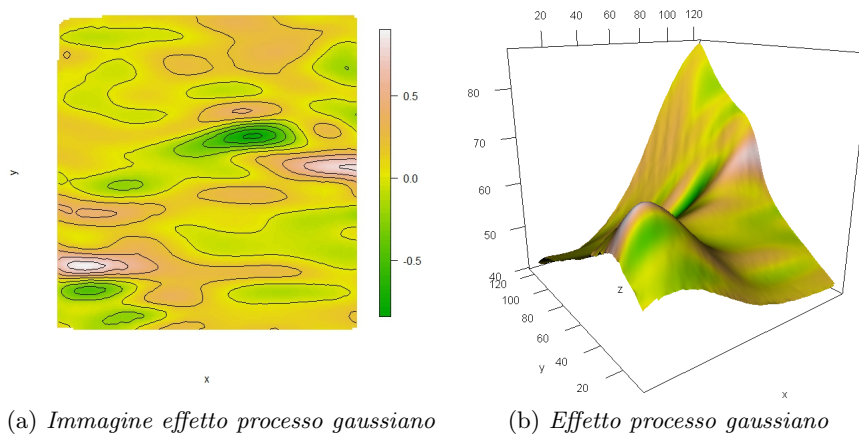


Figura 4.11: Effetto processo gaussiano

La correzione di alcune zone della superficie rispetto ad altre causa delle variazioni nell'incertezza attorno al valore predetto (figura 4.12), infatti il valore della semiampiezza dell'intervallo di confidenza non è più uniforme sull'intera superficie, ad esclusione dei bordi, ma presenta dei picchi all'interno nei punti dove la correzione è stata eseguita dal processo gaussiano. Questo fenomeno è dovuto alla discrepanza tra il valore attribuito all'artefatto fisico nel punto  $\mathbf{x}_i$  dallo strumento meno preciso rispetto dello strumento che commette un errore minore.

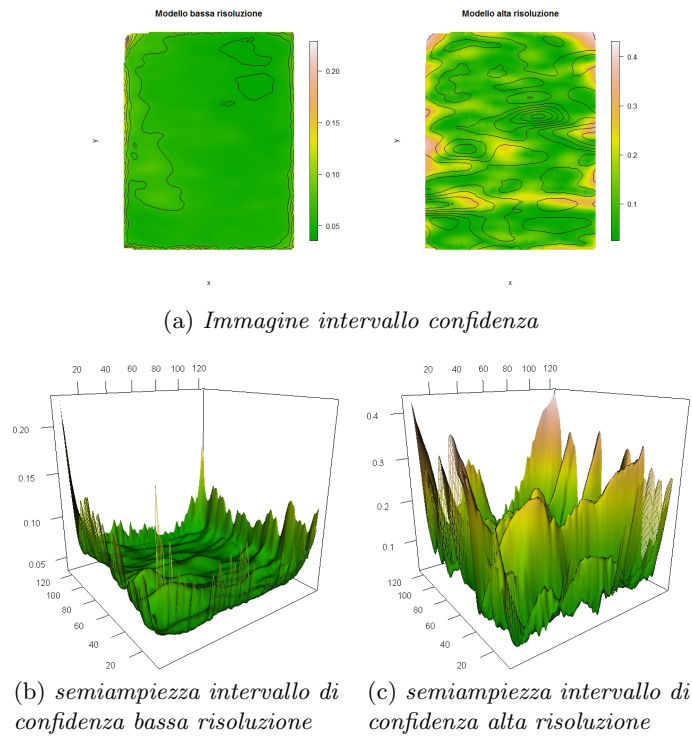


Figura 4.12: Semi-ampiezze intervalli di confidenza

## 4.4 Analisi residui

In questa sezione verranno confrontati i risultati ottenuti dai vari modelli attraverso statistiche sui dataset formati dai residui. Questi sono calcolati come la differenza tra il valore ‘vero’, ottenuto tramite la macchina di misura a coordinate e quello predetto dallo specifico modello. Questo approccio implica però un vantaggio per i punti ad alta risoluzione, ma non essendo possibile avere a disposizione i veri punti della superficie, quelli provenienti dalla macchina di misura a contatto sono quelli che più si avvicinano alla realtà inoltre non può essere utilizzato il modello CAD a causa degli errori introdotti dalla lavorazione di fresatura.

Nella prima parte si illustrano dei grafici che riportano la ricostruzione della superficie grazie alla predizione dei punti test, i cui colori indicano il valore assoluto dei residui in un particolare punto dello spazio.

I residui del modello, i cui parametri sono stimati dai soli punti ad elevata accuratezza, (figura 4.13) sono centrati sul valore di zero e le maggiori differenze rispetto ai punti acquisiti con la CMM si hanno dove la superficie di confidenza presenta un picco, cioè nella zona situata vicino alla frontiera della regione sperimentale e in quelle parti in cui i punti ad alta risoluzione non sono sufficienti per poter realizzare una stima accurata dei parametri.

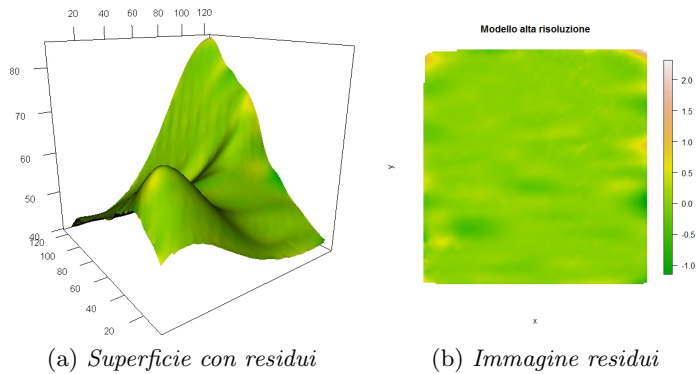
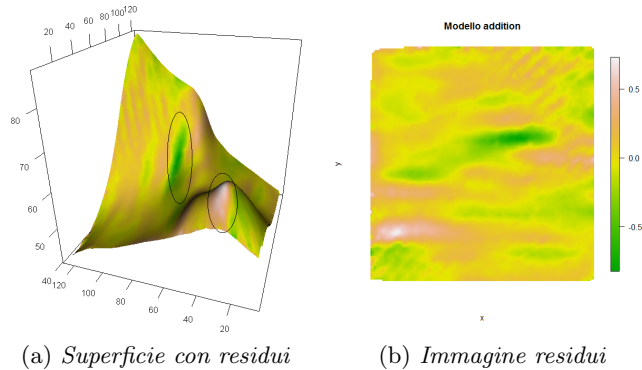


Figura 4.13: Residui modello alta risoluzione

I residui del modello *addition* (figura 4.14) presentano due zone in cui assumono valori elevati, evidenziati nel grafico di sinistra. Questo errore, che non si ha nel modello ad alta risoluzione, è dipendente dalla posizione e come si vedrà in seguito, è tipico dei dati a bassa accuratezza. La presenza di questi errori indica che i punti ad elevata accuratezza, avendo lo stesso peso di quelli a bassa precisione, non sono riusciti a correggerli, per cui questo metodo di integrazione dei dati non è valido per la superficie oggetto dell'analisi.

Figura 4.14: Residui modello *addition*

Nella figura 4.15 sono riportate le immagini e le superfici in cui viene evidenziato il valore dei residui nei modelli a bassa risoluzione (colonna di sinistra) e *fusion* (colonna di destra). Il modello a bassa risoluzione commette degli errori sistematici nella ricostruzione della superficie, soprattutto nelle zone evidenziate nel grafico 4.15b. Le zone colorate di verde attribuiscono un'altezza maggiore di quella che è in realtà, mentre nella zone di colore bianco si verifica una sottostima del valore della funzione. Nel passaggio al modello *fusion* questi errori vengono corretti ed i residui sono casualmente distribuiti lungo il dominio della superficie.

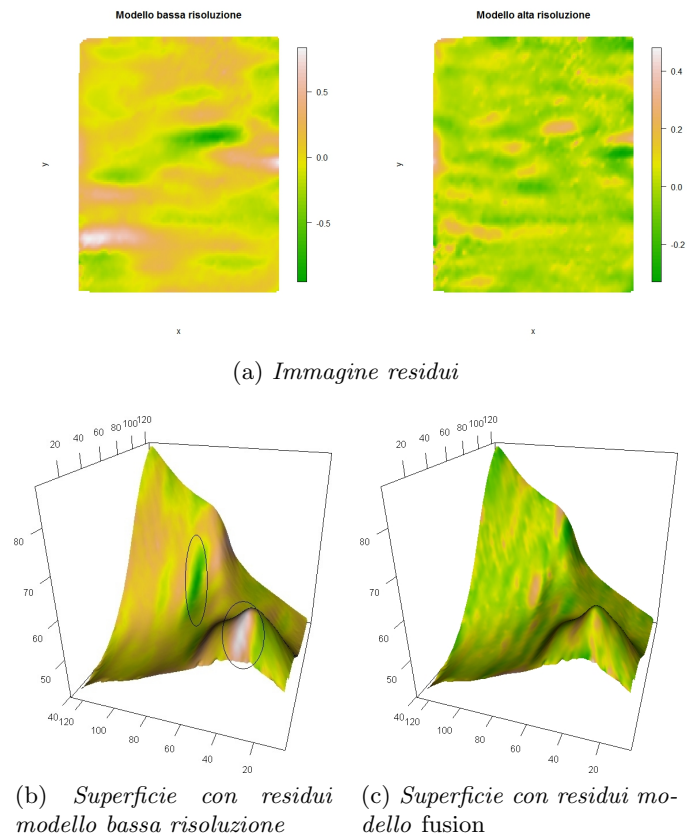


Figura 4.15: Residui modello bassa accuratezza e *fusion*

In figura 4.16 e 4.17 si presentano, rispettivamente, i boxplot dei residui e dei loro quadrati dei vari modelli utilizzati.

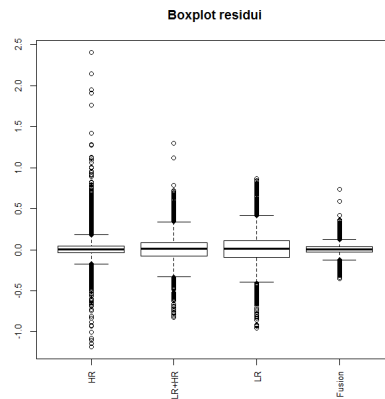


Figura 4.16: Boxplot residui

Con *HR* (*high resolution*) viene indicato il modello ottenuto dai soli punti ad alta risoluzione, mentre con *LR* (*low resolution*) il modello i cui parametri sono stimati dai punti a bassa accuratezza, mentre con il termine *Fusion* si indica il modello che collega i punti a bassa con quelli ad alta risoluzione infine con *LR+HR* si indica il modello *addition*. Dai boxplot dei residui è possibile effettuare le seguenti considerazioni:

- la variabilità dei residui è diminuita passando da un modello a bassa ad uno con fusione;
- il valore dei residui si concentra maggiormente attorno a zero passando da un modello ad uno stadio ad uno con due;
- i residui del modello con i soli punti ad alta risoluzione presentano una grande quantità di outlier con un elevato valore assoluto.

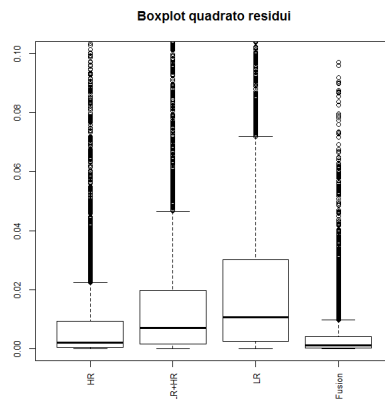


Figura 4.17: Boxplot quadrato residui

Dai boxplot dei quadrati dei residui (figura 4.17) si osserva che la variabilità diminuisce passando da un singolo modello ad un modello gerarchico e si verifica, inoltre, un miglioramento per quel che riguarda le mediane. Si riscontrano valori maggiori anche per la variabilità e la mediana di un modello di soli punti ad alta risoluzione rispetto ad un modello fusion. L'aggiunta di punti ad elevata accuratezza nello stesso dataset senza un modello gerarchico porta buoni miglioramenti, ma non come la fusione mediante modello gerarchico.

Prima di passare ad un'analisi quantitativa si riporta la figura 4.18 in cui vengono raffigurati gli errori del modello a bassa accuratezza (nero) e quelli del modello fusion (rosso) valutati in corrispondenza di tutti i punti del *data set train* (per la stima dei parametri). Si nota come il modello (4.1) commetta degli errori sistematici che vengono successivamente corretti dal modello (4.3) i cui residui sono distribuiti casualmente nello spazio.

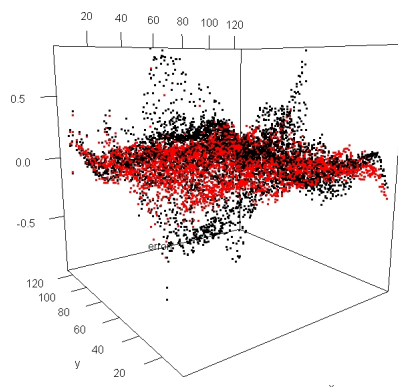


Figura 4.18: Residui modello bassa (nero) ed alta (rosso)

Le statistiche puntuali calcolate sui residui dei modelli sono riassunte nella tabella 4.4.

	<i>HR</i>	<i>LR+HR</i>	<i>LR</i>	<i>Fusion</i>
Media residui	0,0104	0,0077	0,0153	0,0029
Mediana residui	0,0015	0,0087	0,0139	0,0008
Media quadrato residui	0,0283	0,0233	0,0348	0,0051
Mediana quadrato residui	0,0020	0,0071	0,0108	0,0010
Varianza residui	0,0282	0,0232	0,0346	0,0051
Varianza quadrato residui	0,0288	0,0039	0,0068	0,0002

Tabella 4.4: Statistiche puntuali

Le statistiche puntuali confermano le osservazioni fatte sui boxplot dei residui. Passando da un modello ad uno stadio ad uno a due si ha una grande diminuzione nelle varianze dei residui e dei loro quadrati poichè entrambi diminuiscono di un ordine di grandezza. Nel confronto tra un modello di soli punti ad alta risoluzione ed uno *addition* se il parametro di interesse è la media campionaria il modello *addition* ottiene performance migliori, mentre se il metodo di paragone sono le mediane è preferibile utilizzare un modello *HR*.

In figura 4.19 e 4.20 sono rappresentati gli intervalli di confidenza al 99% della mediana, rispettivamente, dei residui e del loro quadrato. Questi intervalli sono stati calcolati attraverso il test non parametrico dei segni poichè, tra le sue ipotesi, non è presente l'assunzione di normalità dei dati, è inoltre robusto rispetto alla presenza di outlier e alla non simmetria dei dati rispetto alla mediana, come nel caso considerato.

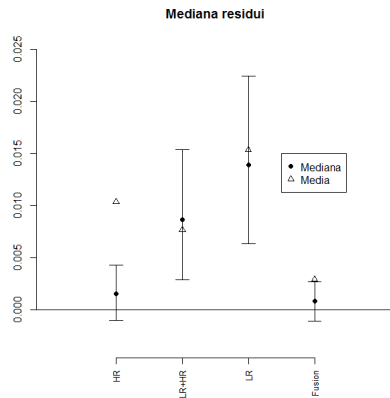


Figura 4.19: Intervalli di confidenza mediana residui

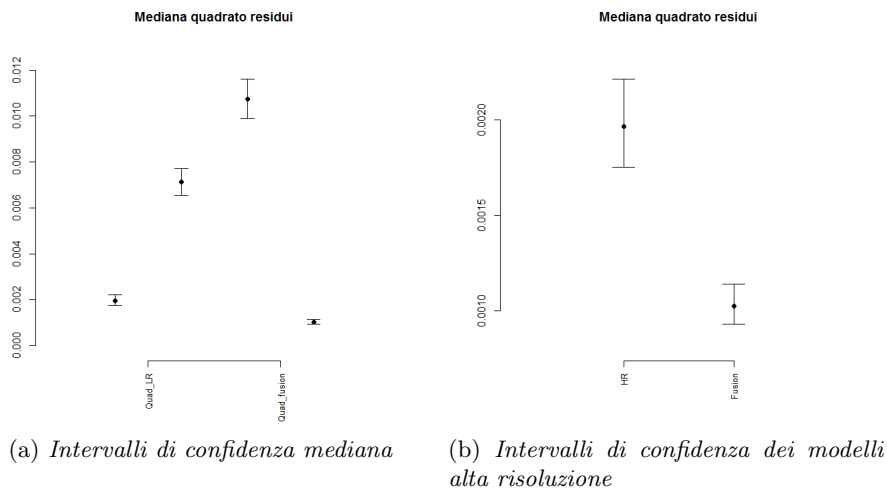


Figura 4.20: Intervalli di confidenza mediana quadrato residui

Gli intervalli di confidenza sulle mediane degli errori indicano che le uniche mediane significativamente diverse da zero sono quelle dei modelli a bassa risoluzione ed *addition* quindi i punti delle superfici predette in questi due casi sono, in media, sottostimati. Quelle degli altri modelli analizzati non sono statisticamente diverse da zero, il che indica una centratura dei residui intorno al valore nullo. Sul grafico è riportato anche il valore puntuale della media, una grande discrepanza dalla mediana indica la presenza di un numero elevato di outlier, la differenza maggiore si ha nel caso *HR*, mentre, nei residui degli altri modelli, le due statistiche non sono molto distanti.



#### 4.5. DECIMAZIONE MEDIANTE METRICA QUADRATICA DEI PUNTI<sup>67</sup>

Vengono riportati due grafici degli intervalli di confidenza per le mediane dei quadrati dei residui (figura 4.20) per poter mettere in evidenza le differenze tra i vari modelli ottenuti con i punti ad alta risoluzione. Per prima cosa si nota che, passando da un modello a bassa risoluzione ad uno a due stadi, si hanno miglioramenti grazie alla compensazione degli errori sistematici presenti sulla superficie misurata attraverso lo scanner a luce strutturata. Invece questa statistica evidenzia che il modello *addition*, pur portando miglioramenti ai punti a bassa accuratezza, ha un valore molto elevato se confrontato con quelli del modello *fusion* e quello di soli punti ad alta accuratezza.

Spostando l'analisi sui soli modelli ad alta risoluzione si conclude che il modello che dà i peggiori risultati è quello di soli punti ad alta risoluzione; rispetto al modello *fusion* questo ha un valore puntuale che corrisponde circa al doppio.

Per il confronto delle varianze dei residui è stato utilizzato il test non parametrico di uguaglianza delle varianze di *Levene*.

I *p-value* dei confronti a coppie delle varianze dei residui sono tutti pari a zero, per cui si può concludere che le varianze dei vari modelli differiscono fra di loro. Dai *p-value* dei confronti a coppie sulle varianze dei quadrati dei residui si può affermare che le uniche varianze che non risultano diverse sono quelle dei modelli *LR* ed *HR*, con un *p-value* pari a 0,198.

### 4.5 Decimazione mediante metrica quadratica dei punti

Solitamente i punti di un artefatto fisico, acquisiti con sistemi a bassa accuratezza, sono in numero molto elevato. Questo porta ad una non applicabilità dei processi gaussiani a fronte del tempo necessario per processare matrici di grandi dimensioni oppure a causa della limitata quantità di memoria RAM che non rende possibile la memorizzazione delle matrici di correlazione dei punti. Per ovviare a questo problema è possibile ricorrere ad una semplificazione della superficie mediante l'utilizzo di un errore quadratico presente in [22].

L'algoritmo è basato sulla contrazione iterativa di coppie di vertici. Ogni vertice è definito come  $\mathbf{v}'_i = (\mathbf{x}'_i \ 1)'$  e la contrazione di due vertici  $\mathbf{v}_1$  e  $\mathbf{v}_2$  in un nuovo vertice  $\bar{\mathbf{v}}$  è indicata con  $(\mathbf{v}_1, \mathbf{v}_2) \rightarrow \bar{\mathbf{v}}$ . Questa operazione muove i due vertici implicati in un nuovo vertice  $\bar{\mathbf{v}}$  e, successivamente, i due vertici e alcune facce vengono eliminate. La procedura termina quando viene raggiunto il numero desiderato di facce o vertici. Per un'implementazione dell'algoritmo è necessario creare la mesh partendo dalla nuvola di punti, questo può essere fatto mediante la triangolazione di Delaunay. Questa tecnica esegue una triangolazione tale che ogni punto appartenente alla nuvola non cada all'interno del cerchio circoscritto ad un qualsiasi triangolo e massimizza il minimo angolo tra tutti gli angoli possibili nella triangolazione.

L'applicazione di questa tecnica al caso in oggetto è presente in figura 4.21.

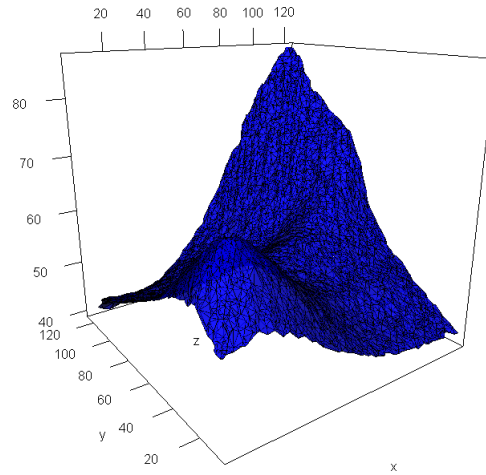


Figura 4.21: Triangolazione di Delaunay

Una volta ottenuti i vertici dei triangoli, è necessario definire i punti che possono essere contratti. Una coppia è considerata valida per l'aggregazione se soddisfa una delle seguenti condizioni:

1.  $(\mathbf{v}_1, \mathbf{v}_2)$  è uno spigolo;
2.  $\|\mathbf{v}_1 - \mathbf{v}_2\|_2 < t$ , con  $t$  definito dall'utente.

La seconda condizione è necessaria nel caso in cui non ci sia congiunzione tra i triangoli, come nel caso in cui siano presenti buchi nella mesh che non sono stati eliminati prima dell'applicazione dell'algoritmo.

Per permettere la contrazione va definita innanzitutto una misura di costo per quantificare l'errore che si commette nella semplificazione della superficie. Gli autori definiscono quindi una matrice  $\mathbf{Q}_i \in (4, 4)$ , mentre l'errore nel generico vertice  $\mathbf{v}_i$  è definito dalla forma quadratica  $\Delta(\mathbf{v}_i) = \mathbf{v}_i' \mathbf{Q}_i \mathbf{v}_i$ . Per ricavare la matrice  $\mathbf{Q}_i$  si parte dalla definizione dell'errore descritto nella (4.4), in cui  $\mathbf{P}_i$  rappresentano i piani, passanti per il vertice  $\mathbf{v}_i$ , ottenuti mediante la triangolazione, mentre  $\mathbf{p}_j = (a_j, b_j, c_j, d_j)$  sono i coefficienti del piano  $j$ -esimo, che soddisfano il vincolo  $a_j^2 + b_j^2 + c_j^2 = 1$ .

$$\Delta(\mathbf{v}_i) = \sum_{\mathbf{p}_j \in \mathbf{P}_i} (\mathbf{p}_j' \mathbf{v}_i)^2 \quad (4.4)$$

Come è possibile osservare dalla definizione, questa forma quadratica è semi-definita positiva poichè è la somma di quantità positive o al più nulle. L'errore definito in precedenza può essere così riscritto mediante una forma quadratica:

$$\begin{aligned}
\Delta(\mathbf{v}_i) &= \sum_{\mathbf{p}_j \in \mathbf{P}_i} (\mathbf{p}'_j \mathbf{v}_i)(\mathbf{p}'_j \mathbf{v}_i) = \\
&= \sum_{\mathbf{p}_j \in \mathbf{P}_i} \mathbf{v}'_i (\mathbf{p}_j \mathbf{p}'_j) \mathbf{v}_i = \\
&= \mathbf{v}'_i \left( \sum_{\mathbf{p}_j \in \mathbf{P}_i} \mathbf{p}_j \mathbf{p}'_j \right) \mathbf{v}_i = \\
&= \mathbf{v}'_i \mathbf{Q}_i \mathbf{v}_i
\end{aligned}$$

Gli autori hanno deciso di scegliere come matrice associata alla forma quadratica dei vertici contratti  $\overline{\mathbf{Q}} = \mathbf{Q}_1 + \mathbf{Q}_2$ , questa matrice, essendo la somma di due matrici semidefinite positive, è anch'essa semidefinita positiva. Il posizionamento del nuovo punto viene scelto attraverso la minimizzazione della funzione  $\Delta(\overline{\mathbf{v}})$ . Per trovare un punto stazionario bisogna trovare le derivate parziali della funzione di costo rispetto ad  $x$ ,  $y$  e  $z$  e porle uguali a zero. Tale operazione è equivalente a risolvere il seguente sistema lineare:

$$\overline{\mathbf{Q}} \overline{\mathbf{v}} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

Risolvibile grazie all'inversione della matrice  $\overline{\mathbf{Q}}$ :

$$\overline{\mathbf{v}} = \overline{\mathbf{Q}}^{-1} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

Nel caso in cui la matrice non sia invertibile un vertice collassa sull'altro, scegliendo come  $\overline{\mathbf{v}}$  quello con funzione di costo minore.

L'algoritmo di decimazione quadratica viene applicato sulla nuvola di punti grazie ad una funzione presente nel software `MeshLab`; il risultato viene riportato nella figura 4.22.

I punti utilizzati sono 1.992 al posto di 4.818 totali, risultando quindi il 40% del totale. La nuvola di punti (figura 4.23a) risulta quindi meno densa rispetto a quella derivata dall'acquisizione dei punti dell'artefatto.

La superficie ottenuta mediante la previsione dei punti attraverso l'utilizzo di un modello di decimazione quadratica (figura 4.23b) ha un andamento simile a quella del modello precedente (figura 4.4) e come questa si discosta da quella ottenuta senza la rimozione dell'errore casuale stimato attraverso il dataset a disposizione.

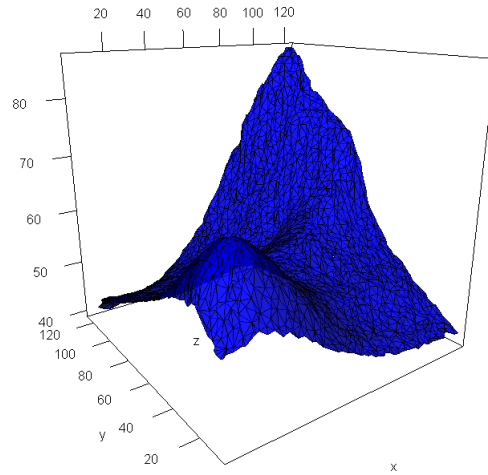


Figura 4.22: Triangolazione di Delaunay

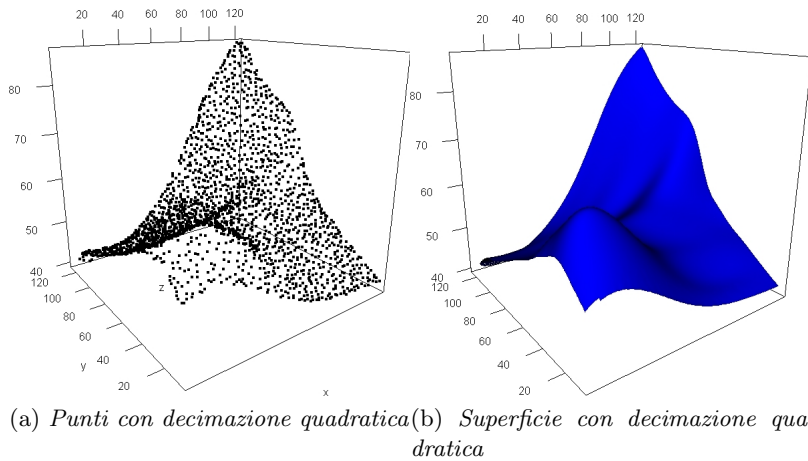


Figura 4.23: Modello con decimazione quadratica

Le stime dei parametri del modello (tabella 4.5) presentano valori simili a quelli stimati utilizzando la totalità dei punti appartenenti al *training set*. Anche in questo scenario c'è un grande effetto del processo gaussiano e la correlazione tra i punti non risulta costante lungo le due direzioni, infatti, i parametri del vettore  $\vartheta$  differiscono di un ordine di grandezza. La stima della varianza dell'errore casuale risulta leggermente superiore rispetto alla situazione precedente poiché, a causa dell'eliminazione di alcuni punti, non è possibile avere le stesse informazioni di partenza; la perdita di informazioni sarà considerata eccessiva se, a valle del modello *fusion*, gli errori di predizione dei due modelli non saranno considerati paragonabili.

#### 4.5. DECIMAZIONE MEDIANTE METRICA QUADRATICA DEI PUNTI 71

Parametro	Valore stima
$\beta_0$	34,8480
$\beta_1$	0,1228
$\beta_2$	0,1451
$\sigma_\eta^2$	49,4922
$\vartheta_1^l$	0,0006
$\vartheta_2^l$	0,0036
$\sigma_{\varepsilon_l}^2$	0,0250

Tabella 4.5: Parametri modello bassa risoluzione

Nel seguito si riporta un'immagine del valore dei residui (4.24b) e del loro posizionamento sulla superficie ricostruita per mezzo della previsione (4.24a) in cui i colori indicano il valore che i residui assumono in un determinato punto nello spazio. Nella ricostruzione non sono presenti errori sistematici, com'era lecito aspettarsi data la somiglianza delle stime dei parametri con quelli del modello precedente. Tuttavia sembra che nella parte più alta dell'artefatto fisico, dove la pendenza è maggiore, il valore assoluto dei residui sia più elevato rispetto alle altre zone della superficie.

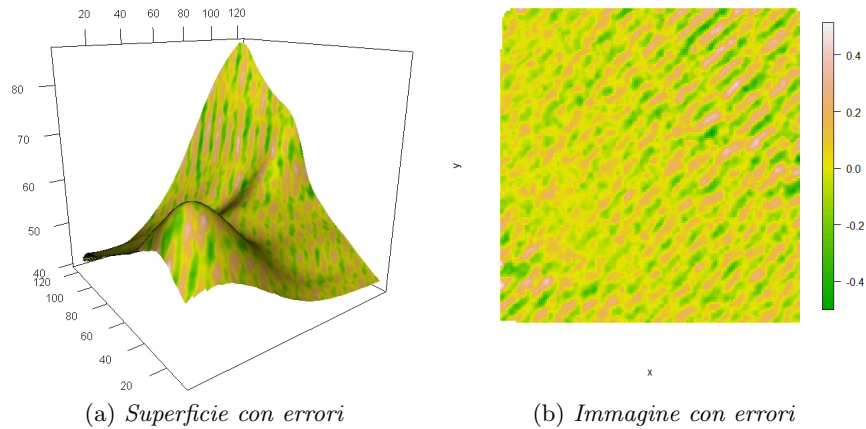


Figura 4.24: Residui modello decimazione quadratica

In figura 4.25 è presente un grafico in cui sono riportati i residui e le ampiezze degli intervalli di confidenza sulla media (blu) e sulla previsione (verde); questi intervalli sono stati calcolati considerando un errore del  $I$  tipo pari al 5%. La percentuale di residui compresa nell'intervallo di previsione è pari al 97,07%, risultato che indica che l'ampiezza dell'intervallo di previsione è leggermente sovrastimata.

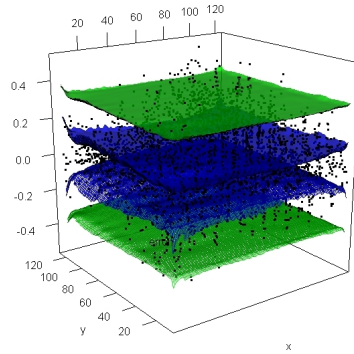
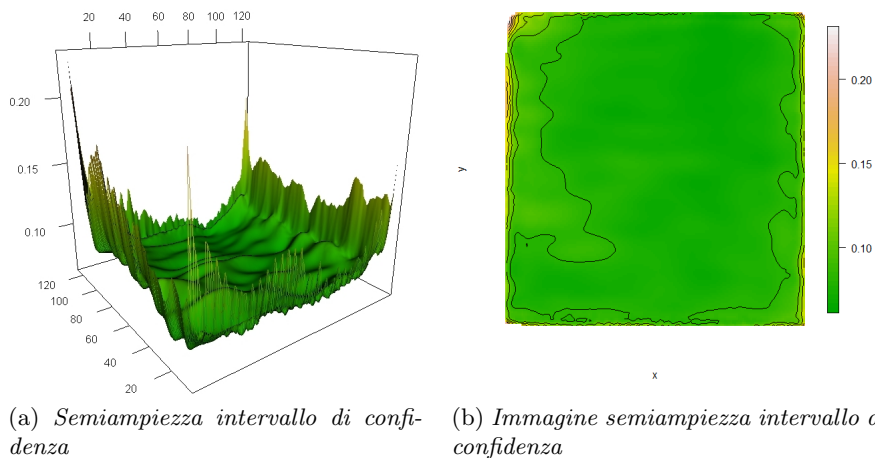


Figura 4.25: Residui e intervalli di confidenza e predizione modello con decimazione quadratica

Da un confronto qualitativo della semiampiezza dell'intervallo di confidenza sulla media (figura 4.26) con quella ottenuta attraverso il modello precedente (figura 4.7) si nota che non sono presenti differenze significative, fatta eccezione degli effetti di bordo che, in questo secondo caso, risultano avere un peso maggiore.



(a) *Semiampiezza intervallo di confidenza*

(b) *Immagine semiampiezza intervallo di confidenza*

Figura 4.26: Semiampiezza intervallo di confidenza sul valore atteso

#### 4.5.1 Modello fusion su modello con decimazione quadratica

A differenza della sezione precedente adesso si utilizza il modello le cui stime sono state eseguite con un numero di punti inferiore, per verificare se la perdita di informazioni avuta nella semplificazione della mesh crea degli errori significativamente diversi. Se così non fosse si potrà affermare che, in questo caso, la perdita di informazioni non ha influenzato il risultato finale della fusione dei

#### 4.5. DECIMAZIONE MEDIANTE METRICA QUADRATICA DEI PUNTI 73

dataset di dati multirisoluzione. Come si nota le stime dei parametri (tabella 4.6) non si discostano molto dal modello di riferimento.

Parametro	Valore stima
$\rho_0$	0,9999
$\delta_0$	0,0226
$\sigma_\delta^2$	0,0396
$\vartheta_1^h$	0,0020
$\vartheta_2^h$	0,0266
$\sigma_{\varepsilon_h}^2$	$5,8425 \cdot 10^{-4}$

Tabella 4.6: Stime modello alta risoluzione decimazione quadratica

I coefficienti del kernel gaussiano  $\vartheta_1^h$  e  $\vartheta_2^h$  hanno un valore molto vicino a quelli del modello precedente, così come le stime delle varianze, mentre si discostano leggermente le stime dei parametri di scala e di traslazione. L'effetto totale della variazione di questi parametri causa delle differenze minime sulla previsione dei valori. I diversi valori assunti dagli stimatori possono essere causati dalle coordinate predette dai due modelli a bassa risoluzione nei punti ad alta accuratezza utilizzati per la stima.

L'effetto del processo gaussiano nel modello fusion con decimazione quadratica, così come in quello fusion precedente, ha un basso effetto nella previsione dei punti test, ed ha degli effetti localizzati in alcune aree per permettere la correzione di errori sistematici presenti nella nuvola di punti con una bassa accuratezza.

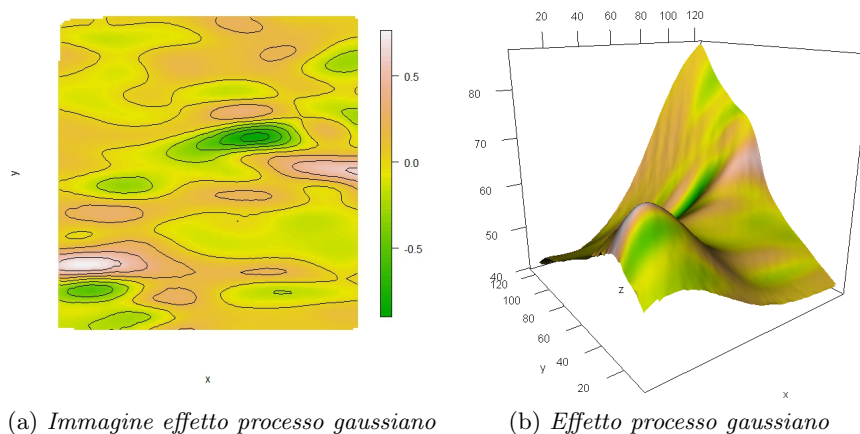


Figura 4.27: Effetto processo gaussiano con decimazione quadratica

I residui dei modelli in cui si sono utilizzati i punti derivanti dalla decimazione quadratica (figura 4.28), hanno lo stesso andamento dei modelli precedenti di

riferimento. Infatti il modello a bassa risoluzione commette errori sistematici nelle zone segnate, che vengono poi corretti dal modello *fusion*.

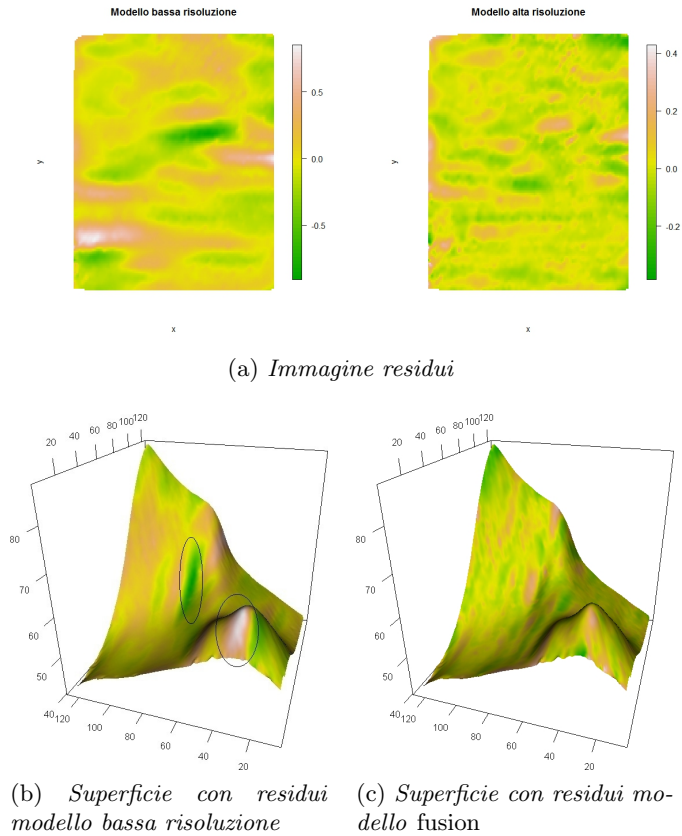


Figura 4.28: Residui modello bassa accuratezza e *fusion* decimazione quadratica

Le statistiche puntuali calcolata sui residui del modello con decimazione quadratica sono riportate in tabella 4.7 in cui *Quad\_LR* si riferisce al modello a bassa risoluzione, mentre *Quad\_fusion* al modello *fusion*.

	<i>Quad_LR</i>	<i>Quad_fusion</i>
Media residui	0,0151	0,0027
Mediana residui	0,0139	-0,0009
Media quadrato residui	0,0360	0,0049
Mediana quadrato residui	0,0111	0,0010
Varianza residui	0,0358	0,0049
Varianza quadrato residui	0,0068	0,0002

Tabella 4.7: Statistiche puntuali modello con decimazione quadratica



#### 4.5. DECIMAZIONE MEDIANTE METRICA QUADRATICA DEI PUNTI75

I valori puntuali dei modelli a bassa risoluzione si discostano poco rispetto a quelli di riferimento 4.4.

Nel seguito si riportano delle stime della funzione di densità dei residui dei modelli oggetto di questa analisi. In figura 4.29 sono riportate le stime sulla distribuzione dei residui, mentre in figura 4.30 sono raffigurate le stime sulla distribuzione del quadrato dei residui.

Le stime delle funzioni di densità dei residui dei modelli a bassa risoluzione si discostano leggermente, infatti la distribuzione del modello a bassa risoluzione di riferimento sembra essere bimodale, mentre quella del modello con decimazione quadratica presenta un andamento unimodale e più simmetrico.

Le stime sulle distribuzioni dei modelli *fusion* sono quasi sovrapponibili, il che indica che hanno praticamente la stessa distribuzione.

Le stime sulle distribuzioni del quadrato dei residui non presentano differenze rilevanti tra i modelli analizzati.

Questa analisi qualitativa suggerisce che non siano presenti differenze tra il modello con decimazione quadratica e quello di riferimento.

Gli intervalli di confidenza sulle mediane (figura 4.31) mostrano che le mediane dei residui dei modelli a bassa risoluzione sono diverse da zero, con valori simili, mentre entrambe le mediane dei modelli *fusion* non si può affermare siano diverse da zero.

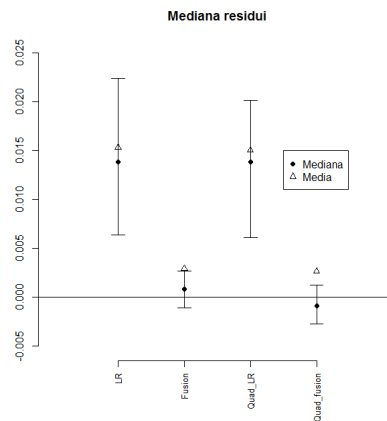


Figura 4.31: Intervalli di confidenza 99% mediana residui

In figura 4.32a sono rappresentati gli intervalli di confidenza sulle mediane dei quadrati dei residui dei modelli a bassa risoluzione. Seppur gli intervalli di confidenza siano sovrapposti con un *p-value* pari a  $2,572e-05$  si può affermare che le due mediane sono statisticamente diverse. Anche se le due statistiche sono risultate diverse la stima della mediana della differenza è pari a  $-1,4223 \cdot 10^4$ ,

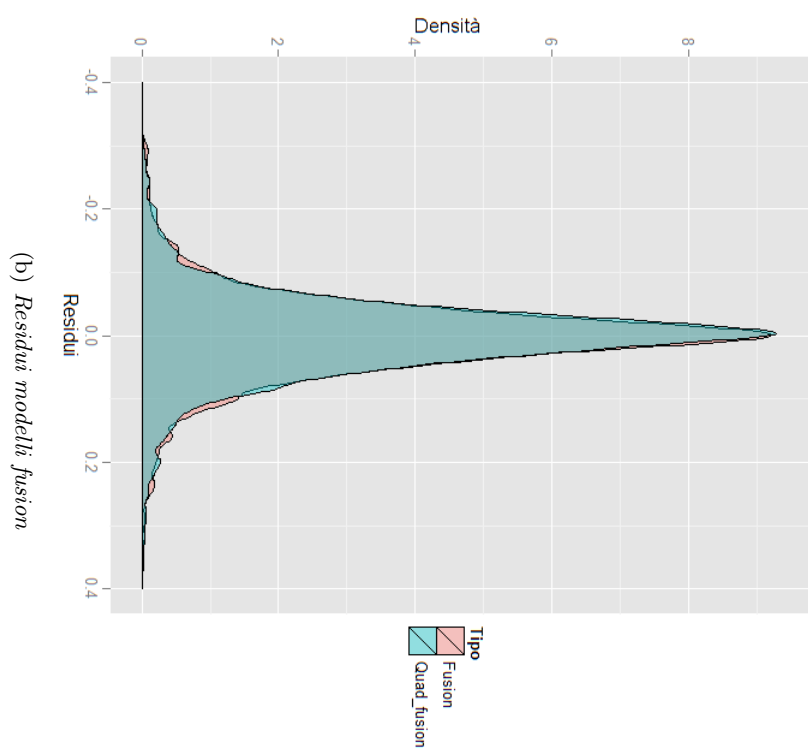
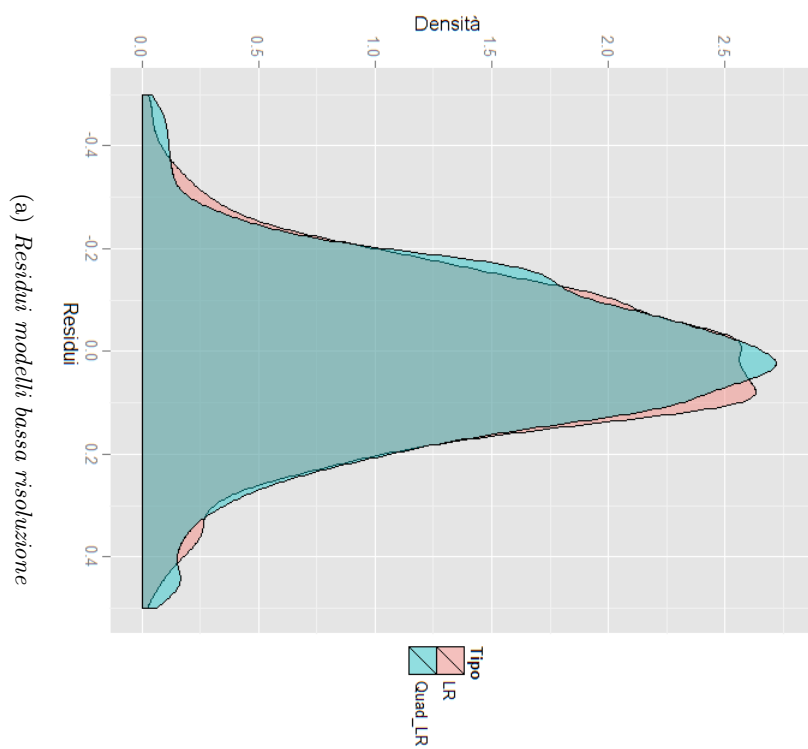
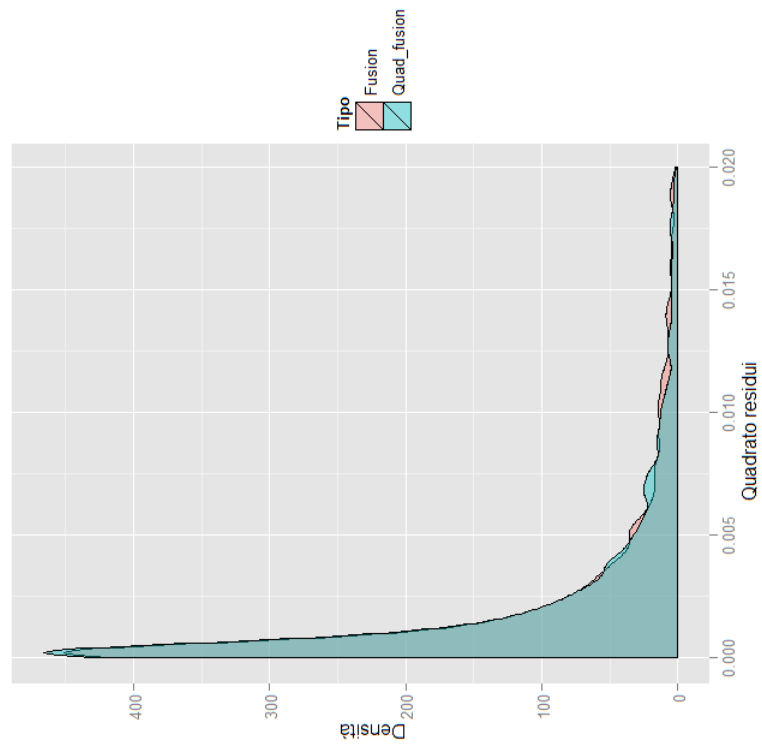
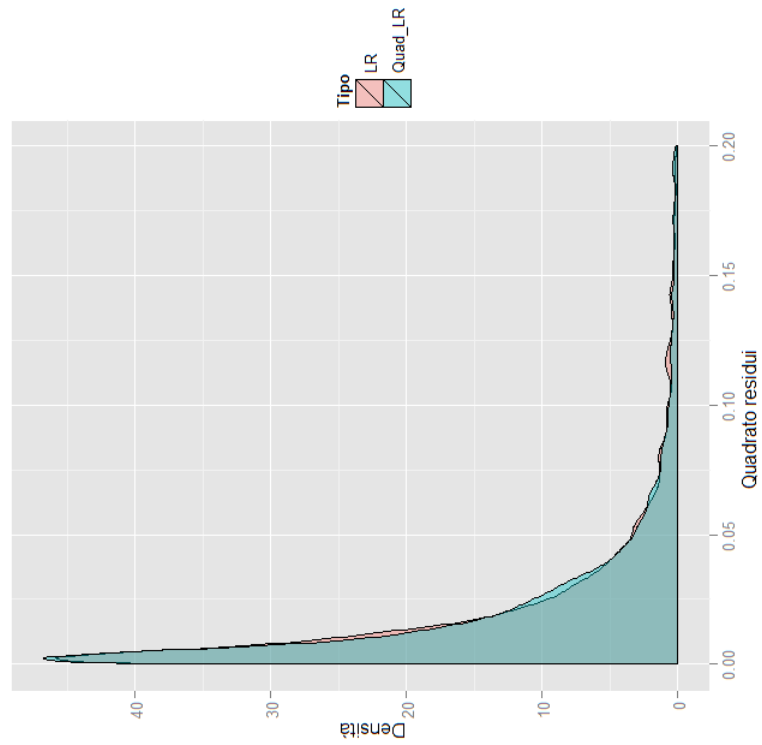


Figura 4.29: Stima densità residui

4.5. DECIMAZIONE MEDIANTE METRICA QUADRATICA DEI PUNTI77



(b) *Quadrati residui modelli fusion*



(a) *Quadrati residui modelli bassa risoluzione*

Figura 4.30: Stima densità quadrati residui

un valore molto basso se paragonato ai singoli valori puntuali.

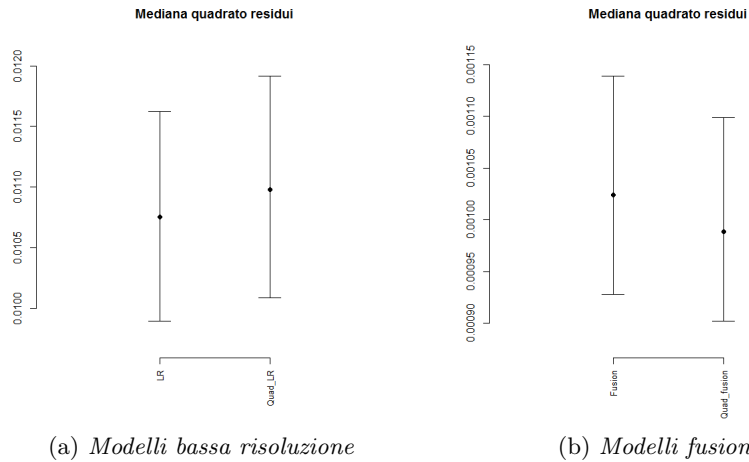


Figura 4.32: Intervalli di confidenza 99% mediana quadrato residui

Le mediane dei quadrati dei residui dei modelli *fusion* (figura 4.32b) non sono statisticamente diverse tra, infatti il  $p$ -value del confronto a coppie è 0,8627. In questo caso il modello gerarchico con decimazione quadratica ottiene risultati simili a quello di riferimento permettendo una riduzione del tempo necessario per la stima dei parametri del modello a bassa risoluzione, senza causare una perdita di prestazioni nel modello *fusion*.

## 4.6 Conclusioni

A valle di questa analisi si può concludere che tramite un modello a due stadi si possono eliminare degli errori sistematici dovuti ad una misura non accurata dei punti di una superficie dovuta ad uno strumento con una precisione non elevata.

Non sono stati ottenuti gli stessi risultati di un modello *fusion* facendo un modello con processo gaussiano sui soli punti ad elevata accuratezza. Questo indica che, anche se i punti del modello a bassa accuratezza presentano un errore sistematico in alcune zone, forniscono un buon punto di partenza per la stima dei parametri di un modello *fusion*.

Un modello *addition* non è adatto nel caso in cui siano presenti errori sistematici nella rilevazione della superficie mediante uno strumento a bassa accuratezza poichè il rapporto tra il numero di punti a bassa accuratezza e quelli ad alta è

molto alto ed i punti più precisi non riescono ad effettuare grandi correzioni perché hanno un peso pari a quello dei punti meno precisi.

L'ultima conclusione che si può trarre è che la decimazione quadratica dei punti del *training set* mediante una metrica quadratica non ha provocato la perdita di informazioni permettendo di eseguire delle buone stime dei parametri del modello a bassa risoluzione.



## Capitolo 5

# Campionamento punti ad alta risoluzione

Dopo aver verificato i vantaggi derivanti dall'utilizzo di un modello gerarchico nel campo della metrologia, si presenta in questo capitolo una proposta di piano di campionamento dei punti ad elevata accuratezza. I fattori che vengono presi in considerazione per un buon posizionamento dei punti affetti da un errore minore sono l'incertezza della predizione derivante da un modello a bassa risoluzione e la curvatura gaussiana della superficie [43].

Nella prima parte del capitolo è presente una breve descrizione della curvatura gaussiana e del piano di campionamento proposto nel presente lavoro di tesi, mentre nella seconda parte è presente un'analisi delle prestazioni di un modello gerarchico basato sui punti campionati secondo l'algoritmo proposto.

### 5.1 Curvatura gaussiana

Sia  $\Phi$  una superficie regolare di classe  $C^k$  con  $k \geq 2$  e sia  $P$  un punto appartenente a tale superficie. Si consideri il piano  $\Pi(P, \boldsymbol{\lambda})$  definito dalla normale a  $\Phi$  nel punto  $P$  ( $\mathbf{n}(P)$ ) e da un vettore  $\boldsymbol{\lambda} \in T\Phi_P$ , dove  $T\Phi_P$  rappresenta il piano tangente alla superficie nel punto  $P$ . L'intersezione di questo piano con la superficie in qualche intorno di  $P$  è una curva regolare  $\gamma$  di classe  $C^k$ . Indicando con  $\tilde{k}(P, \boldsymbol{\lambda})$  la sua curvatura nel punto  $P$ , e se  $\tilde{k}(P, \boldsymbol{\lambda}) \neq 0$ , si indica con  $\mathbf{v}(P, \boldsymbol{\lambda})$  il vettore normale principale di  $\gamma$  nel punto  $P$ . Siccome  $\gamma$  giace nel piano  $\Pi(P, \boldsymbol{\lambda})$ ,  $\mathbf{v}(P, \boldsymbol{\lambda}) = \pm \mathbf{n}(P)$ . Si definisce il numero reale  $k(P, \boldsymbol{\lambda}) = \langle \mathbf{v}(P, \boldsymbol{\lambda}), \mathbf{n}(P) \rangle \tilde{k}(P, \boldsymbol{\lambda})$  assumendo che

$$k(P, \boldsymbol{\lambda}) = \begin{cases} \tilde{k}(P, \boldsymbol{\lambda}) & \text{se } \mathbf{v}(P, \boldsymbol{\lambda}) = \mathbf{n}(P, \boldsymbol{\lambda}), \\ 0 & \text{se } \mathbf{v}(P, \boldsymbol{\lambda}) = \mathbf{0}_2, \\ -\tilde{k}(P, \boldsymbol{\lambda}) & \text{se } \mathbf{v}(P, \boldsymbol{\lambda}) = -\mathbf{n}(P, \boldsymbol{\lambda}), \end{cases}$$

in cui  $\langle \cdot, \cdot \rangle$  rappresenta il prodotto scalare.

**Definizione 5.1.1.** *La curvatura normale della superficie  $\Phi$  nel punto  $P$  nella direzione  $\lambda$  è il numero reale  $k(P, \lambda)$ .*

Il segno della curvatura normale dipende dalla scelta della direzione della normale  $\mathbf{n}(P)$ . Sia  $\mathbf{r} = \mathbf{r}(u, v)$  il raggio vettore che rappresenta la forma parametrica della superficie  $\Phi$ ,  $u = u(t)$  e  $v = v(t)$  le equazioni della curva  $\gamma$ , che dipendono da un parametro  $t$ ,  $\gamma = \mathbf{r}(t) = \mathbf{r}(u(t), v(t))$ , con le proprietà:

$$\begin{aligned}\gamma(0) &= P \\ \boldsymbol{\tau}(0) &= \boldsymbol{\lambda} = \lambda_1 \mathbf{r}_u + \lambda_2 \mathbf{r}_v \\ \lambda_1 &= \frac{du}{dt}(0) \\ \lambda_2 &= \frac{dv}{dt}(0).\end{aligned}$$

Si deriva la curvatura normale  $k(P, \lambda)$ :

$$\begin{aligned}k(P, \lambda) &= \langle \mathbf{v}(P, \lambda), \mathbf{n}(P) \rangle \tilde{k}(P, \lambda) = \left\| \frac{d^2 \mathbf{r}}{dt^2} \right\| \cdot \frac{1}{\|\boldsymbol{\lambda}\|^2} \langle \mathbf{v}(P, \lambda), \mathbf{n}(P) \rangle, \\ \left\| \frac{d^2 \mathbf{r}}{dt^2} \right\| &= \mathbf{r}_{uu} \lambda_1^2 + 2 \mathbf{r}_{uv} \lambda_1 \lambda_2 + \mathbf{r}_{vv} \lambda_2^2 + \mathbf{r}_u \frac{d^2 u}{dt^2} + \mathbf{r}_v \frac{d^2 v}{dt^2}\end{aligned}$$

dove  $\|\cdot\|$  rappresenta la norma euclidea.  
Sostituendo si ottiene

$$k(P, \lambda) = \frac{1}{\|\boldsymbol{\lambda}\|^2} [\langle \mathbf{r}_{uu}, \mathbf{n} \rangle \lambda_1^2 + 2 \langle \mathbf{r}_{uv}, \mathbf{n} \rangle \lambda_1 \lambda_2 + \langle \mathbf{r}_{vv}, \mathbf{n} \rangle \lambda_2^2]. \quad (5.1)$$

Introducendo la notazione

$$\begin{aligned}E &= \left( \frac{dr_1}{du} \right)^2 + \left( \frac{dr_2}{du} \right)^2 + \left( \frac{dr_3}{du} \right)^2, & F &= \frac{dr_1}{du} \frac{dr_1}{dv} + \frac{dr_2}{du} \frac{dr_2}{dv} + \frac{dr_3}{du} \frac{dr_3}{dv}, \\ G &= \left( \frac{dr_1}{dv} \right)^2 + \left( \frac{dr_2}{dv} \right)^2 + \left( \frac{dr_3}{dv} \right)^2, & L &= \langle \mathbf{r}_{uu}, \mathbf{n} \rangle, & M &= \langle \mathbf{r}_{uv}, \mathbf{n} \rangle, & N &= \langle \mathbf{r}_{vv}, \mathbf{n} \rangle\end{aligned}$$

si ha

$$k(P, \lambda) = \frac{L\lambda_1^2 + 2M\lambda_1\lambda_2 + N\lambda_2^2}{E\lambda_1^2 + 2F\lambda_1\lambda_2 + G\lambda_2^2}, \quad (5.2)$$

dove  $E, F, G, L, M, N$  dipendono dallo specifico punto  $P$ .

Siccome tutte le direzioni sul piano tangente  $T\Phi_P$  formano un insieme compatto di trasformazioni omeomorfe alla circonferenza,  $k(P, \lambda)$  ha almeno un minimo ed un massimo, cioè due estremanti.



**Definizione 5.1.2.** *La curvatura principale di una superficie  $\Phi$  nel punto  $P$  è un estremante della curvatura normale  $k(P, \lambda)$  in un dato punto  $P$  appartenente a  $\Phi$ . Le direzioni sul piano  $T\Phi_P$  lungo cui  $k(P, \lambda)$  assume i suoi valori estremi sono i vettori principali della superficie  $\Phi$  nel punto  $P$ .*

Si dimostra che esistono due estremanti di questa funzione, e sono gli zeri dell'equazione di secondo grado (5.3), si indicano con  $k_1$  e  $k_2$  e sono le due curvatures principali.

$$(EG - F^2)k^2 - (EN + GL - 2MF)k + LN - M^2 = 0 \quad (5.3)$$

Si dà ora la definizione della curvatura gaussiana.

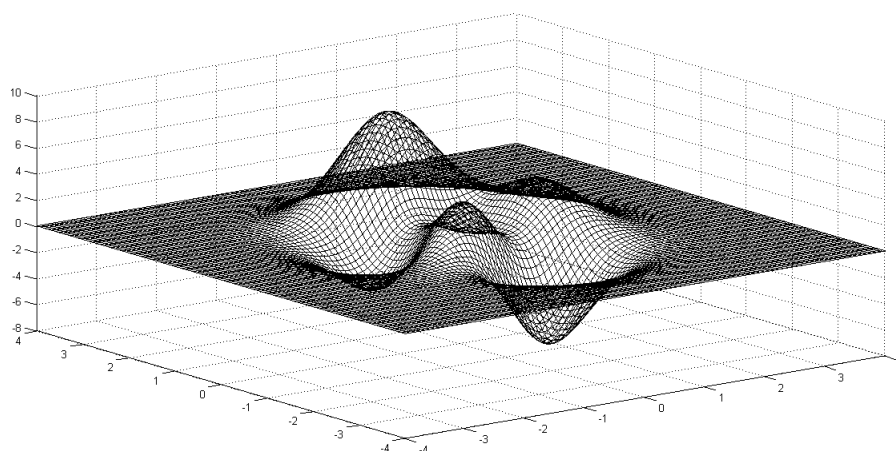
**Definizione 5.1.3.** *La curvatura gaussiana di una superficie,  $K(P)$ , è definita come*

$$K(P) = k_1(P) \cdot k_2(P).$$

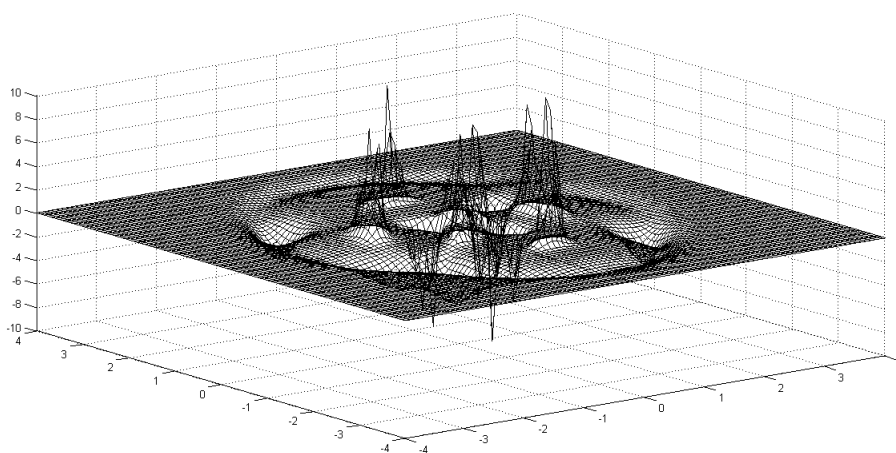
Una proprietà di  $K(P)$  è che dipende solo dal punto della superficie  $P$  ed è una quantità invariante rispetto a traslazioni e rotazioni della superficie  $\Phi$ .

In un caso reale non si ha a disposizione l'equazione parametrica della superficie a causa degli errori dovuti al processo di lavorazione, per cui le derivate sono state calcolate numericamente.

In figura 5.1a è rappresentata la superficie *peaks* (mistura di funzioni gaussiane) con dominio  $[-4, 4]^2$  (figura 5.1a). La curvatura gaussiana, stimata con metodi numerici, è rappresentata in figura 5.1b. Si noti che sono stati eliminati punti con valore in modulo maggiore di dieci per mettere in evidenza la differenza tra le zone piane, in cui il valore è pressoché nullo, e quelle dove la curvatura gaussiana ha un valore diverso da zero.



(a) *Superficie peaks*



(b) *Curvatura gaussiana*

Figura 5.1: Superficie *peaks* con curvatura gaussiana

## 5.2 Piano di campionamento proposto

Dopo aver verificato nel capitolo precedente i vantaggi di un modello gerarchico, ci si pone ora il problema della selezione dei punti ad elevata accuratezza poichè sono solitamente disponibili in quantità limitata. Nel piano di campionamento proposto nel presente lavoro di tesi si considerano come fattori che determinano il buon posizionamento di un punto ad alta risoluzione l'incertezza sulla predizione che si ottiene dal modello a bassa accuratezza e la curvatura gaussiana della superficie scansionata.

Il valore dell'incertezza è un buon indicatore poichè è alto dove il modello a bassa risoluzione non è in grado di predire bene, cioè dove i punti sono stati acquisiti con un errore maggiore per esempio a causa di una disomogeneità in varianza.

Si considera la curvatura gaussiana poichè è elevata nelle zone in cui sono presenti bruschi cambiamenti della superficie, mentre è caratterizzata da valori più bassi più la superficie acquisita si avvicina ad un piano.

La scelta di inserire nel campionamento la curvatura gaussiana si basa sull'assunzione che per correggere la superficie nelle zone che si avvicinano ad un piano siano necessari pochi punti, mentre ne serviranno un numero maggiore dove la variazione della superficie è più brusca.

Il campionamento proposto si basa sui seguenti passi:

1. generazione di punti candidati secondo una griglia regolare;
2. suddivisione del piano in  $n$  classi;
3. calcolo della curvatura gaussiana sui punti candidati mediante differenze finite;
4. calcolo della semiampiezza degli intervalli di confidenza sulla media o sul valore predetto;
5. codifica del valore assoluto della curvatura gaussiana e della semiampiezza dell'intervallo di confidenza in modo tale che il massimo sia uno in entrambi i casi;
6. calcolo dei punti da campionare in ogni classe in modo proporzionale alla combinazione della media del valore assoluto della curvatura gaussiana e della media dell'incertezza;
7. calcolo della densità di probabilità mediante una combinazione lineare della curvatura gaussiana e dell'incertezza in ogni punto;
8. campionamento dei punti da questa distribuzione, mediante il metodo *rejection sampling* [12];

9. aggiunta di un errore casuale, con media zero e varianza piccola, per non avere punti allineati in una direzione. La deviazione standard è fissata ad una quantità pari ad un centesimo del range del piano sperimentale (asse  $x$  ed  $y$ );
10. verifica di non superamento del confine del piano sperimentale, in caso contrario forzatura della coordinata al valore estremo.

Si presentano ora dei modi diversi di calcolare l'incertezza sul valore predetto ed in seguito si illustrano i passi dell'algoritmo per mezzo di un esempio ponendo una particolare attenzione al punto 2. dell'algoritmo.

### Incetezza

Si presentano due metodi di calcolo dell'incertezza e un loro confronto. L'incertezza può essere definita come la semiampiezza dell'intervallo di confidenza oppure dell'intervallo di predizione. Si tratta solo il caso delle varianze asintotiche, ma nel caso si voglia considerare il MSPE i passaggi sono simili a quelli riportati di seguito. Si mostra ora che, considerando la semiampiezza dell'intervallo di confidenza, l'incertezza non dipende dallo specifico livello di confidenza,  $\alpha$ . Infatti considerando  $\alpha_0 \in [0, 1]$ , un generico errore di I tipo, e

$$\sigma_{\hat{\mu}_i} = \sqrt{\sigma_{\eta}^2 - \sigma_{\eta}^4 \mathbf{r}'_{\eta_{12}} \boldsymbol{\Sigma}_l^{-1} \mathbf{r}_{\eta_{12}}} \quad (5.4)$$

si ha che l'incertezza standardizzata sul valor medio (il massimo valore che può assumere corrisponde ad 1),  $c_i^1$ , è calcolata come:

$$c_i^1 = \frac{z_{1-\frac{\alpha_0}{2}} \sigma_{\hat{\mu}_i}}{\max_i \left\{ z_{1-\frac{\alpha_0}{2}} \sigma_{\hat{\mu}_i} \right\}} = \frac{\sigma_{\hat{\mu}_i}}{\max_i \left\{ \sigma_{\hat{\mu}_i} \right\}} \quad (5.5)$$

con  $z_{1-\frac{\alpha_0}{2}}$  il quantile di ordine  $1 - \frac{\alpha_0}{2}$  della variabile aleatoria  $Z \sim \mathcal{N}(0, 1)$ , e  $\mathbf{r}_{\eta_{12}} = \text{Corr}(\mathbf{y}_l, y_l(\mathbf{x}_i))$ .

Si osserva che  $c_i^1$  non dipende dallo specifico  $\alpha_0$  scelto. Se si considera l'incertezza sulla previsione di un nuovo punto si giunge alla stessa conclusione. Nel caso in cui si preferisca utilizzare l'incertezza di previsione al posto di quella sul valore atteso, l'incertezza standardizzata,  $c_i^2$  è data da

$$c_i^2 = \frac{\sigma_{y_{\hat{i}}}}{\max_i \left\{ \sigma_{y_{\hat{i}}} \right\}} = \frac{\sqrt{\sigma_{\varepsilon_l}^2 + \sigma_{\hat{\mu}_i}^2}}{\sqrt{\sigma_{\varepsilon_l}^2 + \max_i \left\{ \sigma_{\hat{\mu}_i}^2 \right\}}} \quad (5.6)$$

$c_i^1 \simeq c_i^2$  quando:

1.  $\sigma_{\varepsilon_l} \ll \sigma_{\hat{\mu}_i} \quad \forall i = 1, \dots, n_1$ ;

2.  $\sigma_{\varepsilon_l} \gg \sigma_{\hat{\mu}_i} \quad \forall i = 1, \dots, n_1;$
3.  $\sigma_{\varepsilon_l} \simeq m\sigma_{\hat{\mu}_i} \quad \forall i = 1, \dots, n_1, \quad m \in \mathbb{R}.$

Anche se i valori di  $c_i^1$  e  $c_i^2$  differiscono, si preserva l'ordinamento, infatti con  $\sigma_{\hat{\mu}_{l_1}} > \sigma_{\hat{\mu}_{l_2}}$  si ha che:

$$c_1^1 = \frac{\sigma_{\hat{\mu}_{l_1}}}{\max_i \{\sigma_{\hat{\mu}_i}\}} > \frac{\sigma_{\hat{\mu}_{l_2}}}{\max_i \{\sigma_{\hat{\mu}_i}\}} = c_2^1$$

e

$$c_1^2 = \frac{\sqrt{\sigma_{\varepsilon_l}^2 + \sigma_{\hat{\mu}_{l_1}}^2}}{\sqrt{\sigma_{\varepsilon_l}^2 + \max_i \{\sigma_{\hat{\mu}_i}^2\}}} > \frac{\sqrt{\sigma_{\varepsilon_l}^2 + \sigma_{\hat{\mu}_{l_2}}^2}}{\sqrt{\sigma_{\varepsilon_l}^2 + \max_i \{\sigma_{\hat{\mu}_i}^2\}}} = c_2^2.$$

Per cui scegliendo  $c^1$  o  $c^2$  i risultati non saranno identici, ma simili.

### Campionamento

La suddivisione del piano in  $n$  classi permette di effettuare un campionamento stratificato, cioè di guidare la scelta dei punti sperimentali nel caso in cui siano pochi, per evitare che vengano estratti quelli con una bassa densità di probabilità.

Si presenta ora l'esempio di tre campionamenti di dieci punti dalla superficie rappresentata in figura 5.2. L'errore è stato simulato attraverso le realizzazioni di un funzione di probabilità normale con media nulla e varianza pari a 0,0012. La funzione a cui è stato aggiunto il rumore è descritta dalla (5.7), con il dominio che corrisponde a  $[100, 200]^2$ .

$$f(x, y) = c_i \sum_{i=1}^3 \left( 2\pi \frac{k_i}{b_i} + a_i \right) - \frac{1}{\sqrt{2\pi} e} \exp \left\{ -\frac{\|\mathbf{d}\|^2}{2e^2} \right\} \quad (5.7)$$

con

$$\begin{aligned} k_i &= y + x \sin \vartheta_i, \quad \forall i = 1, 2, 3 \\ \mathbf{d} &\in \mathbb{R}^2, \\ d_1 &= x - x_0 - \left\langle \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix}, \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix} \right\rangle \cos \varphi, \\ d_2 &= y - y_0 - \left\langle \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix}, \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix} \right\rangle \sin \varphi, \\ \boldsymbol{\vartheta}' &= (0, 5\pi; -0, 2\pi; \pi)', \quad \mathbf{a}' = (0; 20; 0)' \\ \mathbf{b}' &= (300; 200; 600)', \quad \mathbf{c}' = (1; 1; 0, 8)' \\ x_0 &= 128, \quad y_0 = 64, \quad \varphi = \frac{\pi}{2} - \frac{\pi}{8}, \quad e = 4 \end{aligned} \quad (5.8)$$

I punti all'interno del solco sono presenti con minore densità poiché si ipotizza provengano da uno scanner con sensore ottico che, come descritto in [7], può presentare zone in cui sono presenti punti mancanti. La mancanza di punti può essere causata dalla presenza di sottosquadri o da un'imprecisa identificazione delle frange della luce strutturata da parte dell'algoritmo di acquisizione che provoca la generazione di punti fantasma che devono essere eliminati.

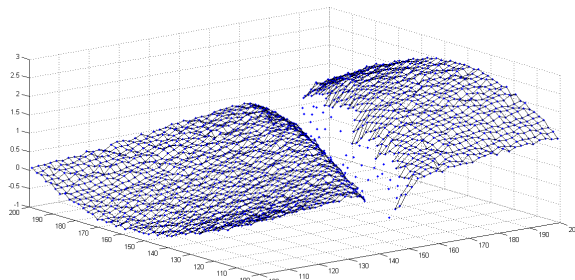


Figura 5.2: Superficie del modello a bassa risoluzione

Si utilizza un modello con processo gaussiano come descritto nella (5.9) per descrivere il comportamento della nuvola di punti.

$$y_l(\mathbf{x}_i) = \beta_0 + \eta(\mathbf{x}_i) + \varepsilon_l(\mathbf{x}_i) \quad (5.9)$$

Successivamente si stima la curvatura gaussiana (figura 5.3a) e la semiampiezza dell'intervallo di confidenza sulla media (figura 5.3b) calcolate nei punti di una griglia regolare di 900 punti. La griglia è stata scelta molto densa per permettere una corretta stima della curvatura gaussiana, che si basa su differenze finite, e per avere una grande quantità di punti da cui poter campionare. Si noti che le funzioni disegnate assumono tutti valori discreti, in figura sono state raffigurate superfici solo per una più semplice comprensione visiva.

La densità di probabilità (figura 5.3c) da cui verranno estratti i punti è proporzionale alla funzione obiettivo calcolata come combinazione convessa del valore assoluto della curvatura gaussiana e dell'incertezza con pesi che equivalgono in entrambi i casi a 0,5:

$$g(\mathbf{x}_i) = \omega_1 k_i + \omega_2 c_i^1 \quad \forall i = 1, \dots, n \quad (5.10)$$

in cui  $g(\mathbf{x}_i)$  rappresenta il valore della funzione obiettivo nel punto  $i$ -esimo,  $c_i^1$  il valore dell'incertezza nel punto  $i$ -esimo (definita nella (5.5))  $k_i = \frac{|K(\mathbf{x}_i)|}{\max_i\{|K(\mathbf{x}_i)|\}}$  nel punto  $i$ -esimo ( $K(\mathbf{x}_i)$ , come introdotto in precedenza, rappresenta la curvatura gaussiana) e  $\omega_1, \omega_2$  i coefficienti che pesano i due valori, in questo caso  $\omega_1 = \omega_2 = 0,5$ .

La densità di probabilità discreta da cui campionare i punti è data da:

$$p(\mathbf{x}_i) = \frac{g(\mathbf{x}_i)}{\sum_{i=1}^n g(\mathbf{x}_i)} \quad \forall i = 1, \dots, n. \quad (5.11)$$

Il campionamento viene eseguito mediante il metodo *rejection sampling*. Considerando che  $p(\mathbf{x}_i) \leq \max_i \{p(\mathbf{x}_i)\} = M$ , con  $M \propto \frac{1}{n}$  e  $j_{max}$  che corrisponde al numero di campioni desiderato, il metodo utilizzato è:

**Passo 1:** definire un contatore  $j = 1$ ;

**Passo 2:** definire un campione candidato generando casualmente un numero compreso tra 1 e  $n$  a cui è associato un punto  $\mathbf{x}_i$ . Generazione di  $u$  che corrisponde ad una realizzazione di  $U \sim \mathcal{U}(0, 1)$ ;

**Passo 3:** accettare il campione candidato se  $u \leq \frac{p(\mathbf{x}_i)}{M}$ ;

**Passo 4:** se  $j < j_{max}$ ,  $j = j + 1$  e tornare al **Passo 2**.

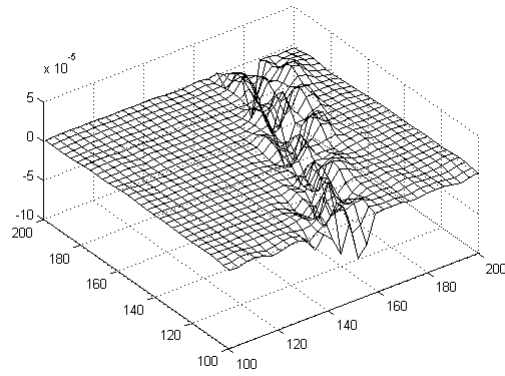
Un vantaggio di questo tipo di campionamento è che, anche nel solco dove si hanno pochi punti a bassa risoluzione, grazie al fatto che l'incertezza è elevata vengono posizionati un gran numero di punti ad alta risoluzione.

Si decide di effettuare un campionamento di dieci punti dalla funzione di probabilità per evidenziare le differenze che si hanno considerando una oppure tante classi, nello specifico il confronto è fatto tra 1 e 25 classi.

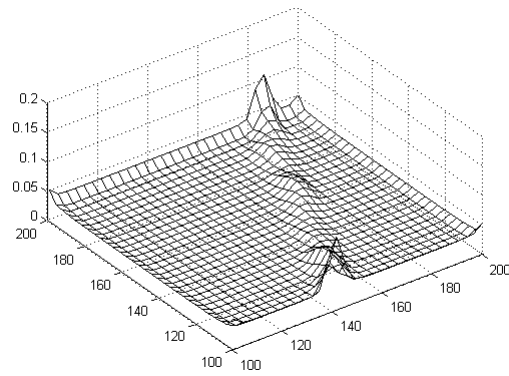
Si è scelto intenzionalmente un numero di punti da campionare così basso per mettere in evidenza le differenze tra un campionamento con tante classi rispetto ad una sola.

Considerando il grafico della superficie (figura 5.2), si desidererebbe che i punti campionati siano il più possibile all'interno della scanalatura ma, come si nota dalla figura 5.4, questo accade raramente nelle tre realizzazioni a causa della piccola dimensione campionaria. Osservando le tre figure congiuntamente è però possibile giungere alla conclusione che i punti presenti nel solco sono in numero maggiore rispetto a quelli estratti nelle altre zone.

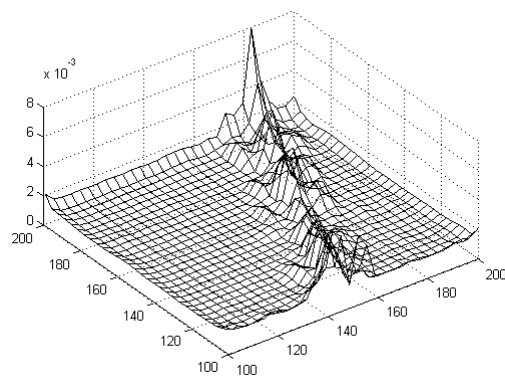
Utilizzando un grande numero di classi (figura 5.5) si riesce, invece, a guidare il campionamento poichè il numero di punti presenti in ogni classe è scelto secondo un metodo deterministico, e solo all'interno della classe vengono eseguiti i campionamenti. Si nota che in questo caso ci sono almeno 15 classi in cui non è presente neanche un punto campionato.



(a) *Curvatura gaussiana*



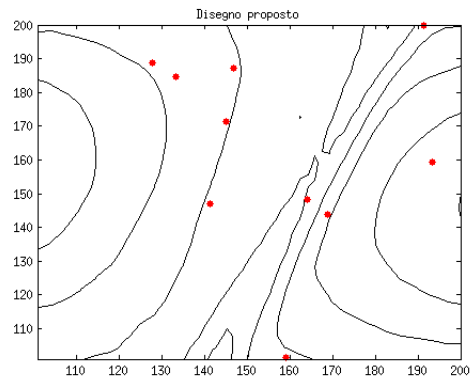
(b) *Semiampiezza intervallo di confidenza sulla media*



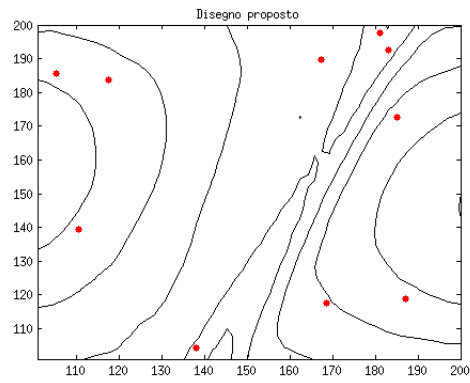
(c) *Densità di probabilità*

Figura 5.3: Densità di probabilità campionamento proposto

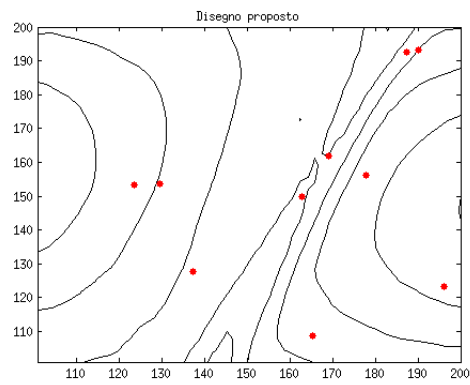




(a) *Disegno 1*

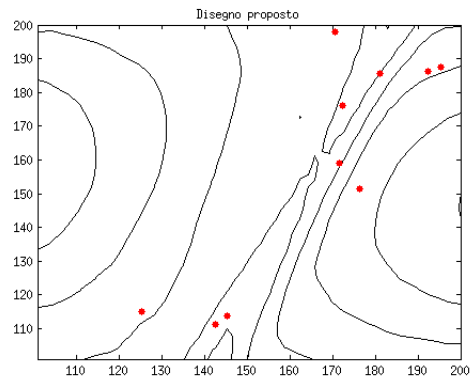


(b) *Disegno 2*

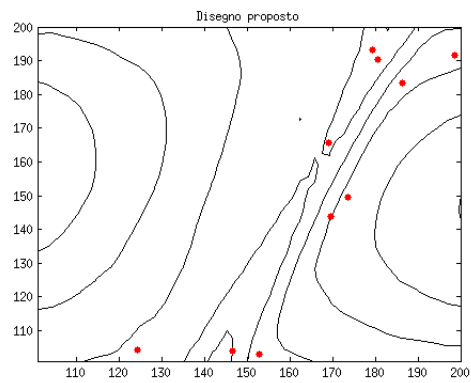


(c) *Disegno 3*

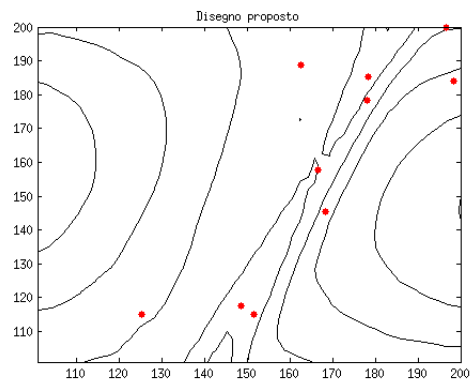
Figura 5.4: Campionamenti con 1 classe



(a) *Disegno 1*



(b) *Disegno 2*



(c) *Disegno 3*

Figura 5.5: Campionamenti con 25 classi

### Effetto dei pesi sulla funzione obiettivo

Per la valutazione dei pesi della funzione obiettivo,  $\omega_1$  e  $\omega_2$  si ricorre al campionamento della nuvola di punti del capitolo 4 poichè la superficie presenta un'elevata curvatura localizzata in una sua parte e, nel caso del modello a bassa risoluzione con decimazione quadratica dei punti, sono presenti dei grandi effetti di bordo, per cui si può valutare il contributo dell'incertezza e della curvatura separatamente. Non sono stati eseguiti dei modelli *fusion* su quel set di dati poichè le nuvole di punti sono state allineate e quindi non è possibile distinguere gli errori introdotti dall'allineamento da quelli propri dello strumento a bassa accuratezza.

L'analisi sarà eseguita in modo tale che  $\omega_1 + \omega_2 = 1$  per valori di  $\omega_1$  pari a 0,1, 0,3, 0,5, 0,7 e 0,9. I valori campionati e i grafici della curvatura gaussiana, dell'incertezza e della densità di probabilità per ognuno di questi valori sono riportati, rispettivamente, nelle figure 5.6, 5.7, 5.8, 5.9 e 5.10; i campionamenti sono stati eseguiti con un totale di 9 classi, ognuna formata da 100 punti candidati.

Con un peso della curvatura molto basso ( $\omega_1 = 0,1$ , figura 5.6), un buon numero di punti è posizionato sulla frontiera del dominio, mentre all'interno appaiono distribuiti uniformemente. Questo accade a causa del peso quasi nullo attribuito alla curvatura gaussiana perché, fatta eccezione ai bordi, l'incertezza è pressoché costante nelle altre zone della superficie.

Diminuendo invece l'importanza dell'incertezza ( $\omega_1 = 0,3$ , figura 5.7), i punti lungo i bordi diventano più radi ed iniziano ad essere presenti con una densità superiore nelle zone dove il valore assoluto della curvatura gaussiana presenta valori elevati.

Se invece si assegna un uguale peso ai due valori standardizzati ( $\omega_1 = \omega_2 = 0,5$ , figura 5.8), aumenta il numero di punti presenti nelle zone in cui la curvatura è massima.

Con  $\omega_1$  pari a 0,7 (figura 5.9) si sono invece ridotti gli effetti di bordo dell'incertezza e i punti sono presenti in quantità maggiori dove il valore assoluto della curvatura gaussiana presenta valori elevati, anche in zone lontane dal picco.

Nell'ultimo campionamento con  $\omega_1$  pari a 0,9 (figura 5.10), sono trascurabili gli effetti di bordo ed i punti si posizionano nelle zone in cui il valore assoluto della curvatura gaussiana presenta dei massimi.

I pesi da attribuire alle due grandezze non dovrebbero mai essere dei valori limite, come 0,9, poichè si privilegierebbe un aspetto della superficie rispetto ad un altro. Nel caso sarebbe attribuito un peso alto alla curvatura gaussiana, non si prenderebbe in considerazione l'incertezza sui punti che sono stati utilizzati per ottenerla, mentre se si darebbe un peso molto elevato all'incertezza non si prenderebbe in considerazione la complessità della forma dell'oggetto. I pesi dovrebbero assumere valori simili, quindi vicini a 0,5, in modo da considerare entrambi gli aspetti.

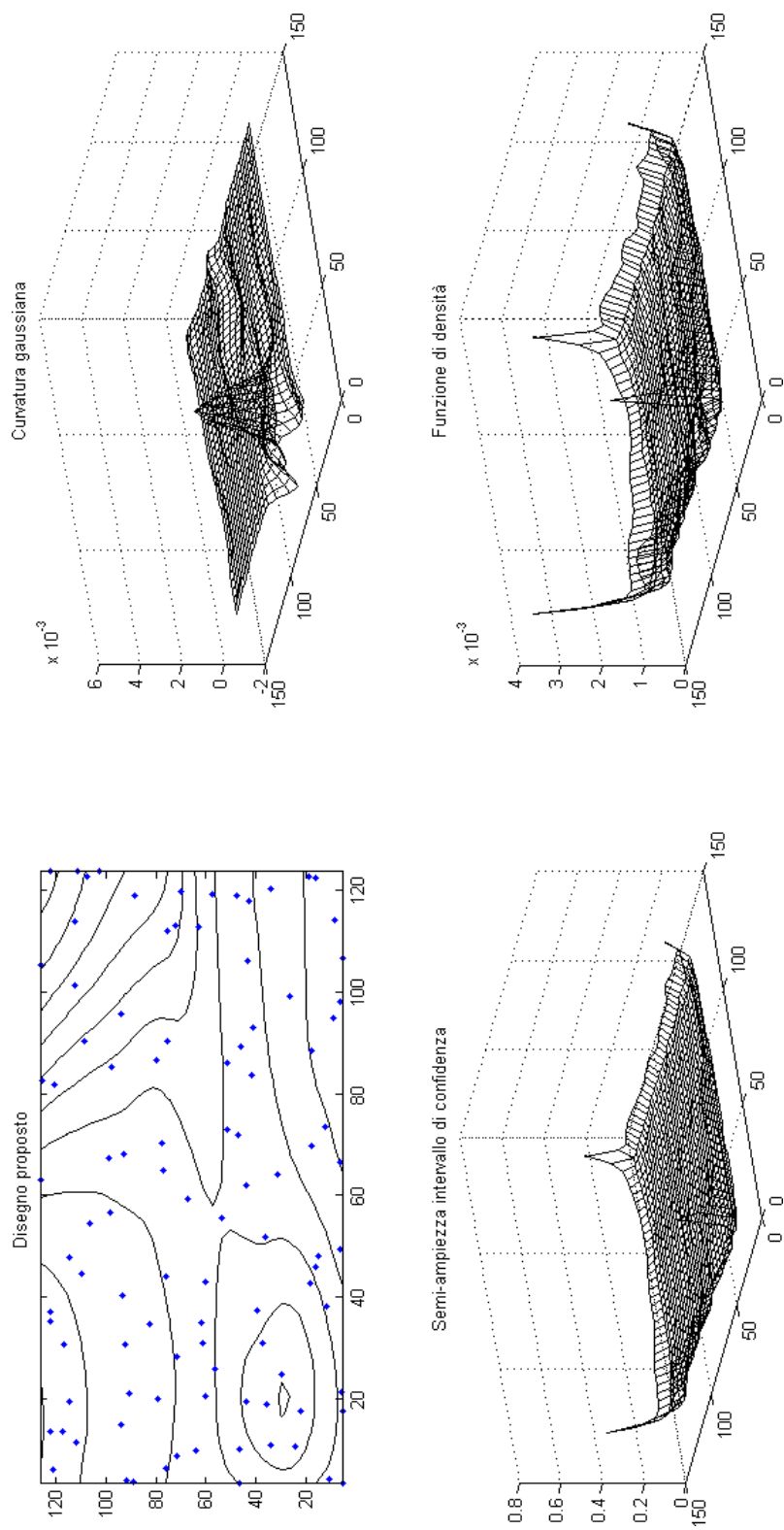


Figura 5.6:  $\omega_1 = 0,1$ ,  $\omega_2 = 0,9$

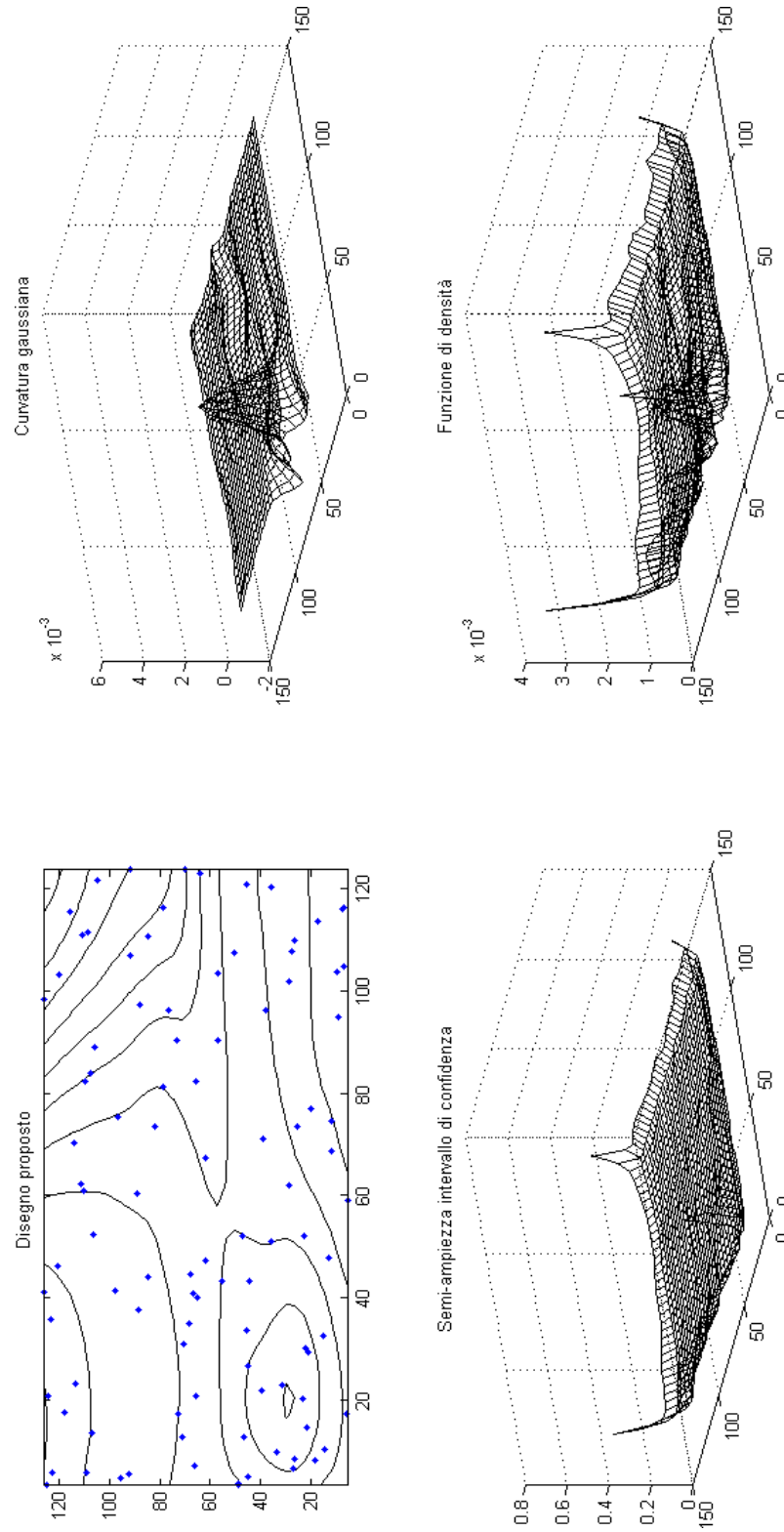


Figura 5.7:  $\omega_1 = 0,3$ ,  $\omega_2 = 0,7$

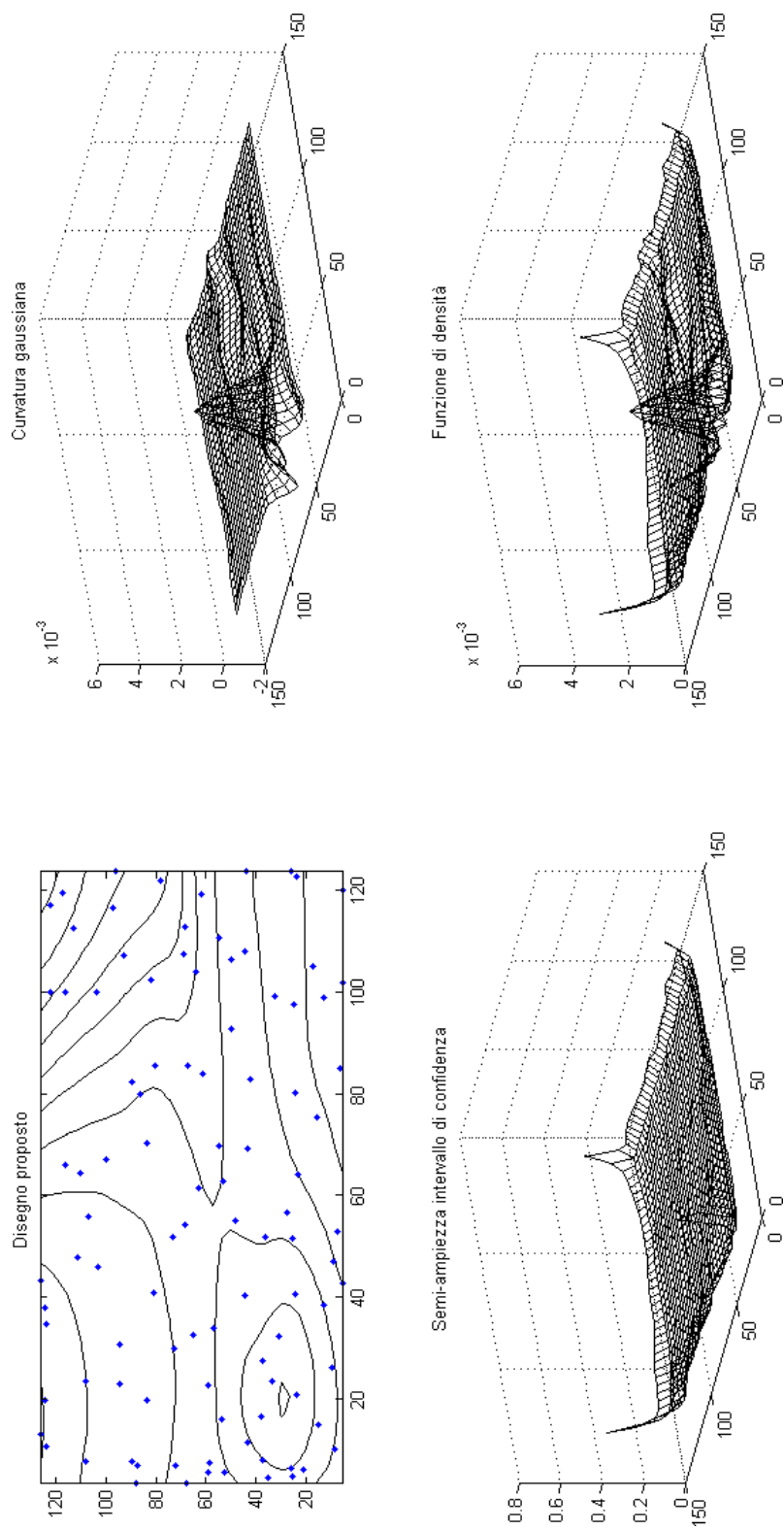


Figura 5.8:  $\omega_1 = \omega_2 = 0,5$

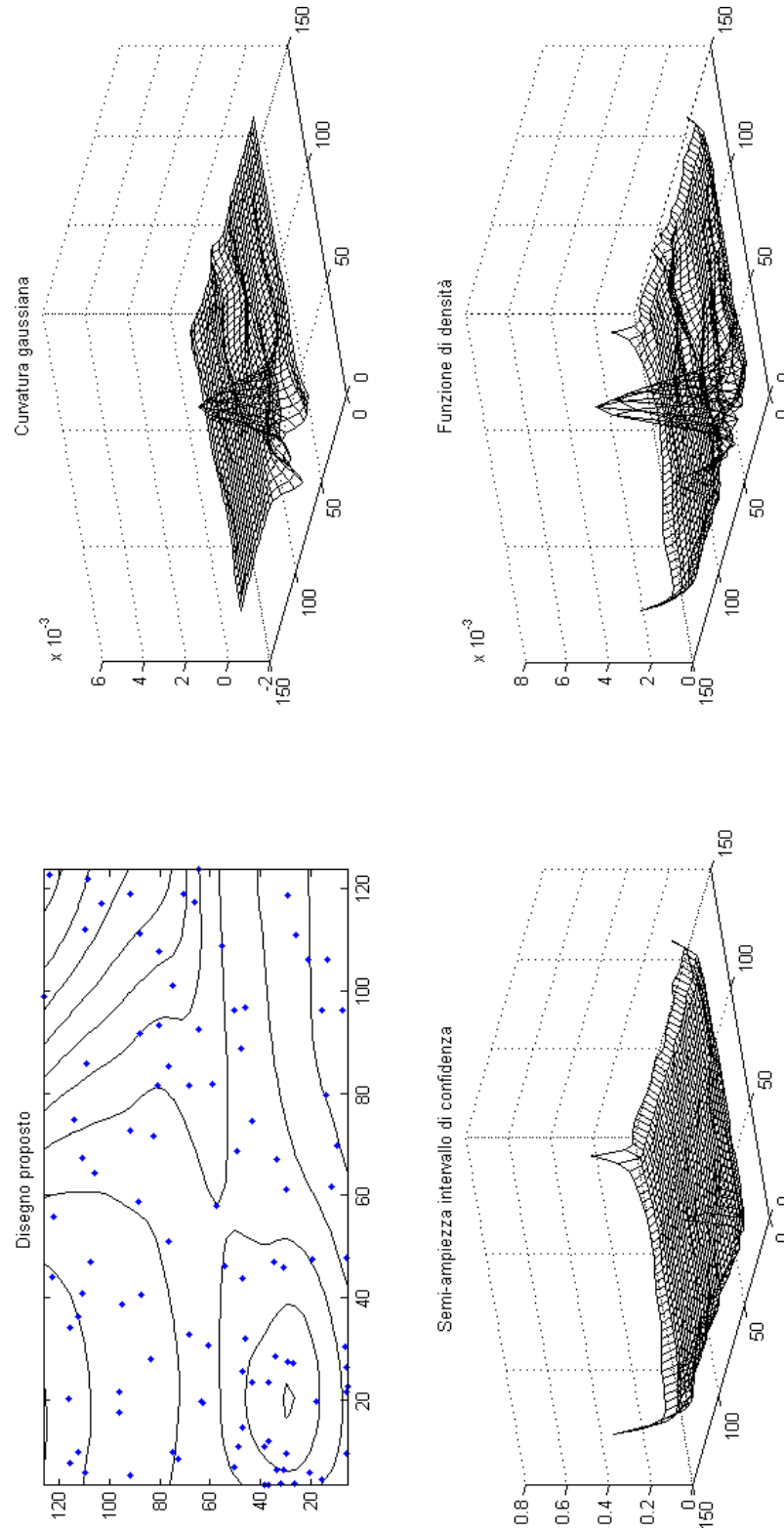


Figura 5.9:  $\omega_1 = 0,7$ ,  $\omega_2 = 0,3$

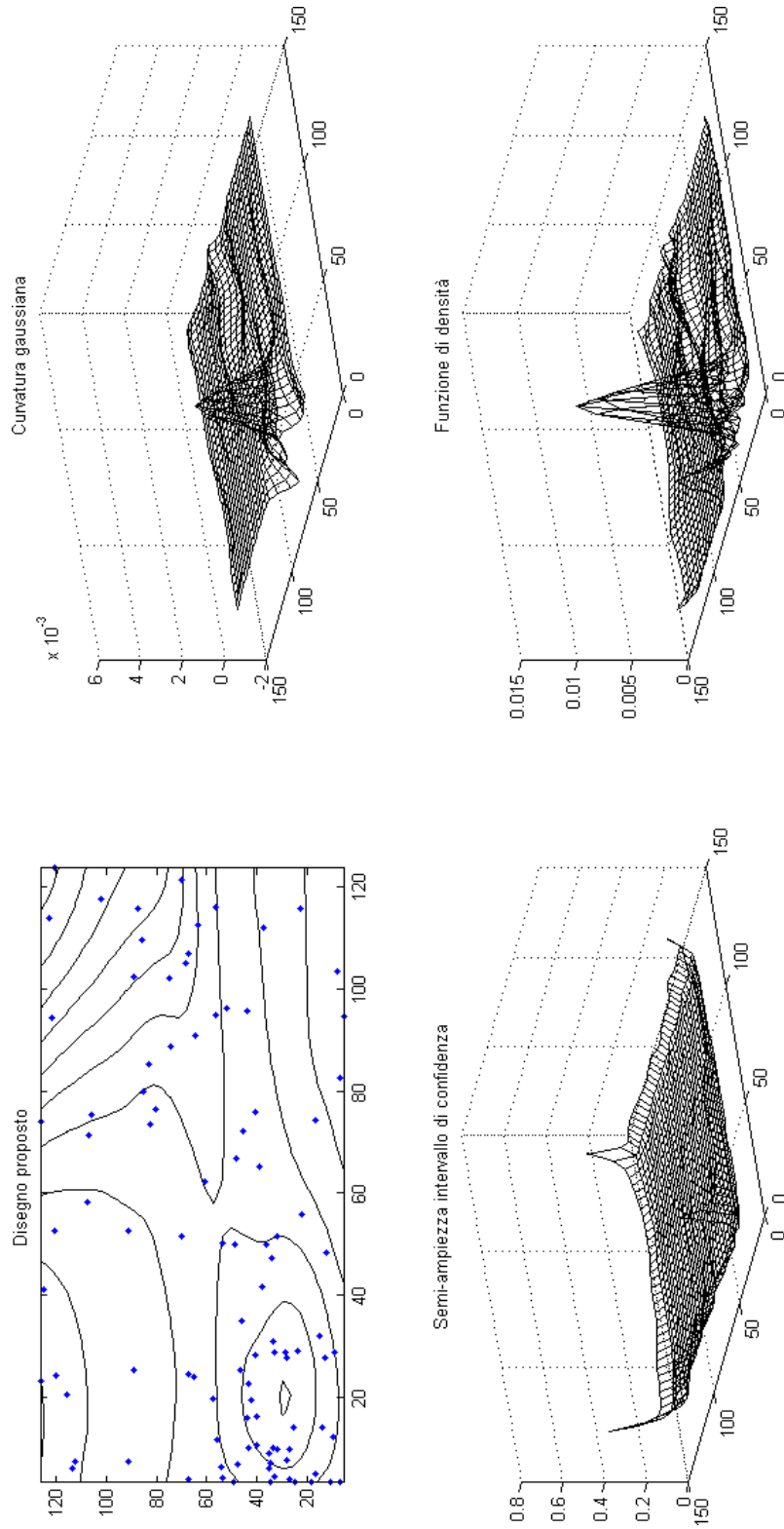


Figura 5.10:  $\omega_1 = 0,9$ ,  $\omega_2 = 0,1$



Un caso in cui sarebbe opportuno assegnare un peso elevato all'incertezza di predizione può essere quando la superficie acquisita è molto regolare, per cui la curvatura gaussiana è approssimabile con una costante.

Il caso opposto, cioè l'attribuzione di un peso elevato alla curvatura gaussiana, si ha quando l'incertezza di misura non varia in tutto il dominio della funzione. In questi due casi limite una caratteristica ha un peso molto maggiore dell'altra poichè se una delle due è costante sarebbe come estrarre un punto a caso, con un campionamento stratificato se si utilizzano delle classi.

### 5.3 Analisi

In questo paragrafo si valutano le prestazioni di un modello *fusion* basato sul campionamento proposto, su un campionamento mediante quadrati latini o su un campionamento misto tra il quadrato latino ed il campionamento proposto. Il campionamento mediante quadrati latini viene spesso utilizzato nel *computer experiment* [38] per permettere di avere informazioni sull'intera regione di sperimentazione ed è inoltre stato utilizzato in [47] nel disegno dei punti ad alta risoluzione per valutare le prestazioni modello gerarchico proposto dagli autori. Si considera solo il caso del modello in cui la fusione avviene in modo gerarchico poichè si vuole valutare la differenza di prestazioni dei modelli citati in precedenza causata da un diverso metodo di campionamento dei punti ad alta accuratezza.

Si sono effettuate quattro analisi indipendenti, ognuna caratterizzata da dieci repliche.

Nelle prime tre analisi vengono utilizzati i punti ottenuti dalla superficie (5.7) mediante una griglia regolare formata da 50 punti lungo ogni direzione, si hanno quindi in totale 2.500 punti teoricamente disponibili per la stima; il dominio della funzione corrisponde a:

$$D : \{(x, y) \in \mathbb{R}^2 \text{ t.c. } 100 \leq x \leq 200, 100 \leq y \leq 200\}.$$

Nella quarta analisi si utilizza la stessa superficie, ma cambia il dominio:

$$D : \{(x, y) \in \mathbb{R}^2 \text{ t.c. } 50 \leq x \leq 250, 50 \leq y \leq 250\}.$$

Per simulare una perdita di informazioni dello scanner a luce strutturata nella zona in cui è presente il solco, si è scelto di utilizzare il 60% dei punti nella fascia in cui l'altezza del solco è compresa tra il 40% e l'80%, mentre il 20% nella parte più profonda della scanalatura.

Si sono utilizzati un totale di 50 punti ad alta risoluzione, che corrispondono ad  $\frac{1}{50}$  di quelli ad alta densità.

Nella prima analisi per simulare i punti a bassa risoluzione si aggiunge un errore casuale con varianza 0,0012 [26] e media nulla, mentre per quelli ad elevata accuratezza l'errore casuale aggiunto ha una varianza pari a  $4 \cdot 10^{-6}$  [1]. Lo scopo

della prima analisi è valutare se il modello *fusion* sui dati ad alta accuratezza del modello proposto ottiene dei risultati migliori di un piano di campionamento comunemente utilizzato nel *computer experiment* nelle condizioni in cui dovrebbe ottenere dei buoni risultati.

Con la seconda analisi si vuole invece valutare se il metodo di campionamento proposto permette di ottenere dei buoni risultati anche in presenza di errori sistematici del modello a bassa risoluzione, che non siano individuabili dalle stime fatte con il modello di primo stadio, per cui possono essere indipendenti dalla curvatura gaussiana e dall'incertezza di predizione del modello di primo stadio.

Nella terza e quarta analisi si considera il caso della metrologia di larga scala in cui si ipotizza che lo strumento utilizzato per l'acquisizione di una nuvola di punti molto densa, ad esempio la luce strutturata, abbia un'incertezza di misura minore dell'altro strumento, come il sistema per metrologia di larga scala MScMs-II descritto precedentemente. Si ipotizza che questo strumento, seppur abbia un'accuratezza peggiore, non commetta degli errori sistematici nella misura, come invece accade per lo scanner a luce strutturata. L'errore casuale dei punti con densità maggiore è stato generato da una variabile aleatoria normale con media zero e varianza pari a 0,00125, come nei primi due casi, mentre quello dei punti con numerosità minore è generato dalle realizzazioni di una normale con media zero e varianza pari a 0,0049, pari a quattro volte quella dello scanner a luce strutturata.

Lo scopo della terza analisi è valutare se il metodo proposto per il campionamento dei punti ottiene risultati simili ai disegni standard quando sono presenti due fattori, incertezza e errore non casuale. La maggiore incertezza renderebbe preferibile lo strumento che permette di acquisire un numero maggiore di punti e l'errore non casuale di quest'ultimo rende però necessario ricorrere ad un altro strumento di misura.

Nella quarta analisi si considera una diversa superficie in cui l'errore dovuto alla bassa densità di campionamento è poco significativo rispetto a quello sistematico della luce struttura. In questo caso il campionamento con un quadrato latino dovrebbe ottenere prestazioni superiori rispetto a quello con campionamento proposto perché sono le sue condizioni ottimali.

### 5.3.1 Analisi 1

Il disegno ottenuto mediante il modello proposto è riportato in figura 5.11. Il campionamento è stato eseguito con un numero di classi pari a nove poiché 50 punti non rappresentano un numero molto elevato se si tiene in considerazione la complessità della superficie analizzata. I punti campionati mediante il disegno proposto sono situati per la maggior parte

all'interno del solco e nella parte alla sua sinistra. Nella parte destra ne sono presenti pochi poichè non è presente una grande variazione della curvatura e in particolare sono situati nella parte vicino alla frontiera del dominio a causa degli effetti di bordo dell'incertezza di predizione. Nella parte di sinistra della superficie invece i punti sono stati campionati in modo all'incirca uniforme grazie alla suddivisione in classi che ha permesso di non avere punti nei bordi della superficie.

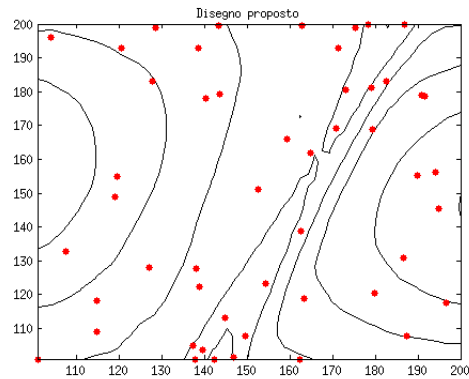


Figura 5.11: Disegno proposto

Il disegno con quadrato latino (figura 5.12) invece campiona i punti senza tenere in considerazione la topologia della superficie poichè l'idea alla base della sua costruzione è che non si hanno a disposizione informazioni sulle caratteristiche della superficie (per cui nessuna zona della superficie sarà privilegiata).

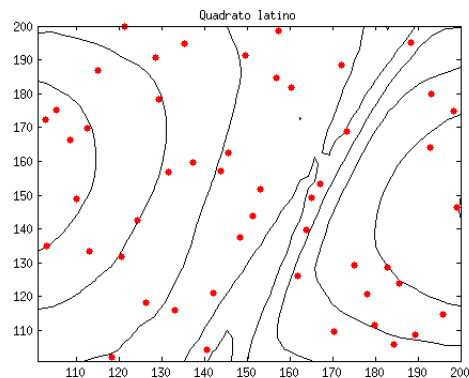


Figura 5.12: Quadrato latino

Il disegno misto (figura 5.13) è formato da 10 punti del campionamento proposto (rosso) e 40 di un quadrato latino (nero), la maggior parte dei punti proviene

da un campionamento mediante quadrato latino poichè il solco rappresenta solo un porzione della superficie.

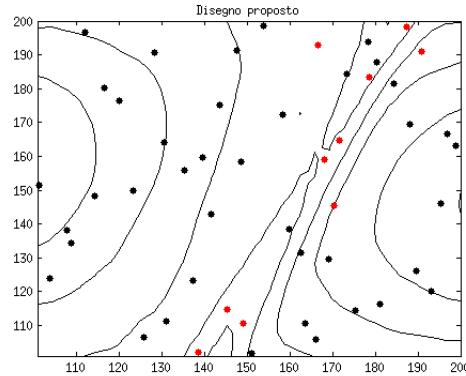


Figura 5.13: Disegno misto

I modelli vengono valutati sulla bontà della previsione di 400 punti, posizionati su una griglia regolare di  $20 \times 20$  punti. I residui sono calcolati come la differenza tra il valore vero, disponibile poichè i dati sono stati generati artificialmente, ed il valore predetto da ciascun modello. I valori su cui si basa il confronto sono la medie dei residui di ogni singola replica e le medie dei loro quadrati. Si utilizza la media al posto della mediana rispetto al caso precedente poichè si può sfruttare il teorema centrale del limite e i valori estremi sono importanti in quanto la parte dove la superficie a bassa risoluzione è meno densa è solo una porzione dell'intera superficie.

I modelli oggetto dell'analisi sono i tre modelli *fusion*, basati sul campionamento proposto, sul campionamento con quadrati latini e su un campionamento misto, ed il modello a bassa risoluzione poichè si vuole verificare se l'utilizzo di misurazioni costose sono giustificate da un miglioramento della capacità predittiva e se un diverso campionamento di questi punti può influenzare il risultato.

Come anticipato si valuteranno due misure:

- le medie degli errori, per verificare se il modello non commette errori sistematici. Infatti se la media degli errori risulta minore (maggiore) di zero si ha una previsione che sovra (sotto) stima il valore reale;
- le medie dei quadrati degli errori, per verificare se i modelli hanno prestazioni uguali tra loro. Infatti elevando gli errori al quadrato i valori con un elevato valore assoluto non si compensano se hanno segno opposto.

Per valutare la media degli errori si utilizza la statistica:

$$T_1^2 = n\bar{\mathbf{x}}_1' \mathbf{S}_1^{-1} \bar{\mathbf{x}}_1 \quad (5.12)$$

dove

$$\bar{\mathbf{x}}_1 = \frac{1}{n} \sum_{i=1}^n \mathbf{x}_i^1, \quad \bar{\mathbf{x}}_1 \in \mathbb{R}^q, \quad \mathbf{S}_1 = \frac{1}{n-1} \sum_{i=1}^n (\mathbf{x}_i^1 - \bar{\mathbf{x}}_1)(\mathbf{x}_i^1 - \bar{\mathbf{x}}_1)', \quad \mathbf{S}_1 \in (q, q)$$

in cui  $\mathbf{x}_i^1$  si usa per indicare che ci si riferisce al dataset delle medie degli errori, le quali sono delle realizzazioni di  $\mathbf{X}_1 \sim \mathcal{N}_q(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1)$ .

Il test di ipotesi è così costruito:

$$\begin{cases} H_0 : \boldsymbol{\mu}_1 = \mathbf{0}_q \\ H_1 : \boldsymbol{\mu}_1 \neq \mathbf{0}_q. \end{cases}$$

Se è vera  $H_0$  allora:

$$T_1^2 \sim \frac{q(n-1)}{n-q} \mathcal{F}(q, n-q) \quad (5.13)$$

Per la valutazione simultanea del quadrato degli errori la statistica utilizzata è:

$$T_2^2 = n(\mathbf{C}\bar{\mathbf{x}}_2)' (\mathbf{C}\mathbf{S}_2\mathbf{C}')^{-1} \mathbf{C}\bar{\mathbf{x}}_2 \quad (5.14)$$

dove

$$\mathbf{C} = \begin{bmatrix} -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{C} \in (q-1, q)$$

$$\bar{\mathbf{x}}_2 = \frac{1}{n} \sum_{i=1}^n \mathbf{x}_i^2, \quad \bar{\mathbf{x}}_2 \in \mathbb{R}^q, \quad \mathbf{S}_2 = \frac{1}{n-1} \sum_{i=1}^n (\mathbf{x}_i^2 - \bar{\mathbf{x}}_2)(\mathbf{x}_i^2 - \bar{\mathbf{x}}_2)', \quad \mathbf{S}_2^{-1} \in (q, q)$$

con  $\mathbf{x}_i^2$  ci si riferisce al set di dati che rappresenta le medie del quadrato degli errori, che sono realizzazioni della variabile aleatoria  $\mathbf{X}_2 \sim \mathcal{N}_q(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2)$ . Si ricorda che la statistica  $T^2$  non dipende dalla particolare matrice di contrasti  $\mathbf{C}$  scelta, il test di ipotesi è così costruito:

$$\begin{cases} H_0 : \mathbf{C}\boldsymbol{\mu}_2 = \mathbf{0}_{q-1} \\ H_1 : \mathbf{C}\boldsymbol{\mu}_2 \neq \mathbf{0}_{q-1}. \end{cases}$$

Sotto  $H_0$  si ha che:

$$T_2^2 \sim \frac{(q-1)(n-1)}{n-q+1} \mathcal{F}(q-1, n-q+1) \quad (5.15)$$

Se uno di questi test non è superato con un errore di I tipo pari all'1% è necessario eseguire dei test sulle singole variabili per capire quale ha causato il

non superamento del test.

Si inizia l'analisi presentando il vettore delle medie:

$$\bar{\mathbf{x}}_1 = \begin{pmatrix} -0,0024 \\ 0,0007 \\ -0,0026 \\ 0,0022 \end{pmatrix}$$

in cui la prima componente rappresenta il modello a bassa risoluzione, la seconda il modello *fusion* con il campionamento proposto, la terza il modello *fusion* con un quadrato latino e la quarta il modello *fusion* con campionamento misto.

Con un *p-value* pari a 2,8677e-06 si può rifiutare l'ipotesi nulla di uguaglianza delle medie a zero. I *p-value* dei *t-test* per i vari modelli sono i seguenti:

$$\begin{aligned} LR &= 2,114e-08 \\ Fusion &= 0,2493 \\ Fusion\_ql &= 9,422e-05 \\ Fusion\_misto &= 0,0002 \end{aligned}$$

Tabella 5.1: *P-value* test uguaglianza media a zero

dove *LR* si riferisce al modello a bassa risoluzione, *Fusion* al modello *fusion* con il campionamento proposto, *Fusion\_ql* al modello *fusion* con il campionamento che deriva da un quadrato latino e *Fusion\_misto* al modello *fusion* con un campionamento misto.

Con un errore di I tipo maggiore o uguale all'1%, mediante un approccio Bonferroni, si può affermare che l'unica media che non risulta statisticamente diversa da zero è quella del modello *fusion* con il campionamento proposto, il che indica che gli altri modelli, in media, effettuano una piccola traslazione dei punti in fase di predizione rispetto al vero valore.

Il vettore delle medie calcolate sui quadrati dei residui risulta:

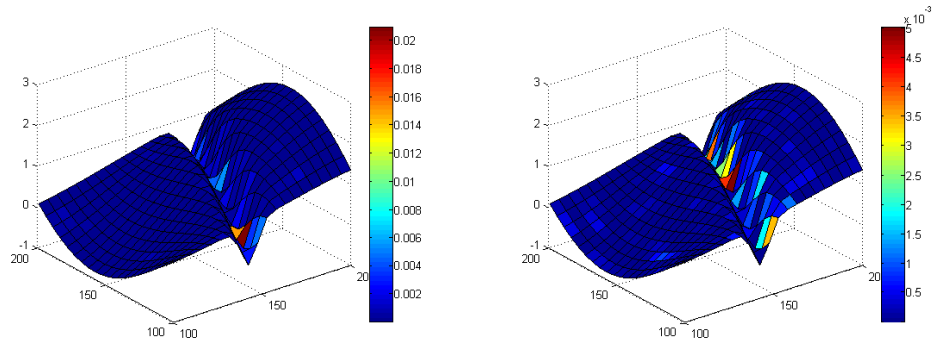
$$\bar{\mathbf{x}}_2 = \begin{pmatrix} 3,8033e-04 \\ 2,3329e-04 \\ 3,5254e-04 \\ 2,2175e-04 \end{pmatrix}$$

Con un *p-value* pari a 1,438e-11 si può rifiutare l'ipotesi nulla di uguaglianza delle medie dei residui nei vari modelli. Si presentano ora i *p-value* dei *t-test* di uguaglianza delle medie dei confronti a coppie:

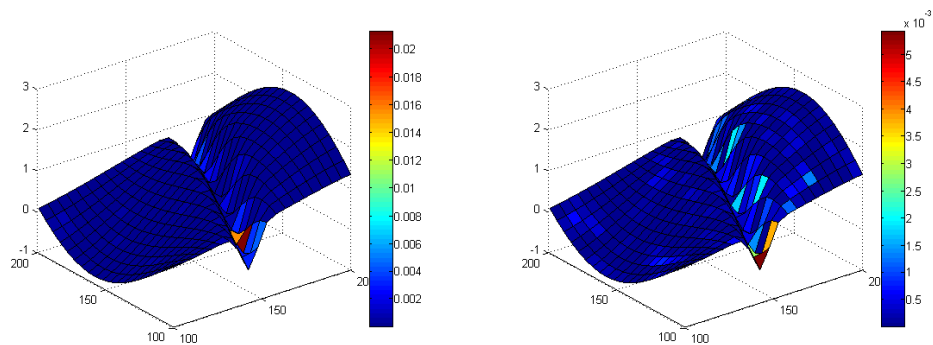
$$\begin{aligned}
 LR \text{ vs } Fusion &= 8,4427e-07 \\
 LR \text{ vs } Fusion_{ql} &= 0,0014 \\
 LR \text{ vs } Fusion_{misto} &= 6,0245e-06 \\
 Fusion \text{ vs } Fusion_{ql} &= 2,3177e-05 \\
 Fusion \text{ vs } Fusion_{misto} &= 0,1833 \\
 Fusion_{ql} \text{ vs } Fusion_{misto} &= 3,4885e-05
 \end{aligned}$$

Tabella 5.2: *P-value* confronti a coppie medie quadrati errori

Con lo stesso errore di I tipo del caso precedente (1%) non si può affermare che i modelli *fusion* con il campionamento proposto ed il modello *fusion* con campionamento misto siano diversi tra loro, per tutti gli altri confronti a coppie c'è significatività statistica per rifiutare  $H_0$ .



(a) Quadrato errori modello bassa risoluzione (b) Quadrato errori modello *fusion* con campionamento proposto



(c) Quadrato errori modello *fusion* con quadrato latino (d) Quadrato errori modello *fusion* con campionamento misto

Figura 5.14: Superficie con quadrato errori

Le differenze tra i vari modelli possono essere spiegate dai grafici di una replica, riportati in figura 5.14, in cui i colori rappresentano il quadrato degli errori

nel punto preso in considerazione. Le differenze tra i vari modelli sono presenti nella parte di superficie in cui è presente il solco poichè i modelli *LR* e *fusion* con quadrato latino commettono degli errori in valore assoluto molto elevati. Gli altri due modelli permettono invece di attenuare questo effetto, anche se non riescono ad eliminarlo del tutto, poichè basati su modelli gerarchici che presumono che le predizioni del modello dello stadio precedente siano buone.

In seguito all'analisi dei risultati si può concludere che il processo gaussiano con il campionamento proposto e con un campionamento misto ha portato dei miglioramenti di prestazioni rispetto al modello meno accurato a bassa risoluzione, mentre con un campionamento mediante un quadrato latino seppur siano presenti alcuni miglioramenti, non sono confrontabili con quelli degli altri due modelli *fusion*.

### 5.3.2 Analisi 2

Per simulare l'errore non casuale si è aggiunto alla superficie (5.7) un paraboloide ellittico con il minimo in corrispondenza del centro della superficie oggetto dell'analisi. Il massimo valore di questo errore è localizzato sui bordi della superficie ed è pari a 5 centesimi di millimetro (figura 5.15).

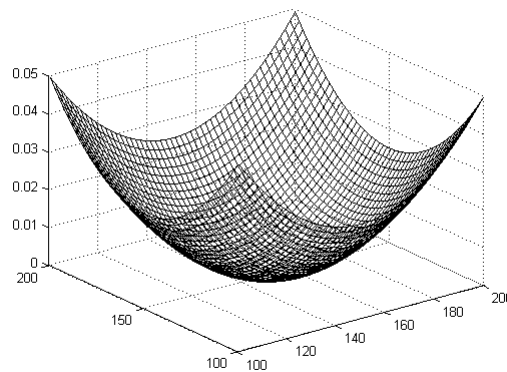


Figura 5.15: Errore aggiunto alla superficie

I test statistici e l'errore di I tipo sono gli stessi utilizzati nell'analisi 1. Si riporta il vettore delle medie dei residui:

$$\bar{\mathbf{x}}_1 = \begin{pmatrix} -0,0208 \\ 0,0011 \\ -0,0020 \\ -0,0011 \end{pmatrix}$$

Si nota subito che la media associata al modello a bassa risoluzione ha un valore che è superiore di un ordine di grandezza rispetto a quelle degli altri



modelli infatti il *p-value* del confronto simultaneo è pari a 1,4742e-09, il che permette di affermare che esiste almeno una media diversa da zero. Per verificare di quale si tratta si eseguono i singoli test:

$$\begin{aligned} LR &= 5,129e-15 \\ Fusion &= 0,0458 \\ Fusion_{ql} &= 0,0003 \\ Fusion_{misto} &= 0,0456 \end{aligned}$$

Tabella 5.3: *P-value* test uguaglianza media a zero

Analizzando *p-value* dell'uguaglianza delle medie a zero si conclude che il modello a bassa risoluzione non si accorge dell'errore che è stato aggiunto, com'era lecito attendersi, mentre il modello con quadrato latino, anche se effettua una correzione, tende ad una piccola sovrastima dei punti della superficie. I modelli in cui è stato utilizzato il campionamento proposto effettuano una correzione dei punti a bassa risoluzione ottenendo una media che non si può affermare sia diversa da zero. Si passa ora all'analisi delle medie dei quadrati degli errori, il valore puntuale risulta:

$$\bar{x}_2 = \begin{pmatrix} 9,9249e-04 \\ 3,0047e-04 \\ 4,0034e-04 \\ 3,2038e-04 \end{pmatrix}$$

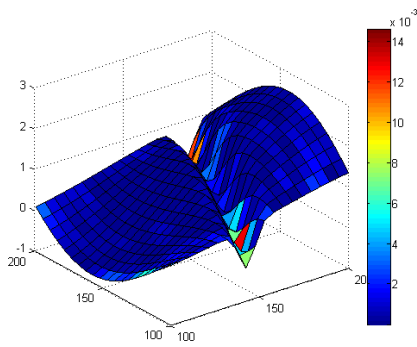
Il *p-value* di uguaglianza delle varie medie risulta 8,9295e-13, quindi si riportano i valori dei *p-value* dei confronti a coppie:

$$\begin{aligned} LR \text{ vs } Fusion &= 9,0110e-12 \\ LR \text{ vs } Fusion_{ql} &= 4,2042e-11 \\ LR \text{ vs } Fusion_{misto} &= 1,1796e-10 \\ Fusion \text{ vs } Fusion_{ql} &= 6,2894e-06 \\ Fusion \text{ vs } Fusion_{misto} &= 0,2444 \\ Fusion_{ql} \text{ vs } Fusion_{misto} &= 4,7245e-04 \end{aligned}$$

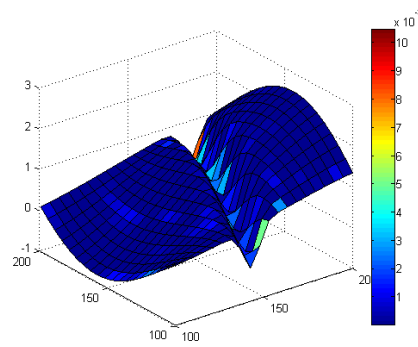
Tabella 5.4: *P-value* confronti a coppie medie quadrati errori

I confronti a coppie delle medie dei quadrati degli errori evidenziano che il modello peggiore, il modello a bassa risoluzione, è statisticamente diverso dagli altri con dei *p-value* molto bassi; mentre con c'è evidenza statistica per affermare che non siano uguali i risultati dei modelli *fusion* con il campionamento proposto e *fusion* con un disegno misto. Risultano invece diversi i valori delle medie dei quadrati degli errori dei modelli *Fusion* e *Fusion<sub>ql</sub>*, mentre, seppur di poco con un approccio Bonferroni, non si può affermare che i modelli *fusion* con campionamento misto e *fusion* con un disegno che deriva da un quadrato latino ottengano risultati diversi.

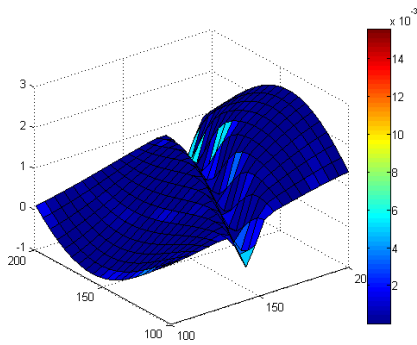
Osservando la distribuzione del quadrato degli errori di una sola replica (figura 5.16), si nota che il modello a bassa risoluzione commette errori sia nella parte del solco che nelle zone esterne della superficie. Il modello *fusion* con i punti ad alta accuratezza che derivano da un quadrato latino riesce a correggere gli errori sistematici (cioè il paraboloide ellittico aggiunto in simulazione), ma non è in grado di individuare gli errori nella parte in cui i punti a bassa accuratezza sono meno densi. I modelli *fusion* con il campionamento proposto e con il campionamento misto, invece, riuscendo ad eseguire entrambe le correzioni, ottengono risultati migliori degli altri modelli.



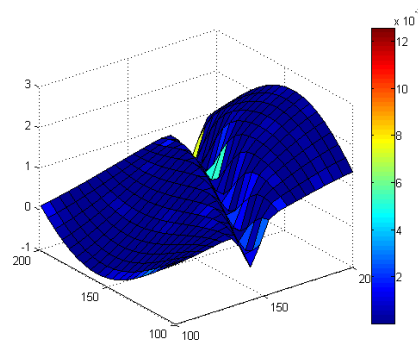
(a) Quadrato errori modello bassa risoluzione



(b) Quadrato errori modello fusion con campionamento proposto



(c) Quadrato errori modello fusion con quadrato latino



(d) Quadrato errori modello fusion con campionamento misto

Figura 5.16: Superficie con quadrato errori

A valle di questa analisi si può concludere che il campionamento proposto, e il campionamento misto tra quello proposto ed un quadrato latino, sono più adeguati per correggere il modello a bassa risoluzione nei casi analizzati. Inoltre tutti i modelli *fusion* individuano e di conseguenza correggono gli errori sistematici simulati nei punti a bassa accuratezza.

### 5.3.3 Analisi 3

Si utilizza la stessa superficie presente nelle due analisi precedenti, ma in questo caso il paraboloide ellittico aggiunto per simulare l'errore dello scanner a luce strutturata ha un valore massimo più elevato che corrisponde a 0,25 millimetri.

Si presenta il vettore delle medie dei residui:

$$\bar{\mathbf{x}}_1 = \begin{pmatrix} -0,0943 \\ 0,0016 \\ -0,0062 \\ -0,0075 \end{pmatrix}$$

osservando i valori puntuali il modello *fusion* con campionamento proposto è quello caratterizzato da una media in valore assoluto più piccola. Con un *p-value* pari a 1,3182e-07 c'è evidenza statistica per affermare che esiste almeno una media il cui valore è diverso da zero, i *p-value* dei singoli *t-test* sono i seguenti:

$$\begin{aligned} LR &= 1,972e-12 \\ Fusion &= 0,6153 \\ Fusion\_ql &= 0,0600 \\ Fusion\_misto &= 0,0296. \end{aligned}$$

Tabella 5.5: *P-value* test uguaglianza media a zero

Analizzando i *p-value* si può concludere che l'unico modello la cui media è diversa da zero è quello a bassa risoluzione. Il vettore della media campionaria dei quadrati degli errori è il seguente:

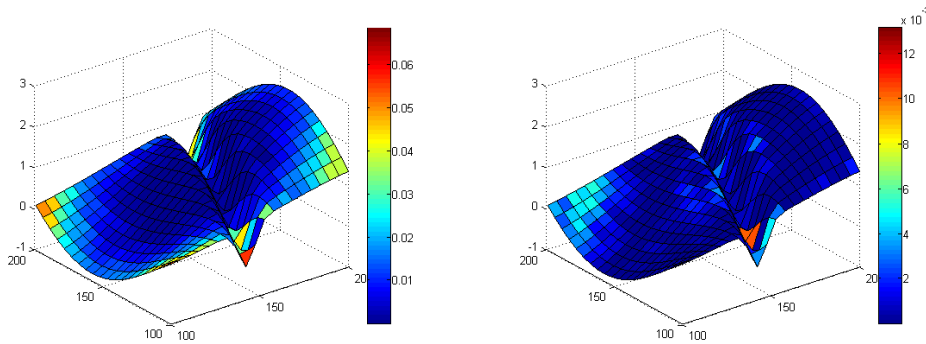
$$\bar{\mathbf{x}}_2 = \begin{pmatrix} 0,0127 \\ 0,0018 \\ 0,0025 \\ 0,0025 \end{pmatrix}$$

Il *p-value* associato alla statistica  $T^2$  di Hotelling è di 1,8757e-10 quindi anche in questo confronto esiste almeno una media diversa dalle altre. Di seguito si riportano i *p-value* dei confronti a coppie:

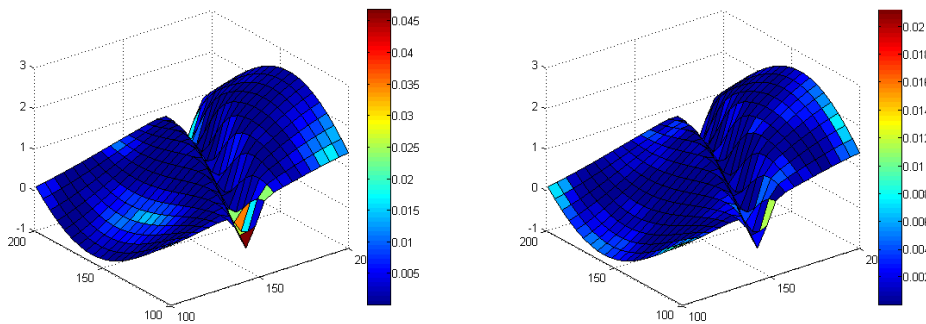
$$\begin{aligned} LR \text{ vs } Fusion &= 1,7058e-08 \\ LR \text{ vs } Fusion\_ql &= 2,6162e-08 \\ LR \text{ vs } Fusion\_misto &= 2,3630e-07 \\ Fusion \text{ vs } Fusion\_ql &= 0,1200 \\ Fusion \text{ vs } Fusion\_misto &= 0,0893 \\ Fusion\_ql \text{ vs } Fusion\_misto &= 0,9416. \end{aligned}$$

Tabella 5.6: *P-value* confronti a coppie medie quadrati errori

L'unica media dei quadrati degli errori che è diversa dalle altre è quella del modello a bassa risoluzione. Questo modello commette dei grandi errori soprattutto nella parte esterna della superficie dov'è maggiore l'errore non casuale inserito artificialmente e anche nella parte dov'è presente il solco. I modelli *fusion* invece, anche se caratterizzati da una maggiore dispersione, effettuano bene le correzioni, come si può notare in figura 5.17 in cui è rappresentata l'immagine del quadrato degli errori in ogni punto relativa ad una replica. Il modello con quadrato latino ha ottenuto le stesse prestazioni dei modelli in cui sono presenti i punti campionati secondo il metodo proposto poichè la maggior parte degli errori del modello *LR* sono dovuti all'errore aggiunto in fase di simulazione e gli errori dovuti alla presenza del solco non sono molto significativi. Tuttavia i modelli *fusion* con campionamento proposto e con un campionamento misto attenuano l'effetto dell'errore anche nella parte meno regolare della superficie.



(a) Quadrato errori modello bassa risoluzione (b) Quadrato errori modello *fusion* con campionamento proposto



(c) Quadrato errori modello *fusion* con quadrato latino (d) Quadrato errori modello *fusion* con campionamento misto

Figura 5.17: Superficie con quadrato errori

### 5.3.4 Analisi 4

L'errore sistematico dello scanner a luce strutturata viene simulato aggiungendo alla superficie un paraboloide ellittico (figura 5.18b) con il minimo in corrispondenza del centro della superficie riportata in figura 5.18a. La superficie utilizzata nella simulazione è presente nella (5.7).

Si è scelto di inserire un errore sistematico con un massimo valore che corrisponde ad una unità perché in questa condizione il modello basato sui punti che derivano dal campionamento mediante quadrato latino ottiene delle buone prestazioni.

La curvatura gaussiana (figura 5.19a), la semiampiezza dell'intervallo di confidenza sulla media (figura 5.19b) e la densità di probabilità da cui sono stati estratti i punti (figura 5.19c) sono state ricavate dalla predizione di un totale di 900 punti a bassa risoluzione.

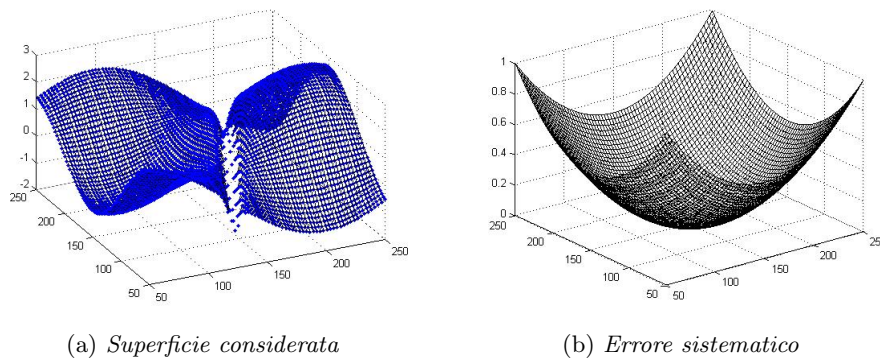
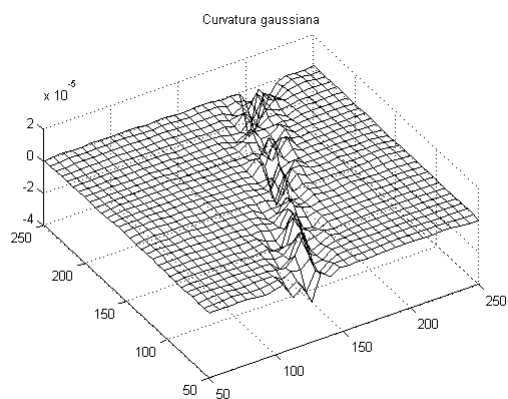


Figura 5.18: Superficie ed errore sistematico

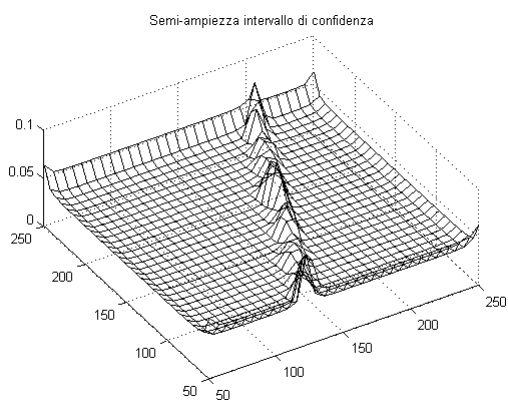
I punti campionati secondo il metodo proposto sono riportati in figura 5.20, la maggior parte sono localizzati nelle zone a minor densità di punti e dov'è presente una curvatura maggiore. Di conseguenza non sono distribuiti in maniera uniforme.

Il campionamento mediante quadrato latino (figura 5.21), a causa della sua progettazione, ha i punti che sono distribuiti lungo tutta la superficie senza una differenziazione tra le varie zone.

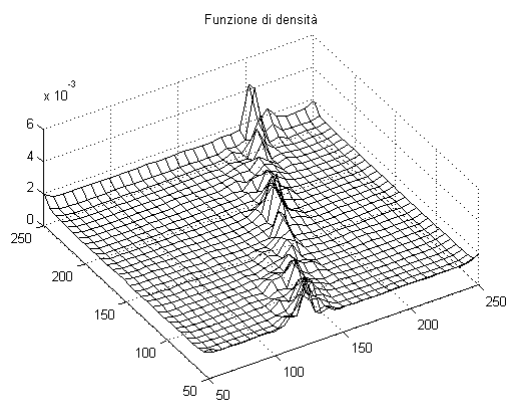
Il campionamento misto (figura 5.22) è stato effettuato come nella superficie precedente con 10 punti (rosso) che corrispondono al campionamento secondo il modello proposto e 40 (nero) che derivano da un campionamento con quadrato latino.



(a) *Curvatura gaussiana*



(b) *Semiampiezza intervallo di confidenza sulla media*



(c) *Densità di probabilità*

Figura 5.19: Densità di probabilità campionamento proposto

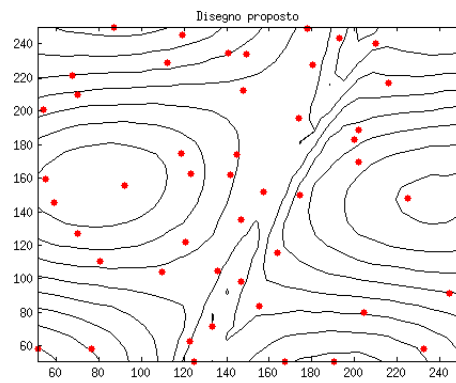


Figura 5.20: Campionamento con disegno proposto

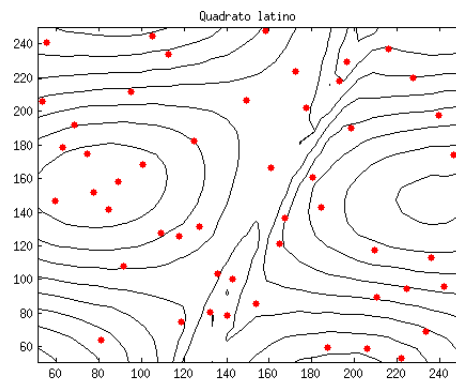


Figura 5.21: Campionamento con quadrato latino

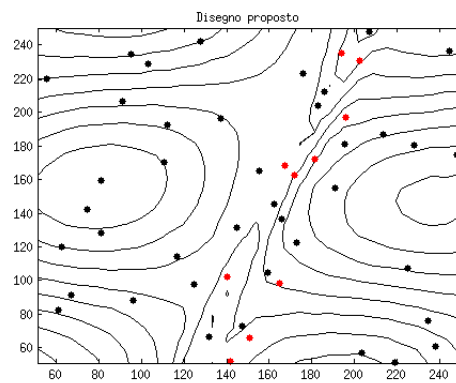


Figura 5.22: Campionamento con metodo misto

Dopo l'esecuzione di 10 repliche, si presentano le medie degli errori:

$$\bar{\mathbf{x}}_1 = \begin{pmatrix} -0,3689 \\ -0,0081 \\ -0,0024 \\ -0,0158 \end{pmatrix}$$

La media degli errori di predizione del modello a bassa risoluzione è l'unica che ha un valore assoluto superiore rispetto alle altre, infatti c'è significatività statistica (tabella 5.7) per affermare che sia diversa da zero. La medesima conclusione si può trarre per il modello che deriva dal campionamento misto.

$$\begin{aligned} LR &= < 2,2e-16 \\ Fusion &= 0,0332 \\ Fusion\_ql &= 0,5939 \\ Fusion\_misto &= 0,0001. \end{aligned}$$

Tabella 5.7: *P-value* test uguaglianza media a zero

Si riportano di seguito le medie dei quadrati degli errori:

$$\bar{\mathbf{x}}_2 = \begin{pmatrix} 0,8915 \\ 0,0079 \\ 0,0050 \\ 0,0040 \end{pmatrix}$$

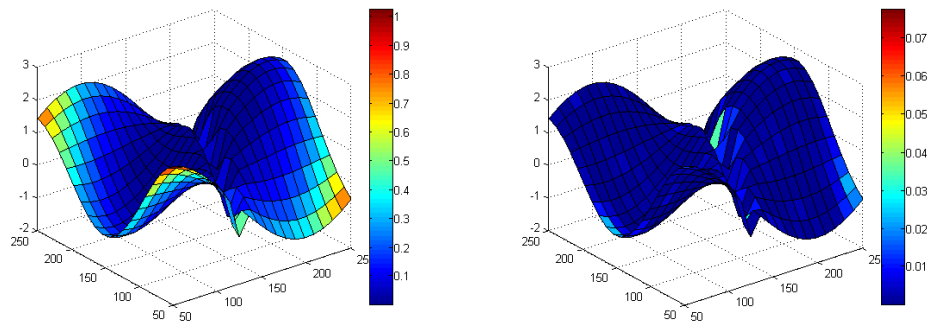
ed i *p-value* dei confronti a coppie:

$$\begin{aligned} LR \text{ vs } Fusion &= 4,9202e-11 \\ LR \text{ vs } Fusion\_ql &= 9,6123e-18 \\ LR \text{ vs } Fusion\_misto &= 2,3553e-19 \\ Fusion \text{ vs } Fusion\_ql &= 0,5975 \\ Fusion \text{ vs } Fusion\_misto &= 0,4846 \\ Fusion\_ql \text{ vs } Fusion\_misto &= 0,4340. \end{aligned}$$

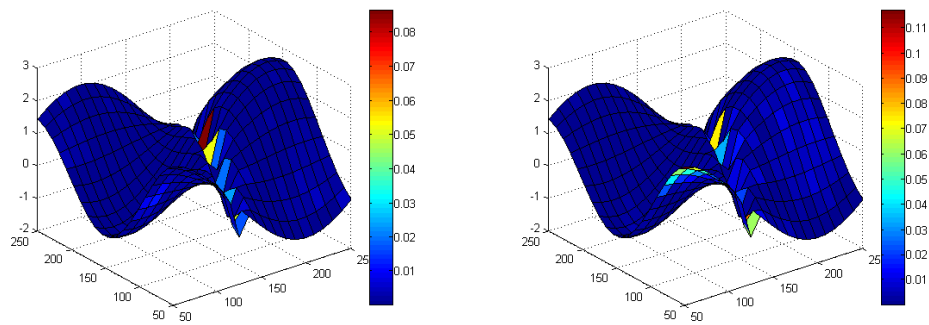
Tabella 5.8: *P-value* confronti a coppie medie quadrati errori

Anche in questo secondo caso della LSM il modello a bassa risoluzione è l'unico che si differenzia dagli altri infatti tutti e tre i modelli *fusion* ottengono risultati, in riferimento al quadrato degli errori, simili; anche se dalla mappa del quadrato degli errori nei singoli punti ottenuta da una replica (figura 5.23), si osserva che il modello con quadrato latino commette degli errori all'interno del solco, mentre mediante il modello proposto gli errori sono più uniformemente distribuiti lungo l'intera superficie. Gli errori del modello a bassa risoluzione sono, invece, per la maggior parte dovuti all'errore non casuale introdotto essendo localizzati negli angoli del dominio.





(a) *Quadrato errori modello bassa risoluzione*      (b) *Quadrato errori modello fusion con campionamento proposto*



(c) *Quadrato errori modello fusion con quadrato latino*      (d) *Quadrato errori modello fusion con campionamento misto*

Figura 5.23: Superficie con quadrato errori

## 5.4 Conclusioni

Analizzando i risultati ottenuti dalle simulazioni si può concludere che il modello con campionamento proposto è migliore di quello con un quadrato latino quando la variabilità dei punti meno densi, di secondo stadio, è minore della variabilità di quelli più densi, sia quando non è presente una errore non casuale, analisi 1, sia quando questo è presente con valori bassi, analisi 2. Nel caso della metrologia di larga scala il modello con campionamento proposto ha ottenuto risultati comparabili a quelli con un disegno standard. Gli errori sono però distribuiti in modo uniforme lungo tutta la superficie, mentre nei modelli con quadrato latino i punti all'interno del solco hanno ottenuto errori, in valore assoluto, maggiori rispetto ai modelli *fusion* con un campionamento proposto.



# Conclusioni

In questo lavoro di tesi si è utilizzato il modello con processo gaussiano [34] per integrare le informazioni provenienti da sensori con diversa accuratezza in modo strutturato. Attraverso una simulazione deterministica sono stati generati i dati a bassa accuratezza mentre quelli ad elevata accuratezza provengono da una simulazione più precisa oppure da un esperimento fisico. Il collegamento tra i dati con diverse accuratezze è eseguito per mezzo di un parametro di posizionamento rappresentato da un processo gaussiano e da un parametro di scala.

Si è utilizzato un modello con kriging universale per caratterizzare il modello a bassa risoluzione. E' stato però necessario introdurre un termine di errore casuale per rappresentare l'incertezza dei dati analizzati. Il modello presente in letteratura è quindi stato ampliato per considerare l'incertezza proveniente dalle stime di predizione del modello a bassa accuratezza. È possibile utilizzare questo modello di collegamento anche nel caso in cui i punti ad elevata accuratezza non appartengono al disegno utilizzato in fase di simulazione.

I modelli sono stati implementati in **Matlab** per poter valutare le prestazioni di dati provenienti da misure di forma di oggetti fisici. L'utilizzo dei processi gaussiani nel campo della metrologia è stato introdotto per la prima volta in [47] e ha permesso di predire punti appartenenti a superfici di forma complessa.

Successivamente è stato proposto un metodo di campionamento dei punti ad elevata accuratezza basato sulla curvatura gaussiana e sull'incertezza di predizione derivante dal modello a bassa accuratezza. Tale metodo è stato poi implementato in **Matlab** per validarlo. Si sono poi confrontate le prestazioni di un modello ad alta risoluzione basato sul campionamento proposto con il medesimo modello basato però sul campionamento comunemente utilizzato nel *computer experiment* e in [47].

## Sviluppi futuri

Un primo sviluppo futuro è l'utilizzo di una funzione di correlazione diversa dal kernel gaussiano utilizzato in questo lavoro. Sebbene la funzione utilizzata sia ragionevole, ne esistono altre che possono portare a risultati migliori. Un

problema da tenere in considerazione è il possibile mal condizionamento della matrice di correlazione che porta ad una stima dei parametri non attendibile poichè i metodi numerici utilizzati per l'inversione della matrice danno risultati non attendibili.

In questo lavoro di tesi non è stata eseguita un'analisi di sensitività sulle stime fatte sui parametri della funzione di correlazione gaussiana, in futuro si ritiene necessario che venga eseguita per indagare sugli effetti che possono portare piccole perturbazioni di questi valori.

Un ulteriore sviluppo riguarda i dataset che provengono da acquisizioni di oggetti fisici che sono spesso formati da una quantità di punti molto elevata. In questo lavoro i punti sono stati decimati mediante una metrica quadratica. Si sarebbe invece potuto mantenere l'intero dataset per poi dividerlo in diverse parti e applicare un modello con processo gaussiano su ciascuna parte. Un metodo per partizionare il set di dati è la regressione ad albero [6], ma dato che le stime dei parametri dei vari modelli sono diverse, non c'è garanzia di continuità nel valore assunto dall'incertezza tra i bordi delle classi comunicanti.

Nel presente lavoro non è stato utilizzato un approccio bayesiano a causa dei tempi lunghi per analizzare i dataset di grande dimensione, come nel caso dei punti a bassa risoluzione. Un processo gaussiano con approccio bayesiano invece potrebbe essere utilizzato per caratterizzare il modello di collegamento tra i punti a bassa ed alta risoluzione, in cui le distribuzioni a priori dei punti ad alta risoluzione sono quelle che derivano dal modello a bassa risoluzione.

# Bibliografia

- [1] URL: [http://www.inspectionengineering.com/Zeiss\\_CMMs.htm](http://www.inspectionengineering.com/Zeiss_CMMs.htm).
- [2] URL: [http://www.primross.com.mx/metrologia/o\\_inspect/o\\_inspect.htm](http://www.primross.com.mx/metrologia/o_inspect/o_inspect.htm).
- [3] URL: <http://www.itpgroup.co.uk/products/probesystems.html>.
- [4] Banerjee S. Carlin B.P. Gelfand A.E. *Hierarchical Model and Analysis for Spatial Data*. A cura di Chapman & Hall. 2003.
- [5] Chan V. Bradley C. “A Complementary Sensor Approach to Reverse Engineering”. In: *Journal of manufacturing science and engineering* 123 (2001), pp. 74–82.
- [6] Stone C. J. Breiman L. Friedman J. H. Olshen R. A. *Classification and Regression Trees*. A cura di Wadsworth. 1984.
- [7] Wagner C. “Informationstheoretische Grenzen optischer 3D-Sensoren”. Tesi di dott. University Erlangen-Nürnberg., 2003.
- [8] Carocci M. Savio E. Sansoni G. De Chiffre L. Carbone V. “Combination of a Vision System and a Coordinate Measuring Machine for the Reverse Engineering of Freeform Surfaces”. In: *International Journal of Advanced Manufacturing Technology* 17 (2001), pp. 263–271.
- [9] Moroni G. Petrò S. Cavallaro M. “Performance evaluation of non contact measuring systems considering bias”. In: *Proceedings of the 4th International Conference on Advanced Research on Rapid Prototyping, Leiria Portugal*. 2009.
- [10] Lin G.C. Chen L.C. “An Integrated Reverse Engineering Approach to Reconstructing Free-form Surfaces”. In: *Computer Integrated Manufacturing Systems* 10/1 (1997), pp. 49–60.
- [11] Neuman H.J. Christoph R. “Coordinate Measuring Machines and Systems”. In: a cura di Paulo H . Pereira. CRC Press, 2007. Cap. 7, pp. 125–151.
- [12] Del Castillo E. Colosimo B. *Bayesian Process Monitoring, Control and Optimization*. CRC/Taylor & Francis, 2006.
- [13] N.A.C. Cressie. *Statistics for Spatial Data*. A cura di John Wiley e Sons. 1993.

- [14] L. Dolci. “Bayesian Hierarchical Gaussian Process Model: an Application to Multi-Resolution Metrology Data”. Tesi di laurea mag. Politecnico di Milano, 2011.
- [15] Rolls C. El Maraghy W. “Design by Quality Product Digitization”. In: *Annals of the CIRP* 50/1 (2001), pp. 93–96.
- [16] Edmundson K. Peggs G. Parker D Estler W. “Large-Scale Metrology - An Update”. In: *CIRP Annals - Manufacturing Technology* 51/2 (2002), pp. 589–609.
- [17] Galetto M. Maisano D. Mastrogiacomo L. Franceschini F. “A Review of Localization Algorithms for Distributed Wireless Sensor Networks in Manufacturing”. In: *International Journal of Computer Integrated Manufacturing, Special Issue: Wireless Manufacturing* 22/7 (2009), pp. 698–716.
- [18] Maisano D. Mastrogiacomo L. Franceschini F. “Mobile Spatial coordinate Measuring System (MScMS). Introduction to the system”. In: *International Journal of Production Research* 47/14 (2009).
- [19] Mastrogiacomo L. Prali B. Galetto M. “An innovative indoor coordinate measuring system for large-scale metrology based on a distributed IR sensor network”. In: *Proceedings of the ASME 2009 International Manufacturing Science and Engineering Conference MSEC2009*. A cura di ASME. 2009.
- [20] Mastrogiacomo L. Pralio B. Galetto M. “MScMS-II: an innovative IR-based indoor coordinate measuring system for large-scale metrology applications”. In: *The International Journal of Advanced Manufacturing Technology* 51 (2010), pp. 291–302.
- [21] Mastrogiacomo L. Pralio B. Galetto M. “The Mobile Spatial coordinate Measuring System II (MScMS-II): system description and preliminary assessment of the measurement uncertainty”. In: *The International Journal of Metrology and Quality Engineering* 1/2 (2010), pp. 111–119.
- [22] Heckbert P.S. Garland M. “Surface simplification using quadratic error metrics”. In: *Proceedings of the 24th annual conference on Computer graphics and interactive techniques*. 1997.
- [23] Kennedy M. Cavendish J.C. Cafeo J.A. Ryne R.D. Higdon D. “Combining Field Data and Computer Simulations for Calibration and Prediction”. In: *SIAM Journal of Scientific Computing* 26/2 (2004), pp. 448–466.
- [24] Mileham A.R. Owen G.W. Jamshidi J. “High Accuracy Laser Scanned View Registration Method for Reverse Engineering Using a CMM Generated CAD model”. In: *IDETC/CIE* (2006), pp. 241–248.
- [25] Owen G.W. Mileham A.R. Jamshidi J. “A New Data Fusion Method for Scanned Models”. In: *Transactions of the ASME* 6 (2006), pp. 340–348.

- [26] Drvar N. Jecić S. “The Assessment of Structured Light and Laser Scanning Methods in 3D Shape Measurement”. In: *4th International Congress of Croatian Society of Mechanics*. 2003.
- [27] Honegger D. Ritter M. Keferstein C.P. “Marktanalyse optischer 1D-, 2D, 3D-Sensoren und optischer KMG”. In: *QZ* 41 (2006), pp. 8–40.
- [28] OHagan A. Kennedy M. C. “Bayesian Calibration of Computer Models”. In: *Journal of the Royal Statistics Society* (2001), pp. 425–464.
- [29] Glombitza M. “Steigerung der Autonomie fertigungsintegrierter Koordinatenmesssysteme durch flexible Bildverarbeitung”. Tesi di dott. RWTH Aachen, 2004.
- [30] C. A. Micchelli. “Interpolation of Scattered Data: Distance Matrices and Conditionally Positive Definite Functions”. In: *Constructive Approximation* 2 (1986), pp. 11–22.
- [31] Suharitdamrong V. Alrashdan A. Motavalli S. “Design Model Generation for Reverse Engineering Using Multi-Sensors”. In: *IIE Transactions* 30/4 (1998), pp. 357–366.
- [32] Bartscher M. Hilpert U. Neuschaefer-Rube U. *Application of Multisensory Measurements and Sensor Data Fusion in Coordinate Metrology. Workshop & Symposium Metrologys Impact on Business*. 2008.
- [33] A. O’Hagan. “Bayes-Hemite quadrature”. In: *Journal of Statistical Planning and Inference* 29 (1991), pp. 145–260.
- [34] Seepersad C.C. Joseph V.R. Allen J.K. Wu C.F.J. Qian Z. “Building Surrogate Models Based on Detailed and Approximate Simulations”. In: *ASME Journal of Mechanical Design* 128 (2006), pp. 668–677.
- [35] Wu C.F.J. Qian Z. “Bayesian Hierarchical Modeling for Integrating Low-Accuracy and High-Accuracy Experiments”. In: *Technometrics* 50/2 (2008), pp. 192–204.
- [36] Alyson G. Hamada W.M. Martz H.F. Kenneth J. R. Reese C. S. “Integrated Analysis of Computer and Physical Experiments”. In: *Technometrics* 46/2 (2004), pp. 153–164.
- [37] H. Schewenke R.G. Whilem R. Hocken. “Task Specific Uncertainty in Coordinate Measurement”. In: *CIRP annals* (2001), pp. 553 –563.
- [38] Welch W.J. Mitchell T.J. Wynn H.P. Sacks J. “Design and Analysis of Computer Experiments”. In: *Satistical Science* 4 (1989), pp. 409–435.
- [39] Williams B.J. Notz W.I. Santner T.J. *The Design and Analysis of Computer Experiments*. A cura di Springer-Verlang. 2003.
- [40] De Chiffre L. Schmitt R. Savio E. “Metrology of Freeform Shaped Parts”. In: *Annals of the CIRP* 56/2 (2007), pp. 810–825.

- [41] Gotway C. A. Schabenberger O. *Statistical Methods for Spatial Data Analysis*. Taylor & Francis Ltd, 2004.
- [42] Huang J. Menq C. Shen T. “Multiple-sensor Planning and Integration for Automatic Coordinate Metrology”. In: *Journal of Computing and Information Science in Engineering* (2001), 167179.
- [43] Huang J. Menq C.H. Shen T.S. “Multiple-Sensor Integration for Rapid and High-Precision Coordinate Metrology”. In: *IEEE/ASME T-MECH* 5/2 (2000), pp. 110–121.
- [44] V.A. Toponogov. *Differential Geometry of Curves and Surfaces: A Concise Guide*. Birkhäuser-Verlag, 2006.
- [45] H. Wackernagel. *Multivariate Geostatistics*. A cura di Springer-Verlang. 2003.
- [46] Jiang X. Sommer K.-D. Neuschaefer-Rube U. Seewig J. Shaw L. Estler T. Weckermann A. “Multisensor data fusion in dimensional metrology”. In: *CIRP annals* 58 (2009), pp. 701–721.
- [47] Hocken R. Schwenke H. Wilhem R.G. “Task Specific Uncertainty in Coordinate Measurement”. In: *CIRP Annals* (2001), pp. 553 –563.
- [48] Yu Din Y. Mallick-B.K. Xia H. “Bayesian hierarchical model for combining misaligned two-resolution metrology data”. In: *IIE Transactions* 43 (2011), pp. 242–258.
- [49] Yu Din Y. Wallik-J. Xia H. “Gaussian Process Model for Form Error Assessment Using Coordinate Measurement”. In: *submitted to IEEE Transactions* (2007).
- [50] Jianguo W. Qiumei Z. Zexiao X. “Complete 3D measurement in reverse engineering using a multi-probe system”. In: *International Journal of Machine Tools & Manufacture* 45 (2005), 14741486.