Alma Mater Studiorum · Università di Bologna

Dipartimento di Informatica — Scienza e Ingegneria
Corso di Triennale in Ingegneria e Scienze Informatiche

Interpolazione di dati gravimetrici per applicazioni Web-based di visualizzazione

Relatore:
Chiar.mo Prof.
Moreno Marzolla

Presentata da:
Edoardo Scorza

Correlatori:
Dott.sa
Maria Rosaria Tondi
Ing.
Giampaolo Zerbinato

Sessione 03 2026
Anno Accademico 2024/2025

Sommario

L'obiettivo proposto in questa tesi è la realizzazione di un applicativo per l'interpolazione di dati geofisici, in particolare di misurazioni gravimetriche. Prima di esplorare le tecniche di interpolazione proposte, viene introdotto l'ambito dell'analisi del sottosuolo e le sue applicazioni; affrontando le sue origini in Italia e come viene analizzato il sottosuolo; in particolare, per misurare le anomalie di gravità. Segue una sezione che introduce la storia e la teoria dietro la gravità e del campo gravitazionale; al fine di collegare il fatto che variazioni nella distribuzione della massa locale, corrispondono ad anomalie di gravità locali. Poi vengono introdotte le tecniche di interpolazione, il perchè sono necessarie e la teoria che le rende rilevanti per il problema posto. Infine, è illustrata l'architettura scelta per l'applicativo e come è gestita la comunicazione e scambio di dati tra client e server.

Indice

Capitolo 1

Introduzione

Lo studio del sottosuolo in Italia risale fino ai tempi di Leonardo da Vinci [1] ma è solo a partire dall'Ottocento che sono partite indagini e cartografie geologiche su larga scala. Prima dell'unità dell'Italia, le mappature del territorio erano molto limitate: principalmente venivano effettuate localmente per ragioni prettamente economiche, principalmente per l'industria mineraria. Oggi la ricerca del sottosuolo è avanzata parecchio è trova numerose applicazioni, sia economiche che culturali.

Nell' archeologia [2], lo studio del sottosuolo consente di individuare siti sepolti, impiegando tecniche avanzate dette metodi geofisici che permettono di analizzare il sottosuolo senza interagire direttamente. Nella vulcanologia [3], conoscere l'interno di un vulcano inattivo può fornire preziose informazioni su di esso e su potenziali cambiamenti, utile sopratutto per prevedere eruzioni. Un uso legato invece al mercato è la ricerca e studio di giacimenti petroliferi e di gas naturali [4]. l'uso dei metodi geofisici consente di risparmiare costose e invasive perforazioni dagli esiti incerti.

1.1 Metodi Geofisici non invasivi

L'evoluzione della ricerca del sottosuolo ha seguito lo sviluppo tecnologico, infatti, come gia menzionato, l'analisi del sottosuolo è effettuata principalmente usando tecniche non invasive; ognuna di esse permette di estrarre dati differenti; per questo spesso si usa una combinazione di esse. Tra le tecniche più diffuse troviamo:

- Il radar a penetrazione del suolo (GPR) usa onde elettromagnetiche a bassa frequenza (10- 3000 MHz) per mappare strutture sotterranee: Un'antenna radar emette impulsi elettromagnetici che, in funzione della frequenza e delle proprietà del mezzo, vengono riflessi; un ricevitore acquisisce l'eco e, misurando il tempo di andata e ritorno, stima la distanza degli oggetti.

- la Tomografia della resistenza elettrica (ERT), sfrutta la misurazione della resistenza del sottosuolo, attraverso degli elettrodi impiantati nel terreno, viene misurata la resistenza fra due punti; con sufficienti misurazioni è possibile creare un modello di resistenze, dal quale è possibile fare inferenze su quale materiale ha tale resistenza.

- La gravimetria, la cui interpolazione dei dati è oggetto di questa tesi, sfrutta le anomalie di gravità per comprendere la distribuzione della massa nel sottosuolo. Le principali applicazioni sono nell'individuazioni di grandi strutture, cavità, e giacimenti petroliferi.

1.2 I fondamenti della gravità

Lo studio della gravità è attribuito per prima a Isaac Newton (1642–1727), universalmente ricordato per i principi della dinamica e la legge di gravitazione universale. Essa afferma che ogni coppia di masse (m1) e (m2) si attraggono con una forza proporzionale al prodotto delle loro masse e inversamente proporzionale al quadrato della loro distanza (r) che le separa.

$$F = G\frac{m_1 m_2}{r^2}$$

Dove G è la costante di gravitazione universale
Questa forza descritta da Newton è la gravità, e la presenza della massa nella formula indica la sua influenza. Per esempio, con questa formula è possibile ottenere la forza che la Terra subisce da una massa in sua prossimità:

- persona: (m = 70kg)
- Terra: (M = 5.97 $\times 10^{24}$kg)
- raggio terrestre: ($r = 6.37 \times 10^6$m)

Risolvendo la formula con i dati presentati, risulta che:

$$F \approx 686\text{N}$$

Prendendo la formula del secondo principio della dinamica ($F = ma$) è combinandola con la legge di gravitazione universale, si ottiene che:

$$ma = G\frac{m_1 m_2}{r^2}$$

risolvendo per a, usando i dati precedenti, si trova la accelerazione di gravità [5]:

$$a = 9.8$$

1.3 La forma della Terra

Sapendo che la gravità è legata alla massa della Terra, capire come è distribuita può fornire un nesso più forte; ma prima c'è un altra domanda altrettanto importante: qual è la forma della Terra? Già in epoca antica filosofi e scienziati si interrogarono a riguardo. Pitagora e la sua scuola furono i primi a ipotizzare la sfericità della Terra. Una prima dimostrazione segue nel III secolo a.C. Eratostene di Cirene stimò la circonferenza terrestre utilizzando:

- La distanza fra Alessandria e Syene, due città posizionate sullo stesso meridiano
- L'altezza del sole raggiunta durante mezzogiorno nelle due città: a Syene i raggi sono quasi perpendicolari al suolo, mentre ad Alessandria ce un angolo di circa 83 gradi, sette in meno rispetto a Syene

Sfruttando la distanza tra le due città e l'angolazione formata dai raggi solari, egli stimo la circonferenza della Terra, pari a circa 40.500 km, molto vicino al valore reale di 40.009 km.

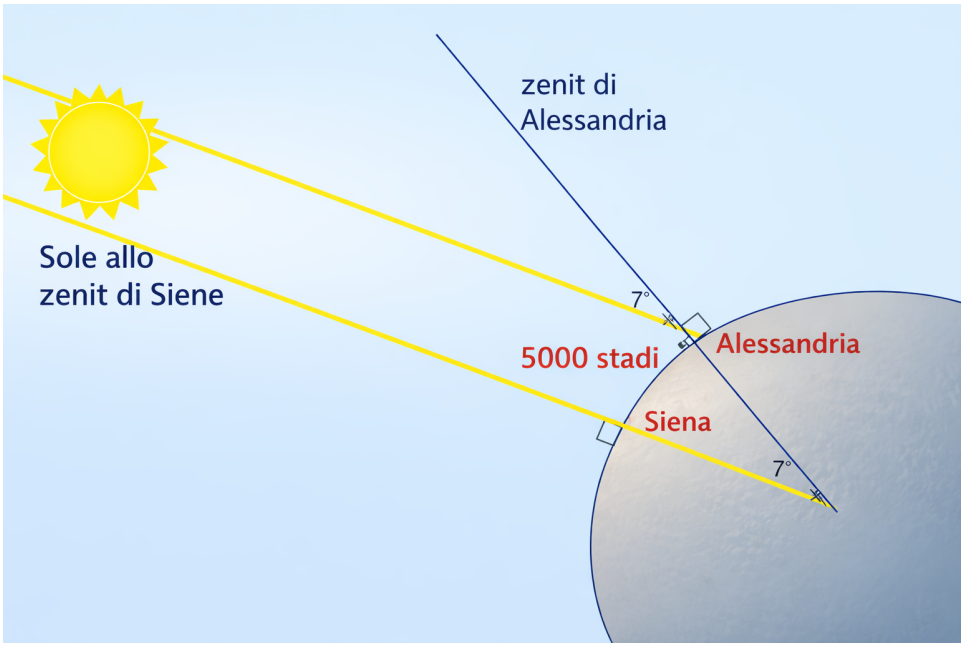

Figura 1.1: Esperimento di Eratostene

Con il progredire degli studi sulla dimensione della Terra, divenne evidente che la Terra non può essere descritta come una sfera perfetta. A partire dal XVII e XVIII secolo, misurazioni sempre più precise indicarono che la Terra è meglio approssimata da un ellissoide, leggermente schiacciato ai poli, conseguenza diretta della rotazione terrestre [6].

1.4 Anomalie di gravità

Attraverso la forma del pianeta, è possibile ottenere buone stime del campo di gravità teorico. Tuttavia, la gravità misurabile sulla superficie non è uniforme.
Una conferma dell'esistenza di anomalie di gravità segue da un curioso episodio: Jean Richer (1671–1673) osservò come un orologio a pendolo, perfettamente calibrato a Parigi, perdesse precisione una volta trasportato a Cayenna, vicino all'equatore. La variazione nel periodo di oscillazione del pendolo indica una differenza nell'accelerazione di gravità, suggerendo che il campo gravitazionale non è uniforme sulla superficie terrestre [6].

1.5 Le leggi della geografia

Una prima ipotesi sul perché il campo gravitazionale presenti variazioni locali si trova nella prima legge della geografia, formulata da Waldo Tobler e successivamente discussa da Miller. Essa afferma che tutti i fenomeni spaziali sono tra loro correlati, ma che tale correlazione tende a diminuire con l'aumentare della distanza.

Applicata alle misure gravimetriche, questa idea suggerisce che le variazioni di gravità non siano valori casuali e scorrelati, ma presentino una correlazione spaziale: i valori misurati in punti vicini tendono a essere simili. Un'anomalia gravimetrica, quindi, non può essere interpretata come un evento isolato, ma come parte di qualcosa di più grande che, complessivamente, porta al valore della gravità ottenuto precedentemente.

Come sottolinea Miller [7], la legge di Tobler non costituisce una legge fisica in senso stretto, ma un principio empirico che descrive il modo in cui i fenomeni naturali si organizzano nello spazio.

1.6 Teoria del Potenziale

Una giustificazione scientifica di questo comportamento, supportata da formulazioni matematiche e dimostrazioni formali, è fornita dalla teoria del potenziale. Essa afferma che la gravità è un campo, una quantità fisica che possiede un valore in ogni punto dello spazio e del tempo. Essa rappresenta la distribuzione di una quantità fisica nello spazio, in questo caso della forza di gravità.

Ogni porzione di materia contribuisce al campo totale, e la misurazione della gravità ottenuta mediante la formula di Newton rappresenta la somma di tutti questi contributi. La gravità osservata in superficie è quindi l'espressione integrata della distribuzione di massa presente nel sottosuolo e nelle regioni circostanti.

In condizioni ideali, quando la distribuzione delle masse è omogenea, il campo di gravità varia in modo regolare. Al contrario, variazioni di densità nel sottosuolo modificano localmente il potenziale gravitazionale, dando origine ad anomalie gravimetriche. Tali anomalie non sono fenomeni isolati, ma contributi locali che, combinandosi tra loro, costruiscono il campo gravitazionale complessivo.

Dal punto di vista teorico, nelle regioni dello spazio prive di sorgenti di massa, il potenziale gravitazionale soddisfa equazioni che sono continue e derivabili. Ne consegue che il campo di gravità non presenta discontinuità improvvise: mettendo vicine molte misurazioni di gravità, osserveremo delle curve.

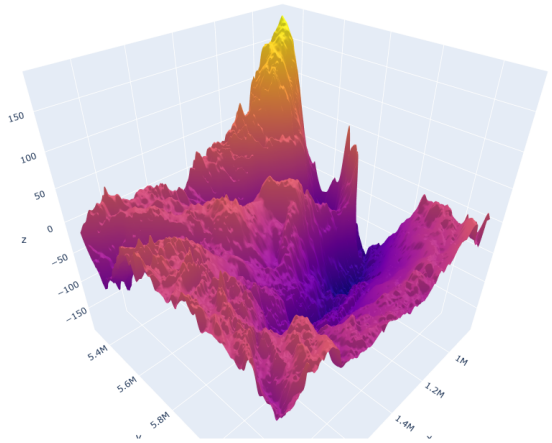

Figura 1.2: Disegno tridimensionale dove sugli assi x-y abbiamo le coordinate, mentre l'asse z è individuato dalle misurazioni di gravità con la correzione di Bougher, si osserva come la distribuzione forma una superficie.

1.7 I gravimetri

La misura della gravità viene effettuata mediante strumenti chiamati gravimetri, progettati per rilevare variazioni molto piccole dell'accelerazione di gravità, dell'ordine di microgal. Esistono due principali categorie di gravimetri:

- Gravimetri assoluti, che misurano direttamente il valore dell'accelerazione di gravità osservando il moto di caduta libera di una massa in un sistema di riferimento controllato.

- Gravimetri relativi, che misurano variazioni di gravità rispetto a un valore di riferimento, generalmente utilizzati nelle campagne di acquisizione sul campo.

Figura 1.3: Esempio di gravimetro Dj krølltopp, CC BY-SA 4.0
https://creativecommons.org/licenses/by-sa/4.0, via Wikimedia Commons

Un gravimetro può essere pensato come una bilancia molto precisa: una massa è sospesa a una molla calibrata e le variazioni del campo gravitazionale provocano piccoli spostamenti della massa, che vengono misurati. Poiché questi strumenti sono sensibili agli effetti ambientali, come temperatura e vibrazioni, vengono utilizzate molle a lunghezza zero, fatte di quarzo o metalli, e sono situati in un contenitore sigillato sottovuoto [8, 9].

1.8 Acquisizione dei dati e prime elaborazioni

Le campagne gravimetriche consistono nella misurazione della gravità in una serie di stazioni distribuite sul territorio di interesse. Per ogni stazione vengono registrati, oltre al valore di gravità, le coordinate geografiche e l'altitudine.
Il valore di gravità misurato in ogni stazione non può essere usato direttamente, poiché include contributi esterni agli effetti del sottosuolo. Per rendere i dati utilizzabili, essi vengono sottoposti a una serie di correzioni note come riduzioni gravimetriche, che permettono di isolare le variazioni di gravità dovute alla distribuzione delle masse sotterranee [9].

Le principali correzioni applicate sono:

- **Correzione di latitudine:** tiene conto della variazione della gravità con la posizione geografica, legata alla forma ellissoidale della Terra e alla sua rotazione. In pratica, dal valore osservato viene sottratto il valore di gravità normale, calcolato secondo formule internazionali.

- **Correzione di aria libera (Free-air):** considera la variazione della gravità con l'altitudine della stazione rispetto al livello del mare. All'aumentare della quota,

il valore della gravità diminuisce e la correzione viene quindi aggiunta al dato misurato.

- **Correzione di Bouguer:** Una possibile necessità quando viene effettuata una rilevazione è ottenere i dati relativi a una precisa profondità. Il valore registrato in superficie è influenzato anche dalle masse immediatamente sotto la stazione di misura.

Ad esempio, se siamo su una collina, parte della gravità misurata dipende dal terreno che sostiene la stazione; Il contributo gravitazionale di questa massa viene calcolato e sottratto dalla misura; ciò che rimane è un'approssimazione dell'anomalia a partire dalla profondità stabilita.

Capitolo 2

Interpolazione Spaziale

A seguito della raccolta dei dati e delle prime elaborazioni, finalizzate a ridurre le influenze esterne dalle misurazioni, si ottiene un insieme discreto di punti che rappresenta il fenomeno osservato. Ora l'obiettivo è trovare un modo per migliorare il dataset; al fine di effettuare delle stime su regioni dove non sono presenti misurazioni di gravità.

Un aumento del numero di osservazioni contribuirebbe a migliorare la qualità della rappresentazione del fenomeno. Tuttavia, il numero di misurazioni rimane necessariamente limitato e il processo di acquisizione dei dati può risultare oneroso sia in termini di tempo sia di costi. La risposta a questo problema è fornita dall'interpolazione, una procedura che comprende un insieme di tecniche basate su strumenti matematici e statistici, il cui obiettivo è stimare i valori del fenomeno in punti non direttamente osservati usando i dati disponibili.

Nell'analisi numerica, l'interpolazione [10] è un insieme di tecniche finalizzate a individuare nuovi punti sul piano cartesiano a partire da un insieme finito di dati, sotto l'ipotesi che tali punti appartengano al grafico di una funzione o, più in generale, a una famiglia di funzioni spesso sconosciuta. Dato un insieme di (n) valori distinti (x_k), detti nodi, e per ciascuno di essi un valore associato (y_k), il problema dell'interpolazione consiste nel determinare una funzione $(f\ )$, appartenente a una certa classe prefissata, tale che:

$$f(x_k) = y_k \quad \text{per} \quad k = 1, \ldots, n. \tag{2.1}$$

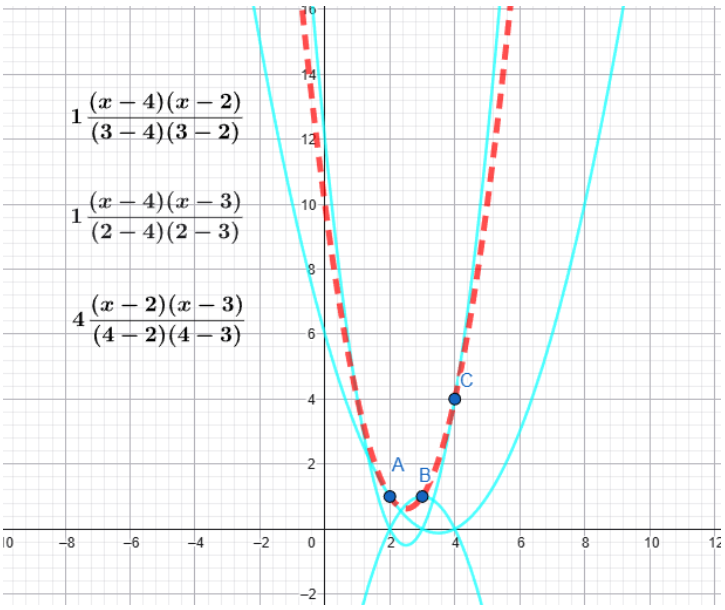

Figura 2.1: Polinomio interpolante di Laplace, per ogni punto c'è una base, la loro somma da origine al polinomio interpolante; realizzato con Geogebra.

Nella Figura 2.1 è rappresentato un insieme di punti che descrive le misurazioni disponibili. La funzione interpolante è la curva che passa esattamente per tali punti e che fornisce una rappresentazione continua del fenomeno osservato, permettendo di associare a ogni valore (x) un corrispondente valore (y), anche in assenza di una misurazione diretta.

L'idea di descrivere un insieme di dati mediante una curva non si limita al solo problema dell'interpolazione, ma si estende a un ambito più ampio noto come curve fitting [11]. In questo contesto, l'obiettivo non è necessariamente quello di far passare la curva esattamente per tutti i punti disponibili, bensì di individuare una funzione che ne rappresenti al meglio l'andamento complessivo, riducendo l'influenza del rumore o delle incertezze sperimentali.

Esempi significativi di curve fitting si ritrovano in numerosi ambiti [12]. In informatica e nei sistemi embedded, funzioni matematiche complesse o computazionalmente onerose vengono approssimate mediante funzioni più semplici, ottenute a partire da un insieme discreto di valori. In ambito statistico e sperimentale, il curve fitting viene utilizzato per modellare relazioni tra grandezze fisiche osservate, come nel caso della regressione lineare o non lineare. Allo stesso modo, nell'intelligenza artificiale e nel machine learning, l'apprendimento di relazioni tra dati, quali immagini, segnali audio o testi, si basa sulla costruzione di modelli matematici che approssimano il legame sottostante tra le osservazioni, generalizzando oltre i dati di partenza [13].

2.1 Sistema di riferimento CRS

In geofisica, i dati analizzati consistono spesso in misurazioni fisiche associate a una specifica posizione nello spazio. In questo contesto, l'interpolazione prende il nome di

interpolazione spaziale [14], poiché i valori da stimare sono punti distribuiti in uno spazio bidimensionale (longitudine e latitudine) o tridimensionale (longitudine, latitudine e altitudine).

Per posizionare correttamente tali punti, è necessario adottare un opportuno sistema di riferimento [15]. Tradizionalmente, la Terra è rappresentata attraverso l'uso di globi; questo approccio preserva le proporzioni della Terra in modo accurato, ma con una scala molto ridotta. Se volessimo trovare una strada per andare da punto A a punto B con un globo, sarebbe praticamente impossibile; per questo sono nate le proiezioni; in altre parole, le proiezioni provano a trasformare la Terra dalla sua forma sferica (tridimensionale) a una forma planare (bidimensionale).

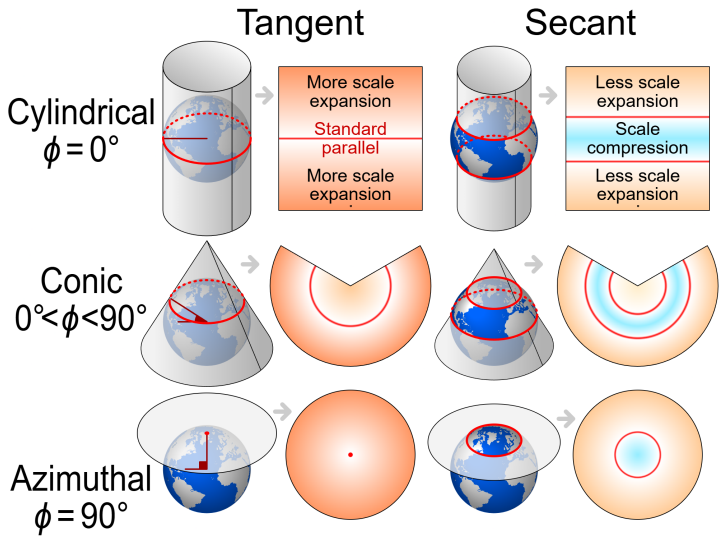

Figura 2.2: Diverse tipologie di superfici usate per fare una proiezione, la scelta dipende da che area del globo si decide di proiettare. By cmglee, US government, Clindberg, Palosirkka - Globe Atlantic.svg, CC BY-SA 4.0, https://commons.wikimedia.org/w/index.php?curid=84845850

Il fatto che esistano più proiezioni e dovuto al fatto che la trasformazione da sistema geografico a cartesiano non è biunivoca: se si prova a distendere la superficie di una sfera su un piano, risulta evidente che ciò non è possibile senza introdurre delle distorsioni; come dimostrato dal teorema Egregium di Gauss. Di conseguenza, ogni proiezione comporta un compromesso: preservare alcune proprietà geometriche a scapito di altre, come distanze, aree o angoli.

Possiamo dividere le tipologie di proiezioni in tre categorie:

- **Proiezioni conformi**. Questa tipologia di proiezione preserva gli angoli e le forme locali, garantendo che le direzioni siano corrette. Sono particolarmente adatte quando l'orientamento e la geometria locale devono essere preservati. La proiezione conforme più conosciuta è la proiezione di Mercatore, ampiamente utilizzata nella navigazione, che, tuttavia, introduce una forte distorsione delle superfici alle alte latitudini.

- **Proiezioni equidistanti**. Le proiezioni equidistanti mantengono corrette le distanze lungo determinate direzioni o rispetto a specifici punti di riferimento. Sono impiegate quando è necessario preservare la misurazione delle distanze. Un esempio è la proiezione azimutale equidistante, nella quale le distanze dal punto centrale della proiezione sono rappresentate in modo fedele.

- **Proiezioni equivalenti**. Questa categoria di proiezioni preserva le aree, assicurando che le superfici rappresentate sulla mappa siano proporzionali a quelle reali sulla Terra. Le forme e gli angoli possono risultare distorti, ma il confronto tra estensioni geografiche rimane corretto. Un esempio è la proiezione di Mollweide, spesso adottata per mappe globali e rappresentazioni tematiche.

L'introduzione dei sistemi di riferimento è necessaria per chiarire quale tipologia di coordinate è più appropriata per l'interpolazione spaziale di una regione. L'uso delle proiezioni risulta, in molti casi, la scelta migliore per varie ragioni [16]:

- Per via delle dimensioni della Terra, le regioni osservate sono approssimabili a superfici.

- Le unità di misura più comunemente usate sono basate sul sistema cartesiano.

- I principali algoritmi di interpolazione prediligono un sistema cartesiano.

La scelta di un algoritmo di interpolazione pone delle sfide [17]; i fenomeni spaziali sono spesso complessi e descritti da dati irregolari, rumorosi o distribuiti in modo non uniforme. Essi devono poter elaborare dataset grandi, essere flessibili ed efficienti. Non esiste l'algoritmo migliore, ma quello che si adatta meglio al problema. Di seguito vengono presentati gli algoritmi scelti, essi sono tra i metodi più utilizzati nei Geographic Information System (GIS).

2.2 Natural Neighbor - Inverse Distance Weighting (IDW)

IDW [18] è un algoritmo di interpolazione che sfrutta direttamente il primo principio di Tobler: il valore di un singolo punto è il risultato dell'influenza dei vicini, che varia in base alla distanza. Dato un insieme di (n) punti con coordinate (x_i) e i rispettivi valori associati $(z(x_i))$, il valore interpolato in un punto (x_0) viene calcolato come:

$$\hat{z}(x_0) = \frac{\sum_{i=1}^{n} w_i(x_0) z(x_i)}{\sum_{i=1}^{n} w_i(x_0)} \tag{2.2}$$

dove $w_i(x_0)$ rappresenta il peso assegnato al punto i-esimo, in funzione della distanza tra x_0 e tutti gli altri punti (x_i) . Nel metodo IDW classico, i pesi vengono calcolati cosi:

$$w_i(x_0) = \frac{1}{d(x_0, x_i)^p}$$

il cui peso è l'inverso della distanza tra i punti, come implica il nome stesso; dove:

- $d(x_0, x_i)$ è la distanza euclidea tra il punto da stimare e il punto campionato.
- $p > 0$ è il parametro di potenza, che controlla l'influenza della distanza e varia da 1 a 4.

Lo scopo del parametro p è regolare il comportamento della funzione interpolante: valori bassi, come uno, producono funzioni lisce, meno reattive alle influenze, mentre per valori alti risultano funzioni più irregolari che tendono ad aderire ai punti.

Per la applicazione dell'algoritmo, è necessario definire:

- **Numero di vicini**: l'interpolazione può essere effettuata utilizzando tutti i punti disponibili oppure solo un numero limitato di punti vicini al punto (x_0).
- **Area di ricerca**: si può definire un raggio massimo oltre il quale i punti non contribuiscono alla stima.
- **Distanza minima**: per evitare problemi numerici quando $d(x_0, x_i)$ tende a 0, viene spesso introdotta una distanza minima o un trattamento speciale per i punti coincidenti.
- **Metriche di distanza alternative**: in alcuni casi si utilizzano distanze anisotrope o ponderate.

I vantaggi sono l'efficienza, la praticità e l'applicazione a n dimensioni senza variazioni; le molte varianti possibili consentono l'adattabilità a molte situazioni. Di contro, il risultato dipende molto dai punti del dataset e spesso non riproduce la forma generale dei dati.

2.3 Spline

I metodi di interpolazione basati su spline [17] approssimano l'andamento dei dati mediante una funzione definita a tratti, composta da polinomi di grado relativamente basso, connessi in modo tale da garantire continuità e regolarità nei punti di connessione, detti nodi. Ci sono diverse tipologie di spline; le più semplici sono le Spline lineari: dato un intervallo, nell'insieme di punti X, a disposizione per l'interpolazione, abbiamo che:

$$P_i(x) = a_i x + b_i, \quad x \in [x_i, x_{i+1}] \tag{2.3}$$

Le spline quadratiche introducono un termine di secondo grado; esso aggiunge curvatura alla funzione interpolante:

$$P_i(x) = a_i x^2 + b_i x + c_i, \quad x \in [x_i, x_{i+1}] \tag{2.4}$$

Le spline cubiche forniscono la migliore approssimazione; quelle quadratiche garantiscono continuità fino alla prima derivata; aggiungendo un altro termine, otteniamo continuità anche nella derivata di secondo ordine:

$$P_i(x) = a_i x^3 + b_i x^2 + c_i x + d_i \quad x \in [x_i, x_{i+1}] \tag{2.5}$$

Per risolvere un problema di interpolazione cubica è necessario:

- Individuare il numero di coefficienti; per ogni equazione ce ne sono 4, e il numero di equazioni è pari a uno meno il numero di punti ($4(n-1)$).
- Trovare le equazioni necessarie per risolvere un sistema lineare allo scopo di ricavare i parametri sconosciuti:
  - imposizione del passaggio per i punti conosciuti ($P_i(x_i) = y_i$), si ottengono $2n$ equazioni.
  - imposizione della derivata prima, solo sui punti interni ($P_i(x_i)' = P_{i+1}(x_{i+1})'$), si ottengono $(n-1)$ equazioni
  - imposizione della derivata seconda, solo sui punti interni ($P_i(x_i)'' = P_{i+1}(x_{i+1})''$), si ottengono $(n-1)$ equazioni.
  - Per ricavare le equazioni mancanti, è necessario fare delle assunzioni sui punti esterni; alcuni approcci sono:
    * La Natural Cubic Spline; essa assume che $P_1(x_1)'' = 0$ e $P_n(x_n)'' = 0$, dove n è il numero di punti.
    * La Clamped Cubic Spline, invece, richiede di fissare manualmente le derivate agli estremi.
- Usare un algoritmo per risolvere il sistema di equazioni, cosi da ricavare le incognite.

I vantaggi di questo approccio sono la qualità e l'uso di polinomi che crea funzioni più conformi; Tra le sue limitazioni si trova la non applicabilità diretta in n dimensioni, e la limitata possibilità di generare punti, in base ai dati a disposizione, come verrà osservato successivamente.

2.4 Kringing

Kriging è un metodo di interpolazione spaziale [14, 19] di tipo geostatistico, ampiamente utilizzato in geofisica e nelle applicazioni GIS. Similmente all'IDW, il Kriging si basa sull'idea che punti spazialmente vicini tendano ad assumere valori simili; in entrambi i metodi, il valore in una posizione non campionata viene stimato come una combinazione lineare pesata dei valori osservati nei punti circostanti.

La differenza consiste nel fatto che, nel Kriging, i pesi non dipendono solo dalla distanza, ma sono determinati dalla correlazione spaziale del fenomeno, ovvero da come i punti sono collegati tra loro. Questo consente al Kriging di fornire non solo una stima del valore interpolato, ma anche una misura dell'incertezza.

Come nell'IDW, la stima nel punto x_0 è espressa come una combinazione lineare dei valori osservati:

$$\hat{z}(x_0) = \sum_{i=1}^{n} \lambda_i z(x_i) \tag{2.6}$$

dove:

- $z(x_i)$ sono i valori misurati.

- λ_i sono i pesi di Kriging.

La correlazione spaziale viene calcolata attraverso il variogramma [20], uno strumento ampiamente usato nella geostatistica. Esso descrive numericamente come la somiglianza tra i valori di una variabile diminuisce all'aumentare della distanza spaziale; in pratica, è l'insieme delle medie delle semivarianze calcolate in funzione della distanza tra tutte le coppie di punti. Il variogramma teorico è definito come:

$$\gamma(h) = \frac{1}{2} E\left[\big(z(x) - z(x+h)\big)^2\right] \tag{2.7}$$

dove h rappresenta il vettore di separazione tra due punti, ed E è il valore atteso o media, calcolata su tutte le varianze di coppie alla stessa distanza (h).

Nella pratica, il variogramma viene stimato a partire dai dati osservati, calcolando, per diverse classi di distanza, la semivarianza:

$$\gamma(h) = \frac{1}{2N(h)} \sum_{(i,j)\in h} \big(z(x_i) - z(x_j)\big)^2 \tag{2.8}$$

dove $N(h)$ è il numero di coppie di punti separate da una distanza approssimativamente pari a h e (i, j) è una coppia di punti che fa parte dell'insieme h, che contiene tutti i punti distanti fra loro h.

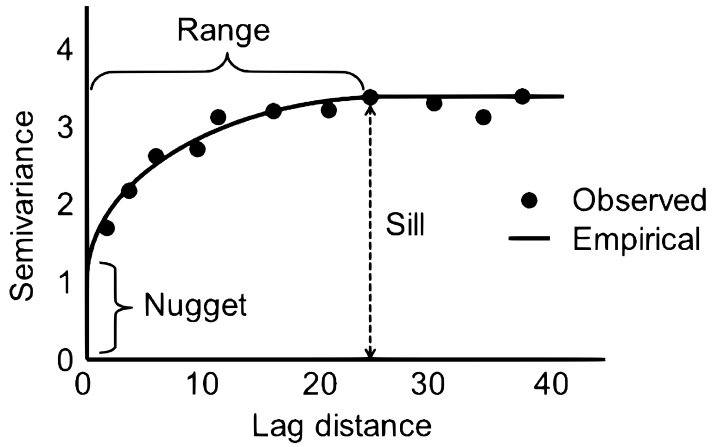

Figura 2.3: Asim Biswas and Bing Cheng Si, CC BY 3.0
https://creativecommons.org/licenses/by/3.0, via Wikimedia Commons

i punti della figura 2.3 rappresentano il variogramma sperimentale; ogni punto è la varianza media (ordinata) stimata rispetto alla distanza (ascissa), la curva è il variogramma teorico, utilizzato per ricavare i pesi e stimare i punti con Kringing. Si nota come la curva ha un andamento logaritmico; ciò significa che prendendo due punti, uno più distante dell'altro, la influenza del più vicino rispetto a quella del più distante tende ad essere simile, per la distanza che tende a infinito. I parametri sono:

- nugget: rappresenta la variabilità a distanza nulla; idealmente dovrebbe essere zero, ma non è quasi mai il caso. Questo è un indicatore della qualità dei dati; dati rumorosi producono un nugget elevato.
- sill: valore a cui il variogramma tende per grandi distanze.
- range: distanza oltre la quale i punti possono essere considerati non correlati; è il valore che inserito nella funzione del variogramma restituisce il sill.

Il modello di variogramma fornisce le informazioni necessarie per il calcolo dei pesi λ_i, e vengono calcolati risolvendo questo sistema:

$$A\lambda = b \tag{2.9}$$

dove:

- A è la matrice che contiene il variogramma calcolato come $\gamma(x_i, x_j)$.
- λ è la matrice che contiene i pesi.
- b contiene il variogramma calcolato con il valore da predire $\gamma(x_n ew, x_i)$.

I vantaggi di questo approccio sono un ottima generazione di punti e la stima della qualità dei punti, grazie all'uso del variogramma.

2.5 Valutazione dell'interpolazione mediante Cross-validation

Il passo successivo all'interpolazione consiste nella valutazione della qualità dei risultati [21]. Tra le tecniche più utilizzate rientrano le procedure di cross-validation, che consentono di stimare la capacità predittiva del modello mediante il confronto tra valori osservati e valori stimati. In particolare è stata adottata la **K-fold cross-validation** [22], una variante della cross-validation classica. Il dataset viene suddiviso in K sottoinsiemi (fold); a ogni iterazione, uno dei fold viene escluso dal training e utilizzato come insieme di test, mentre i restanti $K - 1$ fold sono impiegati per costruire il modello di interpolazione. Il processo viene ripetuto K volte, in modo da testare in modo uniforme il dataset.

Il confronto tra valori osservati y_i e valori stimati $\hat{y}_i$ è effettuato mediante le seguenti metriche:

- **Mean Absolute Error (MAE):**

$$\frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{2.10}$$

  Misura l'errore medio assoluto tra valori osservati e stimati. È espresso nella stessa unità fisica delle anomalie di gravità e fornisce una misura diretta dell'accuratezza media dell'interpolazione.

- **Root Mean Squared Error (RMSE):**

$$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{2.11}$$

Penalizza maggiormente gli errori di grande entità grazie alla quadratura dei residui. Risulta quindi particolarmente sensibile alla presenza di outlier.

- **Mean Error (ME):**

$$\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i) \tag{2.12}$$

Rappresenta la media dei residui e fornisce un'indicazione del bias del modello. Valori prossimi a zero indicano assenza di errore sistematico, mentre valori positivi o negativi evidenziano rispettivamente una sottostima o una sovrastima sistematica delle anomalie.

- **Normalized Mean Squared Error (NMSE):**

$$NMSE = \frac{n\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{(\sum_{i=1}^{n} y_i)\,(\sum_{i=1}^{n} \hat{y}_i)} \tag{2.13}$$

Fornisce una misura adimensionale dell'errore quadratico medio, normalizzata rispetto alla scala globale del fenomeno. A differenza dell'RMSE, la NMSE consente il confronto tra modelli applicati a dataset differenti o a variabili con diversa ampiezza. Valori prossimi a zero indicano un'elevata qualità dell'interpolazione, mentre valori elevati segnalano prestazioni insoddisfacenti.

Capitolo 3

Implementazione e confronto dei modelli

Per l'implementazione è stato scelto il linguaggio Python, grazie alla sua ampia disponibilità di librerie dedicate al calcolo scientifico e all'analisi dei dati. In particolare, sono state utilizzate le seguenti librerie:

- **Geopandas**, per la gestione del database di coordinate e delle operazioni geospaziali;
- **SciPy**, per le funzioni matematiche avanzate e l'interpolatore spline `CloughTocher2D`;
- **PyKrige**, per l'interpolazione dei dati mediante kriging.

3.1 Lettura ed elaborazione

1. Lettura del database contenente i punti.
2. Creazione di una struttura dati geografica con le seguenti colonne:
   - Latitudine
   - Longitudine
   - Anomalia di gravità
3. Costruzione della geometria dei punti utilizzando:
   - la longitudine come coordinata X.
   - la latitudine come coordinata Y.
4. Assegnazione del sistema di riferimento geografico.
5. Selezione dei punti appartenenti alla regione rettangolare definita con massimo e minimo di longitudine e latitudine
6. Trasformazione del sistema di riferimento da coordinate geografiche a coordinate proiettate metriche.

I dati vengono salvati in una struttura chiamata **dataframe**, una matrice indicizzata per righe e colonne su cui è possibile fare operazioni statistiche. In particolare, quella utilizzata si chiama geodataframe, che aggiunge funzioni specifiche ai dati di tipo spaziale, come cambio di coordinate, il plot nello spazio, ecc. Ad esempio possiamo osservare come i dati sono distribuiti nello spazio:

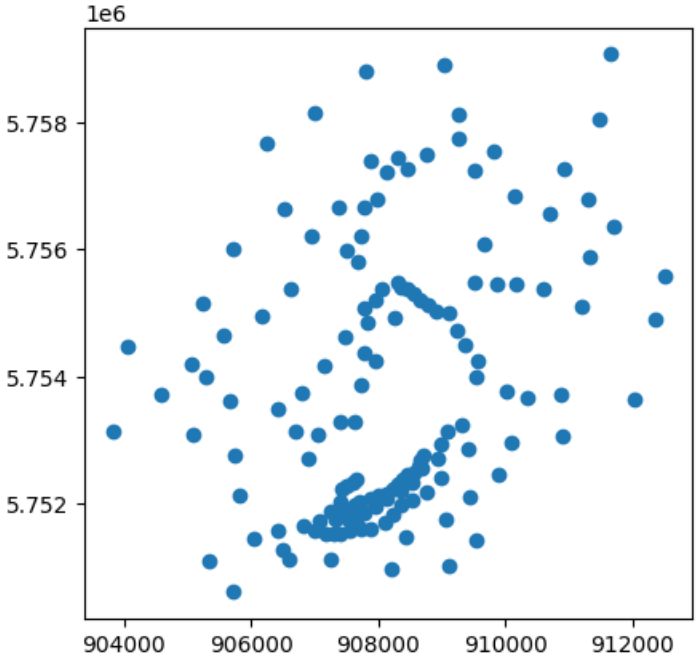

Figura 3.1: immagine generata tramite l'istruzione "plot()", della libreria Geopandas. Rappresenta la distribuzione dei dati nello spazio, gli assi rappresentano le coordinate cartesiane

Osserviamo che nella figura 3.1 ci sono diverse aree vuote e concentrazioni di punti; questo fornisce una buona idea di cosa ci si può aspettare dall'interpolazione; prendiamo ad esempio l'area definita tra min-906000 e 5.756-max; la assenza di punti in una cosi grande regione ci suggerisce che la qualità dell'interpolazione, in quella regione, sarà bassa. Un'ulteriore informazione utile sono le metriche statistiche descrittive fornite dalla libreria Pandas tramite il metodo "describe()", applicato al GeoDataFrame:

	Free_air_80	x	y
count	151.000000	151.000000	1.510000e+02
mean	124.735975	908164.384423	5.753886e+06
std	34.623608	1628.003725	2.101099e+03
min	64.813284	903814.077700	5.750608e+06
25%	92.941167	907348.471532	5.752078e+06
50%	114.874062	908044.218350	5.753276e+06
75%	153.980830	909073.923640	5.755385e+06
max	188.765916	912508.129931	5.759061e+06

Figura 3.2: Le tre colonne sono le coordinate cartesiane e la misurazione di gravità con la correzione Free-air; generato con il metodo "describe()" di Pandas.

Quello che si osserva è un altro punto di vista dell'immagine precedente; le statistiche delle coordinate indicano come i punti sono distribuiti nello spazio, i valori che assumono e la quantità di dati presenti.

3.2 Interpolazione

Per l'interpolazione è stato definito un framework il cui compito è predisporre il dataset e la griglia dei punti da interpolare. Il framework è progettato per accettare moduli che incapsulano le diverse tecniche di interpolazione, ciascuno dei quali implementa la seguente interfaccia:

```
metodoInterpolante(puntiDelDataset, puntiDaGenerare, nomeValoreDaInterpolare)
    -> puntiGenerati
```

L'introduzione di moduli di interpolazione non risponde soltanto all'esigenza di rendere la procedura modulare ed estendibile, ma anche alla necessità di riutilizzare le stesse implementazioni nella fase di valutazione. Curiosamente durante la fase di analisi è stato trovato un altro uso:

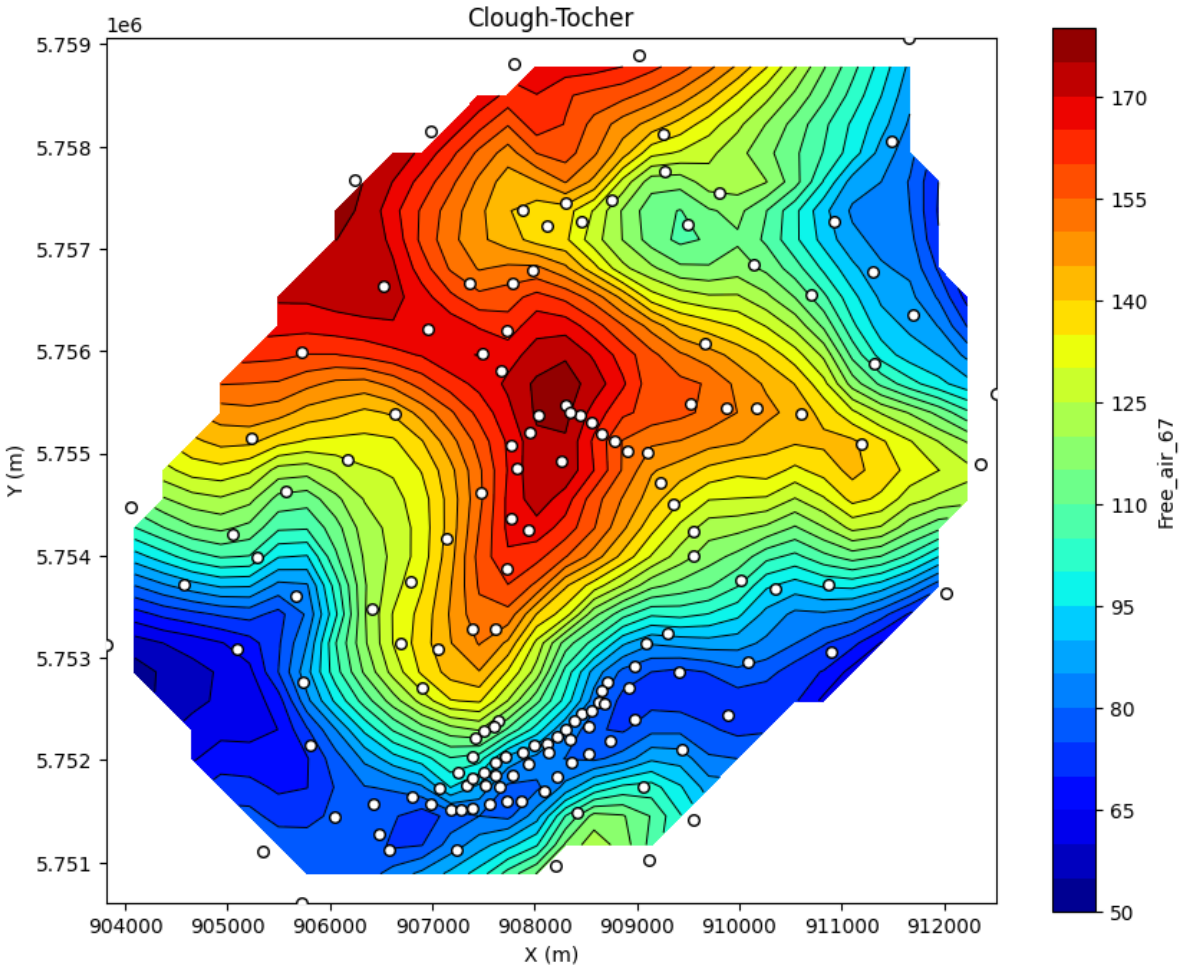

Figura 3.3: L'immagine illustra l'interpolazione effettuata dall'algoritmo Clough Tocher, algoritmo di tipo spline; sugli assi x ed y ci sono le coordinate cartesiane, la scala di colore è la anomalia di gravità, i punti bianchi sono i dati originali forniti.

Come è possibile notare nella figura 3.3, l'interpolazione mediante l'algoritmo spline Clough Tocher non è in grado di interpolare completamente l'area; ciò è dovuto dalla mancanza di punti nelle aree vicino ai bordi della figura. Per ovviare a questo problema è stato adottato un approccio ibrido: il metodo di interpolazione primario genera i punti, se non tutti i punti sono stati elaborati, un metodo secondario, di tipo Nearest neighbor in questo caso, conclude l'interpolazione. Segue la procedura per realizzare l'interpolazione:

1. L'estrazione delle coordinate e del dato da interpolare;
2. generazione della griglia di punti per la quale la funzione interpolante dovrà generare il rispettivo dato;
3. costruzione del modello con i dati forniti;
4. inferenza e restituzione dei dati generati.

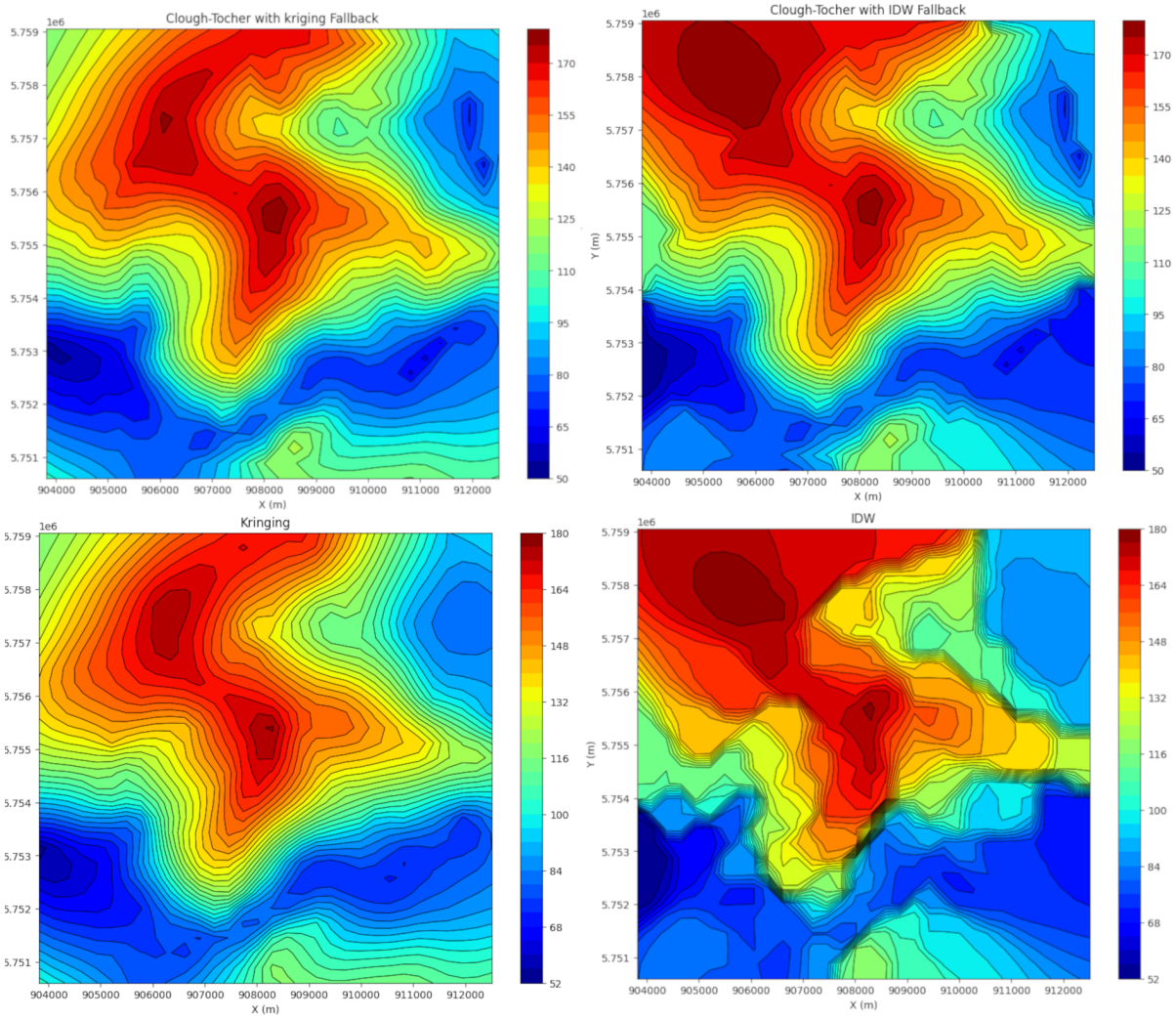

Figura 3.4: Le quattro immagini sono le interpolazioni effettuate con le 3 tecniche; sugli assi ci sono le coordinate; la scala di colore è la anomalia di gravità.

Nella figura 3.4 si osserva come IDW ha un comportamento abbastanza diverso rispetto agli altri, la spigolosità del modello è un risultato del parametro P, nel peso. Si nota anche la similarità tra i modelli spline, i quali adottano equazioni per modellare il fenomeno descritto dalla teoria del potenziale, e il modello kriging; nonostante la diversa natura di essi.

3.3 Valutazione dei modelli con K-fold

Come menzionato precedentemente, la valutazione riusa i moduli dell'interpolazione per prevedere i punti del validation set, secondo l'algoritmo di cross-validation K-Fold 2.5:

```
Funzione K-FoldCrossValidation(
    Database,
    metodoPrimarioInterpolazione,
    metodiSecondari,
    nSeparazioni,
    nRipetizioni):

    inizializza liste delle metriche

    for 1 a nRipetizioni:

        genera partizione K-Fold con nSeparazioni

        for ogni fold:

            dividi i dati in:
                train_set
                test_set

            Z_pred = metodoPrimarioInterpolazione(train_set, test_coords)

            se Z_pred contiene valori NaN:
                per ciascun metodo in metodiSecondari:
                    sostituisci i NaN usando il metodo corrente
                    interrompi se non ci sono più NaN

            residui = valori_test - Z_pred

            calcolo metriche con i dati di fold

            aggiungi le metriche alle rispettive liste


    return metriche calcolate
```

le metriche calcolate, restituite dall'algoritmo, sono un dizionario contenente insiemi di valori che rappresentano scostamenti corretti con le rispettive metriche, per interpretare questi dati, vengono utilizzati due tipologie di plot, presenti nelle pagine seguenti.

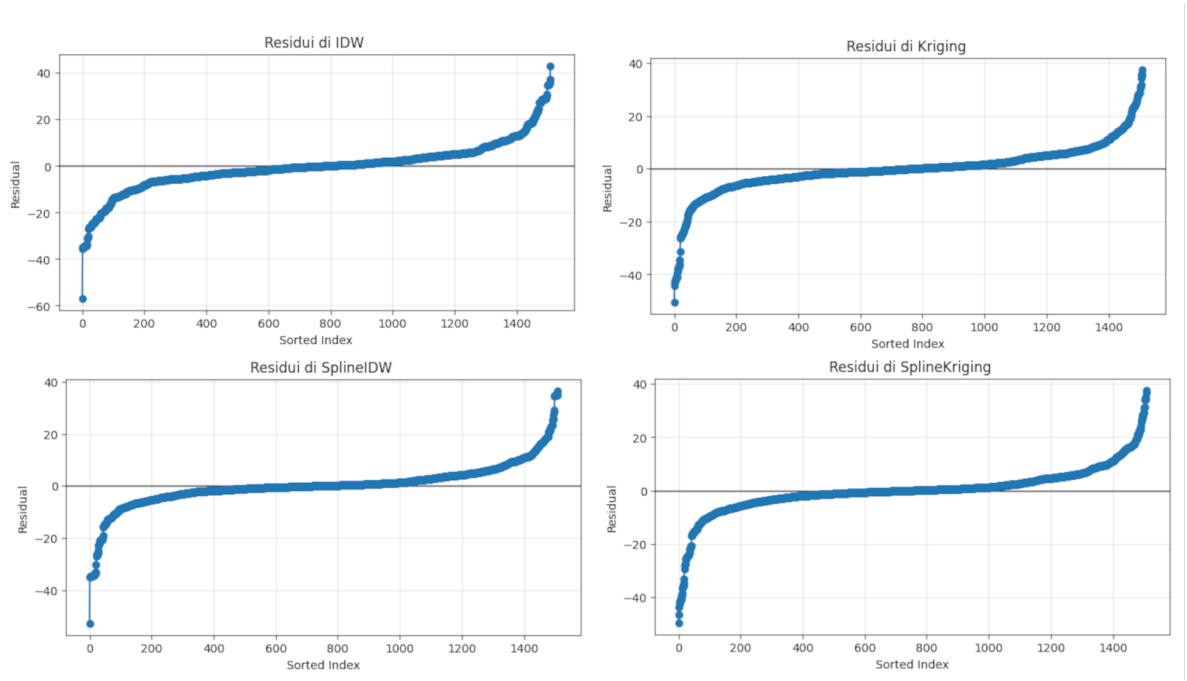

Figura 3.5: Le immagini rappresentano i residui originali ottenuti mediante K-Fold per i quattro metodi interpolanti; sulle ordinate si trovano i punti, mentre sull'ordinata di quanto si scostano dal valore reale.

La figura 3.5 presenta quattro plot, in ogni plot vengono mostrati i residui generati dall'algoritmo K-Fold, calcolati come $y_{pred} - y_{obs}$, ordinati da quelli con maggiore scostamento negativo a quelli con maggiore scostamento positivo. Si noti come tutti e quattro i metodi si comportano in modo simile; la principale differenza è presente agli estremi; i metodi spline tendono ad avere meno punti con un residuo alto: la curva è più vicina allo zero, a pari ascissa; soprattutto SplineIDW, i cui estremi dell'ordinata sono inferiori rispetto alla variante kriging.

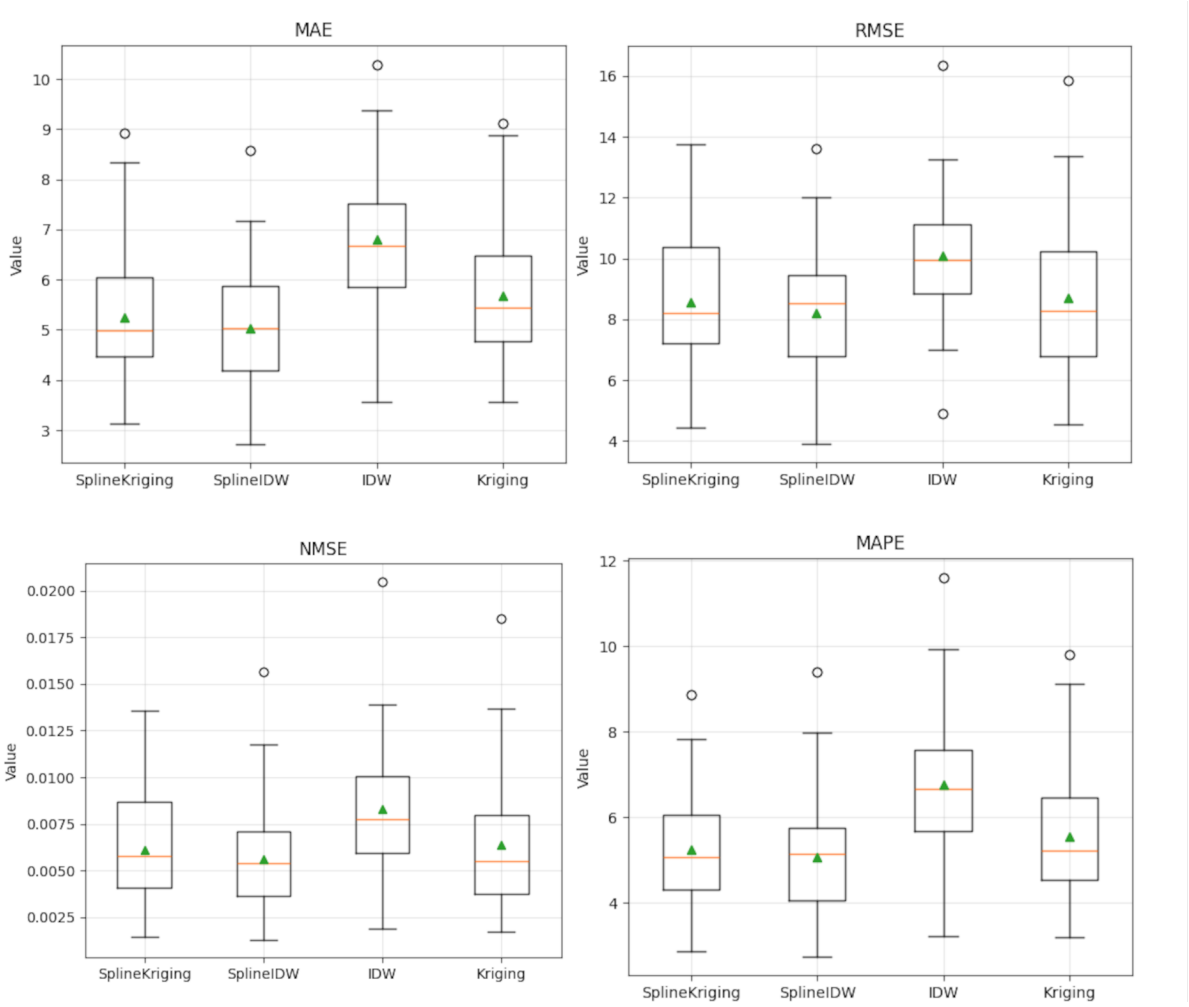

Figura 3.6: Queste immagini mostrano i boxplot delle metriche precedentemente mostrate, raggruppati per algoritmo di interpolazione

La figura 3.6 mostra il confronto tra le tecniche di interpolazione; al suo interno si trovano quattro plot, ognuno di essi raggruppa quattro boxplot, i cui valori sono ottenuti con la tecnica indicata nel titolo dei plot. In ogni boxplot è possibile leggere:

- **Mediana:** rappresenta il valore centrale della distribuzione. Divide il campione in due parti uguali (50% dei dati sopra e 50% sotto); corrisponde alla linea gialla.
- **Primo quartile (Q1):** corrisponde al 25% dei dati. Indica il valore al di sotto del quale si trova il primo quarto delle osservazioni; corrisponde al lato inferiore del rettangolo.
- **Terzo quartile (Q3):** corrisponde al 75% dei dati. Indica il valore al di sotto del quale si trova il 75% delle osservazioni; corrisponde al lato superiore del rettangolo.
- **Intervallo interquartile (IQR):** definito come $IQR = Q3 - Q1$, rappresenta la dispersione dei dati; corrisponde all'ampiezza del box.
- **Baffi (whiskers):** di estendono fino ai valori minimo e massimo non considerati outlier; corrispondono alle linee orizzontali poste agli estremi della linea verticale.

- **Outlier:** punti che si collocano oltre i baffi del boxplot. Indicano osservazioni anomale o valori estremi della distribuzione; corrispondono ai cerchi.
- **Media:** rappresenta il valore medio aritmetico della distribuzione; corrisponde al triangolo verde.

Osservando la figura 3.6, precisamente i boxplot riferiti a **MAE**, si osserva come il metodo IDW presenti valori mediani più elevati rispetto agli altri approcci, indicando un errore medio assoluto maggiore. Le varianti basate su spline, in particolare SplineKriging e SplineIDW, mostrano valori inferiori e una minore dispersione; questo indica una migliore accuratezza e stabilità del modello. Il metodo Kriging si colloca in posizione intermedia.

Considerando **RMSE**, IDW mostra ancora valori più elevati e una maggiore presenza di outlier rispetto agli altri; indicando una maggiore sensibilità agli errori estremi. Le tecniche spline, al contrario, mostrano una distribuzione più compatta e valori medi inferiori.

Su **NMSE** si osservano valori complessivamente contenuti per tutti i metodi, ma IDW presenta ancora una dispersione maggiore e valori medi superiori, mentre le varianti spline mostrano prestazioni relativamente migliori.

MAPE evidenzia che IDW produce errori percentuali medi più elevati, mentre SplineKriging e SplineIDW mantengono valori più contenuti e una distribuzione più stabile. Concludendo, è affermabile che per questo dataset, le tecniche spline, in particolare la combinazione tra CloughTocher2D e IDW ha una resa migliore sul dataset.

Capitolo 4

Messa in produzione dell'applicazione

A seguito dello sviluppo del sistema interpolante descritto nei capitoli precedenti, il passo successivo è rendere le funzionalità implementate accessibili tramite un'interfaccia Web. In particolare, la applicazione deve supportare i seguenti requisiti funzionali:

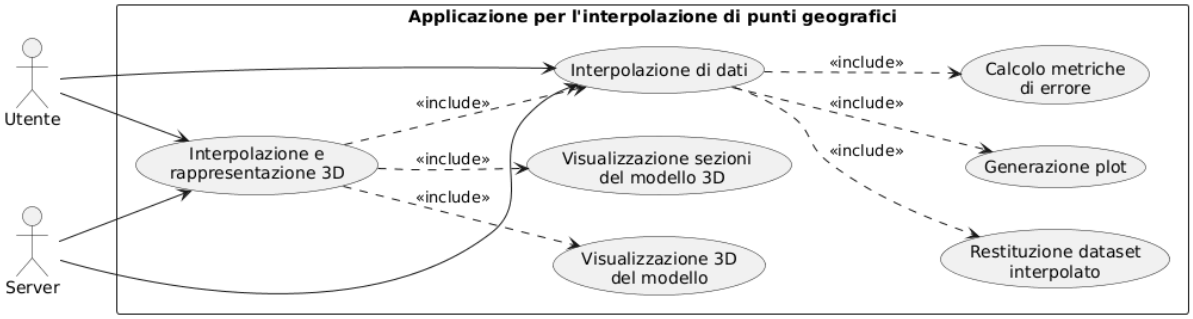

Figura 4.1: diagramma dei casi d'uso, rappresenta ciò che l'utente può fare all'interno dell'applicazione Web; realizzato con plant.uml

L'immagine 4.1 rappresenta il diagramma dei casi d'uso dell'applicativo; il diagramma dei casi d'uso è il primo approccio alla modellizzazione di un sistema software. Ha diversi usi: definire un confine di utilizzo del software, mettere per iscritto le specifiche richieste dal committente, ed è anche la base per il testing del software. In questo diagramma troviamo gli attori (omini stilizzati sulla sinistra), persone fisiche o entità che interagiscono con il sistema, e i casi d'uso, sequenze di azioni che il sistema può eseguire. Sotto si riassume ciò che è rappresentato dal diagramma:

- Esecuzione dell'interpolazione del dataset fornito dall'utente, con successiva:
  - visualizzazione dei plot generati;
  - calcolo e presentazione delle metriche di errore;
  - possibilità di scaricare il dataset interpolato.
- Integrazione del dataset interpolato con un modello di densità fornito dall'utente, al fine di:

- generare una rappresentazione tridimensionale del modello;
- consentire la visualizzazione di sezioni lungo i diversi assi cartesiani.

4.1 Ambiente d'uso e Tecnologie

Oltre ai requisiti funzionali, un altra importante considerazione è l'ambiente d'uso. È richiesto che il sistema sia usufruibile localmente e tramite Web; considerando anche l'implementazione tramite Python, è necessario un sistema client-server, capace di eseguire codice Python sul server e restituire al Web client i dati richiesti.

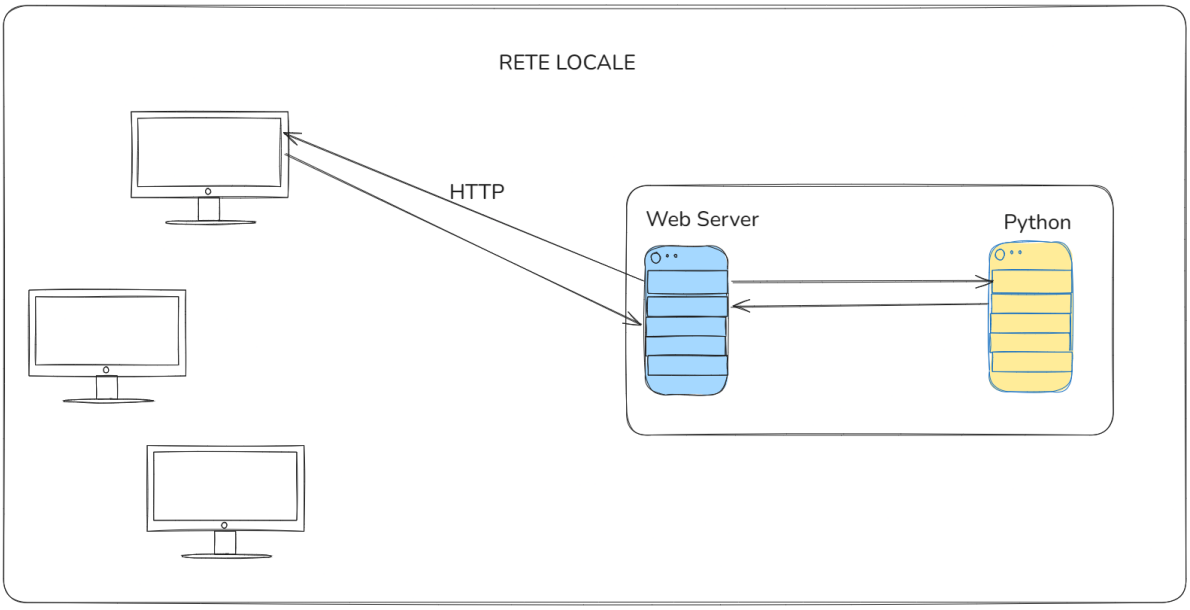

Figura 4.2: Schema che rappresenta in modo generale i componenti necessari per il sistema.

L'immagine 4.2 mostra una semplificazione dello schema di comunicazione con il Web server; a sinistra i client, a destra il server; la comunicazione tra essi avviene secondo gli standard Web; un ulteriore componente, deve occuparsi di eseguire le funzioni precedentemente implementate.
Tra gli approcci valutati, è stato adottata una soluzione uniforme, basata interamente sul linguaggio Python. Tale scelta consente di mantenere coerenza tecnologica tra la logica computazionale, già implementata in Python, e il livello applicativo Web, evitando l'introduzione di ulteriori linguaggi o ambienti di esecuzione per fare da tramite.

Python, analogamente ad altri linguaggi moderni, mette a disposizione numerosi framework per lo sviluppo di applicazioni Web e diversi server applicativi. Per garantire interoperabilità tra questi componenti, è stato adottato un approccio conforme allo standard definito nella Python Enhancement Proposal [23] 3333 (PEP 3333). La PEP 3333 definisce lo standard WSGI (Web Server Gateway Interface), che stabilisce un'interfaccia comune tra applicazioni Web scritte in Python e server Web compatibili. L'obiettivo

principale di questa specifica è disaccoppiare il livello applicativo dal livello server, consentendo a qualsiasi applicazione conforme allo standard di essere eseguita su qualsiasi server WSGI, senza modifiche sostanziali al codice.

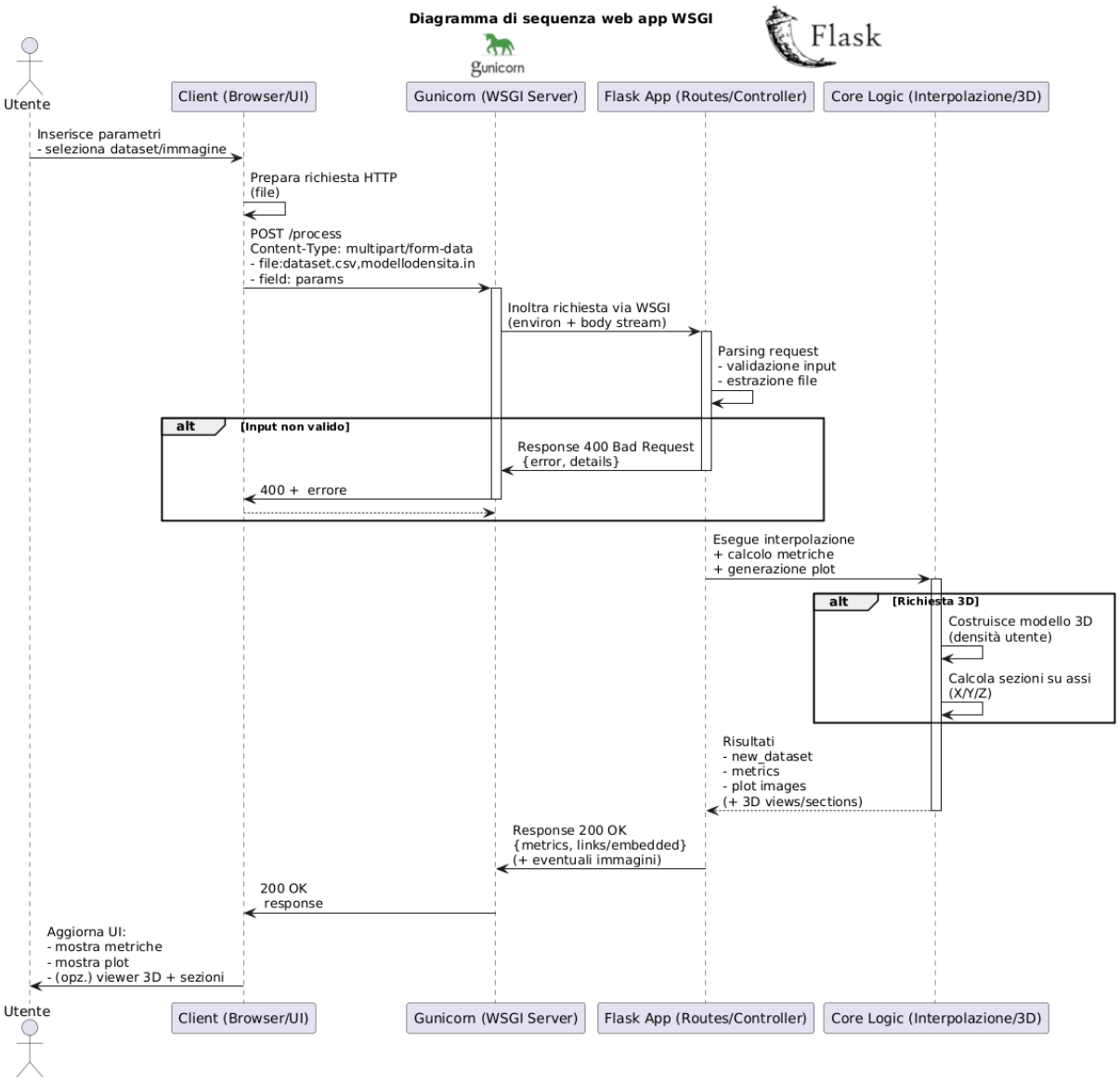

Figura 4.3: Diagramma di sequenza; descrive la comunicazione tra client e server includendo i componenti dell'architettura con WSGI, il Webserver Gunicorn e l'applicazione flask

Nella figura 4.3 viene illustrato come avviene la comunicazione tra client e server; il client, ricevuta la pagina Web dal server, è in grado di fare delle richieste verso il server; che passano prima attraverso il server Gunicorn, che invia i dati rispettando lo standard WSGI e li passa all' applicazione Flask, dove verranno eseguiti i vari algoritmi di interpolazione e plotting in base a quanto richiesto; il tutto viene restituito a Gunicorn, che lo trasforma in un pacchetto HTTP e lo restituisce al client.

4.1.1 Flask e la gestione delle route

Flask e altri framework Web moderni non espongono direttamente il file system tramite URL: le richieste non corrispondono a file statici ma a funzioni (view) registrate sulle route. Il Webserver (es. Gunicorn) costruisce l'oggetto `request` e chiama la view corrispondente; la view elabora parametri, body JSON o upload, applica validazioni e logica applicativa e restituisce una risposta (stringa, `JSON`, file, template, ecc.); tutto ciò previene l'esposizione del file system, aumentando la sicurezza.

```
app = Flask(__name__)

# Route semplice (GET), invocabile, utilizzando
# il server di sviluppo, con http://localhost:5000/
@app.route("/", methods=["GET"])
def index():
    return "Applicazione pronta", 200

# Route con query string: /search?q=term
# invocabile con http://localhost:5000/search
@app.route("/search", methods=["GET"])
def search():
    q = request.args.get("q", "")
    # processa q ...
    return jsonify({"query": q})
```

4.1.2 Gunicorn e i worker

Gunicorn è un server HTTP WSGI che funge da ponte tra il client e l'applicazione Python (Flask in questo caso); esso è composto da un processo master e da più processi worker: il master apre le socket di ascolto, gestisce il ciclo di vita dei worker (fork, restart, reload) e monitora il sistema, i worker invece eseguono effettivamente il codice dell'applicazione e servono le richieste. I worker sono processi (o thread/processi ibridi) che gestiscono le connessioni in arrivo e invocano l'applicazione WSGI per produrre la risposta; esistono diverse classi di worker:

- **sync** (predefinito): ogni worker gestisce una richiesta alla volta; semplice e stabile.
- **gthread**: worker multithreadati; ogni worker può servire più richieste contemporaneamente tramite thread.
- **gevent/eventlet** (async): worker basati su event loop e I/O non bloccante, adatti a molte connessioni I/O-bound.

4.2 Interfaccia Web

Durante la comunicazione fra client è server non vi sono continui scambi di pagine HTML (eccetto per richiesta della pagina index); il tutto avviene mediante chiamate AJAX. Asyncronous Javascript And Xml è un modo di fare richieste HTTP senza passare per il browser, utilizzando Javascript per fare richieste in modo asincrono. Normalmente, all' interazione con un link o l'invio di un form, il browser esegue una richiesta HTTP, aspettandosi come risposta una nuova pagina HTML, che sovrascriverà la pagina corrente. Usando AJ AX, dopo aver eseguito una richiesta, il server risponde restituendo un JSON; che viene intercettato ed elaborato dal Javascript.

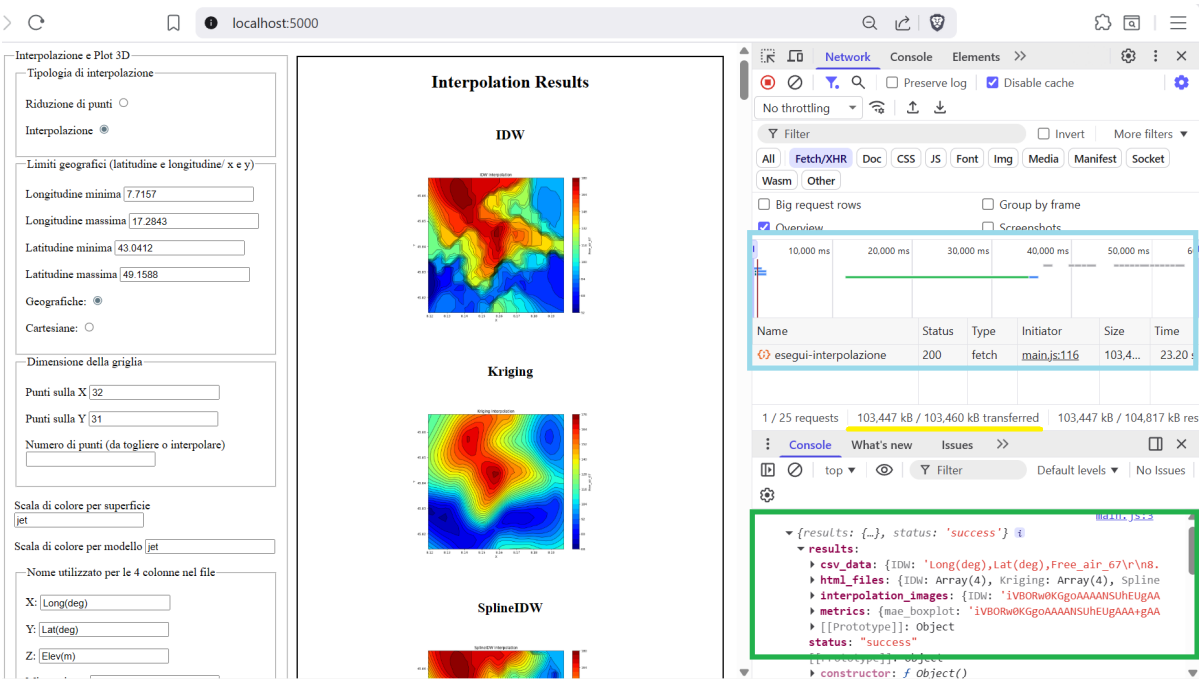

Figura 4.4: Immagine dell'interfaccia Web (Browser Brave); tramite per usufruire della applicazione.

Nell'immagine 4.4 si può vedere l'interfaccia dell'applicativo:

- Sulla sinistra è presente il form, nella quale bisogna specificare tutti i parametri necessari per effettuare l'interpolazione.
- Al centro dell'immagine, si trovano i risultati prodotti dal server, che includono i plot, le visualizzazioni 3D, e le statistiche.
- Sulla destra si trova lo strumento sviluppatore; nel riquadro azzurro, si può osservare le richieste effettuate, in questo caso una, e vedere alcune metriche, come il tempo impiegato, ovvero 23.20 secondi. In basso, nel riquadro verde si trova la console, dove è stato stampato il JSON ricevuto, composto da:
  - **csv_data**: contiene i database di dati interpolati.
  - **html_files**: sono le pagine che consentono la visualizzazione dei modelli 3D e delle sezioni.

– **interpolation_images**: le immagini dell'interpolazione, salvate in base64.
– **metrics**: le immagini delle metriche, anch'esse salvate in base64.

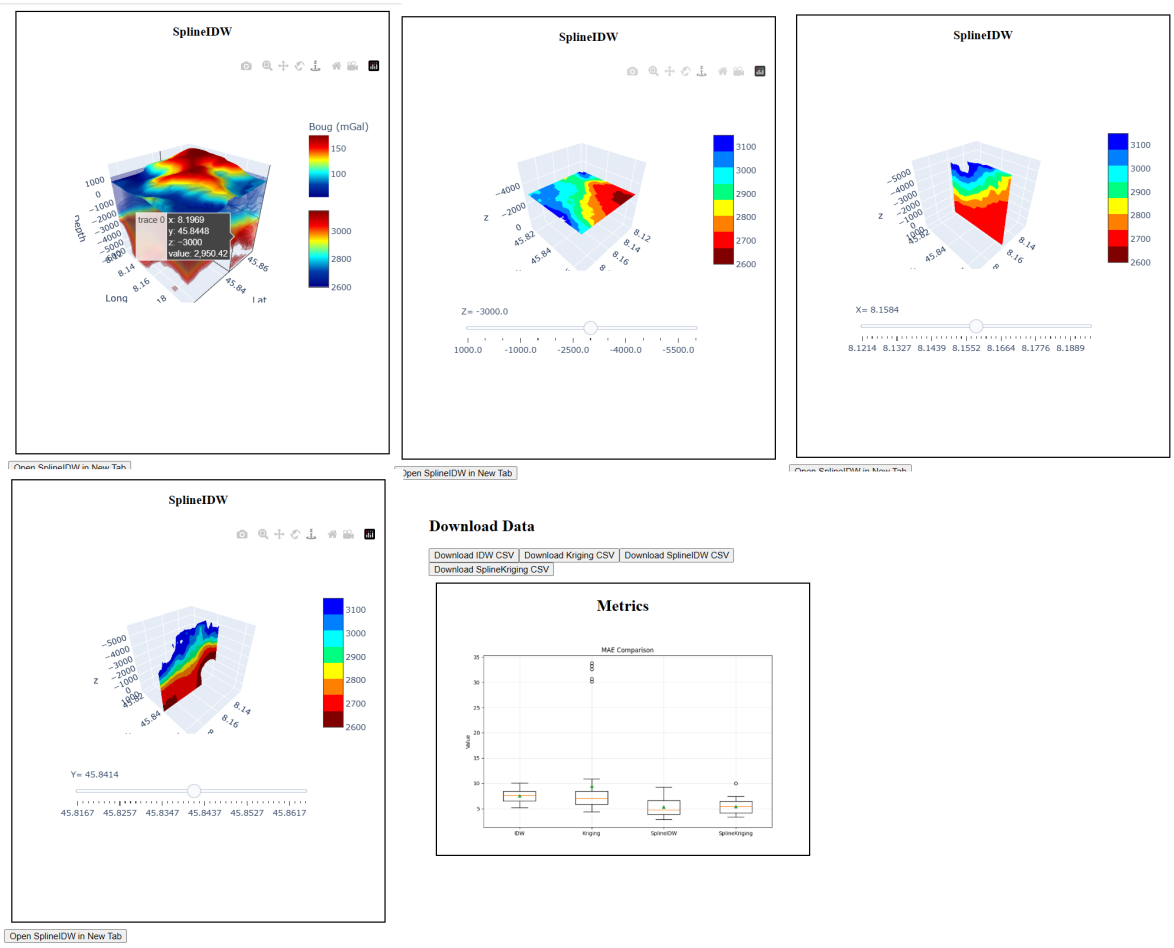

Figura 4.5: Sezioni interattive dell'interfaccia Web, generate con Plotly.

La figura 4.5 mostra le visualizzazioni interattive generate dal server, in particolare quelle generate mediante Spline + IDW, mediante la libreria Plotly. Il primo elemento in altro a sinistra è il modello di densità visualizzato insieme ai dati interpolati, in basso il bottone consente la visualizzazione in un altra pagina del modello; dal secondo fino al quarto si trovano le sezioni del modello; l'ultimo elemento mostra le statistiche dei metodi, sempre in questa sezione, i bottoni servono a scaricare i database generati.

Capitolo 5

Conclusioni e Sviluppi futuri

Le tecniche di interpolazione sperimentate hanno restituito dati all'altezza delle aspettative, insieme alla applicazione per l'usufruizione del servizio, l'obiettivo è raggiunto. Detto questo rimane molto margine di miglioramento, inanzitutto si potrebberò esplorare ulteriormente le tecniche di interpolazione:

- Affinare i parametri con Grid-search.
- Uso di nuovi algoritmi.
- Esplorazione di nuove tecniche per la valutazione delle interpolazioni.

Con l'avvento dell'intelligenza artificiale, una soluzione basata sul machine learning, combinando anche altre informazioni derivanti da altre tecniche di analisi del sottosuolo, potrebbe dare risultati interessanti.
Riguardo l'infrastruttura si può pensare di spostare il sistema all'interno di un container, per facilitare la mobilitazione, aumentare la sicurezza e tolleranza ai guasti. L'adozione di container porterebbe anche a una diversa gestione delle richieste e dei processi, questo porta a rivedere la struttura, allo scopo di spostare la gestione interna, compito di Gunicorn, a livello container; oltre ai benefici già elencati, questo trasformerebbe l'applicazione in un modulo, inseribile in una struttura già esistente.

Bibliografia

[1] Istituto Superiore per la Protezione e la Ricerca Ambientale (ISPRA). The history of the geological survey of italy. https://www.isprambiente.gov.it/en/ispra-services/forms-and-services/the-geological-survey-of-italy/the-history-of-the-geological-survey-of-italy, 2026.

[2] Lev Eppelbaum. High-precise gravity observations at archaeological sites: How we can improve the interpretation effectiveness and reliability? In *EGU General Assembly Conference Abstracts*. European Geosciences Union, 2015.

[3] A. Portal, L.-S. Gailler, P. Labazuy, and J.-F. Lénat. Geophysical imaging of the inner structure of a lava dome and its environment through gravimetry and magnetism. *Journal of Volcanology and Geothermal Research*, 2016.

[4] Vagif Gadirov, Ekrem Kalkan, Adil Ozdemir, Yildiray Palabiyik, and Kamran Gadirov. Use of gravity and magnetic methods in oil and gas exploration: Case studies from azerbaijan. *International Journal of Earth Sciences Knowledge and Applications*, 2022.

[5] Jearl Walker. *Fondamenti di Fisica*. Pearson, Milano, 2015.

[6] W. Lowrie. *Fundamentals of Geophysics*. Cambridge University Press, 1997.

[7] Harvey J. Miller. Tobler's first law and spatial analysis. *Annals of the Association of American Geographers*, 2004.

[8] European Centre for Geodynamics and Seismology (ECGS). Gravimeters. https://www.ecgs.lu/wulg/gravimeters/.

[9] John M. Reynolds. *An Introduction to Applied and Environmental Geophysics*. John Wiley & Sons, 2nd edition, 2011.

[10] Alfio Quarteroni, Riccardo Sacco, Fausto Saleri, and Paola Gervasio. *Matematica Numerica*. Zanichelli, 2014.

[11] Elsevier. Curve fitting. https://www.sciencedirect.com/topics/nursing-and-health-professions/curve-fitting, 2026.

[12] National Instruments. Overview of curve fitting models and methods in labview. https://www.ni.com/en/shop/labview/overview-of-curve-fitting-models-and-methods-in-labview.html, 2024.

[13] Kushal Shah. Is machine learning just glorified curve fitting? *Medium*, Mar 2022.

[14] Libor Mitas and Helena Mitasova. Spatial interpolation. In Paul A. Longley, Michael F. Goodchild, David J. Maguire, and David W. Rhind, editors, *Geographical Information Systems: Principles, Techniques, Management and Applications*. Wiley, 1999.

[15] QGIS Project. Coordinate reference systems. https://docs.qgis.org/3.40/en/docs/gentle_gis_introduction/coordinate_reference_systems.html, 2026.

[16] Esri. 001605: Distances for geographic coordinates (degrees of latitude and longitude) are analyzed using chordal distances in meters. https://pro.arcgis.com/en/pro-app/latest/tool-reference/tool-errors-and-warnings/001001-010000/tool-errors-and-warnings-01601-01625-001605.htm, 2026. ArcGIS Pro Documentation.

[17] GIS and Remote Sensing Education. What is spatial interpolation? what are the different methods of interpolation used in gis? *Medium*, September 30 2025.

[18] Esri. How idw works. https://pro.arcgis.com/en/pro-app/3.4/tool-reference/3d-analyst/how-idw-works.htm.

[19] C. A. Calder and Noel Cressie. Kriging and variogram models. In Rob Kitchin and Nigel Thrift, editors, *International Encyclopedia of Human Geography*. Elsevier, Oxford, 2009.

[20] Elsevier. Variogram. https://www.sciencedirect.com/topics/mathematics/variogram, 2026.

[21] Yahya Darmawan, Munawar, Dwiki Anugerah Atmojo, Huda Wahyujati, and Lamtupa Nainggolan. Accuracy assessment of spatial interpolations methods using arcgis. In *E3S Web of Conferences*. EDP Sciences, 2023.

[22] scikit-learn developers. Cross-validation: evaluating estimator performance. https://scikit-learn.org/stable/modules/cross_validation.html.

[23] Python Software Foundation. Pep 3333 – python web server gateway interface v1.0.1. https://peps.python.org/pep-3333/.