

Guida Completa GeoArchaeo Plugin

Analisi Geostatistica per l'Archeologia in QGIS

Enzo cocca

Indice

1. Introduzione
 2. Installazione e Setup
 3. Concetti Base di Geostatistica
 4. Funzionalità del Plugin
 5. Workflow Completo con Esempi
 6. Interpretazione dei Risultati
 7. Algoritmi Implementati
 8. Casi d'Uso Archeologici
 9. Risoluzione Problemi
 10. Sviluppi Futuri
-

1. Introduzione

GeoArchaeo è un plugin QGIS specializzato per l'analisi geostatistica di dati archeologici. Integra tecniche avanzate di interpolazione spaziale, machine learning e analisi multivariata per supportare archeologi e ricercatori nell'interpretazione di dati geofisici e di scavo.

Caratteristiche Principali:

- **Analisi Variogramma:** Studio della correlazione spaziale
 - **Kriging Avanzato:** Interpolazione ottimale con quantificazione dell'incertezza
 - **Machine Learning:** Pattern recognition e anomaly detection
 - **Campionamento Ottimale:** Design efficiente per nuove campagne
 - **Integrazione Multi-sensore:** Co-kriging per GPR + Magnetometria
 - **Report Automatici:** Documentazione professionale in PDF/HTML
-

2. Installazione e Setup

Requisiti:

- QGIS 3.x o superiore
- Python 3.6+
- Librerie Python richieste:

```
numpy >= 1.19.0
scipy >= 1.5.0
pandas >= 1.1.0
scikit-learn >= 0.23.0 (opzionale, per ML)
plotly >= 4.14.0 (per visualizzazioni interattive)
pykrige >= 1.5.0 (per kriging avanzato)
reportlab (opzionale, per PDF professionali)
```

Installazione:

1. Copia la cartella GeoArchaeo in:
 - Windows: C:\Users\[username]\AppData\Roaming\QGIS\QGIS3\profiles\default\python\plugins
 - macOS: ~/Library/Application Support/QGIS/QGIS3/profiles/default/python/plugins/
 - Linux: ~/.local/share/QGIS/QGIS3/profiles/default/python/plugins/
2. Riavvia QGIS e attiva il plugin dal Plugin Manager

Primo Utilizzo:

1. Clicca sull'icona GeoArchaeo nella toolbar
2. Usa "Carica Layer di Test" per dati di esempio
3. Esplora le varie funzionalità con i dati forniti

3. Concetti Base di Geostatistica

3.1 Variogramma

Il variogramma misura come varia la correlazione tra punti al crescere della distanza.

Parametri chiave: - **Nugget:** Variabilità a distanza zero (errore di misura + variabilità microscopica) - **Sill:** Varianza totale del fenomeno - **Range:** Distanza oltre la quale i punti non sono più correlati

3.2 Kriging

Metodo di interpolazione che: - Fornisce la migliore stima lineare non distorta (BLUE) - Quantifica l'incertezza delle predizioni - Considera la struttura spaziale dei dati

3.3 Pattern Recognition

Tecniche di ML per identificare: - Cluster di siti simili - Anomalie nei dati geofisici - Pattern temporali negli scavi stratificati

4. Funzionalità del Plugin

4.1 Analisi Variogramma

Input: - Layer punti con attributo numerico (es. pottery_count, magnetic_intensity) - Distanza massima di analisi - Numero di lag (intervalli di distanza) - Modello teorico (Sferico, Esponenziale, Gaussiano, Matérn)

Output: - Grafico variogramma empirico vs teorico - Parametri del modello fittato (nugget, sill, range) - Test di anisotropia direzionale - File JSON con parametri salvati

Esempio pratico:

Input: 50 punti con conteggio ceramica
Distanza max: 100m
Modello: Sferico

Output:

- Nugget: 12.3 (variabilità locale)
- Sill: 145.7 (varianza totale)
- Range: 35.2m (correlazione spaziale)

4.2 Kriging (Interpolazione)

Tipi disponibili: 1. **Ordinary Kriging:** Standard, assume media costante ma sconosciuta 2. **Universal Kriging:** Con trend spaziale (deriva) 3. **Co-Kriging:** Usa variabili secondarie correlate

Input: - Layer punti con valori - Variogramma (calcolato o importato) - Risoluzione griglia output (es. 1m) - Estensione area interpolazione

Output: - Raster interpolato (.tif) - Raster varianza predizione - Cross-validation metrics (RMSE, MAE, R^2)

Interpretazione Cross-Validation: - $RMSE < 10\%$ del range dati = Ottimo
- $MAE =$ Errore medio assoluto - $R^2 > 0.7 =$ Buona capacità predittiva

4.3 Machine Learning

Algoritmi disponibili:

K-Means (Clustering)

- **Uso:** Raggruppare siti archeologici simili
- **Input:** Multi-layer con attributi numerici
- **Output:** Layer con cluster ID colorati

DBSCAN (Density-Based Clustering)

- **Uso:** Identificare concentrazioni anomale

- **Vantaggio:** Trova cluster di forma irregolare
- **Parametri:** eps (raggio), min_samples

Random Forest

- **Uso:** Classificare siti/periodi
- **Richiede:** Layer training con etichette
- **Output:** Probabilità di appartenenza

Isolation Forest

- **Uso:** Detectare anomalie (es. sepolture, strutture)
- **Output:** Score anomalia (0=normale, 1=anomalo)

4.4 Campionamento Ottimale

Metodi:

1. **Maximin:** Massimizza distanza minima tra punti
 - Migliore copertura spaziale uniforme
2. **Riduzione Varianza:** Minimizza varianza kriging
 - Ottimale per interpolazione accurata
3. **Stratificato:** Basato su zone/strati
 - Utile per aree eterogenee

Input: - Punti esistenti - Numero nuovi punti desiderati - Area di studio (poligono) - Metodo selezione

Output: - Layer punti suggeriti (stelle gialle) - Priorità campionamento - Riduzione varianza attesa

5. Workflow Completo con Esempi

Esempio 1: Analisi Distribuzione Ceramica

Obiettivo: Mappare densità ceramica in sito romano

Dati: - 45 punti campionamento - Conteggio frammenti per m² - Area 100x150m

Procedura: 1. **Carica dati:** pottery_distribution.csv 2. **Statistiche descrittive:** Media: 45.2 frammenti/m² Dev.Std: 23.4 Skewness: 1.2 (distribuzione asimmetrica)

3. Variogramma:

- Modello: Sferico
- Range: 32m (dimensione tipica strutture)
- Nugget/Sill: 0.15 (bassa variabilità locale)

4. Kriging:

- Risoluzione: 0.5m

- Cross-validation: RMSE=8.2, $R^2=0.83$

5. Interpretazione:

- Alta concentrazione ($>80 \text{ fr/m}^2$) = Aree attività
- Media (40-80) = Occupazione normale
- Bassa (<40) = Aree marginali

Esempio 2: Integrazione GPR + Magnetometria

Obiettivo: Identificare strutture sepolte

Procedura: 1. **Carica entrambi i dataset** 2. **Co-Kriging** con GPR primario, MAG secondario 3. **ML Anomaly Detection** su raster combinato 4.

Risultato: Mappa probabilità strutture

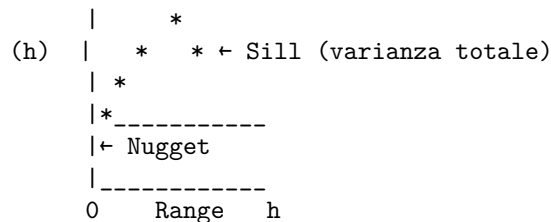
Esempio 3: Pianificazione Scavo

Situazione: 20 saggi esistenti, budget per 10 nuovi

Procedura: 1. **Variogramma** su dati esistenti 2. **Campionamento ottimale** - Metodo varianza 3. **Risultato:** 10 posizioni che massimizzano informazione 4. **Stima miglioramento:** -35% varianza predizione

6. Interpretazione dei Risultati

6.1 Lettura Grafici Variogramma



Interpretazioni comuni: - **Nugget alto:** Molta variabilità locale o errori misura - **Range corto:** Fenomeno molto variabile spazialmente - **Crescita lineare:** Presenza di trend, usa Universal Kriging

6.2 Mappe Kriging

Raster Predizione: - Colori caldi = Valori alti - Colori freddi = Valori bassi
- Contour lines = Isovalore

Raster Varianza: - Bassa vicino ai punti campionati - Alta in aree scoperte - Guida per nuovo campionamento

6.3 Risultati ML

Cluster Map: - Colori diversi = Gruppi diversi - Punti isolati = Possibili outlier - Cluster grandi = Pattern dominanti

Anomaly Score: - 0-0.3 = Normale - 0.3-0.7 = Interessante - 0.7-1.0 = Anomalia forte

7. Algoritmi Implementati

7.1 Variogramma Empirico

$$\gamma(h) = 1/(2N(h)) \sum [z(x_i) - z(x_i+h)]^2$$

Dove $N(h)$ = numero coppie a distanza h

7.2 Ordinary Kriging

Sistema di equazioni:

$$\sum_{i=1}^n \lambda_i z(x_i) + \lambda_0 = z(x_0) \quad \text{per } j=1..n$$
$$\sum_{i=1}^n \lambda_i = 1$$

7.3 DBSCAN

1. Per ogni punto p :
 - Se $|N(p, \epsilon)| \geq \text{minPts}$: p è core point
 - Espandi cluster da p
2. Punti non raggiunti = rumore

7.4 Isolation Forest

$$\text{Score}(x) = 2^{(-E(h(x))/c(n))}$$

Dove $E(h(x))$ = lunghezza media percorso negli alberi

8. Casi d'Uso Archeologici

8.1 Prospezione Geofisica

Problema: Interpretare anomalie magnetiche **Soluzione:** - Kriging per interpolare tra linee survey - Isolation Forest per evidenziare anomalie - Co-kriging con resistività per validazione

8.2 Analisi Distribuzione Manufatti

Problema: Comprendere uso spazio in abitato **Soluzione:** - Variogramma per scale attività - Kriging densità per aree funzionali - Clustering per identificare workshop

8.3 Scavi Stratigrafici

Problema: Correlare strati tra saggi distanti **Soluzione:** - Kriging 3D (quota + attributi) - Pattern recognition su ceramica diagnostica - Interpolazione cronologica

8.4 Survey Territoriale

Problema: Ottimizzare copertura in area vasta **Soluzione:** - Campionamento adattivo basato su primi risultati - Stratificazione per geomorfologia - Predizione aree alto potenziale

9. Risoluzione Problemi

Errori Comuni e Soluzioni

“**Troppi pochi punti per kriging**” - Minimo: 10-15 punti - Soluzione: Aggregare aree o usare metodi deterministici

“**Variogramma non converge**” - Dati troppo irregolari - Prova modelli diversi o riduci lag distance

“**Memory error in ML**” - Dataset troppo grande - Usa batch processing o riduci features

“**Raster vuoto/nero**” - Controlla CRS compatibili - Verifica range valori (no negativi per log) - Estensione oltre limiti dati

Best Practices

1. **Pre-processing:**
 - Rimuovi outlier estremi
 - Trasforma dati skewed (log, sqrt)
 - Standardizza unità misura
2. **Variogramma:**
 - Usa 10-15 lag
 - Max distance = 1/3 diagonale area
 - Minimo 30 coppie per lag
3. **Kriging:**
 - Risoluzione 1/2 distanza minima punti
 - Valida sempre con cross-validation
 - Considera anisotropia se presente

4. **ML:**
 - Bilancia classi per training
 - Normalizza sempre features
 - Valida su subset indipendente
-

10. Sviluppi Futuri

10.1 Miglioramenti Immediati

1. **Kriging Spazio-Temporale Avanzato**
 - Modelli non-separabili
 - Covarianze complesse
 - Integrazione date C14
2. **Deep Learning Integration**
 - CNN per pattern recognition su raster
 - Autoencoder per anomaly detection
 - Transfer learning da altri siti
3. **Visualizzazione 3D**
 - Kriging volumetrico
 - Sezioni interattive
 - Export per Unity/Unreal
4. **Analisi Incertezza**
 - Simulazioni stocastiche
 - Propagazione errori
 - Analisi sensitività

10.2 Nuove Funzionalità

1. **Integrazione Dati Multispettrali**
 - Analisi immagini drone/satellite
 - Indici vegetazione per cropmark
 - Classificazione supervisionata
2. **Network Analysis**
 - Connettività tra siti
 - Least cost path con friction
 - Viewshed cumulativa pesata
3. **Predictive Modeling**
 - Modelli predittivi siti
 - Risk assessment per erosione
 - Scenari cambiamento climatico
4. **Realtà Aumentata**
 - Export per AR app
 - Sovrapposizione ricostruzioni
 - Tour virtuali guidati

10.3 Integrazione con Altri Strumenti

1. **Database:**
 - Connessione diretta PostgreSQL/PostGIS
 - Sync con database scavo (ARK, etc.)
 - Version control per analisi
2. **Cloud Computing:**
 - Processing su Google Earth Engine
 - Distributed kriging per big data
 - Collaborative analysis
3. **Standard Archeologici:**
 - Export CIDOC-CRM
 - Integrazione con ADS
 - Metadata Dublin Core

10.4 Community Features

1. **Template Analysis:**
 - Workflow pre-configurati per tipologie sito
 - Best practices condivise
 - Benchmark dataset
2. **Educational:**
 - Tutorial interattivi integrati
 - Esercizi con soluzioni
 - Certificazione competenze

Conclusioni

GeoArchaeo rappresenta un ponte tra metodi quantitativi avanzati e pratica archeologica. L'obiettivo è democratizzare l'accesso a tecniche geostatistiche mantenendo rigore scientifico e usabilità.

Citazione

Se usi GeoArchaeo in pubblicazioni:

{Enzo Cocca} (2025). GeoArchaeo: Geostatistical Analysis for Archaeology. QGIS Plugin v1.0. DOI: [xxx]

Documento aggiornato: Gennaio 2025 Versione Plugin: 1.0.0