

## ANALISI SPAZIALE DEI DATI AMBIENTALI (T28\_24\_3 E 4)

### MODULO 1 - R (LIVELLO BASE PER GEOSTATISTICA)

31 MAGGIO, 5, 17 GIUGNO 2024, 9.00 - 13.00

### MODULO 2 - TECNICHE DI KRIGING (GEOSTATISTICA LIVELLO BASE)

06, 13, 20 SETTEMBRE, 11 OTTOBRE, 9.00 - 13.00

#### Obiettivi didattici

Il corso ha l'obiettivo di far acquisire ai partecipanti le principali conoscenze, sia di tipo teorico che pratico, necessarie per produrre un'analisi spaziale (geografica) dei dati ambientali tramite l'utilizzo operativo del linguaggio di programmazione R.

L'analisi spaziale (geografica) dei dati riguarda i metodi statistici che consentono la valutazione, la descrizione e la modellizzazione delle variabili ambientali che considerano come elemento fondamentale l'informazione spaziale caratterizzante tutte le basi dati nell'ambito delle Scienze della Terra, dell'Ambiente e della Natura.

Le tecniche di geostatistica, e tra queste il kriging, consentono la descrizione ed il trattamento statistico appropriato di una caratteristica fondamentale dei fenomeni naturali che è rappresentata dalla continuità (o discontinuità) di tipo spaziale (geografico).

Il corso enfatizza gli aspetti pratici attraverso l'utilizzo di un software dedicato (R + pacchetti specifici) e la modellizzazione dei dati spaziali rappresentati da dataset geografici, evitando, per quanto possibile, il formalismo matematico degli aspetti più strettamente teorici (che in ogni caso non possono essere completamente trascurati).

R è un linguaggio di programmazione open-source specializzato per l'analisi statistica, l'elaborazione dati, la creazione di modelli statistici e la visualizzazione grafica. Lo sviluppo del codice sorgente è sostenuto da R Core Team e R Foundation for Statistical Computing (<https://www.r-project.org/>). Il software è installabile su differenti piattaforme (UNIX e similari, Windows, MacOS) e viene ampiamente utilizzato sia in ambito accademico che industriale grazie all'enorme possibilità di estensione fornita dallo sviluppo di librerie altamente specializzate.

R offre degli strumenti potenti e sofisticati in numerosissimi e diversificati ambiti di attività in cui gli utenti possono scrivere script e funzioni personalizzate per automatizzare compiti ripetitivi ed estendere le funzionalità di base.

#### Destinatari

I destinatari sono individuati tra il personale tecnico ARPAV che si occupa di analisi ed elaborazione dati nelle diverse matrici di competenza trattate dalle strutture DRQA, DRST, DRFT, DRL e ATG. Altro personale potenzialmente interessato riguarda le strutture organizzative, sia a carattere regionale che provinciale (Dipartimenti), con un interesse specifico per l'analisi statistica e la visualizzazione grafica dei dati ambientali.

Ripartizione indicativa dei posti disponibili per strutture ARPAV:

Struttura	N. partecipanti
DRQA	6
DRST	6
DRTF	6
DRL	2
Dip. Provinciali	3
ATG	2
<b>Totale partecipanti (max)</b>	<b>25</b>

## Requisiti di accesso

---

Ai partecipanti è richiesta una conoscenza di base delle tecniche di analisi statistica dei dati ambientali, valutazione ed elaborazione grafica dei risultati.

Il modulo 1 è obbligatorio solo per i partecipanti che non hanno una conoscenza base di R applicata alla geostatistica.

## Referente scientifico

---

Massimo Bressan, tecnico ambientale ARPAV presso DRQA, U.O. QUALITA' DELL'ARIA – UMAVO

## Metodologia

---

I contenuti del corso sono di carattere pratico e finalizzati ad acquisire le basi teorico-pratiche delle moderne tecniche di geostatistica per l'analisi dei dati spaziali/geografici in ambiente GIS tramite l'utilizzo dell'ambiente di programmazione R e di alcuni pacchetti dedicati.

Gli incontri si svolgeranno da remoto in modalità videoconferenza su piattaforma Zoom.

## Attività pre-work

---

All'avvio del corso tutti i partecipanti devono essere dotati di un PC con preinstallato il software R, Rstudio e pacchetti dedicati.

### Nota importante:

**Durante il corso non sarà possibile fornire assistenza né dedicare del tempo per effettuare l'installazione del software o la risoluzione di eventuali problemi di installazione e configurazione.**

Tutti i partecipanti sono invitati a contattare preventivamente i propri referenti informatici di struttura per concordare l'installazione sul proprio PC di:

1. R, <https://cran.r-project.org/>, installare ultima versione disponibile R  $\geq 4.0.x$ , sono richiesti i privilegi di amministratore;
2. Rstudio, <https://posit.co/download/rstudio-desktop/>, installare ultima versione disponibile in base al proprio sistema operativo, sono richiesti i privilegi di amministratore;
3. tidyverse, <https://www.tidyverse.org/>, metapacchetto che comprende le seguenti librerie: dplyr, ggplot2, readr, stringr, purrr, tibble, stringr, tidyr, forcats, installazione da effettuare all'interno dell'ambiente R / Rstudio, non richiede privilegi di amministratore, `install.packages("tidyverse")`;
4. gstat: Spatial and Spatio-Temporal Geostatistical Modelling, Prediction and Simulation, <https://cran.r-project.org/web/packages/gstat/index.html>, installazione da effettuare all'interno dell'ambiente R / Rstudio, non richiede privilegi di amministratore, `install.packages("gstat")`.

## Durata e articolazione contenuti del corso

---

Il corso si articola in due moduli:

- Modulo 1 - R (livello base per geostatistica): 3 incontri di 4 ore ciascuno
- Modulo 2 - Tecniche di kriging (geostatistica livello base): 4 incontri di 4 ore ciascuno.

Il modulo 1 è obbligatorio solo per i partecipanti che non hanno una conoscenza di base di R applicato alla geostatistica.

## Programma

---

### Modulo 1 - R (livello base per geostatistica)

- Introduzione
  - R: cos'è, com'è strutturato (base e pacchetti) e potenzialità (panoramica)
  - RStudio: a cosa serve e come funziona (panoramica)
- Concetti di base
  - sintassi riga di comando e utilizzo script, help online
  - assegnazione ed utilizzo variabili
  - principali strutture dati: vettori, matrici, liste, data frame (tibble)
  - operatori matematici, relazionali, logici
  - dataset subsetting, esempi di Importazione, manipolazione ed esportazione dati
- Funzioni definite dall'utente e pacchetti specifici
  - cosa sono e a cosa servono, creazione di una funzione
- Trasformazione dati
  - package tidyverse, tidy dataset, split-apply-combine
  - verbs: select, filter, mutate, summarise
  - format: long and wide data structure
- Elaborazione grafica dei dati
  - grafica base (cenni)
  - ggplot
- Esempio di "data workflow"
  - importazione dataset, manipolazione e creazione nuove variabili
  - computazione statistiche, creazione grafici, esportazione tabellare e grafica

### Modulo 2 - Tecniche di kriging (geostatistica livello base)

- Basi teoriche
  - Richiamo di concetti statistici essenziali: varianza, covarianza, correlazione, autocorrelazione, fondamenti di algebra lineare nella regressione lineare
  - Assunzioni di stazionarietà e isotropia nei processi di kriging e approcci per "rilassare" le condizioni
  - I fondamenti statistici dell'analisi spaziale dei dati e le variabili regionalizzate
- Analisi spaziale dei dati
  - stima per interpolazione e metodi deterministici IWD (inverso pesato sulla distanza)
  - Kriging come tecnica di interpolazione esatta che onora i punti
  - Covarianza, correlogramma e variogramma, Kriging variance, come viene calcolata
  - Studio variografico di una variabile georeferenziata, variogramma sperimentale e modello di variogramma: range, sill e nugget. Effetto nugget.
  - Confronto obiettivi, finalità e limiti dei differenti "flavours": ordinary (OK), universal (UK), co- (CK), external drift (KED), regression (RK), indicator (IK)
  - Block vs. punctual kriging (due approcci alternativi di interpolazione, vantaggi, svantaggi)
- Valutazione della qualità della predizione
  - Analisi dei residui
  - Cross validation (resampling procedure to evaluate models on a limited data sample)



Agenzia Regionale per la Prevenzione  
e Protezione Ambientale del Veneto



REGIONE DEL VENETO

- Esempi e analisi di casi

---

## Docente

**Prof. Francesco Pirotti**, Università degli Studi di Padova, TESAF Department of Land, Environment, Agriculture and Forestry - CIRGEO Interdepartmental Research Center in Geomatics  
<https://www.cirgeo.unipd.it/cirgeo/?about-us/francesco-pirotti>

---

## Materiali didattici

Il materiale didattico del corso sarà fornito su supporto informatico e potrà essere scaricato al seguente indirizzo: <https://formazione.sigmagpaghe.com>, al menu "le mie iscrizioni".

---

## Valutazione del corso

I partecipanti dovranno compilare il questionario di gradimento on-line (customer satisfaction) entro 10 giorni dalla fine del corso al seguente indirizzo: <https://formazione.sigmagpaghe.com>, al menu "le mie iscrizioni".

---

## Attestato di partecipazione

L'attestato di partecipazione verrà rilasciato a coloro che hanno frequentato almeno il 75% delle ore previste e che avranno compilato il questionario di gradimento on-line (customer satisfaction) al seguente indirizzo: <https://formazione.sigmagpaghe.com>, al menu "le mie iscrizioni".