

CAPITOLO 1

INTRODUZIONE ALL'ANALISI GEO-SPAZIALE

Teoria e metodi

Introduzione: l'analisi geo-spaziale e le sue applicazioni

L'analisi geo-spaziale comprende ogni tecnica formale applicata allo studio di entità reali situate sulla superficie terrestre sulla base delle loro proprietà geometriche e geografiche; il suo campo classico di applicazione sono le Scienze della Terra, in particolar modo nella pianificazione, ecologia, urbanistica, agronomia geologia e geologia Ambientale. In altre parole l'analisi spaziale rappresenta un insieme di processi di elaborazione applicati a dati con un significato geografico al fine di produrre nuove informazioni derivate.

Un'analisi si dice geo-spaziale se i suoi risultati dipendono dalla localizzazione geografica degli oggetti, ovvero se cambiano al variare di posizione, forma e dimensione dei suoi dati di input. L'analisi geo-spaziale può essere applicata ad una pluralità di problemi; senza avere la pretesa di essere esaustivi, ma solo allo scopo di dare una idea della diversità dei settori che impiegano queste tecniche si riporta il seguente elenco esemplificativo.

- Suitability models: modelli che individuano le localizzazioni più idonee per una certa attività o per un certo uso del suolo
 - o Siti adatti a certe attività di sviluppo economico
 - o Zonizzazione per politiche di intervento
 - o Zonizzazione per analisi di impatto o per politiche di conservazione
 - o Ecc.
- Modelli di distanza o di costo: Modelli utili all'analisi dell'influenza della posizione spaziale delle risorse
 - o Identificazione bacini di approvvigionamento
 - o Identificazione zone di vendita
 - o Localizzazione servizi
 - o Rendite di posizione
 - o ecc.
- Modelli specifici
 - o Modelli erosione
 - o Modelli disponibilità idrica
 - o Habitat fauna
 - o Ecc.
- Modelli geostatistici di previsione: sulla base di fenomeni passati (serie storiche cartografiche) o attuali, i modelli cercano di prevedere le tendenze future
- Surface modelling: sulla base di dati puntuali vengono stimate superfici continue relative al fenomeno

- Dati metereologici e loro effetti (cambiamento climatico)
- Modelli diffusione inquinanti

Caratteristiche dei dati geo-spaziali

Allo scopo di rappresentare e gestire le informazioni spaziali, è necessario utilizzare il concetto di *dato geo-spaziale*: esso deve essere sufficientemente elastico da permettere di adattarlo alla complessità dell'entità a cui si riferisce ovvero il territorio.

Il concetto di “territorio” ha subito, specialmente negli ultimi decenni, una trasformazione radicale: da semplice risorsa materiale suscettibile di sfruttamento, a spazio controllabile nel quale le differenziazioni sono viste come resistenze alla trasformazione, per arrivare ad una interpretazione in cui è riconosciuto il carattere relazionale e incerto, proprio di un sistema complesso.

Un prezioso contributo è fornito da Magnaghi (2000), per il quale il territorio è un «soggetto vivente ad alta complessità», in quanto rappresenta l'esito di un «processo di territorializzazione», ovvero un processo di strutturazione dello spazio fisico da parte della società insediata.

Claude Raffestin definisce il territorio come spazio caratterizzato dalla presenza di 13 elementi (suddivisi in 3 gruppi) che possono essere combinati in modo differente dall'attore sintagmatico (cioè colui che in esso realizza un piano o un programma): altitudine, atmosfera, bacini d'acqua dolce, coste, mari/oceani, morfologia, caratteristiche del suolo, rocce, terremoti, attività vulcanica (elementi ecologici) copertura vegetale, fauna (elementi biologici) l'uomo e le sue opere (elementi antropologici).

La complessità evidenziata dagli autori citati comporta alcune caratteristiche tipiche dei dati geospaziali.

Necessità di astrazione: il dato territoriale rappresenta una semplificazione di una realtà multiforme, per gli scopi della società umana. Per tali scopi il sistema territoriale può essere scomposto in componenti elementari secondo diversi criteri:

– identificazione di oggetti elementari con caratteristiche geometriche definite: edifici, strade, pozzi, linee elettriche, ecc.

– caratteristiche qualitative diffuse del territorio, delimitabili con confini e perimetrazioni più o meno arbitrarie: substrati geologici, usi del suolo, vegetazione, ecc.

– caratteristiche quantitative diffuse, rappresentabili per gradienti discreti: temperatura al suolo, altimetria, velocità del vento, inquinamento, ecc.

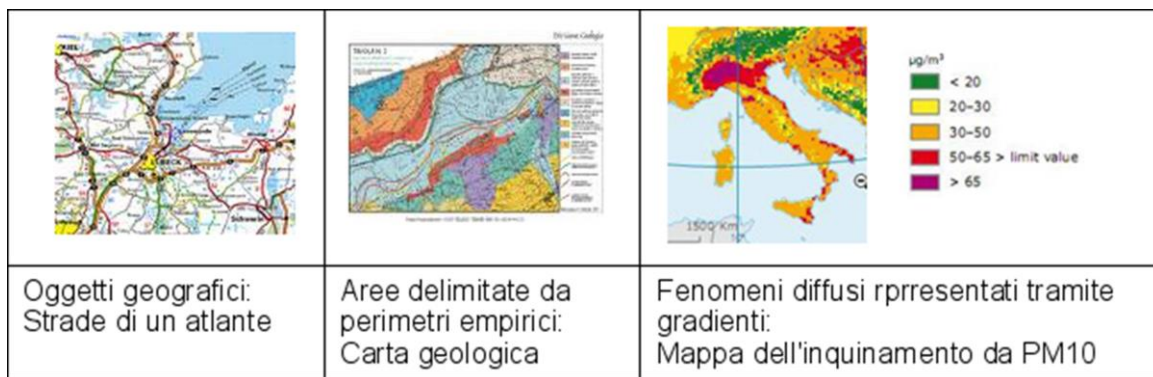


Figura 1: Scomposizione delle informazioni territoriali

Gli stessi fenomeni possono inoltre essere rappresentati in modo diverso a seconda delle necessità. Per esempio (vedi figura seguente) densità di superficie edificata per ettaro oppure territorio urbanizzato sono due concetti strettamente correlati, ma rappresentati in modo diverso (percentuale il primo, metri quadri il secondo).

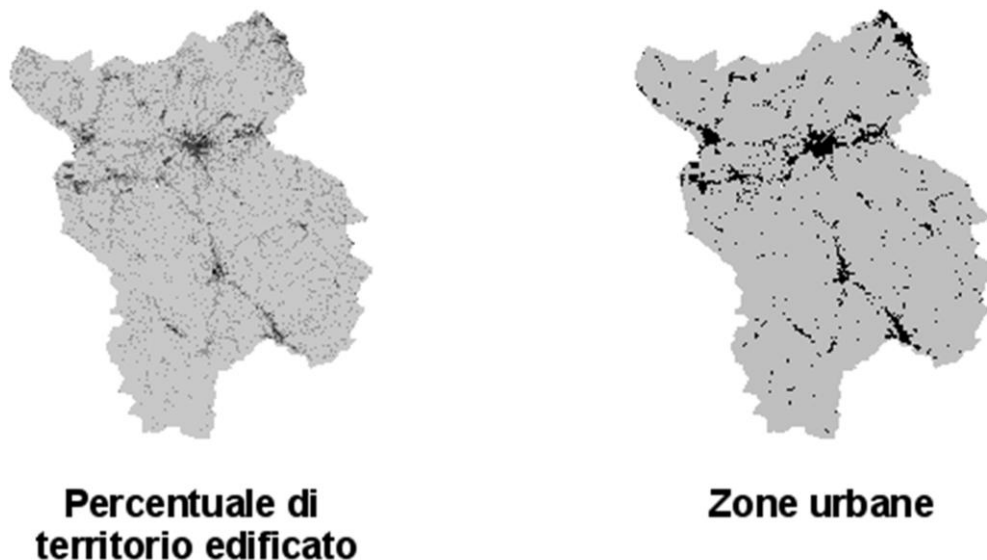


Figura 2: Astrazione del geodato

Tridimensionalità del territorio: Sul territorio possono coesistere nella stessa localizzazione geografica diverse informazioni, sovrapposte fisicamente o logicamente. Per esempio, una sovrapposizione fisico-geometrica di informazioni fra loro fisicamente collegate riguarda il substrato geologico, il suolo (fertilità, umidità, ecc.), le coltivazioni e il clima (temperatura, piovosità, ecc.). Informazioni sovrapposte dal punto di vista normativo sono i confini nazionali, regionali, comunali e catastali. L'informazione territoriale può quindi essere organizzata per strati complessi detti strati informativi o layers. Più in dettaglio quindi per strato informativo (detto anche

tema, tematismo, *coverage layer*...) si intende una collezione di entità, chiamate anche *primitive* concernenti la rappresentazione di uno specifico fenomeno reale con le seguenti caratteristiche: un'entità può appartenere ad uno e un solo strato informativo; uno strato informativo contiene entità omogenee secondo una determinata logica; l'unione di tutti i possibili strati informativi rappresenta in maniera completa il sistema "mondo reale" per gli scopi per cui è costituito il geodatabase e per la scala e la regione geografica di riferimento.

Scala di dettaglio dell'analisi: aumentando il dettaglio cresce la densità dell'informazione, ma parimenti diviene più difficoltoso realizzare astrazioni utili per l'impiego dell'informazione stessa. Una caratteristica tipica dell'informazione geografica è quindi la scala a cui questa è impiegata per prendere decisioni o costruire piani o programmi (vedi definizione del territorio di Raffestin) o per strutturare il territorio stesso (vedi definizione di Magnaghi). La scala, come vedremo, può essere espressa in termini numerici, ma anche con concetti amministrativi (scala comunale, regionale, nazionale, globale).

Regione geografica di riferimento: la definizione della scala comporta generalmente anche la definizione di confini spaziali generali a cui riferire l'insieme di dati. Questi possono essere definiti tramite una coppia di coordinate spaziali che individuano una superficie rettangolare oppure la regione può avere confini complessi derivanti da partizioni di tipo amministrativo, geografico o altro.

Modalità di rilevamento del dato: i dati territoriali possono essere rilevati direttamente percorrendo e misurando l'area oggetto di analisi con strumenti topografici (dai tradizionali strumenti ottici, ai recenti sistemi di posizionamento satellitare, meglio noti come GPS). Sempre più impiegata è però la tecnica di rilevamento a distanza, o telerilevamento, attraverso la ripresa di immagini da mezzi aerei o da satelliti. Infine, dati storici sono generalmente disponibili attraverso fonti derivate, quali mappe cartacee a loro volta derivanti da rilevamenti diretti sul territorio o da telerilevamento (vedi figura seguente).

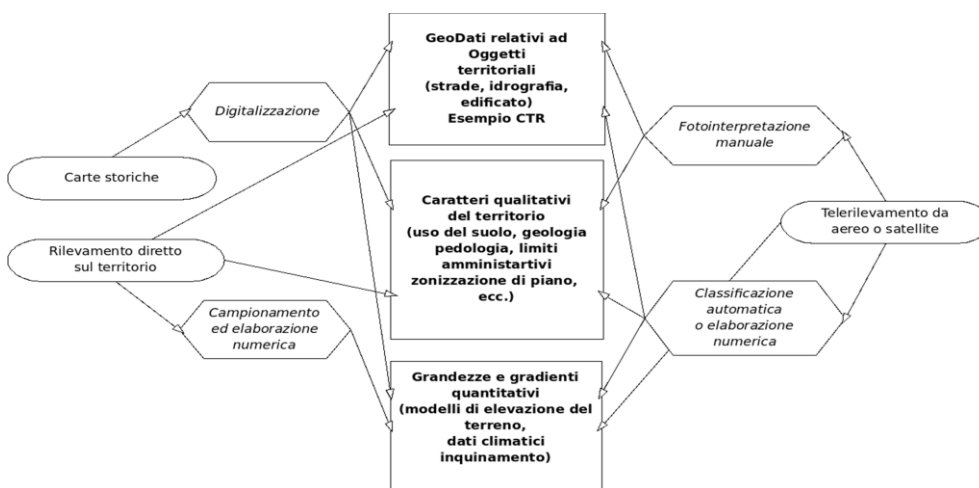


Figura 3: Fonti dei geo-dati territoriali

Il rilevamento può inoltre essere sistematico, quando si rilevano tutti gli oggetti appartenenti ad uno strato informativo (per esempio si rilevano tutte le strade di una certa regione, o quando si realizza un “mosaico” completo di tipologie di uso del suolo), oppure campionario, quando si rilevano le informazioni puntualmente secondo una griglia regolare o irregolare; le informazioni rilevate in modo campionario sono destinate ad essere elaborate in modo da coprire nel complesso l'area di indagine.

Dinamismo temporale: il territorio è un sistema in continua evoluzione, quindi il momento temporale al quale si riferisce il rilievo di una informazione rappresenta un elemento di fondamentale importanza per gli scopi a cui questa è destinata. Una tematica strettamente collegata è poi la possibilità di confrontare informazioni provenienti da momenti temporali diversi (informazioni in serie storica) in modo da poter comprendere le cause dell'evoluzione di fenomeni territoriali.

Per poter far fronte alla complessità insita nell'informazione geografica è necessario adottare strutture informative complesse formate da più dati geo-spaziali, riferiti a porzioni di spazio geografico. Esse sono definite da confini e organizzate in archivi rispondenti a scopi specifici o generali. Tali archivi sono detti *geodatabases* o Sistemi Informativi Territoriali (SIT).

Nell'ambito dei *geodatabase* il territorio nella sua complessità può essere scomposto in sottosistemi, oggetti e relazioni con lo scopo di costruire un modello semplificato del mondo reale tramite dati opportunamente strutturati e con finalità il più possibile circoscritte e definite. Si avranno quindi *geodatabases* relativi alla pianificazione territoriale, alla gestione ambientale o ecologico-naturalistica, contenenti solo gli elementi del sistema territorio necessari; essi saranno strutturati nella modalità più opportuna per lo scopo dell'analisi geospaziale alla quale sono riferiti.

Le caratteristiche fondamentali che un *geodatabase* deve possedere sono le seguenti:

- deve essere caratterizzato da una rappresentazione del territorio reale geometricamente corretta e coerente dal punto di vista spaziale: cioè il dato deve essere *georeferenziato*;
- deve consentire di poter descrivere e classificare le diverse entità territoriali in modo efficiente e flessibile per i diversi ambiti applicativi della analisi geo-spaziale: cioè deve prevedere l'organizzazione di informazioni sulle caratteristiche delle diverse componenti territoriali. Tali informazioni sono organizzate, come vedremo in dettaglio in strati informativi e/o in *attributi*.
- deve consentire di considerare e rappresentare le informazioni relative alle diverse relazioni spaziali fra gli oggetti: connessione, adiacenza, inclusione: deve cioè avere una *struttura topologica*.

Nei paragrafi seguenti si analizzeranno in dettaglio questi tre concetti.

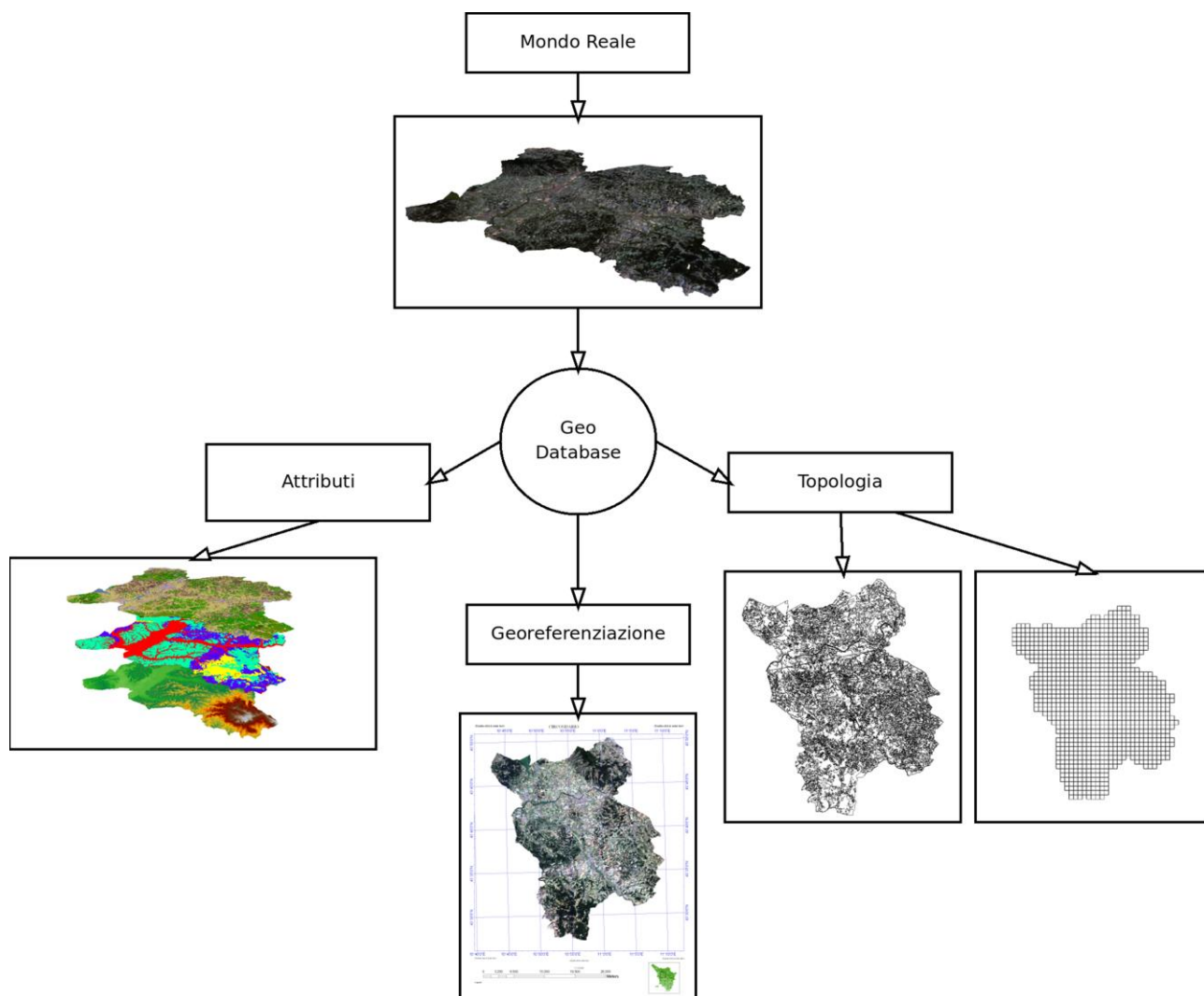


Figura 4: Struttura del geodatabase

La georeferenziazione: le proiezioni geografiche

Per georeferenziazione si intende in generale la possibilità di associare a un qualsiasi oggetto presente sul territorio le sue coordinate geografiche in un sistema di proiezione di riferimento. Lo scopo della georeferenziazione è quindi quello di rappresentare con coppie di coordinate numeriche riferite ad un piano, zone più o meno estese della superficie terrestre (vedi figura).

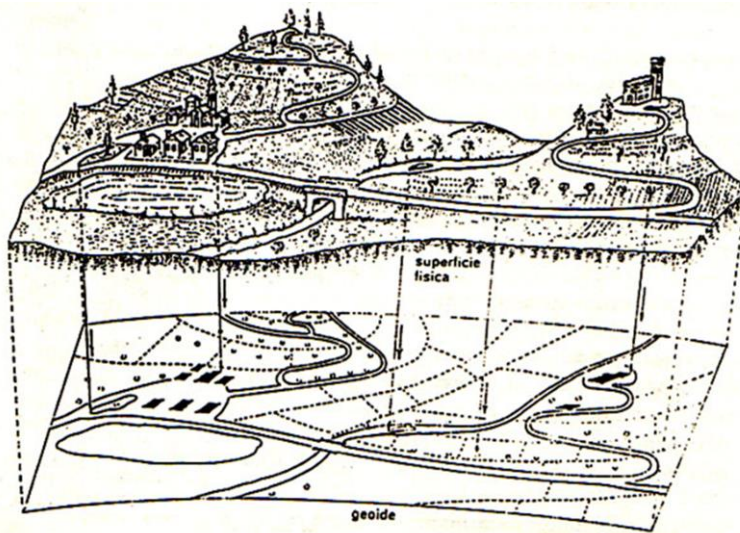


Figura 5: La georeferenziazione del mondo reale

Rimandando a pubblicazioni specializzate per una trattazione formale, per la comprensione del presente lavoro è sufficiente capire che il problema di base della georeferenziazione dei geo-dati è quello di dover rappresentare su un piano cartografico convenzionale ciò che si trova sul nostro pianeta, cioè su di una superficie curva e irregolare. La superficie che maggiormente si avvicina a quella effettiva della Terra è quella secondo la quale si disporrebbero i mari sottratti all'azione del vento, delle maree e delle correnti e sottoposti solamente agli effetti dell'attrazione terrestre e della forza centrifuga di rotazione della Terra. Tale forma è denominata geoide ed è ottenibile effettuando il prolungamento del livello medio dei mari sotto i continenti fino ad avvolgere l'intero pianeta. Tale superficie priva di asperità e ovunque convessa, in ogni suo punto si mantiene normale al vettore gravità ed è una superficie equipotenziale.

Per poter sviluppare i calcoli con lo scopo di determinare la posizione planimetrica di ogni punto della superficie effettiva, non possedendo il geoide una espressione matematica, si introduce la superficie ellissoidica di riferimento. La forma di tale superficie è data dall'appropriata grandezza di due dei parametri principali: il semiasse maggiore e lo schiacciamento determinati entrambi con operazioni di geodesia.

Per definire delle coordinate sull'ellissoide si utilizzano i concetti di meridiano e di parallelo. Per parallelo si intende il luogo dei punti sull'ellissoide in cui la normale all'ellissoide forma un angolo costante con il piano equatoriale, formato dalla rotazione dell'asse maggiore dell'ellisse generatore dell'ellissoide. Si tratta perciò di un cerchio posto su un piano secante l'ellissoide, parallelo al piano equatoriale. Per meridiano si intende invece il luogo dei punti definito dall'intersezione tra un semipiano originato dall'asse di rotazione terrestre (asse minore dell'ellisse) e la superficie dell'ellissoide. Dato un punto P_0 ubicato sull'ellissoide, si definisce latitudine geografica o ellissoidica (φ), l'angolo che la retta normale all'ellissoide passante per P_0 forma con il piano equatoriale. Si definisce longitudine geografica o ellissoidica (λ) l'angolo tra il meridiano su cui giace P_0 ed un meridiano di origine scelto arbitrariamente, in Italia generalmente si adotta il meridiano di Greenwich (GB) o quello passante per Monte Mario, presso Roma.

L'insieme dei meridiani e dei paralleli costituisce il reticolo geografico e la porzione di superficie terrestre delimitata da due meridiani delimita quello che si dice un fuso.

L'utilizzo dell'ellissoide come superficie di riferimento planimetrica rende possibile effettuare il passaggio, pur con delle inevitabili deformazioni, dalla superficie fisica della Terra alla sua proiezione cartografica. Per ottenere ciò una volta definito l'ellissoide di riferimento ed il suo orientamento per la costruzione della cartografia è fondamentale la determinazione della posizione di un limitato numero di punti caratterizzati da elevata precisione (vertici trigonometrici). Si crea perciò una rete geodetica. L'insieme dei dati di forma, dimensione e orientamento relativi all'ellissoide, uniti alla misurazione e calcolo sul terreno della rete geodetica è detto *datum* geodetico o sistema geodetico di riferimento. I datum utilizzati in Italia sono: World Geodetic Datum 1984 (WGS84), European Datum 1950 (ED50) e Roma 1940.

Una volta nota, dalle misurazioni geodetiche, la forma della Terra e scelta una superficie di riferimento che la approssimi, alla quale collegare i singoli punti reali posti sulla sua superficie fisica, si pone il problema di riportare su un piano (la carta o lo schermo del computer), mantenendo le proprietà volute, il dato geografico. Poiché l'ellissoide è una superficie a doppia curvatura, la sua trasposizione su una superficie piana comporta inevitabilmente delle deformazioni. I metodi che si utilizzano sono:

- ✓ proiezione su un piano: si hanno in questo caso proiezioni dette azzimutali che si classificano in polari, equatoriali o oblique a seconda che il piano sia tangente all'ellissoide rispettivamente ad un polo, ad un punto dell'equatore o un qualsiasi altro punto della superficie;
- ✓ proiezioni su una superficie a singola curvatura
 - cilindro: proiezioni cilindriche classificate in dirette, inverse o trasverse e oblique in base all'orientamento del cilindro;
 - nella proiezione cilindrica "diretta tangente" l'asse del cilindro è perpendicolare all'equatore;
 - nella proiezione cilindrica "inversa tangente" l'asse del cilindro è giace sull'equatore;
 - nelle proiezioni cilindriche dirette o inverse "secanti" le caratteristiche dell'asse sono le stesse degli esempi sopra riportati, ma il diametro del cilindro è inferiore rispetto all'asse dell'ellissoide; una proiezione cilindrica inversa a cilindro secante molto importante è la proiezione di Gauss, su cui si basano tra gli altri i sistemi UTM e Gauss-Boaga;
 - cono: proiezioni coniche.

I sistemi di proiezione più utilizzati in Italia sono:

- il sistema Gauss-Boaga che si basa sull'utilizzo del datum geodetico Roma40 e sulla proiezione di Gauss di due fusi, cioè su due proiezioni distinte. Ciascun fuso ha ampiezza pari a circa $6^{\circ} 30'$. Sono rispettivamente indicati con la denominazione fuso "Ovest" e fuso "Est";
- il sistema UTM (Universal Transverse Mercator) si basa come il sistema italiano sulla proiezione di Gauss, il globo è suddiviso in 60 fusi di 6° e per ciascuno di essi è effettuata una proiezione.; i fusi sono numerati a partire dall'antimeridiano di Greenwich (180°) in direzione E e l'Italia si trova sui fusi dal 32 al 34; il sistema UTM è utilizzato con diversi datum di riferimento. In Italia è utilizzato il datum Europeo del 1950 e il WGS84.

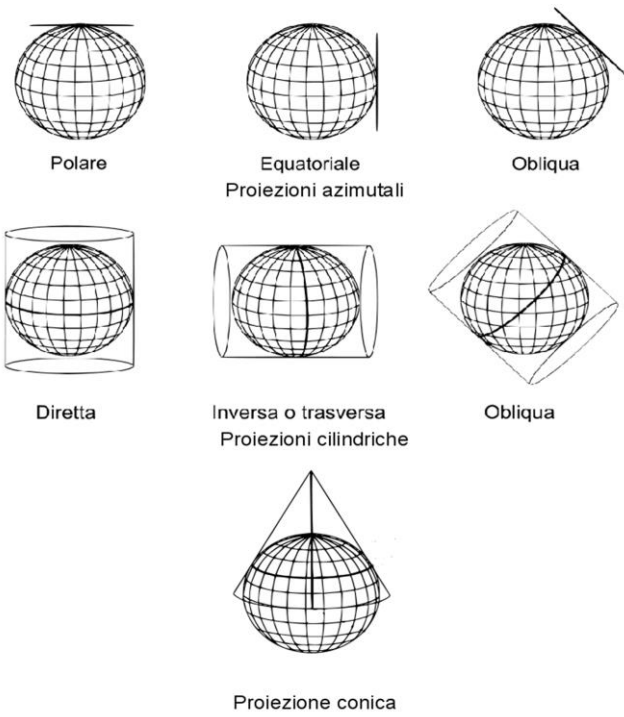


Figura 6: Sistemi di proiezione

Gli attributi: una tassonomia dei dati geospaziali

Nel recente passato sono stati proposti numerosi criteri di classificazione di dati geo-spaziali, con finalità generiche o specifiche. A livello internazionale lo standard che si sta affermando è quello collegato alla direttiva europea INSPIRE. INSPIRE (acronimo di *Infrastructure for Spatial Information in Europe* - Infrastruttura per l'Informazione Territoriale in Europa) è un progetto della Commissione Europea con l'obiettivo di realizzare infrastrutture di dati territoriali nella Comunità europea. I dati territoriali, così come definiti dalla direttiva, sono tutti quei dati che “attengono, direttamente o indirettamente, a una località o un'area geografica specifica”. La classificazione dei dati territoriali proposta dal progetto INSPIRE è stata recepita dalla direttiva europea 2007/2/CE e comprende le seguenti categorie di dati tematici:

1. Sistemi di coordinate

Sistemi per georeferire in maniera univoca le informazioni territoriali nello spazio mediante un sistema di coordinate (x, y, z) e/o latitudine e longitudine e quota, sulla base di un dato geodetico orizzontale e verticale.

2. Sistemi di griglie geografiche

Griglia multi-risoluzione armonizzata con un punto di origine comune, un posizionamento e una dimensione standard delle celle.

3. Nomi geografici

Denominazione di aree, regioni, località, città, periferie, paesi o centri abitati, o qualsiasi elemento geografico o topografico di interesse pubblico o storico.

4. Unità amministrative

Unità amministrative di suddivisione delle zone su cui gli Stati membri hanno e/o esercitano la loro giurisdizione a livello locale, regionale e nazionale, delimitate da confini amministrativi.

5. Indirizzi :

Localizzazione delle proprietà basata su identificativi di indirizzo, in genere nome della via, numero civico, codice postale.

6. Parcelle catastali

Aree definite dai registri catastali o equivalenti.

7. Reti di trasporto

Reti di trasporto su strada, su rotaia, per via aerea e per vie navigabili e relative infrastrutture.

8. Idrografia

Elementi idrografici, comprese le zone marine e tutti gli altri corpi ed elementi idrici ad esse correlati, tra cui i bacini e sub-bacini idrografici.

9. Siti protetti

Aree designate o gestite in un quadro legislativo internazionale, comunitario o degli Stati membri per conseguire obiettivi di conservazione specifici.

10. Elevazione

Modelli digitali di elevazione per superfici emerse, ghiacci e superfici oceaniche. La voce comprende l'altitudine terrestre, la batimetria e la linea di costa.

11. Copertura del suolo

Copertura fisica e biologica della superficie terrestre comprese le superfici artificiali, le zone agricole, i boschi e le foreste, le aree (semi)naturali, le zone umide, i corpi idrici.

12. Orto immagini
Immagini georeferenziate della superficie terrestre rilevate da satellite o da telesensori.
13. Geologia
Classificazione geologica del suolo. Questa voce comprende il basamento roccioso, gli acquiferi e la geomorfologia.
14. Unità statistiche
Unità per la divulgazione o l'utilizzo di dati statistici.
15. Edifici
Localizzazione geografica degli edifici.
16. Suolo
Caratterizzazione del suolo e del sottosuolo in base a profondità, tessitura (texture), struttura e contenuto delle particelle e della materia organica, pietrosità, erosione, eventualmente pendenza media e capacità prevista di ritenzione dell'acqua.
17. Uso del suolo
Classificazione del territorio in base alla dimensione funzionale o alla destinazione socioeconomica presenti e programmate per il futuro (ad esempio ad uso residenziale, industriale, commerciale, agricolo, silvicolo, ricreativo).
18. Salute umana e sicurezza
Distribuzione geografica della prevalenza di patologie (allergie, tumori, malattie respiratorie, ecc.), le informazioni contenenti indicazioni sugli effetti relativi alla salute (indicatori biologici, riduzione della fertilità e epidemie) o al benessere degli esseri umani (affaticamento, stress, ecc.) in relazione alla qualità dell'ambiente, sia in via diretta (inquinamento atmosferico, sostanze chimiche, riduzione dello strato di ozono, rumore, ecc.) che indiretta (alimentazione, organismi geneticamente modificati, ecc.).
19. Servizi di pubblica utilità e servizi amministrativi
Sono compresi sia impianti quali gli impianti fognari, di gestione dei rifiuti, di fornitura energetica, e di distribuzione idrica, sia servizi pubblici amministrativi e sociali quali le amministrazioni pubbliche, i siti della protezione civile, le scuole e gli ospedali.
20. Impianti di monitoraggio ambientale
L'ubicazione e il funzionamento degli impianti di monitoraggio ambientale comprendono l'osservazione e la misurazione delle emissioni, dello stato dei comparti ambientali e di altri parametri dell'ecosistema (biodiversità, condizioni ecologiche della vegetazione, ecc.) da parte o per conto delle autorità pubbliche.
21. Produzione e impianti industriali
Siti di produzione industriale; compresi gli impianti di cui alla direttiva 96/61/CE del Consiglio, del 24 settembre 1996, sulla prevenzione e la riduzione integrate dell'inquinamento e gli impianti di estrazione dell'acqua, le attività estrattive e i siti di stoccaggio.
22. Impianti agricoli e di acquacoltura
Apparecchiature e impianti di produzione agricola (compresi i sistemi di irrigazione, le serre e le stalle).
23. Distribuzione della popolazione — demografia
Distribuzione geografica della popolazione, comprese le relative caratteristiche e i livelli di attività, aggregata per griglia, regione, unità amministrativa o altra unità analitica.

24. Zone sottoposte a gestione/limitazioni/regolamentazione e unità con obbligo di comunicare dati

Aree gestite, regolamentate o utilizzate per la comunicazione di dati a livello internazionale, europeo, nazionale, regionale e locale. Sono comprese le discariche, le zone vietate attorno alle sorgenti di acqua potabile, le zone sensibili ai nitrati, le vie navigabili regolamentate in mare o in acque interne di grandi dimensioni, le zone per lo smaltimento dei rifiuti, le zone di limitazione del rumore, le zone in cui sono autorizzate attività di prospezione ed estrazione, i distretti idrografici, le pertinenti unità con obbligo di comunicare dati e le aree in cui vigono piani di gestione delle zone costiere.

25. Zone a rischio naturale

Zone sensibili caratterizzate in base ai rischi naturali (cioè tutti i fenomeni atmosferici, idrologici, sismici, vulcanici e gli incendi che, per l'ubicazione, la gravità e la frequenza, possono avere un grave impatto sulla società), ad esempio inondazioni, slavine e subsidenze, valanghe, incendi di boschi/foreste, terremoti, eruzioni vulcaniche.

26. Condizioni atmosferiche

Condizioni fisiche dell'atmosfera. Questa voce comprende i dati territoriali basati su misurazioni, su modelli o su una combinazione dei due e comprende i punti di misurazione.

27. Elementi geografici meteorologici

Condizioni meteorologiche e relative misurazioni; precipitazioni, temperatura, evapotraspirazione, velocità e direzione dei venti.

28. Elementi geografici oceanografici

Condizioni fisiche degli oceani (correnti, salinità, altezza delle onde, ecc.).

29. Regioni marine

Condizioni fisiche dei mari e dei corpi idrici salmastri suddivisi in regioni e sottoregioni con caratteristiche comuni.

30. Regioni biogeografiche

Aree che presentano condizioni ecologiche relativamente omogenee con caratteristiche comuni.

31. Habitat e biotopi

Aree geografiche caratterizzate da condizioni ecologiche specifiche, processi, strutture e funzioni (di supporto alla vita) che supportano materialmente gli organismi che le abitano. Sono comprese le zone terrestri e acquatiche, interamente naturali o seminaturali, distinte in base agli elementi geografici, abiotici e biotici.

32. Distribuzione delle specie

Distribuzione geografica delle specie animali e vegetali aggregate per griglia, regione, unità amministrativa o altra unità analitica.

33. Risorse energetiche

Risorse energetiche, compresi gli idrocarburi, l'energia idroelettrica, la bioenergia, l'energia solare, eolica, ecc., ove opportuno anche informazioni, in termini di altezza/profondità, sull'entità della risorsa.

34. Risorse minerarie

Risorse minerarie, compresi i minerali metallici, i minerali industriali, ecc., ove opportuno anche informazioni, in termini di altezza/profondità, sull'entità della risorse.

Le strutture dei dati e la topologia: geodati vettoriali e geodati raster

Introduzione

Come già detto i dati geo-spaziali costituiscono un'astrazione concettuale dei fenomeni e degli oggetti che si trovano nello spazio geografico reale. Ci sono due principali approcci concettuali che portano a modellizzare i geodati in formato cosiddetto vettoriale o in formato raster.

Nel formato vettoriale il territorio è discretizzato in oggetti geometrici elementari, punti, linee spezzate e aree poligonali irregolari chiuse; la georeferenziazione di questi oggetti avviene assegnando un numero sufficiente di coppie di coordinate (vettore di coppie di coordinate) tale da descrivere compiutamente l'oggetto in uno spazio bidimensionale. L'informazione territoriale è associata a ciascun oggetto, infine le relazioni spaziali fra i diversi oggetti derivano dalle loro caratteristiche geometriche. Per esempio, una strada può essere concettualizzata tramite una linea spezzata, la georeferenziazione avviene assegnando a ciascuna strada un vettore di coordinate geografiche dei vertici della spezzata, l'informazione geografica è rappresentata da una "scheda" contenente i dati relativi alla strada (larghezza, fondo stradale, proprietà, stato di manutenzione ecc.), le relazioni spaziali, per esempio gli incroci con altre strade sono ricostruibile sulla base dei vertici comuni a due oggetti oppure calcolando le coordinate di intersezione fra due tratti rettilinei della spezzata.

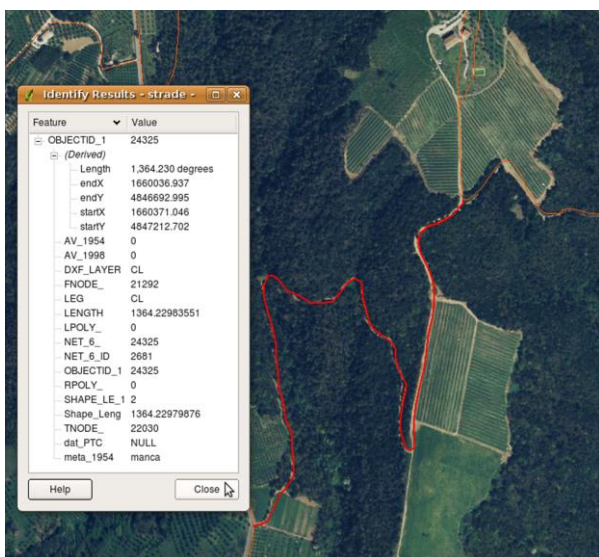


Figura 7: Dato vettoriale lineare

La seconda modalità di concettualizzazione del dato territoriale è il cosiddetto formato *raster*. I geodati raster sono rappresentati tramite una griglia generalmente quadrata che copre l'intera estensione del geo-dato. L'informazione geografica, in genere una informazione per ogni strato informativo, è attribuita sistematicamente a ciascuna cella della griglia. La presenza di oggetti

territoriali, di qualità territoriali o di quantificazione di fenomeni è in altre parole registrata sistematicamente per ciascuna cella (vedi figura seguente).

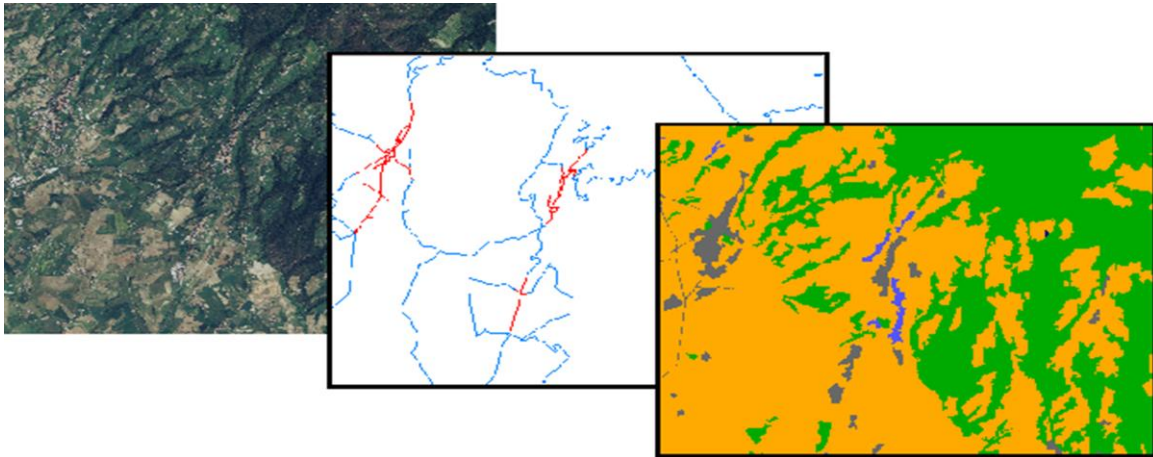


Figura 8: Dati raster: mondo reale, strade e forme di uso del suolo

Le relazioni spaziali avvengono considerando le celle confinanti nel piano o corrispondenti (come localizzazione spaziale) per quanto riguarda relazioni che coinvolgono i diversi strati informativi.

Le due modalità di organizzazione del geodato sono spesso interscambiabili:

- lo stesso fenomeno può quasi sempre essere rappresentato mediante entrambi i formati
- è possibile, con alcune limitazioni, trasformare un formato nell'altro.

Come sarà esaminato in dettaglio alla fine del capitolo, la scelta di una o dell'altra struttura dipende dall'impiego a cui viene destinato il dato.

I paragrafi successivi esaminano in dettaglio le caratteristiche dei due formati di organizzazione dei geodati territoriali.

La struttura vettoriale

Come già accennato, i dati vettoriali sono divisi in tre tipi di oggetti. I punti vengono impiegati per rappresentare oggetti, ma anche fenomeni territoriali che, alla scala del geodatabase, non hanno una dimensione spaziale apprezzabile, ma che sono considerati solo per la loro occorrenza nello spazio. Esempi tipici sono pali di illuminazione, sorgenti, alberi isolati, localizzazione di incidenti stradali, sorgenti di emissione di inquinanti, ecc. I punti sono georeferenziati tramite una singola coppia di coordinate.

Gli oggetti che hanno una delle due dimensioni fortemente prevalente sull'altra, sempre alla scala dell'indagine, sono registrati nel formato vettoriale come oggetti lineari, detti anche polilinee o archi. Gli oggetti lineari sono georeferenziati attraverso una serie ordinata di due o più coordinate geografiche. I punti nello spazio che identificano gli estremi dell'oggetto sono detti *vertici*. Gli archi sono in natura generalmente curvilinei, ma la curvatura viene discretizzata attraverso una

successione di segmenti. I punti che identificano nello spazio i diversi segmenti sono detti vertici. La densità dei vertici dipende ancora una volta dalla scala del geodatabase, ma influenza anche la modalità e i risultati delle elaborazioni che possono essere effettuate. Le informazioni sono generalmente associate all'intero segmento. Gli oggetti lineari possono essere orientati, quando è possibile individuare un vertice iniziale ed uno finale che determina l'orientamento della linea. Esempio di archi orientati sono i fiumi per i quali può essere determinata la direzione di scorrimento dell'acqua. Un reticolo di oggetti lineari orientati è detta *grafo*.

Gli oggetti che hanno una dimensione spaziale complessivamente apprezzabile, sempre alla scala di riferimento del geodatabase, sono rappresentati tramite poligoni irregolari concavi o convessi. Gli oggetti poligonali sono georeferenziati tramite una serie ordinata di coordinate in cui il primo vertice coincide con l'ultimo, chiudendo così il perimetro della figura. Anche nel caso dei poligoni, la densità dei nodi che approssima il confine complesso di una certa superficie territoriale è funzione della scala e delle finalità del geodatabase.

La topologia dei geo-dati vettoriali

La topologia è un procedimento matematico utilizzato per definire esplicitamente le relazioni spaziali esistenti fra diversi oggetti. Le informazioni topologiche, a differenza delle informazioni metriche, descrivono come si rapportano spazialmente tra di loro superfici, linee e punti e non cambiano se lo spazio viene distorto e/o deformato. Ad esempio le relazioni topologiche fra poligoni nella figura seguente non cambiano se lo spazio geometrico è deformato: ciascun poligono confina sempre con gli altri due.

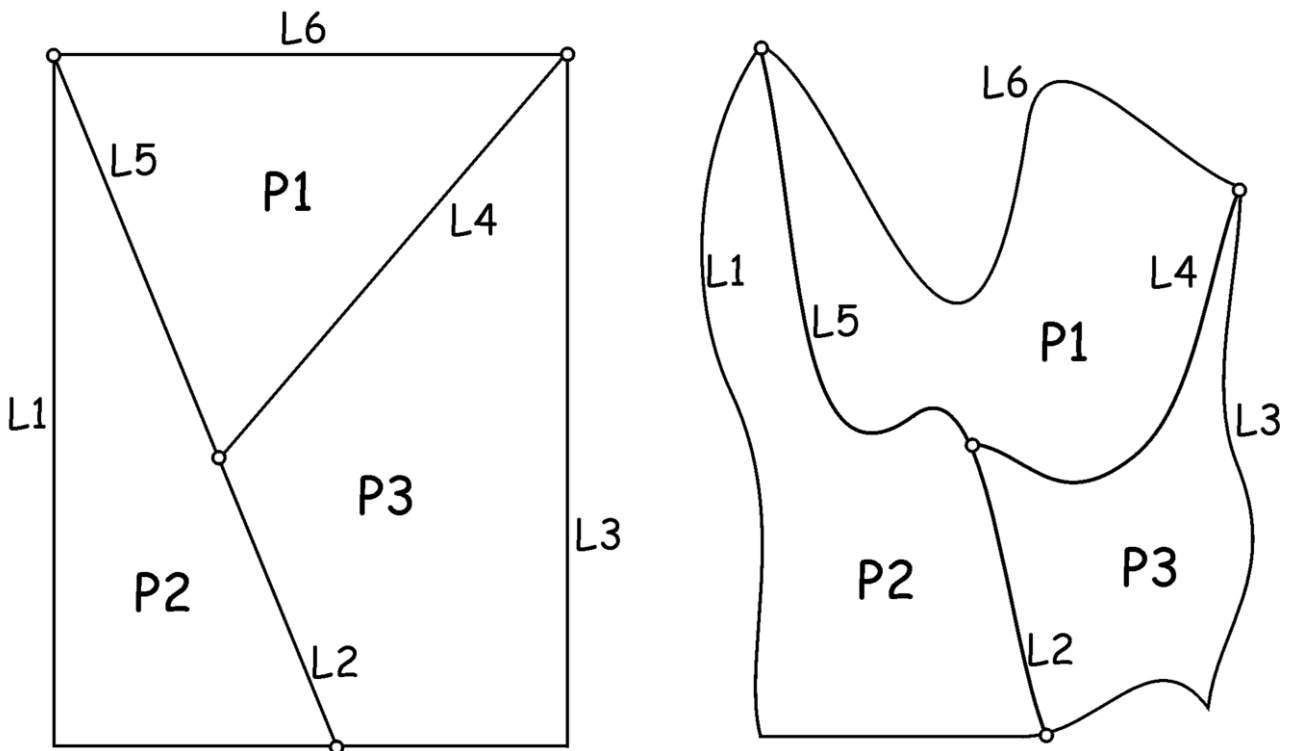


Figura 9: Relazioni topologiche di adiacenza fra poligoni

Per esempio una rete stradale può essere modificata metricamente accorciando determinati tratti tortuosi pur lasciando invariata la sua topologia. La stessa cosa può avvenire per il catasto: i confini possono essere tracciati in modo diverso, ma le informazioni topologiche non devono necessariamente per questo subire delle modifiche.

Oggetti monodimensionali nello spazio bidimensionale (un insieme di segmenti di linea collocati su un piano di riferimento): le relazioni topologiche stabiliscono anche qui quali segmenti si incontrano in quali nodi. Non esistono limitazioni al numero di segmenti che possono convergere in un unico nodo, essendo infinite le possibili direzioni di partenza da un punto nello spazio bidimensionale. Esempi di questi impianti possono essere le reti stradali, le condutture fognarie, i reticoli idrografici.

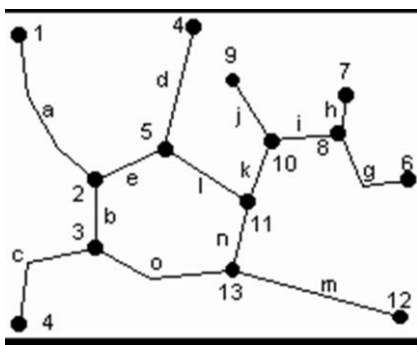


Figura 10: Topologia monodimensionale in spazio bidimensionale

Oggetti bidimensionali nello spazio bidimensionale (un insieme di superfici collocate su un piano di riferimento): le relazioni topologiche stabiliscono quali superfici si incontrano in quali segmenti di confine e, come illustrato al punto 2, quali segmenti di confine si incontrano in quali nodi (di confine). L'esempio più spesso citato nella letteratura del settore è quello di una carta dell'uso del suolo, dove l'intera superficie di riferimento (la carta, l'unità amministrativa descritta o altro) è coperta da un mosaico continuo di superfici. Ma anche impianti discontinui, come le carte degli edifici (delle superfici coperte), ricadono entro questa categoria, sebbene esista un'ampia regione dello spazio descritto che non è definita esplicitamente come poligono (questa regione viene spesso chiamata *poligono universo*).

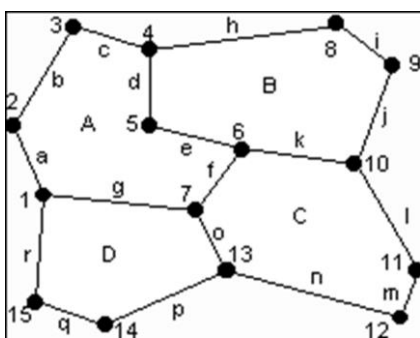


Figura 11: Topologia bidimensionale in spazio monodimensionale

Oltre a questi due tipi di impianti topologici, impiegati nei geodati, la fantasia (e la matematica) non pongono alcun limite logico a concepire oggetti e spazi con tre o più di tre dimensioni.

La topologia è quindi una materia complessa la cui trattazione esula dagli scopi di questo manuale. Dalla topologia derivano però alcune "regole topologiche" derivanti dalla topologia bidimensionale o planare, indispensabili per la costruzione di corretti modelli spaziali basati su dati vettoriali:

- ogni layer deve contenere primitive fra loro omogenee (solo punti, solo poligoni o solo linee);
- in un layer lineare:
 - ogni arco deve essere limitato da due nodi
 - gli archi si devono intersecare solo in corrispondenza dei nodi
- in un layer poligonale:
 - ogni arco limita due poligoni (left, right) ;
 - ogni poligono è costituito da un elemento chiuso fatto da una serie di nodi e archi;
 - due poligoni non possono sovrapporsi, o, in altre parole non può esistere una localizzazione definita da una coppia di coordinate, in cui sono presenti due o più poligoni;
 - attorno ad ogni nodo ci sono archi e poligoni.

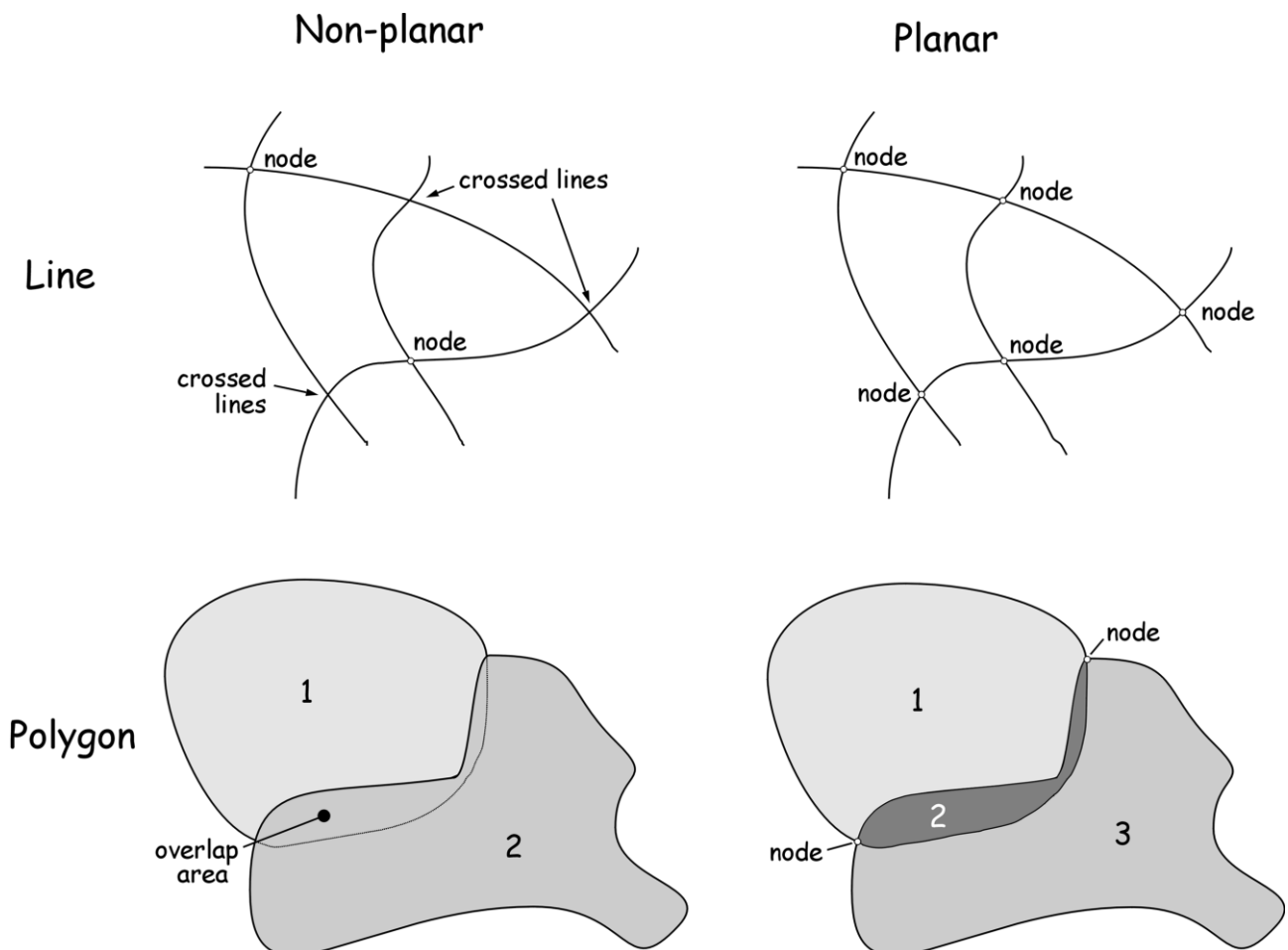


Figura 12: Dati topologici planari e non planari

La struttura raster

Tutti i dati raster sono definibili geometricamente come “tassellazioni” cioè una copertura di uno spazio per mezzo di figure ripetute (i tasselli) senza sovrapposizioni (senza cioè che due o più figure si accavallino l’una sull’altra) e senza lasciare spazi vuoti fra un tassello e l’altro. La tassellazione raster divide lo spazio geografico in una serie di tessere alle quali è associato il dato geografico. Il metodo è detto quantizzazione dello spazio (Kemp, 1993), cioè lo spazio viene diviso in unità elementari (quanta) che vengono successivamente analizzati ed elaborati. La forma delle unità elementari può essere molteplice (triangoli, esagoni, parallelogrammi), anche se la più usata è quella quadrata.

Considerando solo la tassellazione quadrata, il modello matematico che esprime un dato raster è perciò rappresentabile come una matrice di R righe per C colonne contenente $N = R \times C$ valori detti celle o *pixel*. L’elemento $x_{i,j}$ della i -esima riga e j -esima colonna contiene il valore del pixel corrispondente (attributo).

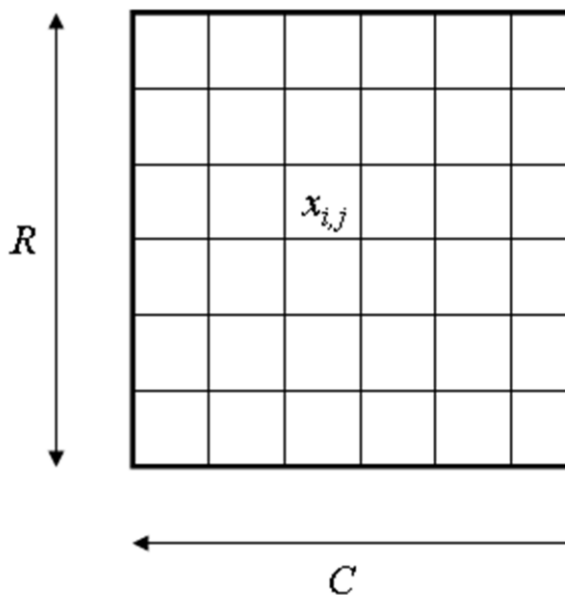


Figura 13: Forma matriciale del modello raster.

Come precedentemente illustrato, la georeferenziazione è la tecnica che permette di associare ad un dato, in formato digitale, una coppia di coordinate che ne fissino la posizione sulla superficie terrestre relativamente ad un sistema di riferimento geografico. Nella georeferenziazione del dato raster è quindi necessario definire univocamente la posizione di ciascuna cella elementare

riferendola ad un sistema di coordinate geografiche di tipo cartesiano. Un dato raster georeferenziato è detto mappa raster.

E' piuttosto facile dimostrare che per georeferenziare un raster con tassellazione quadrata sono sufficienti uno qualunque fra i seguenti insiemi di dati:

- $(X_{min}; X_{max}; Y_{min}; Y_{max}; L)$
- $(X_{min}; Y_{min}; N_c; N_r; L)$
- $(X_{min}; X_{max}; Y_{min}; Y_{max}; N_c; N_r)$

Con:

X_{min} coordinata minima in latitudine,
 X_{max} coordinata massima in latitudine
 Y_{min} coordinata minima in longitudine,
 Y_{max} coordinata massima in longitudine,
 L larghezza della cella,
 N_c numero colonne,
 N_r numero di righe (figura 14).

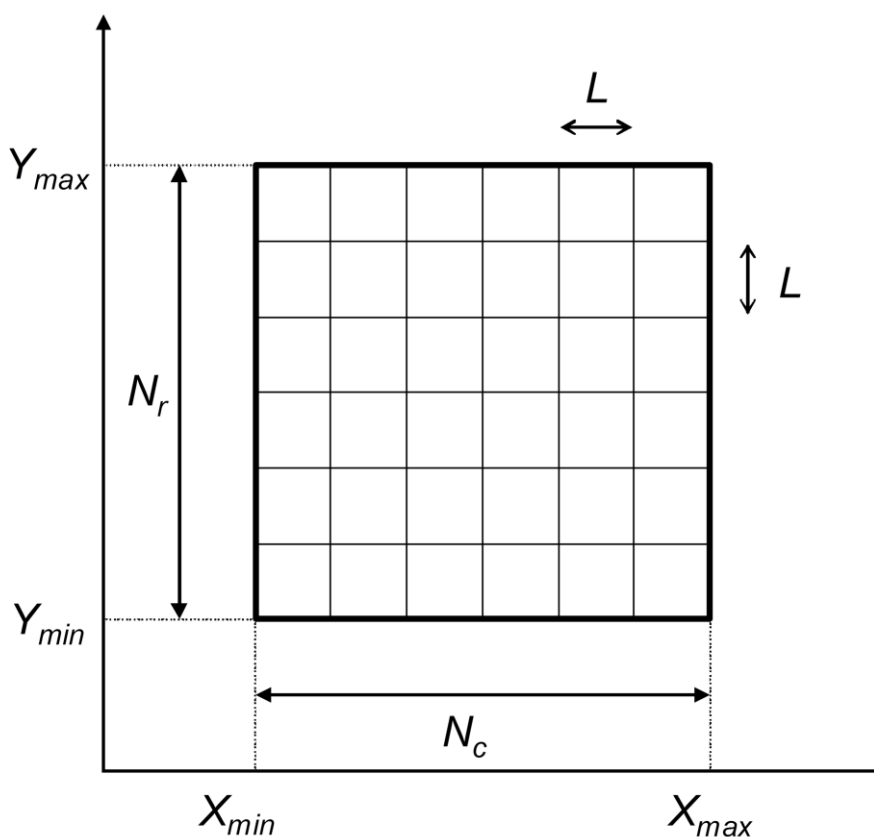


Figura 14: Georeferenziazione del dato raster

Come si evidenzierà nel prossimo capitolo, la scelta dei parametri di georeferenziazione di un dato raster è un passaggio fondamentale nella elaborazione spaziale. L'individuazione di un corretto spazio geografico deriva da due passaggi chiave:

- la definizione di una appropriata “finestra cartografica”;
- la scelta della corretta risoluzione.

Per finestra cartografica si intende la superficie terrestre coperta dal dato raster, così come definita dalle due coppie di coordinate (X_{min} ; Y_{min}) e (X_{max} ; Y_{max}). La sua ampiezza dipende strettamente dal tipo di analisi territoriale che è necessario effettuare e quindi da:

- la superficie da rappresentare: tipicamente nel caso di analisi riferite ad un ambito amministrativo questo sarà dato dai suoi confini;
- la distanza di influenza di fenomeni che interesseranno le elaborazioni da effettuare nel territorio: per esempio, nel caso dell'inquinamento la finestra cartografica dovrà essere allargata fino ad interessare la distanza massima di dispersione degli inquinanti, nel caso di una analisi di rete ecologica bisognerà considerare gli *home range* di spostamento degli animali e la diffusione dei disturbi antropici nello spazio.

La risoluzione di un raster georeferenziato è definita, geograficamente, dalla dimensione L della cella riportata sul terreno¹; per esempio 1 metro di risoluzione significa che riportata in dimensioni reali la cella occupa una superficie di 1 metro per 1 metro, 100 metri di risoluzione vuol dire che la cella rappresenta 1ha di territorio. La risoluzione influenza due caratteristiche importanti nell'analisi spaziale:

- l'accuratezza nel posizionamento e nella restituzione dell'informazione: tanto più la risoluzione è alta tanto più definito risulta il dato sia nell'errore connesso alla sua reale posizione nello spazio sia nel realismo della restituzione cartografica (figura 15).

¹ A livello topografico la risoluzione è spesso definita come punti per pollice (*dot per inch, dpi*).

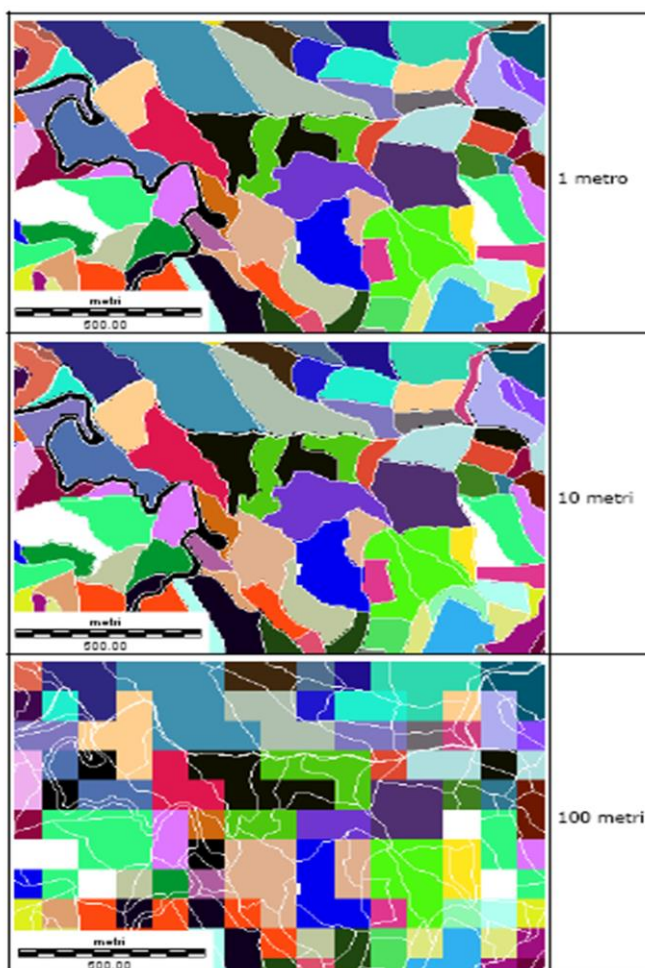


Figura 15: Effetto della diversa risoluzione

- la velocità di elaborazione: quanto più la risoluzione è alta tanto più lenta sarà l'elaborazione; a tale proposito è da considerare che la complessità di elaborazione cresce esponenzialmente con la risoluzione (almeno al quadrato) e può quindi facilmente portare all'impossibilità pratica nel risolvere il problema.

La scelta della risoluzione è quindi un passaggio delicato e deve tenere conto dei seguenti principi.

- Scala della restituzione cartografica: è da considerare che la letteratura oculistica mondiale concorda sul fatto che l'occhio umano, in condizioni ottimali di contrasto, può distinguere alla distanza minima di messa a fuoco (20-30 cm) fino ad un massimo di 10 linee per millimetro - cioè 250 dpi. La risoluzione spaziale dell'occhio umano diminuisce drasticamente all'aumentare della distanza (ad un metro di distanza siamo già sotto ai 75

dpi), e "precipita" in situazioni di scarso contrasto. Per tale motivo, per esempio, è inutile ad una scala 1:10.000 avere risoluzioni sotto il metro.

- Accuratezza/errore nella fonte dei dati: è concettualmente inutile, se non fuorviante, utilizzare risoluzioni più elevate della massima accuratezza posizionale del peggior dato raster che si usa nel modello spaziale. Se il modello spaziale impiega più fonti di dati, la risoluzione da adoperarsi è il massimo fra il dato vettoriale con il più alto errore di posizione (o la più bassa accuratezza) ed il dato telerilevato con la più bassa risoluzione.
- Applicazioni dei dati elaborati nei modelli spaziali e incertezza connessa con la costruzione del modello: per esempio, nel caso di modelli meteorologici è poco utile impiegare elevate risoluzioni, dal momento che i dati di partenza derivano dalla interpolazione di un numero limitato di stazioni a terra; nei modelli di dispersione degli inquinanti la distribuzione di questi a partire da una fonte puntiforme sul territorio è meramente probabilistica ed una elevata risoluzione non aggiunge maggiore certezza al dato.

Chirici, nelle sue "dispense di geomatica" (non pubblicato) propone le relazioni tra risoluzione di un dato e scala della cartografia tematica. Seguendo questo autore, la figura seguente mostra la relazione tra risoluzione di un dato e scala della cartografia tematica derivata (sull'asse delle X la scala è espressa come denominatore del rapporto di scala). In alto a sinistra il rapporto per tutte le scale fino a 1:1000000 a destra in modo più dettagliato fino alla scala 1:10000. In basso il rapporto in forma logaritmica. In tutti i grafici le linee individuano l'area in cui comunemente il rapporto scala/risoluzione ricade per progetti operativi di cartografia tematica.

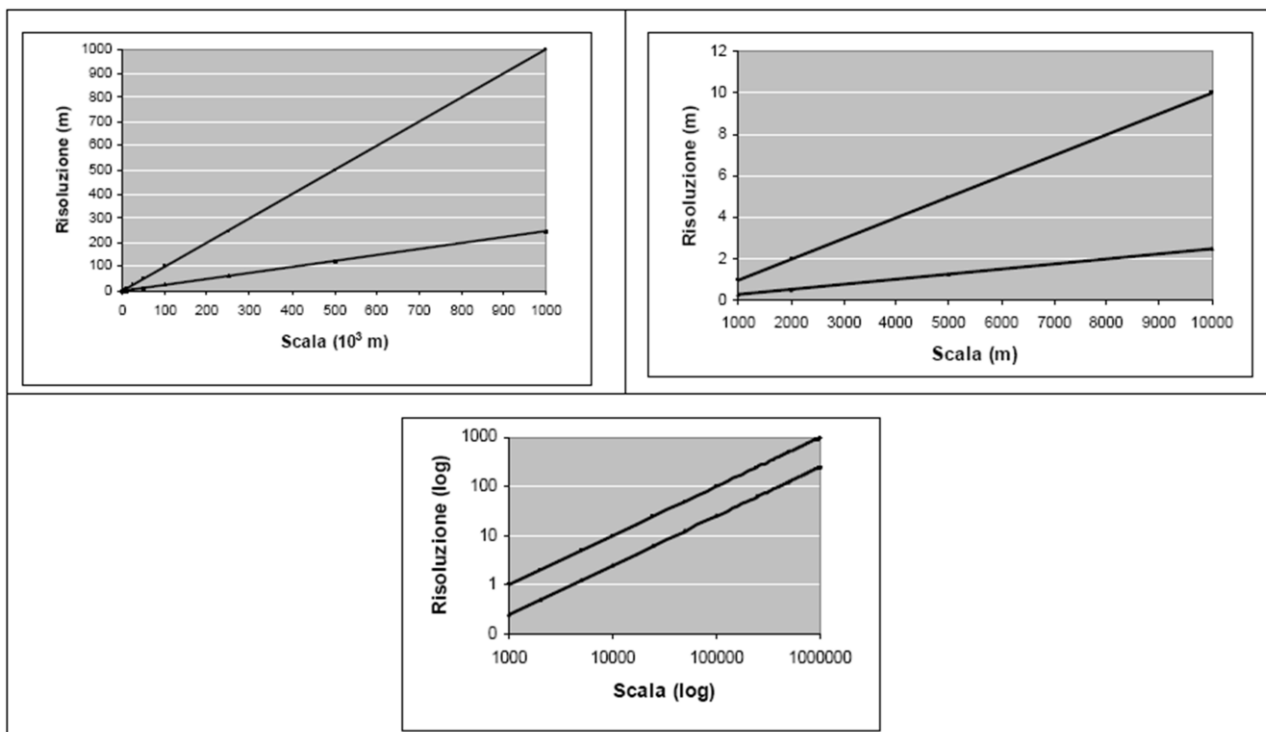


Figura 16: Relazione fra scala e risoluzione (tratto da Chirici)

Il ricampionamento

La modifica della risoluzione di un dato raster, operazione detta ricampionamento, fermi restando gli altri parametri di georeferenziazione è una operazione delicata perché comporta, come abbiamo visto, il cambiamento della scala di riferimento del dato stesso. Il ricampionamento, congruente con la regola generale già indicata, dovrebbe avvenire solo per diminuire la risoluzione del dato (ovvero per aumentare la grandezza del pixel). Il ricampionamento di senso opposto volto ad aumentare la risoluzione nominale del dato è invece una procedura assai complessa e necessita di informazioni e metodologie specifiche per ciascuna applicazione. Nel caso si diminuisca la risoluzione possono essere utilizzati due approcci. Entrambi analizzano la relazione esistente tra gli n pixel originari e il nuovo più grande pixel che deve essere generato.

- l'algoritmo di moda assegna al nuovo pixel generato il valore più frequente (la moda) tra quelli presenti negli n pixel originari.
- l'algoritmo di media assegna al nuovo pixel generato il valor medio tra quelli presenti negli n pixel originari.

Il primo è indicato per dati qualitativi, per esempio la generalizzazione di una cartografia di uso del suolo che deve passare da 250 a 1000 m di risoluzione: al nuovo pixel di 1000 m viene associato il codice di uso del suolo più frequente tra i sedici originari pixel di 250 m. Il secondo è indicato per dati quantitativi, per esempio la generalizzazione di una mappa delle temperature medie annue per passare da 250 a 1000 m di risoluzione. Al nuovo pixel di 1000 m viene associato un valore di temperatura calcolato come media tra i sedici originari pixel di 250 m.

E' importante notare che l'utilizzo dell'algoritmo di media per dati qualitativi è concettualmente sempre sbagliato mentre per i dati quantitativi è possibile scegliere l'algoritmo preferito sulla base delle esigenze di analisi.

La rasterizzazione dei dati vettoriali.

I database geografici vettoriali rappresentano un'importante fonte di dati del modello raster. Si pensi ad esempio ai geodatabase relativi all'uso del suolo, alle carte tecniche regionali, ai grafi stradali e idrografici. L'operazione di conversione fra il modello raster e quello vettoriale è semplice dal punto di vista computazionale, ma delicata per quanto riguarda la necessità di mantenere la coerenza fra i due formati. I due modelli sono infatti concettualmente diversi per le seguenti fondamentali caratteristiche:

- il formato vettoriale è caratterizzato da tre tipi di primitive cartografiche: punti, linee e poligoni, per i quali non si hanno equivalenti concettuali nel raster (figura 17);
- il formato vettoriale associa ad ogni primitiva un record di database; tale database non è presente nel formato raster.

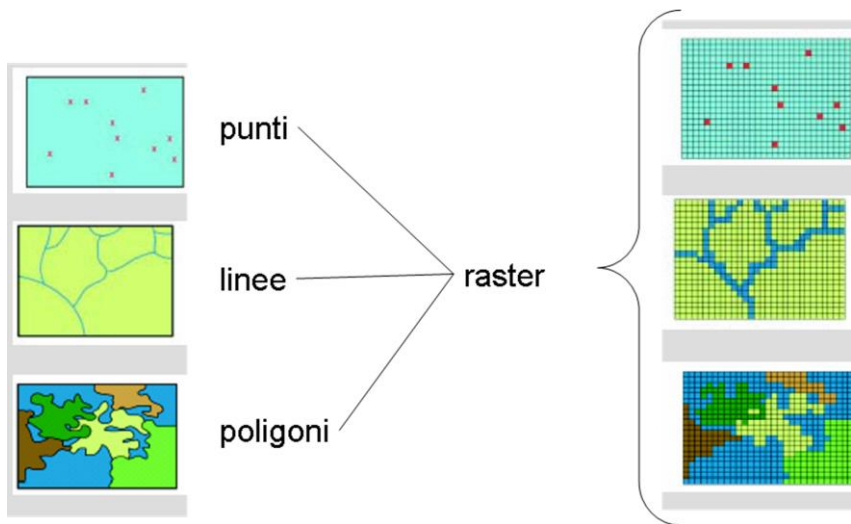
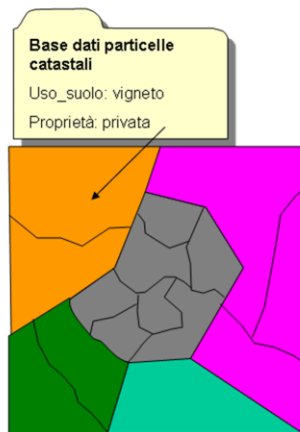


Figura 17: Rasterizzazione di dati vettoriali

Per quanto riguarda la prima differenza, è da considerare che, se la unicità di rappresentazione semplifica i problemi di tipo topologico, essa può generare alcuni inconvenienti da tenere ben presenti.

Innanzitutto ciascuna informazione nel formato raster genera una nuova griglia, le griglie divengono così dati indipendenti (figura 18) e perdono il riferimento concettuale ad uno specifico oggetto.

I dati Vettoriali associano a ciascun oggetto un record di database



Nel modello raster ciascun attributo rappresenta una griglia a parte

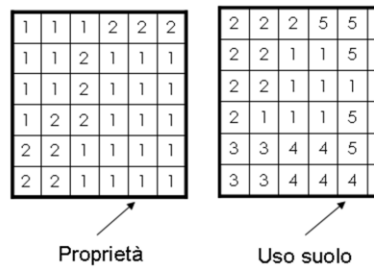


Figura 18: Differenze nell'organizzazione delle informazioni alfanumeriche fra modello vettoriale e modello raster.

Nel modello raster si perde la cosiddetta “dimensione topologica” delle primitive cartografiche vettoriali. I punti sono così rappresentati generalmente, ma non necessariamente, da celle isolate; le linee sono rappresentate generalmente, ma non necessariamente da sequenze concatenate di celle; i poligoni sono rappresentati generalmente, ma non necessariamente da gruppi di celle che condividono determinate caratteristiche (figura 19). Tale problema può essere risolto solamente documentando nell’ambito di un appropriato metadata l’origine del dato raster.

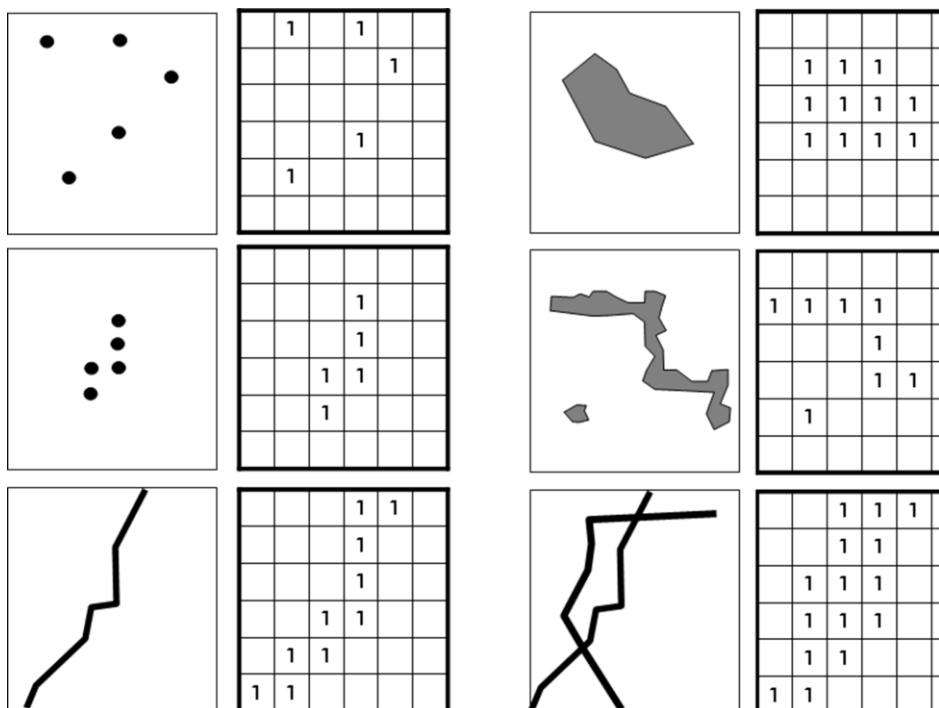


Figura 19: Indeterminazione spaziale del modello raster.

Molte volte nel passaggio da vettoriale a raster è possibile affrontare alcuni problemi di definizione del significato spaziale e di posizionamento del dato riorganizzando in modo diverso i dati con la creazione di strati informativi maggiormente omogenei e quindi coerenti. La figura 20 mostra una situazione problematica nella rasterizzazione di un grafo delle vie di comunicazione che può essere affrontata componendo il dato di partenza in entità più omogenee.

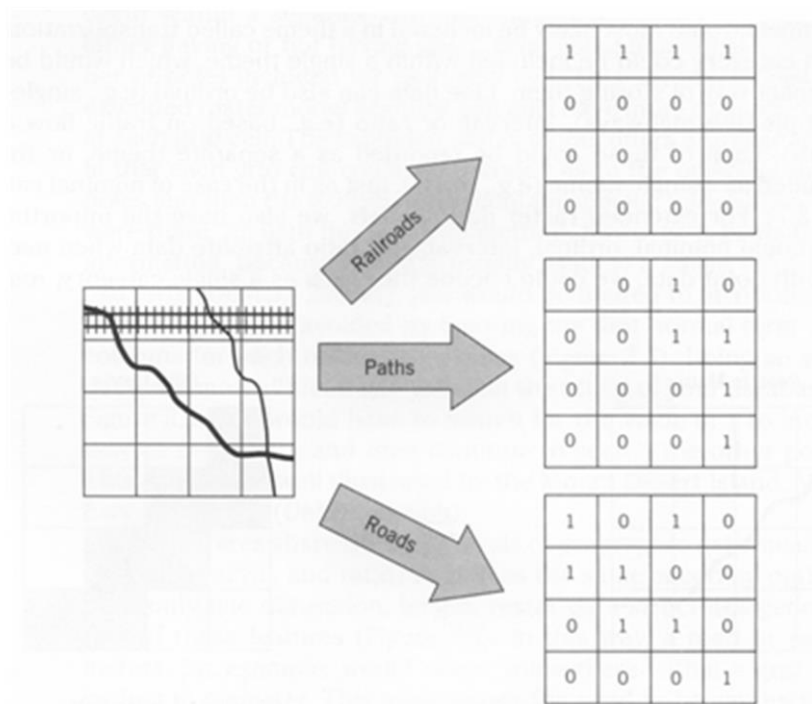
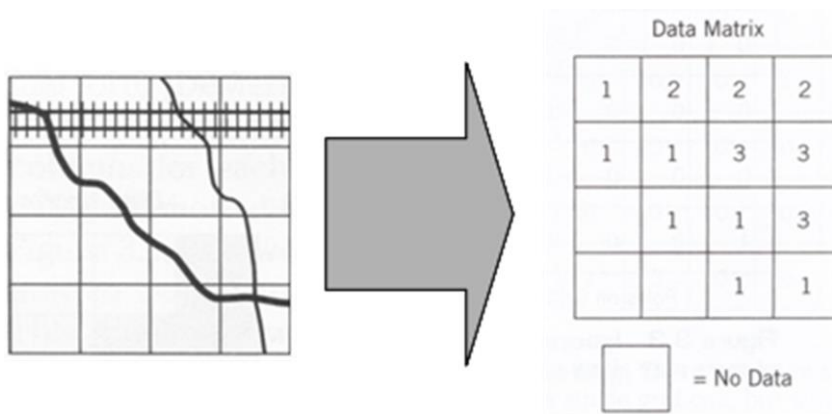
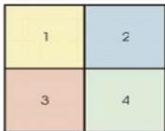
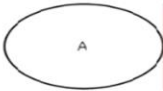

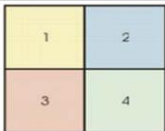
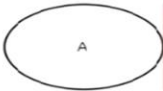
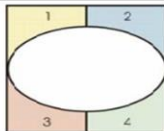
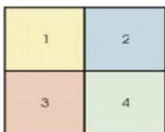
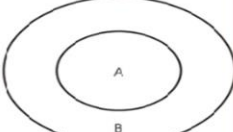

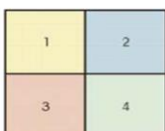
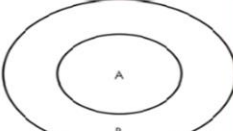
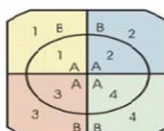
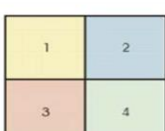
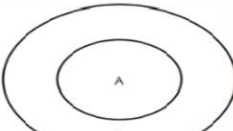




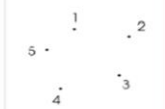



Figura 20: Problemi di rasterizzazione risolvibili con la scomposizione del dato vettoriale.

Modelli di analisi spaziale

Overlay di dati vettoriali

Le operazioni di analisi spaziale che verranno presentate di seguito sono applicate ad un modello (in genere vettoriale) di tipo topologico. Tali operazioni si riferiscono sia ai dati spaziali (primitive) che agli attributi ad essi associati. Le più importanti operazioni spaziali sono rappresentate nella figura seguente:

OPERAZIONE	INPUT	OVERLAY	OUTPUT
CLIP			
ERASE			
IDENTITY			
INTERSECT			
UNION			
MERGE			
BUFFER		-	

Queste operazioni, avvengono utilizzando due coperture una di input ed una di overlay, a parte che per l'operazione di buffer; nei primi quattro casi la copertura di overlay è sempre di poligoni

mentre quella di input può essere di qualsiasi tipo (punti, linee, poligoni). L'operazione di union può avvenire solo su coperture di poligoni. L'operazione di merge può avvenire tra coperture dello stesso tipo generando un nuovo strato informativo che ne rappresenta la fusione. L'operazione di buffer non è di sovrapposizione ma dato un input qualsiasi genera delle fasce di rispetto intorno alle primitive di input producendo in output una copertura di poligoni.

L'operazione di CLIP crea un nuovo livello informativo contenente tutte le informazioni della copertura di input che ricadono nella regione definita dalla copertura di ritaglio. Il nuovo livello informativo conterrà nel database tutti i dati attributo e parte delle primitive della copertura di input. Tali primitive possono essere punti, linee o poligoni, mentre la copertura di ritaglio deve essere necessariamente di poligoni.

L'operazione ERASE crea un nuovo livello informativo contenente tutte le informazioni della copertura di input che ricadono all'esterno della regione definita dalla copertura di cancellazione. Il nuovo livello informativo conterrà nel database tutti i campi attributo e parte delle primitive della copertura di input. Tali primitive possono essere punti, linee o poligoni, mentre la copertura di cancellazione deve essere necessariamente di poligoni. L'operazione di Erase può essere considerata come la complementare all'operazione di CLIP.

L'operazione di IDENTITY crea un nuovo tema dalla sovrapposizione di due strati informativi. In particolare il tema di output conterrà tutte le caratteristiche della copertura di input e per ciascuna di esse le caratteristiche dei poligoni della copertura di identità in cui esse ricadono. Il database del nuovo tema conterrà tutti i campi attributo e geografici sia della copertura di input che di quella di identità, in particolare a ciascuna primitiva di input verranno assegnati i relativi attributi del poligono in cui la primitiva ricade.

L'operazione INTERSECT crea un nuovo livello informativo dalla sovrapposizione di due temi, in particolare il tema di output conterrà solo le primitive che ricadono nell'area comune ad entrambi i livelli informativi; in particolare a ciascuna primitiva di input (poligoni, linee o punti) verranno assegnati i relativi attributi del poligono in cui la primitiva ricade.

L'operazione di UNION crea una nuovo livello informativo a partire da due coperture di poligoni, in particolare il tema di output conterrà le informazioni relative ad entrambe creando nuovi poligoni caratterizzati dagli attributi dell'una e/o dell'altra a seconda che si verifichi o meno sovrapposizione di primitive.

L'operazione di MERGE crea un nuovo livello informativo combinando le primitive del tema di input con quelle della copertura di overlay. Tale operazione a seconda del software utilizzato è possibile tra primitive dello stesso tipo o di tipo diverso, nel primo caso il database di output conterrà tutti i campi relativi ai due temi di input, nel secondo caso saranno accorpate solo le coordinate geografiche delle primitive senza alcun attributo.

L'operazione di BUFFER crea un nuovo tema nel quale intorno alle primitive viene generata una fascia di contorno di una specificata larghezza. Lo strato informativo di output sarà costituito dai poligoni involucro delle primitive selezionate. La dimensione delle fasce può essere o costante o variabile in funzione dei valori assunti da un campo attributo del tema di input.

Overlay di dati raster

Come si è potuto notare, la caratteristica principale di un dato raster è quella di essere costituito da una matrice di valori numerici organizzati per righe e colonne. In termini più specifici, un dato raster può essere considerato come un vettore bidimensionale di attributi del territorio, descritti tramite un valore numerico, la cui localizzazione a terra è implicitamente descritta dalla posizione del valore nelle righe e nelle colonne e dai parametri di georeferenziazione del dato. La sovrapposizione geografica di più dati spaziali è possibile a condizione che si possa ottenere la perfetta corrispondenza geografica di tutti i valori della matrice. In altre parole è necessario che i dati che vengono sovrapposti siano caratterizzati dalla stessa finestra cartografica ($(X_{min} ; Y_{min})$ e $(X_{max} ; Y_{max})$) e dallo stesso numero di righe e colonne (figura 22).

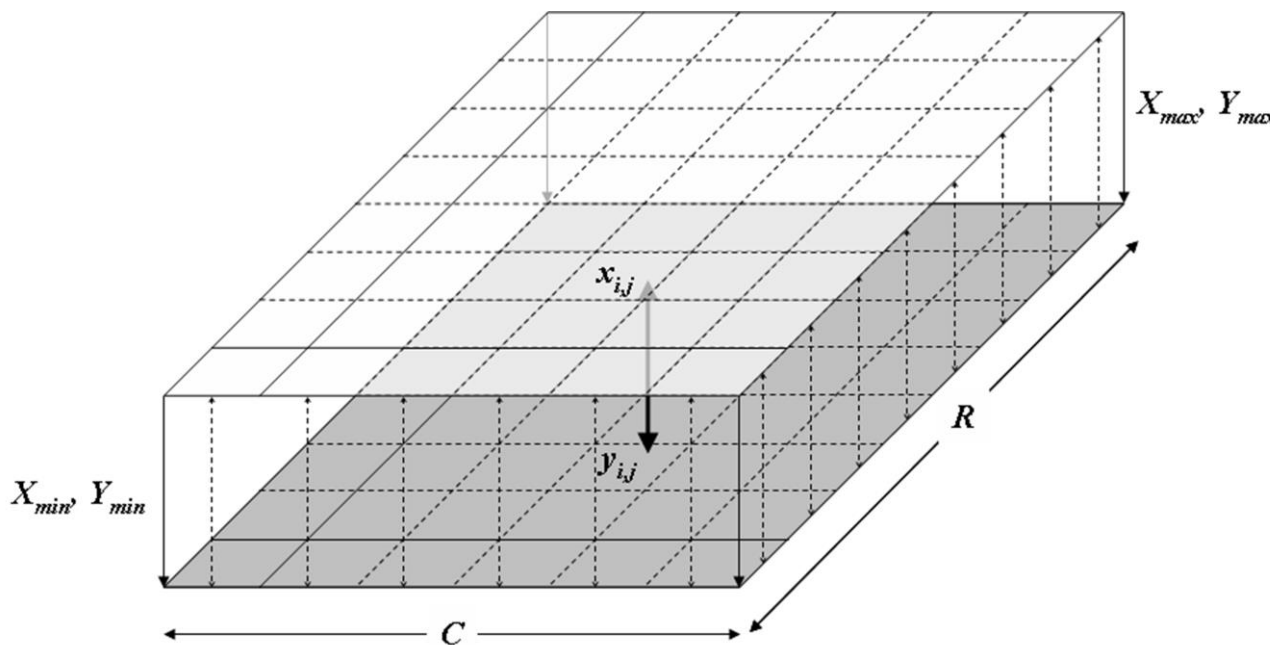


Figura 22: Condizioni di corrispondenza del dato raster.

Tale condizione è fondamentale per la realizzazione di elaborazioni spaziali in quanto permette la costruzione di relazioni e modelli alquanto complessi.

Map algebra e dati raster: gli operatori di Tomlin

Come lo stesso termine lascia intuire, per “*map algebra*” si intende l’uso di operatori elementari in sequenza allo scopo di risolvere problemi spaziali complessi. È sostanzialmente analoga ad

un'espressione algebrica in cui la combinazione di grandezze fornisce un risultato. Prevede l'impiego di espressioni di tipo logico e matematico applicate a dati spaziali.

La terminologia è stata formalizzata da Tomlin nel 1992, il quale ha definito gli operatori e le modalità con cui si poteva realizzare una struttura che si potesse definire di map algebra.

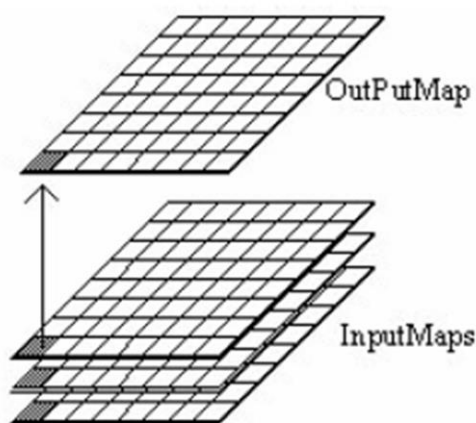
La definizione di map algebra è stata effettuata ed implementata in ambiente raster, dove le potenzialità di calcolo del modello forniscono i migliori risultati. L'applicazione di map algebra come schema è possibile anche in campo vettoriale, dove non è però ottimizzata.

Vediamo ora quali sono gli operatori principali di Map Algebra (**o operatori di Tomlin**).

Si definiscono quattro tipi di operatori:

- operatori locali
- operatori zonali
- operatori focali
- operatori globali

Operatori locali. Secondo la definizione di Tomlin, gli operatori locali consistono in operazioni su celle corrispondenti di uno o più dati raster di input (figura 23). Come si può vedere dall'esempio di figura 24, dove è applicato l'operatore "somma locale", i valori attribuiti alle singole celle esprimono la somma algebrica eseguita tra le corrispondenti celle delle mappe di input. Il discorso è analogo per qualsiasi altra operazione, non solo di tipo matematico: gli operatori possono infatti essere di tipo matematico (aritmetico, trigonometrico, statistico, logaritmico e così via), logico (vero o falso), di confronto (maggiore o minore). Gli impieghi degli operatori locali sono molteplici, dalla semplice conversione di unità di misura (operatore di moltiplicazione) alla attribuzione di valori ottenuti da operazioni algebriche o statistiche su serie di valori, ad operazioni di riclassificazione, come vedremo in seguito.



a. Local functions

Figura 23: Operatore locale di Tomlin

Operatore aritmetico

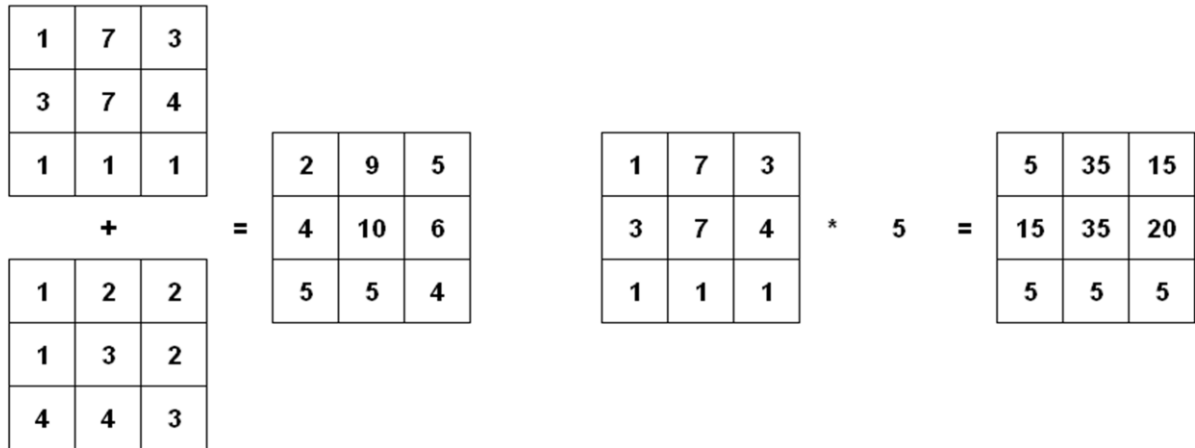


Figura 24: Esempio di operatore locale aritmetico

Operatore relazionale

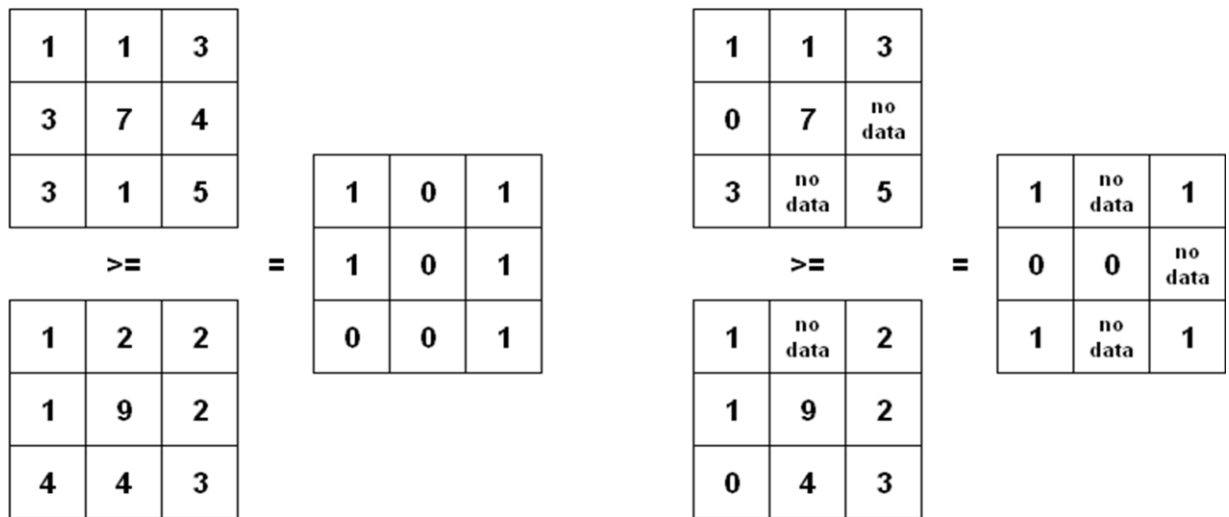


Figura 25: Operatori locali relazionali: la figura mostra anche la differenza concettuale fra il valore zero e l'assenza del dato (valore della cella = "no data").

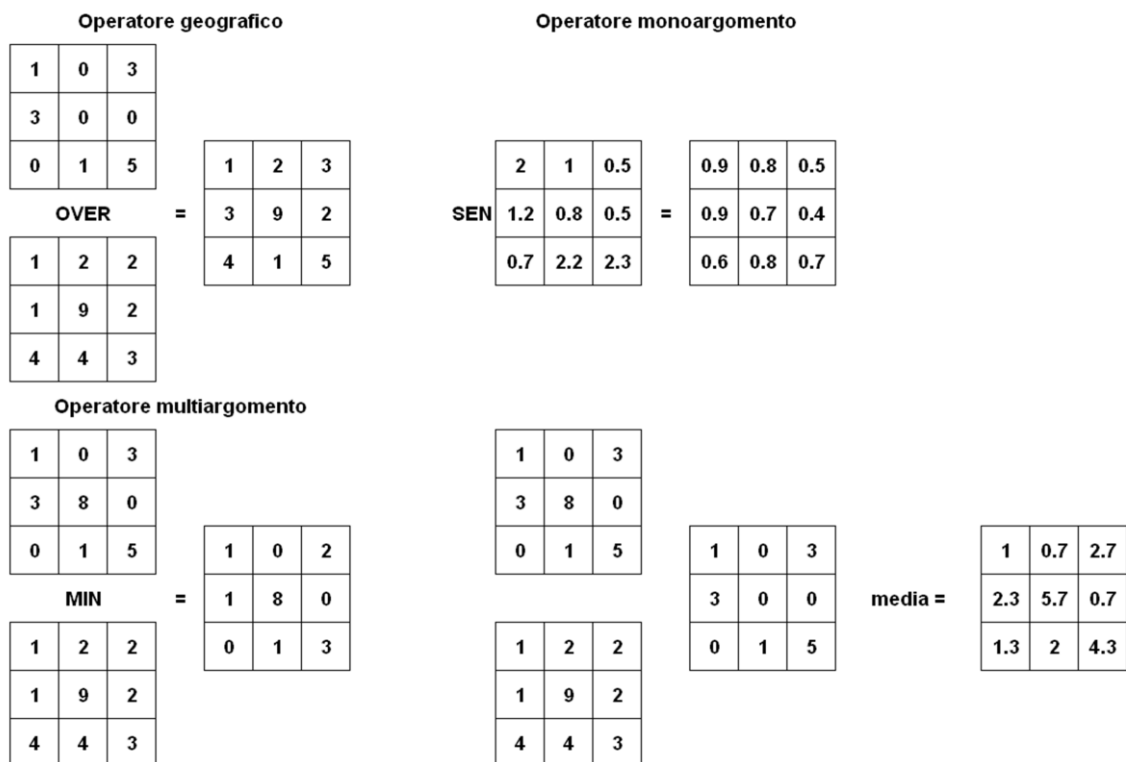
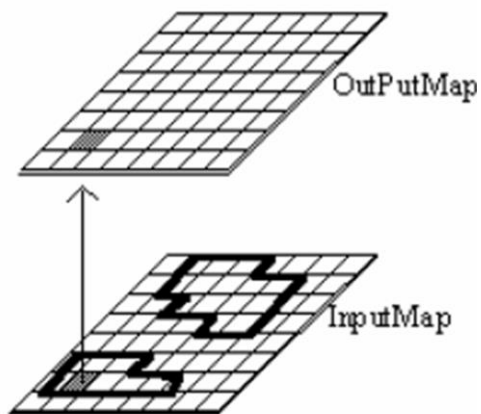


Figura 26: Altri operatori locali: l'operatore OVER restituisce i valori della prima griglia eccetto ove questa ha valore 0.

Operatori zonali: Attraverso l'applicazione di un operatore zonale, il valore delle celle del layer risultato è funzione dei valori delle celle di un layer di input, raggruppate secondo zone definite in un altro layer di input (figura 27). Un esempio di problema risolvibile con operatori zonali è il seguente: *data una mappa dei centri urbani e un modello di elevazione del terreno calcolare la altitudine media di ogni città.*



c. Zonal functions

Figura 27: Operatore Zonale di Tomlin

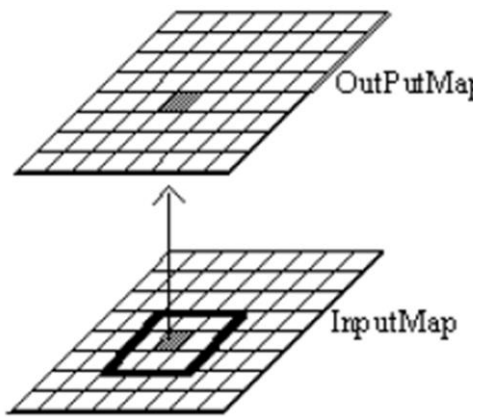
Come si vede in figura 28, nel primo layer compaiono delle celle a cui sono attribuiti determinati valori. Nel secondo layer di input vengono invece definite le zone. Nel primo layer di input vengono sommate tra loro solamente le celle appartenenti ad una stessa zona, così come definita dal secondo layer. Se si osservano le celle appartenenti alla zona 2, si vede che la loro somma algebrica è pari al valore che compare in tutte le celle della zona 2 nella mappa risultato. Oltre alla somma zonale, sono operazioni zonali tipiche il calcolo della media, in cui a tutte le celle di una zona viene attribuito il valore della media dei valori delle singole celle, e il calcolo del valore minimo e massimo o di un range di valori.



Figura 28: Esempio di operatori zonali

Operatori focali: Le funzioni focali si applicano ad un intorno di un singolo dato raster (in altre parole abbiamo un solo layer di input; fig. 29). Il risultato dipende dal valore della cella in esame e dai valori delle celle poste nelle righe e colonne precedenti e seguenti; in pratica viene fatta scorrere sulla matrice dei dati una “finestra mobile” (detta anche *kernel*) che definisce i pixel che verranno elaborati. Il caso più semplice è quello di una finestra mobile di 3 celle per 3 celle, che viene fatta sistematicamente scorrere sulla matrice dei dati, come mostrato in figura 30.

La funzione focale, oltre che dalla forma e dalla dimensione della finestra mobile è definita anche dal tipo di operatore, ad esempio operazioni di media, deviazione standard, range, somma, ecc.



b. Focal functions

Figura 29: Operatore focale di Tomlin

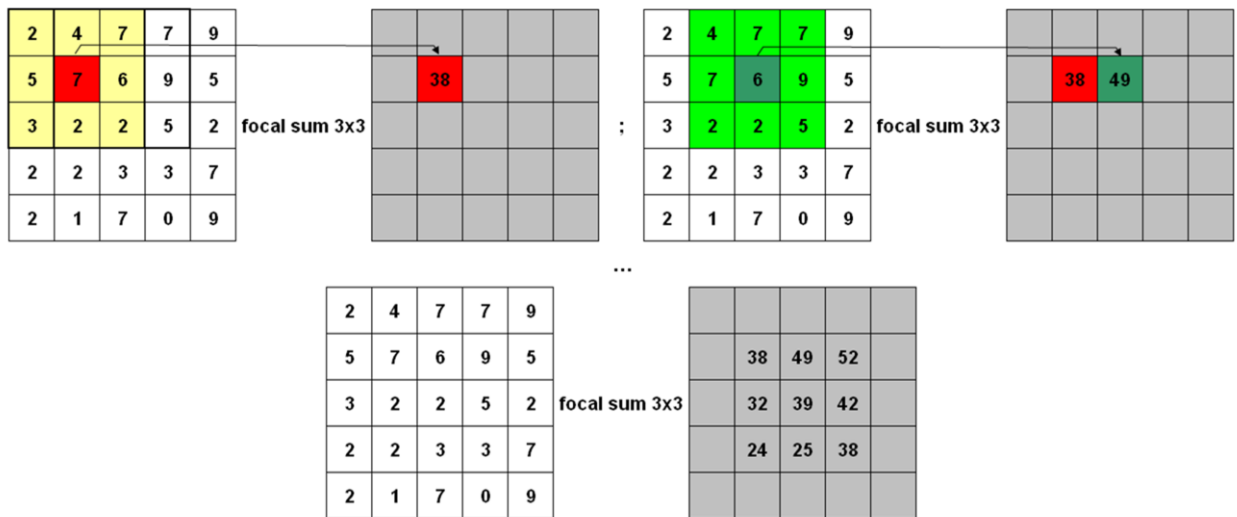


Figura 30: Esempio di somma focale

Negli esempi riportati in figura l'intorno scelto è di forma quadrata, ma sono possibili altre forme, come quella circolare. Le funzioni focali trovano importanti estensioni e applicazioni nei filtri, di cui ci occuperemo in uno specifico capitolo.

Operatori globali: Gli operatori globali vengono utilizzati per determinare dati raster di output in cui il valore delle celle dipende potenzialmente dai valori di tutti gli elementi del dato raster di input. Un esempio di tale analisi è la determinazione del percorso a costo minimo tra una cella di origine e una cella di destinazione. Questa operazione comporta il calcolo di tutti i possibili percorsi al fine di individuare quello a costo minimo. Per poter eseguire questo calcolo è evidentemente necessario conoscere il valore di tutte le celle del dato raster di input: per questo motivo tali operatori vengono definiti globali.

Diagrammi di flusso

I diversi operatori di Map Algebra possono venir combinati in sequenza per risolvere problemi spaziali attraverso modelli più o meno complessi. A tal proposito è opportuno seguire un procedimento suddivisibile in diverse fasi:

- individuazione e descrizione del problema cartografico
- scomposizione del problema nei suoi elementi e descrizione attraverso un flow-chart, del procedimento che ci porta alla risoluzione del problema, fino a individuare le basi dati necessarie;
- verifica della disponibilità dei dati necessari e della loro adeguatezza, in caso contrario verificare se è possibile definire il problema con dati diversi;
- documentazione del procedimento.
- espressione del flow-chart nel linguaggio formale di Map Algebra del sistema in uso
- validazione dei risultati ed eventuali modifiche al modello

Il modo più corretto ed efficiente per concettualizzare un modello è attraverso un procedimento che dallo scopo porta ai dati elementari, cioè quello che viene definita *backward inference*.

Il primo passo è definire nel modo più univoco possibile lo scopo della valutazione:

Per esempio:

individuare le aree idonee per la localizzazione di discariche di rifiuti

Il problema viene poi successivamente scomposto nei suoi elementi

La discarica deve essere situata lontano da centri residenziali

Il substrato geologico deve essere impermeabile

La discarica non deve essere visibile da luoghi di interesse turistico

Il costo di realizzazione deve essere più basso possibile

Gli elementi del problema vengono poi ulteriormente specificati:

La discarica deve essere situata lontano da centri residenziali

La distanza minima deve essere di almeno ...

Il substrato geologico deve essere impermeabile

Sono idonei i seguenti substrati geologici ...

La discarica non deve essere visibile da luoghi di interesse turistico

Le zone da vincolare sono le seguenti:

Visibili da strade turistiche

Visibili da borghi storici

Il costo di realizzazione deve essere più basso possibile

Il costo dipende dai seguenti fattori:

Uso del suolo del terreno da impiegare

Collegamento stradale: a sua volta dipende da

Distanza da strade

Pendenza delle aree attraversate

Infine si giunge ad identificare le basi dati necessarie per l'elaborazione. Nell'esempio:

Mappa dei centri residenziali

Carta geologica

Grafo delle strade turistiche

Localizzazione dei borghi storici

DEM per calcolare l'intervisibilità e la pendenza del terreno

Mappa dell'uso del suolo

Grafi stradale

Successivamente si verifica la disponibilità e l'adeguatezza delle basi dati da impiegare (eventualmente modificando il modello individuato) e si specificano i parametri di georeferenziazione (finestra cartografica e risoluzione) che saranno validi per tutte le basi dati impiegate.

Il passaggio successivo consiste nella costruzione del diagramma di flusso che dovrebbe contenere le basi dati iniziali, finali ed intermedie nonché gli operatori impiegati nella elaborazione, così come mostrato in figura 31.

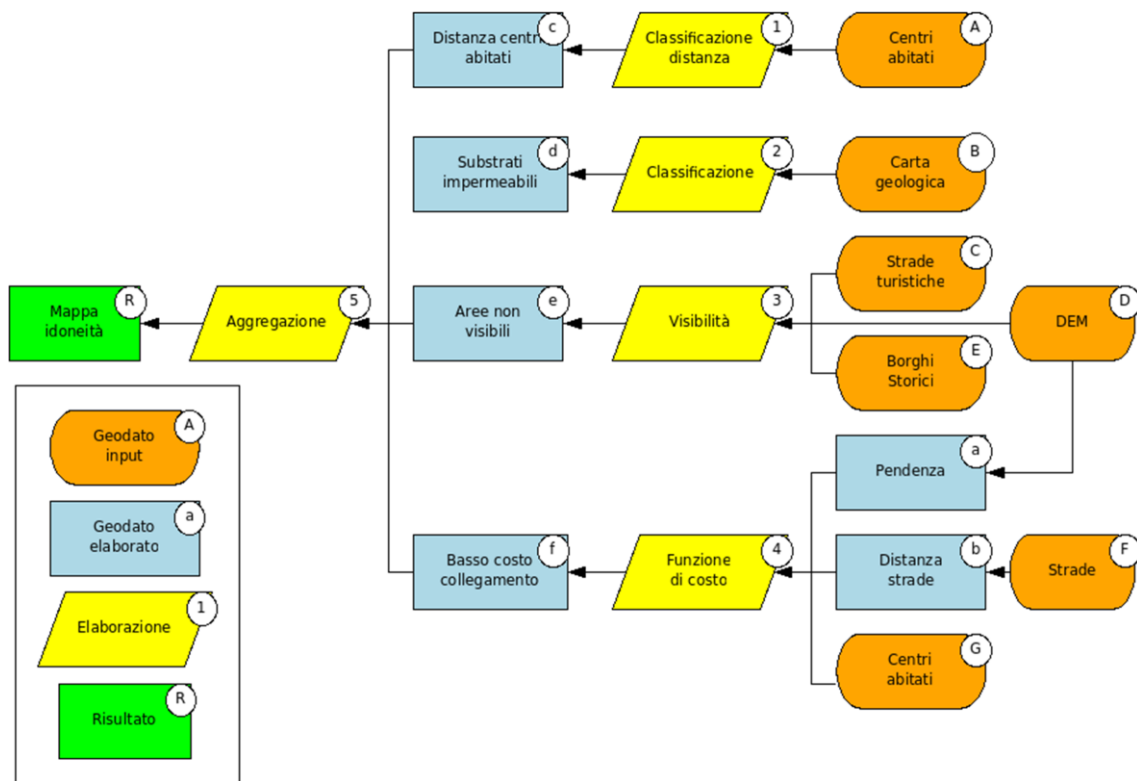


Figura 31: Esempio di diagramma di flusso

È importante inoltre completare la costruzione del modello con una documentazione il più possibile esaustiva di tutto il procedimento adottato, numerando le basi dati iniziali, le basi dati intermedie e gli operatori e documentando, con schemi dettagliati come quelli mostrati in figura 32, tutto il procedimento seguito.

Parametri della elaborazione:

Scopo dell'analisi:

Parametri di georeferenziazione:
 X_{min} Y_{min}
 X_{max} Y_{max}
Risoluzione

Tabella basi dati		
Nome file	Unità misura	Descrizione e/o legenda

Tabella operatori	
Numero	Operatore, Parametri ed eventuale grafico

Figura 32: Esempio di documentazione di un modello di analisi spaziale

Mappe concettuali e reti semantiche

Le **mappe concettuali** sono uno strumento grafico per rappresentare informazione e conoscenza, teorizzato da Joseph Novak, negli anni '70. Secondo l'impostazione originaria di Novak e Gowin, le caratteristiche essenziali di una mappa concettuale sono le seguenti:

- la logica d'insieme è connessionista,
- la mappa si sviluppa dall'alto verso il basso, a partire da un concetto iniziale,
- la struttura generale delle connessioni è di tipo inclusivo, dal generale al particolare, dal sovraordinato al sottordinato,
- oltre a quelle di inclusione, tra i concetti sono possibili anche relazioni trasversali.
- le relazioni tra i nodi sono orientate e soprattutto devono essere esplicitate, in linea di massima attraverso indicatori forti, quali verbi e connettivi sintattici,
- la relazione tra due concetti costituisce una proposizione,
- l'insieme della mappa ha una struttura di significato che è data da concetti, relazioni, proposizioni.
- se le relazioni non sono esplicitate con un'etichetta precisa, la mappa perde di significato, fino a non averne del tutto;
- l'impostazione grafica deve essere omogenea e lineare.

Dagli anni '70 ad oggi questo insieme di requisiti è stato sostanzialmente indebolito e reso più flessibile. La perdita di rigore nella definizione dello strumento si è accompagnata ad una crescente flessibilità operativa. Anche il significato della espressione "mappa concettuale" si è parallelamente evoluto nel tempo. Strettamente parlando, la nozione di mappa concettuale riguarda uno strumento grafico, e va distinta dalle nozioni di mappa mentale e di mappa cognitiva (cognitive map), con i quali si fa riferimento alle strutture mentali elaborate dai soggetti per immagazzinare la conoscenza. Però anche l'uso di questi termini non è del tutto definito e codificato; in particolare, il termine "mappa mentale" si è esteso ad indicare **strumenti grafici di rappresentazione della conoscenza**. In breve, le mappe concettuali vengono oggi comunemente considerate, al pari delle mappe mentali, **un mezzo per la rappresentazione delle strutture cognitive**.

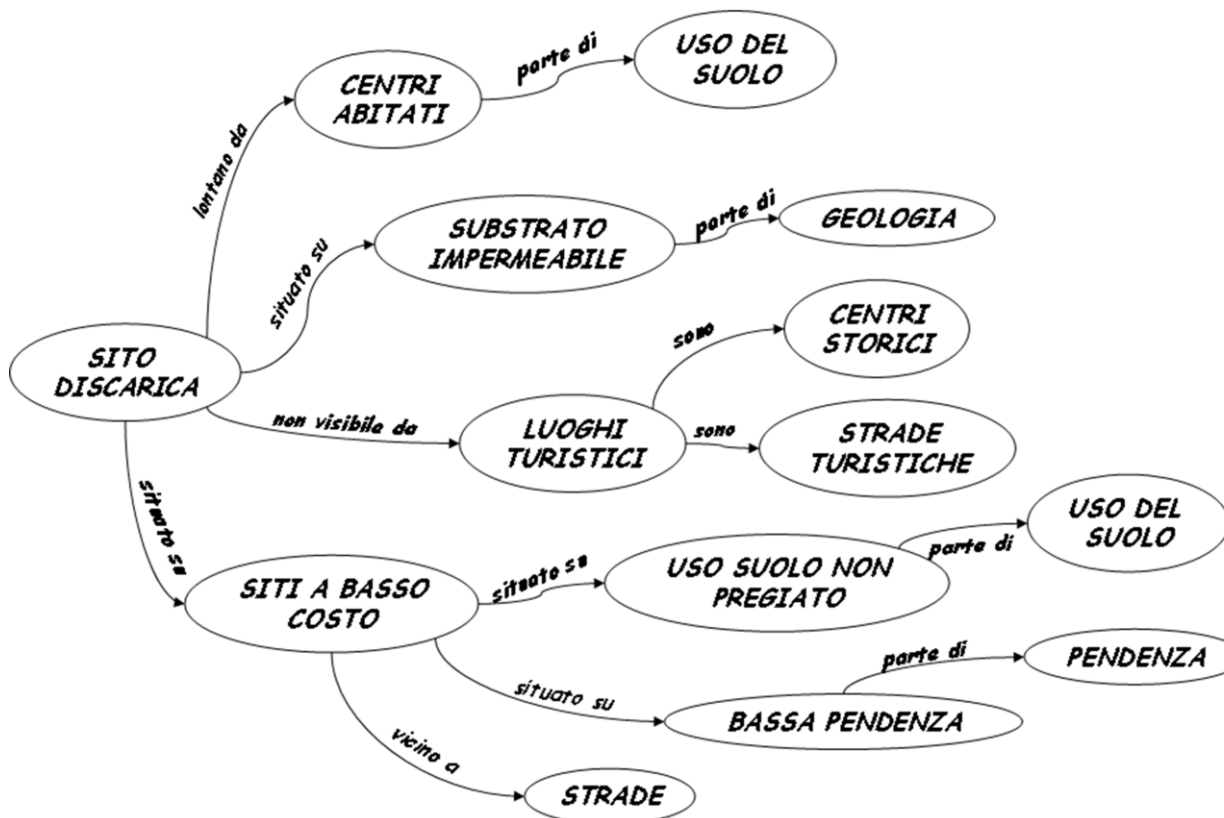
Le mappe concettuali servono per rappresentare in un grafico le proprie conoscenze intorno ad un argomento secondo un principio cognitivo di tipo costruttivista, per cui ciascuno è autore del proprio percorso conoscitivo all'interno di un contesto, e mirano a contribuire alla realizzazione di apprendimento significativo, in grado cioè di modificare davvero le strutture cognitive del soggetto e contrapposto all'apprendimento meccanico, che si fonda sull'acquisizione mnemonica.

Una mappa concettuale è una rappresentazione grafica della conoscenza, un disegno schematico, un quadro riassuntivo di un ragionamento che abbiamo fatto. All'interno della mappa la conoscenza viene rappresentata attraverso due elementi fondamentali:











- i concetti, individuati con parole-chiave e inserite in forme geometriche: i nodi.
- le relazioni, collegamenti tra i diversi concetti che evidenziano la sovraordinazione, la sottoordinazione e la coordinazione, in altre parole la gerarchia, espressi attraverso linee o frecce: gli archi.

Le relazioni si possono nominare per mezzo di parole-legame sistemate sugli archi. Un nodo può rappresentare un punto di partenza o rappresentare l'argomento da affrontare; quindi si individuano i concetti pertinenti e tra questi viene stabilito un ordine di relazioni gerarchiche (diagramma ad albero). Questa operazione genera una cosiddetta **rete semantica**, nel senso che l'arco associato ad una etichetta permette di definire "il tipo" di relazione che si stabilisce fra i nodi.

Una mappa concettuale quindi rappresenta una formalizzazione più libera e informale di un diagramma di flusso e può quindi essere molto utile a concettualizzare un problema al fine di esprimerlo successivamente in termini più rigorosi, come si può riscontrare dalla figura seguente che riporta la mappa concettuale del problema di localizzazione della discarica.



Il passaggio da mappa concettuale a diagramma di flusso consiste nel convertire le relazioni in operazioni spaziali. Per esempio il termine semantico "situato su" corrisponde alla operazione spaziale di overlay, "parte di" corrisponde ad una selezione. Nella tabella seguente è riportata una possibile chiave di conversione fra termini semantici e operatori cartografici (modificato da Albrecht, 1996).

Termini semantico	Operazione spaziale	operatori
“è parte di”	Selezione	 <i>Interpolation</i>  <i>Thematic Search</i>  <i>Spatial Search</i>  <i>(Re-)classification</i>
“situato su”	overlay	 <i>Overlay</i>
“presso” “vicino a” “lontano da”	distanza	 <i>Cost/Diffusion/Spread</i>  <i>Proximity</i>  <i>Nearest Neighbor</i>  <i>Buffer</i>
“visibile”	analisi 3D	 <i>Viewshed Analysis</i>

Chiave di conversione fra termini semantici e operatori cartografici (modificato da Albrecht, 1996).

Bibliografia

J.D. Novak e D.B. Gowin, *Imparando a imparare*, S.E.I., Torino, 1989

Tomlin C.D. (1990). *Geographic information system and cartographic modelling*. Englewood Cliff, Prentice-Hall, New York.

Albrecht, J., 1995b. Semantic Net of Universal Elementary GIS Functions. *Proceedings CSM/ASPRS Annual Convention and Exposition Technical Papers*, Vol. 4 (Auto-Carto 12), pp. 235-244. Bethesda, MD.

Appendice²

Introduzione ai raster GIS

La maggior parte degli ambienti informatici dedicati alla elaborazione di dati raster è formata da un'interfaccia di visualizzazione e gestione files e da un insieme di programmi separati, detti moduli, finalizzati alla elaborazione dei dati. I GIS raster hanno (prevalentemente) le seguenti caratteristiche comuni che li rendono piuttosto differenti dai programmi dedicati alla elaborazione di dati vettoriali:

- ogni elaborazione genera una nuova base dati raster, con conseguente occupazione di memoria; solo raramente l'elaborazione modifica la base dati esistente
- la base dati generata viene immediatamente salvata sul disco, appena calcolata; alcune volte tale salvataggio avviene in una cartella temporanea;
- non sempre la base dati generata viene visualizzata sullo schermo;

Quasi tutti gli ambienti raster GIS sono attualmente in grado di operare con dati vettoriali, se non altro per permettere l'importazione e l'esportazione. Molti ambienti GIS vettoriali sono in grado di operare con dati raster in modo più o meno evoluto; in quel caso per i dati raster valgono, generalmente, le caratteristiche precedentemente elencate.

Nei prossimi paragrafi saranno descritte le procedure di definizione dei parametri del modello raster e di map-algebra per due ambienti piuttosto diffusi nell'ambito della elaborazione raster: quantum-gis-grass ed idrisi.

Grass e QGIS

GRASS è senz'altro il riferimento per le analisi territoriali in ambiente Open Source in quanto permette di effettuare analisi raster evolute. Quantum GIS (QGIS), nato nel giugno 2002 costituisce un ambiente GIS interoperativo con GRASS in grado di superare le difficoltà di utilizzo derivanti da una struttura complessa, pensata per analisi approfondite.

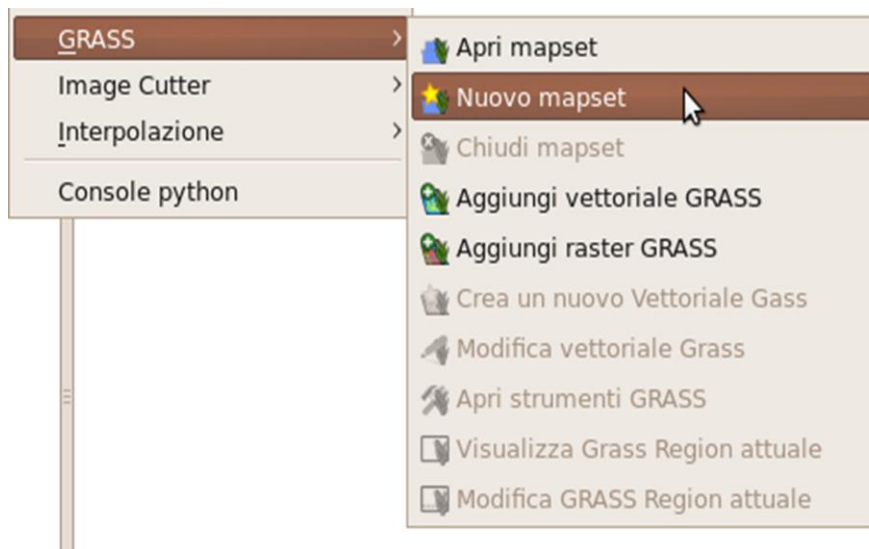
L'organizzazione dei dati: il mapset.

I dati di GRASS sono strutturati in cartelle con una particolare struttura detta di GISDATABASE. La cartella di lavoro deve essere creata prima del lancio del programma GRASS. All'interno della cartella di lavoro i dati sono organizzati per progetti di lavoro e immagazzinati in cartelle chiamate LOCATION. Ogni location è definita dal suo sistema di coordinate, proiezione e finestra cartografica. La struttura delle cartelle e dei dati che definiscono una location sono creati

² L'interfaccia grafica legata ai diversi software GIS può variare a seconda della versione.

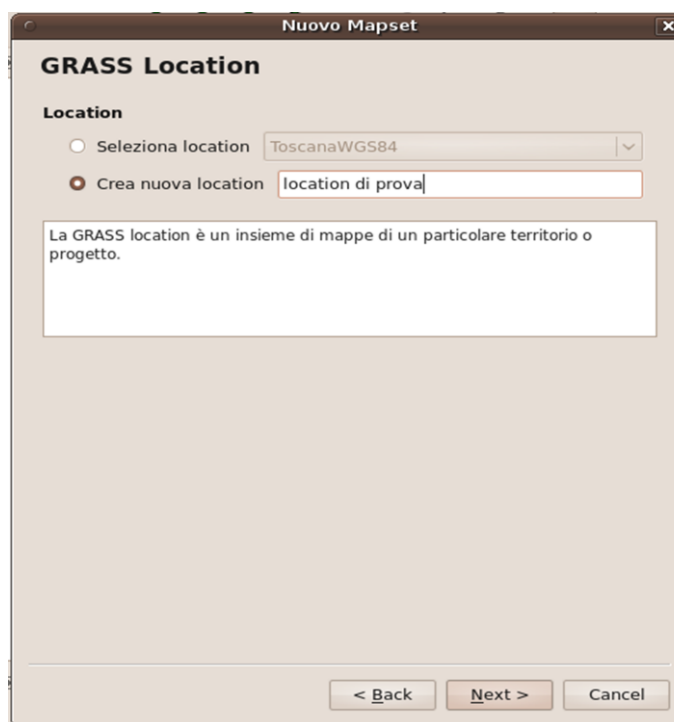
automaticamente dal programma. Ogni location è ulteriormente suddivisa in MAPSET che hanno lo scopo di specificare, per ciascun progetto, cartelle contenenti sottoprogetti. Per esempio, la cartella GRASS costituisce un GISDATABASE; questa può essere suddivisa in due location, definite attraverso la loro proiezione geografica: circondario GBW e ToscanaWGS84; all'interno delle due location si possono avere i mapset "obbligatori" PERMANENT accessibili da qualsiasi utente senza specificare la location e altri mapset definiti dall'utente.

La procedura per creare una location all'interno di QGIS è la seguente. Nel menu Plugin selezionare GRASS e Nuovo mapset.

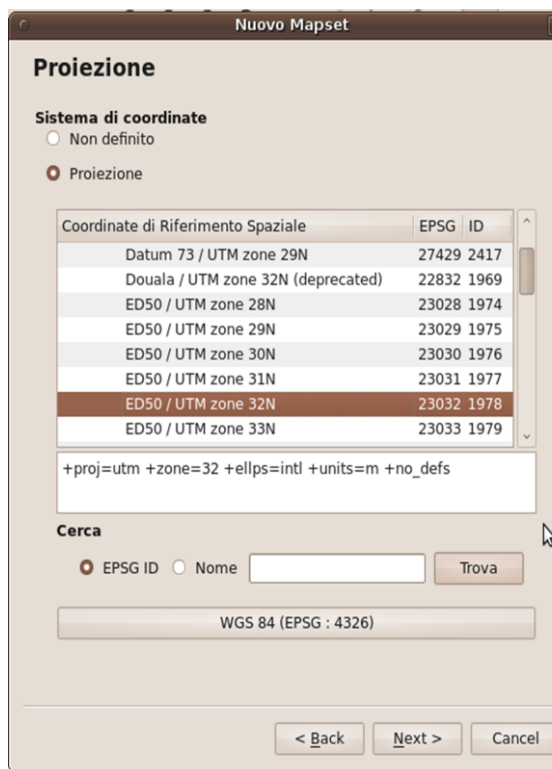


Successivamente specificare la cartella di GRASS e la location.

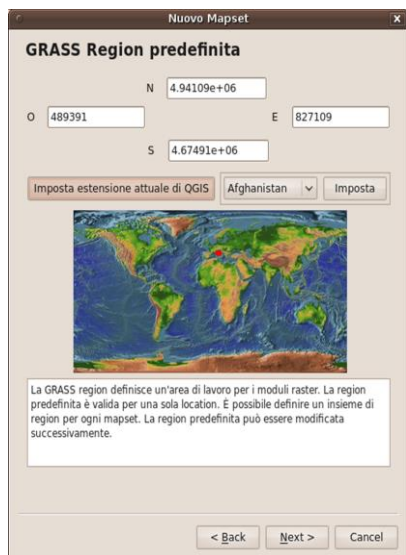


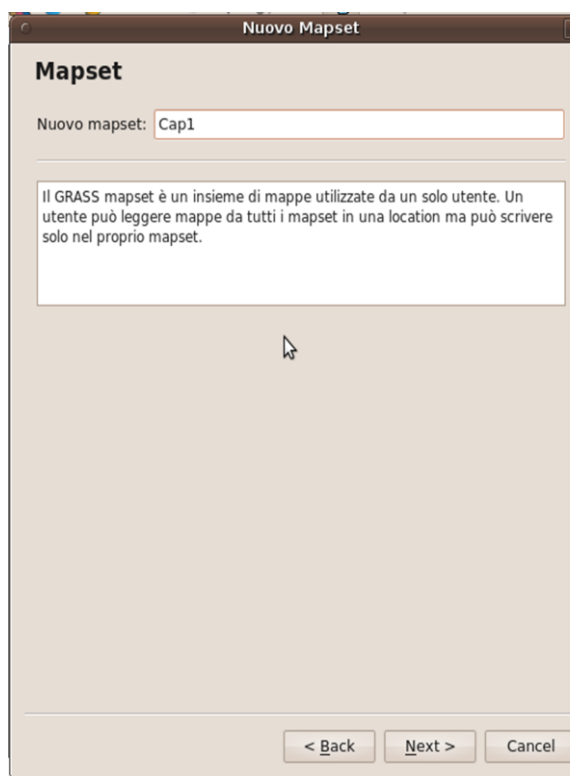


Per definire la proiezione è possibile effettuare una ricerca per nome. In alternativa è possibile specificare il codice del sistema di coordinate.



Infine è necessario specificare i limiti della finestra cartografica come mostrato e il mapset.





Quando viene creata una nuova location, QGIS-GRASS crea automaticamente uno speciale mapset, chiamato PERMANENT finalizzato ad immagazzinare i dati di base della proiezione e del progetto. Le operazioni di gestione delle mappe all'interno di una location debbono essere effettuate all'interno dell'ambiente QGIS-GRASS, in quanto coinvolgono operazioni complesse su file interni alle diverse cartelle servizio che vengono create.

La linea comando

Il modo senz'altro più potente per effettuare operazioni in GRASS è attraverso la linea di comando della Console dei Comandi. I comandi vengono scritti nel linguaggio di GRASS (bash).

La sintassi generica di un comando GRASS è la seguente:

```
module [-flag1] [-flag2] parameter1=map1 [,map2,...]  
[parameter2=number] [--o] [--q] [--v].
```

Dove *module* è il nome del comando, i *flag* opzionali abilitano caratteristiche speciali del comando, *parameter* specificano il nome dei file di input e di output, una costante numerica, il nome di una metodologia caratteristica del comando, ecc., *--o* abilita la sovrascrittura di mappe

esistenti (altrimenti grass non la permette), `--q` e `--v` diminuiscono o aumentano le informazioni che vengono fornite sul flusso di elaborazioni del comando. Per avere una descrizione dei parametri di uno specifico comando è possibile usare l'opzione `help`:

```
d.rast help
```

Il prefisso di ciascun comando, cioè la lettera prima del punto specifica la classe del comando. I comandi sono raggruppati nelle seguenti classi:

- d.* display commands
- db.* database commands
- g.* general commands
- i.* imagery commands
- m.* miscellaneous commands
- ps.* postscript commands
- r.* raster commands
- r3.* raster3D commands
- v.* vector commands

Il significato di null-value

GRASS, come molti ambienti di elaborazione raster (per esempio ARCMAP o ARCVIEW) distinguono, all'interno dei valori che possono assumere le celle di un dato raster, fra il valore zero (0) e l'assenza del dato (NULL value). Infatti il valore 0 può rappresentare un dato quantitativo, mentre il valore NULL è utilizzato laddove il dato non è disponibile. La differenza fra zero e NULL è molto importante nelle elaborazioni matematiche e logiche.

Per specificare in una base dati il valore NULL è possibile utilizzare il modulo `r.null` nei tools di GRASS.

Map algebra con Q-GIS (e GRASS)

Le operazioni di map algebra sono accessibili attraverso diversi possibili comandi. Nell'ambito degli operatori locali GRASS mette a disposizione il modulo `r.mapcalc`.

Il modo però senz'altro più potente per effettuare operazioni di map algebra è attraverso la linea di comando. Attraverso la sintassi specifica di GRASS, ereditata dai linguaggi di programmazione, è così possibile realizzare procedure di calcolo molto complesse.

Una sessione di `mapcalc` inizia con il comando:

```
r.mapcalc "nuovamappa=(... espressione...)"
```

oppure con il comando `r.mapcalc` seguito da una o più espressioni del tipo `nuovamappa1=(...espressione...)`

nuovamappa2= (...espressione...)

...

All'interno dei comandi di `r.mapcalc`, nell'ambito degli operatori locali, è possibile realizzare le operazioni e utilizzare le funzioni riportate in tabella.

operatore	significato
-	negation
~	one's complement
!	not
^	exponentiation
%	modulus
/	division
*	multiplication
+	addition
-	subtraction
<<	left shift
>>	right shift
>>>	right shift (unsigned)
>	greater than
>=	greater than or equal
<	less than
<=	less than or equal
==	equal
!=	not equal
&	bitwise and
	bitwise or
&&	logical and
&&&	logical and[1]
	logical or
	logical or[1]
?:	conditional

Funzioni di `r.mapcalc`

Gli operatori focali saranno analizzati in dettaglio in un apposito capitolo. Per quanto riguarda gli operatori zonali, questi sono accessibili attraverso il modulo `r.statistics`. Il modulo richiede che i dati da sovrapporre sia convertiti in valori interi, questo è realizzabile con il comando `r.rescale` dando come estremi i valori minimi e massimi dei due dataset da sovrapporre, rinvenibili nel metadata del browser. Il comando `r.statistics` consente di sovrapporre mappe sulla base di numerose statistiche.

Funzione	Descrizione
abs(x)	return absolute value of x
acos(x)	inverse cosine of x (result is in degrees)
asin(x)	inverse sine of x (result is in degrees)
atan(x)	inverse tangent of x (result is in degrees)
atan(x,y)	inverse tangent of y/x (result is in degrees)
cos(x)	cosine of x (x is in degrees)
double(x)	convert x to double-precision floating point
eval([x,y,...],z)	evaluate values of listed expr, pass results to
exp(x)	exponential function of x
exp(x,y)	x to the power y
float(x)	convert x to single-precision floating point
graph(x,x1,y1[x2,y2..])	graph
if	decision options:
if(x)	1 if x not zero, 0 otherwise
if(x,a)	a if x not zero, 0 otherwise
if(x,a,b)	a if x not zero, b otherwise
if(x,a,b,c)	a if x > 0, b if x is zero, c if x < 0
int(x)	convert x to integer [truncates]
isnull(x)	check if x = NULL
log(x)	natural log of x
log(x,b)	log of x base b
max(x,y[,z...])	largest value of those listed
median(x,y[,z...])	median value of those listed
min(x,y[,z...])	smallest value of those listed
mode(x,y[,z...])	mode value of those listed
not(x)	1 if x is zero, 0 otherwise
pow(x,y)	x to the power y
rand(a,b)	random value x : a <= x < b
round(x)	round x to nearest integer
sin(x)	sine of x (x is in degrees)
sqrt(x)	square root of x
tan(x)	tangent of x (x is in degrees)
xor(x,y)	exclusive-or (XOR) of x and y

Funzioni di r.mapcalc