

L'analisi geografica Open Source: GRASS GIS parte 1^a - Vettori

di Leonardo Lami, Paolo Cavallini, Radim Blazek

Se il GIS è uno strumento così diffuso nel mondo, lo si deve in gran parte all'attenzione che è stata data all'interoperatività, almeno a livello di formati. Buona parte del merito di questa scelta di fondo, com'è noto, si deve all'Open Geospatial Consortium. Quel che è meno noto, invece, è che l'OGC è nato come evoluzione dello sviluppo del più importante software libero per l'analisi geografica: GRASS, un acronimo di Geographic Resources Analysis Support System. È forse il GIS dalla tradizione più lunga (le sue origini risalgono al 1982) ed ha una storia complessa: all'inizio è stato sviluppato prima dai laboratori di ricerca dell'esercito statunitense, poi è stato rilasciato come software libero e il suo sviluppo è stato coordinato da Markus Neteler, presso alcune Università (Baylor, Hannover). Oggi il centro di sviluppo risiede presso l'ITC-irst di Trento (un vanto nazionale!).

Come tutti i GIS più "antichi", GRASS si porta dietro una fama di programma complesso, senza interfaccia grafica, e di difficile apprendimento. Grazie al lavoro degli ultimi anni, l'ultima versione disponibile (la 6.0, liberamente scaricabile dal sito ufficiale: <http://grass.itc.it>) è invece pienamente usabile e disponibile per le principali piattaforme operative: GNU/Linux, MacOSX, MS-Windows, iPAQ (Fig. 2). Grazie ad una struttura modulare (ogni comando è un'applicazione indipendente), è molto stabile e robusto anche per analisi pesanti e complesse, sia in campo raster che vettoriale e per



Figura 1 - Sviluppo di QGIS. Incremento della dimensione del codice programma nel tempo. Si nota un fortissimo incremento a partire dal gennaio 2004

L'analisi di immagine, sia in 2 che in 3 dimensioni (Fig. 3). I comandi disponibili sono oltre 350, molti dei quali dotati di varie opzioni, per un set complessivo di oltre 600 possibili istruzioni (una lista completa su http://grass.itc.it/grass60/manuals/html60_user/index.html). Lo scopo di questo primo articolo è quella di fornire un'introduzione generica al-

l'ambiente di lavoro di GRASS ed approfondire in particolare le caratteristiche e le possibilità di analisi del formato vettoriale. In futuro tratteremo delle analisi d'immagine e dei raster, delle visualizzazioni 2,5 e 3D e dei modelli specializzati.

Le caratteristiche del modello dati vettoriale

Il nuovo formato vettoriale di GRASS mantiene una caratteristica notevole di questo programma, ovvero la struttura topologica, che gestisce le informazioni che riguardano le mutue relazioni spaziali tra i diversi elementi, come la connessione, l'adiacenza o l'inclusione rendendo il dato in se stesso più corretto e ricco di informazioni. Da un punto di vista dell'utilizzo pratico, una delle conseguenze più ovvie è che

Figura 2 - Le principali piattaforme operative per cui GRASS è disponibile



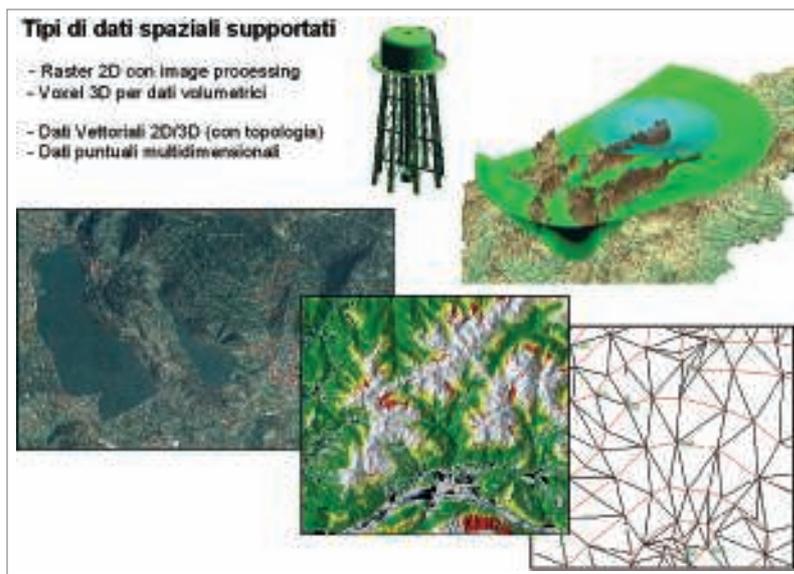


Figura 3 - Le principali tipologie di dati geografici gestiti da GRASS

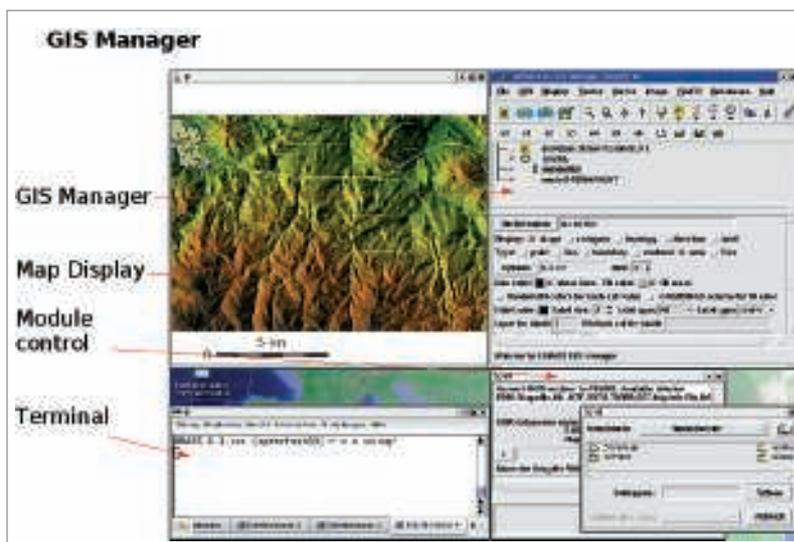


Figura 4 - L'interfaccia grafica di default

durante la digitalizzazione in GRASS non è possibile incorrere nei noti errori di coerenza (presenza di linee intersecate, ecc.) comuni in programmi con modello dati non topologico. Altri software risolvono il problema tramite una correzione a posteriori degli errori, ma l'approccio di GRASS è più potente e razionale. Altra particolarità, propria dei GIS più completi, è la possibilità di gestire dati vettoriali multiformato, ovvero caratterizzati dalla contemporanea presenza di punti, linee e poligono-

Tabella 1 - Elenco e breve descrizione dei principali gruppi di funzioni individuabili nel menù a tendina

MENÙ	CONTENUTO
File	Comandi di import/export (Fig. 5), gestione dei progetti di visualizzazione e della stampa
GIS	Gestione della regione di visualizzazione, delle variabili di sistema, del sistema di proiezione e della conversione tra dati geometrici diversi
Display	Comandi di visualizzazione (2D e 3D), posizione geografica, misure interattive
Raster	Opzioni di visualizzazione e comandi di elaborazione ed analisi dei raster
Vector	Opzioni di visualizzazione e comandi di elaborazione ed analisi dei vettori
Image	Elaborazione ed analisi di immagini: rettificazione, Brovey transformation, classificazione, filtri, risposta spettrale...
Grid3D	Opzioni di visualizzazione e comandi di elaborazione ed analisi dei raster 3D (voxel)
Database	Comandi di gestione dei database: connessione, operazioni sulle tabelle, queries
Help	Un utile manuale in html

ni nello stesso strato informativo; questa peculiarità permette di poter rappresentare oggetti complessi per i quali un unico formato non sia sufficiente. Oltre agli aspetti geometrico e topologico, il modello dei dati, per essere efficace, deve prevedere il collegamento con dati che descrivono i singoli oggetti reali, definibili come attributi. Elemento distintivo di GRASS è il fatto di avere associato ad ogni elemento geometrico del vettore (punto o singolo segmento di linea o poligono) un valore numerico, detto categoria, che permette di collegare in maniera stabile i singoli elementi ai record della tabella alfanumerica associata tramite un determinato campo chiave. Ciò permette da una parte, di poter effettuare semplici selezioni sui diversi elementi del vettore, dall'altra di ottimizzare la costruzione e il rapporto con la tabella associata. In particolare è possibile associare ad ogni dato vettoriale più strati categorici (layer) e più valori di categoria per lo stesso strato; è quindi possibile connettere allo stesso elemento geometrico più records della stessa tabella ed associare allo stesso vettore tabelle diverse che contengono dati relativi a differenti contesti aumentando in modo semplice la quantità di informazioni associate. Naturalmente i vettori possono essere collegati a semplici file dbf o a database relazionali quali PostgreSQL in modo diretto o MySQL, Oracle ed altri tramite driver ODBC. In tale contesto è utile ricordare che per velocizzare la gestione geometrica del vettore e l'in-

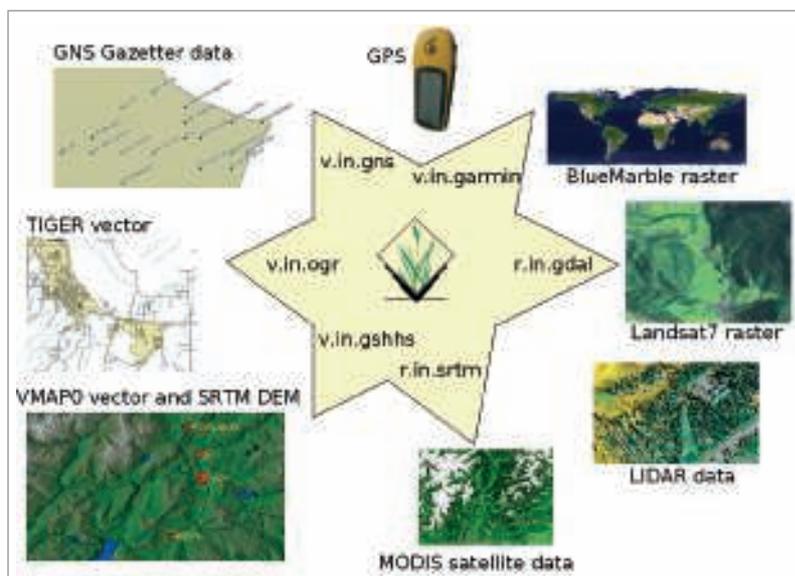


Figura 5 - I più comuni formati vettoriali e raster importabili in GRASS

terrogazione anche complessa dei dati associati, GRASS utilizza un'indicizzazione sia per la parte spaziale, sia per le categorie prima descritte. Da ricordare infine la capacità di gestire anche dati vettoriali tridimensionali (di estrema utilità per numerosi campi applicativi).

Come lavorare in GRASS: interfaccia grafica e shell di testo

GRASS permette di lavorare sia utilizzando varie interfacce grafiche (GUI, ovvero Graphical User Interface) sia tramite la linea di comando.

La GUI di default (Fig. 4), scritta nel linguaggio di programmazione grafica Tcl/Tk, si presenta in un formato abbastanza semplice ma la sua evoluzione è tuttora in corso ed è caratterizzata da una serie di finestre a tendina che raggruppano le funzioni presenti raccolte sotto diversi voci (Tabella 1).

A questo menù si associano una serie di icone per le operazioni più ricorrenti, tra cui quelle di zoom e di pan e di avvio della digitalizzazione, nonché uno spazio per le opzioni di visualizzazione dei vettori e dei raster.

L'interfaccia grafica più sofisticata e moderna è quella di QGIS (www.qgis.org; vedi l'articolo sul numero scorso) che associa alle numerose capacità di questo "desktop GIS" un plugin per la visualizzazione dei dati e per richiamare i comandi di GRASS. Un'altra opzione, particolarmente gradevole all'a-

spetto, è costituita da JGRASS (Fig. 6) (www.hydrologis.com/html/jgrass/jgrass_en.html), programmato in Java e tutt'ora in sviluppo, che mostra sofisticati effetti di trasparenze (esempi su: www.hydrologis.com/html/jgrass/screenshotshots/images.html).

Una pratica interfaccia a linea di comando (detta shell di testo, secondo la tradizione Unix) permette infine di

azionare i comandi in modalità testuale e di avviare programmi (detti script), anche complessi, che rendono possibile eseguire operazioni cicliche; queste possono comprendere anche variazioni dei dati di input in modo completamente automatico.

Qual è la migliore interfaccia? Come spesso accade nel mondo del software libero, la risposta non è univoca; per un primo approccio a GRASS, sicuramente QGIS è l'opzione migliore, per facilità ed immediatezza d'uso (anche se per il momento non tutti i comandi sono disponibili); d'altro lato, l'interfaccia di default (in linguaggio Tcl/Tk) è al momento la più completa, richiede meno risorse, e in alcune condizioni può essere più semplice da installare; in alcuni contesti, dove già si faccia uso di Java, l'interfaccia programmata in questo linguaggio è gradevole e di uso semplice. Per lavori complessi e ripetitivi, la shell offre opportunità eccezionali, fra cui la possibilità di ripetere esattamente la stessa analisi in modo totalmente automatico (utile ad esempio per analisi da ripetersi ad intervalli regolari) e la possibilità di ottenere la completa documentazione delle analisi svolte (utile per render conto della

Tabella 2 - Elenco e breve descrizione delle funzioni principali relative al dato vettoriale

TIPO DI OPERAZIONE	DESCRIZIONE
Operazioni topologiche e geometriche	Ricostruzione della topologia, creazione di polilinee a partire da segmenti adiacenti e viceversa, trasformazione di tipologie geometriche in altre, operazioni sui diversi strati categorici (layer)
Connessioni database e vettori	Connettere vettori a tabelle, creare connessioni a formati vettoriali esterni (shapefile, PostGIS, MapInfo, ecc.) senza necessità di importazione
Proiezioni	Rettificazione e georeferenziazione, riproiezione in altri sistemi di riferimento spaziale
Queries	Query visive, selezione di elementi in base agli attributi o in base alle tipologie geometriche
Reports e statistiche	Aggiornamento delle tabelle associate sulla base delle caratteristiche geometriche del vettore, statistiche univariate sugli attributi, test di normalità sulla distribuzione degli elementi puntiformi
Analisi spaziali	Creazione di buffers, operazioni di overlay, misura di distanze tra i diversi elementi geometrici di due vettori
Analisi di reti	Individuazione del percorso più breve, delle sottozone di competenza, albero di Steiner...
Analisi spaziale con vettori puntiformi	Creazioni di minimi poligoni convessi, triangolazione di Delaunay, creazione di set di punti random, analisi kernel
Creazione di superfici	Interpolazione di superfici tramite i metodi di Voronoi, Inverse Distance Squared Weighting e Regularized Spline with Tension"

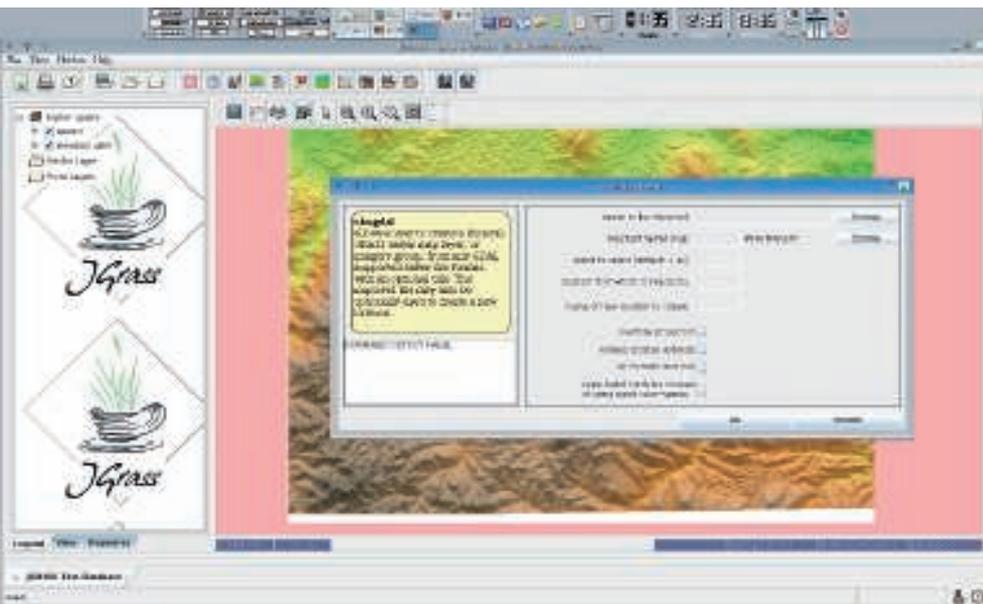


Figura 6 - L'interfaccia grafica di J-GRASS

correttezza delle proprie analisi). L'integrazione con gli strumenti standard Unix inoltre la rende estremamente potente, in quanto è possibile utilizzare tutti gli elementi base della programmazione (cicli, condizioni, espressioni regolari, ecc.). L'apprendimento della linea di comando è reso più agevole dalla GUI standard, che mostra la "traduzione" in linguaggio testuale di un comando dato in via grafica. Copiando ed incollando tali comandi si potranno costruire i propri script in modo molto semplificato. Quindi l'utente, a seconda delle sue esigenze, della propria preparazione e delle proprie inclinazioni, selezionerà di volta in volta la soluzione più adatta.

Alcuni esempi

Com'è ovvio, GRASS contiene un modulo di digitalizzazione per la creazione/modifica di dati vettoriali e moduli di importazione per numerose tipologie vettoriali tramite la libreria OGR (vedi www.gdal.org/ogr/ogr_formats.html) ma anche le possibilità di elaborazione ed analisi per il formato vettoriale sono numerose (Tabella 2). Se ad esempio vogliamo creare un buffer intorno ad un vettore possiamo utilizzare il comando `v.buffer`, la finestra di gestione delle opzioni relative a questa funzione può essere attivata scrivendo il nome del comando nella con-

sole di testo o tramite il seguente percorso sul menù a tendina: *Vector* → *Create vector buffers*.

Nella finestra che si aprirà sarà necessario inserire il nome del vettore di cui vogliamo creare il buffer (name of input vector), il nome che vogliamo dare al nuovo vettore (name of output vector), scegliere quali tra le diverse geometrie di cui è costituito il vettore di input ci interessano per la creazione del buffer (type), definire l'ampiezza

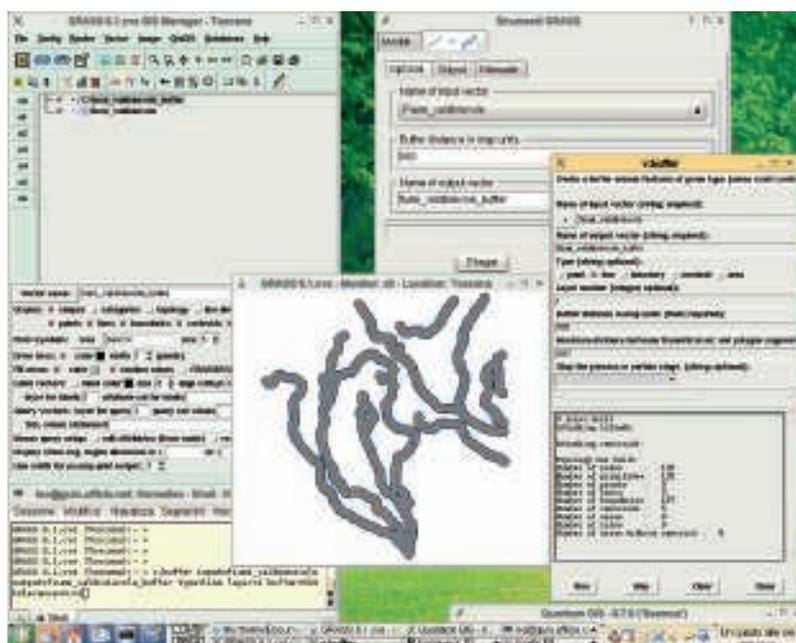
del buffer stesso (buffer distance in map units) ed infine premere il pulsante RUN (Fig. 7).

Nella finestra di output verranno fornite informazioni sulle operazioni in corso, una descrizione geometrica del vettore creato ed eventuali "warnings" od errori riscontrati durante il processo; la prima riga di questo messaggio equivale alla riga di comando, comprensiva delle opzioni scelte, da digitare nella console per lanciare il comando in modalità testuale.

Le capacità di analisi e la varietà di opzioni possono essere apprezzate nella determinazione della minima distanza tra gli elementi di un vettore puntiforme, che per esempio potrebbe rappresentare le segnalazioni GPS di un animale, e gli elementi geometrici di un altro vettore, per esempio quello relativo alle zone boscate; il comando da utilizzare via testo è `v.distance`, da menù *Vector* → *Locate nearest features to points or centroids*.

Prima di descrivere le opzioni di questa funzionalità è utile chiarire il valore aggiunto delle categorie che caratterizzano i diversi elementi vettoriali; in questo esempio i diversi valori delle categorie del vettore puntiforme rappresentano i diversi animali mentre le categorie dell'altro vettore descrivono il tipo di bosco. Grazie a tale struttura

Figura 7 - Il comando `v.buffer`: le finestre di dialogo di GRASS e QGIS, la shell di testo e il risultato grafico



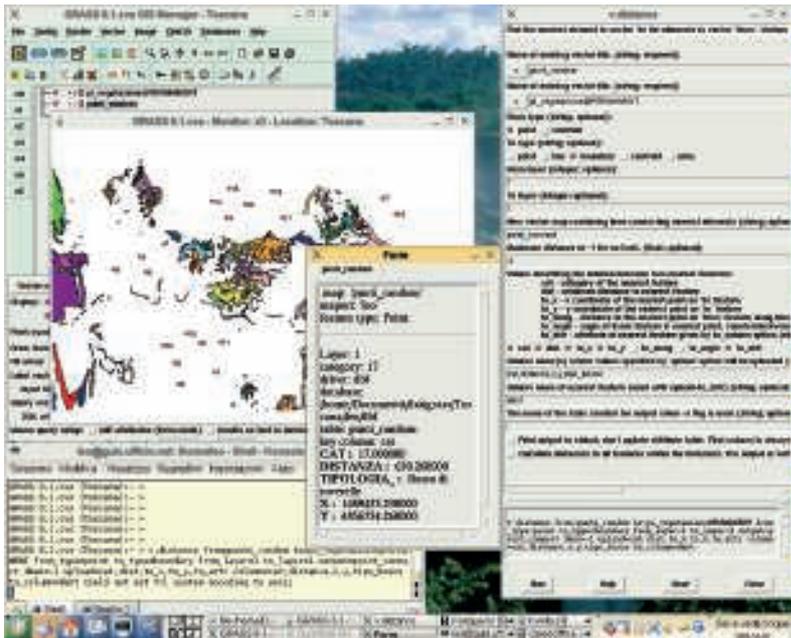


Figura 8 - Il modulo v.distance: le finestre di dialogo di GRASS, la shell di testo e il risultato grafico

il comando v.distance ci fornirà non solo le distanze dal bosco più vicino ma anche le relazioni esistenti tra i singoli animali e le caratteristiche del bosco più vicino incrementando la quantità di informazioni ottenute.

Nella finestra di opzioni occorre inserire: i nomi del vettore puntiforme e del vettore da cui calcolare la distanza (name of existing vector file), la tipologia geometrica dei vettori su cui vogliamo effettuare l'operazione di misura (from type, to type), gli strati categorici da prendere in considerazione nel calcolo delle distanze e nell'upgrade delle tabelle (from layer, to layer), il nome di un possibile nuovo vettore che rappresenti i segmenti di connessione tra gli elementi dei due vettori (new vector...) e una possibile distanza massima da considerare (maximum distance...).

A questo punto è necessario scegliere quali siano le informazioni da ottenere dall'analisi: per ogni categoria (animale) del vettore puntiforme è possibile determinare la categoria (tipo di bosco) più vicina, la distanza da questa, le coordinate del punto del poligono più vicino per ogni punto vettoriale, il valore di un attributo per il record associato al poligono più vicino ed ulteriori relazioni tra le linee di connessione. Infine si definiscono i nomi dei campi

della tabella collegata al vettore puntiforme dove registrare le informazioni ottenute in modo da poterle associare ai singoli elementi e quindi si preme RUN; interrogando il vettore puntiforme è possibile osservare i risultati dell'analisi (Fig. 8).

Analisi complesse

Un'operazione di notevole interesse è la costruzione di isosuperfici (ad esempio temperature, concentrazioni di sostanze) a partire dai valori associati ad un vettore puntiforme; l'esempio più classico è la costruzione di un modello digitale del terreno a partire da un vettore relativo ai punti quotati di un determinato territorio.

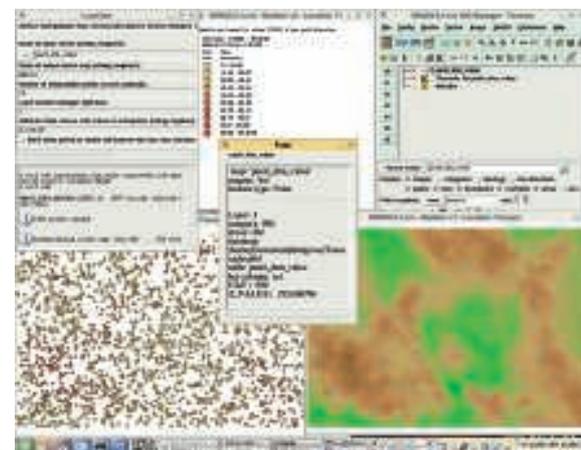
GRASS utilizza diversi metodi di interpolazione che sono basati sia su condizioni di localizzazione (triangolazione di Voronoi, inverso della distanza pesata) sia su funzioni di smoothing (RST). La triangolazione di Voronoi (v.voronoi) permette la creazione di un vettore poligonale dove ai singoli poligoni è associato il valore (quota) del punto corrispondente, il raster ottenuto si presenta come una superficie costituita da piani caratterizzati da diverse altitudini. Il metodo IDW (inverse distance weighted) permette di assegnare ad ogni cella del raster di output la media

pesata sulla distanza dei valori assegnati agli n punti vettoriali più prossimi; tale funzione (v.surf.idw) è consigliata quando la densità dei punti del vettore di input è maggiore della densità delle celle del raster di output.

Il metodo RST (regularized spline with tension) calcola i valori della griglia raster sulla base di una funzione che meglio interpola i punti del vettore di input, tale metodo è applicabile su un set di punti per ottenere una isosuperficie (v.surf.rst), di cui è possibile calcolare anche i vari parametri topografici (esposizione, pendenza...) e derivate parziali di vario grado, od un raster volumetrico (v.vol.srt).

Come esempio proponiamo la creazione di un modello digitale del terreno a partire da un vettore puntiforme utilizzando la tecnica IDW; a tal scopo utilizziamo il comando v.surf.idw da shell o da interfaccia grafica (Raster → Interpolate surfaces → Inverse distance weighed interpolation from vector points) ma prima occorre definire la risoluzione di lavoro e quindi del raster da creare (ad esempio g.region res=50). Nella relativa finestra occorrerà inserire il nome del vettore puntiforme (name of input vector), il nome dell'isosuperficie da creare (name of output raster map), il numero dei punti da considerare per il calcolo delle media pesata (number of interpolation points), nonché il nome dell'attributo della tabella associata al vettore puntiforme che contiene il valore da interpolare (attribute table column..., Fig. 9).

Figura 9 - Il comando v.surf.idw: le finestre di dialogo di GRASS, la shell di testo e il risultato grafico



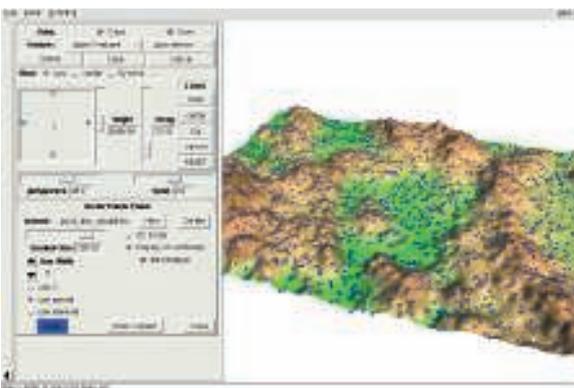


Figura 10 - Il risultato di *v.surf.idw* visualizzato in modalità prospettica tramite il modulo *nviz*

Il risultato può essere visualizzato sia in versione 2D, tramite il monitor standard, sia in versione tridimensionale, tramite il comando *nviz* (*Display* → *Start NVIZ*) che permette inoltre di sovrapporre vettoriali, creare animazioni, e interagire in modo interattivo (*fly-through*) (Fig. 10).

Un altro interessante campo di applicazione è quello dello studio dei network; in GRASS sono possibili ben cinque tipi di analisi:

- individuazione del percorso più breve o a minor costo per congiungere due punti lungo la rete sia in modo interattivo (*d.path*) utilizzando il mouse, sia costruendo un nuovo vettore che descrive il percorso (*v.net.path*);
- divisione della rete in aree ad uguale distanza o tempo di spostamento da un punto definito come ad esempio un servizio medico, una scuola od altro (*v.net.iso*);

- dividere la rete in zone di competenza per ogni servizio dislocato sul territorio in funzione del costo di arrivo (distanza, tempo o costo) lungo la rete stradale dai servizi stessi (*v.net.alloc*);
- determinare la sottorete di minima lunghezza che collega un insieme di punti definiti (albero minimo di Steiner), ad esempio per collegarle degli uffici, posando i cavi lungo le strade, con una rete ad alta velocità (*v.net.steiner*, Fig. 11);
- calcolare il percorso minimo che collega un insieme di punti definiti senza passare per gli stessi più di una volta (problema del commesso viaggiatore) per ottimizzare ad esempio il percorso di fornitori e ispettori (*v.net.salesman*).

Come già in parte illustrato, tutte queste analisi includono la possibilità di definire il percorso migliore sulla base di attributi che rappresentano i "pesi" di attraversamento degli elementi della rete (tempo, costo...) e di effettuare la ricerca dei percorsi utilizzando differenti tipologie di reti quali autostrade, viabilità ordinaria, servizi di trasporto pubblico, ecc. L'individuazione interattiva dei percorsi migliori tra due punti è stato scelto come esempio pratico per analizzare le capacità di analisi dei network da parte di GRASS; a tale scopo è necessario avere a disposizione un semplice vettore lineare che costituisca la rete stradale dove le diverse tipologie di strade siano associate a records che riportino il costo di attraversamento nei due sensi. La finestra interattiva del comando *d.path* (*Vector* → *Network*

analysis → *Shortest route* "visualizzazione only") necessita l'inserimento del vettore che identifica la rete (name of input vector), la tipologia dell'elemento che costituisce l'ossatura della rete (arc type), lo strato categorico dei collegamenti e dei nodi (arc layer, node layer), il nome dei campi che contengono il costo di attraversamento nei due sensi delle diverse tipologie di strade (arc forward/backward direction cost), il nome del campo che contiene il costo di attraversamento dei nodi che potrebbero rappresentare le città di cui non abbiamo a disposizione la rete stradale (node cost column) ed i colori per l'individuazione del percorso. Una volta cliccato sul pulsante RUN si entra nella modalità interattiva: con il tasto centrale del mouse si sceglie il punto di partenza e con il sinistro il punto di arrivo, sul monitor verrà evidenziato il percorso individuato mentre sulla finestra di output comparirà la distanza o il tempo impiegato.

Conclusioni

Questa brevissima introduzione non può certo dar conto della potenza della parte vettoriale e database di GRASS; per un approfondimento, le risorse disponibili sono molte, soprattutto sul Web. Consigliamo soprattutto la consultazione della pagina di documentazione sul sito di GRASS (<http://grass.itc.it/gdp/>), disponibile in molte lingue, incluso l'italiano.

Leonardo Lami, Paolo Cavallini

Faunalia
Piazza Garibaldi 5
Pontedera (PI)
www.faunalia.it

Radim Blazek

ITC-irst - Centro per la Ricerca Scientifica e Tecnologica
MPBA - Predictive Models for Biol. & Environ. Data Analysis
Via Sommarive, 18
38050 Povo (TN)
<http://mpa.itc.it>

Figura 11 - Caratteristiche topologiche ed alfanumeriche del dato rappresentate una rete viaria ed un esempio di creazione di un "albero minimo di Steiner"

