



Impostazione del piano delle elaborazioni in un SIT

Approcci per il processamento dei dati e la creazione di valore informativo aggiunto

Corso di Sistemi Informativi Territoriali per il Planning e l'Urban Design – UD13

prof. Giovanni Borga

Progettare la logica di elaborazione dei dati in un SIT

Mentre le tecniche di base, seppur fondamentali, sono strumenti indipendenti dalla tematica di applicazione, l'approccio da adottare nella progettazione della **logica di elaborazione dei dati del SIT** è invece **in relazione diretta con il tipo di supporto decisionale** che il sistema dovrà offrire all'utilizzatore.

Un SIT ha in generale una sua valenza intrinseca indipendente dal supporto al processo decisionale, derivata dagli ovvi benefici del poter gestire in modo efficiente e strutturato le informazioni necessarie a condurre le proprie attività.

*Tuttavia, un sistema che abbia una **finalità direttamente riconducibile ad uno specifico processo decisionale** nell'ambito del governo del territorio è un sistema concettualmente più evoluto e giustifica maggiormente gli investimenti necessari alla sua realizzazione.*

La strategia della quantificazione dei fenomeni

Le strategie da cui si ottiene il miglior supporto decisionale sono basate su tecniche di quantificazione di aspetti della realtà sotto forma di

INDICI NUMERICI

Il vantaggio di definire una **funzione che quantifica** un aspetto del tema che si vuole affrontare è quello di poter svolgere almeno due tipologie di comparazioni:

1. **su base geografica** (ovvero creare mappe tematiche graduate)
2. **su base temporale** (ovvero calcolare dei trend).

Le comparazioni di indici quantitativi applicati ad un livello di unità geografiche congruo permettono la definizione di **specifiche strategie di analisi** dei dati di un SIT.

Strategie di uso comune nel processamento dei dati geografici

In base all'approccio determinato dalla tematica di applicazione, le **strategie più frequentemente adottate** nel progettare il piano di geoprocessing sono relativamente poche ma hanno tutte la fondamentale caratteristica di essere esplicitamente orientate al supporto decisionale; possiamo classificarle secondo 5 modelli:

1. Modello di analisi domanda-offerta
2. Modello orientato alla valutazione di impatti
3. Modello di stima dei rischi
4. Modello orientato alla comparazione di scenari
5. Modello basato sull'analisi di rete

Le diverse strategie si applicano sostanzialmente combinando le tecniche di base di geoprocessing e visualizzazione definendo un workflow sequenziale.

1) Modello domanda-offerta

Il modello di analisi domanda-offerta è finalizzato a **confrontare due valori/indici** che quantificano, da un lato la disponibilità di un bene o di un servizio, dall'altro la richiesta effettiva o potenziale di tale bene o servizio.

Il modello domanda-offerta è dunque in sé piuttosto elementare.

Gli impieghi più tipici sono ad esempio i casi in cui poter stimare o misurare domanda e offerta di beni e servizi su base territoriale consente di effettuare una migliore allocazione di strutture o di migliorare l'efficacia e l'efficienza dei processi.

Tematiche tipiche di applicazione:

- *Energia*
- *Trasporti*
- *Commercio / Servizi*

1) Modello domanda-offerta: metodologia

1. Produzione di un livello informativo di stima o georeferenziazione della domanda
2. Produzione di un livello informativo di stima o georeferenziazione dell'offerta
3. Individuazione o creazione di un livello di base composto di unità geografiche ottimali per le procedure di processamento e per la restituzione dei risultati
4. Calcolo di un indice sintetico domanda/offerta sulla base delle unità geografiche che potrà essere:
 - a) differenziale (offerta-domanda oppure domanda-offerta)
 - b) di rapporto (offerta/domanda oppure domanda/offerta)

NB: generalmente i livelli di domanda e offerta si ottengono dall'elaborazione integrata di più livelli informativi che costituiscono singole componenti che vanno preventivamente determinate e successivamente combinate tra loro.

1) Modello domanda-offerta: esempio di workflow

Tema: offerta di servizio ferroviario

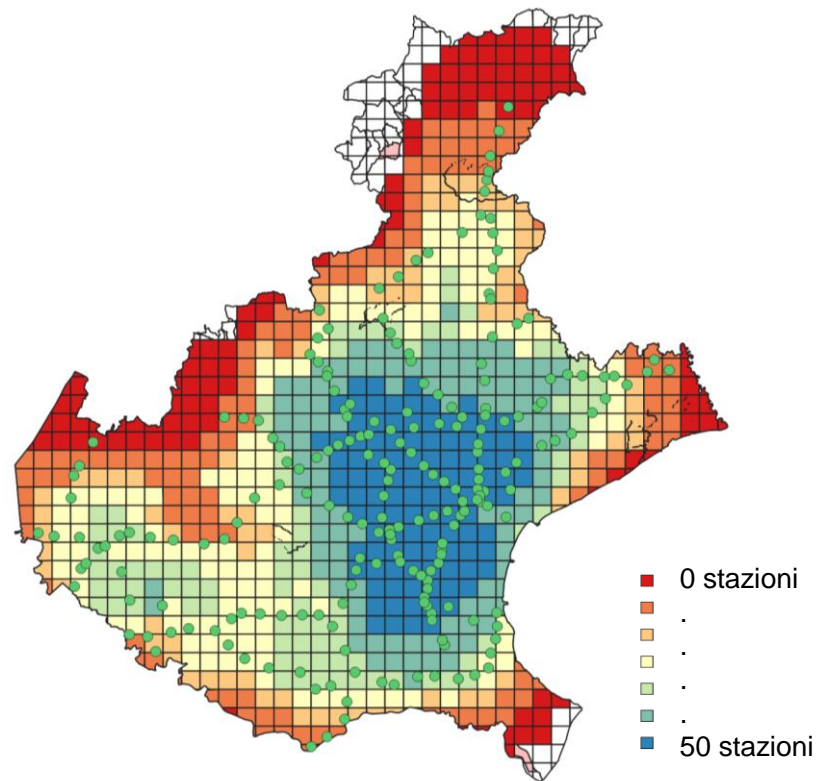
1. *Join tabellare tra tabulato della popolazione e layer dei confini comunali e calcolo della densità demografica*
2. *Creazione di una griglia a maglie regolari (fishnet)*
3. *Intersezione tra layer comuni e griglia*
4. *Aggregazione dei poligoni risultanti dall'intersezione per ID cella e calcolo della media della densità demografica*
5. *Buffer sui baricentri delle stazioni ad una distanza che rappresenta l'area di accessibilità*
6. *Intersezione tra buffers e griglia*
7. *Aggregazione dei poligoni risultanti dall'intersezione per ID cella e calcolo del conteggio dei buffers*
8. *Join tra i due dataset risultati dalle aggregazioni al layer grid per ID cella*
9. *Calcolo del rapporto tra conteggio stazioni e densità demografica per singola cella e realizzazione della relativa mappa tematica graduata*

1) Modello domanda-offerta: esempio di indice di offerta

Rappresentazione su base grid dell'offerta di servizio ferroviario.

L'indice di quantificazione dell'offerta in questo caso è semplificato per finalità didattiche: è infatti calcolato come numerosità di stazioni nel raggio di un intorno fisso che sintetizza il loro bacino di fruizione.

Nella mappa tematica i colori tendenti al verde-blu rappresentano maggiore offerta, mentre quelli tendenti al giallo-rosso una minor offerta.

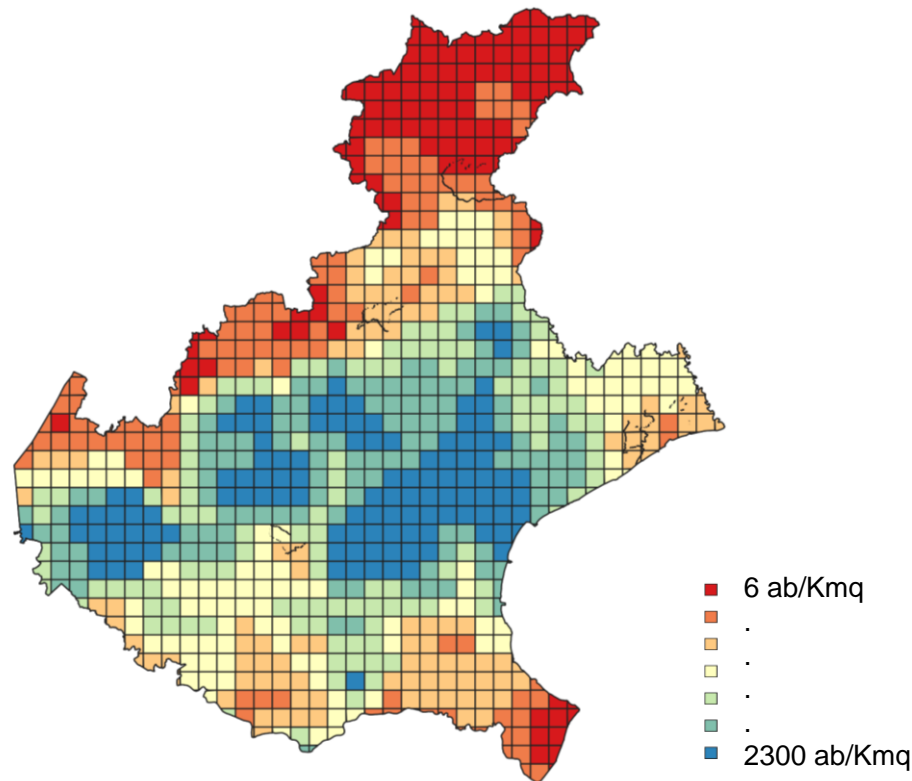


1) Modello domanda-offerta: esempio di indice di domanda

Rappresentazione su base grid della domanda di servizio ferroviario.

Anche l'indice di quantificazione della domanda è volutamente esemplificativo: è equiparato unicamente alla popolazione residente (preventivamente calcolata sotto forma di densità cella per cella).

I colori tendenti al verde-blu rappresentano maggiore densità, mentre quelli tendenti al giallo-rosso una minor densità.

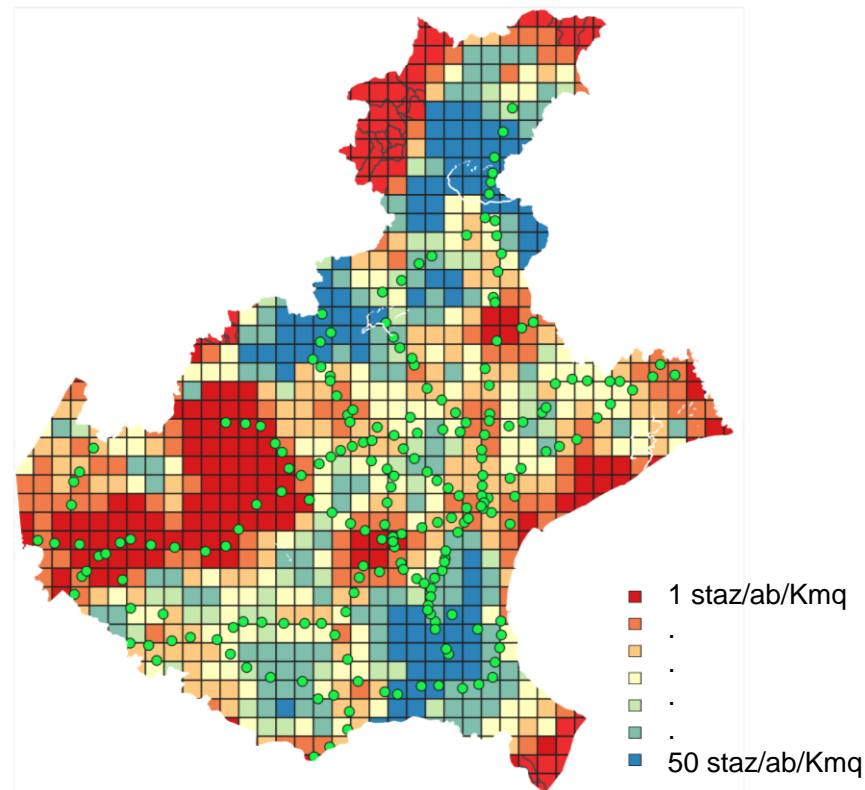


1) Modello domanda-offerta: esempio di indice sintetico

Rappresentazione su base grid del rapporto offerta/domanda di servizio ferroviario.

Come quantificazione del rapporto offerta/domanda si divide semplicemente il numero di stazioni nel bacino per la densità demografica ottenendo un indice quantitativo espresso in *stazioni/ab/Kmq*.

NB: essendo l'offerta a denominatore, anche questa è una mappa dell'offerta, che però non è in valore assoluto ma ponderata in relazione alla domanda potenziale.



1) Modello domanda-offerta: considerazioni

In questo esempio **gli indici sono stati calcolati in modo estremamente semplicistico**. Normalmente i dati permettono di elaborare quantificazioni in modo più sofisticato; ad esempio l'offerta di servizio è più correlata al numero di partenze/arrivi, dimensione dei convogli ecc, più che dalla sola presenza della stazione. Altre considerazioni simili possono essere fatte per la stima della domanda.

Nel caso dell'analisi dell'offerta di servizi ferroviari si è optato per indice di rapporto in quanto **le unità di misura** (n. stazioni e densità demografica) **non sono omogenee**.

Nel caso di unità omogenee per domanda e offerta (es. quantificazioni economiche) può essere opportuno calcolare un indice di differenza.

2) Modello orientato alla valutazione di impatti

La valutazione di come un'azione umana impatta sul territorio circostante è un'operazione necessaria in tantissimi ambiti di analisi.

Questo approccio è quindi adatto ai casi in cui si intenda studiare le esternalità di fenomeni più o meno circoscritti in cui **l'impatto è dato da più fattori singolarmente apprezzabili ma difficilmente valutabili in senso cumulativo o sinergico.**

Il requisito fondamentale è dato dalla disponibilità di informazioni sui singoli fattori di impatto sul territorio.

La caratteristica peculiare di questo modello è determinata dal fatto che i fattori sono generalmente disomogenei e incidono in modo differenziato sull'impatto globale determinato dalla loro concomitanza. Ogni fattore va dunque associato ad un «peso specifico» nel calcolo generale.

2) Modello orientato alla valutazione di impatti

L'impatto di un determinato fattore può essere:

- a) **Misurato** *(nel qual caso la valutazione viene effettuata ex-post e ha più valenza di monitoraggio di efficienza/efficacia dei sistemi);*
- b) **Stimato con un modello probabilistico** *(in questo caso la valutazione avviene ex-ante e prende la forma di un modello di stima, spesso multi-criteri)*

L'approccio della valutazione di impatti può essere adottato anche per aspetti positivi, ad esempio per quanto riguarda ciò che rende attrattiva una zona urbana o territoriale in virtù di attività, beni o servizi presenti all'interno o in un ambito circostante.

Tematiche tipiche di applicazione:

- *Ambiente e biodiversità*
- *Inquinamento e salute umana*
- *Accessibilità / Inaccessibilità ai servizi*
- *Attrattività / Degrado*

2) Modello orientato alla valutazione di impatti: metodologia

1. Produzione di un livello informativo di stima o georeferenziazione di ogni fattore semplice
2. Individuazione o creazione di un livello di base composto di unità geografiche ottimali per le procedure di processamento e per la restituzione dei risultati
3. Intersezione dei diversi livelli relativi ai fattori semplici con il livello base in modo da ottenere un livello base con una colonna e relativi valori per ogni fattore
4. Scelta di un coefficiente di ponderazione per ogni fattore (peso)
5. Calcolo di un indicatore sintetico globale mediante una formula di somma pesata (valore₁ x peso₁ + valore₂ x peso₂ ... + valore_n x peso_n).

In molti casi occorre «normalizzare» i diversi valori per omogeneizzare la variabilità (es. ridefinirli in modo che tutti varino da 0 a 1 o da 0% a 100% ...)

Può essere anche opportuno definire più indici determinati da diversi «set di pesi» per valutare la variazione dell'indicatore globale.

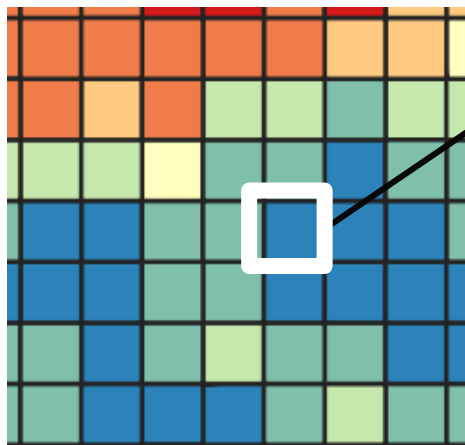
2) Modello orientato alla valutazione di impatti: esempio di workflow

Tema: impatto potenziale determinato da fonti di inquinamento sulla produzione agricola

1. *Creazione di un layer per ogni inquinante (es. traffico veicolare, emissioni da attività industriali ecc.)*
2. *Creazione di layer di buffer (o multi-buffer) per ogni inquinante che rappresenta l'area di potenziale impatto.*
3. *Creazione di una griglia a maglie regolari (in questo caso è fortemente consigliata in quanto la distribuzione sul territorio di questi impatti è omogenea)*
4. *Esecuzione di intersezioni tra buffers degli inquinanti e griglia*
5. *Aggregazione dei poligoni risultanti dall'intersezione per ID cella e calcolo della media (o del valore massimo) di impatto potenziale di ogni inquinante.*
6. *Eventuale normalizzazione dei valori di ogni inquinante (es da 0 a 1) e scelta dei pesi specifici di ciascuno*
7. *Join tra i due dataset risultati dalle aggregazioni al layer grid per ID cella*
8. *Calcolo della somma pesata di ciascun valore di inquinante per singola cella e realizzazione della relativa mappa tematica graduata*

2) Modello orientato alla valutazione di impatti: formulazione

La formulazione finale più utilizzata per il calcolo dell'indice complessivo è una sommatoria pesata in cui a ciascun fattore viene associato un coefficiente di moltiplicazione proporzionale alla sua importanza relativa all'interno dell'indice.



id	fatt1	fatt2	fatt3	$fatt1*w1 + fatt2*w2 + fatt3*w3$
-----------	-------	-------	-------	----------------------------------



Formula dell'indice sintetico

3) Modello di stima dei rischi

L'approccio alle problematiche urbane ha sempre una doppia chiave di lettura: spesso si tratta di individuare **soluzioni a criticità** ma in altri casi l'obiettivo è quello di **valorizzare alcune risorse**.

Nel primo caso, la ricerca di soluzioni a situazioni negative è molto frequente un diretto riferimento al **concetto di rischio**.

Com'è noto, il rischio è per definizione un **concetto probabilistico**:

E' la probabilità che accada un certo evento capace di causare un danno alle persone.

Va dunque affrontato con gli strumenti e le tecniche proprie di questa disciplina.

3) Modello di stima dei rischi

La nozione di rischio implica l'esistenza di:

- a) una sorgente di pericolo
- b) le possibilità che essa si trasformi in un danno.

Il concetto di rischio è associabile a tantissimi domini delle attività umane e dei fenomeni naturali proprio in virtù del fatto che **l'intercorrelazione e le interferenze tra dinamiche differenti quasi sempre è di difficile interpretazione.**

Per la lettura di queste dinamiche occorre dunque applicare **modelli di stima** che fanno uso di indici quantitativi derivati anche da fattori qualitativi.

In un SIT i modelli di stima sono associabili alla componente geografica dell'informazione.

3) Modello di stima dei rischi

I modelli di stima del rischio si basano su una funzione generale ormai nota in letteratura.

$$\text{In generale:}$$
$$\mathbf{R = f (Probabilità, Danno)}$$

In protezione civile, con riferimento alle calamità naturali la funzione è stata declinata in:

$$\mathbf{R = f (Pericolosità, Vulnerabilità, Esposizione)}$$

*in cui la probabilità è sostituita dalla pericolosità mentre il danno diviene
f(Vulnerabilità, Esposizione)*

Tematiche tipiche di applicazione:

- *Tutte le tipologie di rischio*

3) Modello di stima dei rischi: metodologia

1. Produzione di un livello informativo di stima o georeferenziazione della pericolosità
2. Produzione di un livello informativo di stima o georeferenziazione della vulnerabilità
3. Individuazione o creazione di un livello di base composto di unità geografiche ottimali per le procedure di processamento e per la restituzione dei risultati
4. Calcolo di un indice sintetico di prodotto pericolosità * vulnerabilità sulla base delle unità geografiche individuate.

Il calcolo dell'indice sintetico, anche in questo caso è piuttosto semplice. Trattandosi di una formula di rischio i due indici di base vanno MOLTIPLICATI anziché rapportati o addizionati.

3) Modello di stima dei rischi: workflow

Tema: rischio alluvioni

1. *Creazione del layer di stima della pericolosità (es utilizzando le aree storicamente allagate e la rete idrografica).*
2. *Creazione del layer della vulnerabilità (semplificato: ad es. popolazione; oppure ponderato: ad es. proporzionale a popolazione, attività economiche, beni di valore ecc.)*
3. *Creazione di una griglia a maglie regolari (fishnet)*
4. ...
5. ...
6. ...
7. *(la restante parte del workflow è analoga a quanto visto per domanda/offerta)*

Nella stima di pericolosità e vulnerabilità si può applicare il modello di valutazione degli impatti(2) per integrare diverse componenti semplici che concorrono a formare gli indici di stima.

4) Modello orientato alla comparazione di scenari

Come detto più volte, un vantaggio evidente offerto da un SIT ben progettato è dato dalla possibilità di lavorare con **indici quantitativi georiferiti** che possono essere ricalcolati in momenti diversi.

La valutazione di scenari su base multitemporale trova un ottimo strumento nella possibilità offerta dai GIS di **produrre mappe di variazione**.

Le condizioni di base sono il mantenimento della medesima unità geografica di riferimento (anche se entro certi limiti è possibile effettuare trasposizioni e ricalcoli) e la stessa metodologia di calcolo degli indici.

Tematiche tipiche di applicazione:

- *Le più disparate, trattandosi di analisi differenziali. La condizione è data dalla disponibilità di serie di dati georiferiti*

4) Modello orientato alla comparazione di scenari: metodologia

La metodologia di base è costituita appunto dalla produzione di mappe di differenza che possono essere ottenute utilizzando le metodologie viste per gli altri approcci.

Le **varianti** a questo modello sono sostanzialmente due:

1. **Comparazione multitemporale** (si producono due mappe con lo stesso indicatore calcolato in due momenti diversi e si deriva una mappa tematica sulla base della differenza dei due valori).
2. **Comparazione multicriteriale** (si producono due mappe con lo stesso indicatore calcolato con set di pesi diversi e si deriva una mappa tematica sulla base della differenza dei due valori)

5) Modello basato sull'analisi di rete

L'approccio legato all'utilizzo di dati topologici riferiti a strutture a rete è per lo più riconducibile a due tecniche di analisi:

- **Individuazione di percorsi ottimali**
- **Definizione delle aree servite**

Dal punto di vista del valore aggiunto prodotto da queste due tecniche, quella delle aree servite è certamente la tecnica più sofisticata; in particolare l'analisi delle aree servite è più specificatamente orientata ad analizzare la correlazione tra diversi dataset e può essere combinata con i modelli già descritti, in primis il modello domanda/offerta.

5) Modello basato sull'analisi di rete: metodologia

La metodologia prevede sempre una prima fase relativa alla **produzione della topologia di rete** a partire da dati più o meno corretti e ottimizzati.

Successivamente si utilizzano molto spesso le tecniche già viste per gli altri approcci per giungere ad una fase finale in cui si applicano gli algoritmi tipici della **network analysis**.

Questi algoritmi sono stati implementati in modo sostanzialmente diverso da software a software ed esistono dei moduli specifici anche piuttosto sofisticati espressamente dedicati alla produzione e analisi delle topologie di rete.

«Pseudo-modello» basato sull'utilizzo di un GeoDBMS

L'utilizzo di un Geodatabase Management System non comporta cambiamenti di tipo strategico nella progettazione del modello di elaborazione dei dati in un SIT, tuttavia alcuni aspetti incidono particolarmente nell'impostazione del sistema.

Possiamo parlare quindi di «pseudo-modello» in quanto **l'adozione di questa tecnologia ha ricadute significative** nell'impostazione del workflow di elaborazione dei dati che possono avere un'incidenza significativa nella progettazione del piano di elaborazione dati.

«Pseudo-modello» basato sull'utilizzo di un GeoDBMS: ricadute

L'approccio può sembrare prettamente *technology-driven*, tuttavia le ricadute operative sono molteplici e, a volte, sostanziali. Eccone alcune:

Storage di dati geografici

I layers vengono memorizzati sotto forma di tabella in cui le geometrie sono salvate in un campo binario di tipo «geometry». L'accesso al dato multiutente è molto più performante, flessibile, interoperabile e sicuro.

Search by location

La ricerca per posizione viene fatta in modo relativamente semplice con criteri geografici nella clausola SQL «where».

«Pseudo-modello» basato sull'utilizzo di un GeoDBMS: ricadute

Generazione di layers dinamici e join con dati tabellari

Qualsiasi vista memorizzata in cui è presente un campo chiave univoco (anche senza constraint impostato) e un campo geometria può essere visualizzato con un GIS. Il contenuto rimane memorizzato nelle tables riferite nella vista.

Georeferenziazione indiretta

La georeferenziazione indiretta avviene semplicemente impostando un JOIN tra una tabella che contiene geometrie e una che contiene le informazioni da georiferire mediante una foreign key.

«Pseudo-modello» basato sull'utilizzo di un GeoDBMS: ricadute

Query di aggregazione di dati geografici

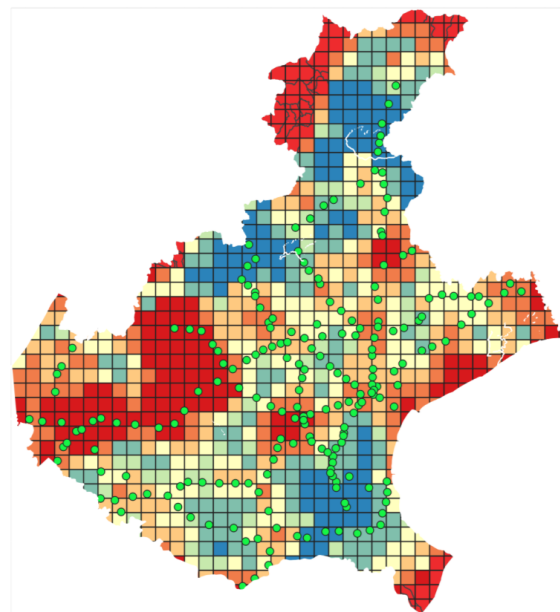
Nelle query con operatore GROUP BY è possibile utilizzare operatori di aggregazione dei dati anche per la componente geografica tra cui ST_UNION che opera il dissolve delle geometrie.

Overlay geoprocessing

In generale sono disponibili tutti gli operatori di geoprocessing offerti dai GIS. Essi tuttavia si applicano con una logica feature-to-feature anziché layer-to-layer (cosa che può non essere di immediata comprensione per chi si approccia per la prima volta ai geoDB).

«Pseudo-modello» basato sull'utilizzo di un GeoDBMS

A titolo di esempio, non esaustivo, riprendiamo il caso dell'analisi dell'offerta di servizi ferroviari utilizzato per esemplificare il modello domanda-offerta e vediamo di seguito la sequenza di query geografiche necessaria per ottenere lo stesso risultato.



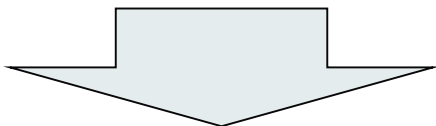
Query geografica per il join dei dati di popolazione

SELECT

```
comuni.id,  
comuni.geom,  
(popolazione.tot2016::integer / ST_AREA(comuni.geom) * 1000000) AS densita
```

FROM

```
comuni  
JOIN popolazione ON comuni.codistat::integer = popolazione.cod_com
```



Vista «comuni_demog»

Questa query effettua un join tra layer dei comuni e dati di popolazione e utilizza l'operatore ST_AREA per calcolare la densità demografica per ciascun comune.

Query geografica per la stima della domanda

SELECT

```
reticolo5km.id,  
reticolo5km.geom,  
avg(comuni_demog.densita) AS densita
```

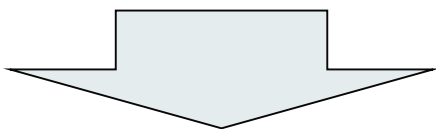
FROM

```
reticolo5km
```

```
JOIN comuni_demog ON ST_INTERSECTS(reticolo5km.geom, comuni_demog.geom)
```

GROUP BY

```
reticolo5km.id,  
reticolo5km.geom
```



Vista «reticolo_demog»

Questa query effettua un'intersezione tra il layer risultante dalla query precedente e un reticolo regolare a maglia 5Km trasferendo il dato di densità demografica. Si usa l'operatore ST_INTERSECTS nel join per correlare i poligoni che reciprocamente si sovrappongono. Il raggruppamento per id/geometria cella raggruppa i dati e consente il calcolo della media.

Query geografica per la stima dell'offerta

SELECT

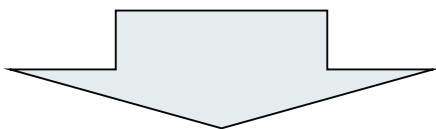
```
reticolo5km.id,  
reticolo5km.geom,  
count(stazioni.id) AS nstaz
```

FROM

```
reticolo5km  
JOIN stazioni ON ST_INTERSECTS(reticolo5km.geom, ST_BUFFER(stazioni.geom, 20000))
```

GROUP BY

```
reticolo5km.id,  
reticolo5km.geom
```



Vista «reticolo_nstaz»

Questa query effettua un'intersezione tra il layer delle stazioni e il reticolo a maglia 5Km utilizzando il già visto operatore ST_INTERSECTS. Dagli elementi del layers stazioni (punti) viene generato contestualmente il buffer di 20Km che rappresenta il bacino potenziale. Il raggruppamento per id/geometria cella raggruppa i dati e consente il conteggio delle stazioni.

Query complessiva con il calcolo dell'indice offerta/domanda

SELECT

```
reticolo5km.id,  
reticolo5km.geom,  
reticolo_nstaz.nstaz / reticolo_demog.densita AS offerta
```

FROM

```
reticolo5km  
JOIN reticolo_nstaz ON reticolo5km.id = reticolo_nstaz.id  
JOIN reticolo_demog ON reticolo5km.id = reticolo_demog.id
```

Questa query effettua un join tra il reticolo e le due viste precedenti tramite id cella e calcola il rapporto tra numero stazioni e densità demografica.

Trattandosi di una vista con un campo id univoco e un campo di tipo «geometry» è visualizzabile e tematizzabile all'interno del GIS.



Guida di riferimento PostGIS

www.postgis.net/docs/reference.html