



UNIVERSITA' DEGLI STUDI DI TRENTO - DIPARTIMENTO DI ECONOMIA

MAPPE DI PROBABILITÀ DI SITO ARCHEOLOGICO: UN PASSO AVANTI

Giuseppe Espa
Roberto Benedetti
Anna De Meo
Salvatore Espa

Discussion Paper No. 2, 2003

The Discussion Paper series provides a means for circulating preliminary research results by staff of or visitors to the Department. Its purpose is to stimulate discussion prior to the publication of papers.

Requests for copies of Discussion Papers and address changes should be sent to:

Prof. Andrea Leonardi
Dipartimento di Economia
Università degli Studi
Via Inama 5
38100 TRENTO ITALY

Mappe di probabilità di sito archeologico: un passo avanti.

G. Espa♣, R. Benedetti♥, A. De Meo♣, S. Espa♣

♣ *Università di Trento, Via Inama, 5 - Trento, gespa@risc1.gelso.unitn.it*

♥ *Università "G. d'Annunzio", Viale Pindaro, 42 - Pescara*

♣ *Istituto per le Tecnologie Applicate ai Beni Culturali, C.N.R. - Via Salaria
km 29.300, C.P.10, 00016 Monterotondo st. (Roma)*

Riassunto. La necessità di disporre di mappe di probabilità di sito archeologico è un tema attualmente al centro di un rinnovato interesse anche grazie alla disponibilità di una moltitudine di informazioni territoriali gestite ed elaborate con l'ausilio della tecnologia GIS. Gli studi che finora si sono occupati del problema, hanno fatto riferimento a modellistiche di tipo regressivo od autoregressivo che, sebbene centrate sulle peculiarità del fenomeno, si sono dimostrate sensibilmente dipendenti dalla presenza di dati anomali nella stima dei parametri e da ipotesi specifiche sui casi in esame. L'intento del presente lavoro è invece quello di proporre una soluzione più robusta e generalizzabile in linea con i recenti sviluppi nel campo della modellistica non parametrica la quale potrebbe dare un forte impulso all'uso congiunto della statistica e della tecnologia GIS. L'informazione geografica ricavabile dai dati fisici e satellitari relativi all'area di studio rappresenta infatti, grazie al dettaglio che la caratterizza, una fonte ausiliaria insostituibile per spiegare la presenza o meno di un sito in una data posizione della mappa. L'approccio proposto è stato sviluppato ed applicato per l'area test di *Cures Sabini* ottenendo dei risultati abbastanza incoraggianti, soprattutto se letti in termini della estrema facilità di interpretazione dei risultati da parte di non statistici e della riscontrata possibilità di rendere automatico il processo di produzione delle mappe di probabilità. Questa evidenza permetterebbe ad operatori del settore non professionalmente preparati da un punto di vista metodologico, di estendere il modello a situazioni anche molto diverse da quella analizzata.

Parole chiave: Prospezione archeologica, Modelli previsivi, Correlazione spaziale, CART, FACT, GIS.

1. Introduzione

Le indagini archeologiche sul campo possono essere ricondotte a due distinte tipologie fondamentali: il campionamento e la prospezione (Benedetti e Espa, 1995).

Per quanto riguarda la prima delle due tipologie, si procede ad un campionamento archeologico qualora in una regione siano eseguiti degli scavi in accordo ad uno specifico disegno campionario. Tale strategia viene adottata principalmente quando l'area di studio è così vasta da non poter essere investigata nella sua totalità. Quindi le strategie campionarie sono motivate dall'esigenza di massimizzare le informazioni deducibili dalle campagne a terra, minimizzando contestualmente i costi che le indagini comportano. Detta esigenza (criterio dell'*efficienza*) si accompagna al desiderio di disporre di un campione che sia rappresentativo della popolazione oggetto di studio (criterio della *totalità*).

Con riferimento alla seconda tipologia di indagine richiamata, si definisce prospezione archeologica un'indagine eseguita avendo come scopo un prefissato obiettivo quale, ad esempio, l'individuazione di una particolare tipologia di siti. Per ciò che concerne la prospezione archeologica, il criterio dell'*efficienza* continua a valere nel senso che la minimizzazione dei costi (tempo, risorse, lavoro) rimane un obiettivo da perseguire. Il criterio della *totalità* diviene però irrilevante. Non è necessario tanto disporre di un campione rappresentativo del totale quanto piuttosto ottenere con il minimo sforzo, i "migliori" esempi per la ricerca che si sta svolgendo.

Mentre la letteratura riguardo alle metodologie di campionamento archeologico è molto voluminosa (cfr., ad esempio, Heizer e Graham, 1967; Mueller, 1974; Redman, 1974; Watson *et al.*, 1971), la prospezione archeologica è stata, tranne alcune rare eccezioni tra le quali si distingue il lavoro di Zubrow e Harbaugh (1978), quasi totalmente ignorata dagli addetti ai lavori.

Appartiene al filone delle tecniche di prospezione archeologica anche la classe dei modelli statistici votati alla previsione della posizione dei siti archeologici sul territorio. Si tratta, come si vedrà meglio nel prosieguo, di una branca di fondamentale importanza e di sicuro sviluppo nell'ambito della ricerca archeologica. A tale riguardo, nel presente lavoro si utilizzerà una classe di modelli che sono, a parere degli autori, in grado di prevedere la probabilità di presenza di siti archeologici in alcune porzioni di territorio.

Tali modelli trovano un naturale supporto alla loro implementazione nell'utilizzo dei Sistemi Geografici Informativi (GIS).

La struttura del lavoro è dunque la seguente. Dopo un paragrafo iniziale dedicato a giustificare l'utilizzo dei modelli previsivi in archeologia, nella sezione 3. si darà risalto alla classe di metodologie che si è deciso di utilizzare nella pratica (le tecniche CART – *Classification And Regression Trees*) e soprattutto alle ragioni di tale scelta. Nel paragrafo 4. verrà descritto in modo puntuale il caso di studio preso in esame che riguarda l'area archeologica di Cures Sabini, situata nella media valle del Tevere. Con riferimento a questa realtà insediativa, la sezione 5. conterrà una presentazione critica dei principali risultati ottenuti dall'implementazione di un modello CART al fine di prevedere la probabilità di sito archeologico di tipologia abitativa. Infine, la sezione 6. sarà dedicata ad alcune brevi considerazioni conclusive.

2. Modelli previsivi della probabilità di sito archeologico

I modelli previsivi sono delle formulazioni, deduttive o induttive, poste in essere allo scopo di prevedere degli eventi non noti. Nello specifico degli studi archeologici è di primaria importanza prevedere la posizione dei siti. Infatti, sono stati scoperti solo una parte dei siti esistenti mentre una grande quantità di siti vengono distrutti ogni anno nel mondo per fare spazio allo sviluppo delle società moderne. Uno dei modi per aiutare la conoscenza e proteggere i siti a rischio, è quello di costruire modelli formali in grado di prevedere dove verosimilmente siano localizzati i siti stessi. A tale scopo possono essere utilizzati modelli deduttivi, ricavati direttamente dalla teoria, oppure modelli induttivi. Questi ultimi si basano fondamentalmente sullo studio di alcune regolarità individuate in un insieme di osservazioni empiriche. Detto studio viene condotto in modo da ottenere indicazioni circa le caratteristiche archeologiche delle zone non direttamente osservate.

Nelle applicazioni pratiche, in ogni modo, è necessario far uso sia della teoria che delle osservazioni empiriche. Ad esempio, modelli di localizzazione di siti archeologici possono essere sviluppati utilizzando un approccio deduttivo che tenga principalmente conto delle esigenze culturali e biologiche di una società. Tali bisogni possono governare la scelta delle variabili indipendenti e condurre così alla costruzione di modelli che non necessitano di informazioni circa la posizione geografica dei siti. Detti

modelli non possono però essere implementati né testati senza osservazioni empiriche. Allo stesso modo si può costruire uno schema puramente induttivo di localizzazione dei siti archeologici ma, in assenza del supporto teorico, il processo di selezione delle variabili diviene inefficiente ed il modello stimato debole e difficilmente interpretabile (Espa e Arbia, 1995).

Nella ricerca archeologica i modelli previsivi solo di recente si sono affermati come componente fondamentale (Kohler e Parker, 1986; Arbia e Espa, 1996 e 1997). Una delle ragioni del successo di tale classe di modelli è da ricercarsi nel fatto che i siti archeologici tendono a presentarsi in ambienti dalle condizioni favorevoli. I modelli previsivi traggono vantaggio da questo genere di informazioni e, sfruttando le differenze tra le caratteristiche ambientali delle zone dove sono presenti i siti e dove invece i siti sono assenti, aiutano lo studioso nell'identificazione di aree in cui è elevata la probabilità di sito archeologico (Warren, 1990a e 1990b).

Uno degli approcci empirici più diffusi alla modellizzazione previsiva è la strategia integrata sviluppata da Kvamme (1983 e 1989), da lui approfondita e perfezionata (Kvamme, 1992; Stančić e Kvamme, 2003), ed utilizzata da numerosi altri ricercatori fra cui Warren (1990a e 1990b) e Warren *et al.* (1987). In estrema sintesi tale strategia integra le enormi potenzialità dei GIS nel creare e processare grandi insiemi di dati con la regressione logistica, classico e flessibile strumento statistico che consente di prevedere la probabilità che un sito archeologico si trovi in una determinata porzione di territorio.

Schematicamente si può procedere applicando i seguenti passi. L'area di studio, discretizzata in M celle quadrate contigue (*pixel*), è tale che per $N < M$ di queste celle la campagna di scavo ha potuto stabilire la presenza o l'assenza di un sito archeologico. In base a tali considerazioni si definisce la seguente variabile binaria:

$$Y_i = \begin{cases} 1 & \text{se la cella } i \text{ contiene un sito} \\ 0 & \text{altrimenti} \end{cases}, i=1,2,\dots,N.$$

Inoltre, per tutte le M celle sono disponibili (per osservazione diretta o per interpolazione; cfr., al riguardo, Benedetti e Espa, 1995) k variabili ausiliarie ambientali (qualitative e quantitative) $\mathbf{X}_i = (X_{1i}, X_{2i}, \dots, X_{ki})$, $i=1,2,\dots,M$, da utilizzarsi per prevedere la probabilità che in ogni *pixel* si presenti un sito archeologico. Le variabili indipendenti, come verrà illustrato

in dettaglio nella sezione 4., forniscono informazioni sulla natura del terreno, possono essere di natura topografica nonché urbanistica e possono includere informazioni ausiliarie quali misure geofisiche, geochimiche ed immagini acquisite mediante telerilevamento da satellite e/o aereo.

La fase successiva alla raccolta dei dati consiste nello stimare un modello logistico lineare su un sottocampione di n *pixel* (campione di addestramento; cfr. sezione 3.) per il quale sia disponibile l'informazione di presenza/assenza di sito. Come è noto, il modello da stimare è il seguente:

$$E(Y_i) = \theta_i = \frac{e^{x_i^T \beta}}{1 + e^{x_i^T \beta}}, i = 1, 2, \dots, n \quad [1]$$

dove $\theta_i = P(Y_i = 1)$ e β è un vettore di parametri incogniti (per maggiori dettagli si veda Cox e Snell, 1989).

Una volta selezionato e stimato il modello “migliore” fra quelli posti a confronto, le sue capacità previsive vengono testate sul campione *test* costituito dalle rimanenti $N-n$ celle osservate nella campagna di scavo. Al termine del processo di *cross-validazione*, il modello può essere applicato per prevedere la probabilità di presenza/assenza di sito nelle zone “non scavate” producendo come *output* una mappa di probabilità di sito.

3. La metodologia utilizzata: *Classification And Regression Trees* (CART)

Modelli previsivi come quello sinteticamente richiamato nella sezione precedente a volte presentano alcuni problemi tecnici che ne possono limitare l'uso (Warren, 1990b).

Comunque, gli studi precedenti condotti dal gruppo di lavoro misto Università-C.N.R. impegnato su tali problematiche hanno riguardato principalmente il tentativo di individuare un modo proficuo di superare il limite più grave che inficia l'uso della regressione logistica: l'ipotesi di indipendenza delle variabili Y_i trattate senza tener conto della loro natura spaziale. Infatti, come accade in tutti gli studi a carattere geografico in cui le osservazioni sono reciprocamente dipendenti, anche i siti archeologici

tendono a *clusterizzare* nello spazio. L'immediata conseguenza di ciò è una più alta probabilità di sito nelle vicinanze di siti rilevati.

Ragionando sulle conseguenze, anche statistiche (Cressie, 1991; Arbia, 1993), che ignorare tale informazione contestuale comporta, il gruppo di lavoro si è mosso cercando di migliorare l'approccio integrato di Kvamme (1983 e 1989).

Le ricerche svolte (si veda, fra i lavori prodotti, Espa e Arbia, 1995) hanno inizialmente imposto di incorporare la nozione di informazione contestuale in un modello per variabili binarie seguendo un approccio di tipo auto-logistico alla Besag (1974). In termini formali il modello usato è il seguente:

$$P(Y_i = y_i | y_j, j \in N(i)) = \frac{\exp \left[y_i \left(\alpha_i + \sum_{j \in N(i)} \beta_{ij} y_j \right) \right]}{\left[1 + \exp \left(\alpha_i + \sum_{j \in N(i)} \beta_{ij} y_j \right) \right]}, \quad i = 1, 2, \dots, n \quad [2]$$

dove $N(i)$ rappresenta l'insieme delle celle vicine all' i -mo *pixel* e dove gli α_i e i β_{ij} sono parametri da stimare. In particolare, gli α_i sono parametri locali e possono essere utilizzati per modellizzare *trend* superficiali o l'influenza di altre variabili mentre i β_{ij} regolano il livello di interazione spaziale ossia l'intensità con la quale le condizioni del contesto locale modificano le "aspettative di sito archeologico". Per la stima dei parametri della [2] l'uso della massima verosimiglianza è precluso. Si può però fare ricorso a due metodi di stima proposti da Besag: una tecnica di codifica dei siti (Besag, 1974) ed una semplice tecnica di pseudo-verosimiglianza (Besag, 1975). Fra le due è senz'altro da preferirsi la seconda dal momento che sfrutta tutti i dati a disposizione e fornisce uno stimatore consistente (Besag, 1977). L'analisi previsiva si può articolare nelle stesse fasi descritte nella sezione 2..

Anche il modello [2] però è stato giudicato inadeguato per gli scopi prefissati a causa della sua incapacità di tenere in considerazione le informazioni ausiliarie disponibili per ogni *pixel* in cui si è discretizzata l'area di studio. Pertanto il naturale passo successivo è stato quello di proporre un approccio congiunto logistico/autologico (Arbia e Espa,

1996) che incorporasse sia l'informazione contestuale che quella ausiliaria. Il modello *log/autolog* può essere espresso nei seguenti termini formali:

$$\theta_i = P(Y_i = 1 | y_j, j \in N(i); \mathbf{X}) = \frac{\exp\left(\beta_0 + \sum_{j \in N(i)} \beta_{ij} y_j + \mathbf{X}^T \boldsymbol{\gamma}\right)}{\left[1 + \exp\left(\beta_0 + \sum_{j \in N(i)} \beta_{ij} y_j + \mathbf{X}^T \boldsymbol{\gamma}\right)\right]}, i = 1, \dots, n \quad [3]$$

dove β_0 è una costante, i parametri β_{ij} sono i coefficienti di interazione spaziale analoghi a quelli visti per la [2], \mathbf{X} è un vettore di variabili indipendenti come quello impiegato nel modello [1] e $\boldsymbol{\gamma}$ la parametrizzazione associata.

Senza dilungarci su tale tipo di modellistica, si rimanda ad Arbia e Espa (1996) per ogni approfondimento e soprattutto per le tecniche di stima.

La proposta formulata nel presente lavoro, la quale accantona i risultati connessi alla [3], potrebbe sembrare un'abiura da tali studi ed un'accettazione incondizionata dell'approccio integrato e sostanzialmente aspatiale proposto da Kvamme (1983 e 1989). In realtà così non è, come verrà mostrato nel prosieguo, e alla luce di ciò sarà meglio affrontare la discussione procedendo per punti.

Innanzitutto vediamo come e perché si è fatto ricorso ad un tipo di metodologia statistica più tradizionale e solo successivamente si discuterà sul fatto che l'approccio seguito nel presente lavoro non esclude il ricorso alla modellistica di stampo spaziale. Quello che qui si è fatto è stato adottare un approccio di classificazione al problema della costruzione di mappe di probabilità di sito archeologico. Fra le numerose tecniche disponibili nella letteratura specialistica (si vedano, per citare solo due riferimenti classici tra i tanti volumi disponibili sull'argomento, Gordon, 1981; Hand, 1981) si è deciso di seguire un approccio di tipo CART, il quale, lungi dall'essere una novità per gli addetti ai lavori, rappresenta piuttosto uno *standard* nella ricerca statistica applicata. A tale riguardo, nel presente paragrafo si suporrà familiarità da parte del lettore con le metodologie e la terminologia proprie degli *alberi di classificazione e regressione*. Questo *background* può essere agevolmente maturato mediante la lettura del manuale di Breiman *et al.* (1984) che costituisce la pietra miliare ma anche il riferimento definitivo

sull'insieme di tecniche note come CART. Comunque prima di procedere oltre è necessario il breve *review* tecnico che segue.

Ogni punto da classificare appartiene ad un campione scavato all'interno della zona indagata (cfr. paragrafo 4.) ed ha associato un vettore \mathbf{x} di misurazioni il quale, almeno in linea teorica, può anche presentare covariate mancanti. Tale circostanza è invero molto frequente nella maggior parte degli studi archeologici. D'ora in avanti, il generico vettore delle osservazioni relative al sito i -mo sarà indicato con $\mathbf{x}_i = (x_1, x_2, \dots)$. Si supponga, poi, che i casi oggetto di studio possano essere attribuiti a J classi numerate mediante gli interi $1, 2, \dots, J$; sia infine $C = \{1, 2, \dots, J\}$ l'insieme delle possibili etichette. Nello specifico dell'applicazione condotta l'insieme C è composto da due soli interi corrispondenti alle etichette "sito" e "non sito". Nel prosieguo di questo breve *review* si manterrà un maggiore livello di generalità in modo da fornire un *frame* globale in cui inquadrare il problema.

Per assegnare una classe ad una unità è necessario stabilire una regola di classificazione, o *classificatore*, cioè una funzione $d(\mathbf{x})$, definita sullo spazio X di tutti i possibili vettori delle misurazioni, tale da essere eguale $\forall \mathbf{x}$ ad uno dei numeri $1, 2, \dots, J$. Nel CART di Breiman *et al.* (1984) il classificatore si presenta nella forma di albero di decisione binario. La costruzione di un classificatore ad albero è basata fondamentalmente sull'esperienza acquisita nel passato¹, esperienza sintetizzata in un campione cosiddetto *di apprendimento*. Si tratta di N casi per i quali, oltre ad essere state condotte le misurazioni, sono note anche le singole classi di appartenenza. Quindi un campione di apprendimento consiste delle coppie $(\mathbf{x}_1, j_1), (\mathbf{x}_2, j_2), \dots, (\mathbf{x}_N, j_N)$ relative ad N casi dove $\mathbf{x}_n \in X$ e $j_n \in \{1, 2, \dots, J\}$, $n=1, 2, \dots, N$. Nel prosieguo tale insieme di dati sarà indicato con $\ell = \{(\mathbf{x}_1, j_1), (\mathbf{x}_2, j_2), \dots, (\mathbf{x}_N, j_N)\}$. Da notare come il vettore \mathbf{x} anche per i casi appartenenti ad ℓ può presentare dati mancanti.

L'albero viene costruito partizionando in modo ricursivo il campione d'apprendimento ℓ cosa che induce un corrispondente partizionamento dello spazio \mathbf{x} . Ad ogni passo il campione ℓ viene bipartito sulla base di

¹ Per attribuire una classe ad una unità è necessario stabilire una regola di classificazione sulla base sia delle informazioni presenti nel campione di apprendimento ma anche delle probabilità a priori e dei costi di errata classificazione stabiliti dall'utente.

risposte fornite a domande del tipo “è $x_i \leq c$?” oppure, per variabili qualitative, “è $x_i \in A$?”.

In estrema sintesi l’algoritmo di Breiman *et al.* (1984) esegue una ricerca (globale ma essenziale) su tutte le costanti c ed i sottoinsiemi A per individuare il miglior *split* sulla base di misure di bontà di bipartizione che valutano essenzialmente il decremento in impurità ΔI indotto dalla bipartizione.

Lo *splitting* verso il basso si arresta al livello del generico nodo t se questo è “puro” (quando tutti i casi che lo compongono afferiscono alla stessa classe) o contiene meno di un numero di casi prestabilito; al limite, il processo di bipartizione è impossibile quando un nodo è costituito da un solo caso.

Quando un nodo viene riconosciuto come terminale bisogna stabilire come classificare i casi che in esso sono contenuti. Un semplice criterio al riguardo consiste nell’invocare una regola di pluralità: al nodo si assegnerà l’etichetta del gruppo maggiormente rappresentato.

Storicamente la costruzione della prima generazione di alberi di classificazione avveniva *splittando* ogni nodo finché non smetteva di essere soddisfatto un qualche criterio di bontà di bipartizione. Quando la qualità di un particolare nodo cadeva sotto una certa soglia, il nodo raggiunto veniva dichiarato terminale. Allorché, inoltre, tutti i rami uscenti dal nodo radice raggiungevano dei nodi terminali, il processo di costruzione dell’albero veniva considerato ultimato.

La filosofia del CART è completamente diversa. Invece di cercare di stabilire se un nodo è terminale o meno, il CART procede diramando l’albero finché è possibile ed esaminando sotto-alberi ottenuti dalla potatura dei rami dell’albero di partenza. Tra tutti i sotto-alberi annidati possibili si sceglierà il migliore in base ad un criterio, detto del *minimo costo-complexità*, che combina il tasso di errore apparente con la dimensione dei sotto-alberi e che è basato su di un processo di *cross-validazione* (CV). Un *by-product* della procedura sommariamente delineata è una stima, non apparente come quella interna di risostituzione ma più *onesta*, del tasso di errata classificazione (stima di CV)².

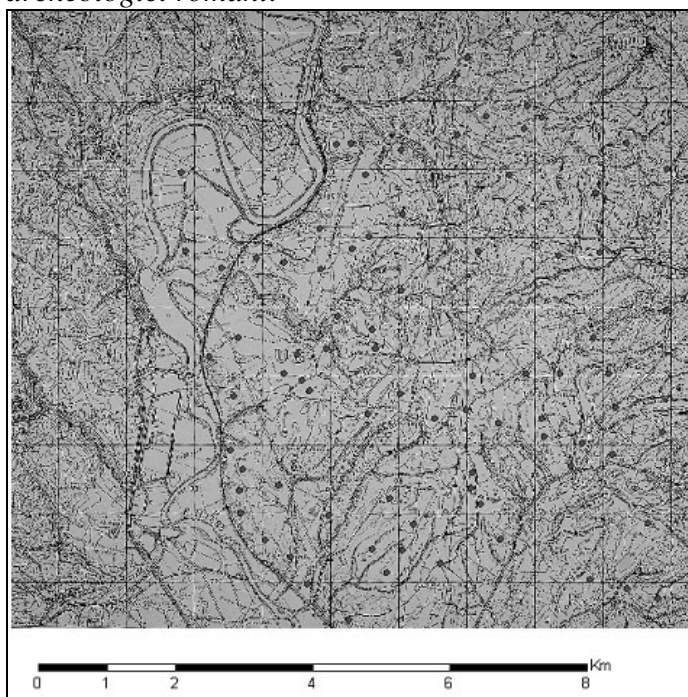
² Oltre allo *split* migliore, il CART valuta anche a livello di ogni nodo delle bipartizioni *surrogate* (o alternative) basate su tutte le altre variabili non coinvolte nello *split* primario. Sulla base dei *surrogati* viene costruito il *ranking* dei predittori di cui si parlerà nella sezione 5..

Con il CART classificare nuovi punti campione per i quali siano disponibili le sole informazioni x (cfr. paragrafo seguente) è un'operazione estremamente semplice e rapida che, fra l'altro, non richiede più l'uso del campione di apprendimento!

4. La zona indagata: l'area di Cures Sabini

L'area presa in esame è quella di Cures Sabini (tavoleta IGM di Montopoli Sabina), dove sono stati individuati siti archeologici di epoca romana e di tipologia abitativa (cfr. Fig. 1). In dettaglio sono stati presi in considerazione 88 siti (Muzzioli, 1980 e 1985). Particolare cura è stata posta nella selezione dei dati per evitare tra l'altro duplicazioni tali da poter inficiare i risultati delle analisi effettuate. In altri termini, le serie sono state depurate da pertinenze di singoli siti abitativi (es. magazzini, apparati produttivi etc.) o da riuso del sito medesimo in epoca storica successiva.

Figura 1. *Cures Sabini: topografia con sovrapposizione dei siti archeologici romani.*



Inoltre, rispetto all'area di Cures nel suo insieme, tali presenze sono situate esclusivamente sulla riva sinistra del Tevere e sono riferibili ad un periodo storico che va dal IV° secolo A.C. fino II° secolo P.C. Questa scelta è dettata in primo luogo da argomentazioni di tipo archeologico, secondo le quali gli insediamenti a destra del Tevere rispondono a logiche differenti rispetto a quelli di sinistra, in quanto si sono insediati su una situazione precedente piuttosto diversa. Infatti il Tevere è considerato la linea di separazione tra la Sabina a sinistra e l'Agro Capenate a destra (Muzzioli 1980 e 1985).

Infine, ci sono motivi legati anche a fattori di omogeneità, nel senso che la presenza di un elemento naturale così incisivo come il Tevere, rende l'area non omogenea per gli studi che si intendono sviluppare e quindi, i risultati potrebbero risultare falsati. Per questo motivo, l'area d'indagine è stata delimitata in modo tale che il fiume costituisca una barriera naturale lungo il lato ovest.

Nell'ambito delle ricerche territoriali riguardanti la media valle del Tevere ed orientate allo studio di fenomeni legati al patrimonio storico archeologico di tale zona, è stato implementato un Sistema Informativo Territoriale (GIS) integrato per l'analisi di dati sia di tipo *raster* che vettoriale (De Meo *et al.*, 1996 e 2000). Tale sistema informativo è stato quindi organizzato inserendo sul *data base* cartografico innanzitutto la cartografia di base della zona. Più precisamente circa cinquanta tavolette in scala 1:25.000 dell'I.G.M. sono state digitalizzate in formato *raster*, mentre le infrastrutture e l'idrografia principale sono state digitalizzate in formato vettoriale. Successivamente, è stata inserita l'orografia numerica in scala 1:25.000 dell'I.G.M. e la geologia regionale in scala 1:100.000. Sono stati inoltre inseriti nel G.I.S. dati geometrici e descrittivi di natura archeologica, in parte desunti da letteratura, in parte ottenuti da ricognizioni sul terreno. Infine, due coperture aeree della zona, rispettivamente della R.A.F. 1945 e della società Sara-Nistri 1979 georeferenziate e opportunamente mosaiccate, sono state inserite nel GIS insieme ad un'immagine da piattaforma orbitante LANDSAT del 1988, nonché la cartografia tematica dell'uso del suolo del Lazio a scala 1:300.000 del progetto "Coriner Land Cover"³.

Il caso preso in esame per il presente studio è quello relativo all'area archeologica di Cures Sabini (De Meo *et al.*, 2002). Questa insiste in una zona storicamente significativa delimitata a nord dal fiume Nera, a sud

³ Dati messi a disposizione dal Ministero dell'Ambiente, S.C.N.

dall'Aniene, ad ovest dal Tevere ed a est dai monti Sabini e definita "Sabina Tiberina". Si tratta di un'area piuttosto vasta ed articolata di cui sono state già indagate diverse località tra cui Eretum, Poggio Somavilla e Magliano Sabina (De Meo *et al.*, 2000; Espa *et al.*, 1995; Colosi *et al.*, 1999).

Nel GIS così organizzato per la Sabina Tiberina sono stati quindi immessi nuovi dati ambientali e altre informazioni di natura strettamente archeologica relativi a Cures Sabini. Quindi, oltre alla cartografia di base (topografia, geologia ed orografia) sono state georiferite e mosaiccate altre fotografie aeree, effettuate da voli in periodi temporali differenti (RAF - 1944, Esacta Company - 1965, Volo Base -1989). E' stata altresì elaborata l'immagine da satellite Landsat precedentemente georiferita e sono state inserite altre variabili, tra cui informazioni archeologiche riguardanti la distribuzione dei siti archeologici, le strade antiche e la ricostruzione della suddivisione agraria romana (Muzzioli, 1980 e 1985) nonché parte della cartografia tematica del Progetto Atlas del Ministero dei BB.CC.⁴

Al fine di stimare un modello previsivo della probabilità di sito archeologico, l'insieme di questi dati è stato opportunamente trattato per la determinazione di alcune variabili che intervengo in fase di stima del modello stesso.

In primo luogo, quindi, sono state create, attraverso procedure GIS di immissione dati e di georeferenziazione, i tematismi di presenza ed assenza di sito, costituiti rispettivamente da 88 e 42 punti (cfr. Fig. 2). Tali coperture sono state quindi successivamente trattate con i livelli delle variabili (cfr. Fig.3).

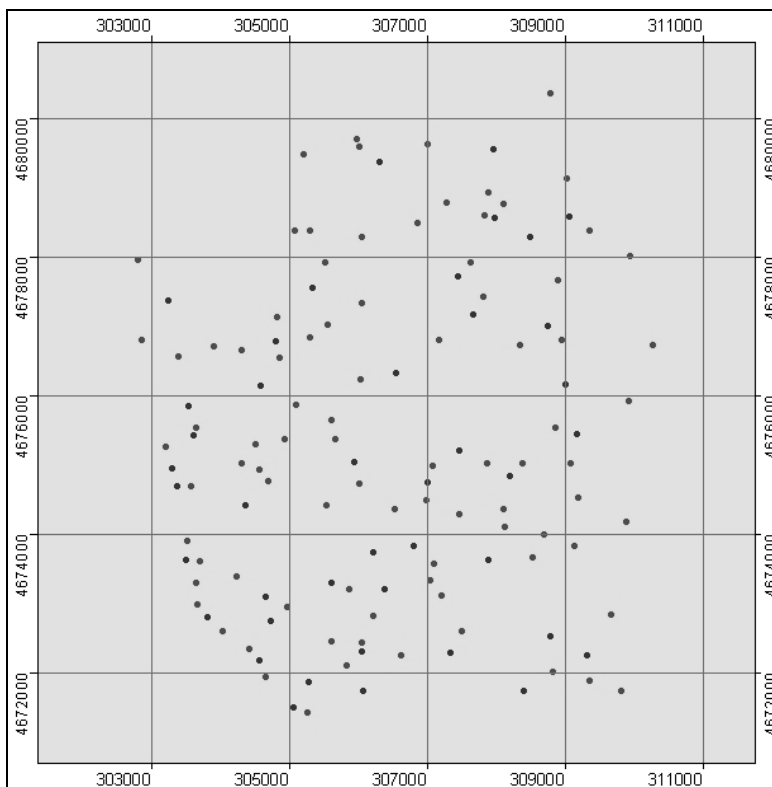
Alcune delle variabili esaminate sono quelle relative alla topografia ed alla morfologia del terreno, intese come quota sul livello del mare e come valori di pendenza e di esposizione dei versanti. Quindi sono state prese in considerazione altre variabili legate all'ambiente naturale, quali la litologia e l'idrografia, e altre relative all'uso del suolo.

Per generare le mappe di distribuzione delle quote altimetriche, quella delle pendenze e quella delle esposizioni (cfr. Figg. 4, 5 e 6), sono stati utilizzati i dati orografici numerici dell'IGM. Quest'ultimi, inizialmente in formato grafico, sono stati trasformati per essere utilizzati all'interno del

⁴ Parte del materiale fotografico (voli RAF ed Esacta) e la cartografia tematica del progetto Atlas è stata fornita dall'I.C.C.D. Ministero BB.CC. all'I.T.A.B.C. - C.N.R. nel quadro di una collaborazione tra i due Istituti e con l'interessamento dalla direzione del Progetto Finalizzato BB.CC.

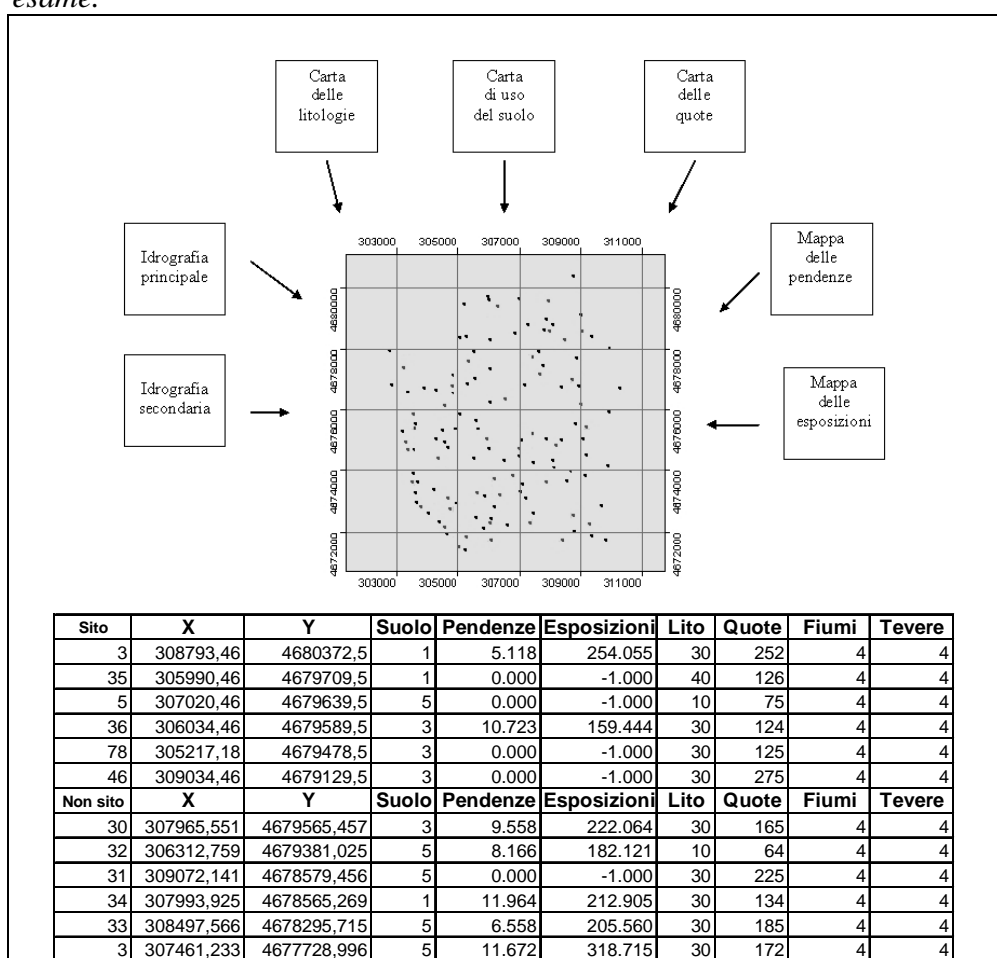
GIS allo scopo di estrarre il modello digitale del terreno. In sostanza, è stata costruita, tramite un processo di triangolazione di punti quotati e curve di livello, una struttura vettoriale di tipo TIN (Triangular Irregular Network) (cfr. Fig. 7) che fornisce informazioni relativamente alle quote, alle pendenze ed alle esposizioni (Gomarasca, 1997). Quest'ultima quindi, opportunamente trasformata in copertura poligonale con associati gli attributi relativi a quota, pendenza ed esposizione, può essere impiegata ai fini della costruzione del modello di cui al paragrafo seguente. Tale passaggio avviene attraverso delle operazioni di *overlay* topologico con i livelli di presenza ed assenza di sito, da cui si ottengono delle coperture di punti e a cui risultano associati, oltre agli attributi che possedevano prima dell'*overlay*, anche quelli della copertura con cui si sono intersecati.

Figura 2. *Cures Sabini: carta dei punti di presenza e assenza di sito archeologico*



Per quanto riguarda le quote, prima di effettuare la sovrapposizione con le carte dei siti e dei non siti, è stata necessaria un'ulteriore elaborazione del TIN. Questo va infatti convertito in un formato griglia (GRID) attraverso un processo di interpolazione di punti (Benedetti e Espa, 1995) e successivamente è possibile generare una copertura in cui ad ogni poligono è associato un valore di quota. Sono state inoltre effettuate altre elaborazioni per i tematismi relativi alla distribuzione delle litologie (cfr. Fig. 8) ed all'uso del suolo (cfr. Fig. 9) i quali sono stati ricavati dal trattamento dell'immagine da satellite LANDSAT.

Figura 3. Schema delle intersezioni tra carta dei punti e variabili prese in esame.



Nel dettaglio, per quanto riguarda la carta delle litologie sono stati utilizzati congiuntamente sia dati satellitari che informazioni desunte da carte geologiche e da fotografie aeree, tutti preventivamente inseriti nel sistema informativo. Sono stati quindi applicati degli algoritmi per enfatizzare la risposta spettrale in particolare delle bande relative al rosso (TM3) e all'infrarosso vicino (TM4), che sono tra quelle maggiormente utilizzate per gli studi ambientali. In sostanza, la combinazione tra le due bande presa in considerazione è stata quella che viene definita come differenza dell'indice di vegetazione normalizzato (NDVI). Da ciò si ricava un'immagine rapporto in cui i vari livelli di grigio corrispondono alle massime differenze di riflettività spettrali tra le bande analizzate. Questo tipo di operazione permette di distinguere la distribuzione della vegetazione, che costituisce spesso un fattore legato alla distribuzione delle litologie presenti.

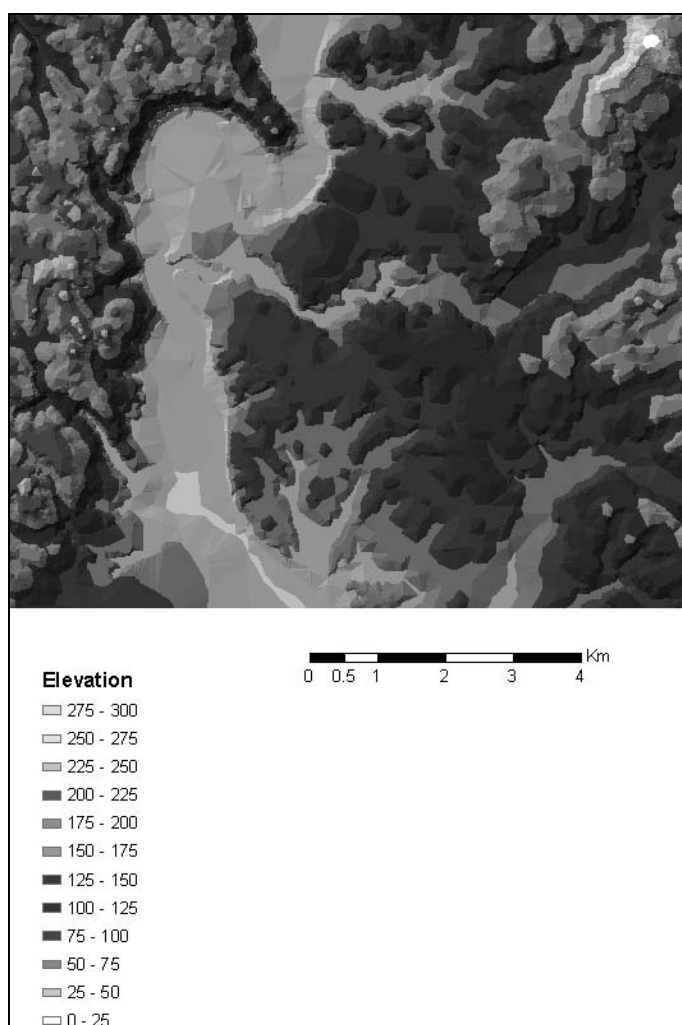
Successivamente, per desumere la carta delle litologie, è stata effettuata una classificazione dei pixels sull'immagine rapporto NDVI congiunta con la TM6 del LANDSAT. Quest'ultima è peraltro tra le più adatte per studi di carattere geologico. La classificazione dei pixels è stata eseguita attraverso una tecnica con supervisore a partire da aree preselezionate sull'immagine con il supporto della cartografia geologica di riferimento (De Meo *et al.*, 2002; Richards, 1986).

Per quanto riguarda la carta di uso del suolo invece, l'immagine da satellite è stata elaborata per identificare le principali configurazioni al suolo. Tale trattamento è stato effettuato tramite una tecnica di classificazione senza supervisore utilizzando l'algoritmo "ISODATA" (GIS ERDAS IMAGINE) applicato sulle bande 2, 3 e 4 che sono tra le più utilizzate per studi di tipo ambientale. In particolare la numero 2 è quella che corrisponde alla lunghezza d'onda del verde ed è quindi utile per misurare il picco di riflessione della vegetazione a quel valore di λ . La numero 3 si usa per identificare le diverse classi di vegetazione ed in corrispondenza della 4 si registrano i valori massimi di riflessione della vegetazione per l'infrarosso vicino. Dall'analisi di queste tre bande si è ottenuta una carta di uso del suolo, distinguendo cinque classi, ovvero aree boschive, aree coltivate, aree a frutteti e oliveti, masse d'acqua e regioni urbane (De Meo *et al.*, 2000; Richards, 1986).

Successivamente, per utilizzare la carta delle litologie e quella di uso del suolo in fase di stima del modello, tali variabili sono state convertite dal

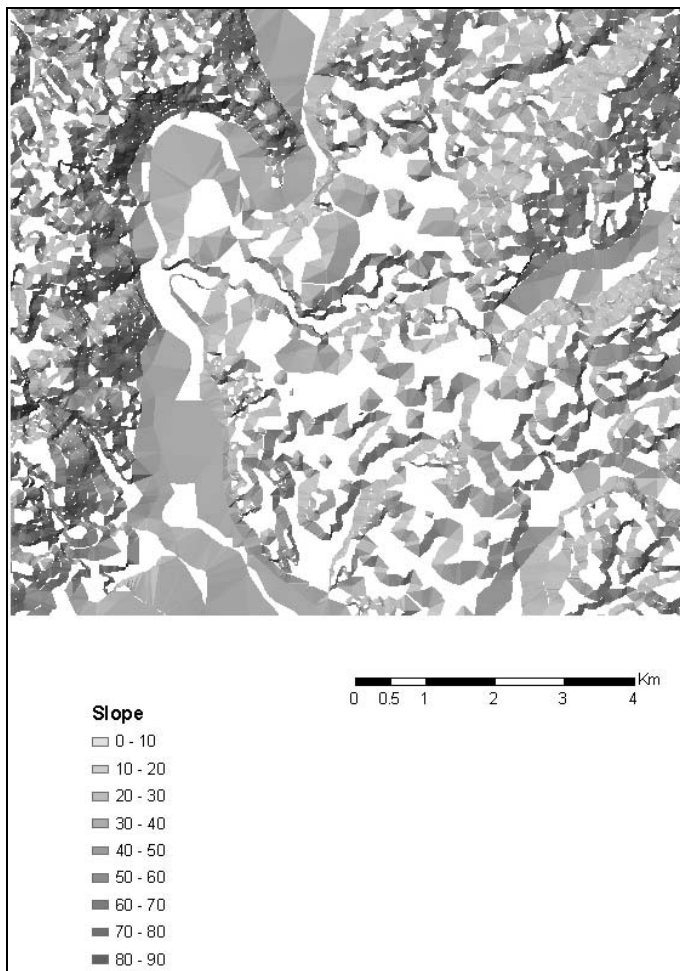
formato griglia a quello vettoriale. Al fine di ottenere una corretta vettorializzazione di tali tematismi è stata necessaria una fase di preparazione sui *raster*. Questa consiste nell'eliminazione dei *pixel* spuri utilizzando dei filtri di tipo modale. Una volta effettuato questo passaggio, è necessario riclassificare l'immagine per ottenere una codifica opportuna delle classi d'interesse.

Figura 4. *Cures Sabini: mappa delle altimetrie.*



L'ultima fase è quella della intersezione topologica delle coperture poligonali ottenute (uso del suolo e litologia) con i livelli vettoriali puntuali dei siti e dei non siti.

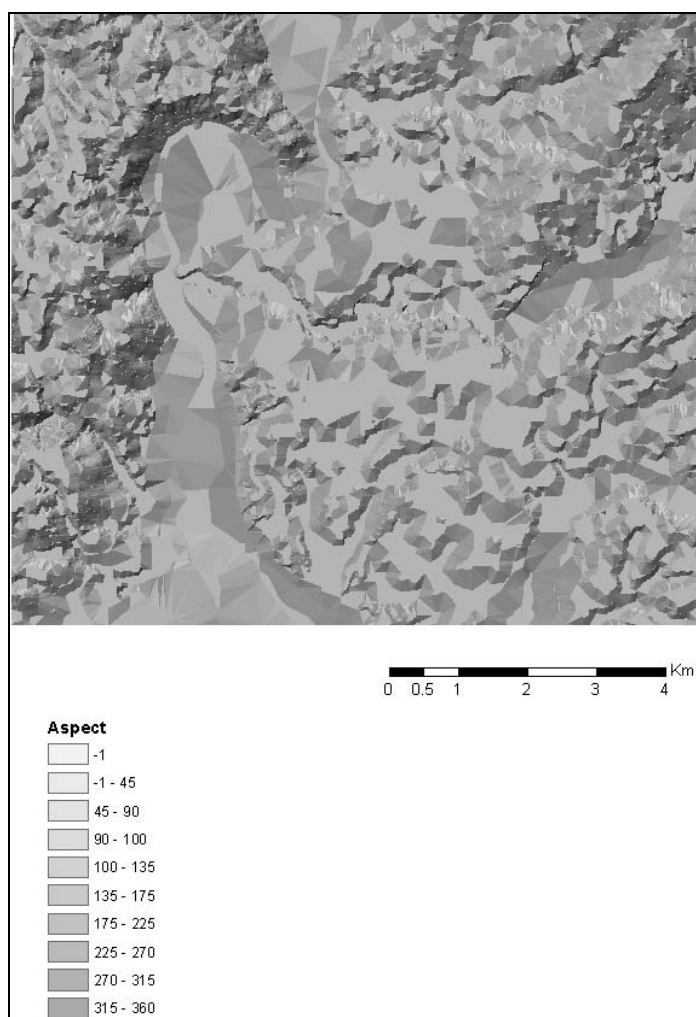
Figura 5. *Cures Sabini: mappa delle pendenze.*



Per maggiore completezza, sono state inoltre individuate non solo la classe di uso del suolo e di litologia del poligono in cui è collocato il punto

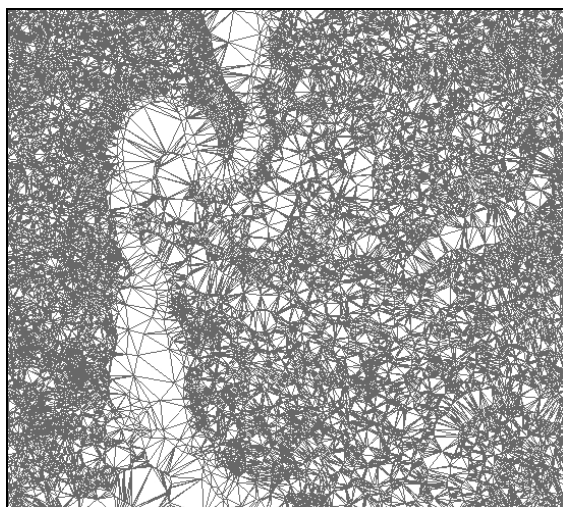
di presenza ed assenza di sito, ma è stata verificata la situazione rispetto a queste variabili in un intorno del punto stesso. Per questo motivo, per ogni punto etichettato come *sito* o come *non sito* è stata individuata un'area circolare con centro nel punto e con raggio di 60 metri.

Figura 6. *Cures Sabini: mappa delle esposizioni.*



Tali intorni sono stati ottenuti tramite operazioni di *buffering* a partire dalle coperture puntuali da cui sono state generate delle coperture poligonali, ovvero i *buffer*. Questi ultimi sono stati intersecati con il tematismo di uso del suolo e quello delle litologie. In questo modo, per ogni punto rispetto alle variabili uso del suolo e litologia, è stato ottenuto, oltre al dato puntuale, anche la situazione al contorno.

Figura 7. *Cures Sabini: struttura vettoriale di tipo TIN.*



Infine, per quanto riguarda la variabile legata all'idrografia, si è proceduto considerando delle classi di distanza dai principali corsi d'acqua presenti nella zona d'indagine ed intesi come risorsa idrica. Sono stati esaminati in un primo tempo tutti i punti rispetto al Tevere, che in quest'area ha un andamento circa nord – sud e poi, i punti rispetto a due affluenti di sinistra del Tevere che sono il Farfa a nord ed il Corese a sud. Dalle fonti storiche è evidente che questi siano i principali corsi d'acqua mentre gli altri sembra siano storicamente meno importanti perché di portata minore e spesso secchi (Muzzioli, 1980).

I tematismi relativi all'idrografia sono stati desunti dai dati Atlas del Ministero dei BB.CC. (cfr. Fig. 10) e sono state quindi individuate quattro fasce rispettivamente a distanza di 100, 200, 300 e >300 metri dal Tevere,

dal Corese e dal Farfa. Anche tali fasce sono state costruite tramite delle operazioni di *buffering* da cui si sono generate delle coperture (cfr. Fig. 11) che sono state successivamente intersecate con le coperture dei siti e dei non siti. Per ulteriori dettagli sulle modalità assunte dalle variabili utilizzate come esplicative nel modello costruito si rimanda alla sezione che segue.

Figura 8. *Cures Sabini: carta delle litologie.*

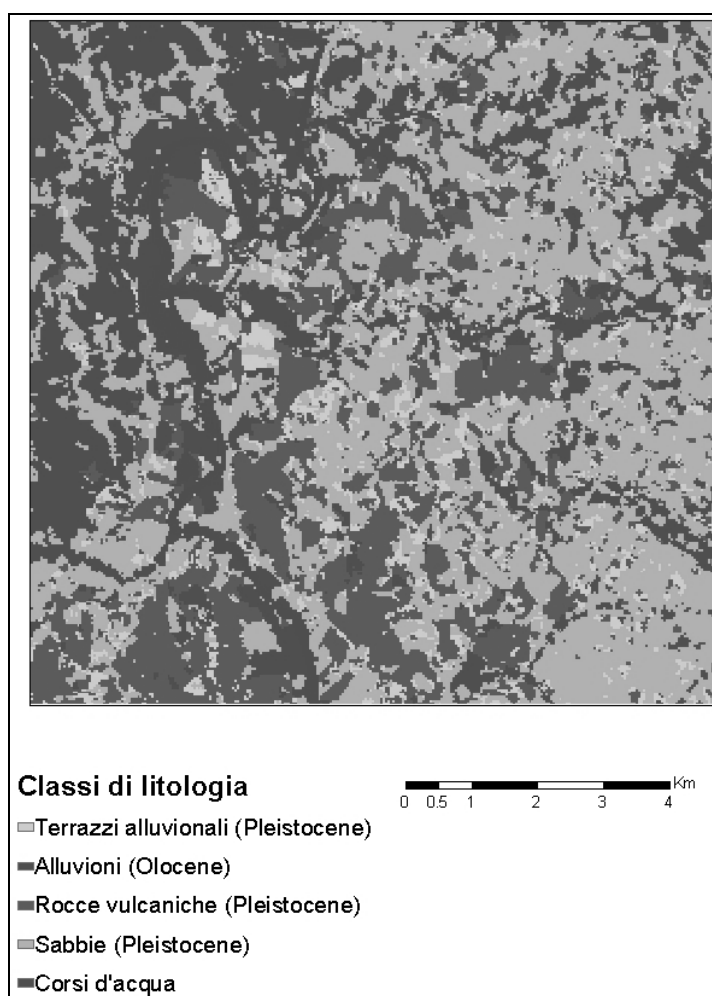


Figura 9. *Cures Sabini: carta di uso del suolo.*

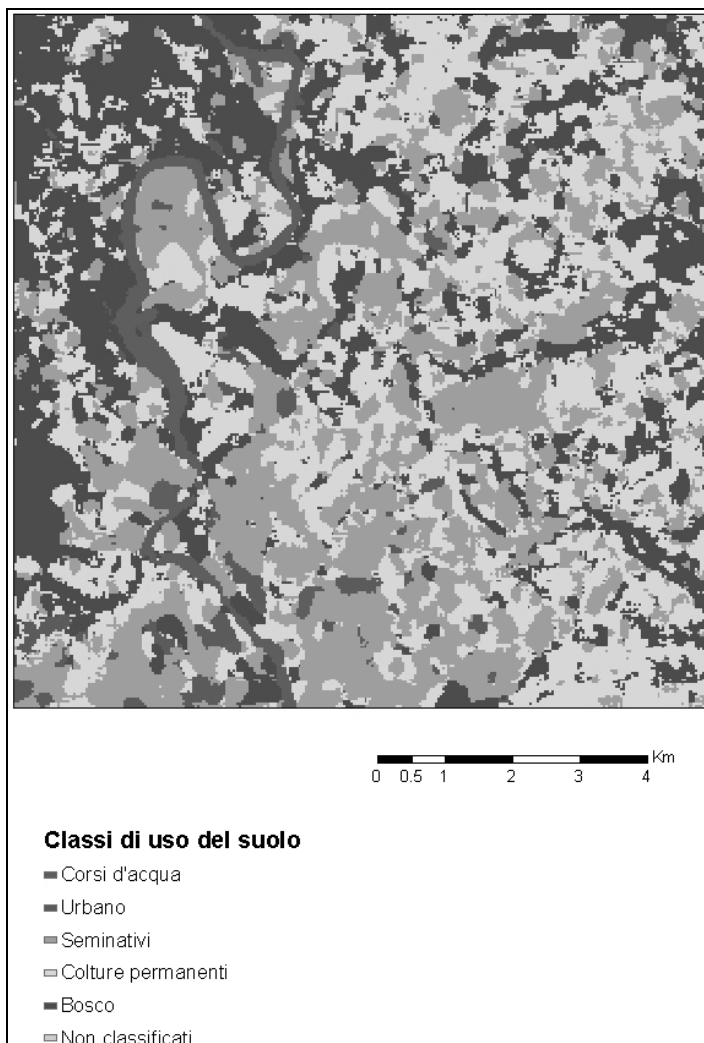


Figura 10. *Cures Sabini: idrografia della zona di indagine.*

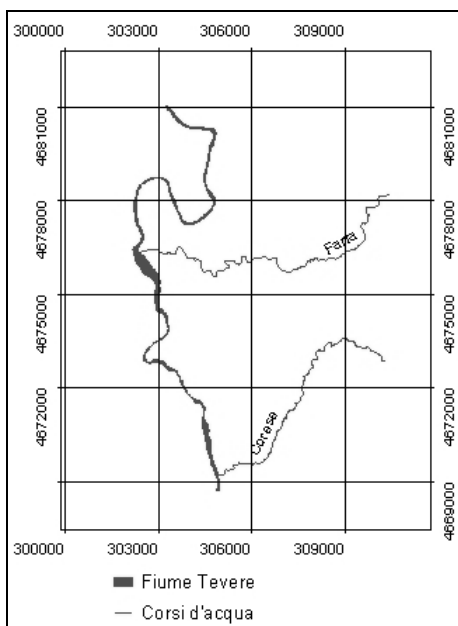
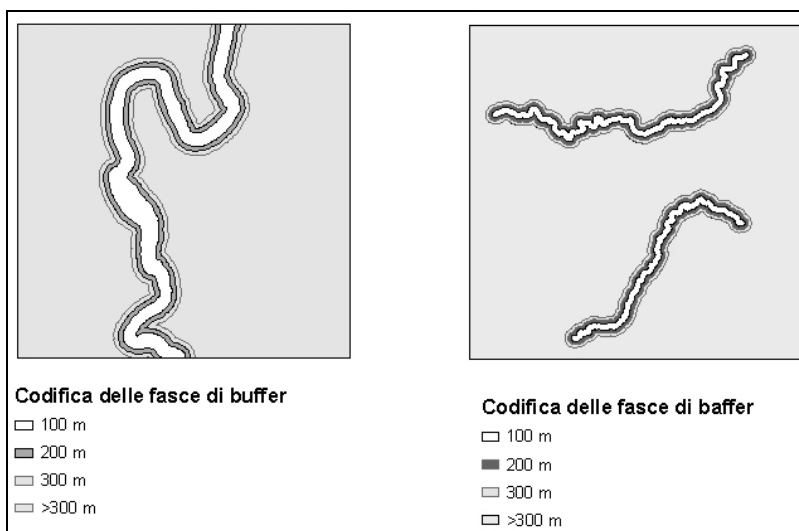


Figura 11. *Cures Sabini: rappresentazione delle fasce di buffer dell'idrografia.*



5. I principali risultati dello studio condotto

L'obiettivo di questa fase dell'analisi è quello di pervenire ad una partizione in gruppi omogenei dei punti campione, a loro volta suddivisi in siti noti e punti in cui si è riscontrata l'impossibilità di presenza del sito, sulla base dell'insieme di variabili esplicative elencate nella Tab. 1⁵.

Tabella 1. Modalità assunte dai principali predittori utilizzati.

Carta di uso del suolo						
Classe	Bosco	Masse d'acqua	Colture permanenti	Urbano	Seminativi	Non classificati
Codifica	1	2	3	4	5	6 / 7

Carta delle litologie					
Classe	Alluvioni	Rocce vulcaniche	Sabbie	Terrazzi alluvionali	Masse d'acqua
Codifica	10	20	30	40	50

Carta delle pendenze		
Valori di pendenza	Pendenza nulla	⇒ Verticalità
Intervallo di valori espressi in gradi	0	⇒ 90

Carta delle esposizioni			
Valori di esposizione	Nord ⇒ Sud	Sud ⇒ Nord	Pendenza nulla
Intervallo di valori espressi in gradi	0 ⇒ 180	180 ⇒ 360	
Codifica			- 1

Mappa delle altimetrie	
Intervallo di valori di quota espressi in metri	0 ⇒ 300

Idrografia (Tevere, Corese, Farfa)				
Codifica in fasce di distanza	1	2	3	4
Intervallo di distanza	100 m	200 m	300 m	> 300 m

⁵ La tabella mostra in dettaglio le codifiche utilizzate per le modalità assunte dai predittori ampiamente presentati nella sezione precedente ed il cui significato non sia di immediata comprensione.

Il passo successivo dell'analisi sarà infatti quello di applicare il modello di segmentazione tipologica costruito sui dati campionari all'intera zona di studio per ottenere la mappa di rischio di presenza di siti archeologici.

Il processo di identificazione e costruzione dell'albero scelto come definitivo è un processo per sua natura iterativo che porta a valutare, come si vedrà meglio nel prosieguo, numerosi modelli alternativi. Fra i vari classificatori ad albero esaminati, quello che si è rivelato avere le caratteristiche migliori per la risoluzione del problema conoscitivo affrontato è tale da determinare il *ranking* dei predittori mostrato in Tab.2.

Tabella 2. *Importanza dei predittori – Albero definitivo: classifica basata su split univariati (0=bassa importanza; 100=importanza alta).*

Variabile	Ranking
SUOLO	63
LITO	58
FIUMI	18
TEVERE	53
ASPECT2	50
X	100
Y	100
SLOPE	65
QUOTE	86

Per completezza di trattazione (per maggiori dettagli si veda Breiman *et al.*, 1984, pag. 146-150) si ricorda che il *ranking* è costruito sulla base del decremento in impurità che deriva dalle cosiddette bipartizioni surrogate⁶. Infatti, alla misura di importanza di ciascuna variabile e quindi alla compilazione della classifica si perviene calcolando la somma, estesa a tutti i nodi dell'albero, del decremento in impurità prodotto, su ogni singola

⁶ Se lo *split* migliore del nodo t è la bipartizione s sulla variabile x_m , si definisce come il miglior surrogato di s stesso quello *split* s' , definito su tutte le variabili eccetto x_m , che sia più simile possibile a s nel senso che sia in grado di prevedere nel modo più accurato possibile l'azione di s .

variabile, dal miglior *split* surrogato. Si tratta di un brillante concetto introdotto per scongiurare il frequente problema pratico del mascheramento delle variabili, cioè della reale *prossimità* di due o più variabili, dovuto al fatto che, negli *split* univariati⁷, può intervenire una sola variabile alla volta.

Tornando alla Tab. 2., sembra importante sottolineare che le variabili di posizione (le coordinate *X* ed *Y*) rivestono, nel processo di costruzione dell'albero, un'importanza superiore a quanto forse poteva essere ipotizzato *a priori*, indicando una forte presenza di un *trend* spaziale. In parte ciò si spiega anche come conseguenza di un naturale limite delle procedure di partizione che è utile segnalare a questo punto: nella costruzione degli *split* gli algoritmi tendono spesso a privilegiare le variabili continue o, comunque, quelle con un maggior numero di modalità.

Lo *spazio* dunque manifesta un effetto! E questo a riprova del fatto che il precedente filone di studi intrapreso dal nostro gruppo ricerca aveva sollevato delle questioni importanti e non trascurabili. Se sia necessario o meno riprendere la modellistica di tipo *log/autolog* (cfr. paragrafo 3.), questo sarà oggetto di future valutazioni. Anche perché, come è noto, l'esistenza di un *trend* spaziale non necessariamente implica la presenza di autocorrelazione spaziale.

Quanto alla variabile relativa alla distanza dai fiumi sorge spontanea un'osservazione: essa non riveste, contrariamente a quanto solitamente si ipotizza (cfr., fra gli altri, Kuiper e Wescott, 1999) una particolare importanza se non come distanza dal Tevere. L'analisi indica quindi o che l'approvvigionamento delle risorse idriche non rappresentava assolutamente un problema nella scelta del sito migliore su cui edificare o che le classi utilizzate per discretizzare la variabile distanza sono poco rispondenti alla distribuzione territoriale del fenomeno.

Un altro aspetto doveroso da precisare è che nella costruzione dell'albero definitivo si è ritenuto opportuno i) assumere sempre costi unitari di errata classificazione e ii) utilizzare probabilità a priori eguali e non stimate sul campione di apprendimento; quest'ultima strategia è stata

⁷ Al riguardo è doveroso precisare una volta per tutte che nel presente lavoro si è deliberatamente stabilito di eseguire esclusivamente *split* univariati e non bipartizioni basate su combinazioni lineari di variabili ("è $\sum_i a_i x_i \leq c$?") dal momento che riteniamo le bipartizioni univariate essere ciò che della rappresentazione ad albero l'occhio umano riesce a comprendere, interpretare e valutare con maggiore facilità.

adottata soprattutto per non “forzare” l’analisi, dal momento che quello in questione è un parametro predeterminato (i punti classificati come “non sito” possono essere aumentati a piacere dal ricercatore); d’altro canto fissare le probabilità a priori pari a quelle del campione di apprendimento avrebbe notevolmente distorto l’analisi visto che il campione di apprendimento stesso non è casuale ed ha una composizione interna a prevalenza di “sito”. Per di più in tale circostanza, da un utilizzo ortodosso del CART e quindi invocando la “regola dell’ 1σ ” basata sulla stima onesta del costo di errata classificazione (CV) (cfr. Tab. 3 (a) e Fig. 12 (a)), si individua come albero ottimo l’albero degenere composto dal solo nodo radice (cfr. Fig. 13 (a)). Procedere nel processo di bipartizione non farebbe altro che incrementare il numero di casi mal classificati (cfr. Tab. 4 (a)).

Tabella 3. *Statistiche per alberi innestati: probabilità a priori (a) dal campione di apprendimento e (b) costanti.*

(a) (l’albero ottimo è indicato con *)

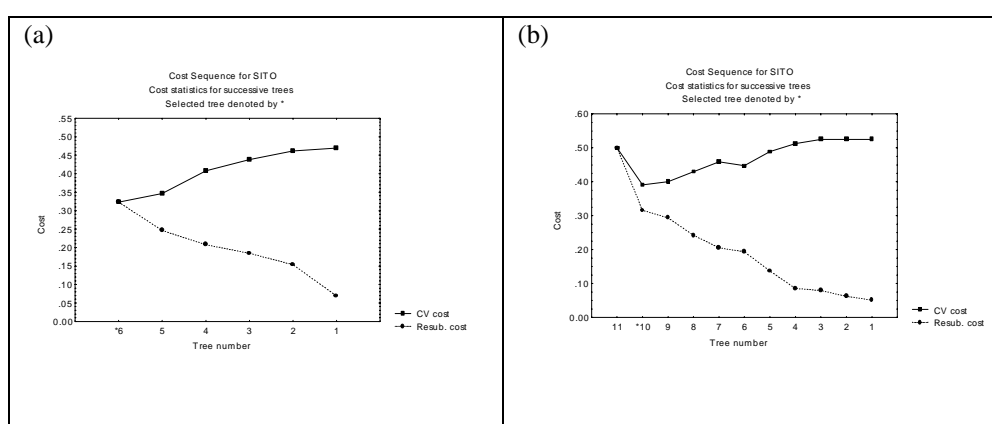
	Nodi terminali	Costo di CV	Std. error	Costo di Risost.	Complessità dei nodi
1	25	0,4692308	0,0437698	0,0692308	0
2	12	0,4615385	0,043723	0,1538462	0,006508876
3	8	0,4384615	0,0435195	0,1846154	0,007692308
4	6	0,4076923	0,0430991	0,2076923	0,011538462
5	3	0,3461538	0,0417254	0,2461538	0,012820513
*6	1	0,3230769	0,0410158	0,3230769	0,038461538

(b) (l’albero ottimo è indicato con *)

	Nodi terminali	Costo di CV	Std. error	Costo di Risost.	Complessità dei nodi
1	31	0,525433	0,045834	0,051136	0
2	27	0,525433	0,045834	0,062500	0,002841
3	24	0,525433	0,045834	0,079545	0,005682
4	23	0,512987	0,046498	0,085768	0,006223
5	15	0,489177	0,046788	0,136905	0,006392
6	10	0,447240	0,046526	0,193723	0,011364
7	9	0,459145	0,046707	0,205628	0,011905
8	6	0,430195	0,046419	0,242424	0,012266
9	3	0,400703	0,045782	0,293831	0,017136
*10	2	0,390963	0,045851	0,316017	0,022186
11	1	0,5	0	0,5	0,183983

Se, al contrario ed in linea con quanto stabilito, si adottano probabilità a priori costanti, analizzando i costi di errata classificazione (cfr. Tab. 3 (b)), come albero ottimo viene individuata, in corrispondenza del $CV \pm 1\sigma$ minore (cfr. anche Fig. 12 (b)), una struttura caratterizzata da un solo split (cfr. Fig. 13 (b)) il quale, fra l'altro come si vedrà meglio più avanti, coincide con la prima bipartizione dell'albero scelto come definitivo (cfr. Fig. 14).

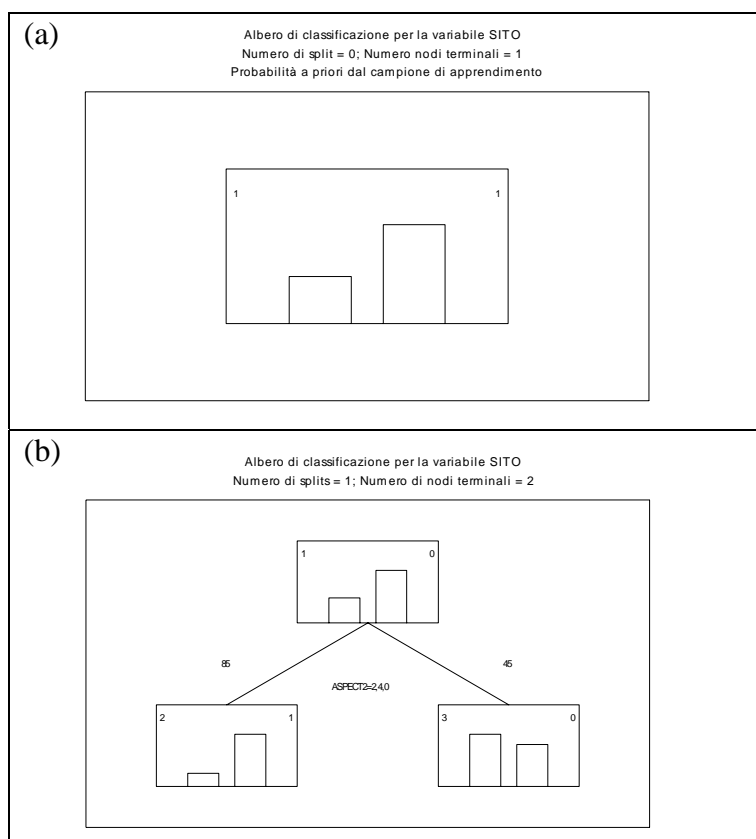
Figura 12. Variabilità e regola dell' 1σ : probabilità a priori (a) dal campione di apprendimento e (b) costanti.



Questa applicazione richiede però che l'albero sia più fitto e ciò perché altrimenti non si produrrebbe alcuna stima di probabilità di sito archeologico; infatti, se si fosse adottato l'albero a due nodi terminali (cfr. Fig. 13 (b)) la mappa di probabilità di sito archeologico corrispondente sarebbe stata una mappa a due soli colori (bianco e nero) e non a toni di grigio e pertanto assolutamente inutile rispetto allo scopo prefissato. L'obiettivo che ci siamo proposti, cioè quello di produrre in pratica una classificazione del territorio in esame, è stato perseguito adottando uno strumento che introduce alcune differenze operative rispetto al CART tradizionale. Fra le varie opzioni disponibili si è deciso di intervenire sulla sola regola di arresto dell'algoritmo adottando una regola cosiddetta "diretta". Questa strategia trasla l'ambito applicativo del presente lavoro dall'alveo ortodosso CART ad una soluzione FACT (*Fast Algorithm for*

Classification Trees). Si veda Loh e Vanichsetakul (1988) per una visione di insieme di tutte le differenze operative che distinguono FACT da CART.

Figura 13. Alberi di classificazione: probabilità a priori (a) dal campione di apprendimento e (b) costanti.



Per quanto concerne la specificità dell'opzione adottata, è noto che la metodologia CART utilizza la *cross-validation* per potare l'albero iniziale fino alla sua dimensione finale (Breiman *et al.*, 1984, 75-78). Nella regola d'arresto diretta adottata invece, lo *splitting* verso il basso si interrompe se:

- i) il tasso di errore apparente (la stima di risostituzione del tasso di errata classificazione) non diminuisce a seguito della bipartizione;

- ii) nel nodo è presente al massimo una classe con dimensione $\geq \text{MINDAT}$, dove il parametro MINDAT è fissato dall'utente.

Tabella 4. *Matrici di errata classificazione*: σ : probabilità a priori (a) dal campione di apprendimento e (b) costanti⁸.

(a)			(b)		
Learning Sample Misclassification Matrix Predicted (row) x observed (column) matrix Learning sample n = 130			Learning Sample Misclassification Matrix Predicted (row) x observed (column) matrix Learning sample n = 130		
	Class 0	Class 1		Class 0	Class 1
0			0		20
1		42	1	17	

Con riferimento al meno ovvio punto ii), il nodo sarà terminale se, etichettato di classe j con il criterio della prevalenza, al suo interno non saranno presenti più di $n_i \times fr$, $i \neq j$ casi dove n_i è la numerosità complessiva (nel campione di apprendimento, quindi!) delle classi diverse da quella attribuita (j) e dove la frazione fr è stabilita dal ricercatore. Ad esempio, riguardo al nostro caso di studio, la composizione del nodo radice è, con ovvietà di significato dei pedici, la seguente: $n_0 = 42$ e $n_1 = 88$. Ipotizzando un generico nodo t etichettato di classe 1, questo sarà terminale se $n_0 \times fr \geq n_0(t)$ ossia se a livello del nodo t il numero di casi etichettati come “non sito” non eccede $n_0 \times fr = 4.2$, essendosi scelta una frazione di casi malclassificati pari al 10% (cfr. Tab. 5).

I risultati del processo di segmentazione sono rappresentati graficamente nell'albero di Fig. 14 e sintetizzati in Tab. 6. L'albero ottenuto possiede 15 nodi terminali, a ciascuno dei quali corrisponde un particolare profilo geografico, profilo caratterizzato da una combinazione di modalità di alcune delle variabili esplicative.

⁸ Per la matrice in (a) è evidente che il numero di casi classificati erroneamente è relativo ai non siti!

Tabella 5. *Minimo numero di casi per classe per la regola di arresto FACT; $fr=.1$.*

Classe	Minimo per classe
0	4
1	9

Il primo *split* è sulla variabile ASPECT (orientamento della superficie su cui ricade il punto, i.e. nord, sud, est, ovest). Il nodo di sinistra viene successivamente suddiviso sulla base dell'uso del suolo (SUOLO). Scendendo lungo l'albero la partizione successiva è sulla base della litologia del terreno (LITO). Solo da questo *split* in poi interviene il *trend* sulla coordinata *Y*.

Risalendo verso l'alto all'interno dell'albero, sul lato destro si ritrovano il minor numero di siti con la maggiore incertezza "sito/non sito". Due ulteriori *split* sulla base della pendenza precedono l'intervento dei *trend* territoriali.

Se ora si sposta l'attenzione sull'aspetto dell'albero, si può affermare che globalmente, il processo di segmentazione perviene ad una stratificazione dei siti che sembra molto simmetrica ed in cui, a delle prime suddivisioni di carattere logico intuitivo, segue in un secondo momento la ricerca di gruppi di dati anomali individuati essenzialmente grazie a partizioni sulle coordinate stesse dei punti.

Nell'insieme, il modello ottenuto riesce a spiegare circa l'81% della presenza di sito, un risultato che, a prima vista, può apparire soddisfacente (cfr. Tab. 7). Si consideri inoltre che un modello di tipo logistico, in cui sarebbe assai facile introdurre dipendenze territoriali (cfr. paragrafo 3.) e che inglobasse tutti gli effetti principali e le interazioni doppie, avrebbe, nel caso di studio in esame, un numero di parametri da stimare di gran lunga superiore al numero di osservazioni disponibili. Per di più, il modello di regressione ad albero incorpora facilmente aspetti quali l'eteroschedasticità dei dati, che, come è noto, influenza l'efficienza delle stime dei modelli lineari ed allo stesso tempo risulta molto robusto in presenza di valori anomali della variabile di risposta. Infine, a questo punto dovrebbe essere anche evidente il fascino dell'approccio proposto al problema della costruzione di mappe di probabilità di sito archeologico: l'analisi viene

condotta principalmente in modo automatico ed i risultati che essa produce sono concisi e facili da comprendere anche per operatori del settore non specialisticamente preparati dal punto di vista metodologico.

6. Conclusioni

Lo scopo del presente lavoro è stato quello di implementare un modello per la costruzione di mappe di probabilità di sito archeologico basato sulla tecnologia GIS. I risultati ottenuti sono incoraggianti, soprattutto alla luce di alcune caratteristiche dell'approccio CART adottato le quali sono, a nostro avviso, particolarmente utili in circostanze applicative simili a quella da noi trattata. Nello specifico ci preme ricordare che il modello stimato i) rappresenta un'alternativa "robusta" alla classica regressione logistica in quanto molto poco dipendente da valori estremi; ii) è facilmente trasformabile in uno strumento automatico; iii) è agevolmente trasferibile ad altre realtà insediative; iv) fornisce degli *output* grafici interpretabili senza sforzo anche da non statistici v) non è in concorrenza con modelli a componenti spaziali come ad esempio quelli *log/autolog* (cfr. paragrafo 3.) ma anzi può rappresentarne un utile complemento in sede di selezione dei predittori e di individuazione di interazioni rilevanti.

Il lavoro presentato è chiaramente *in progress* in quanto, come non sarà certamente sfuggito al lettore attento, perché lo studio possa dirsi completato è necessario applicare il modello di classificazione stimato sui dati campionari all'intera zona di studio per ottenere la mappa finale di probabilità di sito archeologico. Con riferimento a tale mappa è bene precisare sin d'ora che questa sarà da interpretarsi non in termini assoluti ma relativi: infatti, potendo essere del tutto arbitrario il numero di punti etichettati come "non sito" all'interno del campione di apprendimento, aumentandone la presenza si genererebbe un conseguente "schiacciamento" in ogni nodo terminale del livello assoluto della probabilità stimata. L'effetto immediato di ciò è che le probabilità di sito calcolate nei nodi finali dell'albero di classificazione sono poco attendibili in termini di valore assoluto; queste assumono, invece, notevole importanza come rapporti relativi tra aree, possibilmente visualizzati dai toni di grigio di una mappa (più chiaro *vs.* più scuro \Rightarrow meno probabile *vs.* più probabile).

Modelli previsivi integrati come quello proposto in questa sede hanno una lunga e consolidata tradizione nel mondo scientifico statunitense; ciò non

accade in Europa, e ancor meno in Italia, dove gli studi al riguardo sono limitati e troppo spesso semplicistici, non solo per le differenze nelle legislazioni che regolamentano le campagne archeologiche nei due continenti (Stančić e Kvamme, 2003). Ma soprattutto a causa della mancata integrazione tra il mondo tecnologico-quantitativo e le discipline più inerenti l'archeologia. Comunque l'importanza dei modelli previsivi in archeologia continua ad essere sottolineata in modo deciso dagli specialisti (cfr., fra gli altri, Judge e Sebastian, 1988; Wescott e Brandon, 2000) soprattutto in vista di un *management* attento delle risorse culturali. E ciò nonostante le critiche rivolte a questo tipo di approccio (Ebert, 2000) fra le quali spicca il fatto che i modelli previsivi, pur individuando le condizioni ambientali e sociali più adatte per alcune tipologie insediative, non fornirebbero adeguato supporto nella comprensione dello sviluppo della configurazione localizzativa e mancherebbero del sostegno della teoria archeologica. A nostro avviso i modelli previsivi di sito archeologico possono essere proficuamente impiegati, oltre che per studi più tradizionali come quello qui proposto, anche in un approccio simulativo di verifica di ipotesi storico/archeologiche consentendo di valutare l'impatto di variabili ambientali e sociali su alcune realtà insediative. Quindi ci attendiamo, vista anche la domanda sempre più pressante da parte di organismi pubblici e privati di tutto ciò che concerne la tecnologia GIS per siti e monumenti nonché di mappe archeologiche di ogni tipo, un rapido sviluppo della modellistica previsiva all'interno di una branca in dinamica evoluzione quale quella dell'archeologia quantitativa.

Figura 14. Albero di classificazione finale.

Albero di classificazione per la variabile SITO
 Numero di splits = 14; Numero di nodi terminali = 15

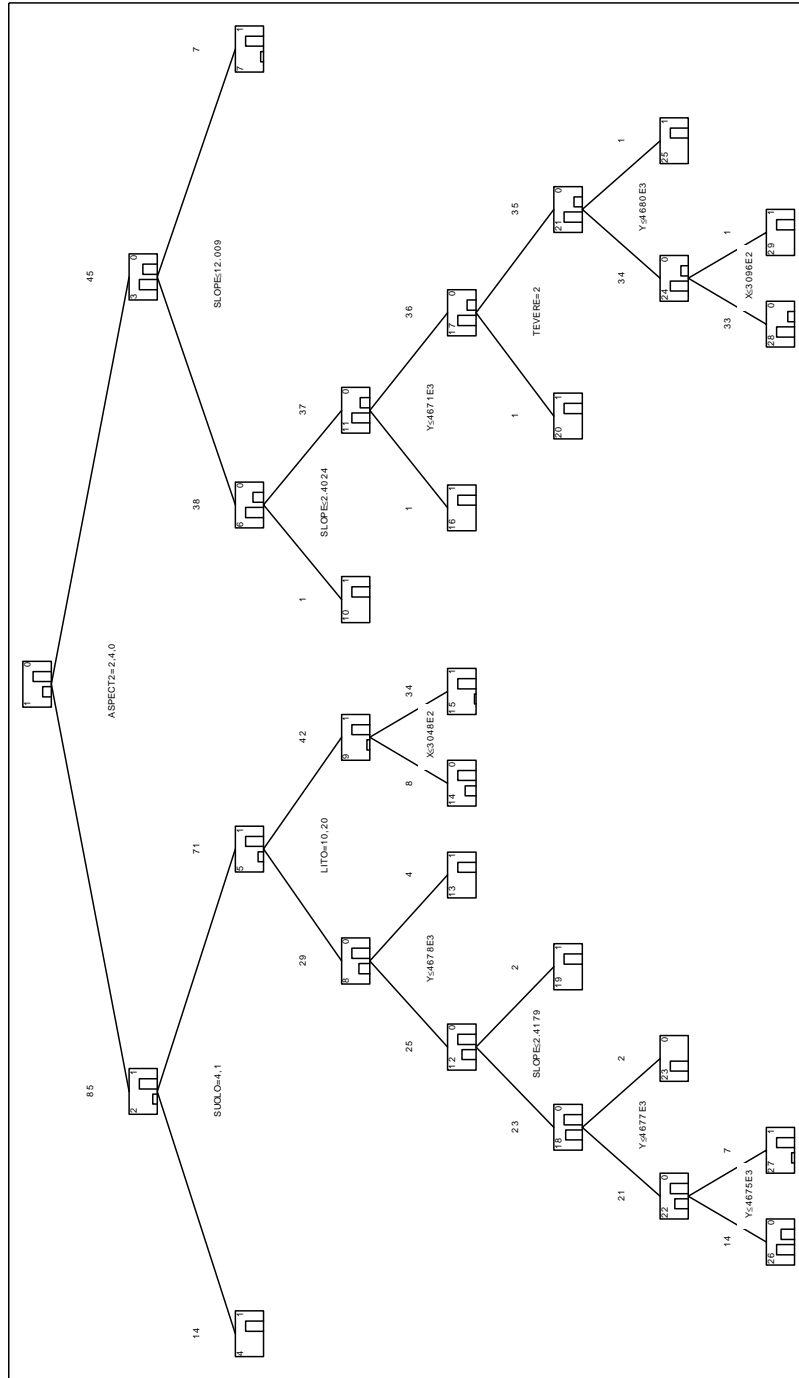


Tabella 6. *Struttura dei nodi dell'albero finale.*

	Codice nodo sx	Codice nodo dx	n. casi di classe 0	n. casi di classe 1	Classe prevista	Variabile di split	Livello dello split	Categorie di split
1	2	3	42	88	0	ASPECT2		0,2,4
2	4	5	17	68	1	SUOLO		1,4
3	6	7	25	20	0	SLOPE	12,0092	
4			0	14	1			
5	8	9	17	54	1	LITO		10,20
6	10	11	24	14	0	SLOPE	2,40243	
7			1	6	1			
8	12	13	11	18	0	Y	4677700	
9	14	15	6	36	1	X	304846	
10			0	1	1			
11	16	17	24	13	0	Y	4671500	
12	18	19	11	14	0	SLOPE	2,41795	
13			0	4	1			
14			3	5	0			
15			3	31	1			
16			0	1	1			
17	20	21	24	12	0	TEVERE		2
18	22	23	11	12	0	Y	4677300	
19			0	2	1			
20			0	1	1			
21	24	25	24	11	0	Y	4680000	
22	26	27	9	12	0	Y	4675000	
23			2	0	0			
24	28	29	24	10	0	X	309607	
25			0	1	1			
26			8	6	0			
27			1	6	1			
28			24	9	0			
29			0	1	1			

Tabella 7. *Albero finale: classi previste (righe) × classi osservate (colonne).*

	Classe 0 non sito	Classe 1 sito
Classe 0	37	20
Classe 1	5	68

Riferimenti bibliografici

- ARBIA G. (1993) Recenti sviluppi nella modellistica spaziale. In S. ZANI (ed.) *Metodi Statistici per le Analisi Territoriali*, Franco Angeli, Milano.
- ARBIA G. ESPA G. (1996) Forecasting statistical models of archaeological site location, *Archeologia e Calcolatori*, 7, 365-372.
- ARBIA G., ESPA G. (1997) Statistical modelling of archaeological sites locations as non-stationary Cox processes and its use in forecasting, *Archeologia e Calcolatori*, 8, 37-41.
- BENEDETTI R., ESPA G. (1995) L'uso di immagini telerilevate come informazione ausiliaria per l'interpolazione di dati archeologici, *Proceedings of the 1st International Congress on: Science and Technology for the Safeguard of Cultural Heritage in the Mediterranean Basin*, Catania, November 27 - December 2, 277-282.
- BESAG J. (1974) Spatial Interaction and the Statistical Analysis of Lattice Systems (with Discussion), *Journal of the Royal Statistical Society*, B, 86, 192-236.
- BESAG J. (1975) Statistical analysis of non-lattice data, *The Statistician*, 24, 3, 179-195.
- BESAG J. (1977) Efficiency of pseudolikelihood estimation for simple Gaussian fields, *Biometrika*, 64, 3, 616-618.
- BREIMAN L., FRIEDMAN J.H., OLSHEN R.A., STONE C.J. (1984) *Classification and Regression Trees*, Wadsworth International Group, Belmont, California.
- COLOSI F., ESPA S., GABRIELLI R., RICCI U., VERGA F. (1999) Statistical methods and computer aided technology for the study of archaeological presence in the Sabina Tiberina area, *Proceedings of Computer applications in archaeology*, Barcellona, March 25-28 1998, Archaeopress Oxford, 239 – 243.
- COX D.R., SNELL E.J. (1989) *Analysis of Binary Data*, 2nd ed., Chapman & Hall, London.
- CRESSIE N.A.C. (1991) *Statistics for Spatial Data*, John Wiley, New York.
- DE MEO A., ESPA G., ESPA S., GABRIELLI R., RICCI U. (1996) Impiego di un Sistema Geografico Informativo per un confronto

- modellistico preliminare tra due aree archeologiche sabine, *Archeologia e Calcolatori*, 7, 103-112.
- DE MEO A., ESPA G., ESPA S., PIFFERI A., RICCI U. (2000) Study of archaeological areas by means of advanced software technology and statistical methods, *Journal of Cultural Heritage*, 1, 233-245
- DE MEO A., ESPA G., ESPA S., PIFFERI A., RICCI U. (2002) Sistema geografico informativo della media valle del Tevere: l'area di Cures Sabini, *Quaderno DISA 75*, novembre, Dipartimento di Informatica e Studi Aziendali, Università di Trento.
- EBERT J.I. (2000) The state of art in "inductive" predictive modeling: seven big mistakes (and lots of smaller ones). In K.L. WESCOTT and R.J. BRANDON (eds.) *Practical Applications of GIS for Archaeologists – A Predictive Modeling Toolkit*, Taylor & Francis, London, 129-134.
- ESPA G., ARBIA G. (1995) Modelli Previsivi della probabilità di sito archeologico, *III International Symposium on Computing and Archaeology*, Roma 22-25 Novembre 1995, Abstracts of papers, 93-95.
- ESPA G., ESPA S., GABRIELLI R., PIFFERI A., RICCI U. (1995) Impiego integrato di metodi statistico-spaziali e tecniche informatiche nell'analisi di aree di interesse storico-archeologico, *Archeologia e Calcolatori*, 6, 35-60.
- GOMARASCA M. (1997) *Introduzione a telerilevamento e GIS per la gestione delle risorse agricole e ambientali*, Arte e Stampa, Daverio, Varese.
- GORDON A.D. (1981) *Classification*, Chapman & Hall, London.
- HAND D.J. (1981) *Discrimination and Classification*, Wiley, New York.
- HEIZER R.H., GRAHAM J.A. (1967) *A Guide to Field Methods in Archaeology*, National Press, Palo Alto.
- JUDGE W.J., SEBASTIAN L. (1988) *Quantifying the present and predicting the past: theory, method, and application of archaeological predictive modeling*. U.S. Government Printing Office, Washington, D.C.
- KOHLER T.A., PARKER S.C. (1986) Predictive models for archaeological resource location. In M.B. SCHIFFER (ed.) *Advances in Archaeological Method and Theory*, Academic Press, New York, 9, 397-452.

- KUIPER J.A., WESCOTT K.L. (1999) A GIS Approach for Predicting Prehistoric Site Locations, http://www01.giscale.com/technical_papers/Papers/paper057/.
- KVAMME K.L. (1983) Computer processing techniques for regional modelling of archaeological locations, *Advances in Computer Archaeology*, 1, 26-52.
- KVAMME K.L. (1989) Geographic information systems in regional archaeological research and data management. In M.B. SCHIFFER (ed.) *Archaeological Method and Theory*, Vol. 1, University of Arizona Press, Tucson, 139-203.
- KVAMME K.L. (1992) A predictive site location model on the High Plains: An example with an independent test. *Plains Anthropologist*, 37, 138, 19-40.
- LOH W.-Y., VANICHSETAKUL N. (1988) Tree-structured classification via generalized discriminant analysis, *Journal of the American Statistical Association*, 83, 403: 715-725.
- MUELLER J.W. (1974) The use of sampling in archaeological survey, *Memoir of the Society for American Archaeology*, 28.
- MUZZIOLI M.P. (1980) Cures sabini, *Forma Italiane*, IV, 2, Firenze.
- MUZZIOLI M.P. (1985) Cures Sabini, Misurare la terra: centuriazione e coloni nel mondo romano. Città, agricoltura, commercio: materiale da Roma e dal suburbio, 48 – 53.
- REDMAN C. (1974) *Archaeological Sampling Strategies*, Addison-Wesley Module in Anthropology, No. 55, Addison-Wesley, Reading, Massachussets.
- RICHARDS J.A. (1986) *Remote Sensing Digital Image Analysis*, Springer-Verlag, Berlino.
- STANČIĆ Z., KVAMME K.L. (2003) Settlement patterns modelling through Boolean overlays of social and environmental variables, <http://www.zrc-sazu.si/pic/pub/predictive/predictive.htm>.
- WATSON P.J., LEBLANC S., REDMAN C. (1971) *Explanation in Archaeology*, Columbia University Press, New York.
- WARREN R.E. (1990a) Predictive modelling in archaeology: a primer. In: K.M.S. ALLEN, S.W. GREEN, E.B.W. ZUBROW (eds.) *Interpreting space: GIS and Archaeology*, Taylor & Francis, New York, 90-111.
- WARREN R.E. (1990b) Predictive modelling of archaeological site location: a case study in the Midwest. In: K.M.S. ALLEN, S.W.

- GREEN, E.B.W. ZUBROW (eds.) *Interpreting space: GIS and Archaeology*, Taylor & Francis, New York, 201-215.
- WARREN R.E., OLIVER S.G., FERGUSON J.A., DRUHOT R.E. (1987) A predictive model of archaeological site location in the Western Shawnee National Forest, Technical Report No. 86-262-17, Quaternary Studies Program, Illinois State Museum, Springfield.
- WESCOTT K.L., BRANDON R.J. (eds.) (2000) *Practical Applications of GIS for Archaeologists – A Predictive Modeling Toolkit*, Taylor & Francis, London.
- ZUBROW E.B.W., HARBAUGH J.W. (1978) Archaeological Prospecting: Kriging and Simulation. In: I. HODDER (ed.) *Simulation Methods in Archaeology*, Cambridge University Press, Cambridge.

Probability maps of archaeological site location: one step beyond.

G. Espa♣, R. Benedetti♥, A. De Meo♣, S. Espa♣

♣ *Università di Trento, Via Inama, 5 - Trento, gespa@risc1.gelso.unitn.it*

♥ *Università "G. d'Annunzio", Viale Pindaro, 42 - Pescara*

♣ *Istituto per le Tecnologie Applicate ai Beni Culturali, C.N.R. - Via Salaria km 29.300, C.P.10, 00016 Monterotondo st. (Roma)*

Abstract. There is an increasing demand, among users of archaeological data, for models useful in the estimation of site probabilities which make use of the large amount of auxiliary variables geographically coded, which are currently available and managed by modern GIS technologies. All the studies on this issue based on advanced statistical methodology refer to regression or autoregression models, which showed to be particularly dependent, in the parameter estimation, on either outliers and on specific hypothesis on the phenomenon under study. The present paper proposes the use of a more robust and generalized solution which arise from the recent literature on nonparametric modelling which, according to authors opinion, could give a great improvement on the joint use of CART and GIS technologies. Satellite and physical landscape data on the analysed area, due to its high detail, could represent an important source of auxiliary information to explain the presence or not of an archaeological site in a given position of a map. The proposed approach was developed and applied using the *Cures Sabini* zone as test area obtaining encouraging products, in particular in terms of the facility of interpretation of the results by not statistical researchers and of the showed possibilities to make automatic the whole process of producing site probability maps. This will allow even to users not methodologically advanced extending the model to situation highly different from our test zone.

Keywords: Predictive models, Spatial correlation, CART, FACT, GIS.

Elenco dei papers del Dipartimento di Economia

1989. 1. *Knowledge and Prediction of Economic Behaviour: Towards A Constructivist Approach.* by Roberto Tamborini.

1989. 2. *Export Stabilization and Optimal Currency Baskets: the Case of Latin American Countries.* by Renzo G.Avesani Giampiero M. Gallo and Peter Pauly.

1989. 3. *Quali garanzie per i sottoscrittori di titoli di Stato? Una rilettura del rapporto della Commissione Economica dell'Assemblea Costituente* di Franco Spinelli e Danilo Vismara.
(*What Guarantees to the Treasury Bill Holders? The Report of the Assemblea Costituente Economic Commission Reconsidered* by Franco Spinelli and Danilo Vismara.)

1989. 4. *L'intervento pubblico nell'economia della "Venezia Tridentina" durante l'immediato dopoguerra* di Angelo Moioli.
(*The Public Intervention in "Venezia Tridentina" Economy in the First War Aftermath* by Angelo Moioli.)

1989. 5. *L'economia lombarda verso la maturità dell'equilibrio agricolo-commerciale durante l'età delle riforme* di Angelo Moioli.
(*The Lombard Economy Towards the Agriculture-Trade Equilibrium in the Reform Age* by Angelo Moioli.)

1989. 6. *L'identificazione delle allocazioni dei fattori produttivi con il duale.* di Quirino Paris e di Luciano Pilati.
(*Identification of Factor Allocations Through the Dual Approach* by Quirino Paris and Luciano Pilati.)

1990. 1. *Le scelte organizzative e localizzative dell'amministrazione postale: un modello interpretativo.* di Gianfranco Cerea.
(*The Post Service's Organizational and Locational Choices: An Interpretative Model* by Gianfranco Cerea.)

1990. 2. *Towards a Consistent Characterization of the Financial Economy.* by Roberto Tamborini.

1990. 3. *Nuova macroeconomia classica ed equilibrio economico generale: considerazioni sulla pretesa matrice walrasiana della N.M.C.* di Giuseppe Chirichiello.

(*New Classical Macroeconomics and General Equilibrium: Some Notes on the Alleged Walrasian Matrix of the N.C.M.* by Giuseppe Chirichiello.)

1990. 4. *Exchange Rate Changes and Price Determination in Polypolistic Markets.* by Roberto Tamborini.

1990. 5. *Congestione urbana e politiche del traffico. Un'analisi economica* di Giuseppe Folloni e Gianluigi Gorla.
(*Urban Congestion and Traffic Policy. An Economic Analysis* by Giuseppe Folloni and Gianluigi Gorla.)

1990. 6. *Il ruolo della qualità nella domanda di servizi pubblici. Un metodo di analisi empirica* di Luigi Mittone.
(*The Role of Quality in the Demand for Public Services. A Methodology for Empirical Analysis* by Luigi Mittone.)

1991. 1. *Consumer Behaviour under Conditions of Incomplete Information on Quality: a Note* by Pilati Luciano and Giuseppe Ricci.

1991. 2. *Current Account and Budget Deficit in an Interdependent World* by Luigi Bosco.

1991. 3. *Scelte di consumo, qualità incerta e razionalità limitata* di Luigi Mittone e Roberto Tamborini.
(*Consumer Choice, Unknown Quality and Bounded Rationality* by Luigi Mittone and Roberto Tamborini.)

1991. 4. *Jumping in the Band: Undeclared Intervention Thresholds in a Target Zone* by Renzo G. Avesani and Giampiero M. Gallo.

1991. 5 *The World Transfer Problem. Capital Flows and the Adjustment of Payments* by Roberto Tamborini.

1992.1 *Can People Learn Rational Expectations? An Ecological Approach* by Pier Luigi Sacco.

1992.2 *On Cash Dividends as a Social Institution* by Luca Beltrametti.

1992.3 *Politica tariffaria e politica informativa nell'offerta di servizi pubblici* di Luigi Mittone

(*Pricing and Information Policy in the Supply of Public Services* by Luigi Mittone.)

1992.4 *Technological Change, Technological Systems, Factors of Production* by Gilberto Antonelli and Giovanni Pegoretti.

1992.5 *Note in tema di progresso tecnico* di Geremia Gios e Claudio Migliarina.

(*Notes on Technical Progress*, by Geremia Gios and Claudio Migliarina).

1992.6 *Deflation in Input Output Tables* by Giuseppe Folloni and Claudio Migliarina.

1992.7 *Riduzione della complessità decisionale: politiche normative e produzione di informazione* di Luigi Mittone

(*Reduction in decision complexity: normative policies and information production* by Luigi Mittone)

1992.8 *Single Market Emu and Widening. Responses to Three Institutional Shocks in the European Community* by Pier Carlo Padoan and Marcello Pericoli

1993.1 *La tutela dei soggetti "privi di mezzi": Criteri e procedure per la valutazione della condizione economica* di Gianfranco Cerea

(*Public policies for the poor: criteria and procedures for a novel means test* by Gianfranco Cerea)

1993.2 *La tutela dei soggetti "privi di mezzi": un modello matematico per la rappresentazione della condizione economica* di Wolfgang J. Irler

(*Public policies for the poor: a mathematical model for a novel means test* by Wolfgang J. Irler)

1993.3 *Quasi-markets and Uncertainty: the Case of General Practice Service* by Luigi Mittone

1993.4 *Aggregation of Individual Demand Functions and Convergence to Walrasian Equilibria* by Dario Paternoster

1993.5 *A Learning Experiment with Classifier System: the Determinants of the Dollar-Mark Exchange Rate* by Luca Beltrametti, Luigi Marengo and Roberto Tamborini

- 1993.6 *Alcune considerazioni sui paesi a sviluppo recente* di Silvio Goglio
(*Latecomer Countries: Evidence and Comments* by Silvio Goglio)
- 1993.7 *Italia ed Europa: note sulla crisi dello SME* di Luigi Bosco
(*Italy and Europe: Notes on the Crisis of the EMS* by Luigi Bosco)
- 1993.8 *Un contributo all'analisi del mutamento strutturale nei modelli input-output* di Gabriella Berloffia
(*Measuring Structural Change in Input-Output Models: a Contribution* by Gabriella Berloffia)
- 1993.9 *On Competing Theories of Economic Growth: a Cross-country Evidence* by Maurizio Pugno
- 1993.10 *Le obbligazioni comunali* di Carlo Buratti (*Municipal Bonds* by Carlo Buratti)
- 1993.11 *Due saggi sull'organizzazione e il finanziamento della scuola statale* di Carlo Buratti
(*Two Essays on the Organization and Financing of Italian State Schools* by Carlo Buratti)
- 1994.1 *Un'interpretazione della crescita regionale: leaders, attività indotte e conseguenze di policy* di Giuseppe Folloni e Silvio Giove.
(*A Hypothesis about regional Growth: Leaders, induced Activities and Policy* by Giuseppe Folloni and Silvio Giove).
- 1994.2 *Tax evasion and moral constraints: some experimental evidence* by Luigi Bosco and Luigi Mittone.
- 1995.1 *A Kaldorian Model of Economic Growth with Shortage of Labour and Innovations* by Maurizio Pugno.
- 1995.2 *A che punto è la storia d'impresa? Una riflessione storiografica e due ricerche sul campo* a cura di Luigi Trezzi.
- 1995.3 *Il futuro dell'impresa cooperativa: tra sistemi, reti ed ibridazioni* di Luciano Pilati.
(*The future of the cooperative enterprise: among systems, networks and hybridisation* by Luciano Pilati).

1995.4 *Sulla possibile indeterminatezza di un sistema pensionistico in perfetto equilibrio finanziario* di Luca Beltrametti e Luigi Bonatti.
(*On the indeterminacy of a perfectly balanced social security system* by Luca Beltrametti and Luigi Bonatti).

1995.5 *Two Goodwinian Models of Economic Growth for East Asian NICs* by Maurizio Pugno.

1995.6 *Increasing Returns and Externalities: Introducing Spatial Diffusion into Krugman's Economic Geography* by Giuseppe Folloni and Gianluigi Gorla.

1995.7 *Benefit of Economic Policy Cooperation in a Model with Current Account Dynamics and Budget Deficit* by Luigi Bosco.

1995.8 *Coalition and Cooperation in Interdependent Economies* by Luigi Bosco.

1995.9 *La finanza pubblica italiana e l'ingresso nell'unione monetaria europea* di Ferdinando Targetti.
(*Italian Public Finance and the Entry in the EMU* by Ferdinando Targetti)

1996.1 *Employment, Growth and Income Inequality: some open Questions* by Annamaria Simonazzi and Paola Villa.

1996.2 *Keynes' Idea of Uncertainty: a Proposal for its Quantification* by Guido Fioretti.

1996.3 *The Persistence of a "Low-Skill, Bad-Job Trap" in a Dynamic Model of a Dual Labor Market* by Luigi Bonatti.

1996.4 *Lebanon: from Development to Civil War* by Silvio Goglio.

1996.5 *A Mediterranean Perspective on the Break-Down of the Relationship between Participation and Fertility* by Francesca Bettio and Paola Villa.

1996.6 *Is there any persistence in innovative activities?* by Elena Cefis.

1997.1 *Imprenditorialità nelle alpi fra età moderna e contemporanea* a cura di Luigi Trezzi.

- 1997.2 *Il costo del denaro è uno strumento anti-inflazionistico?* di Roberto Tamborini.
(*Is the Interest Rate an Anti-Inflationary Tool?* by Roberto Tamborini).
- 1997.3 *A Stability Pact for the EMU?* by Roberto Tamborini.
- 1997.4 *Mr Keynes and the Moderns* by Axel Leijonhufvud.
- 1997.5 *The Wicksellian Heritage* by Axel Leijonhufvud.
- 1997.6 *On pension policies in open economies* by Luca Beltrametti and Luigi Bonatti.
- 1997.7 *The Multi-Stakeholders Versus the Nonprofit Organisation* by Carlo Borzaga and Luigi Mittone.
- 1997.8 *How can the Choice of a Time-Consistent Monetary Policy have Systematic Real Effects?* by Luigi Bonatti.
- 1997.9 *Negative Externalities as the Cause of Growth in a Neoclassical Model* by Stefano Bartolini and Luigi Bonatti.
- 1997.10 *Externalities and Growth in an Evolutionary Game* by Angelo Antoci and Stefano Bartolini.
- 1997.11 *An Investigation into the New Keynesian Macroeconomics of Imperfect Capital Markets* by Roberto Tamborini.
- 1998.1 *Assessing Accuracy in Transition Probability Matrices* by Elena Cefis and Giuseppe Espa.
- 1998.2 *Microfoundations: Adaptive or Optimizing?* by Axel Leijonhufvud.
- 1998.3 *Clower's intellectual voyage: the 'Ariadne's thread' of continuity through changes* by Elisabetta De Antoni.
- 1998.4 *The Persistence of Innovative Activities. A Cross-Countries and Cross-Sectors Comparative Analysis* by Elena Cefis and Luigi Orsenigo

1998.5 *Growth as a Coordination Failure* by Stefano Bartolini and Luigi Bonatti

1998.6 *Monetary Theory and Central Banking* by Axel Leijonhufvud

1998.7 *Monetary policy, credit and aggregate supply: the evidence from Italy* by Riccardo Fiorentini and Roberto Tamborini

1998.8 *Stability and multiple equilibria in a model of talent, rent seeking, and growth* by Maurizio Pugno

1998.9 *Two types of crisis* by Axel Leijonhufvud

1998.10 *Trade and labour markets: vertical and regional differentiation in Italy* by Giuseppe Celi e Maria Luigia Segnana

1998.11 *Utilizzo della rete neurale nella costruzione di un trading system* by Giulio Pettenuzzo

1998.12 *The impact of social security tax on the size of the informal economy* by Luigi Bonatti

1999.1 *L'economia della montagna interna italiana: un approccio storiografico*, a cura di Andrea Leonardi e Andrea Bonoldi.

1999.2 *Unemployment risk, labour force participation and savings*, by Gabriella Berloff e Peter Simmons

1999.3 *Economia sommersa, disoccupazione e crescita*, by Maurizio Pugno

1999.4 *The nationalisation of the British Railways in Uruguay*, by Giorgio Fodor

1999.5 *Elements for the history of the standard commodity*, by Giorgio Fodor

1999.6 *Financial Market Imperfections, Heterogeneity and growth*, by Edoardo Gaffeo

1999.7 *Growth, real interest, employment and wage determination*, by Luigi Bonatti

2000.1 *A two-sector model of the effects of wage compression on unemployment and industry distribution of employment*, by Luigi Bonatti

2000.2 *From Kuwait to Kosovo: What have we learned? Reflections on globalization and peace*, by Roberto Tamborini

2000.3 *Metodo e valutazione in economia. Dall'apriorismo a Friedman*, by Matteo Motterlini

2000.4 *Under tertiarisation and unemployment*. by Maurizio Pugno

2001.1 *Growth and Monetary Rules in a Model with Competitive Labor Markets*, by Luigi Bonatti.

2001.2 *Profit Versus Non-Profit Firms in the Service Sector: an Analysis of the Employment and Welfare Implications*, by Luigi Bonatti, Carlo Borzaga and Luigi Mittone.

2001.3 *Statistical Economic Approach to Mixed Stock-Flows Dynamic Models in Macroeconomics*, by Bernardo Maggi and Giuseppe Espa.

2001.4 *The monetary transmission mechanism in Italy: The credit channel and a missing ring*, by Riccardo Fiorentini and Roberto Tamborini.

2001.5 *Vat evasion: an experimental approach*, by Luigi Mittone

2001.6 *Decomposability and Modularity of Economic Interactions*, by Luigi Marengo, Corrado Pasquali and Marco Valente.

2001.7 *Unbalanced Growth and Women's Homework*, by Maurizio Pugno

2002.1 *The Underground Economy and the Underdevelopment Trap*, by Maria Rosaria Carillo and Maurizio Pugno.

2002.2 *Interregional Income Redistribution and Convergence in a Model with Perfect Capital Mobility and Unionized Labor Markets*, by Luigi Bonatti.

2002.3 *Firms' bankruptcy and turnover in a macroeconomy*, by Marco Bee, Giuseppe Espa and Roberto Tamborini.

2002.4 *One “monetary giant” with many “fiscal dwarfs”: the efficiency of macroeconomic stabilization policies in the European Monetary Union*, by Roberto Tamborini.

2002.5 *The Boom that never was? Latin American Loans in London 1822-1825*, by Giorgio Fodor.

2002.6 *L'economia senza banditore di Axel Leijonhufvud: le ‘forze oscure del tempo e dell'ignoranza’ e la complessità del coordinamento*, by Elisabetta De Antoni.

2002.7 *Why is Trade between the European Union and the Transition Economies Vertical?*, by Hubert Gabrisch and Maria Luigia Segnana.

2003.1 *The service paradox and endogenous economic growth*, by Maurizio Pugno.

2003.2 *Mappe di probabilità di sito archeologico: un passo avanti*, di Giuseppe Espa, Roberto Benedetti, Anna De Meo e Salvatore Espa.

(Probability maps of archaeological site location: one step beyond, by Giuseppe Espa, Roberto Benedetti, Anna De Meo and Salvatore Espa).

PUBBLICAZIONE REGISTRATA PRESSO IL TRIBUNALE DI TRENTO

