

UNIVERSITÀ DEGLI STUDI DI MILANO

SCUOLA DI DOTTORATO

*Scienze della Terra*

DIPARTIMENTO

*Scienze della Terra "A. Desio"*

*Dottorato di Ricerca in Scienze della Terra*

# Tecniche di cartografia digitale per la redazione di mappe di dettaglio del suolo

**Marco Musetti**

Matricola R10535-R28

TUTORE:

**prof. Luca Trombino**

CO-TUTORE

**Dott. Roberto Comolli**

COORDINATORE DEL DOTTORATO:

**prof. Elisabetta Erba**

A.A. 2016/2017



## INDICE

Introduzione	1
Bibliografia	5
1. Digital Soil Mapping	6
1.1 Introduzione	6
1.2 Concetti di base della pedologia	8
1.3 L'evoluzione verso il DSM	11
1.4 Approccio utilizzato	13
Bibliografia	15
2. Inquadramento area di studio	17
2.1 Inquadramento climatico	18
2.2 Breve descrizione geologica dei litotipi	21
2.3 Delineazione dei principali usi del suolo e sintetica caratterizzazione della copertura vegetale	24
2.4 Inquadramento pedologico	28
Bibliografia	32
3. Le variabili geomorfometriche	33
3.1 Introduzione	33
3.2 Materiali e metodi	34
3.3 Analisi e risultati	36
3.4 Conclusioni	69
Bibliografia	71
4. Progettazione dell'attività di campionamento	73
4.1 Introduzione	73
4.2 Materiali e metodi	77
4.2.1 Introduzione di una variabile categoriale	84
4.3 Analisi e risultati	86
4.4 Discussione	93
4.5 Conclusioni	95
Bibliografia	96

5. Spessore del suolo	98
5.1 Introduzione	98
5.2 Materiali e metodi	100
5.3 Analisi e risultati	103
5.4 Discussione e conclusioni	116
Bibliografia	121
6. Erosione del suolo	125
6.1 Introduzione	125
6.2 Descrizione dei fenomeni di erosione del suolo	125
6.3 Mappa dell'erosione del suolo e delle aree soggette a dinamica di tipo calanchivo	147
6.3.1 Materiali e metodi	148
6.3.2 Analisi e risultati	153
6.4 Mappa della propensione alla frane superficiali	159
6.4.1 Materiali e metodi	159
6.4.2 Analisi e risultati	160
6.5 Discussione e conclusioni	169
Bibliografia	174
7. Tipologie di suolo	177
7.1 Introduzione	177
7.2 Caratteristiche generali dell'area di studio	179
7.3 Materiali e metodi	180
7.3.1 Tipologie di suolo: principali caratteristiche e classificazione	183
7.4 Analisi e risultati	189
7.5 Discussione e conclusioni	196
Bibliografia	197
8. Conclusioni generali	198

## ELENCO DELLE ABBREVIAZIONI E DEGLI ACRONIMI

AaLSmp	argilla>45% (Aa), sabbia<50% (L), solum 61-100 cm (Smp)
AaLSms	argilla>45% (Aa), sabbia<50% (L), solum 35-60 cm (Sms)
AbHSms	argilla<25% (Ab), sabbia>50% (H), solum 35-60 cm (Sms)
AbHSp	argilla<25% (Ab), sabbia>50% (H), solum >100 cm (Sp)
AbHSs	argilla<25% (Ab), sabbia>50% (H), solum<35 cm (Ss)
AbLSp	argilla<25% (Ab), sabbia<50% (L), solum >100 cm (Sp)
ADV	Unità dell'Ardivestra
ALLD	Abruptic Luvisol (Loamic, Differentic)
ALLLD	Abruptic Leptic Luvisol (Loamic, Differentic)
Altichan	altitude above channel network (SAGA)
AmHSmp	argilla 25-35% (Am), sabbia >50% (H), solum 61-100 (Smp)
AmHSp	argilla 25-35%, sabbia >50%, (H), solum >100 cm (Sp)
AmLSmp	argilla 25-45% (Am), sabbia<50% (L), spessore solum 61-100 cm (Smp)
AmLSms	argilla 25-45% (Am), sabbia<50% (L), spessore solum 35-60 cm (Sms)
AmLSp	argilla 25-45% (Am), sabbia<50% (L), solum >100 cm (Sp)
AmLSs	argilla 25-45% (Am), sabbia<50% (L), solum<35 cm (Ss)
ANN	artificial neural network
ANT	Formazione di Antognola
ARPA	Agenzia Regionale Protezione Ambiente della Lombardia
AS	aree stabili
asperit	Asperit (QGIS)
aspsharp	aspect (Jenness tool)
AVL	Arenarie di Monte Vallassa
AWC	Available Water Capacity
BAI	Brecce di Baiso
BPE	Unità di Costa Pelata
C	calanchi
CARG	Cartografia geologica e geotematica
catarea	catchment Area (recursive method) (SAGA)
catchslope	catchment Area (max flux method) (SAGA)
CCL	Calcaric Cambisol (Loamic)
ccross	cross-sectional curvature (Jenness tool)
CCS	Conglomerati di Cassano Spinola
CCS1	Membro delle arenarie di Monte Arzolo
cgen	general curvature (Jenness tool)
channelNB	channel network base level (SAGA)
clong	longitudinal curvature (Jenness tool)
CM	calanchi mammellonari
ConvInd	convergence index (SAGA)
cplan	plan curvature (Jenness tool)
cprof	profile curvature (Jenness tool)
CSC	Capacità di Scambio Cationico
ctang	tangential curvature (Jenness tool)
CTG	Formazione di Contignaco
CTIgrass	Compound Topographic Index (GRASS)
ctot	total curvature (Jenness tool)
DA	discriminant analysis
DCL	Dystric Cambisol (Loamic)
DEM	Digital Elevation Mode
difsol	diffuse insolation (SAGA)
dirsol	direct insolation (SAGA)

DSM	Digital Soil Mapping
DUSAF	Destinazione d'Uso dei Suoli Agricoli e forestali
ED	erosione/deposizione
EDTA	acido etilendiamminotetraacetico
EIC	erosione da flusso idrico concentrato
EL	erosione laminare
EPA	Environmental Protection Agency
ER	Error-Rate
ERSAF	Ente Regionale per i Servizi all'Agricoltura e alle Foreste della Lombardia
FAAS	Flame Atomic Absorption Spectroscopy
FAO	Food and Agriculture Organization
flowaccS	flow accumulation (SAGA)
flowdir	flow direction (ArcGIS)
flowdrop	flow drop (ArcGIS)
GIS	Geographic Information System
GNO	Formazione di Sapigno
GRASS	Geographic Resources Analysis Support System
IFFI	Inventario dei Fenomeni Franosi in Italia
LCCC	Leptic, Calcaric Cambisol (Clayic)
LCCL	Leptic Calcaric Cambisol (Loamic)
LCR	Leptic Calcaric Regosol
LCRL	Leptic Calcaric Regosol (Loamic)
LDCM	Leptic Dystric Cambisol (Loamic)
LDR	Leptic Dystric Regosol
LSCR	Leptic Skeletic Calcaric Regosol
LSfactor	LS factor (SAGA)
LVC	Leptic Vertisol (Calcaric)
massbal	mass balance index (SAGA)
MMP	Marne di Monte Piano
MR	Misclassification risk
MR <sub>g</sub>	Misclassification risk per la classe g
MVTa	Litozona Inferiore
NER	Non-Error-Rate
NIR	Near infrared
NIZ	Unità di Nizza
PET	potential evapotranspiration
PLSR	Partial Least Square Regression
POI	Sintema del Po (Unità Postglaciale)
QGIS	Quantum gis
R	roccia affiorante
RAN	Formazione di Ranzano
RAN1	Formazione di Ranzano - Membro del Pizzo d'Oca
RAN1c	Formazione di Ranzano - Litozona pelitico-arenacea
RAN2a	Formazione di Ranzano - Litozona arenaceo-conglomeratica
RAN2b	Formazione di Ranzano - Litozona arenaceo-pelitica
RAN2c	Formazione di Ranzano - Litozona pelitico-arenacea
RAN3	Formazione di Ranzano - Membro di Varano de' Melegari
RAN3a	Formazione di Ranzano - Litozona arenaceo-conglomeratica
RAN3b	Formazione di Ranzano - Litozona arenaceo-pelitica
RAN3c	Formazione di Ranzano - Litozona pelitico-arenacea
S	suffusione
ssharp	slope (Jenness tool, Sharp algorithm)
Sn <sub>g</sub>	Sensibility per la classe g

SOFM	Self-Organizing Feature Maps
Sp <sub>g</sub>	Specificity per la classe g
topPosiS	Topographic position index (SAGA)
TPI	Topographic position index (QGIS)
TRI	Terrain Ruggedness Index (QGIS)
TTS	Unità di Torretta
TWI	topographic wetness index (SAGA)
USDA	United States Department of Agriculture
UTS	Unità Tipologiche del Suolo
VRM	Vector Ruggedness Measure (SAGA)
VRZ	Unità di Varzi
WI	wetness index (SAGA)
WRB	World Reference Base for soil Resources
Zstd	elevation

## ELENCO DELLE FIGURE

Fig. 1 Componenti coinvolte nel processo di mappatura dei suoli che influenzano la qualità del prodotto finale a priori.	6
Fig. 1.1 Generalizzazione spaziale per suoli con limitata estensione.	7
Fig. 1.2 Generalizzazione spaziale per suoli dissimili che ricorrono in modo intermittente.	7
Fig. 1.3 Rappresentazione dello spazio multidimensionale utilizzato per la costruzione delle mappe.	8
Fig. 2.1 Ubicazione dell'area di studio all'interno della provincia di Pavia.	17
Fig. 2.2 Medie mensili di temperatura e precipitazioni per la stazione di Varzi Nivione (elaborazione su dati giornalieri ARPA Lombardia).	18
Fig. 2.3 Precipitazioni totali annuali 2002-2016 per la stazione di Varzi Nivione (elaborazione su dati giornalieri ARPA Lombardia).	19
Fig. 2.4 Temperature medie mensili per le stazioni di Varzi Nivione (2002-2016) e Montemarzino (1933-1955).	19
Fig. 2.5 Bilancio idrico sec. Thornthwaite per la stazione di Varzi Nivione, relativo a differenti riserve idriche (AWC) del suolo.	20
Fig. 2.6 Area di studio con le Unità Geologiche (stralcio del Foglio 178-Voghera, della Carta Geologica d'Italia (1:50.000) del Progetto CARG).	22
Fig. 2.7 Area di studio con le classi d'uso del suolo secondo la carta DUSAF del 2012.	24
Fig. 2.8 Variazione dell'areale adibito ad uso agricolo avvenuta tra il 1954 e il 2012, con perimetro delle due valli	25
Fig. 2.9 Dettaglio della Carta dei Suoli della Regione Lombardia (1:250000)	29
Fig. 2.10 Carta delle Unità geolitologiche (scala 1:10.000) della Val di Nizza (De Wrachien et al.1981).	30
Fig. 2.11 Carta dei suoli della Val di Nizza (scala 1:10.000) (Casati 1985).	31
Fig. 3.1 Box plot relativo alle variabili standard	36
Fig. 3.2 Rappresentazione della variabile Terrain Ruggedness Index (QGIS) (A)+e degli outliers con valore superiore a (media+3 deviazione standard) (B).	38
Fig. 3.3 Rappresentazione della variabile Topographic position index (QGIS) (A) e degli outliers con valore superiore a (media+3 deviazione standard) (B).	38
Fig. 3.4 Rappresentazione della variabile Asperit (QGIS) (A) e degli outliers con valore superiore a (media+3 deviazione standard) (B).	39
Fig. 3.5 Rappresentazione della variabile total curvature (Jenness tool) (A) e degli outliers con valore superiore a (media+3 deviazione standard) (B).	39
Fig. 3.6 Rappresentazione della variabile tangential curvature (Jenness tool) (A) e degli outliers con valore superiore a (media+3 deviazione standard) (B).	39
Fig. 3.7 Rappresentazione della variabile profile curvature (Jenness tool) (A) e degli outliers con valore superiore a (media+3 deviazione standard) (B).	40
Fig. 3.8 Rappresentazione della variabile plan curvature (Jenness tool) (A) e degli outliers con valore superiore a (media+3 deviazione standard) (B).	40

Fig. 3.9 Rappresentazione della variabile longitudinal curvature (Jenness tool) (A) e degli outliers con valore superiore a (media+3 deviazione standard) (B).	40
Fig. 3.10 Rappresentazione della variabile general curvature (Jenness tool) (A) e degli outliers con valore superiore a (media+3 deviazione standard) (B).	41
Fig. 3.11 Rappresentazione della variabile cross-sectional curvature (Jenness tool) (A) e degli outliers con valore superiore a (media+3 deviazione standard) (B).	41
Fig. 3.12 Rappresentazione della variabile slope (Jenness tool) (A) e degli outliers con valore superiore a (media+3 deviazione standard) (B).	41
Fig. 3.13 Rappresentazione della variabile wetness index (SAGA) (A) e degli outliers con valore superiore a (media+3 deviazione standard) (B).	42
Fig. 3.14 Rappresentazione della variabile topographic wetness index (SAGA) (A) e degli outliers con valore superiore a (media+3 deviazione standard) (B).	42
Fig. 3.15 Rappresentazione della variabile Compound Topographic Index (GRASS) (A) e degli outliers con valore superiore a (media+3 deviazione standard) (B).	43
Fig. 3.16 Rappresentazione della variabile catchment Area (recursive method) (SAGA) (A) e degli outliers con valore superiore a (media+3 deviazione standard) (B).	43
Fig. 3.17 Rappresentazione della variabile LS factor (SAGA) (A) e degli outliers con valore superiore a (media+3 deviazione standard) (B).	43
Fig. 3.18 Rappresentazione della variabile convergence index (SAGA) (A) e degli outliers con valore superiore a (media+3 deviazione standard) (B).	43
Fig. 3.19 Rappresentazione della variabile catchment Area (max flux method) (SAGA) (A) e degli outliers con valore superiore a (media+3 deviazione standard) (B).	44
Fig. 3.20 Rappresentazione della variabile altitude above channel network (SAGA) (A) e degli outliers con valore superiore a (media+3 deviazione standard) (B).	44
Fig. 3.21 Rappresentazione della variabile flow drop (ArcGIS) (A) e degli outliers con valore superiore a (media+3 deviazione standard) (B).	44
Fig. 3.22 Rappresentazione della variabile diffuse insolation (SAGA) (A) e degli outliers con valore superiore a (media+3 deviazione standard) (B).	45
Fig. 3.23 Rappresentazione della variabile flow accumulation (SAGA) (A) e degli outliers con valore superiore a (media+3 deviazione standard) (B).	45
Fig. 3.24 Rappresentazione della variabile Topographic position index (SAGA) (A) e degli outliers con valore superiore a (media+3 deviazione standard) (B).	45
Fig. 3.25 Rappresentazione della variabile Vector Ruggedness Measure (SAGA) (A) e degli outliers con valore superiore a (media+3 deviazione standard) (B).	46
Fig. 3.26 Rappresentazione della variabile mass balance index (SAGA) (A) e della variabile channel network base level (SAGA) (B).	46
Fig. 3.27 Rappresentazione della variabile aspect (Jenness tool) (A) e della variabile DEM (B).	46
Fig. 3.28 Rappresentazione della variabile flow direction (ArcGIS) (A) e della variabile direct insolation (SAGA) (B).	47
Fig. 3.29 Rappresentazione degli outliers relativi alla Plan Curvature con valore maggiore di: (media+3 deviazione standard).	47
Fig. 3.30 Rappresentazione degli outliers relativi alla Plan Curvature con valore assoluto maggiore di 100.	48
Fig. 3.31 Rappresentazione di alcuni outliers relativi alla Plan Curvature con valore maggiore di: (media+3 deviazione standard).	49
Fig. 3.32 Normal probability plot relativo alla variabile elevation (A) e suo istogramma di frequenza (B).	50
Fig. 3.33 Normal probability plot relativo alla variabile Terrain Ruggedness Index (QGIS) (A) e suo istogramma di frequenza (B).	50
Fig. 3.34 Normal probability plot relativo alla variabile Topographic position index (QGIS) (A) e suo istogramma di frequenza (B).	50
Fig. 3.35 normal probability plot relativo alla variabile Asperit (QGIS) (A) e suo istogramma di frequenza (B).	51
Fig. 3.36 Normal probability plot relativo alla variabile total curvature (Jenness tool) (A) e suo istogramma di frequenza (B).	51

Fig. 3.37 Normal probability plot relativo alla variabile tangential curvature (Jenness tool) (A) e suo istogramma di frequenza (B).	51
Fig. 3.38 Normal probability plot relativo alla variabile profile curvature (Jenness tool) (A) e suo istogramma di frequenza (B).	52
Fig. 3.39 Normal probability plot relativo alla variabile plan curvature (Jenness tool) (A) e suo istogramma di frequenza (B).	52
Fig. 3.40 Normal probability plot relativo alla variabile longitudinal curvature (Jenness tool) (A) e suo istogramma di frequenza (B).	52
Fig. 3.41 Normal probability plot relativo alla variabile general curvature (Jenness tool) (A) e suo istogramma di frequenza (B).	53
Fig. 3.42 Normal probability plot relativo alla variabile cross sectional (Jenness tool) (A) e suo istogramma di frequenza (B).	53
Fig. 3.43 Normal probability plot relativo alla variabile aspect (Jenness tool) (A) e suo istogramma di frequenza (B).	53
Fig. 3.44 Normal probability plot relativo alla variabile slope (Jenness tool) (A) e suo istogramma di frequenza (B).	54
Fig. 3.45 Normal probability plot relativo alla variabile wetness index (SAGA) (A) e suo istogramma di frequenza (B).	54
Fig. 3.46 Normal probability plot relativo alla variabile topographic wetness index (SAGA) (A) e suo istogramma di frequenza (B).	54
Fig. 3.47 Normal probability plot relativo alla variabile Compound Topographic Index (GRASS) (A) e suo istogramma di frequenza (B).	55
Fig. 3.48 Normal probability plot relativo alla variabile catchment Area (recursive method) (A) e suo istogramma di frequenza (B).	55
Fig. 3.49 Normal probability plot relativo alla variabile mass balance index (SAGA) (A) e suo istogramma di frequenza (B).	55
Fig. 3.50 Normal probability plot relativo alla variabile LS Factor (SAGA) (A) e suo istogramma di frequenza (B).	56
Fig. 3.51 Normal probability plot relativo alla variabile convergence index (SAGA) (A) e suo istogramma di frequenza (B).	56
Fig. 3.52 Normal probability plot relativo alla variabile channel network base level (SAGA) (A) e suo istogramma di frequenza (B).	56
Fig. 3.53 Normal probability plot relativo alla variabile catchment Area (max flux method) (SAGA) (A) e suo istogramma di frequenza (B).	57
Fig. 3.54 Normal probability plot relativo alla variabile altitude above channel network (SAGA) (A) e suo istogramma di frequenza (B).	57
Fig. 3.55 Normal probability plot relativo alla variabile flow direction (ArcGIS) (A) e suo istogramma di frequenza (B).	57
Fig. 3.56 Normal probability plot relativo alla variabile flow drop (ArcGIS) (A) e suo istogramma di frequenza (B).	58
Fig. 3.57 Normal probability plot relativo alla variabile diffuse insolation (SAGA) (A) e suo istogramma di frequenza (B).	58
Fig. 3.58 Normal probability plot relativo alla variabile direct insolation (SAGA) (A) e suo istogramma di frequenza (B).	58
Fig. 3.59 Normal probability plot relativo alla variabile flow accumulation (Jenness tool) (A) e suo istogramma di frequenza (B).	59
Fig. 3.60 Normal probability plot relativo alla variabile Topographic position index (SAGA) (A) e suo istogramma di frequenza (B).	59
Fig. 3.61 Normal probability plot relativo alla variabile Vector Ruggedness Measure (SAGA) (A) e suo istogramma di frequenza (B).	59
Fig. 3.62 Rappresentazione dei valori di Variance-decomposition proportion per ciascun singular values maggiore di 30.	67
Fig. 3.63 Rappresentazione della variabile general curvature (Jenness tool).	68
Fig. 3.64 Rappresentazione della variabile longitudinal curvature (Jenness tool).	69
Fig. 3.65 Rappresentazione della variabile cross-sectional curvature (Jenness tool).	69
Fig. 4.1 Diagramma di Pareto relativo alla percentuale di varianza spiegata dalle prime 10 PCs.	80
Fig. 4.2 Scree plot relativo alla principal component analysis eseguita sul set di 30 variabili geomorfometriche.	80

Fig. 4.3 Loadings plot relativo alla principal component analysis eseguita sul set di 30 variabili geomorfometriche.	80
Fig. 4.4 Scatter plot relativo alla tabella 4.4.	82
Fig. 4.5 Segmentazione del territorio in 7 classi ottenuta applicando una ANN Self-Organizing Feature Maps.	82
Fig. 4.6 Mappa per il rilevamento pedologico con l'indicazione delle unità geologiche semplificate e dei punti di campionamento rilevati.	85
Fig. 4.7 Confronto della densità di probabilità empirica relativa ai punti di campionamento eseguito con il metodo free survey (sample poits 2014) con i punti relativi al campionamento effettuato secondo l'approccio proposto (sample poits 2015) e con la densità di probabilità empirica relativa all'intera area di studio (Total points) per la variabile elevation (A) e per la variabile Topographic position index (SAGA) (B).	86
Fig. 4.8 Confronto della densità di probabilità empirica relativa ai punti di campionamento eseguito con il metodo free survey (sample poits 2014) con i punti relativi al campionamento effettuato secondo l'approccio proposto (sample poits 2015) e con la densità di probabilità empirica relativa all'intera area di studio (Total points) per la variabile Vector Ruggedness Measure (SAGA) (A) e per la variabile wetness index (SAGA) (B).	87
Fig. 4.9 Confronto della densità di probabilità empirica relativa ai punti di campionamento eseguito con il metodo free survey (sample poits 2014) con i punti relativi al campionamento effettuato secondo l'approccio proposto (sample poits 2015) e con la densità di probabilità empirica relativa all'intera area di studio (Total points) per la variabile Topographic position index (QGIS) (A) e per la variabile Terrain Ruggedness Index (QGIS) (B).	87
Fig. 4.10 Confronto della densità di probabilità empirica relativa ai punti di campionamento eseguito con il metodo free survey (sample poits 2014) con i punti relativi al campionamento effettuato secondo l'approccio proposto (sample poits 2015) e con la densità di probabilità empirica relativa all'intera area di studio (Total points) per la variabile slope (Jenness tool) (A) e per la variabile topographic wetness index (SAGA) (B).	87
Fig. 4.11 Confronto della densità di probabilità empirica relativa ai punti di campionamento eseguito con il metodo free survey (sample poits 2014) con i punti relativi al campionamento effettuato secondo l'approccio proposto (sample poits 2015) e con la densità di probabilità empirica relativa all'intera area di studio (Total points) per la variabile LS factor (SAGA) (A) e per la variabile mass balance index (SAGA) (B).	88
Fig. 4.12 Confronto della densità di probabilità empirica relativa ai punti di campionamento eseguito con il metodo free survey (sample poits 2014) con i punti relativi al campionamento effettuato secondo l'approccio proposto (sample poits 2015) e con la densità di probabilità empirica relativa all'intera area di studio (Total points) per la variabile flow direction (ArcGIS) (A) e per la variabile flow drop (ArcGIS) (B).	88
Fig. 4.13 Confronto della densità di probabilità empirica relativa ai punti di campionamento eseguito con il metodo free survey (sample poits 2014) con i punti relativi al campionamento effettuato secondo l'approccio proposto (sample poits 2015) e con la densità di probabilità empirica relativa all'intera area di studio (Total points) per la variabile direct insolation (SAGA) (A) e per la variabile flow accumulation (SAGA) (B).	88
Fig. 4.14 Confronto della densità di probabilità empirica relativa ai punti di campionamento eseguito con il metodo free survey (sample poits 2014) con i punti relativi al campionamento effettuato secondo l'approccio proposto (sample poits 2015) e con la densità di probabilità empirica relativa all'intera area di studio (Total points) per la variabile total curvature (Jenness tool) (A) e per la variabile diffuse insolation (SAGA) (B).	89
Fig. 4.15 Confronto della densità di probabilità empirica relativa ai punti di campionamento eseguito con il metodo free survey (sample poits 2014) con i punti relativi al campionamento effettuato secondo l'approccio proposto (sample poits 2015) e con la densità di probabilità empirica relativa all'intera area di studio (Total points) per la variabile tangential curvature (Jenness tool) (A) e per la variabile Compound Topographic Index (GRASS) (B).	89
Fig. 4.16 Confronto della densità di probabilità empirica relativa ai punti di campionamento eseguito con il metodo free survey (sample poits 2014) con i punti relativi al campionamento effettuato secondo l'approccio proposto (sample poits 2015) e con la densità di probabilità empirica relativa all'intera area di studio (Total points) per la variabile plan curvature (Jenness tool) (A) e per la variabile profile curvature (Jenness tool) (B).	89
Fig. 4.17 Confronto della densità di probabilità empirica relativa ai punti di campionamento eseguito con il metodo free survey (sample poits 2014) con i punti relativi al campionamento effettuato secondo l'approccio proposto (sample poits 2015) e con la densità di probabilità empirica relativa all'intera area di studio (Total points) per la variabile longitudinal curvature (Jenness tool) (A) e per la variabile convergence index (SAGA) (B).	90

Fig. 4.18 Confronto della densità di probabilità empirica relativa ai punti di campionamento eseguito con il metodo free survey (sample poits 2014) con i punti relativi al campionamento effettuato secondo l'approccio proposto (sample poits 2015) e con la densità di probabilità empirica relativa all'intera area di studio (Total points) per la variabile general curvatures (Jeness tool) (A) e per la variabile channel network base level (SAGA) (B).	90
Fig. 4.19 Confronto della densità di probabilità empirica relativa ai punti di campionamento eseguito con il metodo free survey (sample poits 2014) con i punti relativi al campionamento effettuato secondo l'approccio proposto (sample poits 2015) e con la densità di probabilità empirica relativa all'intera area di studio (Total points) per la variabile catchment Area (max flux method) (SAGA) (A) e per la variabile cross-sectional curvatures (Jeness tool) (B).	90
Fig. 4.20 Confronto della densità di probabilità empirica relativa ai punti di campionamento eseguito con il metodo free survey (sample poits 2014) con i punti relativi al campionamento effettuato secondo l'approccio proposto (sample poits 2015) e con la densità di probabilità empirica relativa all'intera area di studio (Total points) per la variabile Asperit (QGIS) (A) e per la variabile catchment Area (recursive method) (SAGA) (B).	90
Fig. 4.21 Confronto della densità di probabilità empirica relativa ai punti di campionamento eseguito con il metodo free survey (sample poits 2014) con i punti relativi al campionamento effettuato secondo l'approccio proposto (sample poits 2015) e con la densità di probabilità empirica relativa all'intera area di studio (Total points) per la variabile altitude above channel network (SAGA) (A) e per la variabile aspect (Jeness tool) (B).	90
Fig. 4.22 Fasi della progettazione dell'attività di campionamento.	92
Fig. 5.1 Unità elementare di una ANN.	101
Fig. 5.2 Scree plot relativo alla principal component analysis eseguito sul set di 30 variabili geomorfometriche.	103
Fig. 5.3 Sequenza di operazioni dei METODI A e B per individuare i sub-set di variabili utilizzati per modellare lo spessore del suolo.	106
Fig. 5.4 Variazione del coefficiente di correlazione per il validation set rispetto ai differenti sub-set di variabili selezionati con il METODO A.	110
Fig. 5.5 Variazione del coefficiente di correlazione per il test set rispetto ai differenti sub-set di variabili selezionati con il METODO A.	110
Fig. 5.6 Variazione del coefficiente di correlazione per il validation set rispetto ai differenti sub-set di variabili selezionati con il METODO B.	111
Fig. 5.7 Variazione del coefficiente di correlazione per il test set rispetto ai differenti sub-set di variabili selezionati con il METODO B.	111
Fig. 5.8 Confronto tra i valori dei coefficienti di correlazione a parità di numero di variabili per i sub-set determinati con entrambi i metodi di selezione (sub-sets METODO A: PC1var4, PC2var10, PC3var14, PC4var18, PC5var21, PC7var23; sub-sets METODO B: DS1VAR4, DS15var10, DS11var14, DS7var18, DS4var21, DS2var23) relativamente a: Validation set (A) e test set (B) per le diverse configurazioni di ANN (10 neuroni, 20 neuroni, 30 neuroni).	112
Fig. 5.9 Percentuale di varianza spiegata rispetto al numero di componenti per il metodo PLSR.	112
Fig. 5.10 Scatter plot relativo ai valori calcolati (VAR30PLS8) rispetto ai valori osservati (obs), del miglior risultato ottenuto con il metodo PLSR, calcolato con 8 componenti.	113
Fig. 5.11 Grafico relativo alla variazione del coefficiente di correlazione R per il metodo PLSR, al variare del numero di componenti trattenute.	113
Fig. 5.12 Scatter plots relativi ai valori calcolati rispetto ai valori predetti, per il set di variabili DS0VAR30 a cui è stata applicata una ANN costituita da 10 neuroni (A) e curva della sua performance.	114
Fig. 5.13 Immagine relativa alla mappa dello spessore del solum, calcolata con il dataset DS0VAR30 e una ANN costituita da 10 neuroni.	115
Fig. 5.14 Istogramma relativo allo spessore del solum dei profili utilizzati per il validation set rispetto ai relativi valori calcolati con il metodo NN10DS0VAR30.	116
Fig. 5.15 Istogramma relativo allo spessore del solum dei profili utilizzati per il validation set rispetto alla differenza ( $\Delta$ ) tra il valore calcolato ed il valore osservato.	116
Fig. 5.16 Area ove è stato campionato il profilo P33_2014, con l'indicazione dell'ubicazione del profilo e dei profili P34 e 53_2014; (A) immagine del versante sul quale è stato campionato il P33_2014 (B).	117
Fig. 5.17 Istogramma relativo alla differenza tra valore osservato e valore predetto per i punti con suolo zero.	118

Fig. 5.18 Ubicazione dei punti con suolo zero con le peggiori predizioni (A, C, E) e indicazione della relativa situazione morfologica.	119
Fig. 5.19 Confronto, per ciascuna misura di profilo osservato (obs), dei residui (osservato - calcolato) relativi ai migliori risultati ottenuti.	119
Fig. 6.1 Formazione delle Arenarie del Monte Vallassa.	127
Fig. 6.2 Profilo rilevato in località Panzini con le caratteristiche tipiche dei suoli presenti sui plateau della Placca di Pizzocorno-Pietragravina.(A). Profilo rilevato su una costa della Placca di Pizzocorno-Pietragravina. (B).	128
Fig. 6.3 Profilo rilevato in una area di accumulo dei blocchi derivanti dalla demolizione della formazione delle Arenarie del Monte Vallassa (A). Profilo con frammenti di roccia della formazione delle Arenarie del Monte Vallassa (B).	129
Fig. 6.4 Profilo rilevato in prossimità di un impluvio, con blocchi di grandi dimensione.	130
Fig. 6.5 Erosione di tipo laminare sui versanti più ripidi in prossimità degli impluvi (A) e (B).	130
Fig. 6.6 Mappa degli areali delle aree soggette a dinamica di tipo calanchivo pubblicata da Philips (Phillips 1998) (A) e da D. E. Alexander (Alexander 1980) (B).	131
Fig. 6.7 Aree soggette a dinamica di tipo calanchivo parzialmente occupate da vegetazione pioniera (A) (B) (C) (D).	136
Fig. 6.8 Il denudamento dell'apparato radicale indica la riattivazione del processo erosivo, successivamente ad un periodo di stabilità (A) e (B).	137
Fig. 6.9 Il particolare portamento della Roverella indica che la presenza di una dinamica erosiva lenta ma continua che costringe la pianta ad uno sviluppo orizzontale (A). Area all'interno di un calanco ove l'accumulo di materiale colluviale ha creato le condizioni per la crescita di vegetazione arborea (B).	137
Fig. 6.10 Profilo rilevato su un'area agricola incolta in una zona soggetta a soliflusso (A) Profilo rilevato a 30 m circa dal P15_2015 su un'area apparentemente più stabile (B).	138
Fig. 6.11 Fenomeni di erosione in corrispondenza della formazione delle Breccie di Baiso in località di Costa Croce. Sul versante con il fianco rivolto a ovest è presente un tipico calanco classificabile di tipo A, sul versante esposto a nord, a circa 200 m, è presente erosione di tipo laminare ed evidenze di rill, è inoltre presente una piccola frana (A). Nella parte soprastante, esposta a nord-ovest, in un'area agricola incolta, è presente erosione da soliflusso a poca distanza si è generato un gully.	139
Fig. 6.12 Profilo rilevato all'interno del gully (P14_2015) (A) Gully (B).	140
Fig. 6.13 Profilo rilevato Su un versante soggetto a erosione di tipo soil creep, in corrispondenza della Formazione delle Breccie di Baiso.	140
Fig. 6.14 Area a a nord-ovest di Sant'Albano (A) Biancana a "pop-corn" (B).	141
Fig. 6.15 Area a a nord-ovest di Sant'Albano, con l'ubicazione della Biancana a "pop-corn" della Fig. 6.14 (B) e del profilo P62_2015.	141
Fig. 6.16 Profilo rilevato sul versante della struttura assimilabile a una biancana (A) Profilo rilevato sulla parte sommitale della struttura assimilabile a una biancana (B).	142
Fig. 6.17 Struttura assimilabile a una biancana ubicata sul fondovalle della Val di Nizza, in località Risaia.	142
Fig. 6.18 Calanco mammellonare (A). Erosione di tipo laminare associata a fenomeni di rill (B).	143
Fig. 6.19 Modello distributivo dei calanchi e delle biancane (A). Modello esplicativo della formazione delle biancane e dei calanchi mammellonari (B).	143
Fig. 6.20 Suffosione in località Sant'Albano.	144
Fig. 6.21 Suffosione in località Lumello (A) e (B).	145
Fig. 6.22 Frane superficiali in Val Ardivestra (A) (B).	147
Fig. 6.23 Punti delle aree pilota, individuate nel corso dell'attività di campionamento.	150
Fig. 6.24 Mappa dell'erosione del suolo prodotta con il metodo pseudoQuadratic, con il perimetro delle due valli.	156
Fig. 6.25 Confronto tra la distribuzione spaziale delle classi, per le mappe dell'erosione del suolo, ottenute con i metodi diagQuadratic (A), diagLinear(B), pseudoLinear (C), Linear (D), con il perimetro delle due valli.	157
Fig. 6.26 Aree soggette a dinamica calanchiva elaborate con il metodo pseudoQuadratic, con i lineamenti delle formazioni geologiche.	159
Fig. 6.27 Aree soggette a dinamica calanchiva elaborate con il metodo diagQuadratic, con i lineamenti delle formazioni geologiche (cfr. fig. 2.6).	159
Fig. 6.28 Scree plot relativo alla principal component analysis eseguito sul set di 30 variabili geomorfometriche.	162

Fig. 6.29 Loadings plot relativo alla principal component analysis eseguito sul set di 30 variabili geomorfometriche.	162
Fig. 6.30 Output della ANN con 10 neuroni, applicata al sub-set di variabili DSPC2: matrice di confusione (A) curva ROC (classe 1 “evento assente” classe 2 “ punto di innesco” (B).	168
Fig. 6.31 Output della ANN con 30 neuroni, applicata al sub-set di variabili DSPC2: matrice di confusione (A) curva ROC (classe 1 “evento assente” classe 2 “ punto di innesco” (B).	168
Fig. 6.32 Performance della ANN applicata al sub-set di variabili DSPC2: ANN con 10 neuroni (A) ANN con 30 neuroni (B).	169
Fig. 6.33 Mappa della PROPENSIONE ALLE FRANE SUPERFICIALI realizzata con il sub-set di variabili DSPC2 e con la ANN con 10 neuroni.	169
Fig. 6.34 Mappa della PROPENSIONE ALLE FRANE SUPERFICIALI realizzata con il sub-set di variabili DSPC2 e con la ANN con 10 neuroni, con l’ubicazione delle frane superficiali impiegate per le elaborazioni.	170
Fig. 6.35 Area soggetta ad erosione del suolo (A) e relativa predizioni ottenuta con il metodo pseudoQuadratic (B).	171
Fig. 6.36 Area soggetta ad erosione del suolo (A) e relativa predizioni ottenuta con il metodo pseudoQuadratic (B).	172
Fig. 6.37 Mappa della PROPENSIONE ALLE FRANE SUPERFICIALI con indicazione delle aree franose dell’inventario IFFI.	174
Fig. 6.38 Indicazione delle aree franose riportate nel catalogo IFFI.	175
Fig. 7.1 Profilo di suolo rappresentativo della classe AaLSmp.	182
Fig. 7.2 Profilo di suolo rappresentativo della classe AaLSms.	183
Fig. 7.3 Profilo di suolo rappresentativo della classe AbLSp.	183
Fig. 7.4 Profilo di suolo rappresentativo della classe AbHSs.	184
Fig. 7.5 Profilo di suolo rappresentativo della classe AbHSms.	184
Fig. 7.6 Profilo di suolo rappresentativo della classe AmHSmp.	185
Fig. 7.7 Profilo di suolo rappresentativo della classe AmHSp.	185
Fig. 7.8 Profilo di suolo rappresentativo della classe AbHSp.	186
Fig. 7.9 Profilo di suolo rappresentativo della classe AmLSp.	186
Fig. 7.10 Profilo di suolo rappresentativo della classe AmLSs.	187
Fig. 7.11 Profilo di suolo rappresentativo della classe AmLSms.	187
Fig. 7.12 Profilo di suolo rappresentativo della classe AmLSmp.	188
Fig. 7.13 Mappa dei suoli secondo la tassonomia WRB ottenuta con il metodo diagQuadratic.	194
Fig. 7.14 Mappa delle tipologie di suolo (basata sulla tessitura e lo spessore del solum), ottenuta con il metodo diagQuadratic.	195

## ELENCO DELLE TABELLE

Tabella 2.1 Elenco delle Unità Geologiche presenti nell’area di studio, con la sigla adottata nel Foglio 178-Voghera e con il relativo areale.	21
Tabella 2.2 Confronto degli areali adibiti ad uso agricolo e occupati da vegetazione spontanea relativi alla carta DUSAF del 1954 con la carta DUSAF del 2012.	25
Tabella 2.3 Areali delle classi d’uso del suolo della carta DUSAF del 1954.	26
Tabella 2.4 Areali delle classi d’uso del suolo della carta DUSAF del 2012.	26
Tabella 3.1 Denominazione e sigla delle variabili geomorfometriche calcolate.	32
Tabella 3.2 A Statistica descrittiva per le variabili geomorfometriche.	35
Tabella 3.2 B	36
Tabella 3.1 Denominazione e sigla delle variabili geomorfometriche calcolate.	32
Tabella 3.2 A Statistica descrittiva per le variabili geomorfometriche.	35
Tabella 3.2 B	36
Tabella 3.3 Numero di valori negativi e positivi per le variabili che presentano numeri con entrambi i segni.	36
Tabella 3.4 A Numero di valori maggiori di: (media+2 deviazione standard).	37
Tabella 3.4 B	37
Tabella 3.5 A Numero di valori maggiori di: (media+3 deviazione standard).	37

Tabella 3.5 B	38
Tabella 3.6 Indice di correlazione lineare di Pearson.	61
Tabella 3.7 Matrice dei p value del test: ipotesi di non correlazione contro l'alternativa di correlazione diversa da zero.	62
Tabella 3.7 Matrice dei p value del test: Continua.	62
Tabella 3.8 Matrice relativa al test di Belsley.	65
Tabella 3.8 Matrice relativa al test di Belsley.	66
Tabella 4.1 Denominazione e sigla delle variabili geomorfometriche utilizzate.	71
Tabella 4.2 Valori dei loadings delle variabili geomorfometriche, distinti secondo l'importanza ed il segno.	78
Tabella 4.3 Percentuale di varianza spiegata dalle prime 10 PCs.	80
Tabella 4.4 Area media e numero di poligoni secondo il numero di classi con cui è stata suddivisa l'area di studio.	82
Tabella 4.5 Unità geologiche presenti nell'area di studio, con l'areale in km <sup>2</sup> e in percentuale riclassificate.	83
Tabella 4.6 Esito del test non parametrico Two-sample Kolmogorov-Smirnov.	92
Tabella 5.1 Valori dei loadings delle variabili geomorfometriche, distinti secondo l'importanza ed il segno. Per ciascuna componente è riportato l'autovalore (eigenvalue) e la percentuale di varianza spiegata (percent_explained).	104
Tabella. 5.2 Risultati dell'applicazione del METODO A. Sono evidenziati i valori dei loadings per le prime sette componenti (PC1-PC7), gli outliers per ciascuna componente (valori $\geq \pm 1,5$ deviazione standard) e i sub-set di variabili selezionate.	107
Tabella. 5.3 Risultati dell'applicazione del METODO B. Nella tabella le variabili sono disposte secondo l'ordinamento decrescente dell'indice P.	108
Tabella. 5.4 Coefficienti di correlazione ottenuti con i sub-set di variabili selezionati con il METODO A, secondo le differenti topologie di ANN.	109
Tabella. 5.5 Coefficienti di correlazione ottenuti con i sub-set di variabili selezionati con il METODO B, secondo le differenti topologie di ANN.	110
Tabella. 5.6 Valori dei coefficiente di correlazione R ottenuti, con il metodo PLSR, variando il numero di componenti trattenute.	113
Tabella. 5.7 Assegnazione della classe di spessore di suolo secondo la scala adottata da ERSF (Regione Lombardia), per i casi che presentano l'errore maggiore.	118
Tabella. 6.1 A Matrice di confusione per le classi di erosione del suolo, relativa al metodo pseudoQuadratic.	155
Tabella. 6.1 B Valori della Sensibility (Sn), Specificity (Sp) e del Misclassification risk (Mr), per ciascuna classe di erosione del suolo, relativi al metodo pseudoQuadratic.	155
Tabella. 6.2 A Matrice di confusione per le classi di erosione del suolo, relativa al metodo diagQuadratic.	155
Tabella. 6.2 B Valori della Sensibility (Sn), Specificity (Sp) e del Misclassification risk (Mr), per ciascuna classe di erosione del suolo, relativi al metodo diagQuadratic.	155
Tabella. 6.3 A Matrice di confusione per le classi di erosione del suolo, relativa al metodo Linear.	155
Tabella. 6.3 B Valori della Sensibility (Sn), Specificity (Sp) e del Misclassification risk (Mr), per ciascuna classe di erosione del suolo, relativi al metodo Linear.	155
Tabella. 6.4 A Matrice di confusione per le classi di erosione del suolo, relativa al metodo diagLinear.	155
Tabella. 6.4 B Valori della Sensibility (Sn), Specificity (Sp) e del Misclassification risk (Mr), per ciascuna classe di erosione del suolo, relativi al metodo diagLinear.	155
Tabella. 6.5 A Matrice di confusione per le classi di erosione del suolo, relativa al metodo pseudoLinear.	156
Tabella. 6.5 B Valori della Sensibility (Sn), Specificity (Sp) e del Misclassification risk (Mr), per ciascuna classe di erosione, relativi al metodo pseudoLinear.	156
Tabella. 6.6 A Matrice di confusione relativa alle classi utilizzate per la redazione della mappa dell'erosione calanchiva, relativamente al metodo pseudoQuadratic.	157
Tabella. 6.6 B Valori della Sensibility (Sn), Specificity (Sp) e del Misclassification risk (Mr), per ciascuna classe utilizzata per la redazione della mappa dell'erosione calanchiva, relativi al metodo pseudoQuadratic.	157
Tabella. 6.7 A Matrice di confusione relativa alle classi utilizzate per la redazione della mappa dell'erosione calanchiva, relativamente al metodo diagQuadratic.	158
Tabella. 6.7 B Valori della Sensibility (Sn), Specificity (Sp) e del Misclassification risk (Mr),	

per ciascuna classe utilizzata per la redazione della mappa dell'erosione calanchiva, relativi al metodo diagQuadratic.	158
Tabella. 6.8 A Matrice di confusione relativa alle classi utilizzate per la redazione della mappa dell'erosione calanchiva, relativamente al metodo Linear.	158
Tabella. 6.8 B Valori della Sensibility (Sn), Specificity (Sp) e del Misclassification risk (Mr), per ciascuna classe utilizzata per la redazione della mappa dell'erosione calanchiva, relativi al metodo Linear.	158
Tabella. 6.9 A Matrice di confusione relativa alle classi utilizzate per la redazione della mappa dell'erosione calanchiva, relativamente al metodo diagLinear.	158
Tabella. 6.9 B Valori della Sensibility (Sn), Specificity (Sp) e del Misclassification risk (Mr), per ciascuna classe utilizzata per la redazione della mappa dell'erosione calanchiva, relativi al metodo diagLinear.	158
Tabella. 6.10 A Matrice di confusione relativa alle classi utilizzate per la redazione della mappa dell'erosione calanchiva, relativamente al metodo pseudoLinear.	158
Tabella. 6.10 B Valori della Sensibility (Sn), Specificity (Sp) e del Misclassification risk (Mr), per ciascuna classe utilizzata per la redazione della mappa dell'erosione calanchiva, relativi al metodo pseudoLinear.	158
Tabella. 6.11 Valori dei loadings dei punti di innesco delle frane superficiali, distinti secondo l'importanza ed il segno. Per ciascuna componente è riportato l'autovalore (eigenvalue) e la percentuale di varianza spiegata (percent_explained).	164
Tabella. 6.12 Determinazione dei sub-sets di variabili. Sono evidenziati i valori dei loadings per le prime quattro componenti (PC1-PC4), gli outliers per ciascuna componente (valori > ± deviazione standard) e i sub-set di variabili selezionate evidenziate in giallo.	166
Tabella. 6.13 Percentuale dei punti di innesco classificati correttamente per ciascun sub-set di variabili, derivanti dai risultati delle matrici di confusione, secondo le differenti topologie di ANN.	167
Tabella. 7.1 Categorie tassonomiche assegnate ai profili campionati sigla utilizzata nella mappa dei suoli.	180
Tabella. 7.2 Categorie tipologiche di suolo assegnate secondo le caratteristiche tessiturali e dello spessore.	181
Tabella. 7.3 Matrice di confusione per le classi tassonomiche, relativa al metodo Linear.	189
Tabella. 7.4 Valori della Sensibility (Sn), Specificity (Sp) e del Misclassification risk (Mr), per ciascuna classe tassonomica, relativi al metodo Linear.	189
Tabella. 7.5 Matrice di confusione per le classi tassonomiche, relativa al metodo diagLinear.	189
Tabella. 7.6 Valori della Sensibility (Sn), Specificity (Sp) e del Misclassification risk (Mr), per ciascuna classe tassonomica, relativi al metodo diagLinear.	190
Tabella. 7.7 Matrice di confusione per le classi tassonomiche, relativa al metodo diagQuadratic.	190
Tabella. 7.8 Valori della Sensibility (Sn), Specificity (Sp) e del Misclassification risk (Mr), per ciascuna classe tassonomica, relativi al metodo diagQuadratic.	190
Tabella. 7.9 Matrice di confusione per le classi tassonomiche, relativa al metodo pseudoLinear (Non-Error-Rate (NER) Error-Rate (ER) e Misclassification risk (MR)).	190
Tabella. 7.10 Valori della Sensibility (Sn), Specificity (Sp) e del Misclassification risk (Mr), per ciascuna classe tassonomica, relativi al metodo pseudoLinear.	191
Tabella. 7.11 Matrice di confusione per le classi tassonomiche, relativa al metodo pseudoQuadratic.	191
Tabella. 7.12 Valori della Sensibility (Sn), Specificity (Sp) e del Misclassification risk (Mr), per ciascuna classe tassonomica, relativi al metodo pseudoQuadratic.	191
Tabella. 7.13 Matrice di confusione per le tipologie di suolo, relativa al metodo diagQuadratic.	192
Tabella. 7.14 Valori della Sensibility (Sn), Specificity (Sp) e del Misclassification risk (Mr), per le tipologie di suolo, relativi al metodo diagQuadratic.	192
Tabella. 7.15 Matrice di confusione per le tipologie di suolo, relativa al metodo Linear.	192
Tabella. 7.16 Valori della Sensibility (Sn), Specificity (Sp) e del Misclassification risk (Mr), per le tipologie di suolo, relativi al metodo Linear.	192
Tabella. 7.17 Matrice di confusione per le tipologie di suolo, relativa al metodo diagLinear.	193
Tabella. 7.18 Valori della Sensibility (Sn), Specificity (Sp) e del Misclassification risk (Mr),	

per le tipologie di suolo, relativi al metodo diagLinear.	193
Tabella. 7.19 Matrice di confusione per le tipologie di suolo, relativa al metodo pseudoLinear.	193
Tabella. 7.20 Valori della Sensibility (Sn), Specificity (Sp) e del Misclassification risk (Mr), per le tipologie di suolo, relativi al metodo pseudoLinear .	193
Tabella. 7.21 Matrice di confusione per le tipologie di suolo, relativa al metodo pseudoQuadratic.	194
Tabella. 7.22 Valori della Sensibility (Sn), Specificity (Sp) e del Misclassification risk (Mr), per le tipologie di suolo, relativi al metodo pseudoQuadratic.	194

## INTRODUZIONE

*“Accurate spatial soil information is urgently needed for dealing with the global issues such as agricultural production, environmental pollution, food security, water security, and human health.”* (Zhang et al. 2016).

*“Early civilizations in Asia, Africa, Europe, and North America had fundamental knowledge of relationships between basic soil properties and the suitability of an area for crop production and settlement. Early soil maps were desirable for purposes of land valuation for taxation, agronomic planning, and even in military operations”* (Brevik et al. 2016).

L'attuale dinamica di crescita della popolazione e di occupazione del territorio comporta scelte oculate nella gestione delle risorse limitate e in particolare di quelle non rinnovabili quali il suolo. La disponibilità di informazioni dettagliate riguardanti la distribuzione spaziale dei suoli e dei loro attributi è, oggi, uno dei fondamenti per attuare una corretta gestione dell'ambiente e del territorio. Tali dati informativi sono già da tempo utilizzati in diversi contesti socio-economici quali: la pianificazione territoriale, la gestione agronomica, il *risk assessment*, la *land evaluation*, la gestione forestale. Queste informazioni sono sollecitate dalle esigenze amministrative dei moderni sistemi antropizzati, le cui dinamiche evolutive sono condizionate da forze antropo-dirette. Queste hanno un effetto superiore a quello dei processi naturali, imponendo variazioni per le quali è difficile prevedere le conseguenze sulle condizioni di vita nel futuro. Inoltre, è recentemente cresciuto l'interesse per dette informazioni, a causa dei rischi introdotti dal cambiamento climatico e dall'inquinamento, nonché per la loro importanza per i servizi ecosistemici, quali gli effetti diretti e indiretti sulla salute dell'uomo. Nelle Scienze della Terra le informazioni sulle caratteristiche dei suoli (es: profondità, numero, spessore, caratteristiche degli orizzonti, proprietà idrauliche) hanno numerose applicazioni quali: la realizzazione di mappe geologiche, le analisi paleoclimatiche, lo studio sulla suscettibilità di un territorio alle frane (Brevik et al. 2016).

Per la mappatura dei suoli, oltre ai processi tradizionali, da tempo sono utilizzati metodi che in letteratura sono associati alla sub-disciplina della Scienza del Suolo nota come *Digital Soil Mapping* (in seguito DSM). Nonostante la ormai copiosa produzione scientifica di settore, nell'applicazione del DSM esistono diversi aspetti che necessitano ulteriori approfondimenti (Brevik et al. 2016), infatti, è tuttora attivo il dibattito per individuare un protocollo di applicazione (Hempel et al. 2008; Zhang et al. 2016). In letteratura è stato sottolineato quali sono i parametri che concorrono alla precisione di un prodotto cartografico: l'accuratezza delle informazioni sul suolo, l'accuratezza delle variabili ambientali che influenzano lo sviluppo pedogenetico, la precisione del sistema di inferenza (Carré et al. 2007). Malgrado i progressi tecnologici e quelli conseguiti nel DSM (Minasny e McBratney 2016), indipendentemente dall'approccio utilizzato, nel moderno processo di mappatura dei suoli vi sono due aspetti - associabili a fattori di natura economica - che possono essere considerati fattori limitanti, i quali influenzano la qualità del prodotto finale a priori. Il primo riguarda lo svolgimento dell'attività di campionamento e delle analisi chimico-fisiche, necessarie per la caratterizzazione dei campioni di suolo e dei profili pedologici, il secondo la qualità delle variabili coinvolte nel processo di mappatura (Fig. 1).

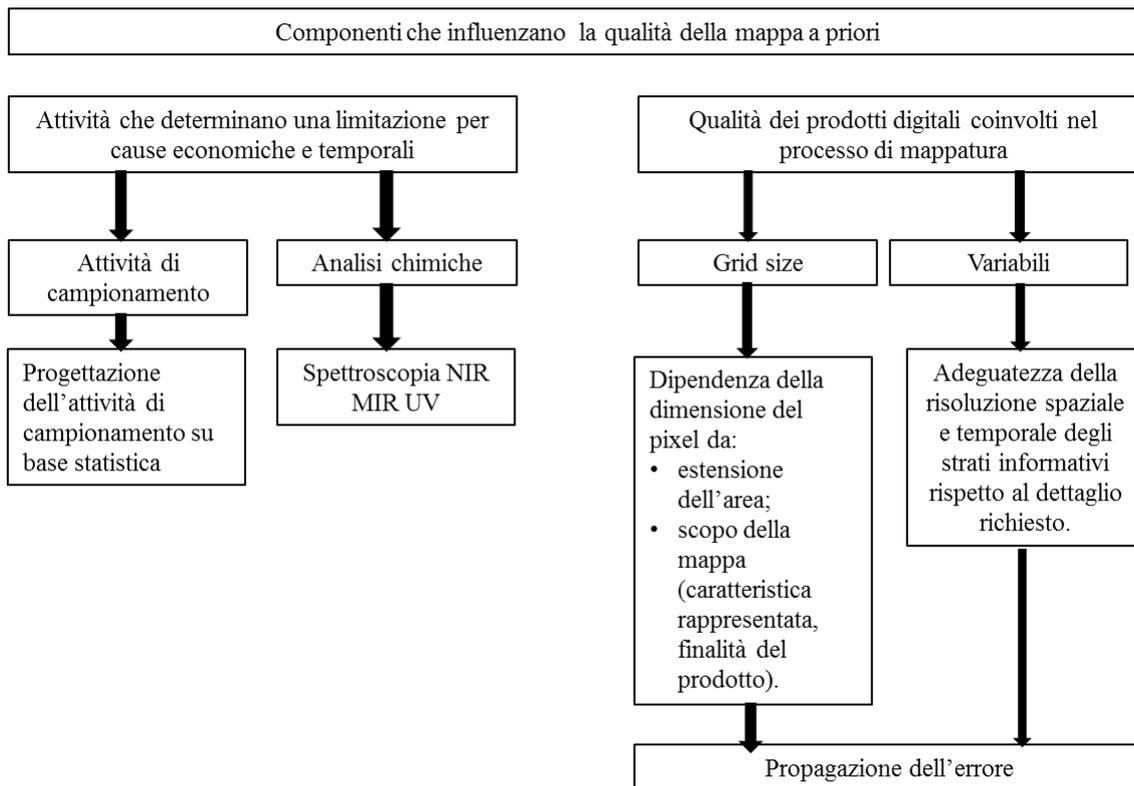


Fig. 1 Componenti coinvolte nel processo di mappatura dei suoli che influenzano la qualità del prodotto finale a priori.

In letteratura, sono pochi i casi in cui è affrontata la gestione dell'attività di campionamento. Questa ha differenti approcci a seconda della dimensione dell'area da cartografare (regionale, di bacino o di parcella), dell'oggetto da cartografare, del dettaglio da raggiungere e del metodo statistico-matematico utilizzato per derivare il prodotto finale (geostatistica o statistica).

Il problema relativo al rilevamento dei suoli è generato dalla necessità di procedere con un approccio efficiente, tenendo conto della variabilità presente nell'area di studio e minimizzando i costi. Applicando le tecniche di DSM, la soluzione può essere costruita nella fase propedeutica alla realizzazione di una mappa, con l'effetto di conoscere la stima dell'incertezza che ne deriva sul risultato finale (Hartemink et al. 2008). Occorre, tuttavia, tener presente l'aspetto pratico: infatti, la progettazione teorica dell'attività di campionamento non può prescindere dalla considerazione degli aspetti logistici che necessariamente si verificano nel corso dello svolgimento di una attività di campo. A tal riguardo, data un'area di studio, in questo lavoro si vuole contribuire a rispondere alle seguenti domande:

- come progettare una campagna di rilevamento contenendo il numero di punti di campionamento?
- come progettare una campagna di rilevamento considerando gli aspetti pratico-logistici?

Indipendentemente dall'approccio utilizzato per la creazione di mappe tematiche, è doveroso evidenziare che occorre considerare la scala geografica di riferimento, dalla quale non si può prescindere per la scelta del dettaglio che si vuole ottenere. Tuttavia il linguaggio del DSM differisce da quello generalmente utilizzato nella cartografia convenzionale e il concetto di scala è sostituito dal concetto di risoluzione spaziale (McBratney et al. 2003).

Per quanto riguarda gli effetti connessi alle caratteristiche degli strati informativi, utilizzati per costruire una mappa, occorre considerare due aspetti: la qualità del prodotto finale è strettamente

legata alla dimensione dei pixel del raster; il risultato finale dipende dalla disponibilità e dalla qualità dei prodotti digitali, costituenti le variabili, coinvolte nel processo di mappatura.

Tenendo presente che la rappresentazione cartografica è solo un'approssimazione della realtà, nel processo di DSM occorre considerare la peculiare variabilità spaziale del suolo. Infatti, indipendentemente dal contesto geografico considerato e dalle caratteristiche del suolo, quest'ultimo presenta generalmente una variazione continua, che può essere rilevata anche su brevi distanze (es. pochi metri) e che può essere interrotta, a volte, da discontinuità. Questa variabilità avviene sia per quanto riguarda le caratteristiche parametriche che gli aspetti qualitativi. Pertanto, la qualità del prodotto finale (la mappa) dipende da quanto la *gride size* è in grado di accordare la reale variabilità del fenomeno che si intende rappresentare, considerando che in un *raster* il suolo entro un dato pixel è percepito come omogeneo.

Riguardo l'applicazione delle tecniche del DSM occorre considerare che per cartografia di dettaglio si intende una cartografia discretizzata con pixel di dimensione inferiore ai 20 m (McBratney et al. 2003; Lagacherie 2008). La scelta della *gride size* è tuttavia il risultato di un compromesso tra il dettaglio che si vuole raggiungere, la dimensione dell'area di studio e la capacità computazionale delle macchine impiegate; queste, a causa dell'alto rapporto prezzo-prestazioni, ancora oggi hanno capacità computazionali che limitano la complessità delle elaborazioni e delle relative produzioni. A riguardo, in questo lavoro le elaborazioni sono state eseguite a scala di bacino con un DEM discretizzato a 10 m.

Per quanto riguarda l'impiego degli strati informativi, tipicamente utilizzati per modellare tematismi relativi al suolo (es: dati climatici, vegetazionali, litologici), occorre considerare alcuni importanti aspetti. Il primo riguarda l'esistenza di queste informazioni: per conoscere alcuni dei fattori che hanno influenza sui processi pedogenetici sono necessarie attività di campionamento e monitoraggio non sempre possibili. Questa banale considerazione ha importanti effetti dal punto di vista applicativo e concettuale riguardo il DSM, in particolare quando si opera su vaste aree geografiche. Infatti, a seconda dell'area geografica una determinata variabile può non essere disponibile o non è detto che la sua qualità sia costante per l'intera area. Un secondo aspetto riguarda la scala o risoluzione spaziale: il dettaglio delle variabili deve essere adeguato alla risoluzione del prodotto che si vuole ottenere. Nel caso in cui la risoluzione spaziale della variabile sia inferiore rispetto a quella del prodotti da ottenere bisognerebbe domandarsi se tale variabile costituisca un fattore limitante o se l'inadeguatezza possa essere trascurata. Il terzo aspetto riguarda la qualità della variabile in merito alla risoluzione temporale: in questo caso l'effetto potrebbe essere ancora più impattante sul risultato finale. Infatti, nel considerare alcune variabili, quali l'uso del suolo, si dovrebbe vagliare la contestualizzazione temporale relativa alla variabile stessa. Un altro aspetto deriva dai precedenti: la qualità delle variabili considerate si riflette sulla qualità del prodotto finale in termini di propagazione dell'errore (Carré et al. 2007).

Mentre gli strati informativi citati possono essere carenti, il Digital Elevation Model (in seguito DEM) – ove disponibile e di qualità – può essere utilizzato per calcolare una serie di variabili, utilizzabili nel processo di derivazione delle mappe. A riguardo, in questo lavoro si cerca di rispondere alla seguente domanda: data un'area di studio appartenente alla catena appenninica, è

possibile costruire mappe di dettaglio relative a tematiche del suolo prescindendo da informazioni ausiliarie e utilizzando la sola informazione derivante dal DEM?

Per quanto concerne gli aspetti metodologici connessi alle applicazioni di DSM occorre considerare che gli studi sino ad ora condotti hanno dimostrato che: *“there is not one method emerging that performs best to estimate multiple soil properties/classes in different geographic regions”* (Grunwald 2010). Riguardo i metodi di modellazione esiste una vasta gamma di approcci (McBratney et al. 2003): la distinzione più importante riguarda la differenza tra i modelli statistici non spaziali e la geostatistica. A riguardo occorre sottolineare che tale scelta solitamente non è discrezionale. Infatti, come sottolineato da Lagacherie *“the application of geostatistics to large areas appears still limited by soil samplings that are not dense enough to capture the patterns of the soil variations. The development of numerical models of soil prediction has reached the point where a lot of functions are now available. However there is not enough comparisons of their performances over the variety of situations that may be practically encountered in Digital Soil Mapping. These comparisons would provide the necessary expertise in selecting the most appropriate function given the objective of the study (spatial extent, resolution) the data context (soil covariates, soil sampling, available pedological knowledge and skills), and the nature of the soil variation. Beside, more sophisticated Digital Soil Mapping functions than the one presented are certainly required (i) to enhance the synergy between soil knowledge and pedometric models (ii) to tackle multiscaled variations of the soil cover and (iii) to deal with evolutions of soil properties with time that are now monitored extensively”* (Lagacherie 2008).

A riguardo, in questo lavoro vengono forniti ulteriori esempi per ampliare la conoscenza riguardo l'applicabilità dei diversi metodi statistici alle tematiche del DSM.

Occorre sottolineare che è di fondamentale importanza, nell'applicare le tecniche del DSM, fornire il grado di accuratezza del prodotto, ancor più se le mappe sono utilizzate nell'ambito dei processi decisionali: *“defining and applying a common accuracy assessment framework is probably the greatest challenge of Digital Soil Mapping on its way toward practical applications”* (Lagacherie 2008). A riguardo, in letteratura è stato evidenziato che nel DSM può essere fornita la qualità dei risultati (es: varianza dell'errore per kriging, coefficiente di determinazione per la regressione), tuttavia è importante rendersi conto che tali indicazioni sono spesso basate su ipotesi del modello che non possono essere verificate nella realtà. Inoltre, detti indici di valutazione sono spesso calcolati con gli stessi dati con cui è stato costruito il modello della mappa. La convalida con un dataset indipendente è altamente preferibile, quando possibile (Lagacherie 2008).

A riguardo in questo lavoro la validazione delle mappe è stata effettuata anche con un set di dati non utilizzato nelle elaborazioni.

## **IN SINTESI**

In sintesi, l'obiettivo di questo lavoro è stato quello di ottenere una base cartografica relativa agli aspetti principali del suolo di un'area appenninica, con un alto livello di dettaglio, con un basso costo sia in termini di tempo che economici e di verificarne l'attendibilità, per apportare informazione di base utile per comprendere i processi evolutivi dell'area di studio.

Nel contempo si è voluto verificare quanto l'applicazione di DSM adottata risponda agli obiettivi di economicità, affidabilità e coerenza dei risultati rispetto agli obiettivi e scopi testé citati. Inoltre, si è voluto verificare quanto i prodotti cartografici siano realizzabili indipendentemente dall'esistenza degli strati informativi relativi ai diversi fattori che concorrono alla pedogenesi (o i loro *proxy*) e che generalmente sono utilizzati nelle procedure di cartografia tradizionale. L'interesse nasce dal voler superare un problema ubiquitario riguardante dette variabili, che spesso sono di dettaglio non idoneo o non sono disponibili, a seconda dell'area geografica considerata. Pertanto, in questo lavoro la produzione cartografica è stata realizzata utilizzando un'informazione elementare: in pratica, sono state realizzate alcune mappe tematiche con la sola informazione derivante dal *digital elevation model* (DEM) e con tecniche statistiche di semplice applicazione. Infatti, sono stati calcolati una serie di parametri geomorfometrici che costituiscono le variabili indipendenti, impiegate nei processi di elaborazione dei prodotti cartografici. Detti parametri non sono stati considerati per il loro specifico significato fisico, né per la loro eventuale influenza sui risultati dei processi di pedogenesi. Di fatto le variabili sono state impiegate come un set di valori associati a ciascun pixel, la cui rilevanza, al fine di ottenere la migliore approssimazione per i prodotti cartografici, è stata valutata con tecniche statistiche.

## **Bibliografia**

- Brevik, Eric C.; Calzolari, Costanza; Miller, Bradley A.; Pereira, Paulo; Kabala, Cezary; Baumgarten, Andreas; Jordán, Antonio (2016): Soil mapping, classification, and pedologic modeling. History and future directions. In: *Geoderma*, 264, pp. 256–274. DOI: 10.1016/j.geoderma.2015.05.017.
- Carré, F.; McBratney, Alex B.; Mayr, Thomas; Montanarella, Luca (2007): Digital soil assessments. Beyond DSM. In: *Geoderma* n. 1-2, 142, pp. 69–79. DOI: 10.1016/j.geoderma.2007.08.015.
- Grunwald, S. (2010): Current state of digital soil mapping and what is next. In: *Digital Soil Mapping*: Springer, pp. 3–12.
- Hartemink, Alfred E.; McBratney, Alex B.; Lourdes Mendonça-Santos, Maria de; Ahrens, Robert J. (2008): *Digital soil mapping with limited data*: Springer Berlin, Germany.
- Hempel, J. W.; Hammer, R. D.; Moore, A. C.; Bell, J. C.; Thompson, J. A.; Golden, M. L. (2008): Challenges to digital soil mapping. In: *Digital Soil Mapping with Limited Data*: Springer, pp. 81–90.
- Lagacherie, Philippe (2008): Digital soil mapping: a state of the art. In: *Digital soil mapping with limited data*: Springer, pp. 3–14.
- McBratney, A.B; Mendonça Santos, M.L; Minasny, B. (2003): On digital soil mapping. In: *Geoderma* n. 1-2, 117, pp. 3–52. DOI: 10.1016/S0016-7061(03)00223-4.
- Minasny, Budiman; McBratney, Alex B. (2016): Digital soil mapping: A brief history and some lessons. In: *Geoderma*, 264, pp. 301–311.
- Zhang, Gan-Lin; Liu, Feng; Song, Xiao-Dong; Zhao, Yu-Guo (2016): Digital Soil Mapping Across Paradigms, Scales, and Boundaries: A Review. In: *Digital Soil Mapping Across Paradigms, Scales and Boundaries*: Springer, pp. 3–10.

# CAPITOLO 1

## DIGITAL SOIL MAPPING

### 1.1. Introduzione

A causa della crescente richiesta di informazioni riguardanti i suoli, negli ultimi vent'anni sono stati numerosi i tentativi di cartografare le loro proprietà con modalità non tradizionali ed in particolare con metodi che sfruttano le tecnologie digitali con il fine, non ultimo, di integrare le informazioni cartografate in più complessi e articolati sistemi di informazione geografica (GIS).

Dette applicazioni trovano la loro collocazione nella sub-disciplina della Scienza del Suolo nota come Digital Soil Mapping (DSM). Per questa disciplina, che si è evoluta nel corso del tempo (Grunwald 2010), il primo workshop mondiale si è tenuto a Montpellier (settembre 2004), ove, tra l'altro, ne è stata formalizzata la definizione: *“Digital Soil Mapping (DSM) can be defined as the creation and population of spatial soil information systems by numerical models inferring the spatial and temporal variations of soil types and soil properties from soil observation and knowledge and from related environmental variables.”* (Hartemink et al. 2008).

Minasny e McBratney hanno formalizzato il *frame* entro cui opera il DSM, il quale necessita di tre componenti:

*“- The input in the form of field and laboratory observational methods, this includes the use of legacy soil observations or soil maps, and collecting new samples using statistical sampling techniques.*

*- The process used in terms of spatial and non-spatial soil inference systems, this includes building mathematical or statistical models relating soil observations with their environmental covariates or SCORPAN factors.*

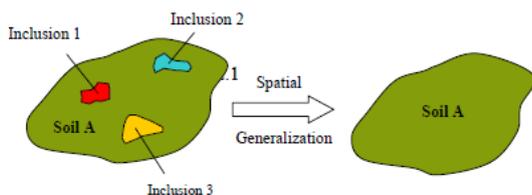
*- The output in the form of spatial soil information systems, which includes outputs in the form of rasters of prediction along with the uncertainty of prediction. This output can be readily updated as new information becomes available.”* (Minasny e McBratney 2016).

Indipendentemente da ciò che viene rappresentato, per la realizzazione di una mappa relativa alle caratteristiche dei suoli è necessaria la disponibilità di dati riguardanti le proprietà chimico-fisiche relative ai suoli dei punti campionati e di informazioni riguardanti le variabili ambientali, che hanno una significativa influenza sulla formazione dei suoli e quindi sulla loro distribuzione. Pertanto, l'attività di mappatura non può prescindere dall'acquisizione dei dati in campo e dipende dalla disponibilità di informazioni riguardo le variabili ambientali connesse alla pedogenesi, chiamate in letteratura *covariate*.

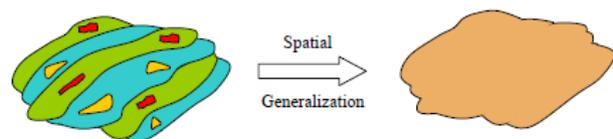
In sintesi il processo di creazione di una mappa, sia adottando il tradizionale processo di mappatura sia utilizzando le tecniche di DSM, comprende 3 step fondamentali:

1. l'attività di campionamento;
2. l'analisi dei campioni di suolo, finalizzata alla loro caratterizzazione;
3. la spazializzazione dei dati puntiformi, rilevati in campo.

La differenza più importante tra il metodo tradizionale di mappatura e l'approccio del DSM sta nel modo di derivare le informazioni dai dati di input, ovvero il metodo di costruzione della mappa partendo dai punti di campionamento. Il metodo tradizionale è principalmente basato su tecniche empiriche e la qualità del prodotto finale è un derivato del costruito mentale dell'operatore. Inoltre, nella attività di rilevamento in campo, che precede la redazione della mappa, la scelta dei punti di campionamento è fortemente influenzata dal modello concettuale di distribuzione dei suoli, basato sull'esperienza dell'operatore e la carta viene realizzata digitalizzando i poligoni, che possono generare una semplificazione della realtà (Zhu et al. 1997). Un tipo di generalizzazione spaziale si verifica quando le caratteristiche di dettaglio sono semplificate nella rappresentazione cartografica a causa della limitazione della scala, la quale determina la dimensione minima del corpo di suolo cartografabile (v. Fig. 1.1). Un altro tipo di generalizzazione si verifica quando due o più componenti dissimili in una ripetizione regolare non possono essere delineati separatamente alla scala di mappatura (v. Fig. 1.2). In questo caso, tutti i componenti sono raggruppati in una singola unità cartografica. Essa quindi può contenere un tipo di suolo prevalente (consociazione), o suoli codominanti dei quali non è noto il modello distributivo, riuniti in associazioni o complessi di suolo. A seconda della scala di rappresentazione della carta finale, altri suoli simili al tipo di suolo dominante e piccole aree di suoli contrastanti (inclusioni), non vengono rappresentati ed eventualmente sono indicati nelle legende o nella descrizione.



**Fig. 1.1** Generalizzazione spaziale per suoli con limitata estensione (Zhu et al. 2007).



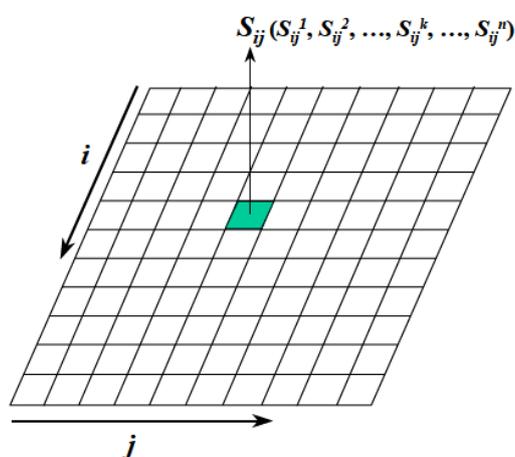
**Fig. 1.2** Generalizzazione spaziale per suoli dissimili che ricorrono in modo intermittente (Zhu et al. 2007).

Utilizzando le tecniche del DSM le mappe sono il risultato dell'applicazione di tecniche statistiche ai dati di input, costituiti dai punti di campionamento georeferenziati e da strati informativi digitali, che permettono di limitare o addirittura eliminare la soggettività dell'operatore (Dobos et al. 2006). Utilizzando i metodi matematici e statistici per derivare la mappa è possibile conoscere l'errore delle stime e quindi la qualità del prodotto finale. Un'altra differenza è legata alla disponibilità di tecnologie (e.g. indagini geofisiche elettromagnetiche o spettroscopia *near infrared* (NIR)) che permettono il rilevamento di caratteristiche fisiche direttamente utilizzabili nel processo di mappatura. Inoltre, l'attività di rilevamento in campo può essere progettata secondo tecniche statistiche, che permettono un sistema di campionamento efficiente, minimizzando i tempi e i costi monetari e nel contempo di ottenere un set campionario in grado di rappresentare adeguatamente la variabilità studiata.

Sostituendo la rappresentazione vettoriale con una rappresentazione in formato raster è possibile migliorare la raffigurazione della distribuzione spaziale dei suoli e delle loro caratteristiche; infatti, seppur con le limitazioni che l'informazione digitale discreta permette, è possibile ottenere quella variabilità a toni sfumati che caratterizza il *continuum* tipico della variabilità delle componenti naturali. Quando è rappresentato un parametro, il gradiente cromatico fornisce l'informazione riguardo la sua misura, mentre nel caso di variabili categoriali è possibile ottenere un gradiente a

toni sfumati riguardo al grado di assegnazione di una certa classe ad un certo pixel, come avviene applicando la logica *fuzzy* o gli algoritmi probabilistici.

L'approccio di DSM utilizzato nel presente lavoro è basato sul concetto di discretizzazione del territorio in una griglia, o matrice multidimensionale (v. Fig. 1.3). Suddividendo una determinata area geografica secondo una griglia è possibile strutturare, in ambiente GIS, un geodatabase. In tal caso, il geodatabase è, di fatto, lo spazio multidimensionale con cui si opera per la costruzione della mappa, al cui interno a ciascun punto del territorio corrisponde un determinato pixel e a ciascun pixel è possibile associare un vettore, il cui array è costituito dalle variabili considerate utili per determinare la distribuzione spaziale dei suoli (Zhu et al. 2007). D'altro canto il medesimo geodatabase può essere pensato come un tensore costituito dall'array di matrici, ciascuna delle quali è relativa ad una variabile.



**Fig. 1.3** Rappresentazione dello spazio multidimensionale utilizzato per la costruzione delle mappe: a ciascuna porzione del territorio corrisponde un pixel di coordinate  $ij$  a cui è associato il vettore  $S_{ij}$  (Zhu et al. 2007).

## 1.2. Concetti di base della pedologia

Per la presente trattazione, si assume che:

1. la pedosfera di una determinata area geografica è quel sistema naturale complesso, frutto dell'interazione tra i domini della litosfera, idrosfera, atmosfera e biosfera (Schaeztl e Anderson 2005);
2. le proprietà del suolo sono correlate alla variabilità spaziale e temporale di questi tre domini;
3. erosione, ricoprimento, cambi climatici, movimenti biomeccanici, variazione del livello delle acque sotterranee, effetti microclimatici e topografici e innumerevoli altri fattori interagiscono nella formazione del suolo (pedogenesi).

Nel tempo, i pedologi hanno creato diversi modelli concettuali per comprendere la formazione dei suoli, le loro caratteristiche e la loro distribuzione sul territorio. Generalmente i modelli concettuali sono strumenti che possono essere utilizzati come ausilio per la descrizione semplificata dei sistemi naturali. Infatti, se efficaci, forniscono strutture attraverso le quali è più semplice osservare i

fenomeni. Il modello costruito può essere complesso ma a volte traduce funzionamenti complessi in equazioni semplici.

Come tipicamente avviene nello studio dei comparti ambientali, in un approccio teorico occorrerebbe definire il sistema esaminato, individuando le componenti che agiscono e/o interagiscono. Il funzionamento del sistema viene determinato dal tipo di interazioni esogene ed endogene delle stesse componenti. Si può considerare, inoltre, che i sistemi ambientali (normali) sono sistemi termodinamici (Stengers 1999); pertanto, in linea teorica, per lo studio del comparto suolo si potrebbe procedere alla quantificazione dello scambio di energia e di materia, che avviene attraverso il contorno (scambio del sistema con l'ambiente) per determinare, infine, la dinamica dei processi che contribuiscono alla pedogenesi e quindi al suo risultato. A riguardo sono stati concepiti diversi modelli concettuali riguardanti la pedogenesi, i quali hanno in comune, più o meno sottesi, i concetti di scambio di materia e di energia.

Tra i modelli concettuali che hanno cercato di spiegare la pedogenesi, occorre, in primis, citare il modello basato su fattori di stato (*state factor model*), attribuito ad Hans Jenny, ma pensato prima da Dokuchaev e da Sergey Zakharov (Florinsky 2012).

In questo modello il suolo è immaginato come il frutto di diversi fattori di formazione che interagiscono e, generalmente, è scritto con una forma algebrica del tipo:

$$S = f(X_1, X_2, X_3, \dots)$$

Ciò che è argomento della funzione, secondo la congettura di Jenny, non deve essere considerato una forza propulsiva o una causa da cui scaturisce la pedogenesi ma, piuttosto, una serie di variabili indipendenti o fattori che concorrono a definire il sistema suolo e che impattano sulla dinamica e sulla sua formazione (Jenny 1994).

Detto modello è generalmente scritto nella seguente forma:

$$S = f(cl, o, r, p, t, \dots)$$

Dove i fattori attivi sono *cl*, *climate*, e *o*, *organic activity* (componente biotica nel senso più ampio del termine); i fattori passivi sono: *r*, *relief*, *p*, *parent material*, *t*, *time* e i puntini stanno ad indicare la necessità di dover includere nel modello eventuali particolari fattori caso-specifici.

L'equazione non può essere risolta se non parzialmente, poiché il sistema suolo è così complesso che risulta impossibile definire completamente i fattori in gioco, dal momento che quest'ultimi sono anche interdipendenti. E' eventualmente più semplice tradurre l'equazione di Jenny in una serie di funzioni di pedogenesi, mantenendo di volta in volta costanti tutti i fattori tranne uno.

Roy Simonson (Simonson 1978) ha sviluppato e presentato un modello che, diversamente da quello di Jenny, è interamente basato sui processi. Egli ha considerato che ciascuno stadio di evoluzione del suolo può apparire, scomparire, accentuarsi o affievolirsi: tale visione, che enfatizza le variazioni degli stadi di sviluppo dei suoli, ha portato a concepire il modello sistema-processo. Simonson, basandosi sulle differenze e sulle analogie tra i suoli, ha osservato che certe discrepanze esibite dai suoli sono dovute ad una differente forza con cui ha operato lo stesso tipo di processo su

materiale simile. Nel modello sistema-processo la genesi dei suoli è costituita da due step: 1) l'accumulo di materiale parentale; 2) la differenziazione del materiale parentale in orizzonti (Schaetzl e Anderson 2005).

Yaalon (Yaalon 1983) ha evidenziato come l'operatività di un modello processo-risposta come quello di Simonson possa essere caratterizzata in termini di tre determinanti:

1. lo stato iniziale del sistema;
2. il processo cui il sistema è assoggettato;
3. lo stadio o la durata del processo.

Il modello di Simonson è basato sul secondo punto: ciò che viene considerato è il grado a cui è il processo, piuttosto che considerare il tipo di processo, e può essere scritto nel seguente modo:

$$S=f(a, r, t_1, t_2)$$

Dove con  $a$  si intende addizione, con  $r$  rimozione, con  $t_1$  trasferimento/traslocazione e con  $t_2$  trasformazione. Il bilancio e il carattere di questi processi, che avvengono contemporaneamente, determinano l'espressione ultima del processo. Il modello di Simonson così formulato può essere considerato parte di un'ampia classe di modelli di bilancio di massa.

Runge (Runge 1973) ha sviluppato un modello fattoriale di sviluppo del suolo che è un ibrido tra il modello stato-fattore di Jenny e il modello di Simonson.

Runge ha enfatizzato l'importanza di due fattori del modello di Jenny, clima e rilievo, e li ha combinati in un unico fattore che ha definito come l'ammontare dell'acqua disponibile per la lisciviazione ( $w$ ); quindi i due fattori producono il vettore processo che esprime potenzialmente quanto l'acqua percola in un profilo. Questo nuovo vettore ( $w$ ) è chiamato vettore organizzante, in quanto utilizza l'energia gravitazionale per organizzare il profilo diminuendo l'entropia del profilo stesso e creando orizzonti. Runge poi ha pensato di combinare la variabile materiale parentale e quella relativa al contributo degli organismi viventi in un altro unico fattore, chiamato fattore di produzione di materia ( $o$ ) (o specularmente: mancanza di mineralizzazione); il fattore ( $o$ ) è visto come un vettore di ringiovanimento. Detto modello può essere scritto nella forma:

$$S=f(w, o, t) \quad \text{dove } t \text{ è il tempo}$$

Un'ulteriore prospettiva, derivata dalla comune accettazione del concetto di pedogenesi *top down* (secondo cui la differenziazione dei profili in orizzonti e la profondità di alterazione cresce con il tempo), è quella concepita da Donald Johnson e Watson-Stegner (Johnson e Watson-Stegner 1987). Nel loro modello lo spessore ( $T$ ) di suolo è visto come una relazione dinamica che coinvolge i processi di variazione del materiale nel profilo, in superficie e in profondità e la loro rimozione

$$T= D+U+R \quad ((D) \text{ deepening}, (U) \text{ upbuilding}, (R) \text{ removal})$$

Il modello di Johnson e Watson-Stegner è basato su due assunti:

1. lo sviluppo del suolo è continuamente influenzato da processi progressivi;
2. i processi progressivi non operano in modo costante.

Più che a un processo unidirezionale, i due geografi dell'Università dell'Illinois hanno pensato che la pedogenesi è un processo con fasi sia di evoluzione che di involuzione. Ciascun processo che contribuisce al cambiamento del suolo inizia, raggiunge l'apice e poi declina.

$S=f(P,R)$  (dove P sta per *progressive pedogenesis* ed R sta per *regressive pedogenesis*)

Il modello pone l'accento sul fatto che l'evoluzione del suolo procede lungo due vie genetiche che interagiscono e che riflettono le variabili dei processi esogeni ed endogeni, i fattori e le condizioni. Ciascun suolo ha una via lungo la quale si "muove" in avanti e una via lungo la quale regredisce. Il tipo di evoluzione dipende dalla preponderanza di una delle due spinte. Inoltre, ciascuna via è stata pensata come costituita da tre componenti, ognuna delle quali consiste di due opposti vettori:

1. vettori che comportano stratificazione/semplificazione;
2. vettori che comportano ritardo o aumento di velocità nella evoluzione;
3. vettori che comportano aumento o diminuzione degli spessori.

Jonathan Phillips (Phillips 1993) ha supportato e ulteriormente sviluppato l'idea di evoluzione concepita da Johnson e Watson-Stegner, focalizzandosi, però, su una spiegazione teorico-matematica di un modello basato sulla dinamica non lineare e sulla teoria del caos.

Indipendentemente dall'approccio teorico cui fare riferimento, si può considerare il suolo come un sistema naturale termodinamico aperto, sul quale influiscono l'apporto di energia e di materia, secondo dinamiche relazionate al contesto. Nel caso in esame il sistema suolo è posto in un ambiente appenninico, dove le dinamiche evolutive (pedogenesi) sono fortemente influenzate dai processi di versante e dal substrato litologico.

### 1.3. L'evoluzione verso il DSM

Dal punto di vista teorico, il processo di pedogenesi è esprimibile con una equazione che, secondo l'approccio classico fattore-suolo, può essere espressa come segue:

$$S(t) = \int_{t_0}^t f(E) d\tau$$

L'equazione è una modifica di quella formulata e pubblicata da Xing Zhu nel suo approccio al DSM (Zhu et al. 2007), (l'intervallo di integrazione è determinato dal periodo di biostasia che dura dal termine del periodo di rexistasia ( $t_0$ ) sino al momento in cui si procede alla mappatura ( $t$ )). Essa assume un aspetto differente rispetto ai modelli concettuali precedentemente citati, in quanto la funzione  $f$  è il metodo statico o matematico utilizzato per derivare la carta del suolo e l'argomento della funzione ( $E$ ) è rappresentato dalle variabili numeriche o categoriali, costituite dagli strati informativi digitali relativi all'area considerata, utili per comprendere i processi pedogenetici.

In considerazione della velocità di cambiamento del suolo, delle sue evoluzioni progressive e/o regressive, è possibile procedere ad una approssimazione dell'equazione fondamentale sopra riportata, che viene accettata anche in considerazione della difficoltà di procedere esplicitamente all'integrazione dei fattori ambientali, formativi del suolo, nel tempo. Pertanto,  $t$  viene considerata

parte di ( $E$ ) e l'espressione del processo pedogenetico che viene considerata è quella semplificata in:  $S = f(E)$  (Zhu 1999). La funzione  $f$  è relativa ad una situazione statica e fornisce come risultato una caratteristica relativa ai suoli. In ambito cartografico questa approssimazione è adeguata, in quanto la mappa rappresenta comunque una situazione in un determinato istante.

Una ulteriore evoluzione di questo modello concettuale sono le equazioni generali formalizzate da McBratney (McBratney et al. 2003), il quale ha integrato il precedente criterio di Jenny aggiungendo il fattore *posizione geografica* e distingue:

(1) modelli di inferenza spaziale che producono *soil class map*:  $Scl = f(S, C, O, R, P, A, N)$

(2) modelli di inferenza spaziale che producono *soil attribute map*:  $Sa = f(S, C, O, R, P, A, N)$

( $S$ : soil, other properties of the soil at a point;  $C$ : climate, climatic properties of the environment at a point;  $O$ : organisms, vegetation or fauna or human activity;  $R$ : topography, landscape attributes;  $P$ : parent material, lithology;  $A$ : age, the time factor;  $N$ : space, spatial position)

Inoltre, ha proposto una formulazione generale dei modelli di inferenza spaziale nella forma di:

$S = f(Q) + e$  (dove  $e$  è l'errore derivante dall'applicazione del metodo di inferenza ai predittori  $Q$ )

L'approccio prevede, indipendentemente da ciò che si vuole rappresentare in mappa (variabili categoriali o numeriche), di considerare un numero di osservazioni ( $n$ ) e di mettere in relazione i valori della variabile indipendente, rilevati in campo, con i predittori  $Q$ , secondo una funzione costruita con un processo di inferenza. Una volta che la funzione è stata modellata in corrispondenza dei punti di osservazione, la previsione può essere estesa ai punti non noti ( $N$ ). L'efficacia del metodo si basa sul fatto che  $n \ll N$  (McBratney et al. 2003). I predittori possono essere variabili categoriali o numeriche, relative a tematiche ambientali o derivanti da misurazioni effettuate con strumentazioni (es: *remote sensing*). Il successo dipende da:

1. avere un numero di predittori sufficienti per l'intera area di studio;
2. avere abbastanza osservazioni sul suolo (*data points*), per poter ottenere un buon *fitting* nel processo di inferenza;
3. avere funzioni  $f()$  abbastanza flessibili per creare eventuali relazione non lineari;
4. avere una buona relazione tra il suolo e il suo ambiente (McBratney et al. 2003).

Come sottolineato da Grunwald (Grunwald 2010) anche se i modelli concettuali citati sono da tempo accettati, la relazione tra variabili ambientali e proprietà del suolo differiscono da regione a regione e i metodi per derivare mappe tematiche, partendo dai punti di osservazione, sono ancora oggetto di investigazione. Infatti, vi sono numerosi esempi presenti nella letteratura del DSM con differenze sia nell'uso dei predittori sia nei metodi di inferenza, utilizzati per la spazializzazione del dato puntuale (McBratney et al. 2003; Minasny e McBratney 2016).

Per quanto concerne i predittori che possono essere utilizzati come *inputs* per il DSM, occorre anche considerare che la disponibilità di dati spaziali e la loro qualità variano a seconda della regione geografica. Riguardo l'impiego di detti *layer* digitali relativi a variabili ambientali occorre tener presente che possono verificarsi alcuni aspetti controproducenti. A volte, tematismi

considerati fondamentali per spiegare la distribuzione spaziale delle caratteristiche dei suoli sono debolmente correlate o addirittura incorrelate con l'obiettivo. Ad esempio una mappa d'uso del suolo di recente redazione o una immagine satellitare attuale possono nascondere un secolare utilizzo del territorio. Ancora più rischioso è l'uso delle mappe litologiche come variabile per il *parent material*. Lungo un pendio, in un determinato punto, il *parent material*, da cui ha avuto origine il suolo, può essere totalmente diverso dal substrato litologico, presente in quel determinato punto e rilevato dal geologo. Inoltre, le mappe relative ai temi ambientali possono avere una definizione non idonea per il grado di dettaglio scelto per la redazione della mappa dei suoli. Questo problema tipicamente si verifica con l'impiego di immagini satellitari, la cui risoluzione spaziale è inversamente proporzionale alla risoluzione spettrometrica. Le informazioni di *remote sensing* su vaste aree spesso non hanno la risoluzione spaziale richiesta per la moderna cartografia dei suoli, se non con elevati costi monetari. Le operazioni di *data fusion*, utili per rendere compatibili i *layers* con differente risoluzione spaziale, possono deteriorare l'informazione che veicolano. Inoltre, a parte i casi di scarsa qualità, gli strati informativi possono essere di scala non adeguata alla precisione richiesta per il prodotto finale. Nel DSM sono spesso impiegate variabili categoriali che derivano da mappe di tipo vettoriale. In tal caso può essere presente imprecisione per le aree prossime ai bordi dei poligoni, specialmente quando si utilizzano mappe redatte con scale non di dettaglio (minore della scala 1:50.000). Per questi casi occorre tener presente che assegnare un determinato valore categoriale ad un punto è una forte assunzione. Utilizzare detti tematismi, nel corso delle elaborazioni statistiche per la costruzione di mappe, può generare errori che possono inficiare la qualità del prodotto finale, specialmente in territori morfologicamente complessi.

#### **1.4. Approccio utilizzato**

Il metodo adottato in questo lavoro considera che in molti casi le variabili geomorfometriche (*terrain attributes*) possono fornire le migliori indicazioni per la variabilità spaziale delle proprietà dei suoli, specialmente in aree dove la variabilità dei fattori ambientali è piccola; pertanto, è di fondamentale interesse l'analisi della topografia (Florinsky 2012; Maynard e Johnson 2014).

In pratica l'approccio utilizzato ha come base teorica il paradigma suolo-paesaggio di Hudson, e la caratterizzazione della topografia sulla base della geomorfometria.

Per il suo sviluppo sono state considerate le seguenti relazioni: suolo-paesaggio-topografia-geomorfometria-suolo.

Per spiegare la distribuzione spaziale del suolo è ragionevole considerare il paradigma suolo-paesaggio (Hudson 1992): Hudson sostiene, in sintesi, che l'interazione tra i fattori (clima, organismi, materiale parentale e topografia in un certo periodo di tempo) porta alla formazione di un unico suolo o gruppo di suoli in una determinata area della superficie terrestre. Si può supporre, quindi, che due siti con condizioni edafiche paragonabili abbiano suoli simili; inoltre, tanto più i siti presentano affinità tra loro, tanto maggiore sarà la probabilità che essi abbiano suoli simili. In altre parole, la somiglianza tra due suoli in due punti diversi può essere approssimata per astrazione al concetto di similitudine tra condizioni edafiche stazionali.

Il concetto di paesaggio, che di per sé è uno specifico livello dello spettro biologico (Odum et al. 1971), non è scevro di ambiguità nella sua definizione (Ingegnoli 2013). Nel caso del paradigma di Hudson, il paesaggio è contestualizzabile considerando la nota equazione di Jenny, relativa alla

concettualizzazione del processo pedogenetico. Nell'ambito della pedologia, infatti, il paesaggio è l'espressione dell'interazione di quei fattori naturali che concorrono allo sviluppo pedogenetico di un suolo. Hudson considera che la distribuzione spaziale delle diverse tipologie di suolo presenta una variabilità continua con la presenza a volte abrupta di discontinuità e che tale distribuzione può essere associata a strutture territoriali individuabili, chiamate *soil-landscape unit*. Hudson ha voluto evidenziare che l'identificazione del paesaggio coincide con l'individuazione di *patches* che sono frutto dell'interazione dei fattori naturali che concorrono alla pedogenesi, tenendo presente che tali fattori hanno un peso diverso a seconda dell'area geografica considerata. D'altro canto secondo Florinsky *“Topography is one of the main factors controlling processes taking place in the near-surface layer of the planet. Topography is result of the interaction of endogenous and exogenous processes of different scales, topography can reflect the geological structure of a terrain“* ed il Digital Elevation Model (DEM) è la principale informazione per lo studio della topografia. Inoltre, *“In soil science, conventional geomorphometric techniques have been used to investigate relationships between soil cover and topography; to predict quantitative soil properties; to produce soil maps; and to study regularities in the structure of the soil cover and its relations with geological features. It is well known that topography is one of the soil-forming factors. Topography influences (micro)climatic and meteorological characteristics, which affect the hydrological and temperature regimes of soils, the prerequisites of the gravity-driven lateral overland and intrasoil transport of water and other substances, as well as the spatial distribution of the vegetation cover. Thus, it is natural that topography directly or indirectly controls the spatial distribution of physical, chemical, and biological soil properties. An in-depth understanding of this control is required for further modeling and mapping of soil properties based on topographic data. Topography influences soil properties through two main “tools”: 1. The gravity-driven lateral migration and accumulation of water; and 2. Spatial differentiation of the temperature regime of slopes.”* (Florinsky 2016).

Si suppone, pertanto, che il paesaggio (Hudson *sensu*) possa essere descritto con una serie di variabili geomorfometriche costruite a partire dal DEM, le quali complessivamente danno origine a un iperspazio che è una approssimazione del corrispondente multidimensionale paesaggio reale. Si ipotizza, infatti, che dette variabili contengano informazioni succedanee ai fattori che concorrono a determinare l'espressione del processo pedogenetico, e che la loro interazione fornisce un surrogato dei reali processi. In altre parole, alcune variabili utilizzate sono associabili ai fenomeni fisici legati al flusso di massa e di energia che spiegano la distribuzione spaziale dei tipi di suolo, altre variabili simulerebbero i fattori che possono rappresentare i termini presenti nell'equazione di Jenny. Complessivamente, l'iperspazio determinato dalle variabili utilizzate concorre a delineare il paesaggio e, di conseguenza, a spiegare la relazione tra suolo e variabilità del territorio.

In sostanza, l'approccio per lo studio dei suoli presentato in questo lavoro è basato sull'impiego delle variabili geomorfometriche primarie e secondarie, calcolate a partire dal DEM. In questo modo, se le dimensioni dei pixel sono ragionevoli rispetto alla reale variabilità ambientale e se il processo di inferenza è idoneo, le variazioni della variabile mappata possono simularne la distribuzione nel dominio geografico. La rappresentazione che ne deriva è una matrice di dati costituita da righe i cui *records* sono associati ai pixel, che costituiscono il raster relativo all'area geografica, e da colonne che sono le variabili utilizzate come argomento delle funzioni, nel processo di creazione del modello matematico statistico, usato per la redazione della mappa. Per

quanto riguarda l'impiego di detta informazione, occorre evidenziare che nell'approccio impiegato in questo lavoro per non viene data alcuna importanza al significato fisico; bensì con l'obiettivo di produrre mappe tematiche che siano la migliore rappresentazione della realtà; la valutazione circa la rilevanza di ciascuna variabile è demandata alle elaborazioni statistiche.

Il grado di successo nell'uso di questo approccio dipende dalla dimensione del pixel, da quanto il *dataset* è popolato, cioè da quanto la matrice contiene l'informazione necessaria per le elaborazioni finalizzate al processo di inferenza, e dalle applicazioni statistiche utilizzate per la costruzione cartografica. Di fatto, le coordinate nello spazio multidimensionale di ciascun punto vettore-suolo, che per costruzione corrisponde al pedon, dovrebbero contenere l'informazione necessaria per mappare tematiche concernenti le caratteristiche del suolo, oltre alle informazioni rilevate in campo.

## **Bibliografia**

Dobos, E.; Carré, F.; Hengl, T.; Reuter, H. I.; Tóth, G. (2006): Digital soil mapping as a support to production of functional maps. In: *Office for Official Publications of the European Communities, Luxemburg. EUR, 22123.*

Florinsky, Igor (2016): Digital terrain analysis in soil science and geology: Academic Press.

Florinsky, Igor V. (2012): Digital Terrain Analysis in Soil Science and Geology: Elsevier.

Florinsky, IV (2012): The Dokuchaev hypothesis as a basis for predictive digital soil mapping (on the 125th anniversary of its publication). In: *Eurasian soil science* n. 4, 45, p. 445.

Grunwald, S. (2010): Current state of digital soil mapping and what is next. In: *Digital Soil Mapping*: Springer, pp. 3–12.

Hartemink, Alfred E.; McBratney, Alex B.; Lourdes Mendonça-Santos, Maria de; Ahrens, Robert J. (2008): Digital soil mapping with limited data: Springer Berlin, Germany.

Hudson, Berman D. (1992): The soil survey as paradigm-based science. In: *Soil Science Society of America Journal* n. 3, 56, pp. 836–841.

Ingegnoli, Vittorio (2013): Landscape ecology: a widening foundation: Springer Science & Business Media.

Jenny, Hans (1994): Factors of soil formation: a system of quantitative pedology: Courier Corporation.

Johnson, Donald L.; Watson-Stegner, Donna (1987): Evolution model of pedogenesis. In: *Soil Science* n. 5, 143, pp. 349–366.

Maynard, J. J.; Johnson, M. G. (2014): Scale-dependency of LiDAR derived terrain attributes in quantitative soil-landscape modeling. Effects of grid resolution vs. neighborhood extent. In: *Geoderma*, 230-231, pp. 29–40. DOI: 10.1016/j.geoderma.2014.03.021.

McBratney, A.B; Mendonça Santos, M.L; Minasny, B. (2003): On digital soil mapping. In: *Geoderma* n. 1-2, 117, pp. 3–52. DOI: 10.1016/S0016-7061(03)00223-4.

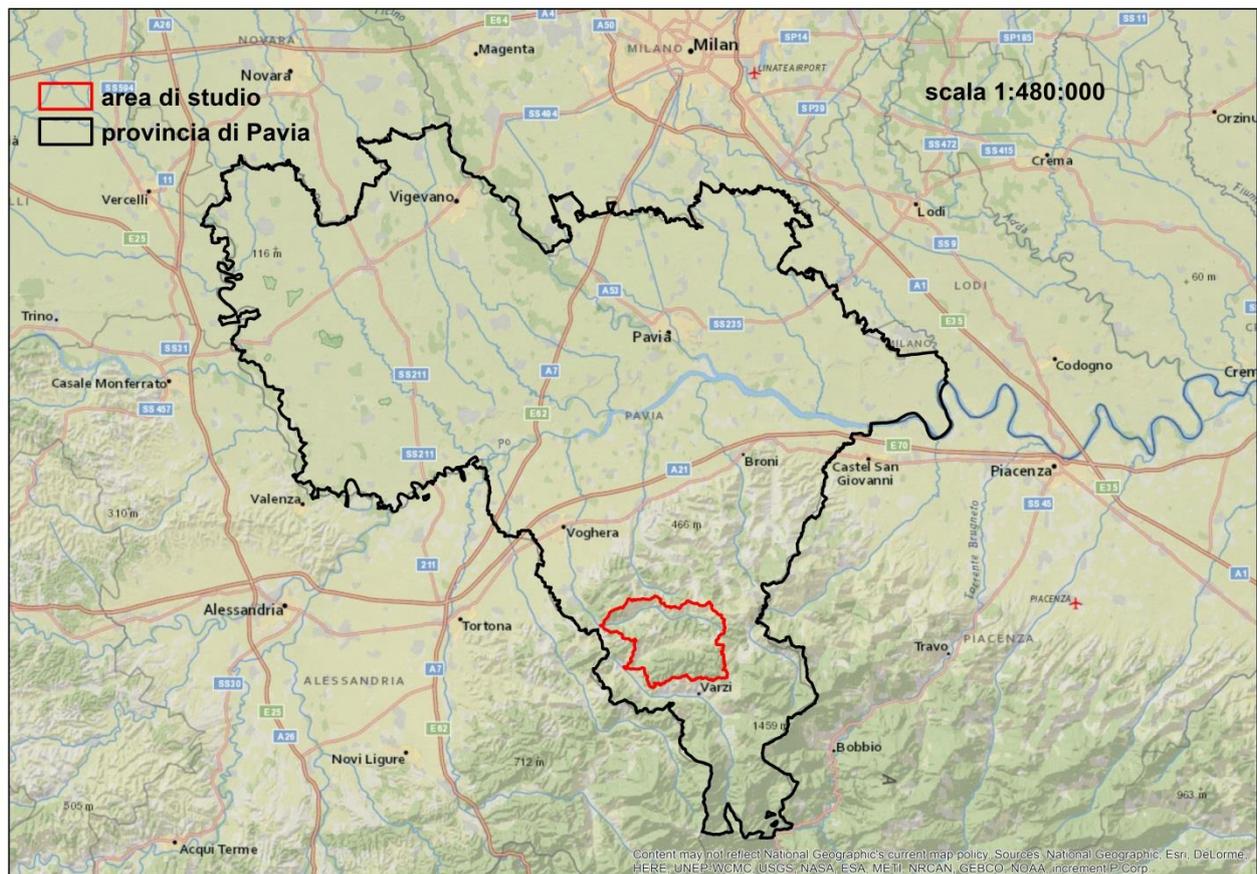
Minasny, Budiman; McBratney, Alex B. (2016): Digital soil mapping: A brief history and some lessons. In: *Geoderma*, 264, pp. 301–311.

- Odum, Eugene Pleasants; Odum, Howard T.; Andrews, Joan (1971): Fundamentals of ecology: JSTOR.
- Phillips, Jonathan D. (1993): Progressive and regressive pedogenesis and complex soil evolution. In: *Quaternary Research* n. 2, 40, pp. 169–176.
- Runge, E. C.A. (1973): Soil development sequences and energy models. In: *Soil Science* n. 3, 115, pp. 183–193.
- Schaetzl, R. J.; Anderson, S. (2005): *Soil Genesis and Morphology*: Cambridge University Press, New York, NY.
- Simonson, Roy W. (1978): A multiple-process model of soil genesis. In: *Quaternary soils*, pp. 1–25.
- Stengers, Isabelle (1999): *La nuova alleanza: metamorfosi della scienza*: Einaudi.
- Yaalon, D. H. (1983): Climate, time and soil development. In: *Pedogenesis and soil taxonomy*/edited by LP Wilding, NE Smeck, and GF Hall.
- Zhu, A. (1999): Fuzzy inference of soil patterns: implications for watershed modeling. In: *Assessment of Non-Point Source Pollution in the Vadose Zone*, pp. 135–149.
- Zhu, A. X.; Band, Lawrence; Vertessy, Robert; Dutton, Barry (1997): Derivation of soil properties using a soil land inference model (SoLIM). In: *Soil Science Society of America Journal* n. 2, 61, pp. 523–533.
- Zhu, A-Xing; Burt, James E.; Moore, Amanda C.; Smith, Michael P.; Liu, Jian; Qi, Feng (2007): SoLIM: A New Technology For Soil Mapping Using GIS, Expert Knowledge & Fuzzy Logic. In: *Overview of SoLIM Programme. University of Wisconsin*.

## CAPITOLO 2

### INQUADRAMENTO AREA DI STUDIO

L'area di studio<sup>1</sup> (fig. 2.1) costituisce parte dell'Appennino Pavese ed è ubicata nella fascia latitudinale compresa tra 44° 55' 20" e 44° 49' 44" N e longitudinale compresa tra 9° 03' 37" e 9° 14' 15" E, l'altitudine varia da 193 metri a 879 metri s.l.m... E' composta dagli areali relativi alla "placca" di Pizzocorno-Pietragavina (17,086966 Km<sup>2</sup>) e ai bacini idrografici della Val di Nizza (27,306982 Km<sup>2</sup>) e Val Ardivestra (47,319974 Km<sup>2</sup>). Diversamente dagli altri bacini idrografici presenti nell'Oltrepo Pavese collinare, caratterizzati da solchi vallivi con direzione da sud-nord, i due bacini studiati e la "placca", in argomento, hanno assunto a seguito dell'evoluzione tettonica, una orientazione est-ovest.

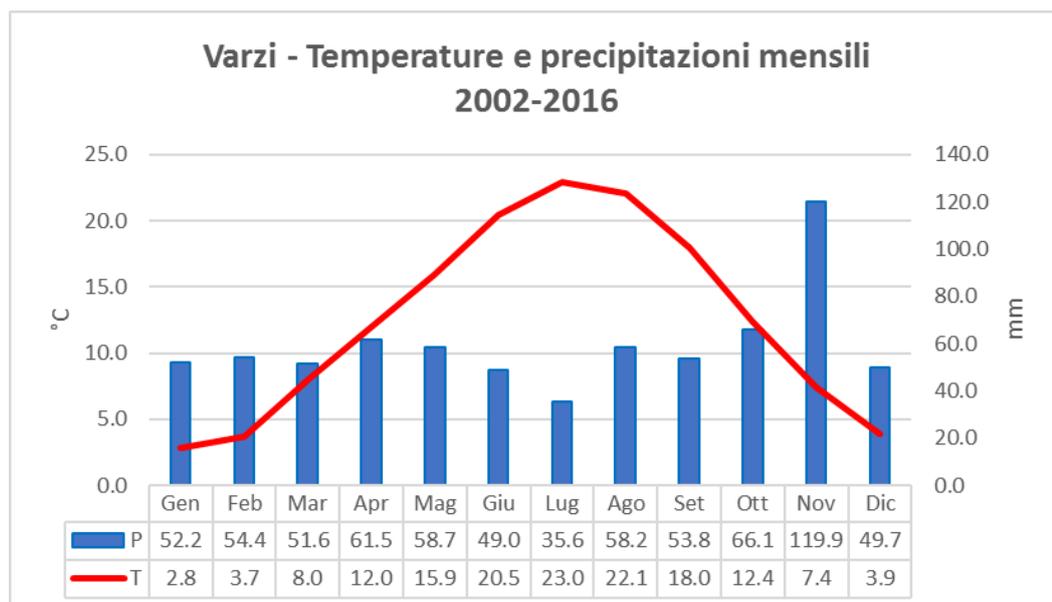


**Fig. 2.1** Ubicazione dell'area di studio (in rosso) all'interno della provincia di Pavia.

<sup>1</sup> L'area di studio è stata delimitata utilizzando lo *shape file* della Regione Lombardia (Carrara, 2004), per quanto riguarda l'individuazione dei bacini idrografici relativi alle due valli, il quale è stato unito alle linee che delimitano le formazioni delle Arenarie del Monte Vallassa e di Contignaco, derivate dallo *shape file* del Foglio 178-Voghera, della Carta Geologica d'Italia (1:50.000).

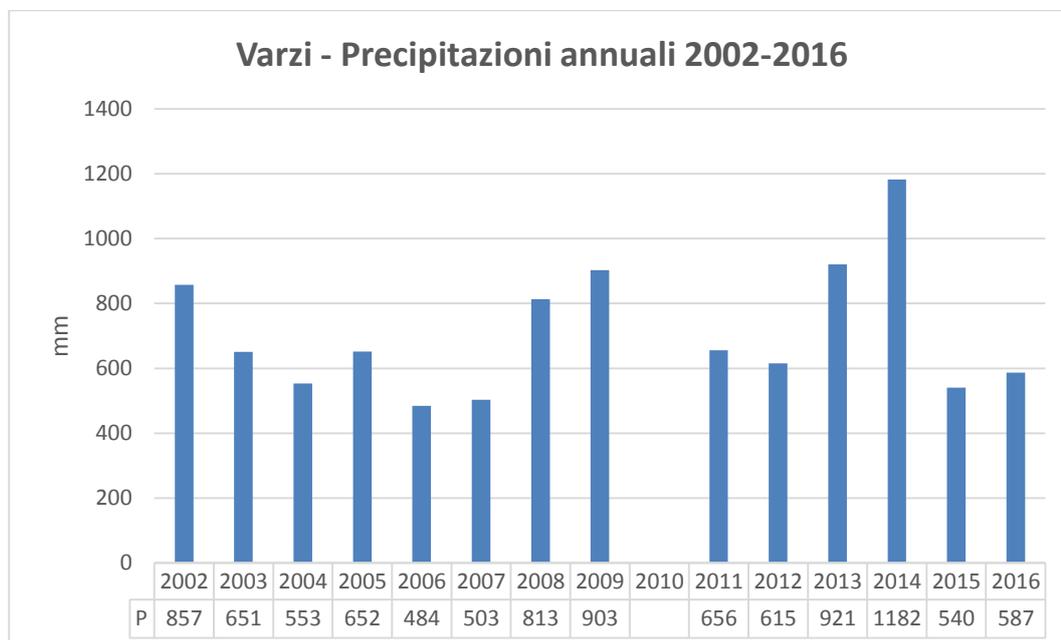
## 2.1 Inquadramento climatico

La collocazione nella fascia latitudinale media e la relativa prossimità sia al Mediterraneo che alle Alpi determinano un regime climatico caratterizzato da temperature medie annue comprese tra 11.0 e 13.5 °C, con gennaio mese più freddo e luglio mese più caldo. A ciò si aggiunga che in inverno la fascia collinare dell'Oltrepò manifesta il caratteristico fenomeno del *thermal belt*, in virtù del quale le temperature invernali sono significativamente più miti rispetto a quelle della pianura sottostante. La piovosità annua media è di 700 mm, ripartiti su un totale di 80–90 giorni di pioggia; il mese di gran lunga più piovoso è novembre, quello meno piovoso luglio. La piovosità estiva (segnale di mediterraneità) più ridotta di quella invernale (segnale di oceanicità) è sintomo del fatto che l'influsso mediterraneo, legato alla vicinanza al Golfo di Genova, domina su quello atlantico, che prevale invece a nord del Po. Per gli andamenti termopluviometrici, importanti per i processi pedogenetici, si è fatto riferimento alla serie di misure della stazione ARPA Lombardia di Varzi Nivione (m 500 s.l.m.), per la quale si dispone dei dati giornalieri dal 2002 al 2016 e che può essere presa a rappresentante del clima dell'area di studio, essendo collocata solo pochi km all'esterno del suo perimetro. In fig. 2.2 sono presentate le medie mensili di temperatura e precipitazioni.

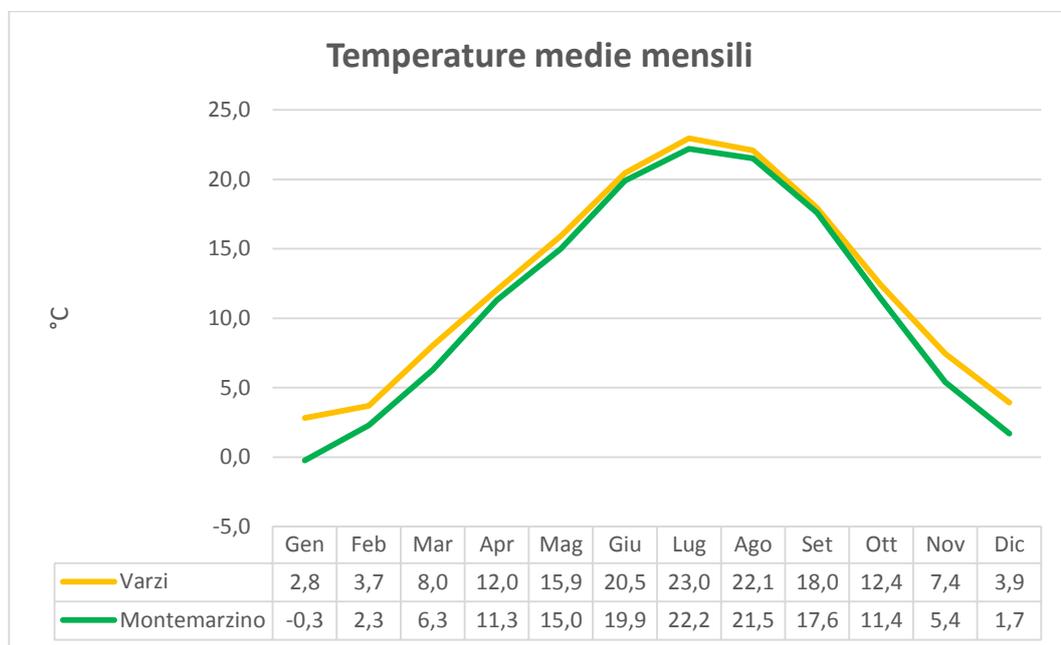


**Fig. 2.2** Medie mensili di temperatura e precipitazioni per la stazione di Varzi Nivione (elaborazione su dati giornalieri ARPA Lombardia).

In fig. 2.3 è rappresentato l'andamento delle precipitazioni annuali a Varzi Nivione nel periodo di misura 2002-2016: è evidente la forte variabilità interannuale, con minimi inferiori a 500 mm e massimi che giungono quasi a 1200 mm.



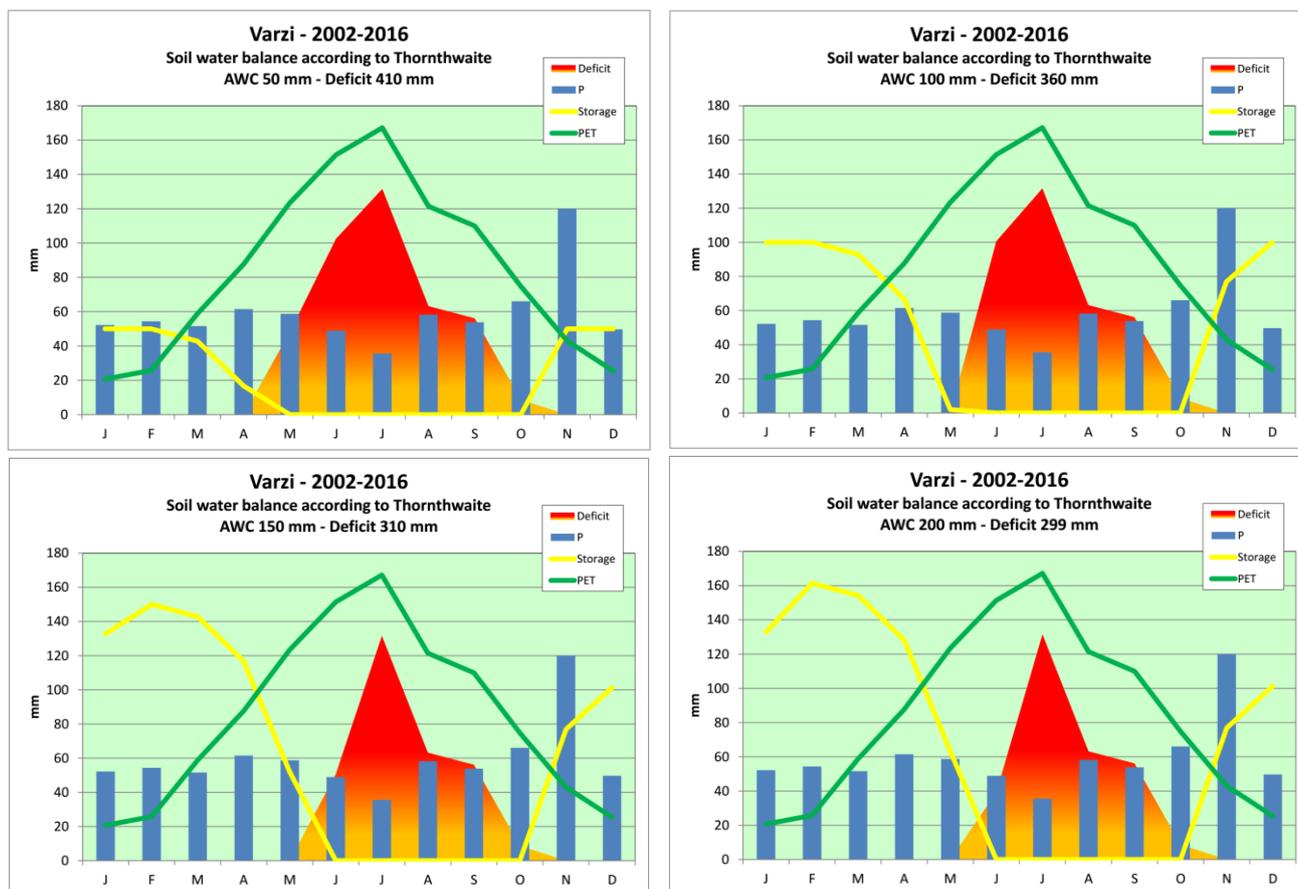
**Fig. 2.3** Precipitazioni totali annuali 2002-2016 per la stazione di Varzi Nivione (elaborazione su dati giornalieri ARPA Lombardia). Il dato 2010 non è riportato, a causa dei troppi valori giornalieri mancanti.



**Fig. 2.4** Temperature medie mensili per le stazioni di Varzi Nivione (2002-2016) e Montemarzino (1933-1955).

Riguardo alle temperature, è interessante notare il deciso incremento che si è verificato negli ultimi decenni. In fig. 2.4 si possono confrontare le temperature medie mensili della stazione di Varzi Nivione e quelle di Montemarzino, distanti fra loro solo pochi km e poste quasi alla stessa quota

(Varzi Nivione 500 m, Montemarzino 468 m). Nel periodo 1933-1955 la temperatura media annuale si attestava a 11.2 °C, mentre per il 2002-2016 è salita a 12.5 °C. L'incremento è presente in tutti i mesi, ma è più sensibile in quelli invernali, rispetto agli estivi.



**Fig. 2.5** Bilancio idrico sec. Thornthwaite per la stazione di Varzi Nivione, relativo a differenti riserve idriche (AWC) del suolo.

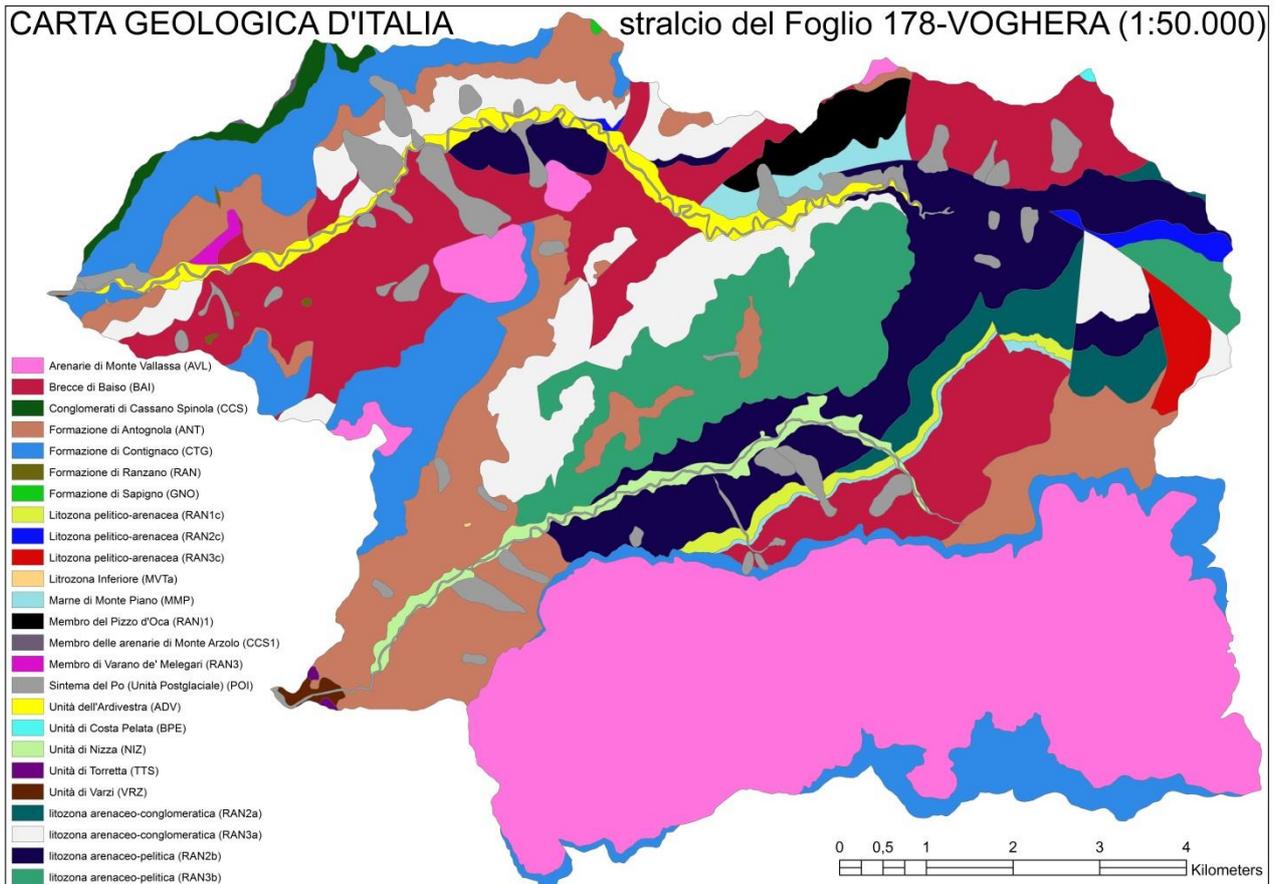
In fig. 2.5 sono visibili i risultati delle elaborazioni per il calcolo del bilancio idrico dei suoli. È stato adottato il modello di Thornthwaite, che stima l'evapotraspirazione potenziale (*potential evapotranspiration* (PET)) a partire unicamente dai dati di temperatura media mensile e di latitudine del sito. Il bilancio è stato svolto per suoli a diversa riserva idrica (*available water capacity* (AWC)), da 50 a 200 mm. Sono riportati lo stato della riserva idrica nel corso dell'annata media e l'entità del deficit annuale. Per i suoli a più limitata riserva (sottili, a tessitura sciolta, oppure ricchi di scheletro) il deficit idrico è molto sensibile e si aggira sui 300 mm/anno. Tuttavia, anche i suoli dotati di una consistente riserva (profondi, a tessitura media e privi di scheletro) vanno soggetti a un deficit importante, di circa 300 mm/anno. Questi dati climatici e di bilancio idrico fanno ipotizzare il manifestarsi nell'area di processi pedologici di vertisolizzazione, laddove le caratteristiche dei suoli (tessitura argillosa e argille di tipo smectitico) li rendano possibili. Va ricordato che il bilancio idrico è stato calcolato per una annata media (precipitazioni di circa 700 mm): per quelle più siccitose (meno di 500 mm) il deficit idrico e la probabilità di incremento del processo di vertisolizzazione sono senz'altro maggiori.

## 2.2 Breve descrizione geologica e dei litotipi

L'area di studio è inclusa nel Foglio 178-Voghera, della Carta Geologica d'Italia (1:50.000), (Progetto CARG) (Vercesi et al. 2015) (Fig. 2.6). Secondo la legenda sono presenti i litotipi appartenenti alle Unità Epiliguri dell'Eocene Medio e Miocene Superiore e ai depositi del Quaternario pleistocenico-olocenico (Unità Di Nizza e Unità dell'Ardivestra, sistema del Pò), indicati in tabella. Tuttavia, considerando l'areale relativo, la presenza di alcune formazioni è esigua, come desumibile dalla seguente tabella.

**Tab. 2.1** Elenco delle Unità Geologiche presenti nell'area di studio, con la sigla adottata nel Foglio 178-Voghera e con il relativo areale.

UG_DESCR25	UG_SIGLA25	area km <sup>2</sup>	area %
Arenarie di Monte Vallassa	AVL	25,2044	27,48
Brecce di Baiso	BAI	12,2938	13,4
Formazione di Antognola	ANT	10,7329	11,7
Formazione di Contignaco	CTG	9,8681	10,76
Formazione di Ranzano - Litozona arenaceo-pelitica	RAN <sub>2b</sub>	8,1217	8,86
Formazione di Ranzano - Litozona arenaceo-conglomeratica	RAN <sub>3a</sub>	6,5243	7,11
Formazione di Ranzano - Litozona arenaceo-pelitica	RAN <sub>3b</sub>	6,4703	7,05
Sistema del Po (Unità Postglaciale)	POI	3,9847	4,34
Formazione di Ranzano - Litozona arenaceo-conglomeratica	RAN <sub>2a</sub>	1,925	2,1
Unità dell'Ardivestra	ADV	1,2387	1,35
Formazione di Ranzano - Membro del Pizzo d'Oca	RAN <sub>1</sub>	1,0283	1,12
Unità di Nizza	NIZ	0,9232	1,01
Marne di Monte Piano	MMP	0,8253	0,9
Formazione di Ranzano - Litozona pelitico-arenacea	RAN <sub>3c</sub>	0,6513	0,71
Formazione di Ranzano - Litozona pelitico-arenacea	RAN <sub>1c</sub>	0,5828	0,64
Conglomerati di Cassano Spinola	CCS	0,5825	0,64
Formazione di Ranzano - Litozona pelitico-arenacea	RAN <sub>2c</sub>	0,3906	0,43
Unità di Varzi	VRZ	0,1291	0,14
Formazione di Ranzano - Membro di Varano de' Melegari	RAN <sub>3</sub>	0,1177	0,13
Unità di Torretta	TTS	0,0322	0,04
Formazione di Ranzano	RAN	0,0288	0,03
Membro delle arenarie di Monte Arzolo	CCS <sub>1</sub>	0,0222	0,02
Unità di Costa Pelata	BPE	0,0187	0,02
Formazione di Sapigno	GNO	0,0143	0,02
Litozona Inferiore	MVT <sub>a</sub>	0,0031	0



**Fig. 2.6** Area di studio con le Unità Geologiche (stralcio del Foglio 178-Voghera, della Carta Geologica d'Italia (1:50.000) del Progetto CARG).

Le Breccie argillose di Baiso (BAI) sono la prima formazione della successione Epiligure. Secondo le ipotesi recenti esse sono costituite da accumuli di depositi di frane e colate sottomarine. Nell'area in esame la Formazione è presente in entrambe le Valli. Comprende una serie di corpi rocciosi a tessitura caotica (del tipo *block in matrix*), costituiti da breccie poligenetiche a matrice argillosa grigia o varicolori, in cui vi si ritrovano clasti spigolosi eterometrici, principalmente di arenarie, argille e flysch (Vercesi et al. 2015).

La Formazione di Ranzano (RAN) è caratterizzata dai 5 membri informali: - membro di Pizzo d'Oca; - membro della Val Pessola, - membro di Varano de' Melegari; - membro di S. Sebastiano Curone; - membro di Albergana. Ogni membro risulta costituito da uno o più sistemi torbiditici.

I primi due, verticalmente sovrapposti non eteropici, definiscono gli intervalli stratigrafici RAN<sub>1</sub> e RAN<sub>2</sub> rispettivamente; gli ultimi tre, eteropici tra loro, insieme identificano l'intervallo superiore, RAN<sub>3</sub>. Quindi da un punto di vista stratigrafico RAN<sub>3</sub> risulta sovrapposto a RAN<sub>2</sub> che a sua volta sovrasta RAN<sub>1</sub>. All'interno di ogni membro e/o sub-unità stratigrafica si osserva una notevole variabilità laterale e verticale delle facies, ma le principali distinte sono: arenaceo-conglomeratica (a), arenaceo-pelitica(b) e pelitico-arenacea(c) (Martelli et al., 1998, e Mutti et al., 1995).

Nell'area in studio, in entrambe le Valli, sono stati rilevati punti di campionamento su aree appartenenti alle formazioni geologiche RAN<sub>2</sub> o RAN<sub>3</sub>, in particolare, in base alla suddivisione delle *litofacies* (a, b, c), RAN<sub>2a</sub> e RAN<sub>2b</sub>, RAN<sub>3a</sub>, RAN<sub>3b</sub> e RAN<sub>3c</sub>. Di seguito si descrivono solo le sottocategorie di interesse .

Membro del Pizzo d'Oca (RAN<sub>1</sub>). Arenarie e peliti in proporzioni variabili con presenza di detrito derivante da erosione di un basamento cristallino.

Membro di Val Pessola (RAN<sub>2</sub>). Arenarie, conglomerati e peliti in proporzioni molto variabili; sono presenti le *litofacies arenaceo-conglomeratica* (RAN<sub>2a</sub>), *litofacies arenaceo-pelitica* (RAN<sub>2b</sub>) e *litofacies pelitico-arenacea* (RAN<sub>2c</sub>), che si trovano verticalmente sovrapposte a delineare un sistema torbido, con alla base i sedimenti conglomeratici (a) e al tetto quelli pelitico-arenacei (c): ciò si verifica anche in RAN<sub>1</sub> e in RAN<sub>3</sub>. Abbondante presenza di detrito ofiolitico, che conferisce alle arenarie un tipico colore verdastro.

*Litofacies arenaceo-conglomeratica* (RAN<sub>2a</sub>): affiora lungo il margine nord-orientale della placca epiligure delle valli di Nizza e Ardivestra; è costituita da arenarie conglomeratiche e conglomerati in strati spessi e molto spessi, amalgamati o separati da sottili livelli pelitici, con basi erosive e scarsa continuità laterale. I ciottoli raggiungono diametri di 50 cm e sono di solito ben arrotondati.

*Litofacies arenaceo-pelitica* (RAN<sub>2b</sub>): sta nel mezzo tra la litofacies a (sotto) e c (sopra); affiora in varie zone del Foglio, tra cui in Val di Nizza. La litofacies è costituita da arenarie da grossolane a fini in strati medi e sottili, con geometria piano-parallela e rapporto arenaria/pelite maggiore o uguale a 1;

Membro di Varano de' Melegari (RAN<sub>3</sub>). Costituito da arenarie, conglomerati e peliti in proporzioni molto variabili. I conglomerati sono particolarmente ricchi di ciottoli tipo flysch ad Elmintoidi, e le arenarie hanno un caratteristico colore grigio.

*Litofacies arenaceo-conglomeratica* (RAN<sub>3a</sub>): affiora in diverse località, lungo il margine settentrionale delle placche epiliguri in particolare della Val di Nizza e Val Ardivestra (area di Piumesana - M.Gaggiolo - Zuccarello, area di Fornace - M. Chiaro). È costituita da conglomerati ed arenarie conglomeratiche in strati spessi e molto spessi, con basi erosive, quasi sempre amalgamati (rapporto arenaria/pelite >>1), con ciottoli di diametro fino ai 50-60 cm ed arenarie da molto grossolane a grossolane.

*Litofacies arenaceo-pelitica* (RAN<sub>3b</sub>): affiora soprattutto in Val di Nizza e Val Ardivestra ed è costituita di strati da sottili a spessi di arenarie da grossolane a fini.

*Litofacies pelitico-arenacea* (RAN<sub>3c</sub>): affiora in un'area molto limitata nei pressi di Calghera, tra la Val di Nizza e la Val Ardivestra; è costituita da strati da molto sottili a medi di arenarie medie e fini, con geometria piano-parallela e rapporto arenaria/pelite <1. La successione stratigrafica della Formazione di Ranzano è variabile, da zona a zona: nella figura successiva è riportata quella dell'area della Val di Nizza - Val Ardivestra (Vercesi et al. 2015).

La Formazione di Antognola (ANT) è presente in entrambe le valli, consta di marne e marne argillose verdognole o grigie, con caratteristiche patine di manganese (colore scuro-nero) sulle superfici di frattura. Presenta al suo interno associazioni di nanofossili e foraminiferi planctonici eocenici e oligocenici. Il limite inferiore è graduale con il membro di Varano de' Melegari (RAN<sub>3</sub>) della Formazione di Ranzano, sottostante; quello superiore è sempre graduale, con la Formazione di Contignaco (Vercesi et al. 2015).

La Formazione di Contignaco (CTG) è interpretata come sedimento pelagico deposto in un bacino non molto profondo, è costituita da marne siltose grigie e marne calcaree chiare, alle quali si intercala, nella parte inferiore, una successione di strati arenaceo-pelitici. Caratteristica della formazione è la presenza di silice, di origine sia biogenica che vulcanica. Il limite inferiore è transizionale con la Formazione di Antognola; il limite superiore è, per discordanza, con le arenarie

di Monte Vallassa. Alcuni punti di campionamento hanno interessato questa formazione in entrambe le valli (Vercesi et al. 2015). Nell'area di studio la grande "placca" di Pizzocorno-Pietragavina è costituita dalla formazione delle Arenarie di Monte Vallassa (AVL). E' presente anche nella val Ardivestra con tre piccole finestre che costituiscono "placche" morfologicamente rilevate. La formazione è costituita da arenarie bioclastiche e biocalcareni; in subordine sono presenti conglomerati fini a bioclasti e biocalciruditi, con clasti di rocce metamorfiche verdi alternate da marne sabbiose grigio-azzurre bioturbate (Vercesi et al. 2015). L'Unità di Nizza (NIZ) e l' Unità dell'Ardivestra (ADV) sono costituite da depositi alluvionali: ghiaie con matrice sabbioso-limosa, ciottoli da sub-arrotondati ad arrotondati di alimentazione strettamente locale. L'unità è attribuibile al Pleistocene superiore, con una continuazione della sedimentazione anche per parte dell'Olocene (Vercesi et al. 2015).

### 2.3. Delineazione dei principali usi del suolo e sintetica caratterizzazione della copertura vegetale

In relazione agli obiettivi del presente lavoro, l'analisi delle classi di uso del suolo dell'area, si è rivelata esplicativa. Sia la più recente carta della *Destinazione d'Uso dei Suoli Agricoli e forestali* (DUSAF Regione Lombardia e ERSAF, 2012) che quella relativa al recente passato (uso del suolo storico (1954) (Regione Lombardia) sono concordi nell'individuare all'interno dell'area di studio tre fondamentali destinazione d'uso, desunte dalla semplificazione della legenda della carta (Fig. 2.7): aree adibite ad uso agricolo, aree ove vi è una crescita di vegetazione e aree antropizzate.

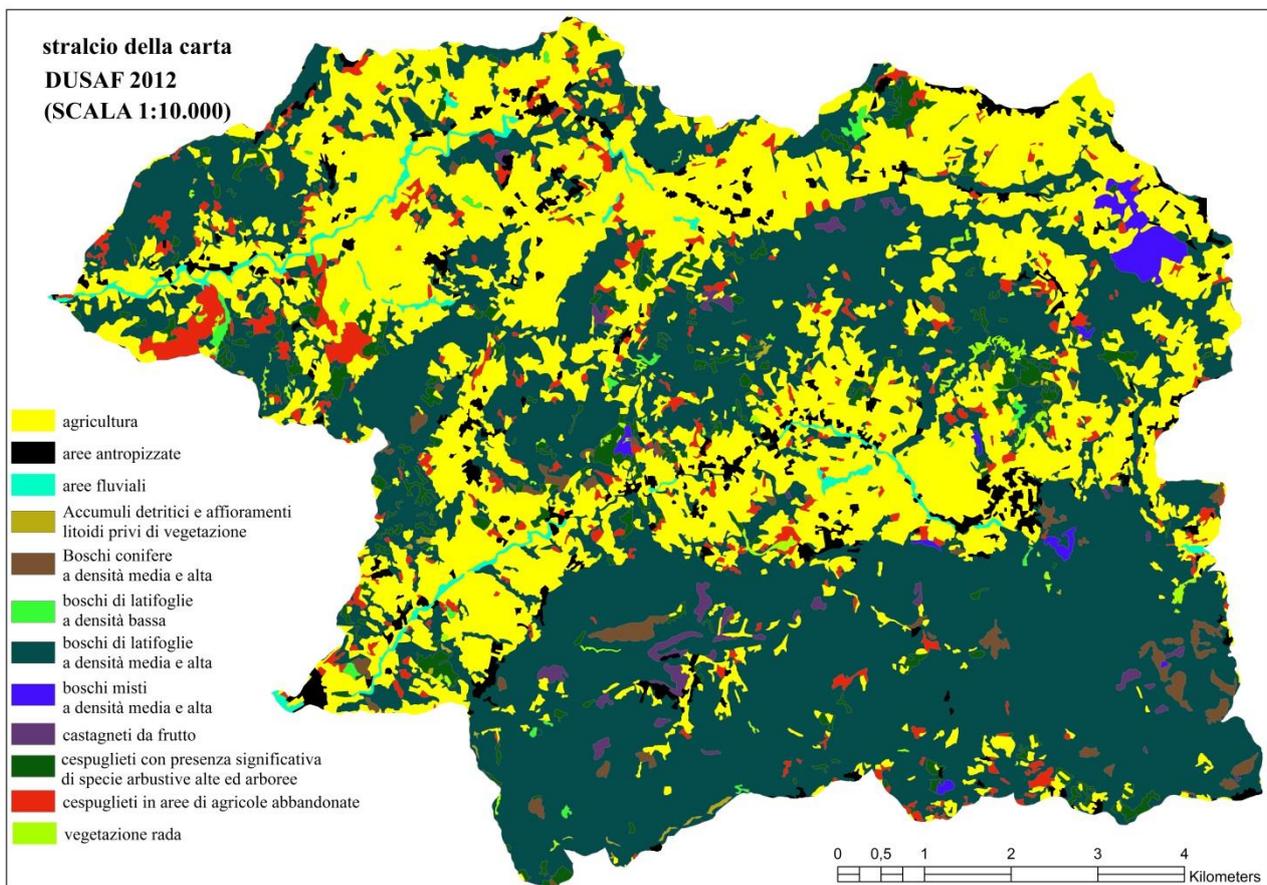
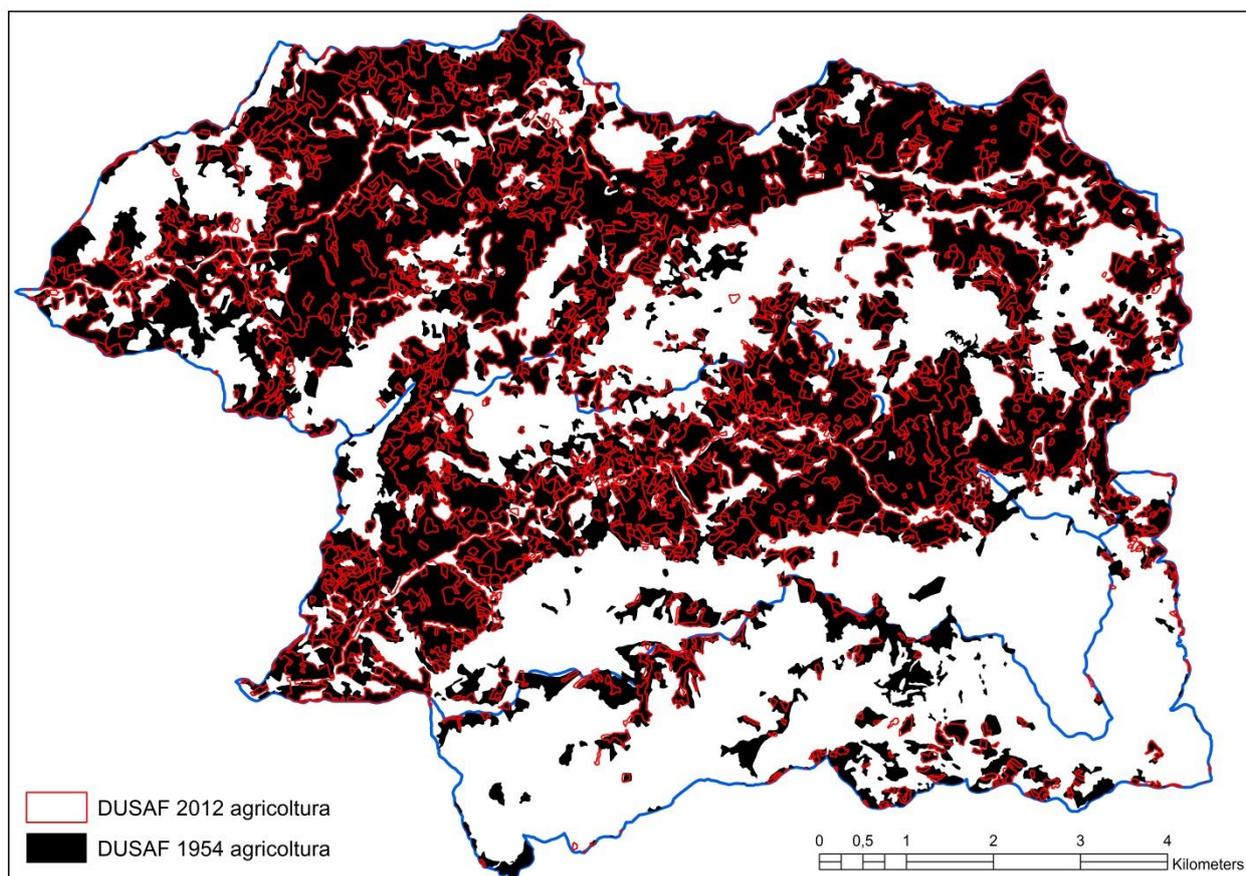


Fig. 2.7 Area di studio con le classi d'uso del suolo secondo la carta DUSAF del 2012.

Confrontando questa partizione del territorio della mappa DUSAF dell'anno 2012 con quella desumibile dalla carta d'uso del suolo del 1954, emerge che nel corso del tempo è avvenuta una variazione d'uso del territorio, dovuta all'abbandono di pratiche agricole (Fig. 2.8) che ha determinato un incremento delle aree con vegetazione arbustiva o arborea. Tale variazione è pari a circa il 15% dell'intera superficie considerata, come risulta dalle seguenti tabelle (Tab. 2.2 e Tab. 2.3, 2.4).

**Tab. 2.2** Confronto degli areali adibiti ad uso agricolo e occupati da vegetazione spontanea relativi alla carta DUSAF del 1954 con la carta DUSAF del 2012.

aree ad uso agricolo		area (Km <sup>2</sup> )	% area sul totale dell'area di studio
1954	Totale complessivo	46,9	51,2
2012	Totale complessivo	31,7	34,6
aree ricoperte da vegetazione a crescita spontanea e affioramenti litoidi		area (Km <sup>2</sup> )	% area sul totale dell'area di studio
1954	Totale complessivo	42,6	46,4
2012	Totale complessivo	56,4	61,5



**Fig. 2.8** Variazione dell'areale adibito ad uso agricolo avvenuta tra il 1954 e il 2012, con perimetro delle due valli (linea blu).

**Tab. 2.3** Aree delle classi d'uso del suolo della carta DUSAF del 1954.

<b>DUSAF 1954</b>		
<b>aree ad uso agricolo</b>	area Km <sup>2</sup>	area %
Altre legnose agrarie	0,01	0,02
Arboricoltura da legno	0,01	0,02
Colture orticole	0,08	0,17
Frutteti e frutti minori	0,28	0,6
Prati permanenti con presenza di specie arboree ed arbustive sparse	0,99	2,12
Prati permanenti in assenza di specie arboree ed arbustive	2,2	4,7
Seminativi arborati	28,45	60,64
Seminativi semplici	14,23	30,32
Vigneti	0,66	1,4
<b>Totale</b>	<b>46,92</b>	<b>51,16</b>
<b>aree ricoperte da vegetazione spontanea</b>	area Km <sup>2</sup>	area %
Boschi conifere a densità media e alta	0,34	0,8
Boschi di latifoglie a densità bassa	2,88	6,76
Boschi di latifoglie a densità media e alta	34,46	80,89
Boschi misti a densità media e alta	0,51	1,2
Castagneti da frutto	0,02	0,05
Cespuglieti con presenza significativa di specie arbustive alte ed arboree	3,76	8,84
Vegetazione dei greti	0,18	0,43
Vegetazione rada	0,42	0,98
<b>Totale</b>	<b>42,6</b>	<b>46,45</b>

**Tab. 2.4** Aree delle classi d'uso del suolo della carta DUSAF del 2012.

<b>DUSAF 2012</b>		
<b>aree ad uso agricolo</b>	area Km <sup>2</sup>	area %
altre legnose agrarie	1,97	6,22
colture orticole a pieno campo	0,03	0,08
frutteti e frutti minori	1,29	4,06
imboschimenti recenti	0,01	0,02
oliveti	0	0,02
pioppeti	0,03	0,1
prati permanenti con presenza di specie arboree ed arbustive sparse	0,42	1,32
prati permanenti in assenza di specie arboree ed arbustive	2,08	6,57
seminativi semplici	25,08	79,13
vigneti	0,79	2,48
<b>Totale</b>	<b>31,69</b>	<b>34,56</b>
<b>aree ricoperte da vegetazione spontanea</b>	area Km <sup>2</sup>	area %
Boschi conifere a densità media e alta	1,03	1,83
boschi di latifoglie a densità bassa	0,28	0,49
boschi di latifoglie a densità media e alta	47,18	83,6
boschi misti a densità media e alta	0,67	1,19
castagneti da frutto	0,66	1,17
cespuglieti con presenza significativa di specie arbustive alte ed arboree	2,83	5,02
cespuglieti in aree di agricole abbandonate	3,29	5,84
vegetazione rada	0,39	0,69
<b>Totale</b>	<b>56,43</b>	<b>61,53</b>

Gli insediamenti antropici sono la componente meno rilevante dell'area (meno del 5% del territorio) e risulta frammentaria. La popolazione è distribuita in nuclei abitati di piccole dimensioni e case sparse, ciononostante non può essere considerata una zona scarsamente antropizzata. Occorre ricordare, infatti, che la presenza dell'uomo, oltre a incidere sul territorio con lo sfruttamento per scopi agricoli, influenza anche la distribuzione della vegetazione, sia con la gestione della ceduzione che con la piantumazione di specie alloctone. Inoltre, occorre sottolineare che con le lavorazioni del terreno, propedeutiche alla semina, a volte predispongono il territorio a fenomeni franosi di tipo superficiale.

Le aree destinate ad uso agricolo, comprendono seminativi principalmente coltivati a erba medica, in avvicendamento con il frumento, o aree con alberi da frutto quali vite e meli.

La copertura boschiva occupa la maggior parte del territorio ed è costituita principalmente da boschi misti di latifoglie decidue e, a seconda delle condizioni, di latifoglie decidue con elevate coperture delle componenti arbustive ed erbacee. Nell'area in esame la presenza e la distribuzione della vegetazione è fortemente dipendente dalle condizioni edafiche e dalle condizioni termo-climatiche. Questi due fattori sono influenzati dalla morfologia in combinazione con l'effetto dei substrati litologici. In particolare, la morfologia determina un orientamento prevalentemente est-ovest delle due valli e della "placca" di Pizzocorno-Pietragavina e conseguentemente una esposizione dei versanti o a Nord o a Sud. I litotipi presenti nell'area originano prevalentemente suoli a pH alcalino e a volte suoli a pH neutro o sub-acido. Gli effetti delle combinazioni, dei fattori citati, influenzano il paesaggio vegetale locale in termini di dominanza di specie e composizione floristica.

Per quanto riguarda le condizioni edafiche la litologia dell'area determina tre tipologie differenti:

- associazioni di marne e peliti o siltiti con intercalazioni di livelli sabbiosi e/o arenacei, quali le Marne di M. Piano, Arenarie di Ranzano (facies marnoso-arenacea), Marne di Antognola, Formazione di Contignaco; su questi litotipi si generano suoli con pH alcalino.

*Fisionomia della vegetazione.* La cenosi forestale che si trova in corrispondenza di tali substrati è costituita da querceti a roverella (*Quercus pubescens*) i quali, date le caratteristiche termofile, si stabiliscono principalmente sui versanti esposti a sud. In condizioni meno termofile la roverella si trova maggiormente in consorzi misti associata all'orniello (*Fraxinus ornus*), al carpino nero (*Ostrya carpinifolia*) e, talvolta, al cerro (*Quercus cerris*). Tra gli arbusti, sono frequenti il ginepro (*Juniperus communis*), i biancospini (*Crataegus monogyna*), il citiso (*Cytisus sessilifolius*), il ligustro (*Ligustrum vulgare*) e il prugnolo (*Prunus spinosa*) (Andreis e Sartori 2009; Assini et al. 2010).

- rocce coerenti arenacee, costituite dalla formazione delle Arenarie di M. Vallassa ("placca" Pizzocorno-Pietragavina); su questo litotipo si possono generare sia suoli a pH neutro e acido che suoli con pH alcalino.

*Fisionomia della vegetazione.* Su questo litotipo, su versanti esposti a nord, sono presenti consorzi forestali mesofili dominati da cerro. In tali boschi il cerro è associato, talvolta, alla roverella, rovere (*Quercus petraea*) e/o carpino nero. Tra gli arbusti, sono frequenti il citiso e varie specie di ginestre (*Cytisus scoparius*), ma anche le rose (*Rosa canina*), il biancospino (*Crataegus monogyna*), il caprifoglio (*Lonicera caprifolium*), il ligustro, il prugnolo, il pero selvatico (*Pyrus pyraster*). In condizioni edafiche ad alto contenuto di carbonato di calcio

e/o su versanti esposti a sud, troviamo cenosi con prevalenza di carpino nero, pianta relativamente più mesofila. In tali boschi il carpino nero può essere accompagnato, da roverella, orniello. Tra gli arbusti, sono frequenti il maggiociondolo (*Laburnum anagyroides*), il citiso, l'acero campestre (*Acer campestre*), il biancospino, il pero selvatico. Sulle parti sommitali della "placca" o sui versanti esposti a nord si trovano anche i castagneti (*Castanea sativa*), introdotto dall'uomo in tempi storici. Tra gli arbusti, sono frequenti il biancospino, l'orniello, il nocciolo (*Corylus avellana*), il ciavardello (*Sorbus torminalis*), che vanno a costituire con il castagno consorzi mesofili.

- argilliti con stratificazione anche obliterata, inglobanti frammenti e blocchi lapidei di natura diversa, dispersi disordinamene nella matrice corrispondente alla Formazione delle Breccie di Baiso; solitamente generano suoli con pH alcalino.

*Fisionomia della vegetazione.* Le aree corrispondenti a questo substrato litologico sono quasi interamente occupati da seminativi a erba medica in rotazione col frumento o da campi incolti.

In tutte le litologie citate, nelle aree degradate o molto influenzate dalla presenza antropica, sono presenti formazioni a robinia. Altre formazioni boschive di origine antropica impiantate in tempi storici per scopi forestali sono: i boschi a pino silvestre (*Pinus sylvestris*) e /o pino nero (*Pinus nigra*).

#### *La vegetazione delle aree in erosione sui calanchi e sulle biancane*

Nell'area di studio, la vegetazione di calanchi e biancane e delle loro morfologie intermedie è sempre caratterizzata da boschi misti di roverella in associazione con carpino nero. Tuttavia, il diverso grado di attività dei processi erosivi (inattivi, quiescenti, e attivi) determina importanti cambiamenti nel grado di copertura vegetale e nella composizione floristica degli strati arbustivo ed erbaceo:

- processi inattivi: copertura dello strato arboreo generalmente superiore al 60%; roverella dominante;
- processi quiescenti: copertura dello strato arboreo dal 20% al 60%; elevata frequenza di carpino e frassino e nello strato arbustivo di ginestra odorosa (*Spartium junceum*), specie tipica di versanti più instabili e in via di colonizzazione;
- processi attivi: copertura dello strato arboreo inferiore al 20% ; elevata frequenza di ginestra odorosa e nelle aree a maggior grado di attività strato erbaceo caratterizzato da elevata frequenza di graminacee (*Brachypodium rupestre*, *Dactylis glomerata*, *Festuca* sp.).

## **2.4 Inquadramento pedologico**

Riguardo le conoscenze pregresse relative alle informazioni pedologiche, per l'area di studio, occorre considerare la carta dei suoli della Regione Lombardia (scala 1:250.000) (Fig. 2.9), la quale è stata realizzata nell'ambito di un progetto finalizzato alla realizzazione della "Prima approssimazione della base dati georeferenziata dei suoli d'Italia". Secondo l'organizzazione gerarchica di questa carta, l'area di studio appartiene alla Regione pedologica (Soil Region) delle colline appenniniche; alle Provincie pedologiche (Soil Sub-Regions) della Bassa e media collina pavese e Alta collina: pendici submontane e montane. Più in particolare, la val di Nizza e la val

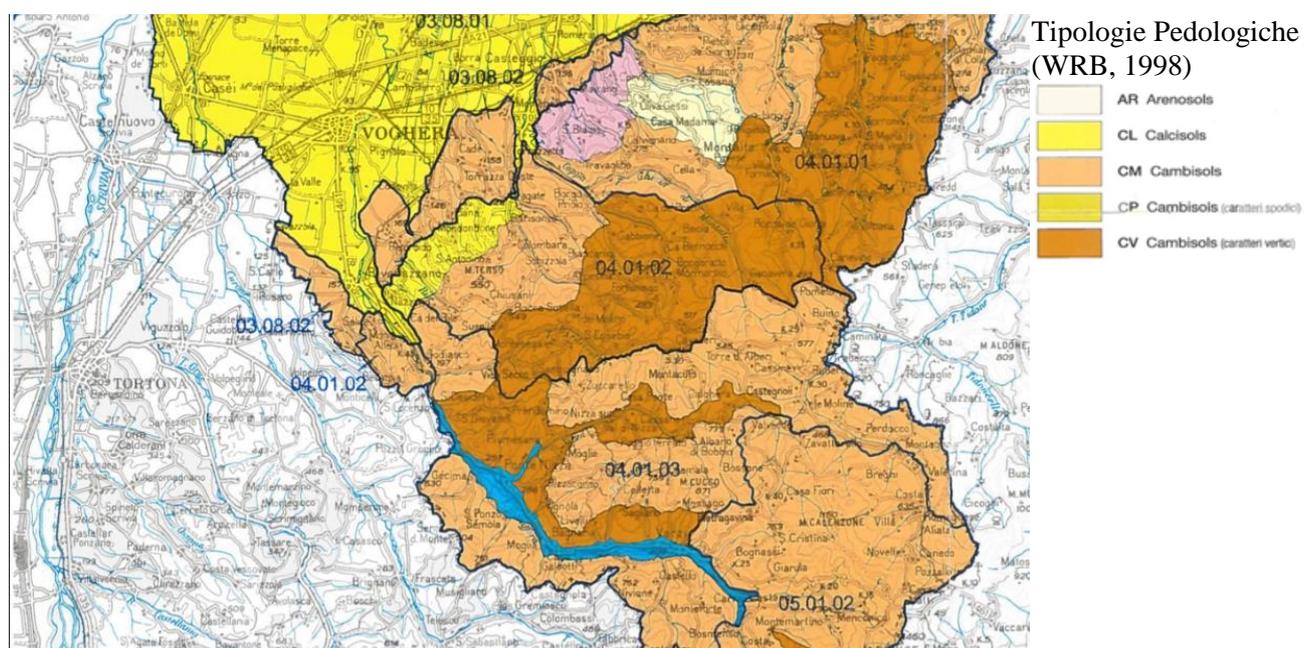
Ardivestra ricadono nei distretti (Grat Soilscape) denominati: Basse colline interne ad est della Val Staffora (04.01.02) e Alta collina e media montagna 04.01.03.

Il distretto 04.01.02, secondo i suoli prevalenti rilevati, contiene le Unità Tipologiche del Suolo (UTS), le quali classificate secondo il World Reference Base (WRB) (FAO, 1998), sono così descritte:

- suoli argillosi con scheletro scarso, alcalini saturi e profondi. Presentano orizzonte vertico con slickensides tra 130 e i 100 cm di profondità. hanno un profilo tipico A-Bss-C, sono classificati prevalentemente Calcaric-Vertic Cambisols (CV-01);
- suoli franchi con scheletro, alcalini o molto alcalini, calcarei, con drenaggio mediocre, segni di idromorfia e permeabilità da moderata in superficie a moderatamente bassa in profondità, hanno un profilo tipico Ap-Bg-Cgk; sono classificati prevalentemente Hypercalci-Gleyic Calcisols (Skeletal) (CL-02);
- suoli mediamente franco-limosi-argillosi, con scheletro comune, acidi e saturi, molto profondi; hanno un profilo tipico Ap-B<sub>w</sub>-BC-C; sono classificati prevalentemente Calcaric Cambisol (CM-14).

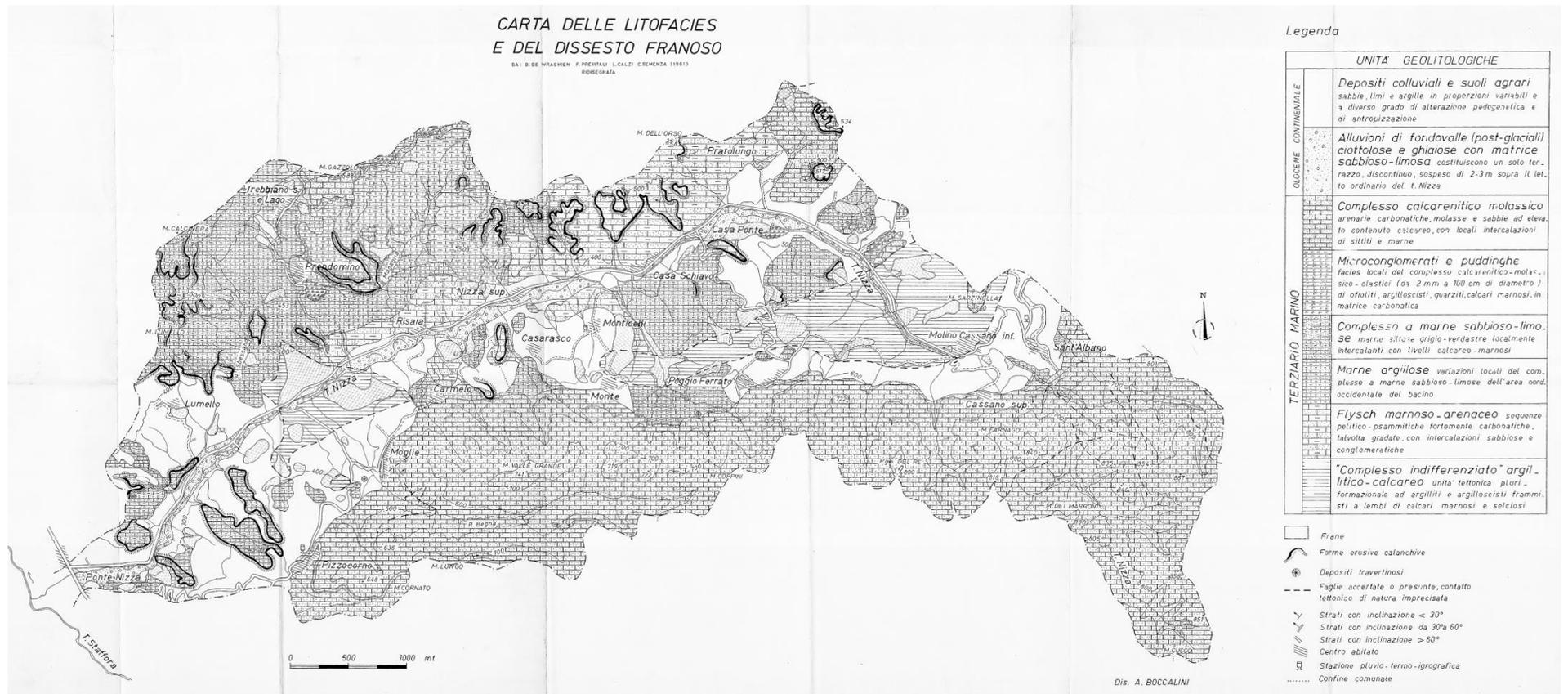
Il distretto 04.01.03, secondo i suoli prevalenti rilevati, contiene, oltre all'UTS CV-01, anche l'Unità Tipologia del Suolo (UTS), di seguito descritta:

- Suoli argillosi subalcalini calcarei molto calcarei poco profondi limitati da sub-strati litoidi coerenti entro i 60 cm dalla superficie; hanno un profilo tipico A-B<sub>w</sub>-C-R sono classificati prevalentemente come Calcaric-Leptic Cambisol (CM-15).

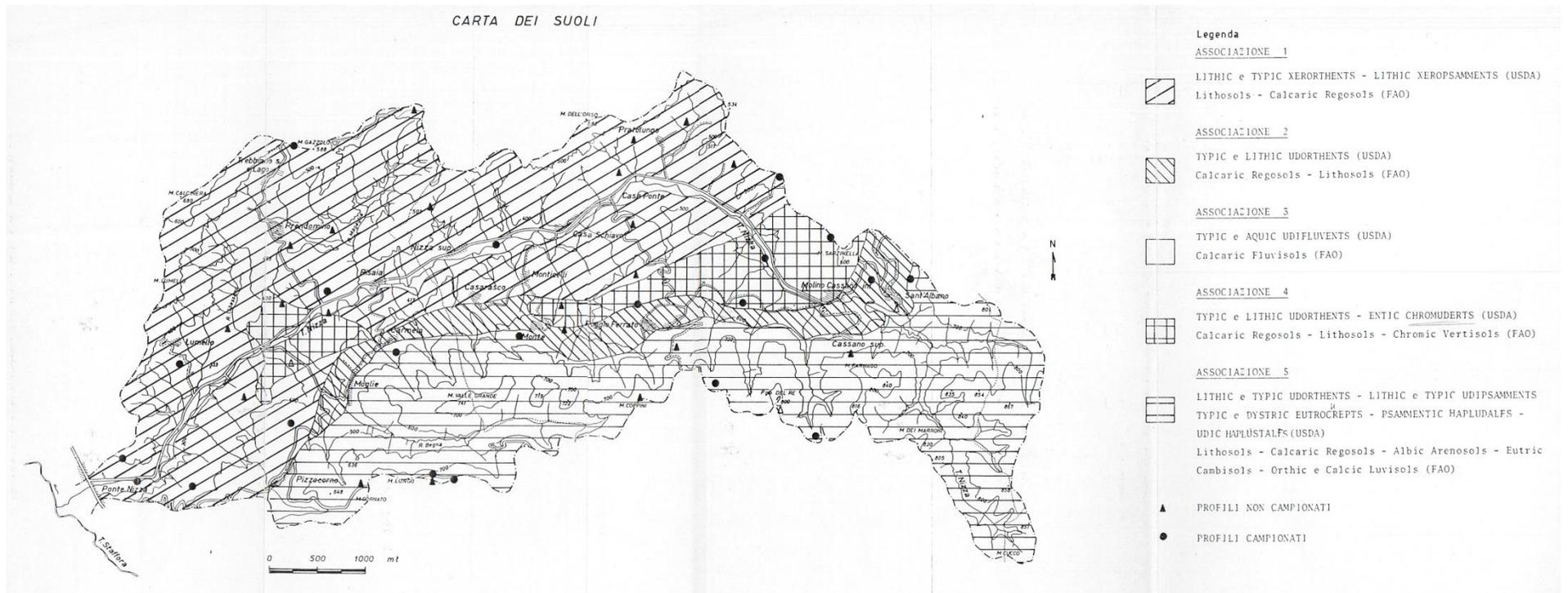


**Fig. 2.9** Dettaglio della Carta dei Suoli della Regione Lombardia (1:250000) e dettaglio della legenda.

Riguardo i suoli e i dissesti franosi, nel territorio della Val di Nizza è stato condotto uno specifico studio a seguito del quale sono state realizzate, tra l'altro, una mappa (1:10.000) riportante le unità geolitologiche denominata *Carta delle Litofacies e del dissesto franoso* (De Wrachien et al.1981) (Fig. 2.10) ed una *Carta dei suoli* (1:10.000) (Casati, 1985), nella quale sono state rappresentate cinque associazioni di suoli, classificati secondo la Soil Taxonomy (Fig. 2.11).



**Fig. 2.10** Carta delle Unità geolitologiche (scala 1:10.000) della Val di Nizza (De Wrachien et al.1981).



**Fig. 2.11** Carta dei suoli della Val di Nizza (scala 1:10.000) (Casati 1985).

## **Bibliografia**

Andreis, Carlo; Sartori, Francesco (2009): Sintassonomia dei boschi lombardi. In: *Natura Bresciana, Ann. Mus. Civ. Sc. Nat., Brescia*, 36, pp. 173–178.

Assini, Silvia; Bracco, Francesco; Carrea, Augusta; Villani, Mariacristina (2010): Stato delle conoscenze floristico-vegetazionali sul corso planiziale del Fiume Po e dei suoi affluenti. In: *Biol Ambientale*, 24, pp. 1–18.

Carrara, Alberto. (2004): Generazione delle linee d'impluvio e di displuvio e calcolo dei parametri morfologici dei sottobacini elementari appartenenti al territorio collinare-montano della Regione Lombardia. Consiglio Nazionale delle Ricerche; IEIIT - Istituto di Elettronica e di Ingegneria dell'Informazione e delle Telecomunicazioni - Sezione di Bologna - Università di Bologna.

Casati, Enrico. (1985): Analisi Geomorfologica e valutazione territoriale di un'area appenninica a intenso degrado fisico-ambientale: il bacino del Torrente Nizza (Oltrepo Pavese).

ERSAF – Regione Lombardia (2012) *La Carta delle Destinazioni d'uso dei suoli agricoli e forestali DUSAF4*

Vercesi P.L., Faletti P., Pasquini C., Papani L., Parotti C., Tucci. G. (2015) - Note Illustrative della carta Geologica d'Italia in scala 1:50.000, foglio178 "Voghera". Servizio Geologico d'Italia - ISPRA, 166 pp.

De Wrachien D., Previtali F, Calzi L., Semenza C. (1981) Lineamenti Geologici-Agrari del bacino del torrente Nizza (Oltrepo Pavese).

## CAPITOLO 3

### Le variabili geomorfometriche

#### 3.1 Introduzione

Ricordando che l'intento in questo lavoro è applicare le tecniche di DSM, utilizzando semplice informazione e considerando quanto esposto (v. Cap. 1), riguardo i concetti contenuti nel paradigma suolo-paesaggio, per questa area di studio, sono stati calcolati 30 parametri geomorfometrici (v. Tab. 3.1), utilizzando un Digital Elevation Model (DEM), discretizzato a 10 metri, elaborato da ricercatori dell'Università di Pavia<sup>1</sup>. Questi parametri sono le variabili indipendenti utilizzate per le elaborazioni statistiche finalizzate alla costruzione di mappe tematiche relative al suolo. Esistono svariati metodi in letteratura per misurare le irregolarità del terreno (Moore et al. 1991; Pike et al. 2009; Behrens et al. 2005; Florinsky 2012; Evans 2012; Gallant 2000; Hengl e Reuter 2009; Shary et al. 2002). La scelta delle variabili non è stata basata su una ricerca per individuare quelle che abbiano un particolare significato fisico o geometrico (Florinsky 2012) ma sulla possibilità di eseguire il calcolo senza dover introdurre ulteriore informazione oltre al DEM. Sono stati considerati alcuni parametri geomorfometrici, di semplice costruzione, che possono fornire l'informazione sostitutiva per caratterizzare un territorio, riguardo gli aspetti topografici e ambientali che potenzialmente possono influenzare i processi pedogenetici. Non ci si aspetta che queste variabili procurino l'informazione per esplicitare funzioni che possano spiegare i processi di formazione del suolo né che siano i migliori *proxy* dei fattori pedogenetici: rievocando gli obiettivi di questo lavoro, si vuole testare se una caratterizzazione paesaggistica basata sulla sola geomorfometria è in grado di fornire una sufficiente informazione per derivare, con metodi statistici, mappe tematiche relative al suolo.

**Tabella 3.1** Denominazione e sigla delle variabili geomorfometriche calcolate.

	nome variabile	sigla		nome variabile	sigla
1	elevation	Zstd	16	Compound Topographic Index (GRASS)	CTIgrass
2	Terrain Ruggedness Index (QGIS)	TRI	17	catchment Area (recursive method) (SAGA)	catarea
3	Topographic position index (QGIS)	TPI	18	mass balance index (SAGA)	massbal
4	Asperity (QGIS)	asperit	19	LS factor (SAGA)	LSfactor
5	total curvature (Jenness tool)	ctot	20	convergence index (SAGA)	ConvInd
6	tangential curvature (Jenness tool)	ctang	21	channel network base level (SAGA)	channelNB
7	profile curvature (Jenness tool)	cprof	22	catchment Area (max flux method) (SAGA)	catchslope
8	plan curvature (Jenness tool)	cplan	23	altitude above channel network (SAGA)	altichan
9	longitudinal curvature (Jenness tool)	clong	24	flow direction (ArcGIS)	flowdir
10	general curvature (Jenness tool)	cgen	25	flow drop (ArcGIS)	flowdrop
11	cross-sectional curvature (Jenness tool)	ccross	26	diffuse insolation (SAGA)	difsol
12	aspect (Jenness tool)	aspsharp	27	direct insolation (SAGA)	dirsol
13	slope (Jenness tool)	slsharp	28	flow accumulation (SAGA)	flowaccS
14	wetness index (SAGA)	WI	29	Topographic position index (SAGA)	topPosiS
15	topographic wetness index (SAGA)	TWI	30	Vector Ruggedness Measure (SAGA)	VRM

<sup>1</sup> Il DEM è stato elaborato utilizzando come dato base le ISOIPSE, i punti quotati e l'IDROGRAFIA estrapolate dalla Carta Tecnica (CT10) della Regione Lombardia. Per le operazioni di interpolazione a 10 metri (intervallo di *contour line*), è stata utilizzata la funzione di ARCFINFO - Topo to Raster - con burning della rete fluviale per perfezionare il prodotto DTM. E' stata, inoltre, eseguita un' analisi per i potenziali errori con un'analisi idrologica sul prodotto ottenuto (ricerca pit e sink), con successiva correzione e rianalisi con procedura ARCFINFO (passo precedente).

### 3.2 MATERIALI E METODI

Per calcolare le variabili geomorfometriche sono stati utilizzati i seguenti programmi open source: SAGA GIS (ver. 2.0.8), QGIS (ver. 2.4.0.), Grass (ver. 6.4.3.). E' stato utilizzato anche il software proprietario ArcGIS della ESRI, per due variabili idrologiche e come piattaforma per poter utilizzare il *tool* realizzato da Jeff Jenness, disponibile in rete (Jenness 2011). Senza proporre una trattazione rigorosa, che esula da questo lavoro, di seguito vengono brevemente presentate le variabili che sono state utilizzate. Riguardo l'argomento occorre sottolineare che in letteratura spesso non vi è corrispondenza tra le denominazioni delle variabili e le formule di calcolo, come per esempio accade per le curvature (Jenness s.d.) e le variabili *Topographic Wetness Index* e *Wetness Index* (Yang, X.; 2005; Zhao et al. 2009; Jenness s.d.; Sørensen et al. 2006; Zawawi et al. 2014). Pertanto, le variabili sono state calcolate con i *tools* implementati sui diversi software, indipendentemente dalla denominazione, e senza considerare a priori l'eventuale ridondanza di informazione che potrebbe sorgere. L'eventuale concordanza tra i risultati, dei diversi *tools*, è stata verificata con metodi statistici, come di seguito indicato. Si precisa, inoltre, che tutte le variabili sono state calcolate su un'area più estesa di quella oggetto di studio al fine di evitare ottenere valori nulli, per alcune variabili, lungo i bordi dell'area investigata.

Tipicamente nell'analisi del paesaggio sono ampiamente utilizzate una serie di variabili topografiche primarie, tra le quali si possono annoverare: la pendenza, l'esposizione e le curvature del terreno. Dette variabili geomorfometriche sono state calcolate con il *tool* messo a punto e reso disponibile, da Jeff Jenness (Jenness s.d.; Jenness 2011). Questo strumento offre tre metodi per calcolare sia la pendenza che l'esposizione, denominati: 4-Cell, Sharpnack & Akin (utilizzato in questo lavoro) e Horn (utilizzato anche nel toolbox *Spatial Analyst* di ArcGIS). I tre metodi differiscono nella formula utilizzata, ma tutti utilizzano una matrice 3x3 per calcolare i gradienti, secondo le direzioni est-ovest e nord-sud. La pendenza è l'aspetto più rilevante per i movimenti gravitativi in un territorio, infatti concorre nei calcoli di molte variabili secondarie. L'esposizione (*aspect*) è una variabile che fornisce informazioni circa l'incidenza della radiazione solare, ma anche riguardo la giacitura che i versanti hanno assunto, in conseguenza sia delle geodinamiche locali (a scala di bacino), sia degli effetti della neotettonica (su scala più ampia). La variabile *aspect* assume sempre valori positivi ad eccezione dei punti perfettamente in piano ove questa variabile assume valore -1.

Vi sono diversi algoritmi che calcolano le curvature del territorio: Jenness ha implementato i calcoli relativi a sette tipi di curvature basandosi sugli scritti di Moore (Moore et al. 1993), Wilson e Gallant (Gallant 2000), e di Porres de la Haza & Pardo Pascual (Porres de la Haza, M et al. 2002). Le formule si basano sull'approccio di Evans, il quale utilizza un polinomio con 6 parametri per approssimare le curvature del territorio a differenza di quello di Zevenbergen & Thorne (approccio utilizzato dallo *Spatial Analyst* di ArcGIS) che impiega un polinomio con 9 parametri (Jenness 2011). I due differenti approcci possono produrre valori di curvatura in conflitto (Jenness 2011).

La *Profile Curvature* descrive come una massa fluida accelera o decelera mentre scorre su un punto. E' definita come la curvatura lungo la linea di massima pendenza. La *Plan Curvature* indica se un fluido converge o diverge quando scorre su un punto ed è definita come la curvatura lungo la linea di intersezione tra la superficie ed il piano XY. La *Tangential Curvature* è definita come la curvatura lungo la linea ortogonale a quella con pendenza più ripida. Come la *Plan Curvature*,

questo valore indica se un fluido convergerà o divergerà mentre scorre su un determinato pixel. La *Longitudinal Curvature* può essere definita come la curvatura derivata dalla intersezione della superficie terrestre con un piano individuato dalla pendenza e dalla direzione dell'esposizione e anch'essa indica se un fluido che scorre sopra un punto verrà accelerato o decelerato (Jenness s.d.). La *Cross-Sectional Curvature* è concettualmente simile alla *Plan Curvature* ed è derivata dalla intersezione tra la superficie e il piano definito dalla normale alla pendenza e dalla direzione dell'esposizione (Jenness s.d.). Essa viene interpretata nello stesso senso della *Plan Curvature*: indica se un fluido che scorre su un determinato punto convergerà o divergerà, quando nel suo migrare va oltre tale punto. La *Total Curvature* e la *General Curvature* sono una descrizione generale della rugosità del paesaggio. La prima ha sempre un valore positivo e quindi non dice se il paesaggio è convesso o concavo, ma può essere usata come una misura della persistenza di un certo tipo di rugosità dello stesso. La seconda può avere valori positivi o negativi, e quindi fornisce informazioni circa la convessità o la concavità del paesaggio (Jenness s.d.). Un valore di *General Curvature* vicino allo zero indica una superficie piana oppure una zona dove la convessità in una direzione è bilanciata dalla concavità nella direzione perpendicolare, ad esempio su una sella (Jenness s.d.).

La variabile *Flow direction* (ArcGIS) e la sua associata *Flow drop* sono tipicamente utilizzate per comprendere il flusso idrologico. La prima consente di mappare la direzione del flusso idrologico determinata dalla massima pendenza (Zhao et al. 2009). Utilizzando l'opzione *Raster Drop* viene creato un raster di output che mostra il valore (in termini percentuali) del rapporto tra la massima variazione dell'elevazione, individuata tra i pixel lungo la direzione del flusso, e la lunghezza del percorso tra i centri dei pixel considerati. Lo strumento che calcola la variabile *Flow accumulation* (SAGA GIS) determina l'accumulo del flusso idrologico, riportando per ciascun pixel il peso cumulato di tutte le celle che vi confluiscono, considerando la linea di massima pendenza del versante (Olaya e Conrad 2009).

La variabile *Terrain Ruggedness Index* (TRI QGIS), fornisce indicazione sulla eterogenità del terreno (Wilson et al. 2007). La variabile *asperity* è una ulteriore misura della rugosità del terreno analoga alla *Total Curvature* (Wilson et al. 2007). La variabile *Topographic Position Index* (TPI QGIS) misura la differenza tra l'altezza del pixel centrale e l'altezza media dei pixel nel suo intorno, considerati secondo un determinato raggio (Wilson et al. 2007; Gallant 2000). Detta variabile calcolata con QGIS fornisce risultati differenti rispetto alla medesima variabile calcolata con SAGA, come viene successivamente evidenziato. In letteratura il nome *Compound Topographic Index* (GRASS) è spesso considerato sinonimo di *Wetness Index* o *Topographic Wetness Index* (Yang, X.; 2005). E' uno dei parametri più comunemente utilizzati in letteratura nelle elaborazioni, è esplicitivo per quantificare il controllo topografico sul movimento idrologico. Per il calcolo del tool di GRASS occorre fare riferimento allo scritto di Moore (Moore et al. 1991).

Per le variabili calcolate con il software SAGA GIS, di seguito brevemente richiamate, occorre fare riferimento alla letteratura citata in SAGA Module Library Documentation. La variabile *Direct radiation* (SAGA) è usata per descrivere la radiazione solare che incide perpendicolarmente sulla superficie terrestre. La variabile *Diffuse radiation* (SAGA) misura la luce solare ricevuta indirettamente a causa dello *scattering* creato dal materiale presente in atmosfera (es: polveri, nuvole, vapore acqueo nelle sue diverse forme di aggregazione). Il *Topographic Wetness Index* (TWI) è il parametro che fornisce informazioni riguardo il *run off* delle acque meteoriche, fornendo

una misura di quanto un'area tende a saturarsi (Olaya e Conrad 2009). Il *Wetness Index* (WI) di SAGA GIS è simile al precedente, ma il calcolo è basato su una modifica della variabile *catchment area* contenuta nella formula. La variabile denominata *Channel Network Base Level* (SAGA) considera la distanza verticale tra una superficie interpolata con gli impluvi che costituiscono la rete di deflusso idrico e la superficie del DEM. La variabile *Vector Ruggedness Measure* fornisce una ulteriore misura della rugosità del terreno basandosi sulla pendenza e l'esposizione. L'equazione denominata *Universal Soil Loss Equation* (USLE) è il modello più utilizzato per stimare il rischio di erosione del suolo. Tra i parametri utilizzati nell'equazione vi è l'*LS factor*, (Moore et al. 1991) che descrive l'effetto della topografia sull'erosione del suolo. L è il fattore che tiene conto della lunghezza del versante mentre S rappresenta la pendenza. L'algoritmo di calcolo del tool di SAGA è basato sulla pendenza e su *specific catchment area* come sostituto della lunghezza del versante. *Catchment area (recursive method)* (SAGA), *Catchment area (max flux method)* (SAGA), *convergence index* (SAGA), sono variabili che forniscono informazioni riguardo il flusso idrologico superficiale (Gruber e Peckham 2009). La variabile *Mass Balance Index* (SAGA) descrive le potenziali zone di accumulo di materiale colluviale e di deplezione del suolo: assume valori negativi per le aree di deposizione, valori positivi per quelle di erosione e valori prossimi allo zero per le aree dove vi è un equilibrio tra erosione e deposizione, come per le zone di pianura (Möller et al. 2008).

In prospettiva dell'applicazione di metodi statistici per la modellazione, sono state effettuate alcune analisi statistiche di base sui predittori, al fine di evidenziare particolari tendenze. Infatti, occorre tenere presente che molti metodi statistici sono basati su particolari assunzioni relativamente alle variabili sia dipendenti che indipendenti.

Per ciascuna variabile sono stati evidenziati gli *outliers*, è stata verificata la normalità, con il test di normalità di Shapiro-Wilk, e rappresentata la distribuzione di frequenza, utilizzando i *normal probability plots* e gli istogrammi di frequenza. E' stato determinato, inoltre, l'indice di correlazione di Pearson ed effettuato il test di Belsley per diagnosticare l'eventuale collinearità tra le variabili.

### 3.3 ANALISI E RISULTATI

Sui dati grezzi sono stati calcolati gli indici di dispersione, gli indici di tendenza centrale e gli indici relativi alla forma della distribuzione (v. Tab. 3.2 A e B). Da questi indici emerge che variabili con lo stesso nome (ad es.: *Topographic Position Index*) sono state ottenute con algoritmi differenti. Nonostante diverse variabili siano accomunate da un valor medio prossimo allo zero, ciascuna di esse è caratterizzata da un proprio ordine di grandezza e unità di misura, derivante dalla formula di calcolo, che ne rende impossibile il confronto.

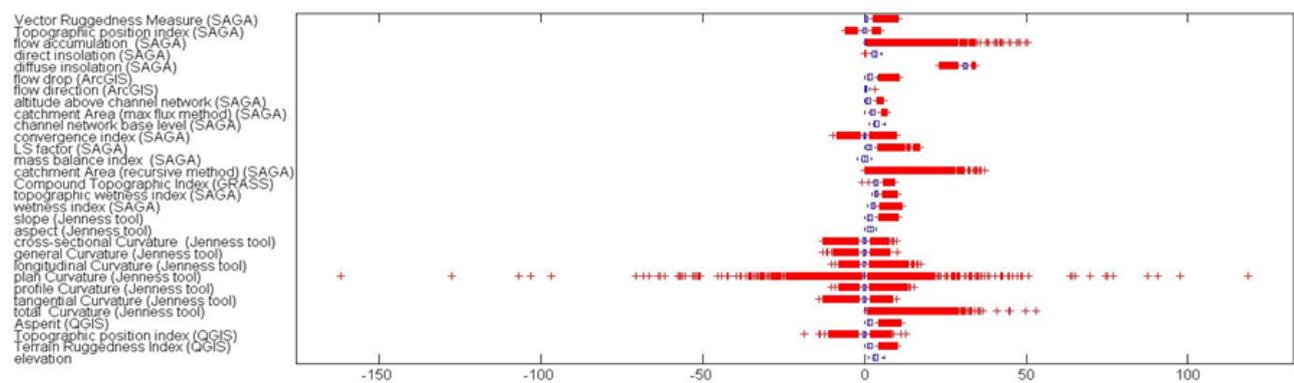
**Tabella 3.2 A** Statistica descrittiva per le variabili geomorfometriche.

	elevation	TRI	TPI	aspefit	ctot	ctang	cprof	cplan	clong	cgen	ccross	aspect (sharp)	slope (sharp)	WI	TWI
Arithmetic average	501,3982	2,599281	0,003213	8,582151	0,015375	0,024313	0,019284	0,0055928	0,027152	0,017123	0,044275	188,3544	33,36844	5,612363	8,734357
Harmonic mean	460,5685	0	0	0	0	0	0	0	0	0	0	20,82103	0	5,148182	8,302752
50 th percentile	478,246	2,371361	0,016479	7,789032	0,005683	0,037796	0,005355	0,1491501	-0,00095	0,081014	0,072055	192,6693	30,31617	5,105622	8,014681
Most frequent value	344,195	0	-3,05E-05	0	0	0	0	0	0	0	0	0	-1	4,016121	22,54347
Interquartile range	213	1,989632	0,442627	6,545929	0,01289	0,642433	0,585988	2,3684032	1,196591	2,330505	1,305976	174,6443	26,07785	1,899335	2,150882
Mean Absolute deviation	118,0301	1,16005	0,327641	3,851638	0,016141	0,472509	0,462586	2,0937093	0,942175	1,699673	0,961385	88,85573	15,18568	1,36529	1,641476
Range	686,008	14,80039	14,92239	55,65854	1,922715	16,83779	18,52191	1537,4078	40,42817	56,4966	32,30311	360,9994	199,0679	21,15416	17,71592
Standard deviation	143,9334	1,458754	0,480162	4,868025	0,036304	0,703586	0,721341	5,4868093	1,458954	2,457798	1,422466	102,8497	19,0691	1,985585	2,281998
Variance	20716,82	2,127963	0,230556	23,69766	0,001318	0,495034	0,520333	30,105076	2,128548	6,040773	2,023411	10578,07	363,6305	3,942546	5,207517
Kurtosis	2,416	4,531	9,009	4,991	173,332	11,645	14,222	2559,499	14,035	7,888	10,957	1,839	4,511	12,413	7,068
Skewness	0,420	0,932	-1,018	1,002	9,687	-1,159	1,450	-9,790	1,380	-0,895	-1,057	-0,128	0,927	2,472	1,914
max	879,794	14,80039	6,064392	55,65854	1,922715	6,944271	11,07682	650,16412	25,28983	25,05636	14,24618	359,9994	199,0679	23,21426	22,88869
min	193,786	0	-8,858	0	0	-9,89352	-7,44508	-887,24371	-15,1383	-31,4402	-18,0569	-1	0	2,060095	5,17277

**Tabella 3.2 B**

	CTIgrass	catarea	massbal	LSfactor	ConvInd	chanelNB	catchslope	altichan	flowdir	flowdrop	difsol	dirsol	flowaccs	topPosiS	VRM
Arithmetic average	5,759142	52303,39	0,034045	4,899395	0,062869	455,7549	0,304554	45,643303	30,11405	31,03637	0,86024	3,241453	4799,689	0,099218	0,032852
Harmonic mean	5,618791	737,2892	0	1,048707	0	422,0597	0	0	4,2735	18,18554	0,85933	0	1308,278	-6,33501	0
Arithmetic average	5,759142	52303,39	0,034045	4,899395	0,062869	455,7549	0,304554	45,643303	30,11405	31,03637	0,86024	3,241453	4799,689	0,099218	0,032852
50 th percentile	5,46855	994,0889	0,047019	4,372883	0,425504	428,6462	0,296741	37,12146	16	26,66667	0,861189	3,35951	1867,976	-0,07845	0,017003
Most frequent value	-1	100	0	4,190271	0,342256	324,1869	0	0	16	20	0,863735	0	100	-9,15437	0
Interquartile range	1,815748	2343,596	0,74763	4,199463	7,203924	191,1338	0,168148	56,795441	28	26,66667	0,030367	1,457721	2804,987	7,59759	0,040113
Mean Absolute deviation	1,172639	91514,48	0,3779	2,453087	6,03376	103,8633	0,097383	31,758136	29,70019	15,14973	0,020464	0,855668	4870,806	5,086321	0,029574
Range	15,49426	12809100	1,904356	53,04394	193,5551	607,7504	0,848732	229,60217	127	203,0833	0,305825	5,714776	584375,1	74,30879	0,426084
Standard deviation	1,564872	344031,6	0,435668	3,103675	9,888425	126,0901	0,120629	39,021947	39,14984	18,98268	0,02768	1,05586	11583,86	6,81098	0,041048
Variance	2,448824	1,18E+11	0,189807	9,632798	97,78094	15898,71	0,014551	1522,7124	1532,71	360,3423	0,000766	1,11484	1,34E+08	46,38945	0,001685
Kurtosis	5,799	231,257	1,874	6,367	12,838	2,443	2,852	3,512	4,353	4,552	5,219	2,818	208,482	4,199	9,810
Skewness	0,769	12,988	-0,107	1,130	-0,559	0,480	0,314	0,944	1,592	0,961	-0,546	-0,454	10,390	-0,197	2,258
max	14,49426	12809200	0,976558	53,04409	98,68102	801,4783	0,848732	229,60217	128	203,3333	0,946685	5,714776	584475,1	34,73267	0,426084
min	-1	100	-0,9278	0,000151	-94,87411	193,7278	0	0	1	0,25	0,64086	0	100	-39,5761	-1,51E-09

Per svincolare le variabili dalla loro unità di misura è stata effettuata la standardizzazione senza procedere alla centratura dei valori. Quest'ultima è stata evitata in quanto distorcerebbe l'effetto del segno, che conferisce un preciso significato fisico ai valori assunti dalle variabili.



**Fig. 3.1** Box plot relativo alle variabili standard.

Il box plot relativo alle variabili standard (Fig. 3.1) fornisce poche informazioni, in quanto l'ordine di grandezza dei valori estremi della variabile *Plan Curvature* compromette l'osservazione dei parametri statistici di base (media, deviazione standard, ecc). Tuttavia il grafico permette di osservare l'andamento generale delle variabili, evidenziandone in particolare la simmetria e la variabilità rispetto allo zero, nonché la presenza di outliers.

**Tabella 3.3** Numero di valori negativi e positivi per le variabili che presentano numeri con entrambi i segni.

	ctang	cprof	cplan	clong	cgen	ccross	aspsharp	CTIgrass	massbalan	convergind	topPosiS
valori positivi	493677	464288	493677	457962	479756	491061	917051	913142	483444	491453	453037
valori negativi	423374	452763	423374	459089	437286	425990	106	4015	433689	425700	464120

Particolare interesse è stato rivolto agli outliers. A riguardo, occorre considerare che la discretizzazione del territorio con un modello digitale del terreno (DEM), indipendentemente dalla misura del pixel, comporta una approssimazione della realtà, che dipende dalla variabilità o conformazione del rilievo. Per questo è necessario comprendere se gli outliers, generalmente esclusi dalle elaborazioni statistiche, in quanto potenzialmente dannosi per le predizioni dei modelli, siano dovuti alla mancata accuratezza del DEM o alla sua inadeguatezza. Quest'ultima è una caratteristica del DEM legata alla sua discretizzazione (dimensione del pixel) che lo rende non idoneo a fornire una rappresentazione realistica della superficie terrestre. L'inaccuratezza è un difetto che esula dalla dimensione del pixel, è legata al processo di modellazione del DEM ed è causata di errori indotti (es: quote errate, numero di punti quotati insufficiente o mal distribuito) che creano discontinuità eccezionali. Queste potenzialmente possono determinare risultati anomali nel calcolo delle variabili

geomorfometriche. In questi casi sarebbe opportuno procedere ad una correzione dei dati o addirittura ad una loro esclusione dalle elaborazioni, utilizzate per costruire i modelli. A tal riguardo, vengono riportate due tabelle ove sono riportati, per ciascuna variabile, il numero di *outliers* (individuati quando il valore supera la media, dei valori assoluti, di 2 (v. Tab. 3.4 A e B) e di 3 volte la deviazione standard (v. Tab. 3.5 A e B)). Nel caso di variabili con valori anche negativi (v. Tab. 3.3), per queste determinazioni è stato scelto di utilizzare il valore assoluto considerando trascurabile l'eventuale influenza dell'asimmetria della distribuzione di frequenza. Tale approssimazione è stata valutata considerando che dette variabili hanno un numero di valori ed il loro intervallo in quote paragonabili rispetto allo zero (v. Tab. 3.3). Mentre per le variabili asimmetriche, *aspect* e *Compound Topographic Index*, il numero dei valori con segno negativo e la loro dimensione è trascurabile al fine di determinarne la media. Dai calcoli emerge che quasi tutte le variabili presentano *outliers*, con porzioni che vanno da 0,42 ad un massimo del 5,86% dei dati, nel caso di valori che superano la media di oltre due volte la deviazione standard, e con un *range* compreso tra 0,05 a 2,52% nel caso di quelli che superano la media di oltre di tre volte la deviazione standard. Dalle prime considerazioni su queste analisi emerge la singolarità della variabile *Plan Curvature*, la quale ha un particolare intervallo dei dati, dovuto alla presenza di dati che possono essere definiti "aberranti" (v. Fig. 3.1). Infatti questa variabile, a fronte di una media aritmetica prossima allo zero (0,00559) e di una deviazione standard pari a 5,487, presenta un minimo e un massimo rispettivamente di -887,24 e 650,16. La seconda evidenza riguarda le variabili che non presentano outliers: la variabile *aspect* (v. Fig. 3.2 A), nel caso di *range* determinato dalla media sommata a 2 volte la deviazione standard, e le variabili *elevation* (v. Fig. 3.26 B), *Mass Balance Index* (SAGA) (v. Fig. 3.27 A), *Channel Network Base Level* (SAGA) (v. Fig. 3.27 B), *flow direction* (ArcGIS) (v. Fig. 3.28 A), *direct insolation* (SAGA) (v. Fig. 3.28 B), per il caso di *range* determinato dalla media sommata a 3 volte la deviazione standard. Per tutte le variabili, i valori superiori alla media sommata a 3 volte la deviazione standard sono stati rappresentati in una mappa, per agevolarne la comprensione ed individuare la causa della loro presenza (v. Fig. dalla 3.2 alla 3.29).

**Tabella 3.4 A** Numero di valori maggiori di: (media+2 deviazione standard).

outlier	elevation	tri	TPI	asperit	ctot	ctang	cprof	cplan	clong	cgen	ccross	aspsharp	slsharp	wetindex	topowetind
>mean+2std	31210	34860	40806	34854	24301	38779	38101	13544	38531	41813	39456	0	34616	38200	53707
CTIgrass	catarea	massbalan	LSfactor	convergind	channelNB	catchslope	altichan	flowdir	flowdrop	difsol	dirsol	flowaccS	topPosiS	vectorRS	
	42122	17129	11743	34754	40989	31853	27744	40293	98045	38124	21892	3882	26009	46113	49416

**Tabella 3.4 B**

outlier%	elevation	tri	TPI	asperit	ctot	ctang	cprof	cplan	clong	cgen	ccross	aspsharp	slsharp	wetindex	topowetind
>mean+2std	3,40	3,80	4,45	3,80	2,65	4,23	4,15	1,48	4,20	4,56	4,30	0,00	3,77	4,16	5,86
CTIgrass	catarea	massbalan	LSfactor	convergind	channelNB	catchslope	altichan	flowdir	flowdrop	difsol	dirsol	flowaccS	topPosiS	vectorRS	
	4,59	1,87	1,28	3,79	4,47	3,47	3,02	4,39	10,69	4,16	2,39	0,42	2,84	5,03	5,39

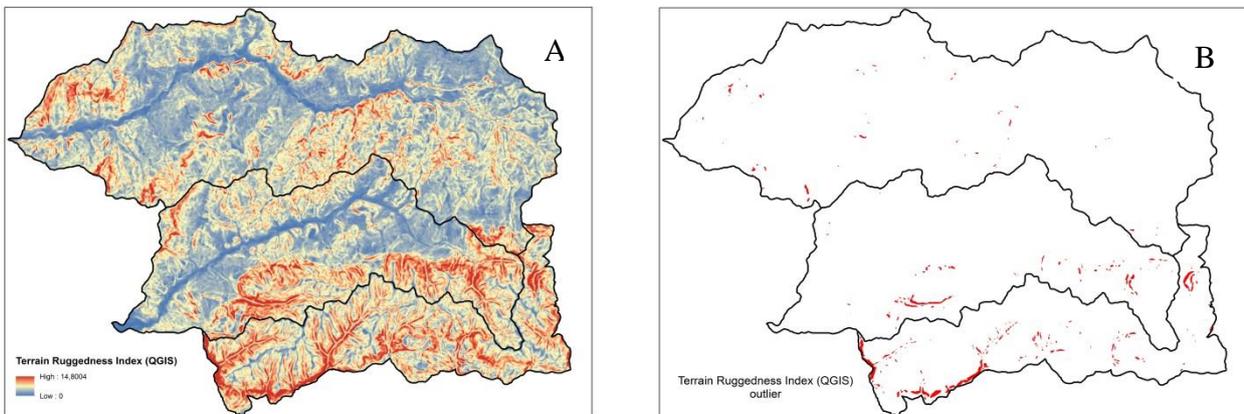
**Tabella 3.5 A** Numero di valori maggiori di: (media+3 deviazione standard).

outlier	elevation	tri	TPI	asperit	ctot	ctang	cprof	cplan	clong	cgen	ccross	aspsharp	slsharp	wetindex	topowetind
>mean+3std	0	7050	17122	7458	14287	16868	17633	7160	17413	17158	16829	0	6892	18773	23120
CTIgrass	catarea	massbalan	LSfactor	convergind	channelNB	catchslope	altichan	flowdir	flowdrop	difsol	dirsol	flowaccS	topPosiS	vectorRS	
	11816	11950	0	7589	23602	0	2079	7011	0	7053	495	0	15122	13724	19710

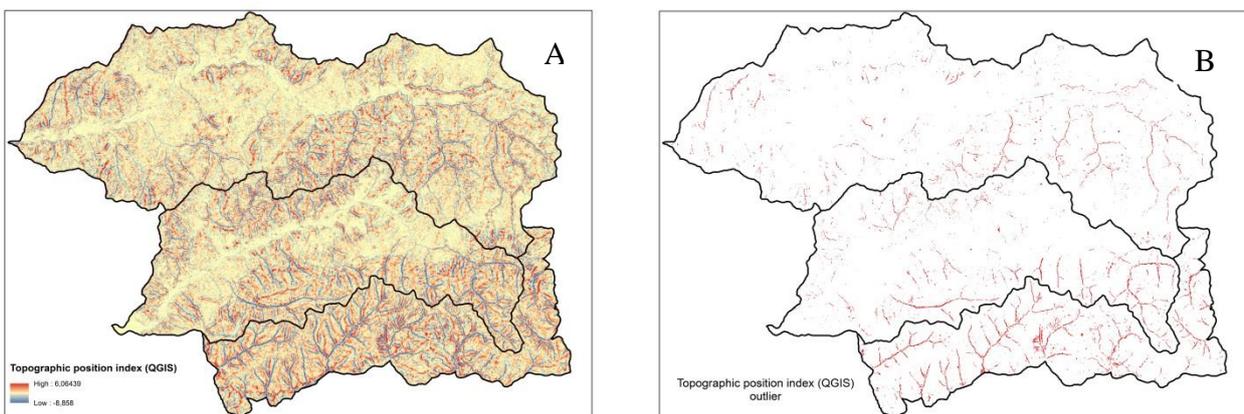
**Tabella 3.5 B**

outlier%	elevation	tri	TPI	asperit	ctot	ctang	cprof	cplan	clong	cgen	ccross	aspsharp	slsharp	wetindex	topowetind
>mean+3std	0,00	0,77	1,87	0,81	1,56	1,84	1,92	0,78	1,90	1,87	1,83	0,00	0,75	2,05	2,52
	CTIgrass	catarea	massbalan	LSfactor	convergind	channelNB	catchslope	altichan	flowdir	flowdrop	difsol	dirsol	flowaccS	topPosIS	vecterRS
	1,29	1,30	0,00	0,83	2,57	0,00	0,23	0,76	0,00	0,77	0,05	0,00	1,65	1,50	2,15

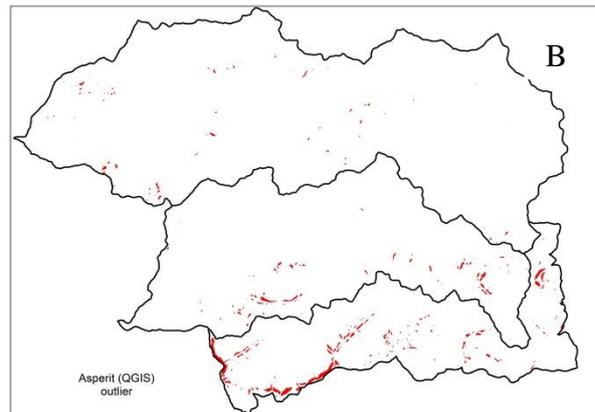
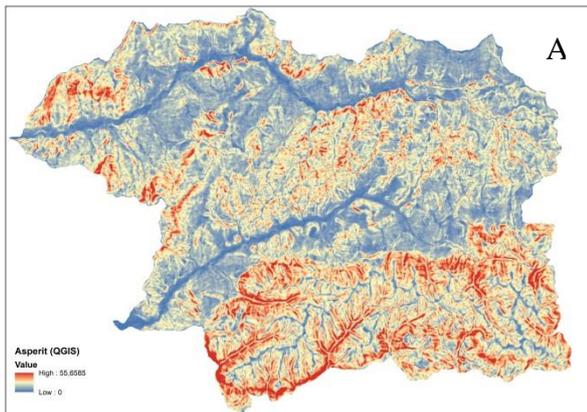
Dai risultati emerge che per una parte delle variabili i punti relativi agli *outliers* appartengono ad aree ove sono presenti pendenze particolarmente elevate o discontinuità del pendio, per altre corrispondono alle zone di impluvio. Per la variabile *Altitude above channel network*, i valori estremi sono relativi alle zone sommitali, dei monti che presentano le maggiori altitudini: Monte Calcinera, Monte dell'Orso, Monte Valle Grande, Monte Lungo, Poggio del Re, Monte Farnago, Monte Guardia, Monte Tre Termini, Castello Verde.



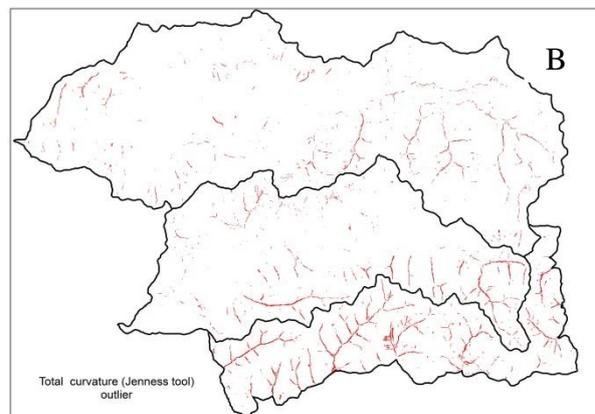
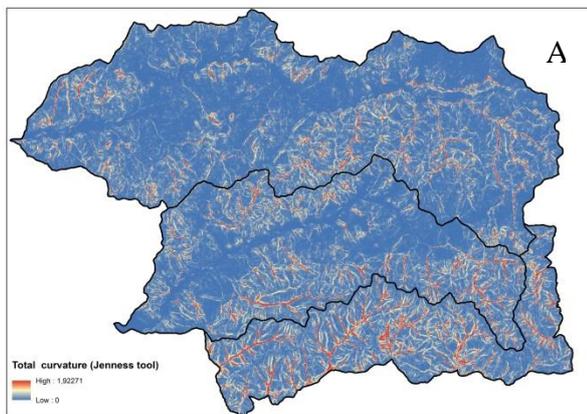
**Fig. 3.2** Rappresentazione della variabile *Terrain Ruggedness Index* (QGIS) (A) e degli *outliers* con valore superiore a (media+3 deviazione standard) (B).



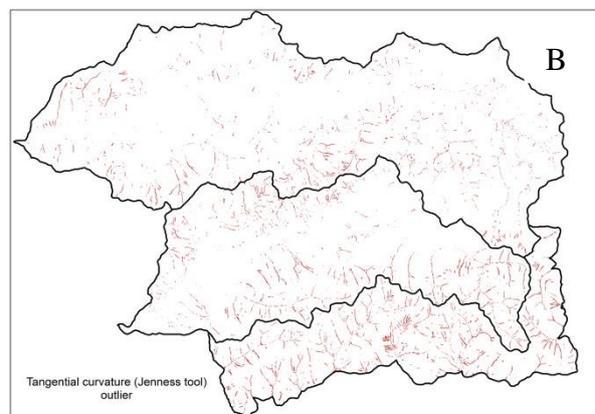
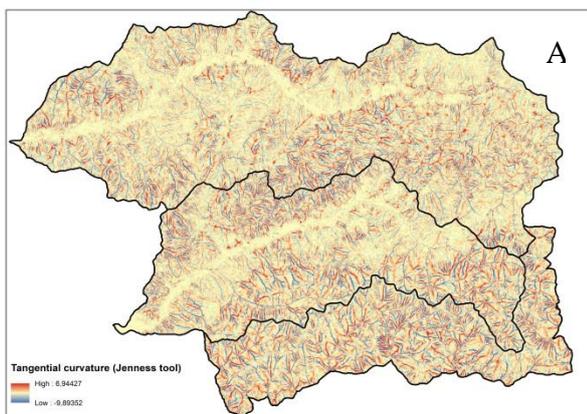
**Fig. 3.3** Rappresentazione della variabile *Topographic position index* (QGIS) (A) e degli *outliers* con valore superiore a (media+3 deviazione standard) (B).



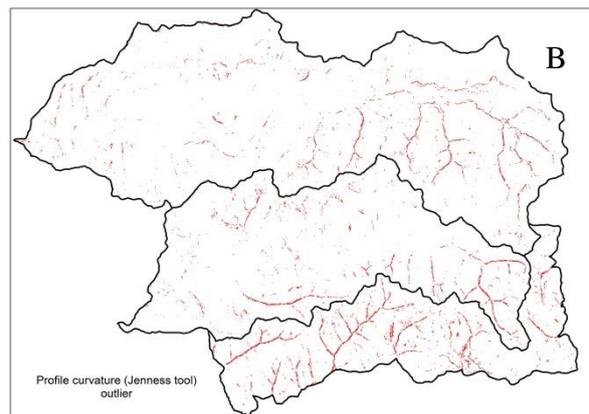
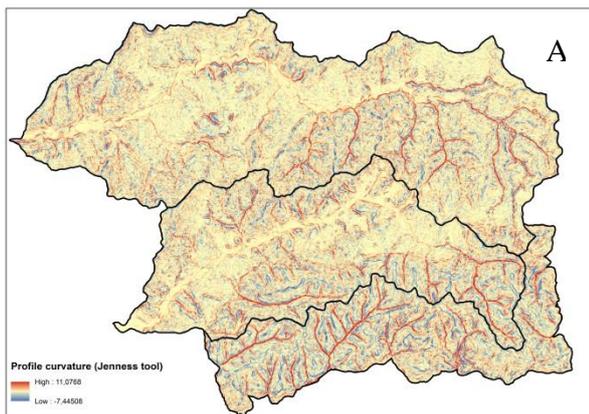
**Fig. 3.4** Rappresentazione della variabile *Asperit* (QGIS) (A) e degli *outliers* con valore superiore a (media+3 deviazione standard) (B).



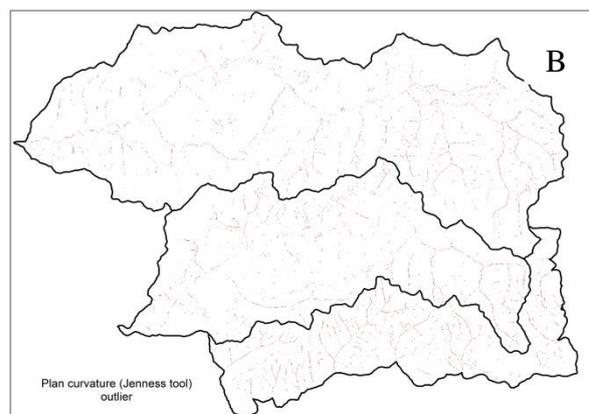
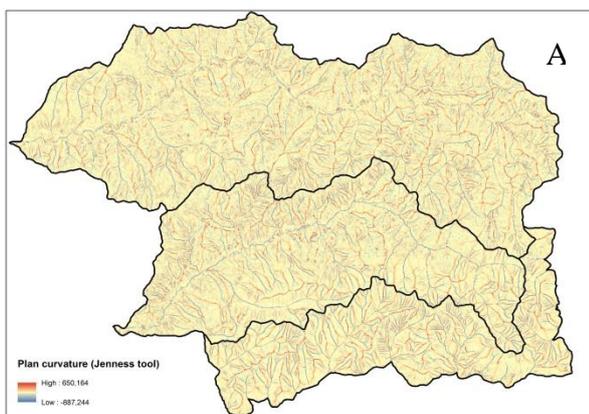
**Fig. 3.5** Rappresentazione della variabile *total curvature* (Jenness tool) (A) e degli *outliers* con valore superiore a (media+3 deviazione standard) (B).



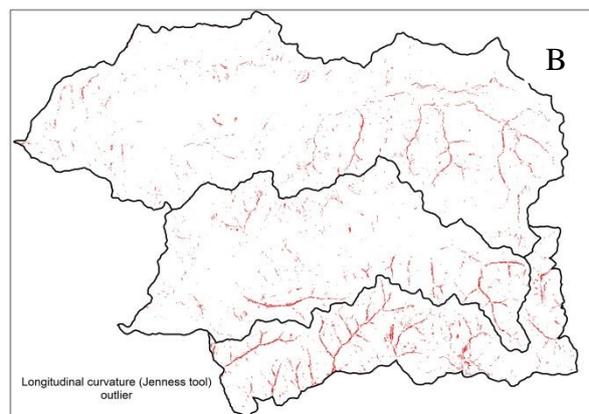
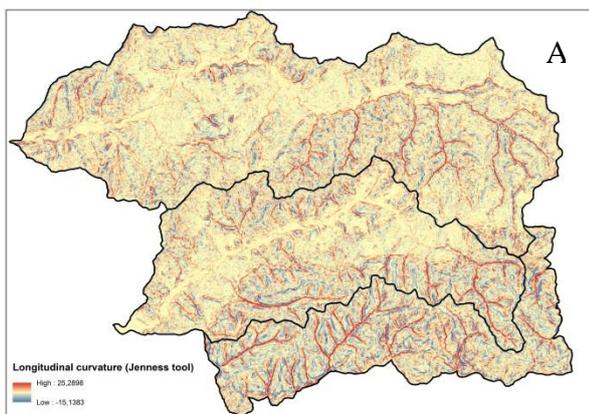
**Fig. 3.6** Rappresentazione della variabile *tangential curvature* (Jenness tool) (A) e degli *outliers* con valore superiore a (media+3 deviazione standard) (B).



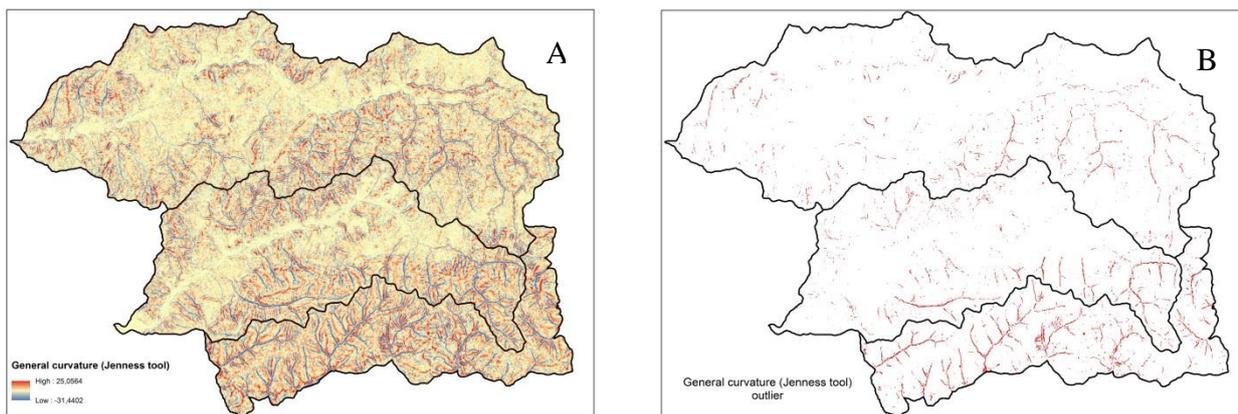
**Fig. 3.7** Rappresentazione della variabile *profile curvature* (Jenness tool) (A) e degli *outliers* con valore superiore a (media+3 deviazione standard) (B).



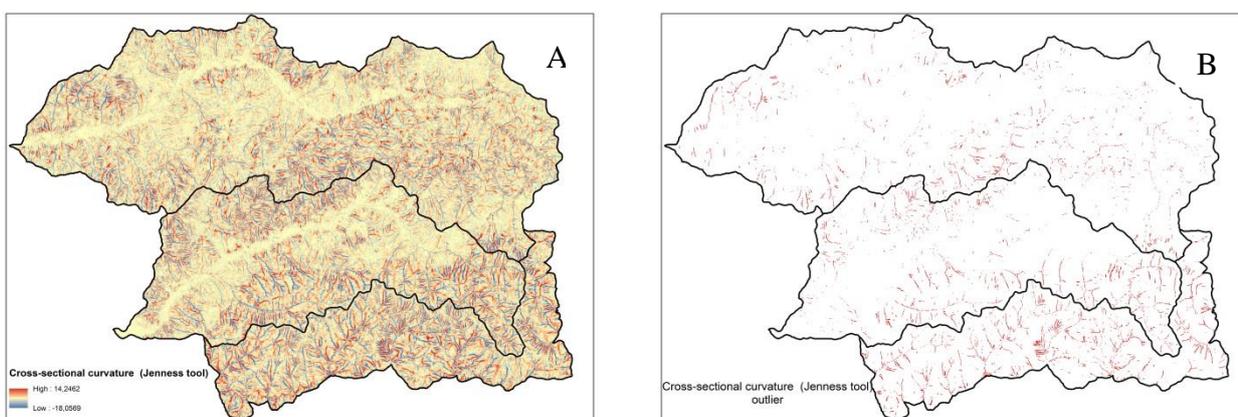
**Fig. 3.8** Rappresentazione della variabile *plan curvature* (Jenness tool) (A) e degli *outliers* con valore superiore a (media+3 deviazione standard) (B).



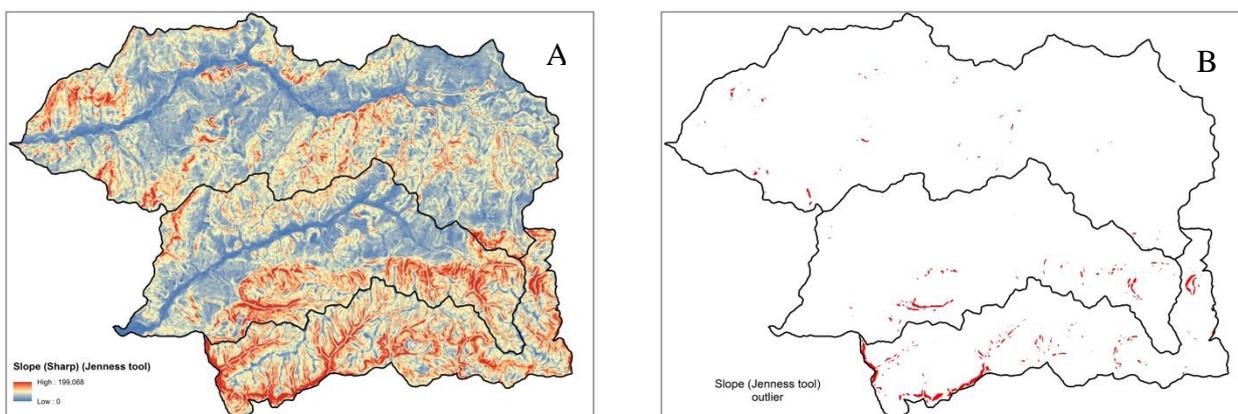
**Fig. 3.9** Rappresentazione della variabile *longitudinal curvature* (Jenness tool) (A) e degli *outliers* con valore superiore a (media+3 deviazione standard) (B).



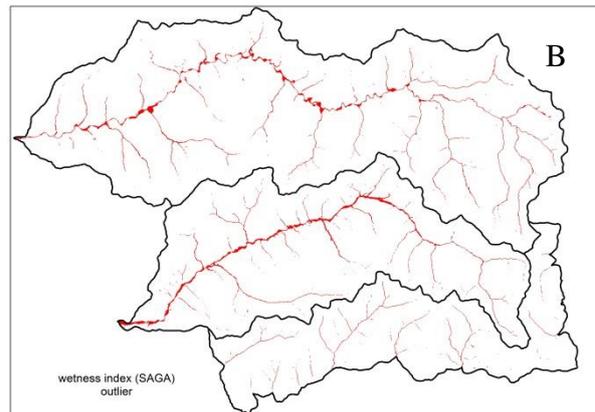
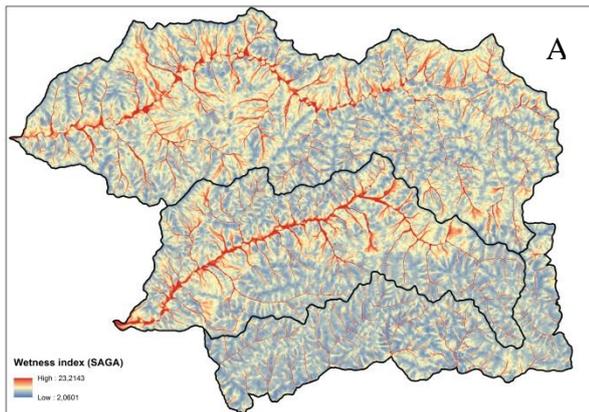
**Fig. 3.10** Rappresentazione della variabile *general curvature (Jenness tool)* (A) e degli *outliers* con valore superiore a (media+3 deviazione standard) (B).



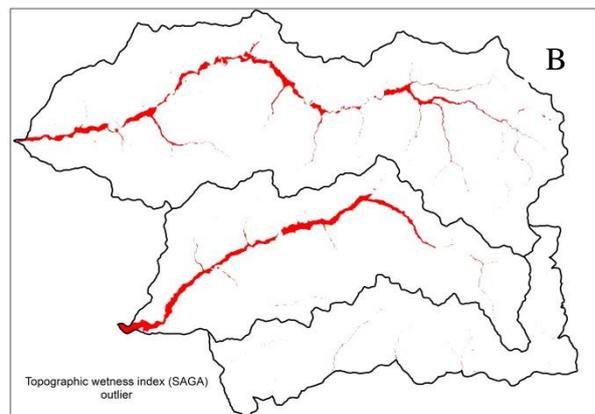
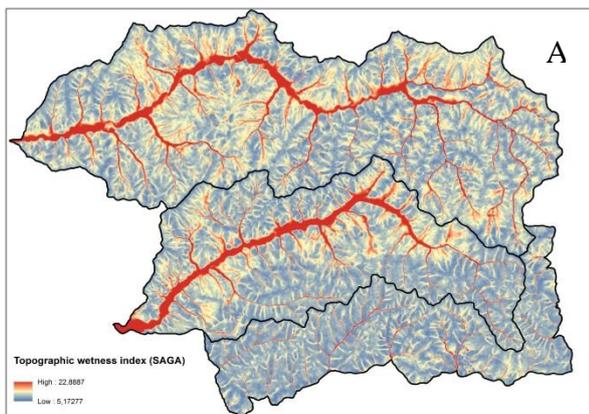
**Fig. 3.11** Rappresentazione della variabile *cross-sectional curvature (Jenness tool)* (A) e degli *outliers* con valore superiore a (media+3 deviazione standard) (B).



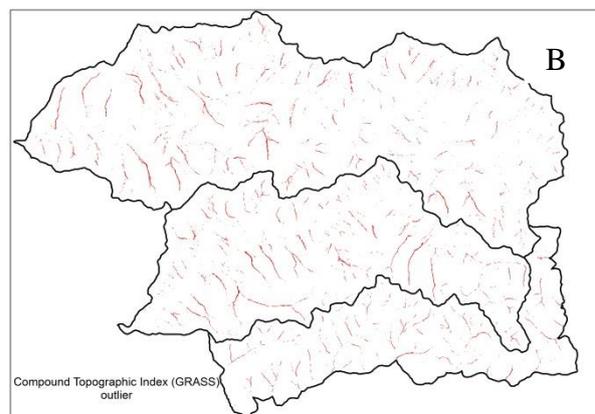
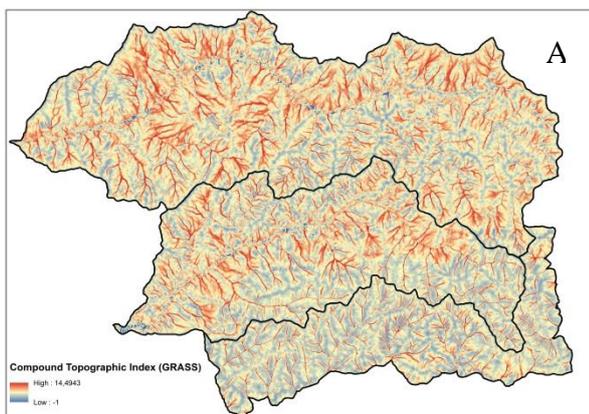
**Fig. 3.12** Rappresentazione della variabile *slope (Jenness tool)* (A) e degli *outliers* con valore superiore a (media+3 deviazione standard) (B).



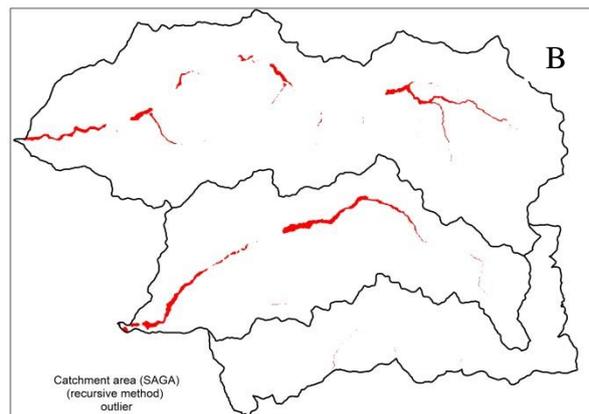
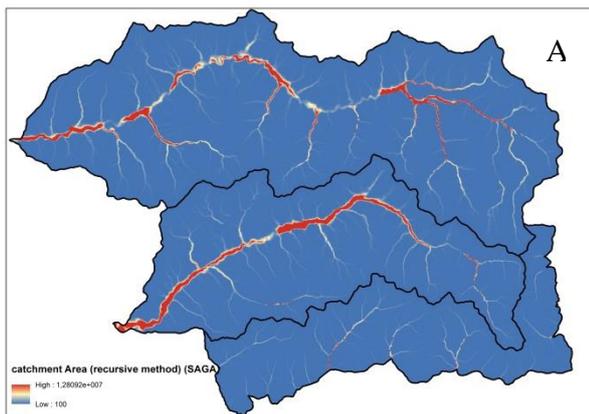
**Fig. 3.13** Rappresentazione della variabile *wetness index* (SAGA) (A) e degli *outliers* con valore superiore a (media+3 deviazione standard) (B).



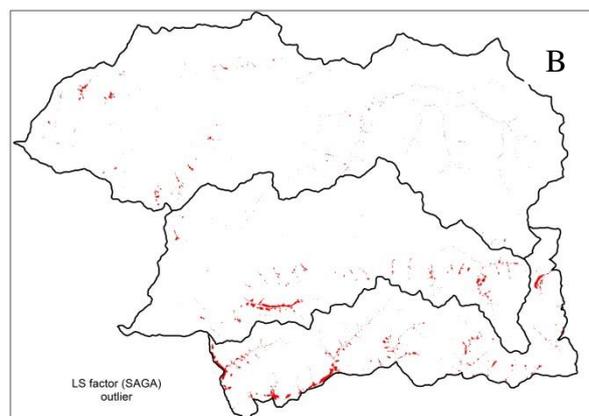
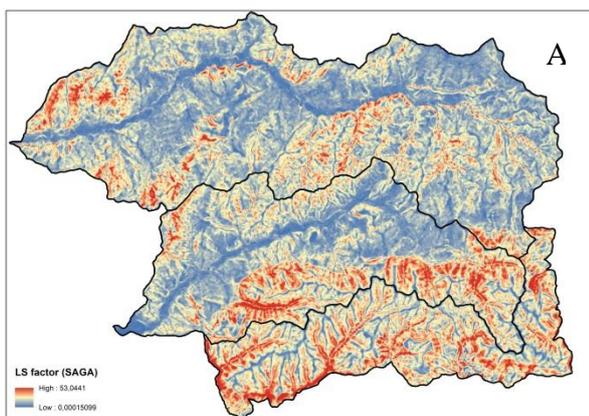
**Fig. 3.14** Rappresentazione della variabile *topographic wetness index* (SAGA) (A) e degli *outliers* con valore superiore a (media+3 deviazione standard) (B).



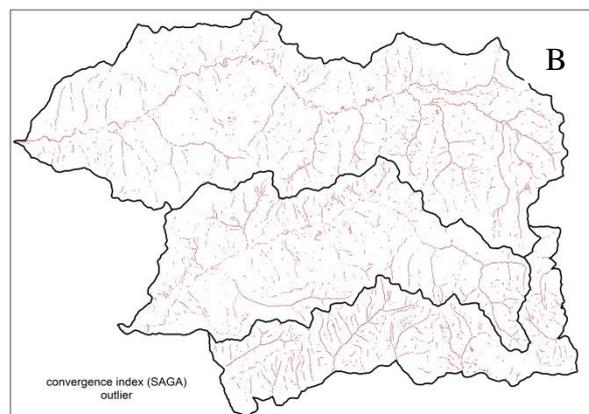
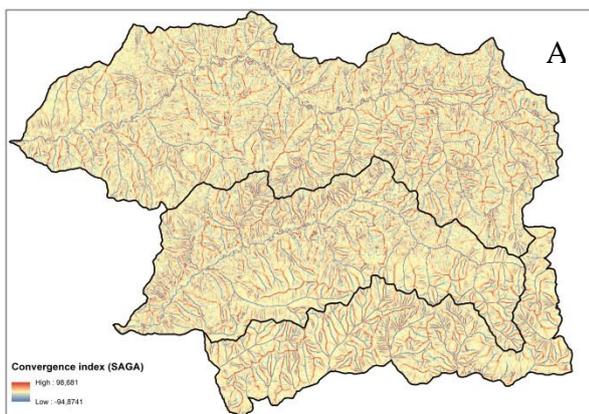
**Fig. 3.15** Rappresentazione della variabile *Compound Topographic Index* (GRASS) (A) e degli *outliers* con valore superiore a (media+3 deviazione standard) (B).



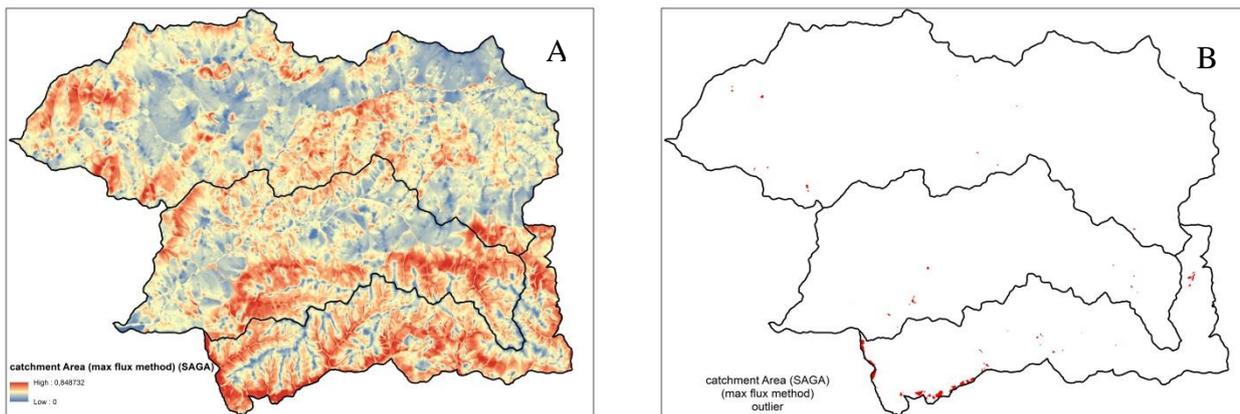
**Fig. 3.16** Rappresentazione della variabile *catchment Area (recursive method) (SAGA)* (A) e degli *outliers* con valore superiore a (media+3 deviazione standard) (B).



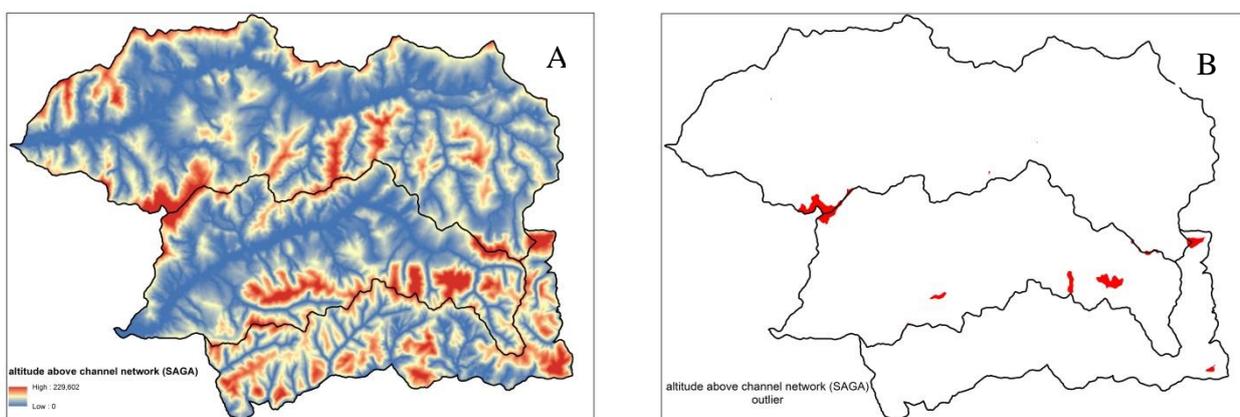
**Fig. 3.17** Rappresentazione della variabile *LS factor (SAGA)* (A) e degli *outliers* con valore superiore a (media+3 deviazione standard) (B).



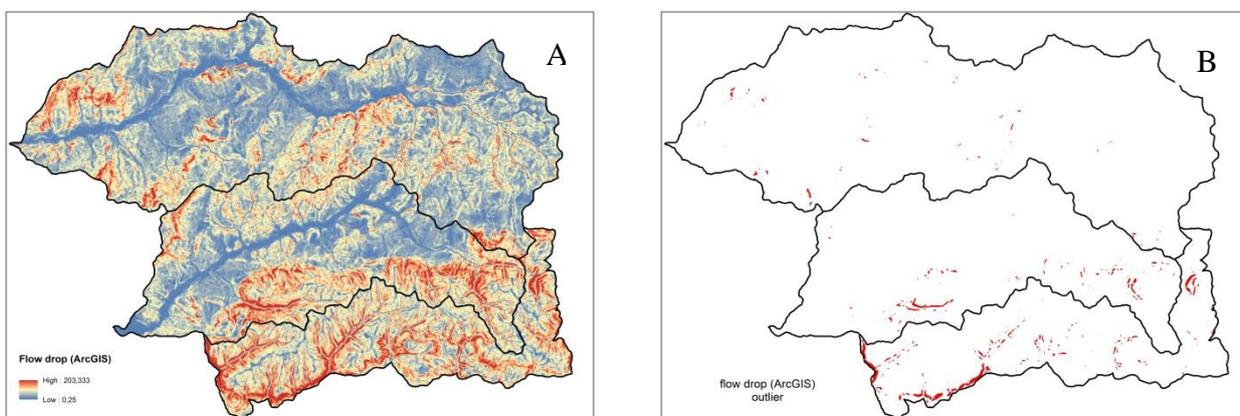
**Fig. 3.18** Rappresentazione della variabile *convergence index (SAGA)* (A) e degli *outliers* con valore superiore a (media+3 deviazione standard) (B).



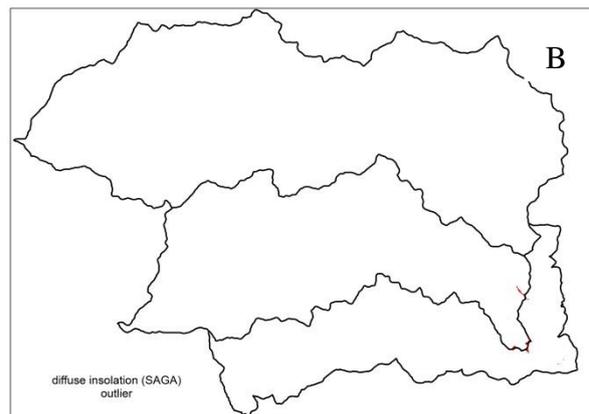
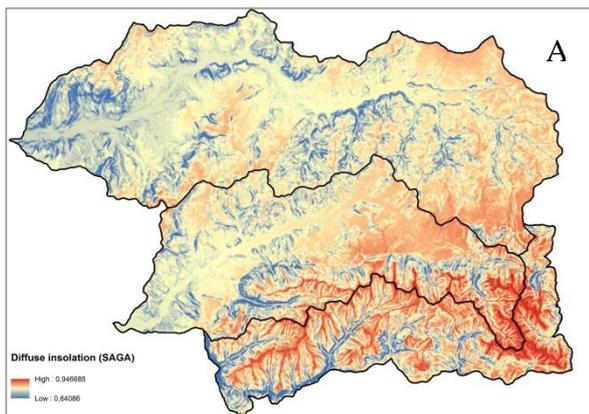
**Fig. 3.19** Rappresentazione della variabile *catchment Area (max flux method) (SAGA)* (A) e degli *outliers* con valore superiore a (media+3 deviazione standard) (B).



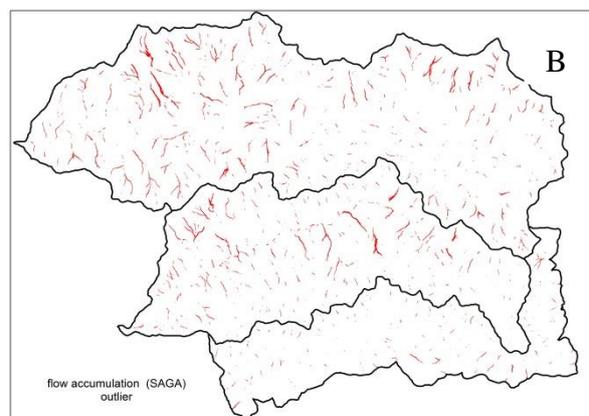
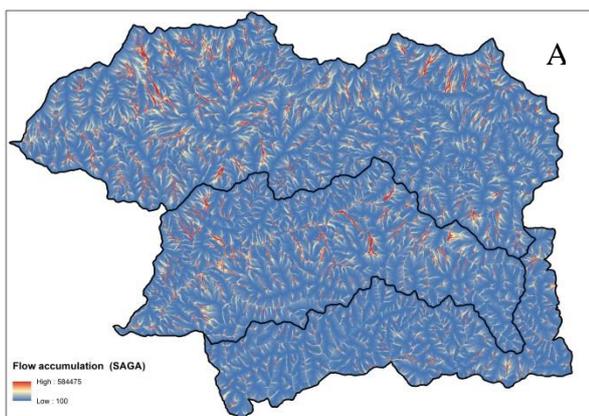
**Fig. 3.20** Rappresentazione della variabile *altitude above channel network (SAGA)* (A) e degli *outliers* con valore superiore a (media+3 deviazione standard) (B).



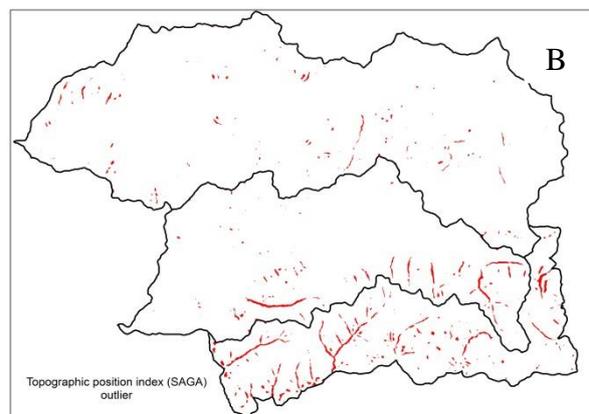
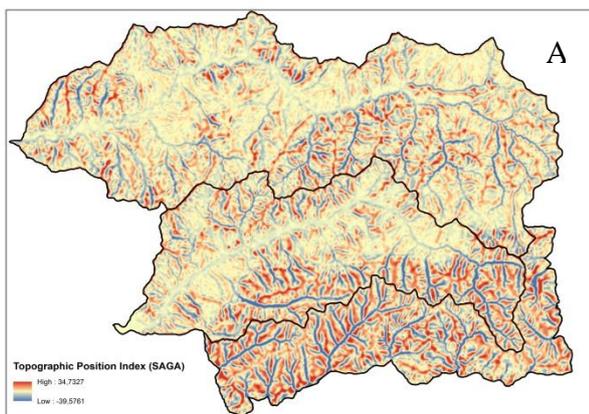
**Fig. 3.21** Rappresentazione della variabile *flow drop (ArcGIS)* (A) e degli *outliers* con valore superiore a (media+3 deviazione standard) (B).



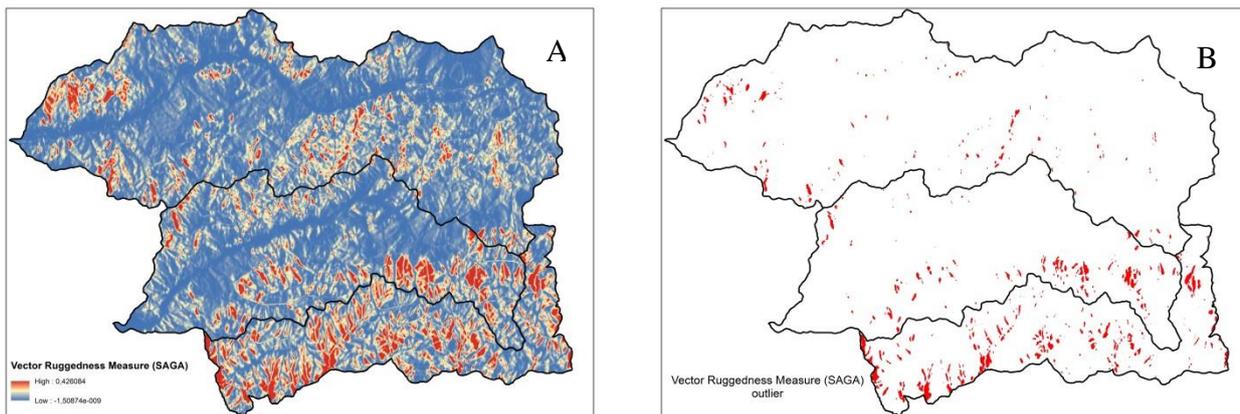
**Fig. 3.22** Rappresentazione della variabile *diffuse insolation* (SAGA) (A) e degli *outliers* con valore superiore a (media+3 deviazione standard) (B).



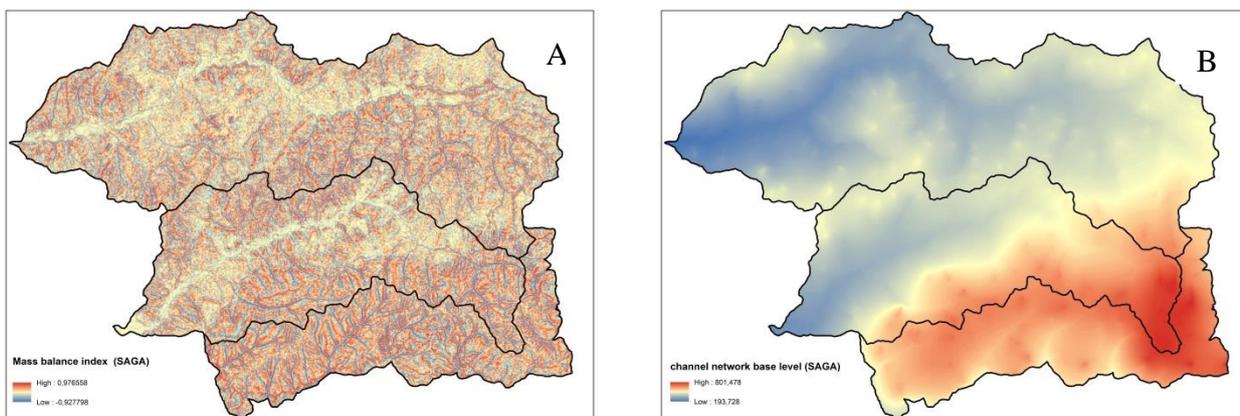
**Fig. 3.23** Rappresentazione della variabile *flow accumulation* (SAGA) (A) e degli *outliers* con valore superiore a (media+3 deviazione standard) (B).



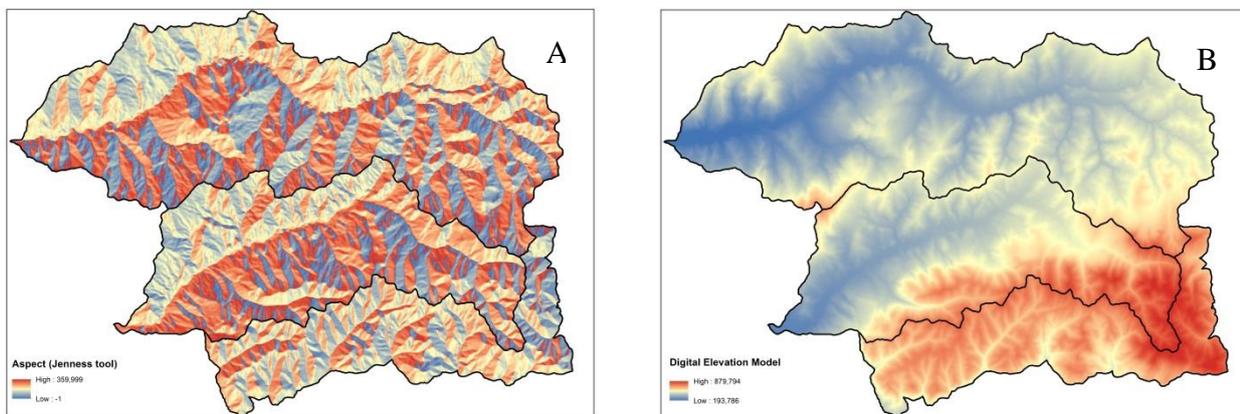
**Fig. 3.24** Rappresentazione della variabile *Topographic position index* (SAGA) (A) e degli *outliers* con valore superiore a (media+3 deviazione standard) (B).



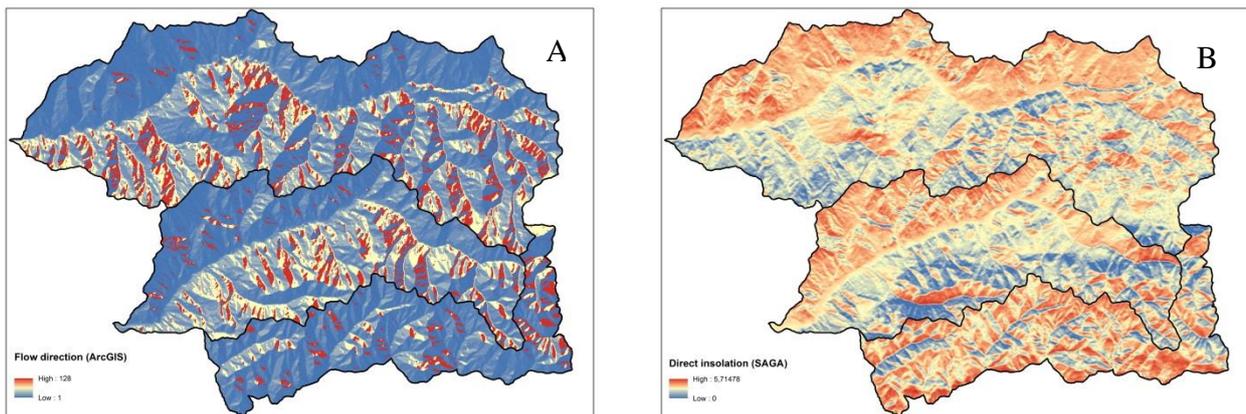
**Fig. 3.25** Rappresentazione della variabile *Vector Ruggedness Measure (SAGA)* (A) e degli *outliers* con valore superiore a (media+3 deviazione standard) (B).



**Fig. 3.26** Rappresentazione della variabile *mass balance index (SAGA)* (A) e della variabile *channel network base level (SAGA)* (B).

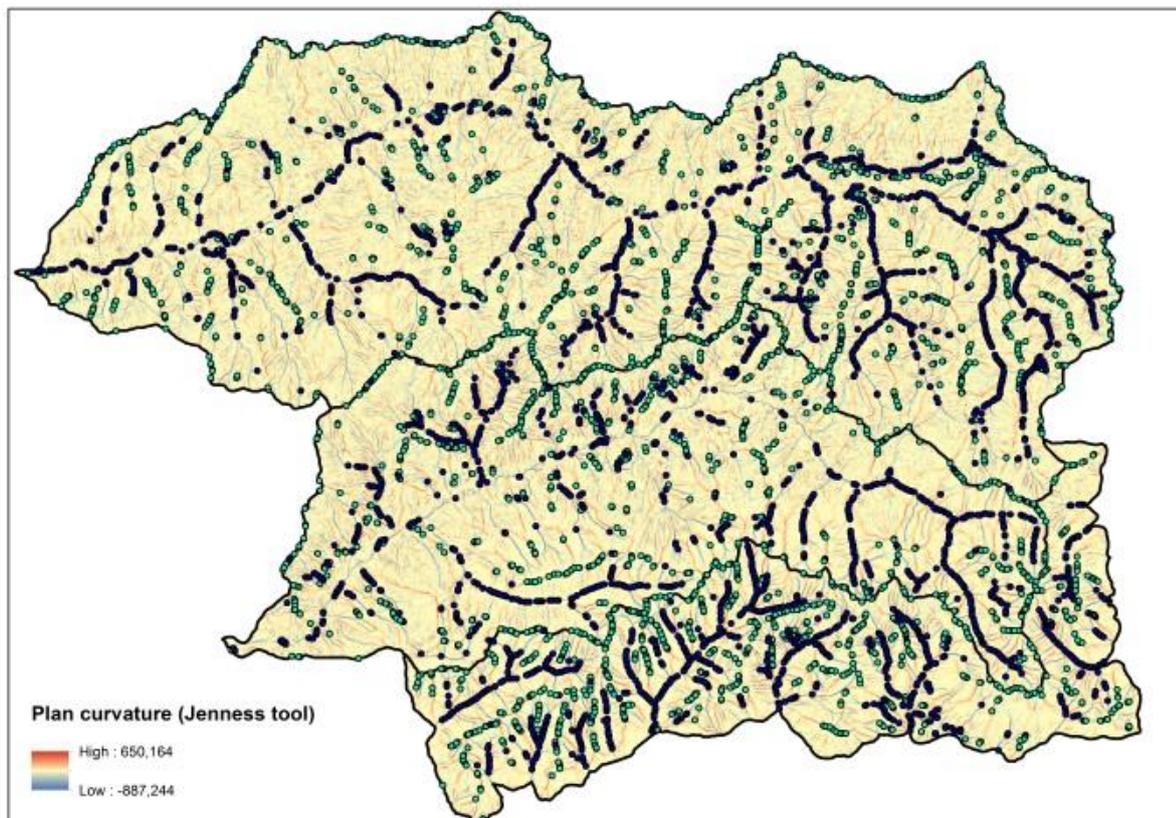


**Fig. 3.27** Rappresentazione della variabile *aspect (Jenness tool)* (A) e della variabile *DEM* (B).

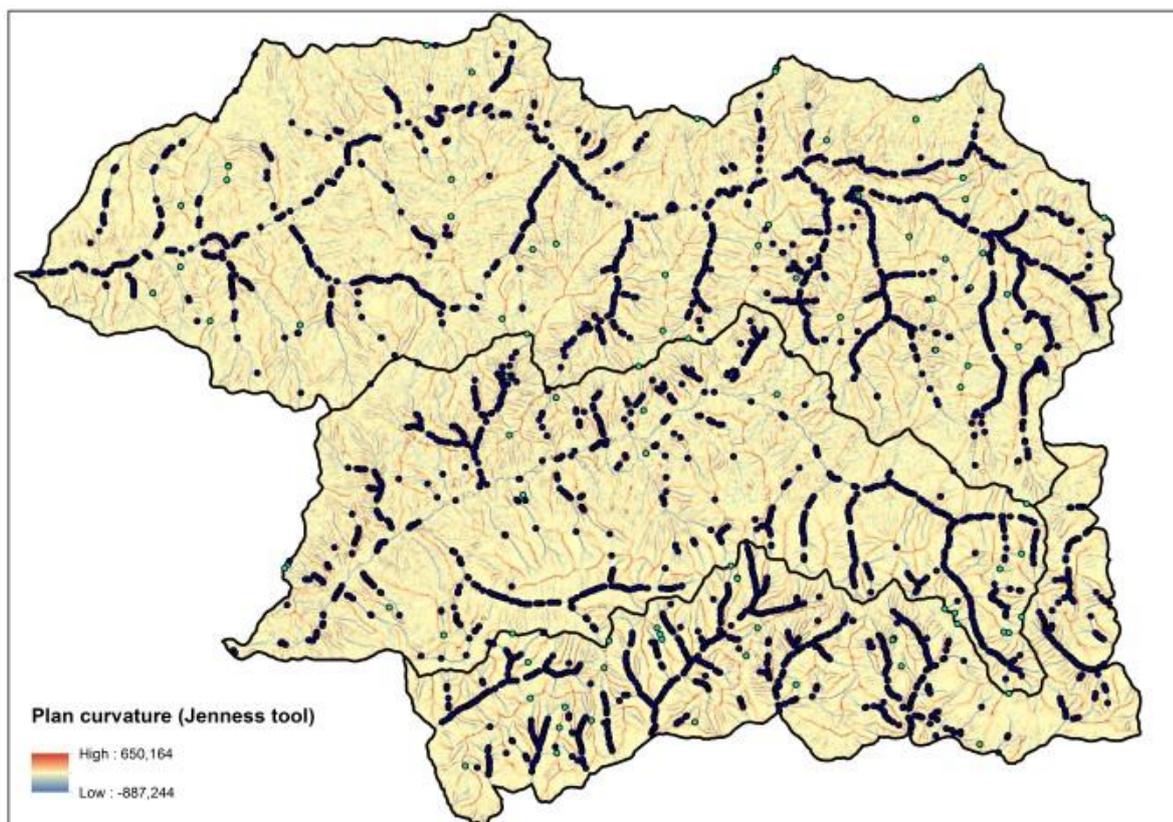


**Fig. 3.28** Rappresentazione della variabile *flow direction* (ArcGIS) (A) e della variabile *direct insolation* (SAGA) (B).

Come già evidenziato la *Plan Curvature* è un caso particolare: infatti, mentre il numero degli *outliers* è paragonabile a quello delle altre variabili, i suoi valori estremi presentano un ordine di grandezza apparentemente spropositato (valori nell'ordine delle centinaia rispetto ai valori tipici inferiori a dieci). Per comprendere tali anomalie, in Figura 3.29 sono stati mappati gli *outliers* di questa variabile, distinti secondo il segno matematico (punti verdi positivi - punti neri negativi), ed in Figura 3.30, sono stati evidenziati quelli con valori assoluti superiori a cento.

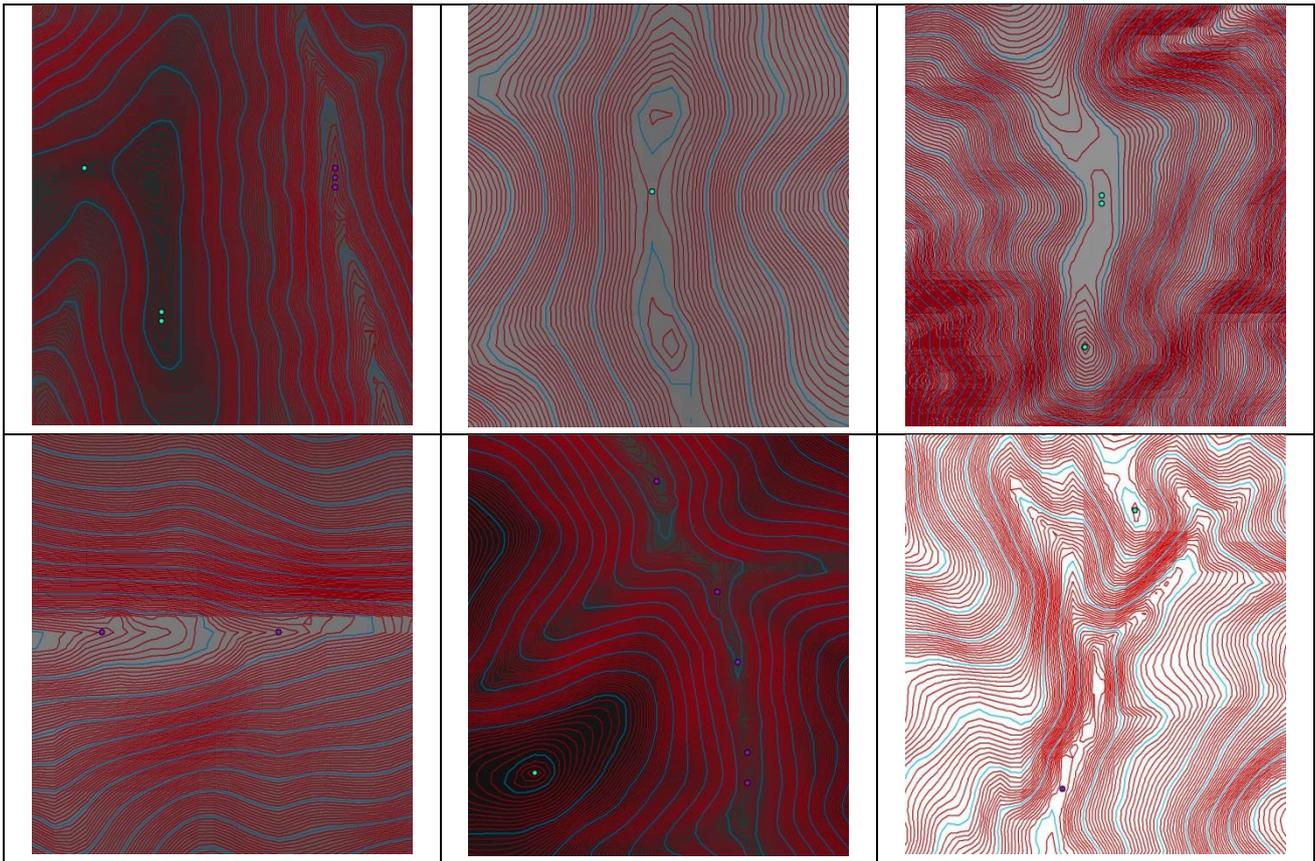


**Fig. 3.29** Rappresentazione degli *outliers* relativi alla *Plan Curvature* con valore maggiore di: (media+3 deviazione standard). I punti verdi hanno valore positivo e appartengono alle zone sommitali dei versanti, i punti neri hanno valori negativi e appartengono agli impluvi.



**Fig. 3.30** Rappresentazione degli *outliers* relativi alla *Plan Curvature* con valore assoluto maggiore di 100. (I punti verdi hanno valore positivo e appartengono alle zone sommitali dei versanti, i punti neri hanno valori negativi e appartengono agli impluvi).

Per comprendere la natura di questi *outliers* occorre considerare che la *Plan Curvature* è una misura della curvatura calcolata lungo le isoipse. In altre parole, un pixel assume il valore della curvatura dell'isoipsa in quel punto. Detta curvatura può essere associata ad una circonferenza completa ed i valori che essa assume sono inversamente proporzionali al raggio di quest'ultima. Il valore della curvatura si ottiene misurando il numero di radianti per una unità lineare di tale circonferenza (*radians/Linear Unit*) (Jenness). Come evidenziato nelle seguenti figure i punti corrispondenti agli *outliers* della *Plan Curvature* sono relativi a pixel ubicati sulla sommità dei versanti (con valori positivi) e nelle zone di impluvio (con valori negativi) ove la topografia è particolarmente articolata, come evidenziato con le isoipse nella seguente figura (Fig.3.31).



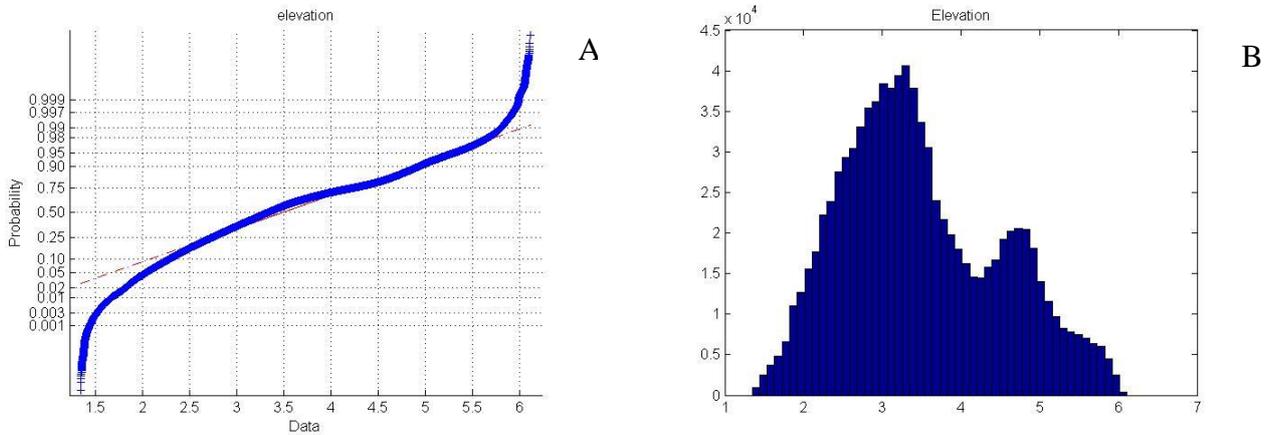
**Fig. 3.31** Rappresentazione di alcuni *outliers* relativi alla *Plan Curvature* con valore maggiore di: (media+3 deviazione standard) (punti verdi positivi - punti viola negativi), con l'indicazione delle isoipse a 1 m (linee rosse) e 5 m (linee blue).

La visualizzazione di detti *outliers* induce a considerare la possibilità che i valori statisticamente anomali sono connaturati con il calcolo delle variabili morfometriche. Più semplicemente, se i dati eccezionali derivassero dalla non accuratezza del DEM, osservando la variazione della pendenza, interpretando le isoipse, si riscontrerebbero delle discontinuità (es: un *sink*, determinato da un pixel con valore anomalo dell'altitudine rispetto all'intorno).

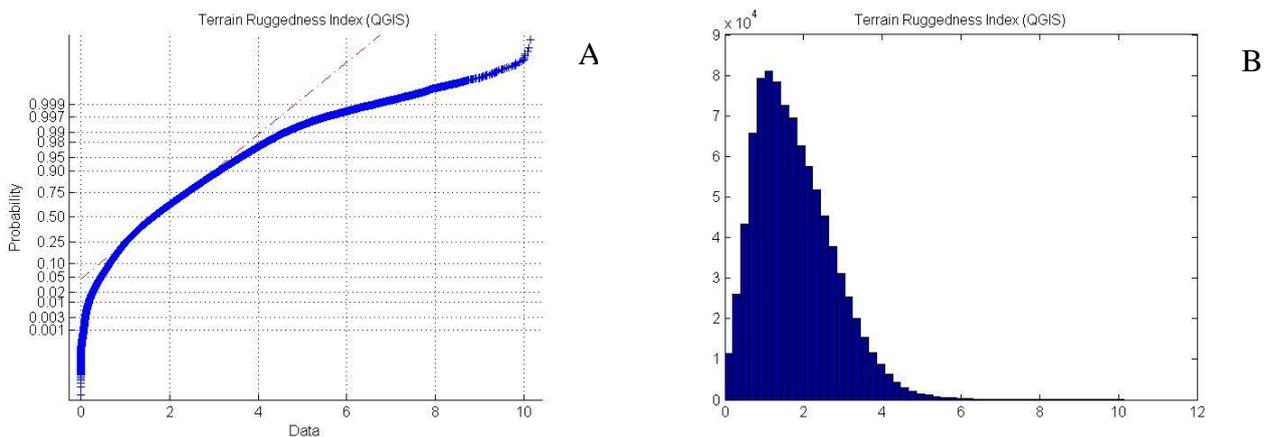
Una importante caratteristica delle variabili utilizzate come predittori è la distribuzione: infatti, molte procedure statistiche sono basate su alcune particolari assunzioni ed in molti casi è richiesto che le variabili esibiscano una distribuzione di tipo normale. Per ciascuna variabile è stato effettuato il test di normalità Shapiro-Wilk, con un 5% di livello di significatività, dal quale è risultato che nessuna delle variabili segue una distribuzione gaussiana. Un'efficace rappresentazione per osservare quanto la distribuzione di una variabile si discosta dalla funzione gaussiana è il *Normal Probability Plot*, costruito mettendo in relazione ciascun valore assunto dalla variabile con il valore di probabilità empirica. Sul grafico è tracciata anche una retta che unisce il 25° ed il 75° percentile dei dati (retta disegnata piena tra il 25° ed il 75° percentile e tratteggiata fino all'estensione del *range* di dati) Questi grafici, unitamente agli istogrammi di frequenza, permettono di osservare contestualmente *kurtosis* e *skewness*, oltre alla normalità.

L'andamento dei grafici (Figure da 3.32 a 3.61) conferma i risultati del test e quanto evidenziato con le statistiche di base. Tali andamenti confermano la non conformità ad una distribuzione gaussiana, ma rendono comunque possibili elaborazioni statistiche che consentono l'utilizzo di variabili con scostamenti non piccoli dalla normalità (come nel caso della *discriminant analysis*).

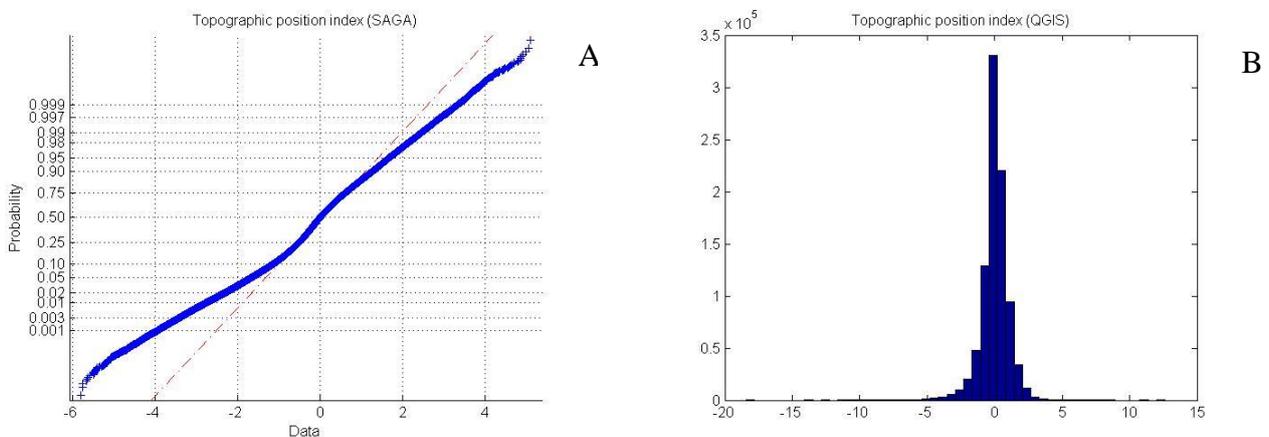
Considerando questi risultati e l'analisi effettuata sugli *outliers*, si è scelto di non procedere ad alcuna trasformazione finalizzata alla completa normalizzazione dei dati. Tale decisione è in accordo con l'idea di non voler cambiare le proprietà numeriche delle variabili, che rispecchiano il comportamento di specifici aspetti fisiografici del territorio.



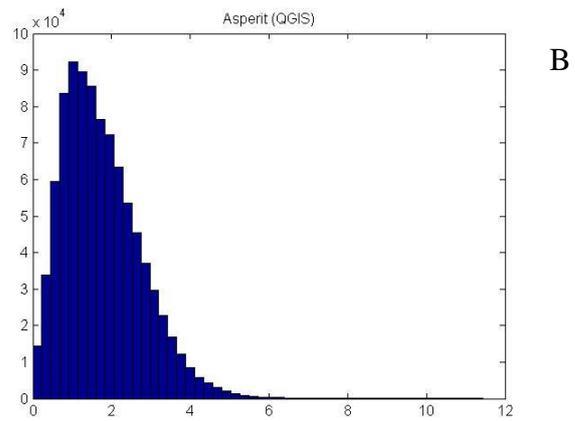
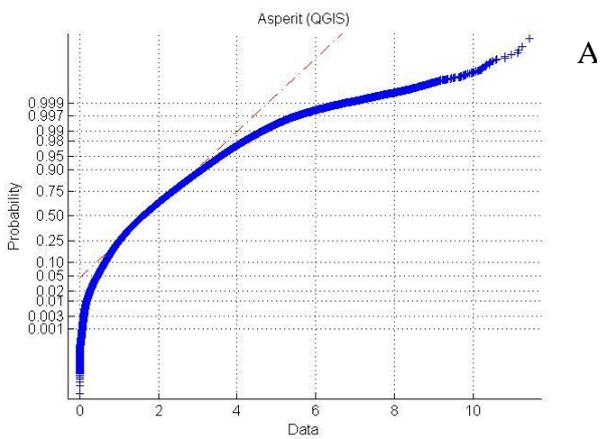
**Fig. 3.32** Normal probability plot relativo alla variabile *elevation* (A) e suo istogramma di frequenza (B).



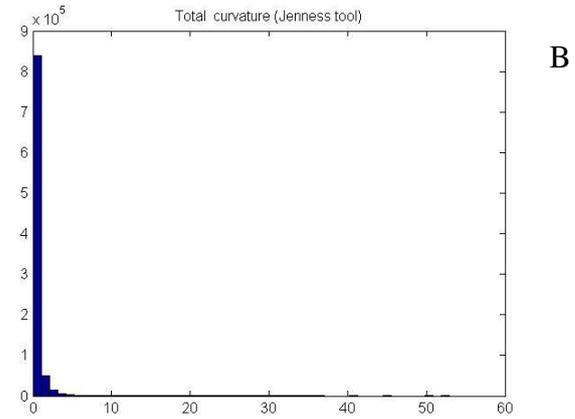
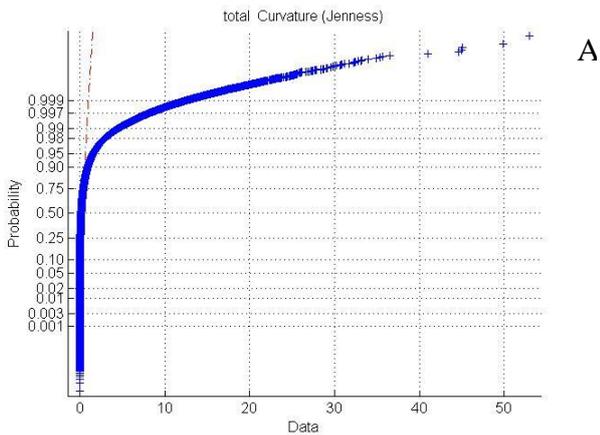
**Fig. 3.33** Normal probability plot relativo alla variabile *Terrain Ruggedness Index (QGIS)* (A) e suo istogramma di frequenza (B).



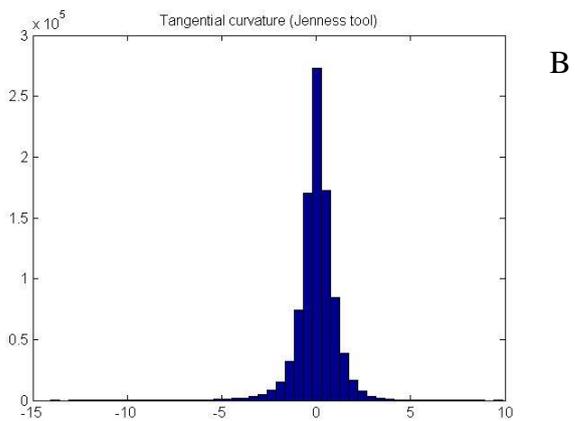
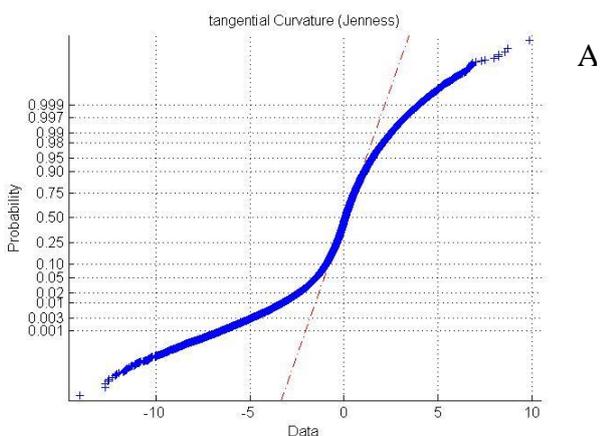
**Fig. 3.34** Normal probability plot relativo alla variabile *Topographic position index (QGIS)* (A) e suo istogramma di frequenza (B).



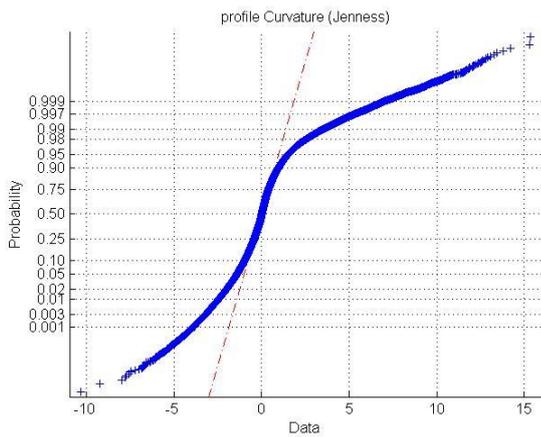
**Fig. 3.35** normal probability plot relative alla variabile *Asperit* (QGIS) (A) e suo istogramma di frequenza (B).



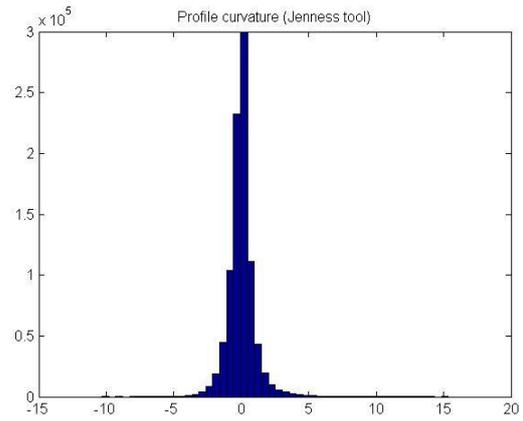
**Fig. 3.36** Normal probability plot relative alla variabile *total curvature* (Jenness tool) (A) e suo istogramma di frequenza (B).



**Fig. 3.37** Normal probability plot relative alla variabile *tangential curvature* (Jenness tool) (A) e suo istogramma di frequenza (B).

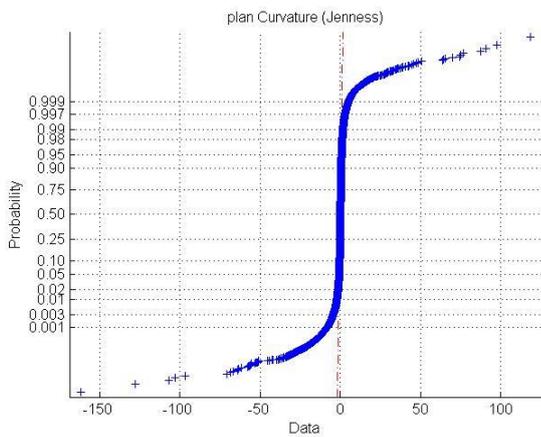


A

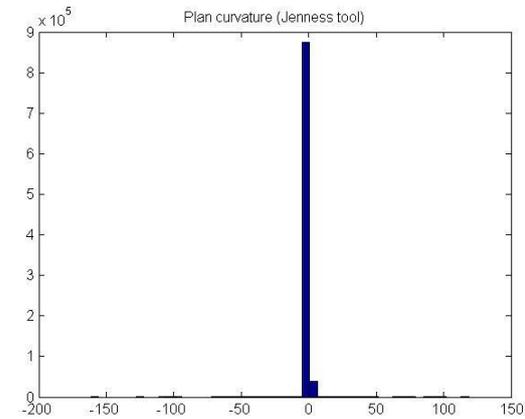


B

**Fig. 3.38** Normal probability plot relativo alla variabile *profile curvature* (Jenness tool) (A) e suo istogramma di frequenza (B).

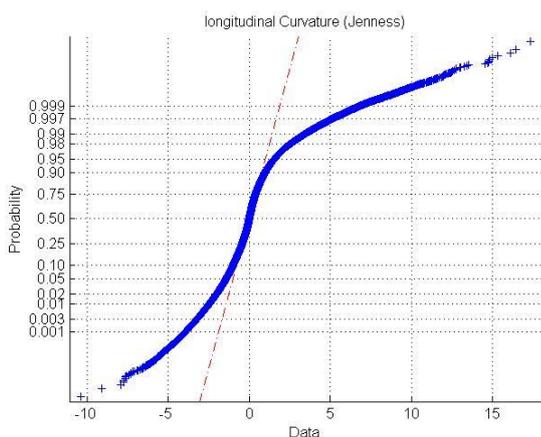


A

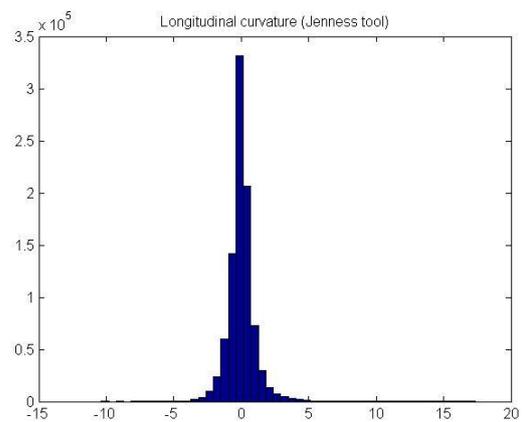


B

**Fig. 3.39** Normal probability plot relativo alla variabile *plan curvature* (Jenness tool) (A) e suo istogramma di frequenza (B).

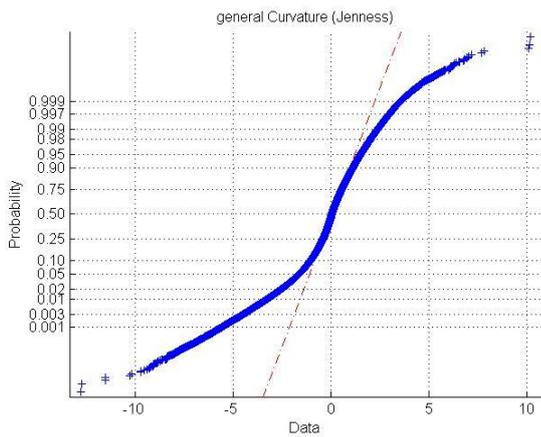


A

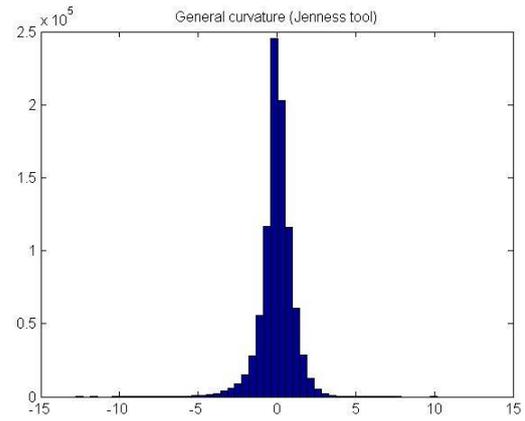


B

**Fig. 3.40** Normal probability plot relativo alla variabile *longitudinal curvature* (Jenness tool) (A) e suo istogramma di frequenza (B).

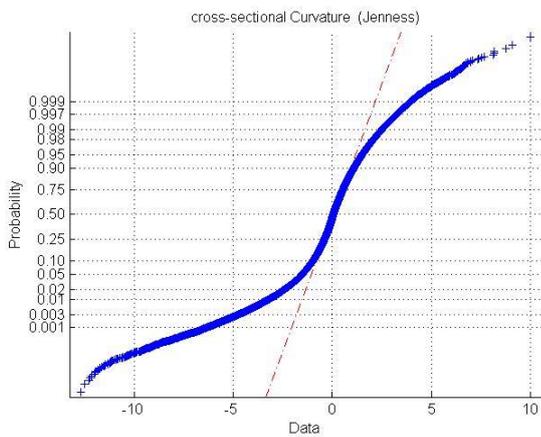


A

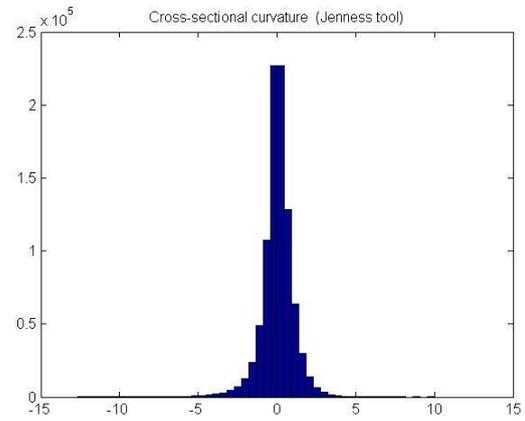


B

**Fig. 3.41** Normal probability plot relativo alla variabile *general curvature* (Jenness tool) (A) e suo istogramma di frequenza (B).

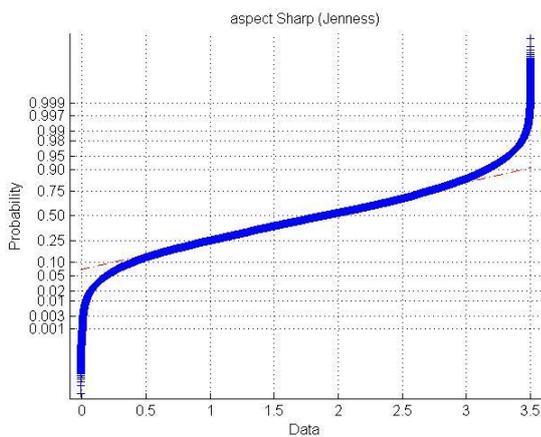


A

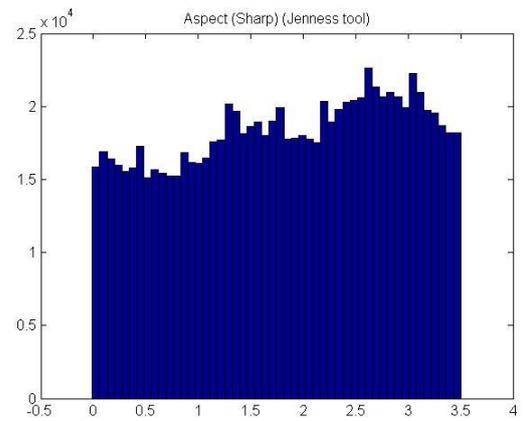


B

**Fig. 3.42** Normal probability plot relativo alla variabile *cross sectional* (Jenness tool) (A) e suo istogramma di frequenza (B).

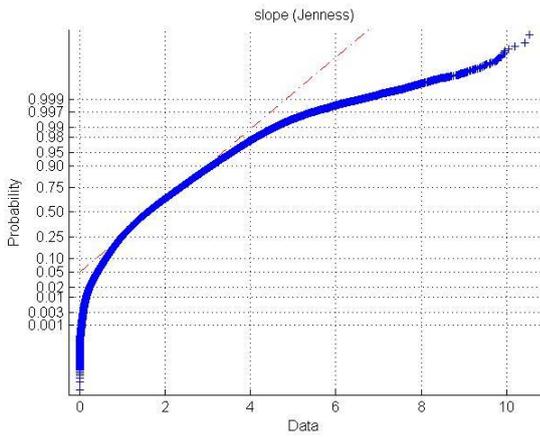


A

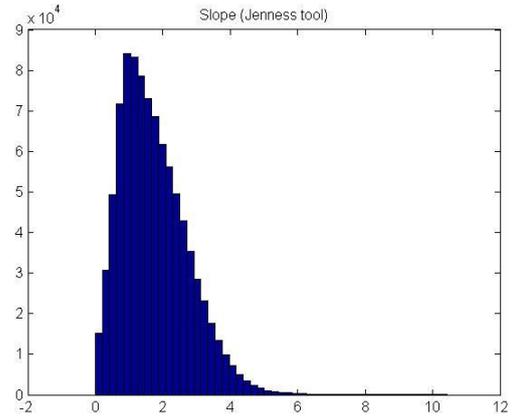


B

**Fig. 3.43** Normal probability plot relativo alla variabile *aspect* (Jenness tool) (A) e suo istogramma di frequenza (B).

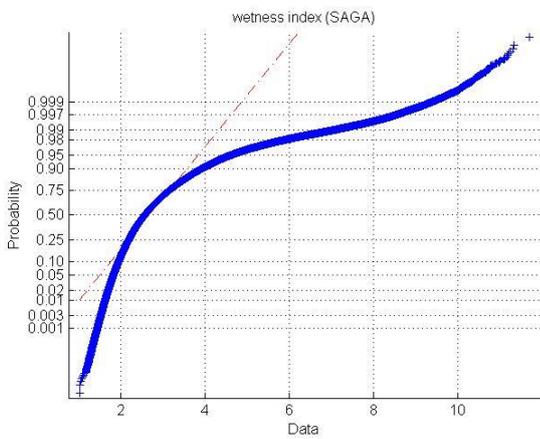


A

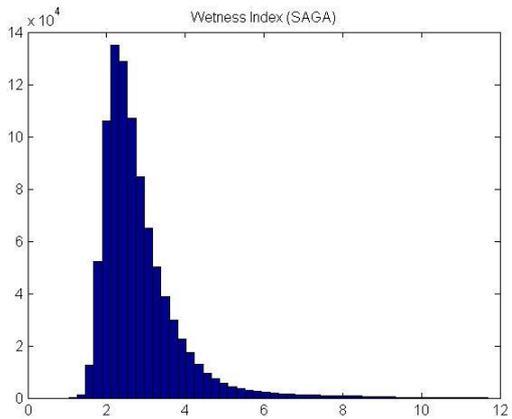


B

**Fig. 3.44** Normal probability plot relative alla variabile *slope* (Jenness tool) (A) e suo istogramma di frequenza (B).

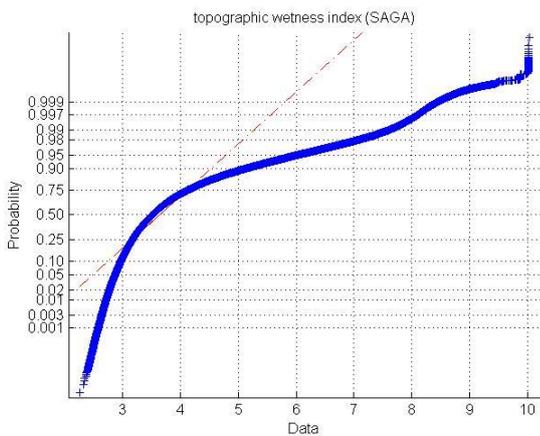


A

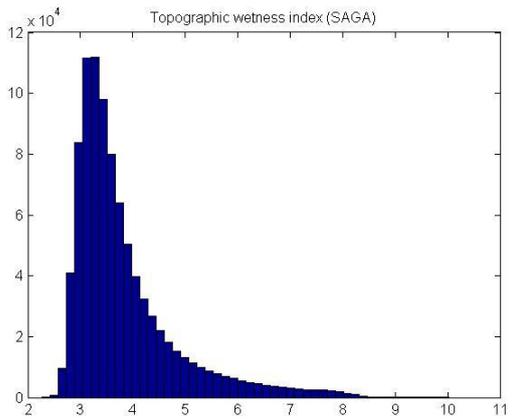


B

**Fig. 3.45** Normal probability plot relative alla variabile *wetness index* (SAGA) (A) e suo istogramma di frequenza (B).

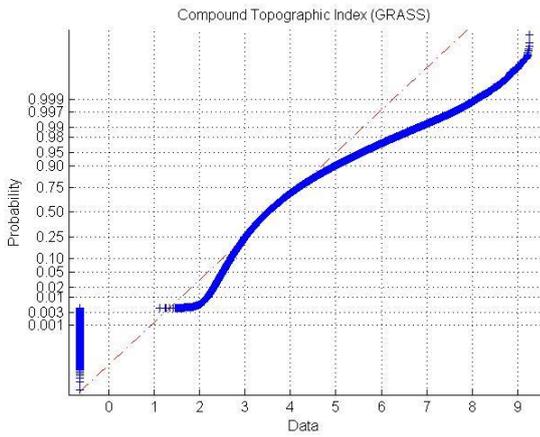


A

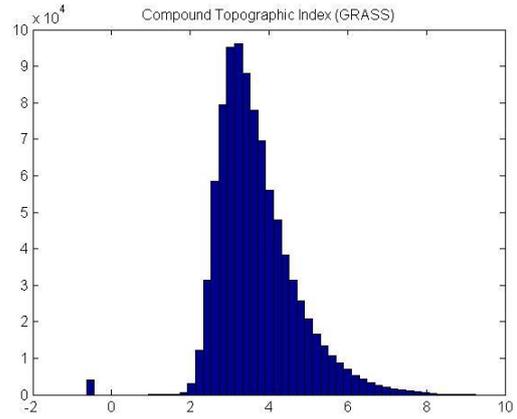


B

**Fig. 3.46** Normal probability plot relative alla variabile *topographic wetness index* (SAGA) (A) e suo istogramma di frequenza (B).

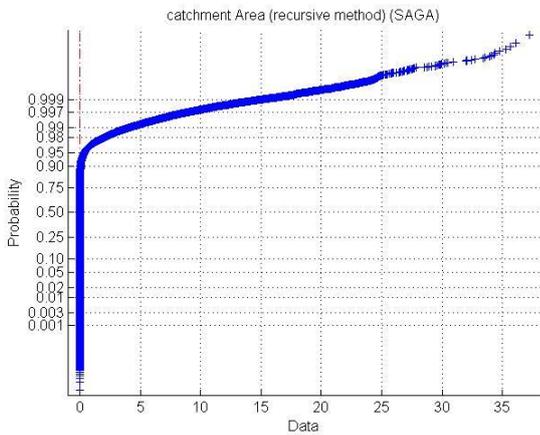


A

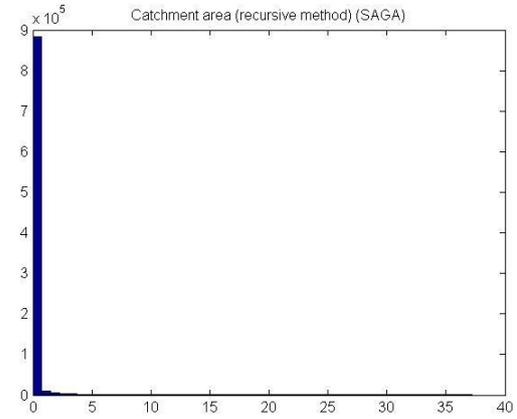


B

**Fig. 3.47** Normal probability plot relativo alla variabile *Compound Topographic Index (GRASS)* (A) e suo istogramma di frequenza (B).

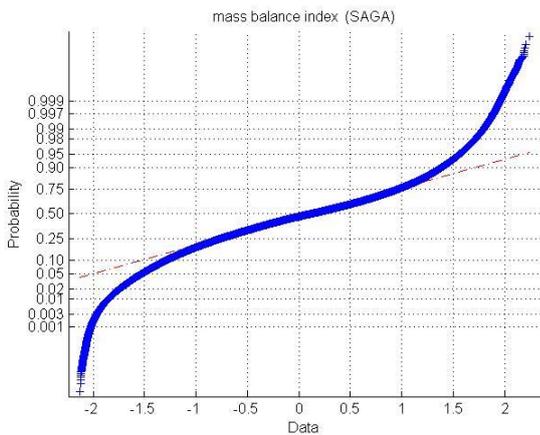


A

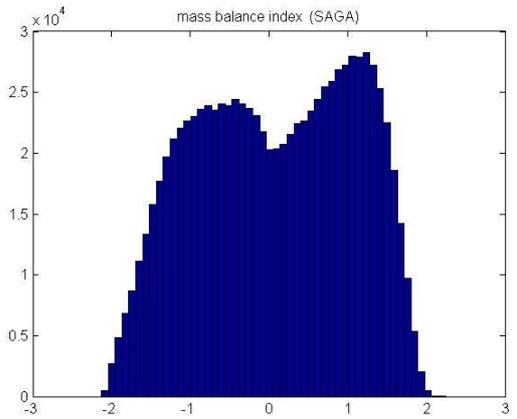


B

**Fig. 3.48** Normal probability plot relativo alla variabile *catchment Area (recursive method) (SAGA)* (A) e suo istogramma di frequenza (B).

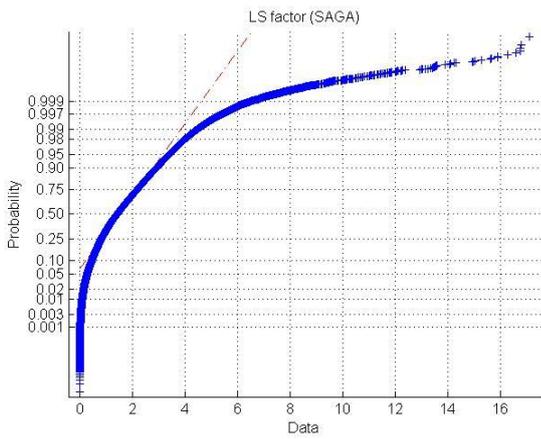


A

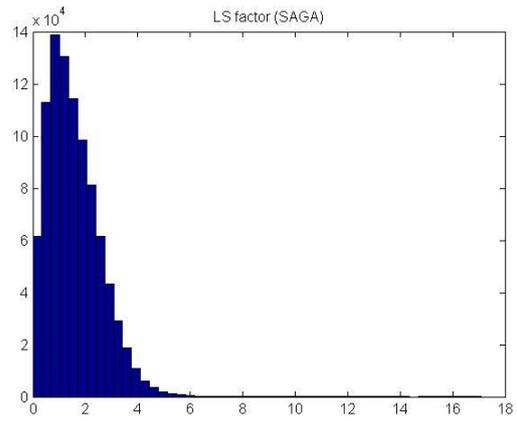


B

**Fig. 3.49** Normal probability plot relativo alla variabile *mass balance index (SAGA)* (A) e suo istogramma di frequenza (B).

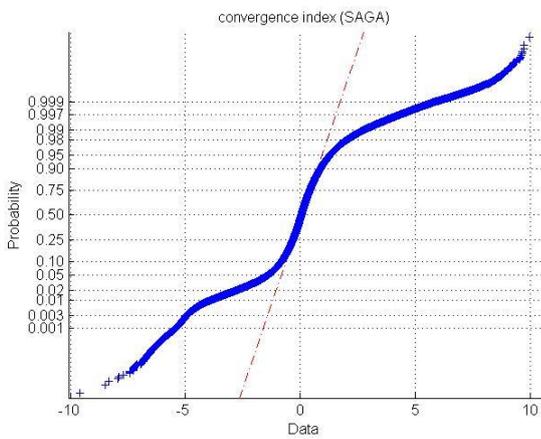


A

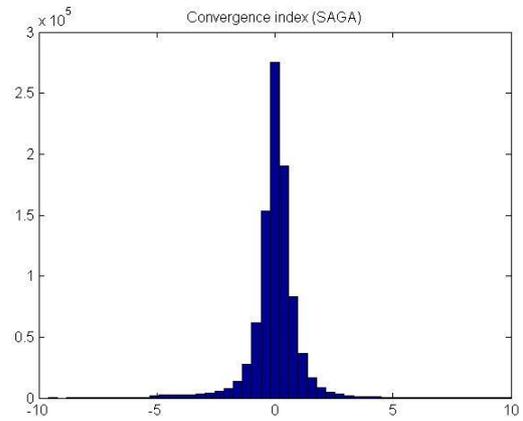


B

**Fig. 3.50** Normal probability plot relative alla variabile *LS Factor* (SAGA) (A) e suo istogramma di frequenza (B).

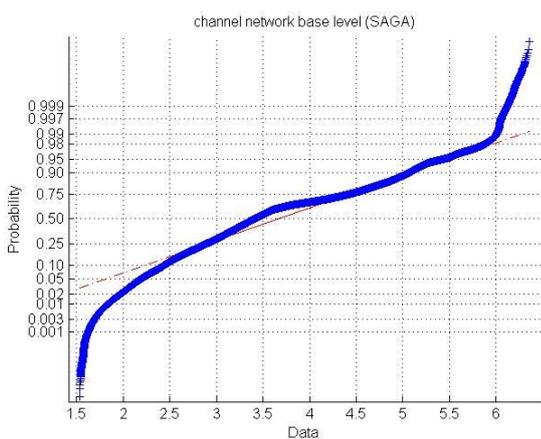


A

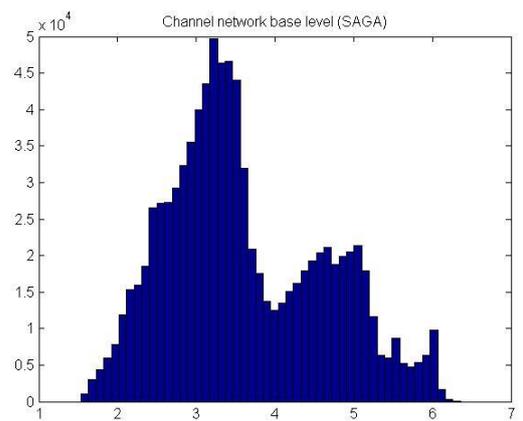


B

**Fig. 3.51** Normal probability plot relative alla variabile *convergence index* (SAGA) (A) e suo istogramma di frequenza (B).

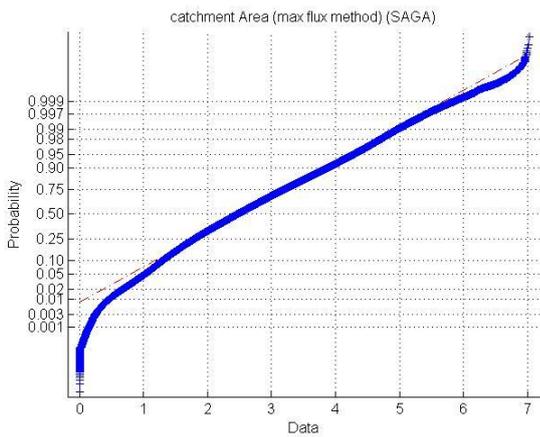


A

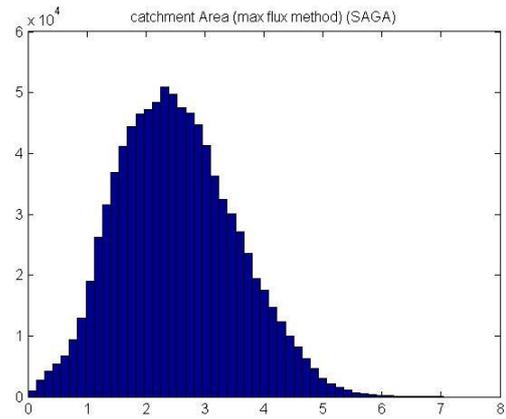


B

**Fig. 3.52** Normal probability plot relative alla variabile *channel network base level* (SAGA) (A) e suo istogramma di frequenza (B).

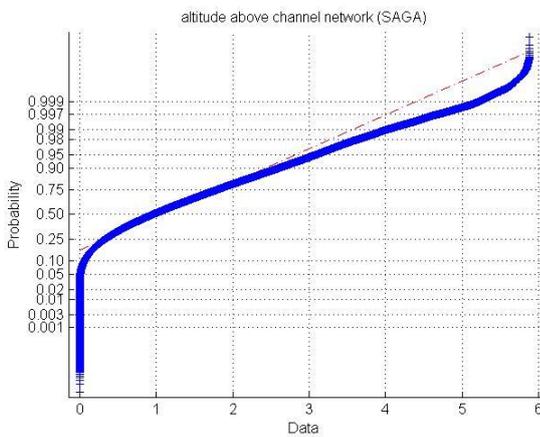


A

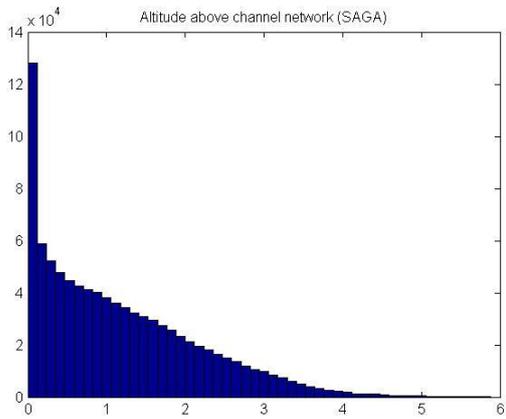


B

**Fig. 3.53** Normal probability plot relativo alla variabile *catchment Area (max flux method) (SAGA)* (A) e suo istogramma di frequenza (B).

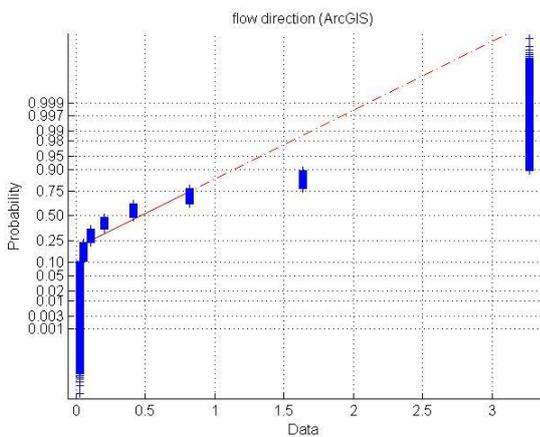


A

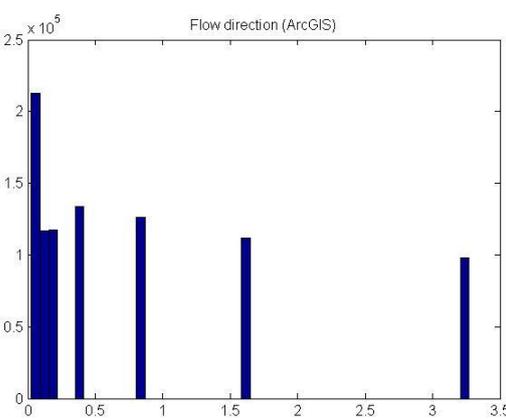


B

**Fig. 3.54** Normal probability plot relativo alla variabile *altitude above channel network (SAGA)* (A) e suo istogramma di frequenza (B).

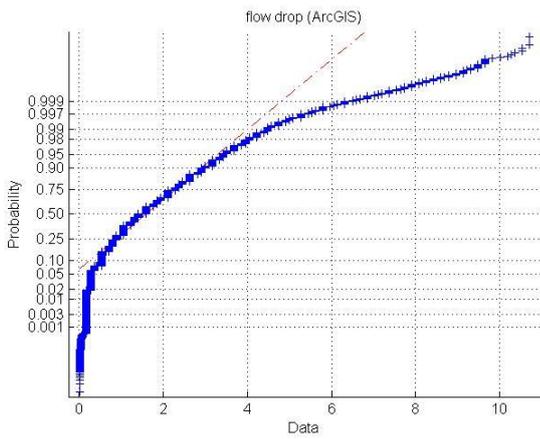


A

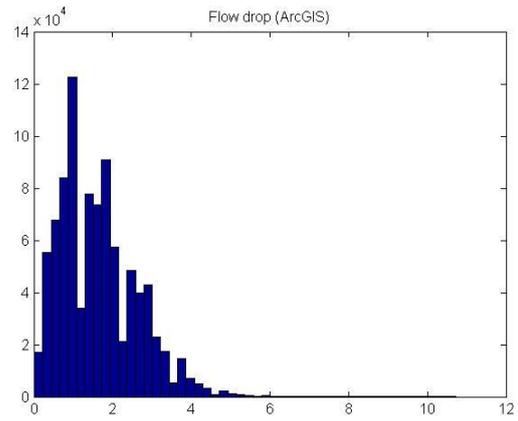


B

**Fig. 3.55** Normal probability plot relativo alla variabile *flow direction (ArcGIS)* (A) e suo istogramma di frequenza (B).

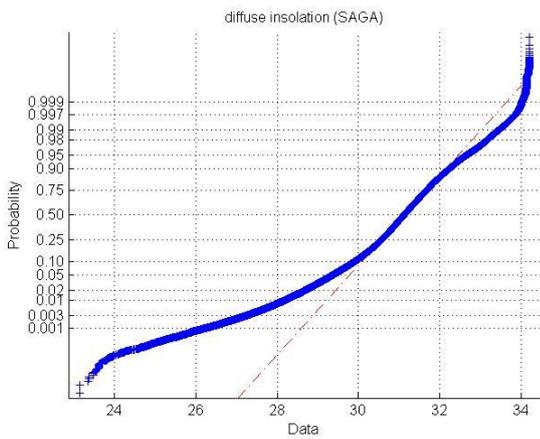


A

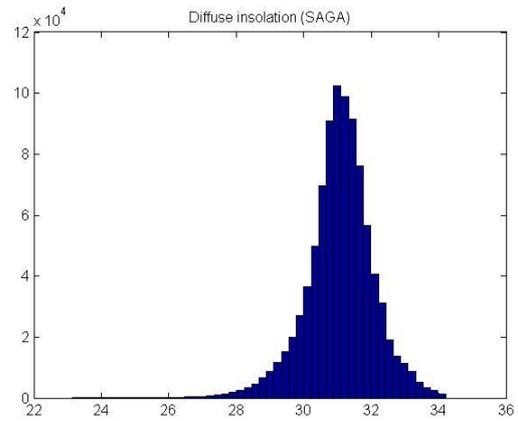


B

**Fig. 3.56** Normal probability plot relative alla variabile *flow drop* (ArcGIS) (A) e suo istogramma di frequenza (B).

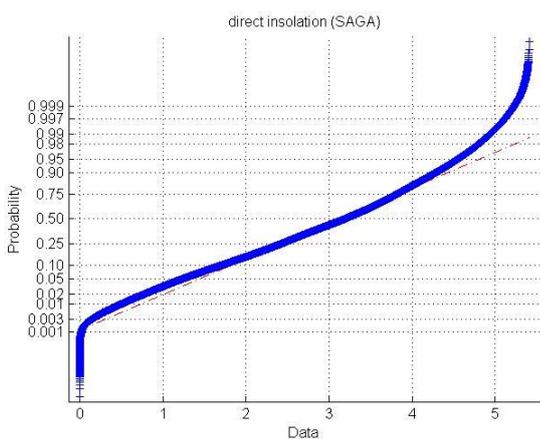


A

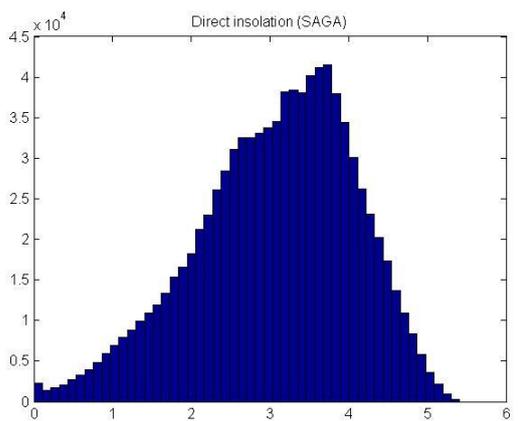


B

**Fig. 3.57** Normal probability plot relative alla variabile *diffuse insolation* (SAGA) (A) e suo istogramma di frequenza (B).

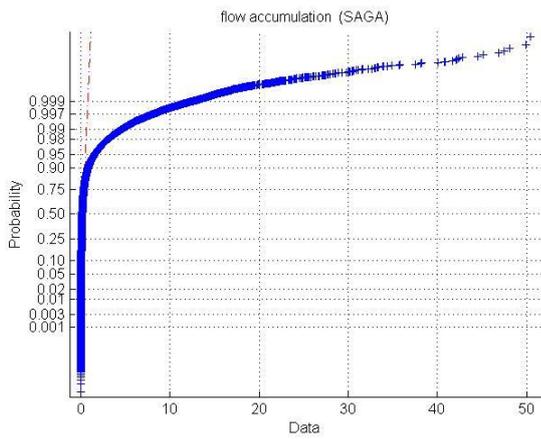


A

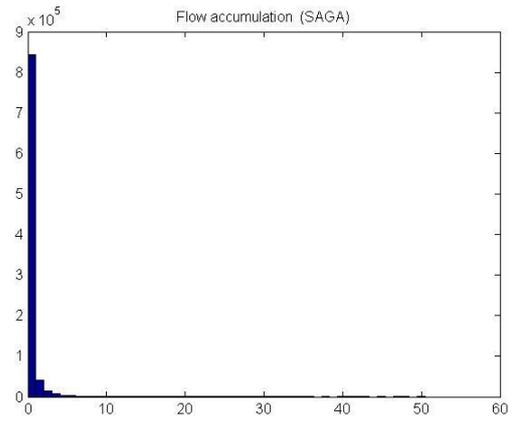


B

**Fig. 3.58** Normal probability plot relative alla variabile *direct insolation* (SAGA) (A) e suo istogramma di frequenza (B).

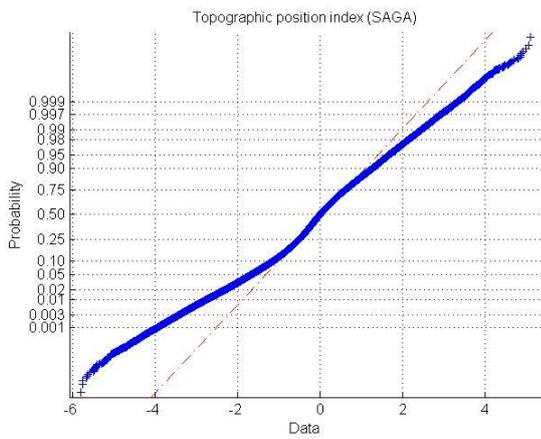


A

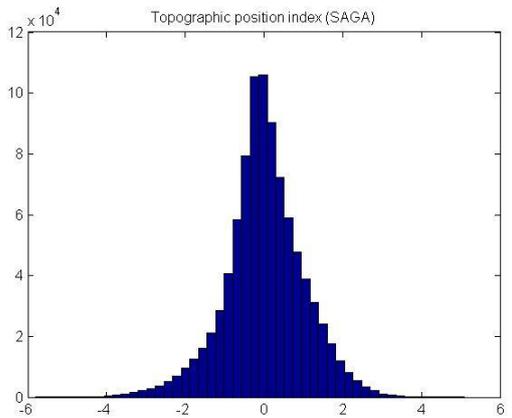


B

**Fig. 3.59** Normal probability plot relativo alla variabile *flow accumulation (SAGA)* (A) e suo istogramma di frequenza (B).

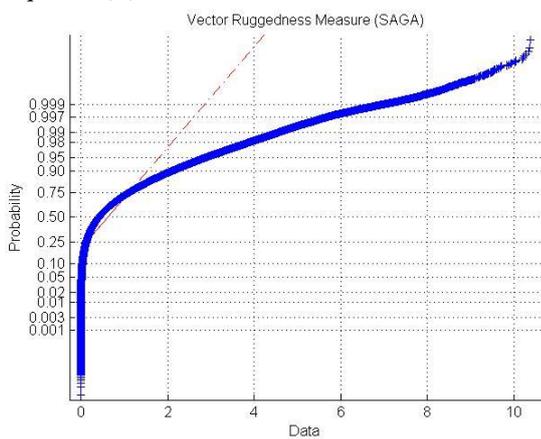


A

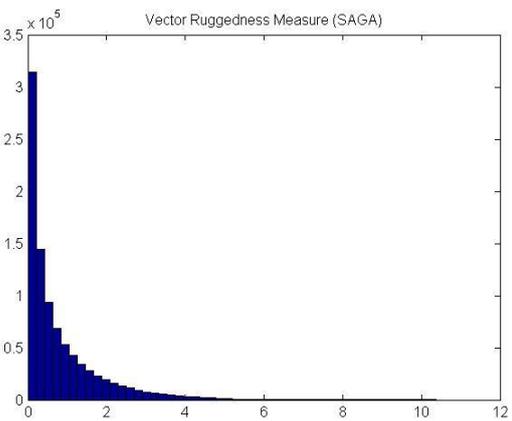


B

**Fig. 3.60** Normal probability plot relativo alla variabile *Topographic position index (SAGA)* (A) e suo istogramma di frequenza (B).



A



B

**Fig. 3.61** Normal probability plot relativo alla variabile *Vector Ruggedness Measure (SAGA)* (A) e suo istogramma di frequenza (B).

Nell'applicazione dei metodi statistici risulta interessante conoscere il tipo di relazione esistente tra le variabili, in particolare se è lineare o meno, nonché la sua forma (entità e direzione). L'indice di Pearson permette di determinare contestualmente sia l'entità che la direzione della correlazione lineare tra le variabili: quanto più il valore dell'indice si avvicina ad 1, tanto più forte è la relazione lineare tra due variabili; viceversa il segno dell'indice indica la direzione del legame (se il segno è positivo le variabili risultano correlate, altrimenti anticorrelate). Le seguenti tabelle contengono l'indice di correlazione lineare di Pearson (Tab. 3.6) e i valori dei *p value* (Tab. 3.7), calcolati sull'intero dataset delle variabili indipendenti. La matrice dei *p value* riporta i valori derivanti dal test: ipotesi di non correlazione contro l'alternativa di correlazione diversa da zero. Se il *p value* delle variabili corrispondenti alla combinazione (i, j) è piccolo (meno di 0,05) la correlazione tra le due variabili i e j è significativamente diversa da zero. In tabella, sono evidenziati i valori di correlazione (intensità del colore rossa) e di anticorrelazione (intensità del colore verde) che indicano una significativa relazione lineare tra alcune variabili. Risulta interessante osservare che vi sono variabili che esibiscono una correlazione lineare minore con quelle per cui ci si aspetterebbe di osservare un forte legame (per la tipologia di misura che veicolano o addirittura per la similitudine nel nome), rispetto ad altre variabili che hanno un significato fisico differente (es: *Topographic Position Index* (QGIS) ha una correlazione maggiore con la *Tangential Curvature* (Jenness tool), la *General Curvature* (Jenness tool) e la *Cross-Sectional Curvature* (Jenness tool) rispetto a quella con *Topographic Position Index* (SAGA)).

**Tabella 3.6** Indice di correlazione lineare di Pearson.

corr_coef	Zstd	TRI	TPI	asperit	ctot	ctang	cprof	cplan	clong	cgen	ccross	aspsharp	slsharp	WI	TWI	CTgrass	catarea	massbal	LSfactor	ConvInd	channelNB	catchstpe	altchan	flowdir	flowdrop	difsol	dirsol	flowaccs	topPosis	VRM
elevation	1,00	0,35	0,09	0,35	0,10	0,03	-0,11	0,04	-0,11	0,09	0,05	-0,03	0,35	-0,31	-0,39	-0,23	-0,16	0,12	0,28	0,07	0,97	0,20	0,56	0,03	0,36	0,47	-0,05	-0,12	0,16	0,25
Terrain Ruggedness Index (QGIS)	0,35	1,00	0,02	0,98	0,16	0,03	0,01	0,00	0,00	0,02	0,03	-0,03	1,00	-0,47	-0,54	-0,34	-0,19	0,04	0,89	0,00	0,32	0,78	0,26	0,01	0,96	-0,64	-0,17	-0,15	-0,03	0,71
Topographic position index (QGIS)	0,09	0,02	1,00	0,02	-0,39	0,78	-0,80	0,45	-0,85	0,99	0,84	-0,01	0,05	-0,50	-0,41	-0,57	-0,08	0,90	-0,27	0,70	0,01	-0,17	0,27	0,02	0,22	0,03	0,04	-0,26	0,75	-0,01
Asperit (QGIS)	0,35	0,98	0,02	1,00	0,13	0,04	0,00	0,01	0,00	0,02	0,04	-0,03	0,98	-0,47	-0,54	-0,34	-0,19	0,04	0,87	0,00	0,31	0,77	0,26	0,04	0,95	-0,63	-0,17	-0,15	-0,02	0,68
total Curvature (Jenness tool)	0,10	0,16	-0,39	0,13	1,00	-0,27	0,35	-0,23	0,36	-0,38	-0,29	0,00	0,09	0,18	0,10	0,15	0,03	-0,17	0,29	-0,30	0,13	0,17	-0,04	-0,01	0,05	0,00	-0,05	0,09	-0,25	0,16
tangential Curvature (Jenness tool)	0,03	0,03	0,78	0,04	-0,27	1,00	-0,29	0,59	-0,38	0,79	0,98	0,00	0,05	-0,51	-0,37	-0,58	-0,05	0,73	-0,25	0,77	-0,01	-0,05	0,16	0,02	0,13	-0,02	0,02	-0,34	0,59	0,00
profile Curvature (Jenness tool)	-0,11	0,01	-0,80	0,00	0,35	-0,29	1,00	-0,17	0,98	-0,81	-0,39	0,01	-0,01	0,32	0,31	0,35	0,09	-0,70	0,19	-0,38	-0,04	0,24	-0,29	-0,02	-0,21	-0,08	-0,05	0,10	-0,64	0,02
plan Curvature (Jenness tool)	0,04	0,00	0,45	0,01	-0,23	0,59	-0,17	1,00	-0,21	0,46	0,57	-0,01	0,03	-0,41	-0,28	-0,36	-0,10	0,40	-0,13	0,66	0,00	-0,11	0,14	0,01	0,07	0,01	0,02	-0,24	0,36	-0,01
longitudinal Curvature (Jenness tool)	-0,11	0,00	-0,85	0,00	0,36	-0,38	0,98	-0,21	1,00	-0,86	-0,46	0,01	-0,02	0,35	0,32	0,39	0,09	-0,75	0,21	-0,42	-0,03	0,23	-0,29	-0,01	-0,22	-0,07	-0,05	0,13	-0,68	0,02
general Curvature (Jenness tool)	0,09	0,02	0,99	0,02	-0,38	0,79	-0,81	0,46	-0,86	1,00	0,85	-0,01	0,04	-0,50	-0,42	-0,57	-0,08	0,89	-0,27	0,70	0,02	-0,18	0,28	0,03	0,22	0,04	0,05	-0,26	0,77	-0,01
cross-sectional Curvature (Jenness tool)	0,05	0,03	0,84	0,04	-0,29	0,98	-0,39	0,57	-0,46	0,85	1,00	0,00	0,06	-0,51	-0,39	-0,58	-0,05	0,77	-0,25	0,78	-0,01	-0,07	0,18	0,03	0,16	-0,01	0,03	-0,32	0,63	0,01
aspect (Jenness tool)	-0,03	-0,03	-0,01	-0,03	0,00	0,00	0,01	-0,01	0,01	-0,01	0,00	1,00	-0,03	0,04	0,05	0,01	0,04	-0,01	-0,02	-0,01	-0,02	-0,01	-0,04	-0,30	-0,03	0,00	-0,04	0,00	-0,01	-0,02
slope (Jenness tool)	0,35	1,00	0,05	0,98	0,09	0,05	-0,01	0,03	-0,02	0,04	0,06	-0,03	1,00	-0,50	-0,55	-0,36	-0,20	0,05	0,88	0,02	0,32	0,78	0,26	0,02	0,97	-0,65	-0,17	-0,16	-0,01	0,70
wetness index (SAGA)	-0,31	-0,47	-0,50	-0,47	0,18	-0,51	0,32	-0,41	0,35	-0,50	-0,51	0,04	-0,50	1,00	0,85	0,65	0,41	-0,48	-0,16	-0,61	-0,22	-0,16	-0,44	-0,03	-0,54	0,16	0,04	0,49	-0,51	-0,29
topographic wetness index (SAGA)	-0,39	-0,54	-0,41	-0,54	0,10	-0,37	0,31	-0,28	0,32	-0,42	-0,39	0,05	-0,55	0,85	1,00	0,54	0,54	-0,40	-0,29	-0,42	-0,29	-0,24	-0,51	-0,04	-0,59	0,14	0,05	0,38	-0,48	-0,33
Compound Topographic Index (GRASS)	-0,23	-0,34	-0,57	-0,34	0,15	-0,58	0,35	-0,36	0,39	-0,57	-0,58	0,01	-0,36	0,65	0,54	1,00	-0,01	-0,55	-0,04	-0,55	-0,15	-0,09	-0,38	-0,01	-0,44	0,15	0,02	0,50	-0,57	-0,24
catchment Area (recursive method) (SAGA)	-0,16	-0,19	-0,08	-0,19	0,03	-0,05	0,09	-0,10	0,09	-0,08	-0,05	0,04	-0,20	0,41	0,54	-0,01	1,00	-0,08	-0,13	-0,15	-0,13	-0,07	-0,17	-0,03	-0,19	0,00	0,02	0,07	-0,11	-0,10
mass balance index (SAGA)	0,12	0,04	0,90	0,04	-0,17	0,73	-0,70	0,40	-0,75	0,89	0,77	-0,01	0,05	-0,48	-0,40	-0,55	-0,08	1,00	-0,21	0,65	0,06	-0,17	0,27	0,02	0,22	0,05	0,03	-0,27	0,70	0,02
LS factor (SAGA)	0,28	0,89	-0,27	0,87	0,29	-0,25	0,19	-0,13	0,21	-0,27	-0,25	-0,02	0,88	-0,16	-0,29	-0,04	-0,13	-0,21	1,00	-0,24	0,28	0,76	0,11	0,01	0,81	-0,59	-0,18	0,03	-0,31	0,63
convergence index (SAGA)	0,07	0,00	0,70	0,00	-0,30	0,77	-0,38	0,66	-0,42	0,70	0,78	-0,01	0,02	-0,61	-0,42	-0,55	-0,15	0,65	-0,24	1,00	0,01	-0,16	0,22	0,03	0,10	0,04	0,03	-0,36	0,55	-0,02
channel network base level (SAGA)	0,97	0,32	0,01	0,31	0,13	-0,01	-0,04	0,00	-0,03	0,02	-0,01	-0,02	0,32	-0,22	-0,29	-0,15	-0,13	0,06	0,28	0,01	1,00	0,21	0,33	0,02	0,32	0,47	-0,04	-0,07	0,03	0,24
catchment Area (max flux method) (SAGA)	0,20	0,78	-0,17	0,77	0,17	-0,05	0,24	-0,11	0,23	-0,18	-0,07	-0,01	0,78	-0,16	-0,24	-0,09	-0,07	-0,17	0,76	-0,16	0,21	1,00	0,06	0,01	0,71	-0,56	-0,17	-0,01	-0,27	0,55
altitude above channel network (SAGA)	0,56	0,26	0,27	0,26	-0,04	0,16	-0,29	0,14	-0,29	0,28	0,18	-0,04	0,26	-0,44	-0,51	-0,38	-0,17	0,27	0,11	0,22	0,33	0,06	1,00	0,05	0,32	0,23	-0,03	-0,20	0,48	0,15
flow direction (ArcGIS)	0,03	0,01	0,02	0,04	-0,01	0,02	-0,02	0,01	-0,01	0,03	0,03	-0,30	0,02	-0,03	-0,04	-0,01	-0,03	0,02	0,01	0,03	0,02	0,01	0,05	1,00	0,02	0,02	-0,61	-0,01	0,02	-0,11
flow drop (ArcGIS)	0,36	0,96	0,22	0,95	0,05	0,13	-0,21	0,07	-0,22	0,22	0,16	-0,03	0,97	-0,54	-0,59	-0,44	-0,19	0,22	0,81	0,10	0,32	0,71	0,32	0,02	1,00	-0,61	-0,15	-0,18	0,13	0,68
diffuse insolation (SAGA)	0,47	-0,64	0,03	-0,63	0,00	-0,02	-0,08	0,01	-0,07	0,04	-0,01	0,00	-0,65	0,16	0,14	0,15	0,00	0,05	-0,59	0,04	0,47	-0,56	0,23	0,02	-0,61	1,00	0,12	0,04	0,15	-0,47
direct insolation (SAGA)	-0,05	-0,17	0,04	-0,17	-0,05	0,02	-0,05	0,02	-0,05	0,05	0,03	-0,04	-0,17	0,04	0,05	0,02	0,02	0,03	-0,18	0,03	-0,04	-0,17	-0,03	-0,61	-0,15	0,12	1,00	0,02	0,07	-0,08
flow accumulation (SAGA)	-0,12	-0,15	-0,26	-0,15	0,09	-0,34	0,10	-0,24	0,13	-0,26	-0,32	0,00	-0,16	0,49	0,38	0,50	0,07	-0,27	0,03	-0,36	-0,07	-0,01	-0,20	-0,01	-0,18	0,04	0,02	1,00	-0,27	-0,10
Topographic position index (SAGA)	0,16	-0,03	0,75	-0,02	-0,25	0,59	-0,64	0,36	-0,68	0,77	0,63	-0,01	-0,01	-0,51	-0,48	-0,57	-0,11	0,70	-0,31	0,55	0,03	-0,27	0,48	0,02	0,13	0,15	0,07	-0,27	1,00	-0,04
Vector Ruggedness Measure (SAGA)	0,25	0,71	-0,01	0,68	0,16	0,00	0,02	-0,01	0,02	-0,01	0,01	-0,02	0,70	-0,29	-0,33	-0,24	-0,10	0,02	0,63	-0,02	0,24	0,55	0,15	-0,11	0,68	-0,47	-0,08	-0,10	-0,04	1,00

**Tabella 3.7** Matrice dei *p value* del test: ipotesi di non correlazione contro l'alternativa di correlazione diversa da zero. Quando il *p value* delle variabili corrispondenti alla combinazione (i, j) è piccolo (meno di 0,05) la correlazione tra le due variabili i e j è significativamente diversa da zero.

p_value	Zstd	TRI	TPI	asperit	ctot	ctang	cprof	cplan	clong	cgen	ccross	aspsharp	slsharp	WI	TWI
elevation	1	0	0	0	0	3,29E-215	0	0	0	0	0	2,65E-173	0	0	0
Terrain Ruggedness Index (QGIS)	0	1	9,43E-55	0	0	1,61E-171	1,45E-17	0,901691	0,000545	2,86E-54	2,14E-202	1,37E-171	0	0	0
Topographic position index (QGIS)	0	9,43E-55	1	1,84E-124	0	0	0	0	0	0	0	1,10E-09	0	0	0
Asperit (QGIS)	0	0	1,84E-124	1	0	9,71E-280	0,000153	8,43E-16	0,000299	3,81E-117	1,23E-284	1,41E-163	0	0	0
total Curvature (Jenness tool)	0	0	0	0	1	0	0	0	0	0	0	0,000322	0	0	0
tangential Curvature (Jenness tool)	3,29E-215	1,61E-171	0	9,71E-280	0	1	0	0	0	0	0	0,000851	0	0	0
profile Curvature (Jenness tool)	0	1,45E-17	0	0,000153	0	0	1	0	0	0	0	8,90E-12	1,02E-29	0	0
plan Curvature (Jenness tool)	0	0,901691	0	8,43E-16	0	0	0	1	0	0	0	1,23E-16	9,45E-132	0	0
longitudinal Curvature (Jenness tool)	0	0,000545	0	0,000299	0	0	0	0	1	0	0	1,01E-12	7,04E-67	0	0
general Curvature (Jenness tool)	0	2,86E-54	0	3,81E-117	0	0	0	0	0	1	0	1,79E-09	0	0	0
cross-sectional Curvature (Jenness tool)	0	2,14E-202	0	1,23E-284	0	0	0	0	0	0	1	0,00205	0	0	0
aspect (Jenness tool)	2,65E-173	1,37E-171	1,10E-09	1,41E-163	0,000322	0,000851	8,90E-12	1,23E-16	1,01E-12	1,79E-09	0,00205	1	5,19E-176	0	0
slope (Jenness tool)	0	0	0	0	0	0	1,02E-29	9,45E-132	7,04E-67	0	0	5,19E-176	1	0	0
wetness index (SAGA)	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
topographic wetness index (SAGA)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Compound Topographic Index (GRASS)	0	0	0	0	0	0	0	0	0	0	0	2,72E-23	0	0	0
catchment Area (recursive method) (SAGA)	0	0	0	0	8,16E-187	0	0	0	0	0	0	2,26E-277	0	0	0
mass balance index (SAGA)	0	2,08E-281	0	0	0	0	0	0	0	0	0	1,39E-14	0	0	0
LS factor (SAGA)	0	0	0	0	0	0	0	0	0	0	0	1,94E-109	0	0	0
convergence index (SAGA)	0	0,054643	0	0,000319	0	0	0	0	0	0	0	1,38E-12	4,83E-122	0	0
channel network base level (SAGA)	0	0	2,56E-45	0	0	1,34E-33	1,73E-265	2,56E-05	1,14E-196	5,20E-49	1,48E-07	7,57E-97	0	0	0
catchment Area (max flux method) (SAGA)	0	0	0	0	0	0	0	0	0	0	0	3,88E-17	0	0	0
altitude above channel network (SAGA)	0	0	0	0	0	0	0	0	0	0	0	2,29E-284	0	0	0
flow direction (ArcGIS)	1,82E-205	2,44E-28	9,88E-111	0	8,71E-23	1,79E-84	2,45E-100	8,41E-40	9,90E-41	6,18E-139	3,58E-193	0	3,11E-53	6,46E-192	5,8134700
flow drop (ArcGIS)	0	0	0	0	0	0	0	0	0	0	0	7,26E-185	0	0	0
diffuse insolation (SAGA)	0	0	7,23E-194	0	0,001059	7,38E-88	0	3,28E-46	0	4,93E-266	1,52E-31	0,775299	0	0	0
direct insolation (SAGA)	0	0	0	0	0	8,35E-120	0	5,73E-78	0	0	1,29E-163	5,89E-304	0	0	0
flow accumulation (SAGA)	0	0	0	0	0	0	0	0	0	0	0	0,020141	0	0	0
Topographic position index (SAGA)	0	5,46E-140	0	1,00E-82	0	0	0	0	0	0	0	2,43E-08	1,96E-17	0	0
Vector Ruggedness Measure (SAGA)	0	0	1,05E-13	0	0	4,99E-05	6,06E-84	3,30E-27	8,56E-58	6,94E-11	2,39E-07	2,70E-51	0	0	0

**Tabella 3.7** (continua).

p_value	CTIgrass	catarea	massbal	LSfactor	ConvInd	channelNB	catchslope	altichan	flowdir	flowdrop	difsol	dirsol	flowaccS	topPosiS	VRM
elevation	0	0	0	0	0	0	0	0	1,82E-205	0	0	0	0	0	0
Terrain Ruggedness Index (QGIS)	0	0	2,08E-281	0	0,054643	0	0	0	2,44E-28	0	0	0	0	5,46E-140	0
Topographic position index (QGIS)	0	0	0	0	0	2,56E-45	0	0	9,88E-111	0	7,23E-194	0	0	0	1,05E-13
Asperit (QGIS)	0	0	0	0	0,000319	0	0	0	0	0	0	0	0	1,00E-82	0
total Curvature (Jenness tool)	0	8,16E-187	0	0	0	0	0	0	8,71E-23	0	0,001059	0	0	0	0
tangential Curvature (Jenness tool)	0	0	0	0	0	1,34E-33	0	0	1,79E-84	0	7,38E-88	8,35E-120	0	0	4,99E-05
profile Curvature (Jenness tool)	0	0	0	0	0	1,73E-265	0	0	2,45E-100	0	0	0	0	0	6,06E-84
plan Curvature (Jenness tool)	0	0	0	0	0	2,56E-05	0	0	8,41E-40	0	3,28E-46	5,73E-78	0	0	3,30E-27
longitudinal Curvature (Jenness tool)	0	0	0	0	0	1,14E-196	0	0	9,90E-41	0	0	0	0	0	8,56E-58
general Curvature (Jenness tool)	0	0	0	0	0	5,20E-49	0	0	6,18E-139	0	4,93E-266	0	0	0	6,94E-11
cross-sectional Curvature (Jenness tool)	0	0	0	0	0	1,48E-07	0	0	3,58E-193	0	1,52E-31	1,29E-163	0	0	2,39E-07
aspect (Jenness tool)	2,72E-23	2,26E-277	1,39E-14	1,94E-109	1,38E-12	7,57E-97	3,88E-17	2,29E-284	0	7,26E-185	0,775299	5,89E-304	0,020141	2,43E-08	2,70E-51
slope (Jenness tool)	0	0	0	0	4,83E-122	0	0	0	3,11E-53	0	0	0	0	1,96E-17	0
wetness index (SAGA)	0	0	0	0	0	0	0	0	6,46E-192	0	0	0	0	0	0
topographic wetness index (SAGA)	0	0	0	0	0	0	0	0	5,8134700	0	0	0	0	0	0
Compound Topographic Index (GRASS)	1	7,58E-28	0	5,24E-248	0	0	0	0	5,69E-09	0	0	1,67E-122	0	0	0
catchment Area (recursive method) (SAGA)	7,58E-28	1	0	0	0	0	0	0	6,99E-199	0	0,006967	3,46E-54	0	0	0
mass balance index (SAGA)	0	0	1	0	0	0	0	0	1,28E-122	0	0	9,32E-194	0	0	3,39E-70
LS factor (SAGA)	5,24E-248	0	0	1	0	0	0	0	9,76E-27	0	0	0	1,15E-214	0	0
convergence index (SAGA)	0	0	0	0	1	2,04E-38	0	0	6,93E-243	0	3,44E-273	2,89E-170	0	0	1,36E-66
channel network base level (SAGA)	0	0	0	0	2,04E-38	1	0	0	1,97E-101	0	0	0	0	2,77E-240	0
catchment Area (max flux method) (SAGA)	0	0	0	0	0	0	1	0	1,43E-08	0	0	0	4,51E-16	0	0
altitude above channel network (SAGA)	0	0	0	0	0	0	0	1	0	0	0	3,91E-246	0	0	0
flow direction (ArcGIS)	5,69E-09	6,99E-199	1,28E-122	9,76E-27	6,93E-243	1,97E-101	1,43E-08	0	1	5,36E-114	3,21E-59	0	2,71E-08	4,35E-107	0
flow drop (ArcGIS)	0	0	0	0	0	0	0	0	5,36E-114	1	0	0	0	0	0
diffuse insolation (SAGA)	0	0,006967	0	0	3,44E-273	0	0	0	3,21E-59	0	1	0	0	0	0
direct insolation (SAGA)	1,67E-122	3,46E-54	9,32E-194	0	2,89E-170	0	0	3,91E-246	0	0	0	1	5,32E-48	0	0
flow accumulation (SAGA)	0	0	0	1,15E-214	0	0	4,51E-16	0	2,71E-08	0	0	5,32E-48	1	0	0
Topographic position index (SAGA)	0	0	0	0	0	2,77E-240	0	0	4,35E-107	0	0	0	0	1	0
Vector Ruggedness Measure (SAGA)	0	0	3,39E-70	0	1,36E-66	0	0	0	0	0	0	0	0	0	1

Un ulteriore aspetto da considerare in relazione ai predittori è rappresentato da una loro eventuale collinearità o multicollinearità (Stewart 1987): tale verifica è stata effettuata con il test di Belsley (Belsley 1991). In generale è probabile che una correlazione lineare tra i regressori comporti un certo grado di dipendenza lineare, mentre l'assenza di correlazione elevata non esclude la collinearità e, d'altro canto, correlazioni elevate non corrispondono a collinearità. Questo tipo di test è di fondamentale importanza nel caso in cui si utilizzino le variabili con i metodi di regressione lineare. In tali tipi di regressione si assume l'esistenza di una relazione lineare tra le variabili dipendenti e quelle indipendenti, oltre alla mancanza di collinearità tra le variabili indipendenti e alla distribuzione normale e indipendenza dei residui (Osborne e Waters 2002). In dette applicazioni, se due o più variabili indipendenti sono in combinazione lineare tra loro il determinante della matrice utilizzata per la il calcolo dei coefficienti di regressione tende a zero e ciò comporta la determinazione di modelli instabili.

Il test di Belsley fornisce come *output*:

- i *singular values* che sono gli elementi della diagonale della matrice che si ottiene decomponendo la matrice dei predittori con la tecnica *singular-value decomposition*, indicati nella prima colonna della Tabella 3.8;
- i *condition indices* che derivano dai *singular values*, per il quale viene utilizzato un valore soglia (30 in questo caso) per diagnosticare gruppi di variabili che possono essere in combinazione lineare (seconda colonna della Tabella. 3.8);
- la matrice relativa alla *Variance-decomposition proportion* che indica per ciascun *condition index* il grado di presenza di collinearità.

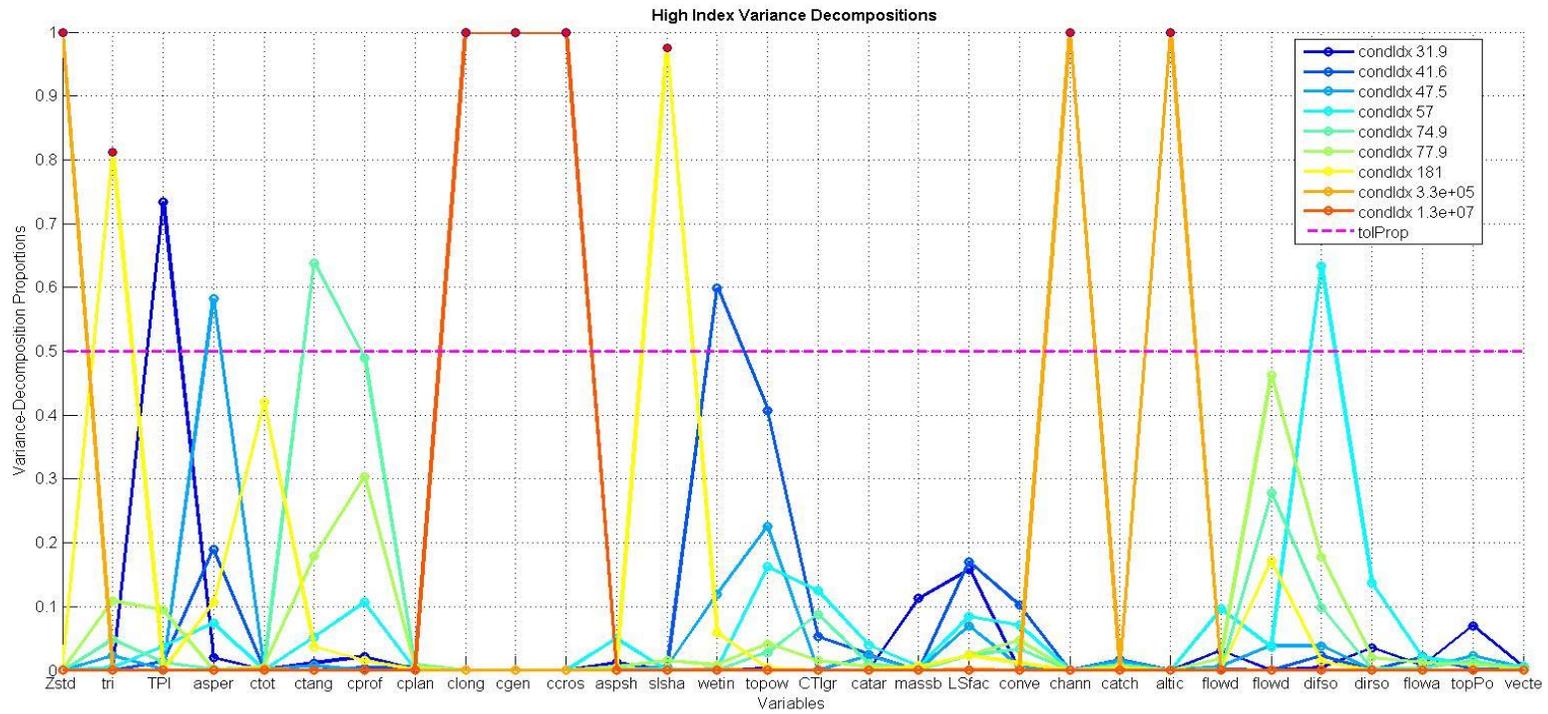
Le variabili in combinazione lineare sono individuate osservando per ciascun *condition index* (che supera il valore soglia di 30) il valore della *Variance-decomposition proportion* (v. Tab. 3.8). Quando in corrispondenza di un *condition index* maggiore di 30 si osservano più variabili con un valore della *Variance-decomposition proportion* che supera quello di soglia (0,5 in questo caso) (evidenziati in rosso nella Tabella 3.8) queste sono in combinazione lineare, proporzionalmente a detto valore (v. Fig. 3.62). Dai risultati emerge come le variabili *General* (cgen), *Longitudinal* (clong) e *Cross-Sectional Curvatures* (ccross) sono collineari, così come *Channel Network Base Level* (channelNB ) e *Altitude Above Channel Network* (altichan). Ugualmente, seppur in misura minore, lo sono *Terrain Ruggedness Index* (TRI) e *slope* (slsharpp). A riguardo è interessante osservare che i coefficienti di correlazione lineare di Pearson in corrispondenza delle tre curvature collineari non risultino essere quelli più alti in assoluto; è emblematico che l'indice di correlazione tra *Cross-Sectional Curvature* e *Longitudinal Curvature* sia pari a 0,46.

**Tabella 3.8** Matrice relativa al test di Belsley: nella prima colonna sono riportati i *singular values*, da cui derivano i *condition indices* riportati nella seconda colonna, ove in grassetto sono indicati quelli che superano il valore soglia. Nelle colonne relative alle variabili sono riportati i valori della *Variance-decomposition proportion*. Quando lungo una riga, in corrispondenza di un *singular values* maggiore di 30, si osservano valori della *Variance-decomposition proportion* maggiori di 0,5 (evidenziati in rosso) le corrispondenti variabili sono in combinazione lineare, proporzionalmente a detto valore.

singular values	condition indices	Zstd	tri	TPI	asperit	ctot	ctang	cprof	cplan	clong	cgen	ccross	aspssharp	slsharpp	wetindex	topowetind
3,769	1,00	1,10E-12	5,11E-06	2,18E-07	3,45E-05	0,00031	1,81E-08	1,68E-07	4,59E-06	5,68E-20	1,36E-20	9,40E-21	0,00068	2,87E-06	6,02E-05	4,61E-05
2,712	1,39	3,78E-14	1,24E-07	0,00028	9,43E-07	0,00061	9,37E-05	7,50E-05	0,00291	3,31E-17	1,78E-17	4,25E-17	7,37E-06	1,01E-07	5,86E-06	9,89E-07
1,306	2,89	5,79E-14	5,37E-06	9,90E-05	3,59E-05	0,00159	0,00042	0,00089	0,04063	2,90E-16	6,52E-18	1,06E-16	0,00015	3,23E-06	0,00014	4,67E-05
1,194	3,16	3,32E-13	3,20E-05	8,95E-07	0,00021	0,00081	0,00019	0,00015	0,01115	4,44E-17	5,12E-20	5,89E-17	0,00272	1,86E-05	0,00066	0,0005
0,956	3,94	1,53E-12	8,70E-06	5,26E-05	4,73E-05	0,01936	3,75E-05	7,07E-08	0,02717	1,19E-18	2,99E-18	1,61E-17	0,00029	4,07E-06	2,77E-05	1,06E-05
0,870	4,33	1,11E-12	1,11E-05	1,15E-05	8,60E-05	0,25282	6,91E-06	3,60E-05	0,0216	1,01E-17	2,64E-19	3,26E-18	0,00028	1,17E-05	3,41E-06	6,96E-07
0,841	4,48	1,25E-12	6,61E-10	8,43E-05	3,17E-06	0,03522	0,00011	5,17E-06	0,00307	1,73E-19	4,57E-18	4,65E-17	0,00254	1,46E-07	6,72E-06	1,95E-06
0,821	4,59	4,68E-12	8,14E-06	0,00015	7,22E-05	0,00166	0,00014	3,75E-07	0,00037	5,89E-21	8,12E-18	7,36E-17	0,01896	4,73E-06	3,32E-05	5,51E-05
0,729	5,17	5,10E-14	2,37E-10	0,00029	7,39E-08	0,03036	0,00112	0,00039	0,4973	9,18E-17	1,46E-17	4,60E-16	0,00319	3,02E-08	0,00011	5,26E-05
0,672	5,61	7,42E-13	1,58E-06	0,00024	1,23E-05	0,02385	0,00024	0,00069	0,0293	2,37E-16	5,51E-18	8,46E-17	0,00515	1,36E-06	0,00036	0,00039
0,502	7,50	5,36E-12	7,24E-06	0,0001	7,72E-05	0,00068	0,0005	0,00022	0,26303	8,75E-18	1,98E-17	1,03E-16	0,00372	5,12E-06	0,00044	1,08E-05
0,480	7,84	2,40E-13	0,00013	7,51E-05	0,00101	0,00153	3,37E-06	2,77E-06	0,00239	7,88E-20	1,96E-19	1,05E-18	0,00031	8,01E-05	0,0013	0,00084
0,458	8,23	3,46E-11	2,29E-05	0,00032	0,0002	0,00234	0,00018	1,16E-05	0,00752	2,80E-18	8,41E-18	4,75E-17	0,50407	1,36E-05	0,00044	0,00013
0,423	8,90	1,49E-11	7,11E-05	0,00048	0,00049	0,00342	0,00018	1,61E-05	0,00958	4,99E-21	6,78E-18	5,92E-17	0,21697	4,08E-05	0,00025	0,00052
0,353	10,68	1,05E-12	2,61E-06	0,00147	4,07E-05	0,15722	0,00202	0,0018	0,0007	7,80E-16	4,19E-16	1,01E-15	0,00011	3,56E-06	6,32E-05	4,04E-06
0,295	12,79	8,23E-10	5,43E-08	7,02E-06	9,61E-06	0,00014	3,11E-05	2,72E-06	2,28E-05	6,87E-18	1,75E-18	1,41E-18	0,0003	3,47E-09	0,00695	0,00364
0,215	17,51	1,61E-11	1,39E-05	0,00013	0,00243	0,00417	5,53E-08	7,38E-05	0,00084	5,12E-17	3,74E-19	3,25E-17	0,12462	2,83E-05	0,03649	0,00974
0,191	19,75	5,85E-15	3,08E-05	0,00517	1,22E-05	0,00583	2,94E-06	0,00231	0,01414	8,12E-16	7,14E-17	3,04E-15	0,0006	1,41E-05	0,01547	0,00103
0,165	22,89	1,78E-12	1,11E-05	0,00016	0,0017	5,32E-05	0,06195	0,04944	0,03531	2,69E-14	1,03E-18	2,87E-14	3,30E-05	1,34E-05	0,00139	0,0022
0,149	25,30	3,16E-12	9,69E-06	0,00153	2,33E-06	0,00224	0,00017	9,01E-05	9,26E-05	1,14E-17	7,58E-18	1,38E-16	0,00143	1,45E-05	0,14191	0,03333
0,142	26,57	1,98E-11	0,00162	0,09274	0,01792	0,01982	0,00573	0,00218	0,00804	1,64E-15	3,48E-16	1,73E-16	0,04467	0,00106	0,00287	0,07541
0,118	<b>31,88</b>	9,39E-12	0,00121	0,73445	0,02019	0,00148	0,01169	0,02196	0,0016	3,89E-15	2,11E-15	5,10E-15	0,012	0,00039	0,00063	0,00323
0,091	<b>41,62</b>	7,42E-12	0,00029	0,01402	0,18904	0,0014	0,0004	0,00431	0,00635	1,85E-16	2,66E-17	8,80E-16	3,10E-05	0,00016	0,59966	0,40728
0,079	<b>47,51</b>	3,43E-11	0,02267	0,00247	0,58287	0,00081	0,00812	5,48E-05	0,00042	1,35E-16	1,46E-16	2,33E-15	3,29E-05	0,0053	0,11951	0,2262
0,066	<b>57,05</b>	2,00E-10	0,00454	0,03746	0,07411	0,00024	0,05221	0,10654	0,00099	1,91E-14	5,83E-15	6,82E-15	0,05022	0,00029	0,0028	0,16253
0,050	<b>74,92</b>	2,71E-11	0,04858	0,01227	0,00193	0,0047	0,63823	0,48904	0,00837	6,15E-14	3,14E-14	7,19E-14	0,00123	0,00153	0,00033	0,02807
0,048	<b>77,86</b>	3,87E-11	0,10905	0,09508	0,00016	0,00539	0,17876	0,30419	0,00585	4,20E-14	1,68E-14	2,94E-14	0,00562	0,01574	0,00861	0,04168
0,021	<b>180,55</b>	3,45E-12	0,81152	0,00034	0,10725	0,4198	0,03747	0,01552	0,00122	4,34E-15	2,08E-15	4,47E-15	2,82E-05	0,97511	0,05867	0,00303
0,000	<b>330064,84</b>	0,99992	0,00014	0,00051	5,88E-05	0,00212	2,57E-07	4,89E-08	1,16E-05	1,52E-15	1,69E-15	1,39E-15	9,96E-06	0,00017	0,0008	4,59E-05
0,000	<b>154357320</b>	7,62E-05	2,15E-08	3,12E-06	5,91E-09	3,67E-05	8,56E-06	1,12E-08	3,82E-05	1	1	1	4,04E-06	2,48E-06	7,04E-07	2,69E-08

**Tabella 3.8** (continua).

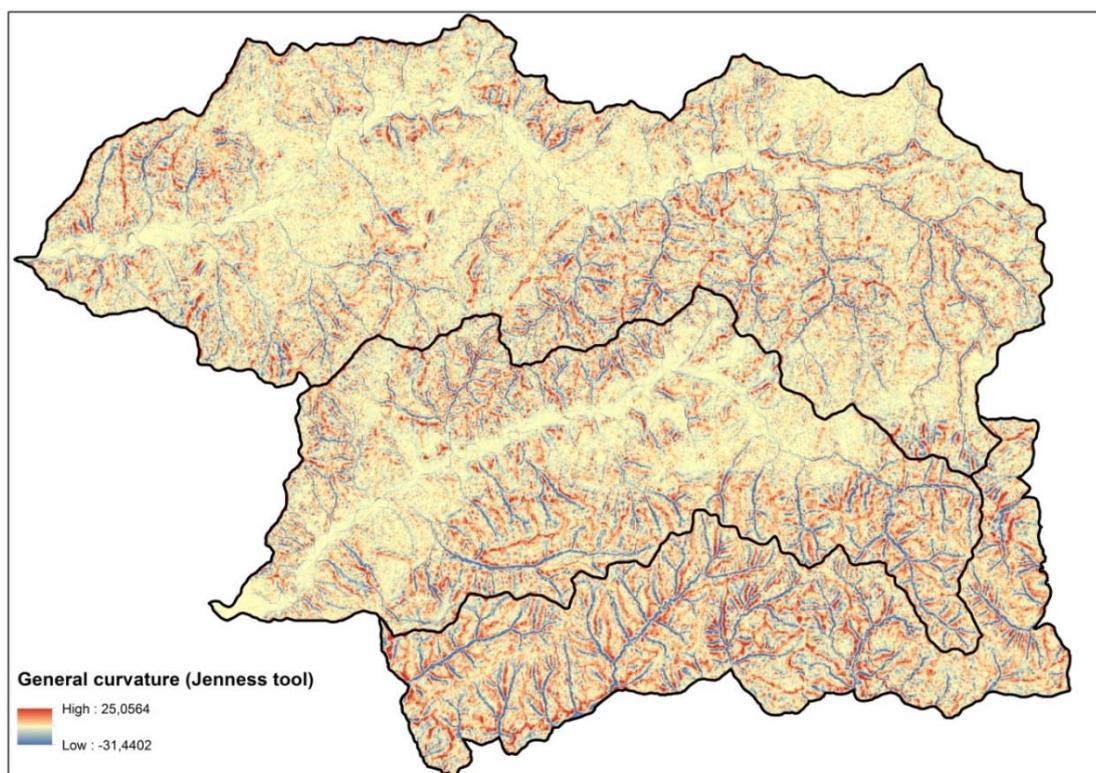
singular values	condition indices	CTIgrass	catarea	massbalan	LSfactor	convergind	channelNB	catchslope	altichan	flowdir	flowdrop	difsol	dirsol	flowaccS	topPosiS	vectorRS
3,769	1,00	9,47E-05	4,59E-05	7,28E-07	8,48E-05	3,36E-06	1,35E-12	0,00017	5,43E-11	0,00052	1,32E-05	2,21E-05	0,0002	0,00044	2,13E-06	0,00077
2,712	1,39	4,96E-06	0,00015	0,00208	2,67E-06	0,00234	2,48E-14	4,29E-08	2,42E-11	2,02E-05	1,79E-06	3,03E-07	4,06E-06	0,00094	0,00296	1,32E-05
1,306	2,89	9,29E-05	0,0049	0,00035	6,65E-05	0,00946	5,40E-14	8,35E-05	1,41E-11	7,60E-05	4,60E-06	4,64E-06	7,42E-05	0,02852	0,00189	0,00314
1,194	3,16	0,00045	0,08345	1,03E-06	0,00054	0,00204	7,29E-13	3,22E-05	8,90E-11	0,00194	0,00011	6,70E-05	0,00081	0,01731	0,00018	0,02271
0,956	3,94	0,00054	0,33879	0,00238	0,00012	0,00387	1,63E-12	4,96E-05	2,04E-10	0,01597	3,16E-05	3,95E-05	0,00019	0,03324	0,00013	0,02069
0,870	4,33	2,08E-06	0,0211	0,00965	0,00021	0,00031	8,22E-13	0,00031	5,43E-10	0,01834	4,28E-05	4,27E-06	3,63E-05	0,00871	0,01877	0,01084
0,841	4,48	1,54E-05	0,03107	0,00323	3,43E-05	3,84E-09	7,30E-13	3,03E-06	9,87E-10	0,15769	1,79E-07	8,14E-06	0,00092	0,2672	0,00369	0,00859
0,821	4,59	8,79E-05	0,00254	0,00144	0,00035	0,00024	3,85E-12	0,00014	1,73E-09	0,15822	3,00E-05	6,27E-05	0,00511	0,20168	0,01693	0,00343
0,729	5,17	0,00016	0,06242	0,00106	0,00014	0,00476	7,77E-14	0,00047	1,18E-09	0,00027	1,54E-06	2,07E-05	0,00042	0,0231	0,00601	0,00292
0,672	5,61	0,00069	0,01591	0,00765	6,26E-05	0,00628	2,14E-13	2,96E-05	8,78E-09	0,00129	3,65E-07	3,89E-05	0,00035	0,18329	0,17732	0,00339
0,502	7,50	6,10E-05	0,01738	0,00053	0,00014	0,52969	3,85E-12	0,00051	2,82E-09	0,00025	3,09E-05	3,07E-08	1,12E-05	0,0349	0,04081	0,02674
0,480	7,84	0,00104	0,0185	0,00149	0,0039	0,00183	1,09E-13	0,00191	3,30E-09	0,03264	0,00042	0,00012	0,00062	0,01152	0,05064	0,78644
0,458	8,23	0,0004	0,00108	0,009	0,00021	0,03933	3,34E-11	4,58E-07	7,28E-09	0,01948	5,64E-05	4,68E-05	0,01364	0,00706	0,12041	0,00787
0,423	8,90	0,00084	3,85E-05	0,02686	0,0007	0,05767	4,68E-12	0,00102	2,47E-08	0,00387	0,00016	0,00015	0,02897	0,00342	0,28266	0,05217
0,353	10,68	2,14E-06	0,00109	0,76773	0,00037	0,00097	3,65E-12	0,00137	1,70E-09	2,40E-05	7,93E-06	5,79E-06	0,00025	0,00605	0,03114	6,30E-05
0,295	12,79	0,00725	0,01253	0,0154	0,00126	0,00493	1,40E-09	0,0064	2,43E-08	0,00377	2,44E-07	3,25E-05	0,01303	0,01492	0,00839	0,00382
0,215	17,51	0,01961	0,02373	0,00071	0,01478	0,0382	2,43E-11	0,06677	1,07E-11	0,32006	0,00021	2,08E-06	0,57161	0,04829	0,09335	0,00496
0,191	19,75	0,01055	0,00076	0,00682	0,11564	0,00032	1,57E-14	0,63795	1,83E-10	0,0048	0,00022	5,52E-05	0,00511	0,01647	8,02E-05	0,00134
0,165	22,89	0,00028	1,36E-07	0,00171	0,00488	0,01377	3,28E-12	0,05135	1,35E-10	0,00015	0,00052	0,00051	0,00051	0,00862	0,00096	4,34E-05
0,149	25,30	0,67594	0,23797	6,87E-07	0,00172	0,01005	4,12E-12	0,00041	6,72E-11	0,0073	0,00023	0,00065	0,01234	0,00433	0,00081	0,00227
0,142	26,57	6,17E-05	0,0219	0,01321	0,29803	0,00184	3,19E-11	0,18199	2,08E-10	0,09835	0,00965	0,00877	0,15059	0,0112	0,0126	0,01578
0,118	<b>31,88</b>	0,00013	4,62E-05	0,11264	0,1586	0,00052	1,49E-11	0,00093	6,21E-11	0,03086	0,00017	0,0049	0,03532	0,00746	0,07045	0,00621
0,091	<b>41,62</b>	0,0532	0,02539	0,00056	0,17043	0,10203	1,18E-11	0,01694	6,13E-11	0,00167	3,48E-06	0,02251	2,01E-05	0,02136	0,00216	0,00055
0,079	<b>47,51</b>	2,90E-05	0,02314	6,48E-05	0,07008	0,00674	4,71E-11	0,0145	2,68E-10	0,00625	0,03879	0,03827	4,89E-05	0,00341	0,02391	0,00542
0,066	<b>57,05</b>	0,12452	0,03987	0,00555	0,08409	0,07078	2,69E-10	0,00035	2,48E-09	0,09624	0,03708	0,63291	0,13738	0,0203	0,00832	0,00749
0,050	<b>74,92</b>	0,08832	0,01051	0,00092	0,02374	0,03344	3,97E-11	0,00859	2,99E-12	0,00094	0,27807	0,09874	0,00263	0,00371	0,01538	7,98E-06
0,048	<b>77,86</b>	0,01522	0,00547	0,00174	0,02472	0,04709	5,32E-11	0,00294	3,07E-10	0,01859	0,46304	0,17813	0,01973	0,01227	0,00965	0,00231
0,021	<b>180,55</b>	0,0002	0,0001	0,0068	0,02376	0,01144	4,04E-12	0,00478	2,61E-10	0,00041	0,17099	0,01391	6,38E-05	0,00027	0,00024	2,19E-06
0,000	<b>330064,84</b>	0,00021	0,00013	0,00043	0,00135	6,08E-05	0,99992	5,74E-06	0,99992	2,62E-06	9,26E-05	1,01E-05	1,06E-06	1,92E-05	0,00015	7,36E-06
0,000	<b>154357320</b>	3,75E-06	2,98E-07	1,61E-10	3,43E-06	2,23E-06	7,62E-05	4,28E-06	7,62E-05	9,50E-07	2,03E-05	2,21E-06	1,83E-07	1,51E-07	4,08E-07	4,22E-06



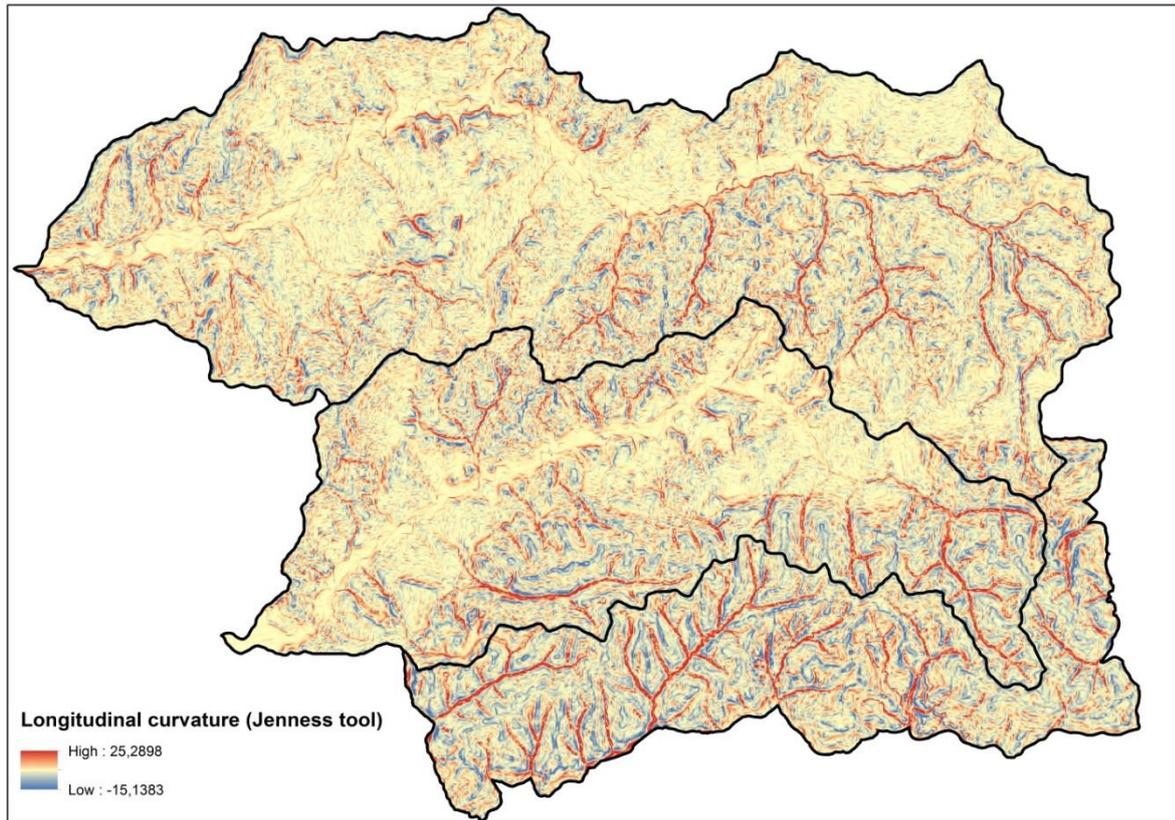
**Fig. 3.62** Rappresentazione dei valori di *Variance-decomposition proportion* per ciascun *singular values* maggiore di 30; emerge la collinearità tra *General* (cgen), *Longitudinal* (clong) e *Cross-Sectional Curvatures* (ccross), oltre a quella tra *Channel Network Base Level* (channelNB ) e *Altitude Above Channel Network* (altichann) e tra *Terrain Ruggedness Index* (TRI) e *slope* (slsharpp).

### 3.4 CONCLUSIONI

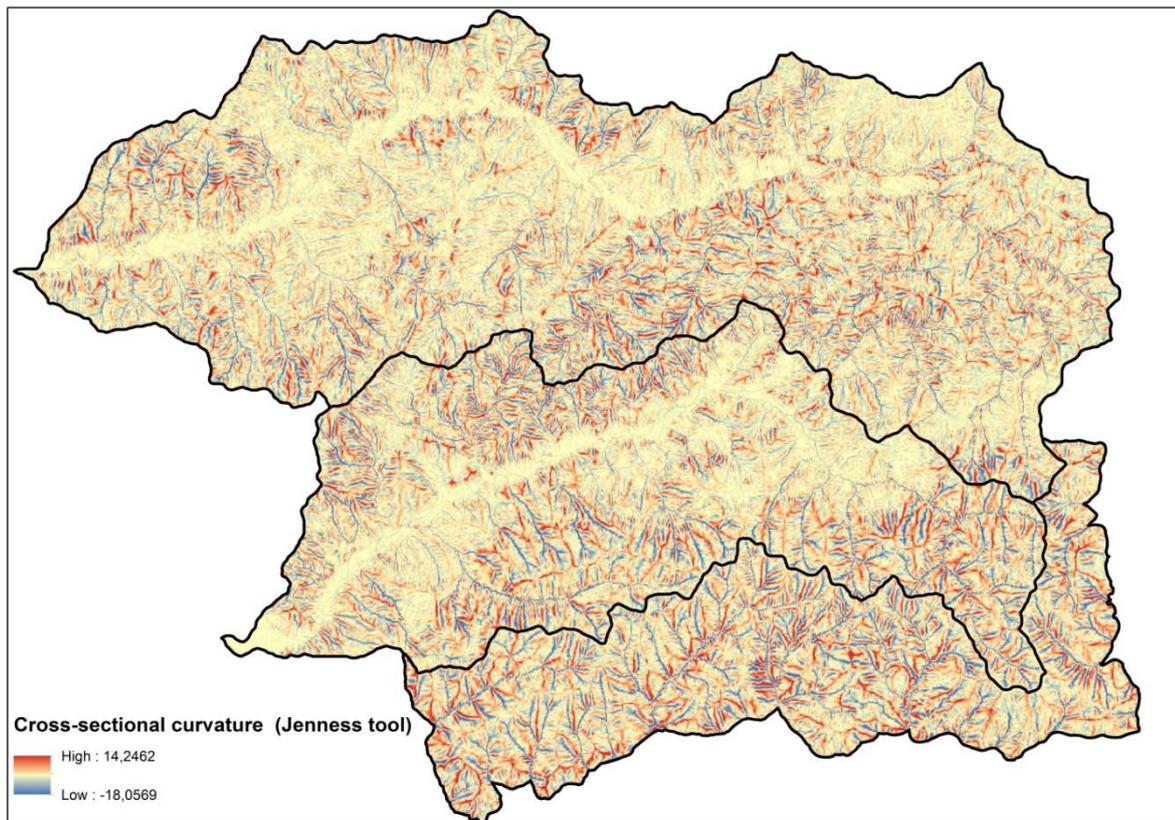
Indipendentemente dall'eventuale inadeguatezza che il DEM veicola nel modellare la superficie terrestre, sembrerebbe che i valori anomali siano connaturati con il calcolo delle variabili morfometriche, perlomeno per territori che presentano una topografia complessa e non siano dovuti a inaccuratezza del DEM. Pur considerando che dette irregolarità sono rilevanti per la fase computazionale dei modelli, e che potrebbero portare a incertezze nell'applicazione dei metodi statistici, gli *outliers* non sono stati esclusi nelle successive applicazioni. Si ritiene che i cosiddetti valori "aberranti" concorrono a riprodurre la variabilità territoriale mentre le particolari distribuzioni di frequenza delle variabili sono legate alla morfologia del territorio. Considerando il comportamento non uniforme di queste ultime, è stato considerato che effettuare una trasformazione delle variabili, al fine di normalizzarle, comporterebbe l'applicazione di metodi variabile-specifici. Questa operazione provocherebbe una deformazione della realtà snaturando il significato delle variabili. Pertanto per le successive elaborazioni non è stata applicata alcuna trasformazione alle variabili ma è stata applicata una scalatura rispetto alle deviazioni standard, al fine da rendere i valori confrontabili e svincolarli dalle unità di misura. Per i casi che esibiscono una collinearità statisticamente verificata prossima a uno, occorre considerare che si tratta di variabili ognuna con un proprio preciso significato morfologico che può fornire una differente versione dell'informazione, riguardo la descrizione del territorio basato sulla geomorfometria (cfr. Fig. 3.63, Fig. 3.64, Fig. 3.65) . Nell'impiego di questa tipologia di predittori, per la modellazione statistica, occorre considerare quali sono le assunzioni su cui si basa il metodo da utilizzare, evitando i metodi come la regressione lineare e prendendo invece in considerazione metodi non parametrici come le tecniche di machine learning o metodi che consentono l'uso di variabili che si discostano non poco dalla normalità come la *discriminant analysis*.



**Fig. 3.63** Rappresentazione della variabile *general curvature* (Jenness tool).



**Fig. 3.64** Rappresentazione della variabile *longitudinal curvature* (Jenness tool).



**Fig. 3.65** Rappresentazione della variabile *cross-sectional curvature* (Jenness tool).

## Bibliografia

- Behrens, Thorsten; Förster, Helga; Scholten, Thomas; Steinrücken, Ulrich; Spies, Ernst-Dieter; Goldschmitt, Michael (2005): Digital soil mapping using artificial neural networks. In: *Z. Pflanzenernähr. Bodenk.* n. 1, 168, pp. 21–33. DOI: 10.1002/jpln.200421414.
- Belsley, David A. (1991): Conditioning diagnostics: Wiley Online Library.
- Evans, Ian S. (2012): Geomorphometry and landform mapping. What is a landform? In: *Geomorphology* n. 1, 137, pp. 94–106. DOI: 10.1016/j.geomorph.2010.09.029.
- Florinsky, Igor V. (2012): Digital Terrain Analysis in Soil Science and Geology: Elsevier.
- Gallant, John C. (2000): Terrain analysis: principles and applications: John Wiley & Sons.
- Gruber, S.; Peckham, S. (2009): Land-surface parameters and objects in hydrology. In: *Developments in Soil Science*, 33, pp. 171–194.
- Hengl, Tomislav; Reuter, Hannes I. (2009): Geomorphometry. Concepts, software, applications. 1st ed. Amsterdam Netherlands, Oxford UK, Boston Mass.: Elsevier (Developments in soil science, 33).
- Jenness, J. (2011): Tools for graphics and shapes: Extension for ArcGIS. Jenness Enterprises.
- Jenness, Jeff (s.d) : DEM Surface Tools for ArcGIS.
- Möller, Markus; Volk, Martin; Friedrich, Klaus; Lymburner, Leo (2008): Placing soil-genesis and transport processes into a landscape context: A multiscale terrain-analysis approach. In: *Journal of Plant Nutrition and Soil Science* n. 3, 171, pp. 419–430.
- Moore, Ian D.; Gessler, P. E.; Nielsen, G. A.; Peterson, G. A. (1993): Soil attribute prediction using terrain analysis. In: *Soil Science Society of America Journal* n. 2, 57, pp. 443–452.
- Moore, Ian Donald; Grayson, R. B.; Ladson, A. R. (1991): Digital terrain modelling: a review of hydrological, geomorphological, and biological applications. In: *Hydrological processes* n. 1, 5, pp. 3–30.
- Olaya, V.; Conrad, O. (2009): Geomorphometry in SAGA. In: *Developments in Soil Science*, 33, pp. 293–308.
- Osborne, Jason; Waters, Elaine (2002): Four assumptions of multiple regression that researchers should always test. In: *Practical assessment, research & evaluation* n. 2, 8, pp. 1–9.
- Pike, R. J.; Evans, I. S.; Hengl, T. (2009): Geomorphometry: a brief guide. In: *Developments in Soil Science*, 33, pp. 3–30.
- Porres de la Haza, M; Pardo Pascual, J.; Sobrino, J. A. (2002): Comparison between the different curvature models of terrain for determining the degree of soil humidity. In: *Recent Advances in Quantitative Remote Sensing*, p. 238.
- Shary, Peter A.; Sharaya, Larisa S.; Mitusov, Andrew V. (2002): Fundamental quantitative methods of land surface analysis. In: *Geoderma* n. 1-2, 107, pp. 1–32. DOI: 10.1016/S0016-7061(01)00136-7.
- Sörensen, Rasmus; Zinko, Ursula; Seibert, Jan (2006): On the calculation of the topographic wetness index: evaluation of different methods based on field observations. In: *Hydrology and Earth System Sciences Discussions* n. 1, 10, pp. 101–112.
- Stewart, Gilbert W. (1987): Collinearity and least squares regression. In: *Statistical Science*, pp. 68–84.
- Wilson, Margaret F. J.; O’Connell, Brian; Brown, Colin; Guinan, Janine C.; Grehan, Anthony J. (2007): Multiscale Terrain Analysis of Multibeam Bathymetry Data for Habitat Mapping on the Continental Slope. In: *Marine Geodesy* n. 1-2, 30, pp. 3–35. DOI: 10.1080/01490410701295962.

Zawawi, Azita Ahmad; Shiba, Masami; Jemali, Noor Janatun Naim (2014): Landform classification for site evaluation and forest planning: integration between scientific approach and traditional concept. In: *Sains Malaysiana* n. 3, 43, pp. 349–358.

Zhao, Guang-ju; Gao, Jun-feng; Tian, Peng; Tian, Kun (2009): Comparison of two different methods for determining flow direction in catchment hydrological modeling. In: *Water Science and Engineering* n. 4, 2, pp. 1–15.

## CAPITOLO 4

### Progettazione dell'attività di rilevamento

#### 4.1 Introduzione

Indipendentemente da ciò che viene rappresentato in una mappa, l'acquisizione dei dati in campo è una attività dalla quale non si può prescindere per la cartografia nelle Scienze della Terra. Questo vale in particolare per la realizzazione di mappe tematiche relative al suolo, utilizzando i metodi contemplati nella disciplina del *Digital Soil Mapping*. “*In digital soil mapping (DSM), the prediction of soil properties and soil classes is based on forming relationships between observed soil attributes and ancillary soil and environmental variables*” (McBratney et al. 2003). Nonostante i progressi nelle tecnologie di telerilevamento, delle tecniche spettroscopiche, radiometriche e delle loro applicazioni, e l'incremento della disponibilità di informazione ausiliaria, tipicamente utilizzata nella cartografia tradizionale, rimane la necessità di dover procedere al rilevamento in campo per la raccolta di un campione di suolo, in un determinato punto del territorio, al fine di poterne determinare le caratteristiche. Queste raramente sono deducibili a priori, sulla base delle sole caratteristiche territoriali o paesaggistiche, oppure basandosi sull'informazione sintetizzata in mappe ausiliarie alla conoscenza del suolo. La realizzazione di una cartografia di dettaglio relativa al suolo, non può essere delegata ad una semplice combinazione di mappe tematiche relative a variabili ambientali o a *proxy* dei fattori dell'equazione di Jenny, anche se di alta qualità. Inoltre, occorre considerare che l'informazione territoriale disponibile, seppur copiosa, non è detto che sia di scala adeguata o rappresentativa dello stato delle cose al momento della mappatura.

Poiché l'attività di campionamento è il punto di partenza per giungere alla conoscenza della distribuzione spaziale delle caratteristiche dei suoli, il suo svolgimento è ciò che più pesantemente condiziona la qualità finale dei prodotti cartografici, sia nella cartografia tradizionale che nel DSM. Lagacheire ha evidenziato che “*It was largely recognised that a good soil dataset is a key factor to build an accurate DSM function and to evaluate the quality of its outputs [...]. The collection of soil data has been (and still remain) a limiting factor that can severely brake the Digital Soil Mapping progresses*” (Lagacherie 2008).

Nonostante i progressi nella scienza del suolo ed in particolare nel *Digital Soil Mapping* (McBratney et al. 2003; Grunwald 2009; Brevik et al. 2016) attualmente l'attività di campionamento è rimasta sostanzialmente invariata rispetto a quella ereditata dalla cartografia tradizionale.

Dato l'elevato costo sia in termini di tempo che monetari è conveniente utilizzare strategie per organizzare preventivamente detta fondamentale attività, con l'obiettivo di operare con efficienza. Operare in modo efficace significa ottenere un duplice obiettivo: minimizzare i costi e nel contempo assicurarsi che la variabilità studiata sia rappresentata adeguatamente (Minasny e McBratney 2006). Tale caratterizzazione deve, inoltre, essere in grado di fornire informazioni coerenti con il livello di accuratezza e precisione richiesto per il prodotto finale (Hengl et al. 2003). Oltre a ciò, per applicare le tecniche del DSM, Minasny e McBratney hanno evidenziato che: “*The purpose of sampling is to obtain data that enable the estimation of some statistical parameter, or spatial predictions of some properties over an area*” (Minasny e McBratney 2006). Questa necessità ha assunto una particolare rilevanza con le tecniche di cartografia basate su metodi statistici. Infatti, la qualità di una mappa

riportante la distribuzione spaziale delle caratteristiche dei suoli, siano esse categoriali o quantitative, è condizionata dal tipo di tecnica statistica applicata nel processo di inferenza utilizzato per derivare la mappa e dalla distribuzione e dalla significatività dei punti di campionamento nel dominio studiato, sia esso geografico che iperspaziale.

Nell'attività tradizionale di campionamento, chiamata da White *free survey* (White 2013), i punti da rilevare sono scelti dall'operatore che, sulla base della propria esperienza, individua i punti ove effettuare il campionamento, seguendo un proprio modello mentale circa la distribuzione della variabilità spaziale dei suoli. Adottando tale metodo soggettivo e non casuale è possibile introdurre un *bias* o informazione dannosa per i metodi statistici utilizzati per la spazializzazione dei dati (Hengl et al. 2003; Brungard e Boettinger 2010) Di contro, in statistica, la teoria classica assume che i dati derivanti dal campionamento siano identicamente e indipendentemente distribuiti.

La progettazione dell'attività di campionamento consiste in una programmazione teorica e preventiva volta all'individuazione dei punti che devono essere visitati per esplorare e caratterizzare i suoli. La determinazione della posizione geografica di un punto (o un transetto) oggetto del campionamento avviene utilizzando algoritmi, che possono prescindere dalla conoscenza del territorio. Tale programmazione include concetti quali: intensità dei punti di campionamento (dimensione della griglia), variabilità spaziale, scala della mappa da ottenere (Hengl et al. 2003).

In letteratura vi sono alcune applicazioni per la progettazione dell'attività di campionamento che Brus e Gruijter hanno distinto in: metodi *design-based*, che utilizzano la teoria classica dei disegni sperimentali e *model-based*, che utilizzano la geostatistica e le sue assunzioni (Brus e Gruijter 1997). In generale gli approcci mirano ad ottenere una ottimale copertura dello spazio geografico (Royle e Nychka 1998; Brus et al. 2006) o la copertura dello spazio geografico e delle variabili ausiliarie considerate (Heuvelink et al. 2006; Dobermann e Simbahan 2006; Brungard e Boettinger 2010; Hengl et al. 2003; Minasny e McBratney 2006).

Riguardo le applicazioni esistenti in letteratura, Lagacheire ha evidenziato che: "*They do not take into account more sophisticated sampling criteria that are often considered in classical soil surveying e.g. the optimisation of real field costs of soil data collection (e.g. accessibility) or the use of a priori hypothesis about the expected patterns of soil variations (e.g. introduction of sampling density constraints to capture an expected short-range variability, identifications of representative transects or areas)*" (Lagacherie 2008).

Di fatto l'approccio teorico-statistico, che predetermina la posizione dei punti di campionamento, sia esso basato sui disegni sperimentali che sulla geostatistica, può generare alcuni problemi al momento della sua applicazione. Certamente la non accessibilità a un punto, predeterminato teoricamente, è una delle cause ostative che più facilmente si possono incontrare nel corso dell'esecuzione dell'attività di campionamento, specialmente in un ambiente antropizzato. Inoltre, indipendentemente dal fatto che si proceda con metodi di campionamento *design-based* o *model-based*, campionare con punti prefissati nello spazio geografico comporta un vincolo che presenta ulteriori svantaggi, *in primis* individuare l'esatto punto dove effettuare il campionamento, sulla base delle coordinate, può risultare arduo. L'errore strumentale di un GPS o semplicemente la copertura

vegetale possono determinare errori che possono superare la dimensione della griglia adottata, introducendo una imprecisione a priori, che potrebbe compromettere il risultato finale dell'elaborazione statistica, utilizzata per derivare la mappa. A seconda della dimensione del territorio e di quella della griglia adottata per discretizzarlo, si può facilmente incorrere nel problema di campionare in un luogo diverso da quello predefinito, nel migliore dei casi in un pixel adiacente a quello che realmente andava campionato.

L'impiego dell'approccio geostatistico presenta ulteriori inconvenienti, come hanno evidenziato i vari autori che hanno creato metodi per progettare l'attività di campionamento. Infatti, mentre Hengl ha affermato che un'allocatione dei punti di campionamento in modo uniforme è auspicabile quando non si conosce il modello di regressione, per dedurre le caratteristiche dei suoli dai predittori (Hengl et al. 2003), Gerard B.M. ha sottolineato che: *“In practice, regular grid sampling may be suboptimal for several reasons. Sampling may be hampered by enclosures that are inaccessible for sampling. The study area may have an irregular shape or the dimensions of the area may be small compared to the grid spacing. The latter may result in prediction error variances that are relatively large near the boundaries of the area. Also, we may want to use previously made observations at locations that cannot be matched with the grid. In these situations, minimization of the kriging variance criterion will yield a sample with irregular spacing.”* (Heuvelink et al. 2006). Riguardo l'applicazione della geostatistica, è stato inoltre evidenziato che: *“Direct observations of the soil are important for two main reasons. First, they are used to establish the character and strength of the relationship between the soil property of interest and the auxiliary information. Second, they are used to improve the predictions based on the auxiliary information, by spatial interpolation of the differences between the observations and predictions. However, the two uses of the soil observations generally impose conflicting requirements on the sample configuration. Estimation of the relationship between the soil property and the auxiliary information benefits from a large spread of the observations in feature space, while spatial interpolation of the differences between observations and predictions gains from a uniform spreading of the observations in geographic space”* (Heuvelink et al. 2006).

Inoltre, per tale approccio, occorre un numero minimo di punti di campionamento per il quale abbia senso applicare il metodo (Webster e Oliver 2001) e su aree vaste tale numero potrebbe raggiungere un numero inapplicabile, dal lato pratico, indipendentemente dal parametro da mappare. Inoltre, per questi casi sarebbe ancor più difficoltoso l'impiego di griglie di campionamento regolari.

Al fine di ovviare alle limitazioni testé argomentate, è stata cercata un'alternativa più semplice dal punto di vista concettuale e computazionale. Il metodo per la progettazione dell'attività di campionamento, presentato in questo lavoro, ha portato alla costruzione di una mappa che è uno strumento per la programmazione dell'attività di campionamento e nel contempo è di ausilio per la sua l'organizzazione e il suo svolgimento pratico.

Lo strumento è stato pensato con l'obiettivo di raccogliere una informazione statisticamente sufficiente per elaborare mappe tematiche relative al suolo con le tecniche di *digital soil mapping*. Un ulteriore obiettivo è stato quello di verificare se un numero contenuto di punti di campionamento fosse in grado di fornire un'informazione sufficiente per le elaborazioni cartografiche.

L'approccio ha come base teorica il paradigma suolo-paesaggio di Hudson (v. Cap.1) e come criterio per la sua realizzazione la *landform segmentation*, basata sulle caratteristiche topografiche. La *landform segmentation* permette di discretizzare il territorio in *spatial units*, caratterizzate dal possedere attributi topografici simili (Pennock e Corre 2001; Pennock 2003; Martin e Timmer 2006; Stepinski e Bagaria 2009). Gli attributi topografici e le proprietà della superficie terrestre possono essere descritte da un *set* di misure numeriche che possono essere derivate dal DEM (Stepinski e Bagaria 2009).

Nel caso del paradigma di Hudson il paesaggio è contestualizzabile considerando la nota equazione di Jenny. Per Hudson il paesaggio è l'espressione dell'interazione dei fattori naturali che concorrono allo sviluppo del processo pedogenetico e considera che la distribuzione spaziale del suolo presenta una variabilità continua con la presenza, a volte, di discontinuità. La distribuzione spaziale dei suoli può essere associata a strutture territoriali individuabili, chiamate da Hudson *soil-landscape units*. Queste, tuttavia, non dovrebbero essere associate, o addirittura confuse, con la *spatial units*, tipicamente utilizzate nella *landform segmentation*.

Secondo l'approccio qualitativo di Hudson, il paesaggio coincide con l'individuazione di *patches* ove l'interazione dei fattori naturali che concorrono alla pedogenesi è omogenea, tenendo presente che tali fattori hanno un peso diverso a seconda dell'area geografica considerata. Diversamente, l'individuazione di forme simili, secondo l'approccio quantitativo tipico della *landform segmentation*, è legata a fattori geologico-strutturali e genetico-evolutivi del territorio, che concorrono a determinare la geomorfologia di quest'ultimo.

La differenza tra l'individuazione di una *spatial unit* e una *patch* corrispondente ad una *soil-landscape unit* (Hudson *sensu*) sta nella diversità della dimensione dei relativi iperspazi, determinati dalle variabili che vengono considerate per la rappresentazione della realtà. In altre parole, le variabili che concorrono all'individuazione dei due differenti oggetti, sul territorio, sono diverse sia nel numero sia nel tipo, anche se alcune di esse possono essere presenti in entrambi i casi.

Nell'approccio proposto, pur non escludendo l'uso di variabili ausiliarie relative a temi ambientali, il territorio è suddiviso in unità, basate sullo studio topografico del territorio, come nella *landform segmentation*: in questo caso, però, ciò che si è cercato non sono le forme del territorio (es: cresta, versante, fondovalle): di fatto la segmentazione del territorio effettuata su base topografica mira a suddividere il territorio in unità che al loro interno sono uniformi per una serie di caratteristiche. Tale insieme di caratteristiche, dedotto per via computazionale, è associabile a variabili geomorfometriche primarie che concorrono a descrivere la forma del territorio e a variabili secondarie che sono la sintesi di un insieme di caratteristiche del territorio (a queste ultime sono associabili i fenomeni fisici legati al flusso di massa e di energia, oppure i fattori della pedogenesi presenti nell'equazione di Jenny). Complessivamente, dunque, l'iperspazio determinato da dette variabili concorre a delineare un *proxy* del corrispondente paesaggio multidimensionale di Hudson e, di conseguenza, a spiegare la relazione tra suolo e variabilità del territorio e quindi la distribuzione spaziale delle caratteristiche dei suoli, rispetto allo spazio geografico. In sostanza, l'approccio presentato è basato sull'impiego delle variabili primarie e secondarie, calcolate a partire

dal DEM, mediante le quali si procede alla suddivisione del territorio in aree omogenee, dette *unità geomorfometriche*.

## 4.2 Materiali e metodi

Precedentemente all'attività di campionamento svolta utilizzando il metodo presentato in questo capitolo, è stata eseguita un'attività esplorativa nell'area di studio che ha portato, tra l'altro, al rilevamento di 63 profili pedologici. Detta attività di campionamento è stata appositamente svolta con il metodo *free survey*, con l'intenzione di costituire un *dataset* di validazione per le successive elaborazioni, finalizzate alla realizzazione di mappe tematiche di dettaglio. In questo caso, pur essendoci una pregressa conoscenza riguardo l'area, la progettazione dell'attività di campionamento è stata svolta come se dette conoscenze non esistessero.

La base informativa è un DEM discretizzato a 10 m, con il quale sono state calcolate 30 variabili morfometriche primarie e secondarie con i *tools* implementati nei *softwares* GIS sia *open source* che proprietari (v. Cap. 3). Sono state calcolate le variabili geomorfometriche indicate in tabella 4.1, ottenendo così un tensore costituito da 30 *rasters*, uno per ciascuna variabile geomorfometrica calcolata. In altre parole, ciascun punto del territorio è associato a una tessera di dimensione 10x10 m, le cui coordinate sono costituite da un *array* di 30 elementi. La matrice, costituita dai 917.147 casi in cui è stato discretizzato il territorio (righe) e da 30 variabili geomorfometriche (colonne), è l'oggetto elaborato con il software MatLab. Occorre sottolineare che per ciascun *pixel*, l'informazione veicolata dalle variabili - anche se in alcuni casi può sembrare ridondante, per lo meno dal punto di vista del nome - è costituita da specifici valori numerici tutti diversi tra loro (v. Cap.3).

**Tabella 4.1** Denominazione e sigla delle variabili geomorfometriche utilizzate.

	nome variabile	sigla		nome variabile	sigla
1	elevation	Zstd	16	Compound Topographic Index (GRASS)	CTIgrass
2	Terrain Ruggedness Index (QGIS)	TRI	17	catchment Area (recursive method) (SAGA)	catarea
3	Topographic position index (QGIS)	TPI	18	mass balance index (SAGA)	massbal
4	Asperity (QGIS)	asperit	19	LS factor (SAGA)	LSfactor
5	total curvature (Jenness tool)	ctot	20	convergence index (SAGA)	ConvInd
6	tangential curvature (Jenness tool)	ctang	21	channel network base level (SAGA)	channelNB
7	profile curvature (Jenness tool)	cprof	22	catchment Area (max flux method) (SAGA)	catchslope
8	plan curvature (Jenness tool)	cplan	23	altitude above channel network (SAGA)	altichan
9	longitudinal curvature (Jenness tool)	clong	24	flow direction (ArcGIS)	flowdir
10	general curvature (Jenness tool)	cgen	25	flow drop (ArcGIS)	flowdrop
11	cross-sectional curvature (Jenness tool)	ccross	26	diffuse insolation (SAGA)	difsol
12	aspect (Jenness tool)	aspsharp	27	direct insolation (SAGA)	dirsol
13	slope (Jenness tool)	slsharp	28	flow accumulation (SAGA)	flowaccS
14	wetness index (SAGA)	WI	29	Topographic position index (SAGA)	topPosiS
15	topographic wetness index (SAGA)	TWI	30	Vector Ruggedness Measure (SAGA)	VRM

Le variabili sono state uniformate rispetto alla corrispondente deviazione standard, per svincolarle dalle unità di misura derivanti dai rispettivi algoritmi di calcolo. E' stato scelto questo tipo di standardizzazione, rispetto a quello con autoscalatura, tipicamente utilizzato per normalizzare i dati, al fine di evitare la variazione del segno dei valori numerici: quest'ultimo, infatti, a seconda dell'algoritmo assume un significato fisico specifico (es: concavità) (v. Cap. 3). Sulla matrice standardizzata del *dataset* è stata effettuata l'analisi delle componenti principali, al fine di valutare la relazione tra le variabili e la loro rilevanza. Con l'analisi delle componenti principali, i predittori originali vengono trasformati in nuove variabili con un sistema di riferimento ruotato rispetto al quello originario, in modo tale che i nuovi assi siano orientati rispetto alla massima varianza dei dati e ortogonali tra loro al fine ottenere nuove variabili non correlate tra loro. La procedura comporta la diagonalizzazione della matrice di covarianza del set di dati. Detta operazione porta alla determinazione di una matrice diagonale degli autovalori e di una matrice dei *loadings*, le cui colonne sono gli autovettori della matrice di covarianza e le cui righe corrispondono alle variabili originali. I valori dei *loadings* sono compresi tra -1 e +1 e forniscono indicazione di quanto una variabile originale è importante per rappresentare ciascuna componente principale. Un valore nei *loadings*, relativo ad una variabile, vicino a 1 in valore assoluto nel nuovo sistema di riferimento, indica che detta variabile è rappresentata da tale componente (v. Tab. 4.2).

Gli elementi della diagonale della matrice degli autovalori forniscono la misura della varianza associata a ciascuna componente principale o autovettore della matrice di covarianza.

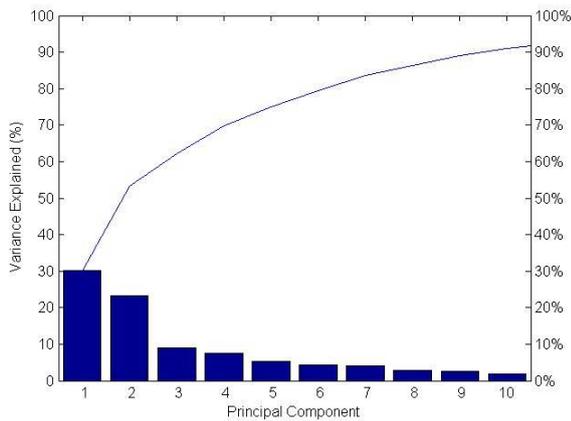
Esistono diversi metodi di *rank analysis*, sia qualitativi che analitici, per valutare il numero di componenti da ritenere. Un metodo qualitativo consiste nell'osservazione dello *scree plot* (autovalori *versus* numero delle componenti) (Fig. 4.2): nel caso in esame si osservano le più importanti variazioni della varianza cumulata spiegata in corrispondenza della quarta componente principale e successivamente intorno alla settima. Di fatto le prime 4 componenti principali spiegano il 68,17% della varianza dell'intero *dataset*, le prime 7 componenti spiegano l'81,45%, mentre si arriva al 90% con le prime 10 componenti (v. Fig. 4.1 e Tab 4.3).

**Tab. 4.2** Valori dei *loadings* delle variabili geomorfometriche, distinti secondo l'importanza (intensità del colore) ed il segno (verdi negativi, rossi positivi). Per ciascuna componente è riportato l'autovalore (*eigenvalue*) e la percentuale di varianza spiegata (*percent\_explained*).

		PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12	PC13	PC14	PC15
<b>percent_explained</b>		<b>30,056</b>	<b>23,452</b>	<b>8,688</b>	<b>5,978</b>	<b>5,528</b>	<b>4,335</b>	<b>3,414</b>	<b>3,311</b>	<b>2,879</b>	<b>2,284</b>	<b>1,867</b>	<b>1,544</b>	<b>1,327</b>	<b>0,948</b>	<b>0,920</b>
<b>eigenvalue</b>		<b>9,017</b>	<b>7,036</b>	<b>2,606</b>	<b>1,793</b>	<b>1,658</b>	<b>1,301</b>	<b>1,024</b>	<b>0,993</b>	<b>0,864</b>	<b>0,685</b>	<b>0,560</b>	<b>0,463</b>	<b>0,398</b>	<b>0,284</b>	<b>0,276</b>
1	elevation	0,110	0,115	0,528	0,014	-0,027	0,191	-0,035	0,088	-0,103	-0,057	-0,008	-0,009	-0,006	0,010	-0,014
2	Terrain Ruggedness Index (QGIS)	0,141	0,336	-0,050	-0,042	-0,011	0,028	-0,013	0,006	-0,007	0,013	-0,008	-0,080	-0,033	0,017	-0,006
3	Topographic position index (QGIS)	0,285	-0,156	-0,070	-0,150	0,020	0,088	-0,037	0,081	0,032	-0,087	-0,058	-0,013	0,083	-0,033	0,015
4	Asperity (QGIS)	0,142	0,330	-0,051	-0,040	0,004	0,021	-0,014	0,009	-0,030	0,017	-0,020	-0,114	-0,043	0,098	0,008
5	total curvature (Jenness tool)	-0,094	0,123	0,103	0,083	-0,027	0,181	-0,018	-0,154	0,881	0,018	0,165	-0,141	-0,014	-0,041	0,042
6	tangential curvature (Jenness tool)	0,249	-0,126	-0,133	0,260	-0,025	0,206	-0,085	0,142	0,060	0,008	-0,291	0,062	0,101	0,040	-0,062
7	profile curvature (Jenness tool)	-0,213	0,129	-0,028	0,474	-0,056	0,077	-0,029	0,018	0,016	0,105	-0,222	0,073	-0,009	0,056	-0,070
8	plan curvature (Jenness tool)	0,170	-0,088	-0,074	0,336	-0,031	0,056	-0,103	0,266	-0,114	0,266	0,680	-0,172	0,090	0,071	-0,349
9	longitudinal curvature (Jenness tool)	-0,228	0,135	-0,011	0,441	-0,048	0,048	-0,021	0,013	0,003	0,112	-0,183	0,069	-0,039	0,055	0,017
10	general curvature (Jenness tool)	0,286	-0,157	-0,066	-0,152	0,021	0,086	-0,036	0,076	0,037	-0,068	-0,059	-0,007	0,086	0,003	-0,022
11	cross-sectional curvature (Jenness tool)	0,261	-0,133	-0,125	0,189	-0,013	0,199	-0,084	0,145	0,067	-0,003	-0,290	0,059	0,109	0,062	-0,021
12	aspect (Jenness tool)	-0,010	-0,010	-0,030	-0,031	-0,283	0,111	0,839	0,304	0,039	0,033	-0,019	-0,068	-0,042	0,283	0,093
13	slope (Jenness tool)	0,150	0,330	-0,059	-0,040	-0,008	0,014	-0,013	0,016	-0,067	0,013	-0,014	-0,081	-0,039	0,026	-0,007
14	wetness index (SAGA)	-0,267	-0,069	-0,047	-0,200	0,024	0,276	-0,060	0,073	0,017	0,027	-0,025	-0,001	0,296	0,127	-0,282
15	topographic wetness index (SAGA)	-0,246	-0,113	-0,119	-0,110	0,010	0,367	-0,052	0,012	-0,076	-0,028	0,101	0,003	0,200	0,018	-0,025
16	Compound Topographic Index (GRASS)	-0,252	-0,012	0,024	-0,135	0,044	-0,121	-0,105	0,363	0,053	-0,035	0,052	-0,058	0,489	0,129	0,293
17	catchment Area (recursive method) (SAGA)	-0,082	-0,059	-0,124	-0,089	-0,019	0,655	0,029	-0,391	-0,234	0,174	0,096	-0,094	-0,164	-0,021	0,236
18	mass balance index (SAGA)	0,267	-0,135	-0,035	-0,113	0,014	0,134	-0,036	0,031	0,268	-0,113	-0,021	-0,040	0,075	-0,040	0,044
19	LS factor (SAGA)	0,031	0,353	-0,036	-0,073	0,004	0,049	-0,042	0,084	0,027	0,029	0,131	-0,177	0,069	0,019	0,034
20	convergence index (SAGA)	0,242	-0,128	-0,069	0,263	-0,016	0,025	-0,047	0,141	-0,023	0,027	0,175	-0,004	-0,075	-0,160	0,617
21	channel network base level (SAGA)	0,075	0,118	0,503	0,027	-0,035	0,249	-0,073	0,177	-0,102	-0,264	0,017	-0,010	-0,144	0,037	-0,073
22	catchment Area (max flux method) (SAGA)	0,042	0,308	-0,102	0,015	-0,005	0,148	-0,055	0,125	-0,074	-0,011	-0,284	-0,161	0,288	-0,166	-0,076
23	altitude above channel network (SAGA)	0,162	0,042	0,322	-0,033	0,014	-0,097	0,106	-0,248	-0,048	0,643	-0,085	-0,002	0,441	-0,082	0,185
24	flow direction (ArcGIS)	0,014	0,006	0,029	0,076	0,716	0,017	-0,016	-0,021	-0,001	-0,012	-0,020	-0,087	-0,064	0,614	0,178
25	flow drop (ArcGIS)	0,192	0,293	-0,052	-0,132	0,003	0,032	-0,011	0,000	-0,024	-0,010	0,030	-0,103	-0,026	0,027	-0,013
26	diffuse insolation (SAGA)	-0,046	-0,214	0,486	0,062	-0,010	0,082	-0,010	0,060	-0,006	-0,055	-0,019	0,034	0,034	-0,036	0,013
27	direct insolation (SAGA)	-0,010	-0,076	0,015	-0,066	-0,615	-0,108	-0,405	-0,111	-0,028	0,008	-0,062	-0,212	-0,043	0,544	0,175
28	flow accumulation (SAGA)	-0,145	-0,001	-0,007	-0,307	0,054	0,079	-0,227	0,526	0,105	0,519	-0,140	0,115	-0,457	-0,092	0,056
29	Topographic position index (SAGA)	0,246	-0,145	0,062	-0,094	0,001	-0,059	0,065	-0,175	0,110	0,289	-0,112	0,079	-0,092	0,269	-0,367
30	Vector Ruggedness Measure (SAGA)	0,096	0,257	-0,049	-0,057	-0,107	0,100	-0,049	-0,025	0,050	-0,037	0,232	0,865	0,124	0,195	0,082

Tab. 4.2 (segue).

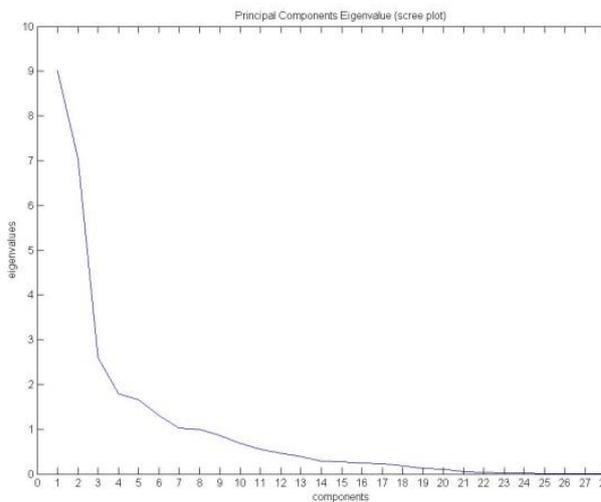
		PC16	PC17	PC18	PC19	PC20	PC21	PC22	PC23	PC24	PC25	PC26	PC27	PC28	PC29	PC30
percent explained		<b>0,818</b>	<b>0,744</b>	<b>0,628</b>	<b>0,409</b>	<b>0,371</b>	<b>0,185</b>	<b>0,104</b>	<b>0,085</b>	<b>0,053</b>	<b>0,032</b>	<b>0,027</b>	<b>0,009</b>	<b>0,006</b>	<b>0,000</b>	<b>0,000</b>
eigenvalue		0,245	0,223	0,188	0,123	0,111	0,055	0,031	0,025	0,016	0,010	0,008	0,003	0,002	1,71E-09	8,39E-14
1	elevation	0,003	-0,067	0,047	-0,008	0,053	-0,098	-0,057	-0,115	-0,044	-0,212	-0,093	0,025	-0,021	0,737	0,000
2	Terrain Ruggedness Index (QGIS)	0,066	-0,106	0,041	-0,038	0,052	-0,117	-0,177	-0,188	-0,210	0,256	0,626	0,080	-0,494	0,000	0,000
3	Topographic position index (QGIS)	-0,022	0,001	-0,095	-0,108	0,059	-0,018	-0,064	-0,014	0,773	-0,284	0,348	-0,053	-0,013	0,000	0,000
4	Asperity (QGIS)	0,047	-0,107	0,013	-0,032	0,107	-0,247	0,709	0,482	0,012	-0,106	0,014	0,017	-0,074	0,000	0,000
5	total curvature (Jenness tool)	-0,038	0,116	0,033	-0,222	0,039	-0,056	-0,011	-0,017	0,013	-0,028	-0,007	-0,016	0,041	0,000	0,000
6	tangential curvature (Jenness tool)	0,081	-0,104	-0,102	-0,209	0,001	0,120	0,284	-0,358	-0,097	0,113	-0,160	-0,538	-0,161	0,000	0,000
7	profile curvature (Jenness tool)	0,095	-0,106	-0,041	0,154	0,075	0,014	0,223	-0,337	0,284	0,125	0,027	0,547	0,135	0,000	0,000
8	plan curvature (Jenness tool)	-0,005	0,142	-0,145	0,019	0,043	-0,057	-0,023	0,023	-0,007	-0,002	0,016	0,003	-0,001	0,000	0,000
9	longitudinal curvature (Jenness tool)	0,065	-0,118	0,015	0,165	-0,011	-0,072	-0,337	0,397	0,099	-0,143	0,057	-0,319	-0,094	0,000	0,457
10	general curvature (Jenness tool)	0,005	0,007	-0,048	-0,228	-0,012	0,049	0,025	0,004	-0,187	0,052	-0,036	0,365	0,105	0,000	0,770
11	cross-sectional curvature (Jenness tool)	0,075	-0,109	-0,067	-0,225	-0,032	0,011	-0,303	0,415	-0,221	-0,056	-0,004	0,303	0,085	0,000	-0,446
12	aspect (Jenness tool)	-0,082	0,039	-0,048	0,003	0,008	0,013	-0,013	-0,018	0,000	0,001	0,002	0,000	0,000	0,000	0,000
13	slope (Jenness tool)	0,069	-0,111	0,025	-0,027	0,071	-0,092	-0,103	-0,100	-0,070	0,152	0,237	-0,251	0,806	0,000	0,000
14	wetness index (SAGA)	-0,226	-0,245	0,135	0,027	-0,590	-0,347	0,043	-0,033	0,037	0,072	0,047	-0,018	0,033	0,000	0,000
15	topographic wetness index (SAGA)	-0,277	-0,257	0,332	0,000	0,634	0,201	0,002	0,061	-0,025	0,061	0,030	-0,021	0,000	0,000	0,000
16	Compound Topographic Index (GRASS)	0,614	0,087	0,060	-0,009	0,091	-0,097	-0,006	-0,027	0,002	0,008	-0,017	-0,015	-0,004	0,000	0,000
17	catchment Area (recursive method) (SAGA)	0,354	0,194	-0,147	-0,002	-0,107	-0,015	0,002	-0,007	0,001	-0,011	-0,001	0,005	-0,001	0,000	0,000
18	mass balance index (SAGA)	0,020	-0,144	-0,254	0,820	0,074	-0,049	0,006	0,004	-0,113	0,028	-0,015	-0,006	0,009	0,000	0,000
19	LS factor (SAGA)	0,064	-0,281	0,030	0,051	-0,337	0,752	0,012	0,035	0,045	-0,167	-0,043	0,042	-0,031	0,000	0,000
20	convergence index (SAGA)	-0,168	-0,052	0,522	0,076	-0,226	-0,115	0,056	-0,035	0,030	0,067	0,006	0,031	0,019	0,000	0,000
21	channel network base level (SAGA)	0,055	-0,032	0,098	0,002	0,038	-0,099	-0,059	-0,117	-0,043	-0,219	-0,096	0,026	-0,022	-0,646	0,000
22	catchment Area (max flux method) (SAGA)	-0,265	0,706	0,142	0,159	-0,013	0,073	-0,015	0,006	-0,018	-0,041	-0,034	0,002	-0,010	0,000	0,000
23	altitude above channel network (SAGA)	-0,167	-0,144	-0,143	-0,035	0,071	-0,043	-0,020	-0,049	-0,022	-0,074	-0,031	0,010	-0,005	-0,200	0,000
24	flow direction (ArcGIS)	-0,191	0,102	-0,081	0,013	0,010	0,033	-0,035	-0,049	0,000	0,001	0,007	-0,001	0,001	0,000	0,000
25	flow drop (ArcGIS)	0,030	-0,112	0,008	-0,041	0,070	-0,146	-0,270	0,066	0,361	0,515	-0,553	0,025	-0,146	0,000	0,000
26	diffuse insolation (SAGA)	-0,014	0,148	-0,085	0,005	-0,101	0,274	0,164	0,321	0,122	0,597	0,262	-0,074	0,062	0,000	0,000
27	direct insolation (SAGA)	-0,180	0,073	-0,069	0,013	0,011	0,025	-0,026	-0,036	-0,002	-0,001	0,005	0,000	0,000	0,000	0,000
28	flow accumulation (SAGA)	-0,068	0,035	-0,080	0,014	0,043	0,011	0,004	-0,009	-0,003	0,004	-0,005	0,004	0,000	0,000	0,000
29	Topographic position index (SAGA)	0,336	0,117	0,610	0,164	0,002	0,097	0,003	0,017	0,048	-0,004	0,011	-0,013	-0,002	0,000	0,000
30	Vector Ruggedness Measure (SAGA)	-0,056	0,137	-0,061	0,019	-0,028	0,048	0,016	0,015	0,011	0,003	0,001	-0,001	0,001	0,000	0,000



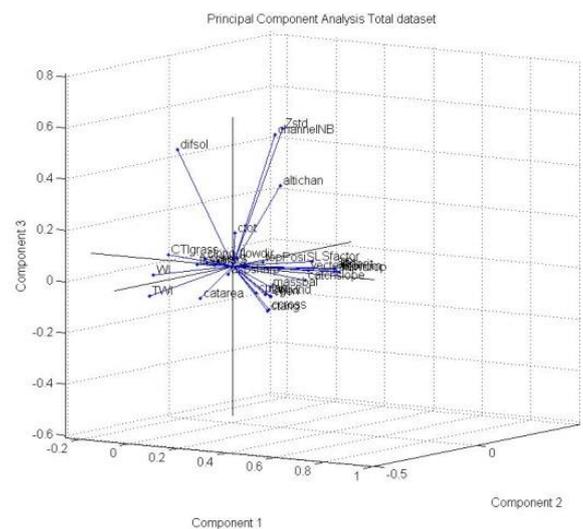
**Fig. 4.1** Diagramma di Pareto relativo alla percentuale di varianza spiegata dalle prime 10 PCs.

**Tabella 4.3** Percentuale di varianza spiegata dalle prime 10 PCs.

Principal component	Percent explained
PC1	30,056
PC2	23,452
PC3	8,688
PC4	5,978
PC5	5,528
PC6	4,335
PC7	3,414
PC8	3,311
PC9	2,879
PC10	2,284



**Fig. 4.2** Scree plot relativo alla *principal component analysis* eseguita sul set di 30 variabili geomorfometriche.



**Fig. 4.3** Loadings plot relativo alla *principal component analysis* eseguita sul set di 30 variabili geomorfometriche.

Osservando i valori della matrice dei *loadings* (Tab. 4.2) relativamente alle prime 4 componenti e con l'ausilio del *loadings plot* relativo alle prime 3 componenti (Fig. 4.3), è stata effettuata una scelta delle variabili da utilizzare. Si precisa che tale scelta non è l'unica possibile. Infatti, osservando i valori dei *loadings* si potrebbe procedere con diverse combinazioni di variabili, sulla base del numero e della tipologia di queste ultime. Per quanto riguarda il numero è stato scelto, arbitrariamente, di procedere con un numero piccolo, rispetto a quello delle variabili disponibili. L'intento è verificare se la segmentazione del territorio basata su un numero ridotto di variabili permette di ottenere dall'attività di campionamento un *sample set* in grado di esaurire la variabilità dell'intero *set* di variabili geomorfometriche considerate.

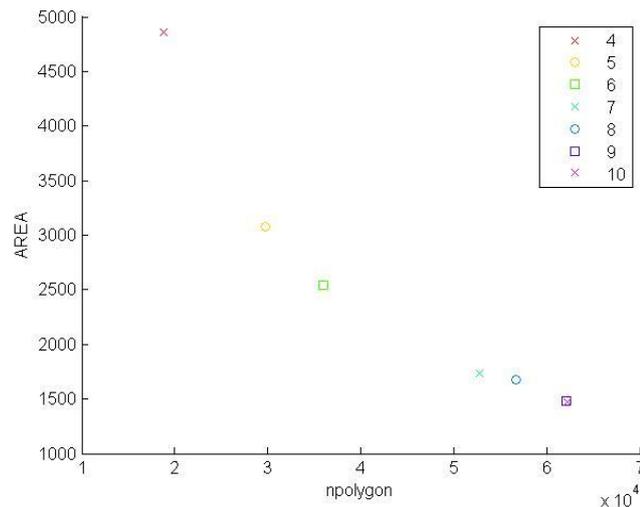
Considerando i valori dei *loadings* più elevati e osservando il comportamento delle variabili nel *loadings plot*, sono state scelte 8 variabili geomorfometriche (*elevation*, *longitudinal curvature* (Jenness tool), *general curvature* (Jenness tool), *wetness index* (SAGA), *Compound Topographic Index* (GRASS), *LS factor* (SAGA), *flow drop* (ArcGIS), *diffuse insolation* (SAGA)), due per ciascuna delle 4 componenti considerate.

Per la segmentazione del territorio è stato utilizzato un metodo di classificazione *unsupervised* sfruttando le *Self-Organizing Feature Maps* (SOFM), una particolare classe di *artificial neural network* (ANN) basate sull'apprendimento competitivo, in grado di classificare i vettori in *input* individuando similarità e correlazione. Notoriamente, le reti neurali sono metodi il cui funzionamento è ispirato a quello dei sistemi nervosi biologici: tali reti, infatti, sono costituite da componenti elementari detti neuroni, connessi in architettura modulabile a seconda dello scopo. Nelle SOFM i neuroni costituiscono i nodi di una rete, generalmente bidimensionale, organizzata topologicamente secondo una certa funzione a cui sono associati dei valori (pesi). L'algoritmo iterativo di addestramento comporta diversi processi: nella fase di inizializzazione tutti i pesi delle connessioni possono assumere valori casuali, successivamente per ciascun vettore di ingresso tali valori vengono aggiornati ed il neurone con il valore più piccolo diventa il *winning neuron*. Ad ogni *input* successivo il peso associato al *winning neuron* e quello degli altri neuroni topologicamente vicini sono alterati e viene determinato il nuovo *winning neuron*: in questo modo la rete subisce un continuo arrangiamento in modo tale che neuroni vicini tra loro rispondono a vettori in *input* simili. Applicando il medesimo tipo di SOFM (*topology function:hextop*, *distance function:linkdist*) sono state costruite 7 mappe, ognuna delle quali è costituita da un diverso numero di unità geomorfometriche (da 4 a 10).

La scelta del numero di classi in cui suddividere il territorio è uno dei passaggi fondamentali di questo approccio. Questa decisione, insieme al numero di repliche da effettuare per ciascuna unità geomorfometrica, determina quanto intenso debba essere il campionamento. Tenendo sempre presenti gli obiettivi di economicità e accuratezza, nell'applicazione di questo criterio di campionamento occorre considerare i seguenti aspetti teorici e pratici. Il primo deriva dalla relazione tra il numero di classi geomorfometriche e la scrupolosità nel voler cogliere le differenze del territorio: maggiore è il numero di classi geomorfometriche con cui è discretizzato il territorio, più accuratamente sarà rilevata la variabilità territoriale dell'area considerata. D'altro canto, all'aumentare del numero di classi aumenta anche il numero di punti di campionamento (per i quali occorre considerare anche le repliche per ciascuna classe) e, contemporaneamente, diminuisce esponenzialmente la dimensione media dell'areale delle unità geomorfometriche con le quali è stato suddiviso il territorio (v. Fig. 4.4. e Tab. 4.4).

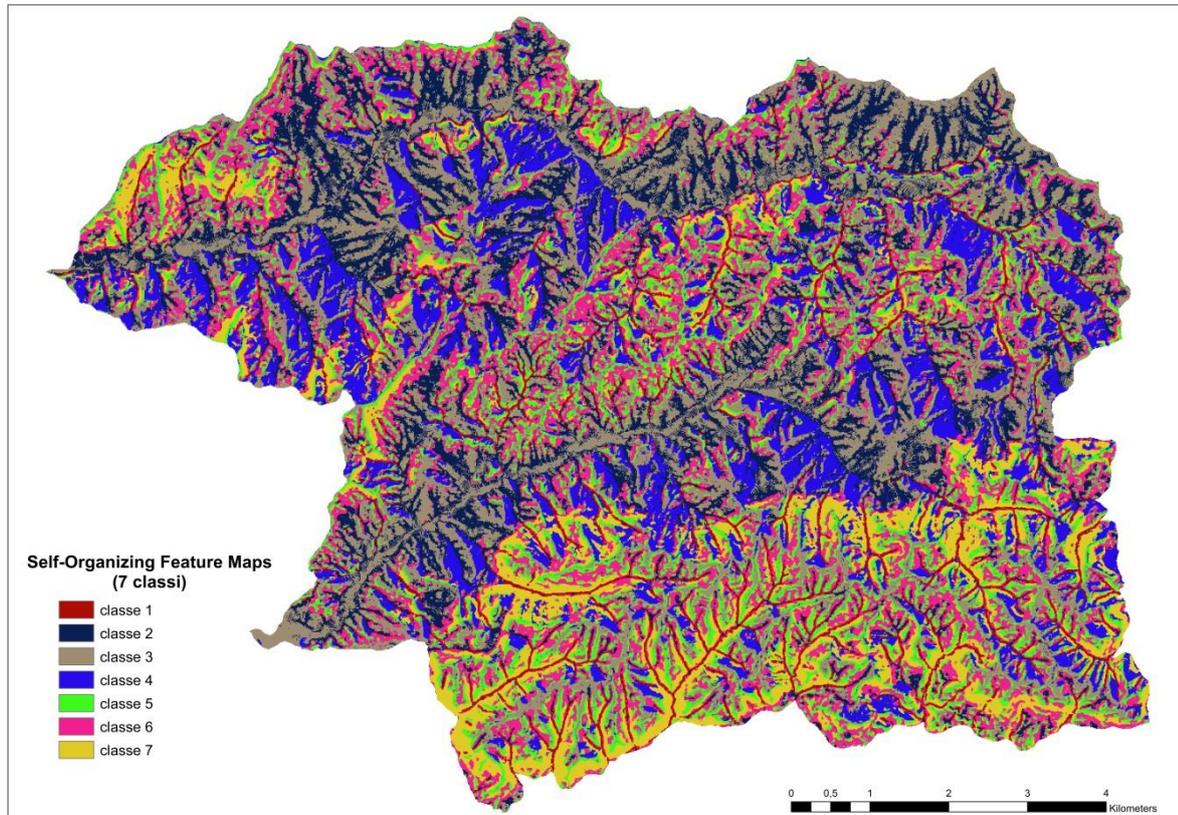
**Tabella 4.4** Area media e numero di poligoni secondo il numero di classi con cui è stata suddivisa l'area di studio.

n° classi	area media (m <sup>2</sup> )	n° di poligoni
4	4860,87	18866
5	3081,10	29763
6	2546,04	36018
7	1736,23	52818
8	1678,22	56645
9	1477,54	62067
10	1473,97	62218



**Fig. 4.4** Scatter plot relativo alla tabella 4.4.

Per di più, occorre tenere in considerazione da una parte il concetto di minima unità cartografabile e dall'altra la necessità pratica di dover individuare in campo la *patch* corrispondente al poligono mappato. In questo caso, dato l'utilizzo del DEM con pixel a 10 m, è stata scelta come minima unità cartografabile una *patch* costituita da 2x2 pixel (400 m<sup>2</sup>), una delle opzioni indicate da McBratney (McBratney et al. 2003). E' stata scelta in maniera arbitraria la mappa con 7 categorie, la quale presenta un comportamento intermedio rispetto alle opzioni considerate (Fig 4.5) e presenta unità geomorfometriche di dimensione non inferiore alla minima unità cartografabile.



**Fig. 4.5** Segmentazione del territorio in 7 classi ottenuta applicando una ANN *Self-Organizing Feature Maps*.

## 4.2.1 Introduzione di una variabile categoriale

Considerando le conoscenze pregresse riguardo l'area di studio (v. Cap.2), è stato scelto di introdurre una variabile categoriale che fornisca informazione riguardo il materiale parentale. Non essendo disponibile una carta litologica si è fatto riferimento alla carta geologica - foglio n° 178 Voghera (scala 1:50.000) (Vercesi et al. 2015), redatta nell'ambito del progetto CARG. Riguardo questo strato informativo occorre precisare che non esiste una relazione certa tra il materiale parentale di un suolo in una determinata porzione di territorio e l'unità geologica assegnata alla stessa area. Questo può accadere per due cause: la prima riguarda le dinamiche evolutive dei versanti, le quali possono fornire al processo di pedogenesi materiale parentale diverso dal substrato roccioso sottostante; la seconda è determinata dalla scala cartografica utilizzata: infatti, quella 1:50.000 non è sufficientemente dettagliata per poter assegnare con certezza la classe geologica corretta, su un'area discretizzata con pixel a 10 m, in corrispondenza dei bordi dei poligoni. Tuttavia, lo strato informativo geologico è stato introdotto, per procedere con maggior scrupolosità al campionamento dell'area. Secondo la legenda della mappa, nell'area di studio sono presenti 25 unità geologiche, di cui 7 occupano l'87% del territorio (v. Tab. 4.5).

**Tabella 4.5** Unità geologiche presenti nell'area di studio, con l'areale in km<sup>2</sup> e in percentuale riclassificate.

UG_DESCR25	UG_SIGLA25	area km <sup>2</sup>	area %	nuove classi
Arenarie di Monte Vallassa	AVL	25,2044	27,48	1
Brecce di Baiso	BAI	12,2938	13,4	2
Formazione di Antognola	ANT	10,7329	11,7	3
Formazione di Contignaco	CTG	9,8681	10,76	4
Formazione di Ranzano - Litozona arenaceo-pelitica	RAN <sub>2b</sub>	8,1217	8,86	5
Formazione di Ranzano - Litozona arenaceo-conglomeratica	RAN <sub>3a</sub>	6,5243	7,11	5
Formazione di Ranzano - Litozona arenaceo-pelitica	RAN <sub>3b</sub>	6,4703	7,05	5
Sintema del Po (Unità Postglaciale)	POI	3,9847	4,34	0
Formazione di Ranzano - Litozona arenaceo-conglomeratica	RAN <sub>2a</sub>	1,925	2,1	5
Unità dell'Ardivestra	ADV	1,2387	1,35	6
Formazione di Ranzano - Membro del Pizzo d'Oca	RAN <sub>1</sub>	1,0283	1,12	5
Unità di Nizza	NIZ	0,9232	1,01	6
Marne di Monte Piano	MMP	0,8253	0,9	esclusa
Formazione di Ranzano - Litozona pelitico-arenacea	RAN <sub>3c</sub>	0,6513	0,71	5
Formazione di Ranzano - Litozona pelitico-arenacea	RAN <sub>1c</sub>	0,5828	0,64	5
Conglomerati di Cassano Spinola	CCS	0,5825	0,64	esclusa
Formazione di Ranzano - Litozona pelitico-arenacea	RAN <sub>2c</sub>	0,3906	0,43	5
Unità di Varzi	VRZ	0,1291	0,14	6
Formazione di Ranzano - Membro di Varano de' Melegari	RAN <sub>3</sub>	0,1177	0,13	5
Unità di Torretta	TTS	0,0322	0,04	esclusa
Formazione di Ranzano	RAN	0,0288	0,03	5
Membro delle arenarie di Monte Arzolo	CCS <sub>1</sub>	0,0222	0,02	esclusa
Unità di Costa Pelata	BPE	0,0187	0,02	esclusa
Formazione di Sapigno	GNO	0,0143	0,02	esclusa
Litozona Inferiore	MVT <sub>a</sub>	0,0031	0	esclusa

Date le caratteristiche dei substrati rocciosi (Vercesi et al. 2015), e il relativo areale, è stata effettuata una semplificazione della carta geologica, suddividendo il territorio in 7 classi geologiche, sufficientemente omogenee dal punto di vista delle caratteristiche mineralogiche (v. Cap.6). L'originale carta geologica è stata semplificata nelle seguenti 7 classi: 1) Arenarie di Monte Vallassa (arenarie bioclastiche e biocalcareni, intercalate da marne sabbiose grigio-azzurre); 2) Breccie di Baiso (corpi rocciosi a tessitura caotica costituiti da breccie poligeniche a matrice argillosa grigia o varicolori); 3) Formazione di Antognola (marne e marne argillose); 4) Formazione di Contignaco (marne siltose grigie, marne calcaree chiare, intercalate nella parte inferiore da strati arenaceo-pelitici), 5) Formazioni di Ranzano, con i suoi diversi membri, nelle loro differenti *litofacies*: membro Pizzo d'Oca (arenarie e peliti in proporzione variabile), membri di Val Pessola e di Varano de' Melegari (caratterizzati entrambi da arenarie, conglomerati e peliti in proporzione molto variabile), RAN1<sub>a</sub> arenaceo-conglomeratica, RAN1<sub>b</sub> arenaceo-pelitica, RAN1<sub>c</sub> pelitico-arenacea), 6) unità che costituiscono il fondo delle valli (Nizza, Ardivestra, Varzi); 7) il cosiddetto sintema del Po, che delinea aree con depositi di frana, detriti di versante ed aree fluviali. Quest'ultima classe viene esclusa dall'attività di campionamento dei suoli, perché non di interesse dal punto di vista pedologico. Le unità che nella cartografia occupano meno dell'1% del territorio non sono state considerate per il tematismo semplificato.

Per tenere in considerazione questa nuova variabile categoriale è stato costruito il prodotto, per intersezione, della mappa relativa alle classi geomorfometriche con quella costituita dalle nuove classi geologiche. La sovrapposizione tra i due tematismi ha originato la mappa finale nella quale il territorio è suddiviso in 26 differenti unità, chiamate *unità territoriali*, potenzialmente oggetto di campionamento. Si sottolinea che la mappa delle classi geologiche non è stata utilizzata come variabile categoriale nel calcolo delle unità geomorfometriche, effettuato con le SOFM, ma ha concorso alla realizzazione della mappa finale per il campionamento, “*splittando*” in più parti le unità geomorfometriche calcolate: infatti, alcune di esse, appartenendo a differenti unità geologiche, determinano un incremento del numero di unità da sottoporre al rilevamento pedologico. In altre parole, la variabile categoriale introdotta fornisce elementi utili per l'esecuzione di un campionamento più accurato (procurando, dunque, un maggior dettaglio circa la variabilità territoriale).

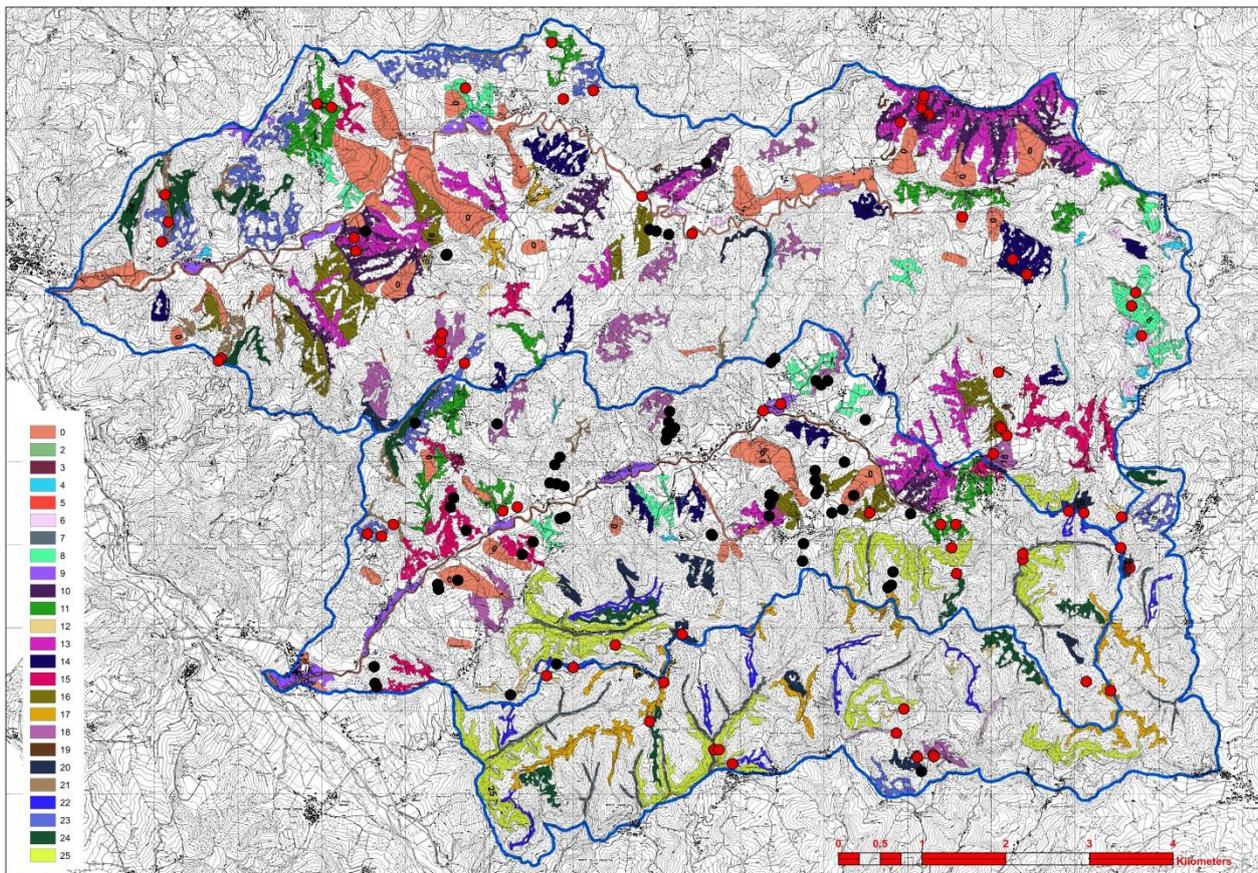
Allo scopo di facilitare l'attività di campo, la mappa finale per il campionamento è stata prodotta selezionando, per ciascuna delle unità territoriali, i poligoni più estesi (12 per ciascuna tipologia): essa è stata inoltre sovrapposta alla Carta Tecnica Regionale (1:10.000) della Regione Lombardia. Questi accorgimenti permettono di individuare in campo con maggior facilità la *patch* da sottoporre a campionamento (v. Fig. 4.6).

Per quanto riguarda l'esecuzione dell'attività di campionamento ed in particolare il numero di repliche da effettuare, si sarebbe potuto considerare il concetto di “*gradi di libertà*” utilizzato nell'applicazione dei disegni sperimentali:

$$LD = E - P - R$$

Dove *LD* sono i gradi di libertà del modello, *E* è il numero totale di esperimenti, *P* il numero dei parametri che devono essere stimati per il modello ed *R* le repliche.

Nell'applicazione dei disegni sperimentali, teoricamente è necessario mantenere un numero di gradi di libertà sufficiente per determinare sia l'errore sia il *lack of fit* di un modello di regressione. Non conoscendo a priori l'approccio statistico che verrà utilizzato per la redazione della mappa, in via precauzionale si è scelto di rilevare almeno tre profili per ciascuna tipologia di *unità territoriale*. Si evidenzia che la classe 0 corrispondente al Sistema del Po (Unità Postglaciale) non è stata oggetto di campionamento, in quanto relativa a depositi di frana, detriti di versante ed aree fluviali. Compatibilmente con le possibilità offerte dalla realtà del territorio, per 21 delle 25 unità di rilevamento, rimanenti, sono state campionate randomicamente un minimo di tre *patches* diverse, scegliendo per ciascuna di esse un punto casuale all'interno dell'area, garantendo l'indipendenza e la casualità degli oggetti che costituiscono il *dataset*, come richiesto dalla statistica.



**Fig. 4.6** Mappa per il rilevamento pedologico con l'indicazione delle unità geologiche semplificate e dei punti di campionamento: punti neri (profili rilevati con il metodo *free survey*) e punti rossi (profili rilevati utilizzando la mappa).<sup>3</sup>

### 4.3 Analisi e risultati

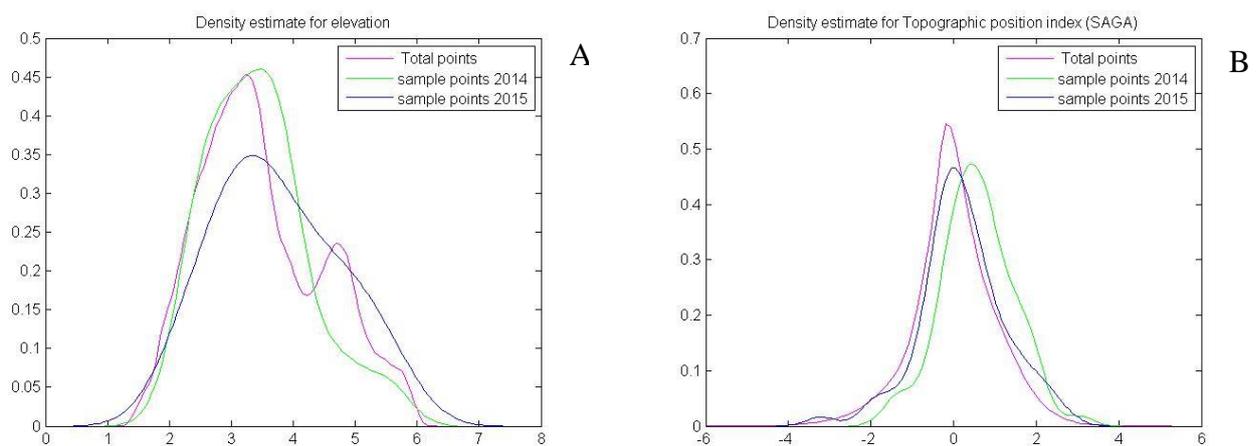
Per quanto riguarda la mappa ottenuta, occorre sottolineare che l'applicazione di una tecnica di *clustering unsupervised* determina una suddivisione del territorio che può portare alla creazione di unità di campionamento non utili allo scopo per cui si procede. In altre parole, poiché la segmentazione avviene sull'intero territorio, può succedere che alcune unità territoriali, determinate dal calcolo, si riferiscano a superfici non interessate dai processi di pedogenesi e che pertanto non devono essere oggetto del rilevamento pedologico, oppure con particolari circostanze per le quali è reso impossibile il campionamento. Nel corso dell'attività di campionamento, particolari situazioni

che hanno impedito il rilevamento si sono verificate per le aree corrispondenti all'unità 3, le quali raggiungono al massimo una superficie pari a 400 m<sup>2</sup>, che corrispondono a depressioni situate in prossimità del letto dei fiumi. Analogamente le unità 5 e 19 (rispettivamente con estensione massima pari a 5000 e a 10.000 m<sup>2</sup>) che sono aree di piccole dimensioni che corrispondono rispettivamente al fondo di stretti impluvi o a zone in erosione costante (es: calanchi), non sono state oggetto di campionamento. Può anche accadere, specialmente in aree antropizzate, che per alcune unità non si possa procedere al campionamento, per differenti cause, come è accaduto nel caso delle unità 4 (minimo 10.000, massimo 42.000 m<sup>2</sup>), che corrispondono a zone situate in posizioni inaccessibili oppure occupate da colture agricole in atto, dunque non campionabili. Per le rimanenti 21 unità, come già detto, il campionamento è avvenuto rilevando un minimo di tre *patches* per ogni *unità territoriale* di campionamento: complessivamente sono stati rilevati 70 profili pedologici.

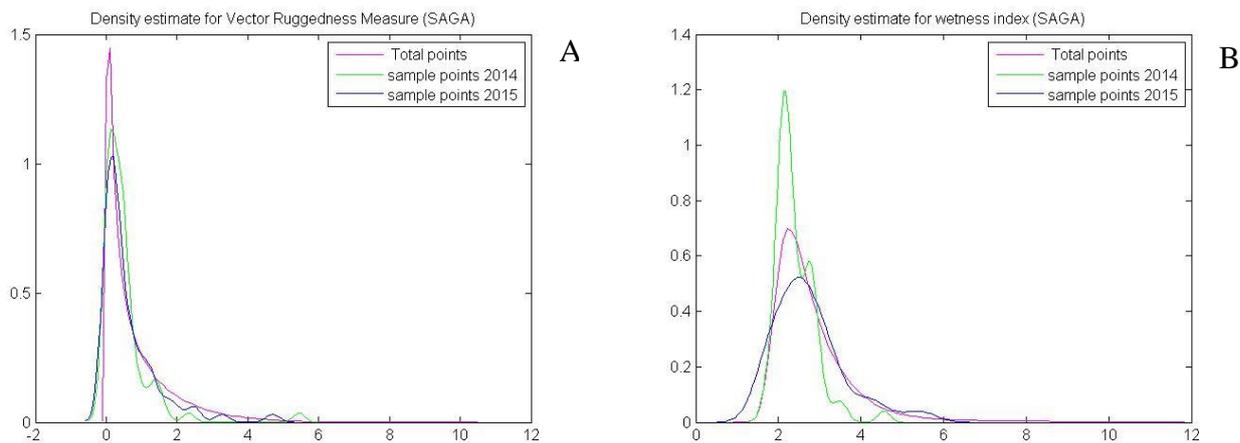
Al fine di valutare il risultato dell'attività di campionamento e contestualmente comprendere la validità del metodo, sono state eseguite due verifiche, in modo da stabilire quanto i punti di rilevamento siano in grado di rappresentare la variabilità del territorio. Ricordando che quest'ultima è determinata dalle 30 variabili geomorfometriche che potenzialmente possono essere utilizzate nelle successive elaborazioni statistiche per la creazione di mappe tematiche, verificare se l'attività di campionamento è esaustiva rispetto alla variabilità territoriale significa appurare se l'insieme dei punti di campionamento è in grado di riprodurre la stessa informazione contenuta nelle variabili geomorfometriche, relativamente all'intera area.

Tal verifica è stata effettuata sia per l'insieme di punti di campionamento rilevati utilizzando il metodo proposto (*dataset* 2015), sia sull'insieme dei profili rilevati con il metodo *free survey* (*dataset* 2014).

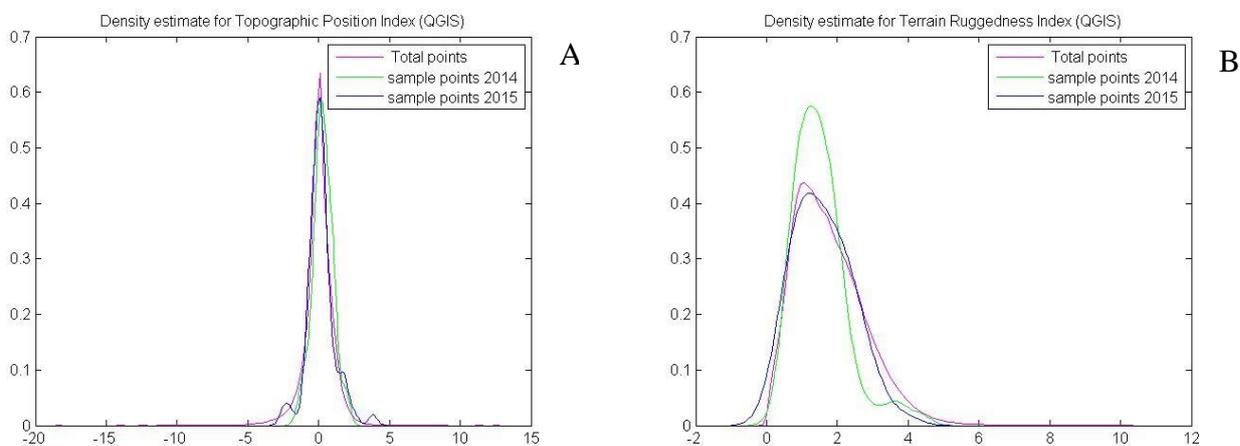
Analalogamente a quanto fatto da Brungard e Boettinger (Brungard e Boettinger 2010), per ogni variabile è stata "plottata" una versione empirica della funzione di densità di probabilità. I grafici riportati da Fig. 4.7 a Fig. 4.21 permettono di confrontare qualitativamente, per ciascuna variabile, la similitudine delle distribuzioni.



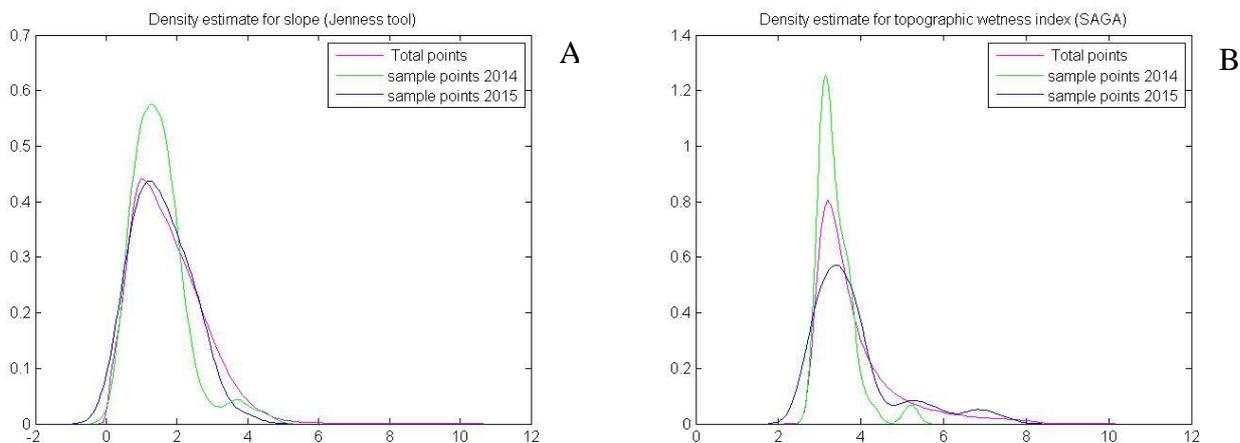
**Fig. 4.7** Confronto della densità di probabilità empirica relativa ai punti di campionamento eseguito con il metodo *free survey* (sample points 2014) con i punti relativi al campionamento effettuato secondo l'approccio proposto (sample points 2015) e con la densità di probabilità empirica relativa all'intera area di studio (Total points) per la variabile *elevation* (A) e per la variabile *Topographic position index* (SAGA) (B).



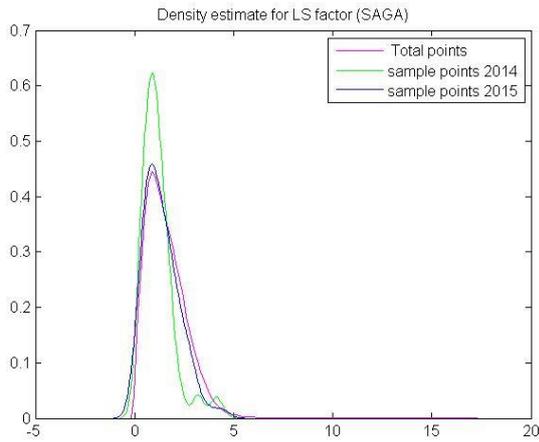
**Fig. 4.8** Confronto della densità di probabilità empirica relativa ai punti di campionamento eseguito con il metodo *free survey* (sample poits 2014) con i punti relativi al campionamento effettuato secondo l'approccio proposto (sample poits 2015) e con la densità di probabilità empirica relativa all'intera area di studio (Total points) per la variabile *Vector Ruggedness Measure* (SAGA) (A) e per la variabile *wetness index* (SAGA) (B).



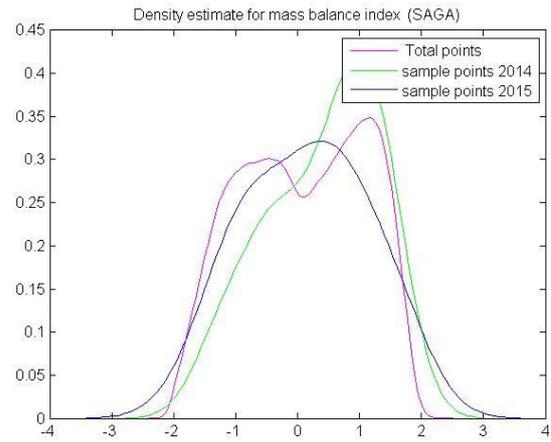
**Fig. 4.9** Confronto della densità di probabilità empirica relativa ai punti di campionamento eseguito con il metodo *free survey* (sample poits 2014) con i punti relativi al campionamento effettuato secondo l'approccio proposto (sample poits 2015) e con la densità di probabilità empirica relativa all'intera area di studio (Total points) per la variabile *Topographic position index* (QGIS) (A) e per la variabile *Terrain Ruggedness Index* (QGIS) (B).



**Fig. 4.10** Confronto della densità di probabilità empirica relativa ai punti di campionamento eseguito con il metodo *free survey* (sample poits 2014) con i punti relativi al campionamento effettuato secondo l'approccio proposto (sample poits 2015) e con la densità di probabilità empirica relativa all'intera area di studio (Total points) per la variabile *slope* (Jenness tool) (A) e per la variabile *topographic wetness index* (SAGA) (B).

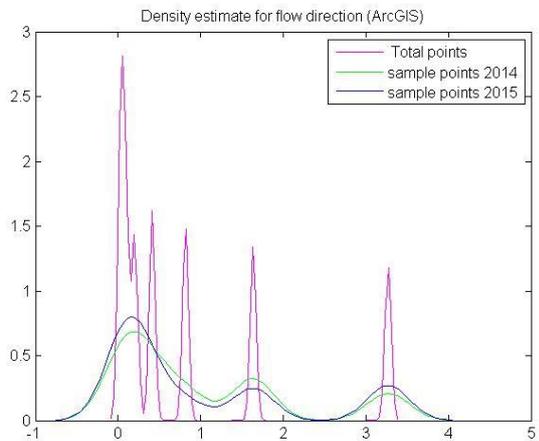


A

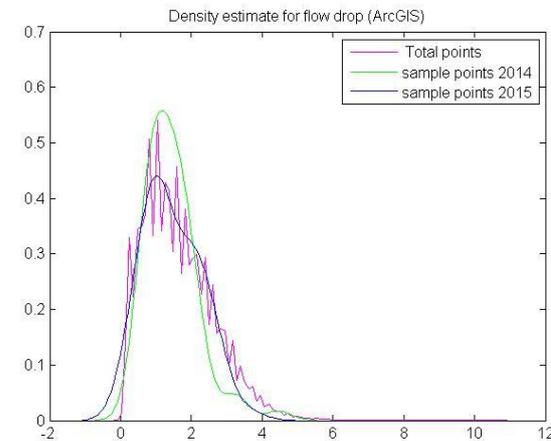


B

**Fig. 4.11** Confronto della densità di probabilità empirica relativa ai punti di campionamento eseguito con il metodo *free survey* (sample poits 2014) con i punti relativi al campionamento effettuato secondo l'approccio proposto (sample poits 2015) e con la densità di probabilità empirica relativa all'intera area di studio (Total points) per la variabile *LS factor* (SAGA) (A) e per la variabile *mass balance index* (SAGA) (B).

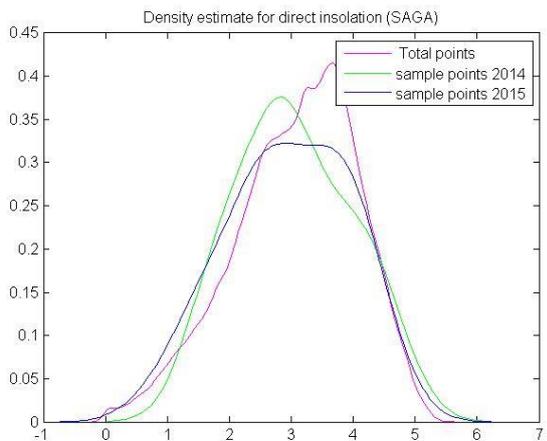


A

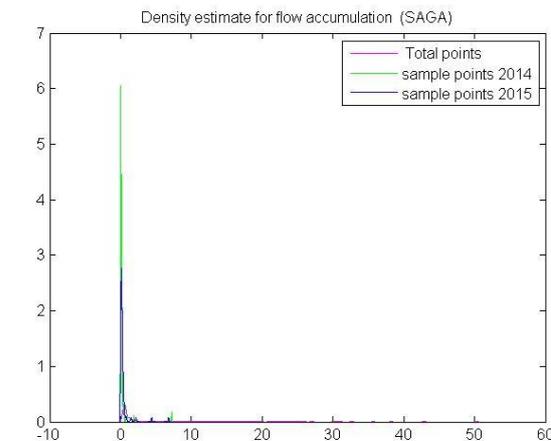


B

**Fig. 4.12** Confronto della densità di probabilità empirica relativa ai punti di campionamento eseguito con il metodo *free survey* (sample poits 2014) con i punti relativi al campionamento effettuato secondo l'approccio proposto (sample poits 2015) e con la densità di probabilità empirica relativa all'intera area di studio (Total points) per la variabile *flow direction* (ArcGIS) (A) e per la variabile *flow drop* (ArcGIS) (B).

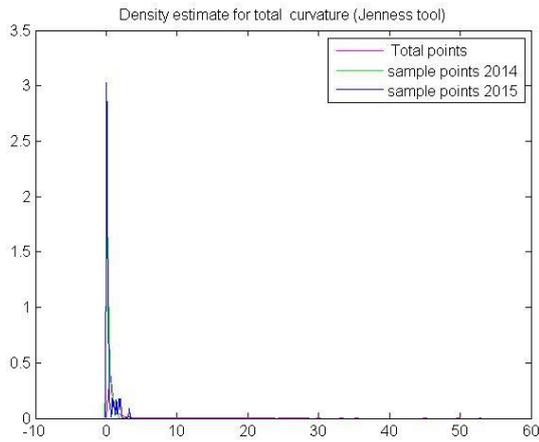


A

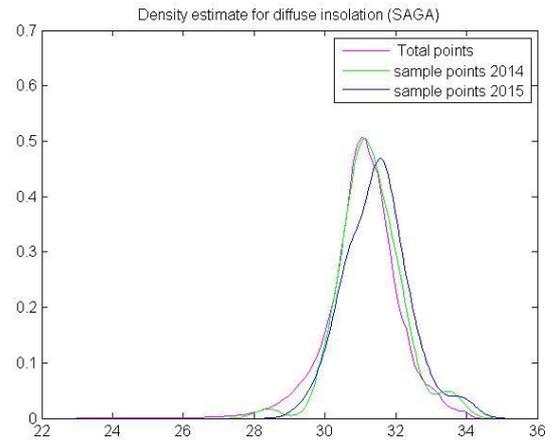


B

**Fig. 4.13** Confronto della densità di probabilità empirica relativa ai punti di campionamento eseguito con il metodo *free survey* (sample poits 2014) con i punti relativi al campionamento effettuato secondo l'approccio proposto (sample poits 2015) e con la densità di probabilità empirica relativa all'intera area di studio (Total points) per la variabile *direct insolation* (SAGA) (A) e per la variabile *flow accumulation* (SAGA) (B).

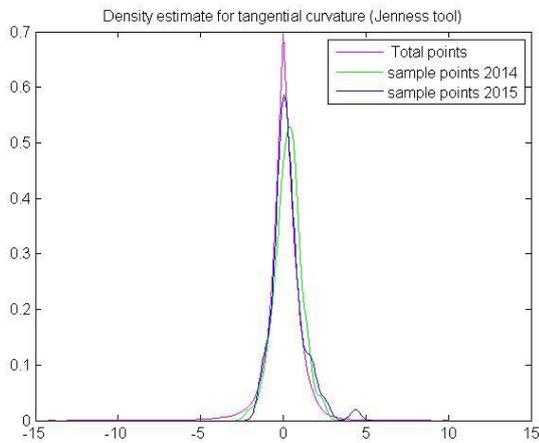


A

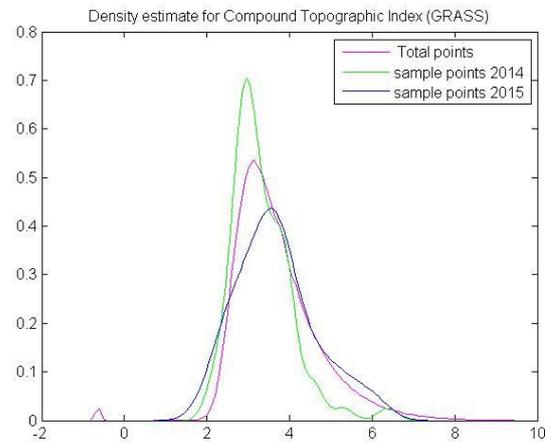


B

**Fig. 4.14** Confronto della densità di probabilità empirica relativa ai punti di campionamento eseguito con il metodo *free survey* (sample poits 2014) con i punti relativi al campionamento effettuato secondo l'approccio proposto (sample poits 2015) e con la densità di probabilità empirica relativa all'intera area di studio (Total points) per la variabile *total curvature* (Jenness tool) (A) e per la variabile *diffuse insolation* (SAGA) (B).

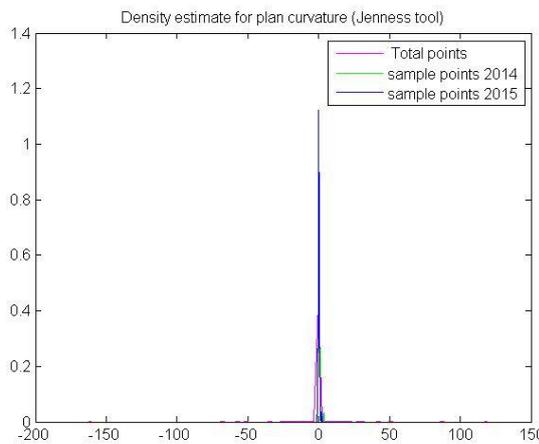


A

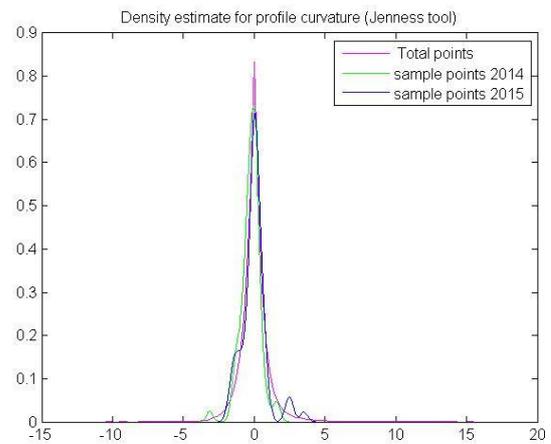


B

**Fig. 4.15** Confronto della densità di probabilità empirica relativa ai punti di campionamento eseguito con il metodo *free survey* (sample poits 2014) con i punti relativi al campionamento effettuato secondo l'approccio proposto (sample poits 2015) e con la densità di probabilità empirica relativa all'intera area di studio (Total points) per la variabile *tangential curvature* (Jenness tool) (A) e per la variabile *Compound Topographic Index* (GRASS) (B).

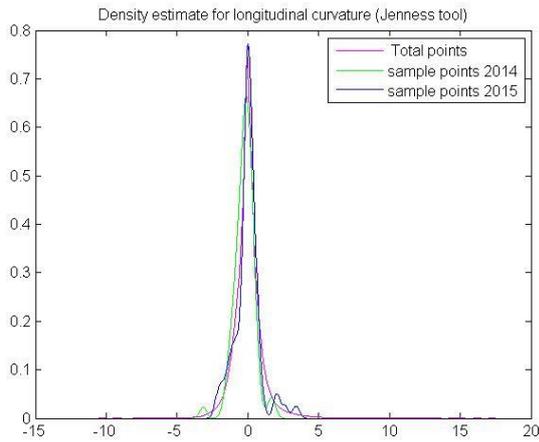


A

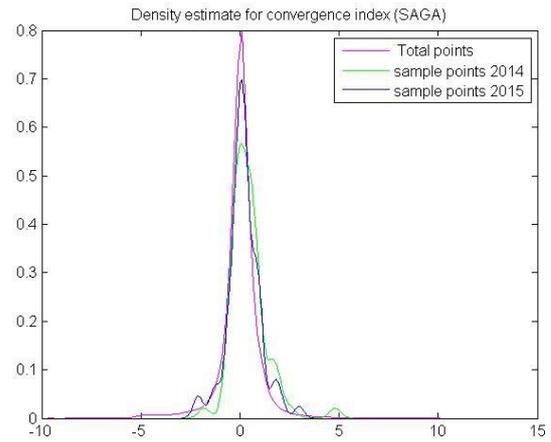


B

**Fig. 4.16** Confronto della densità di probabilità empirica relativa ai punti di campionamento eseguito con il metodo *free survey* (sample poits 2014) con i punti relativi al campionamento effettuato secondo l'approccio proposto (sample poits 2015) e con la densità di probabilità empirica relativa all'intera area di studio (Total points) per la variabile *plan curvature* (Jenness tool) (A) e per la variabile *profile curvature* (Jenness tool) (B).

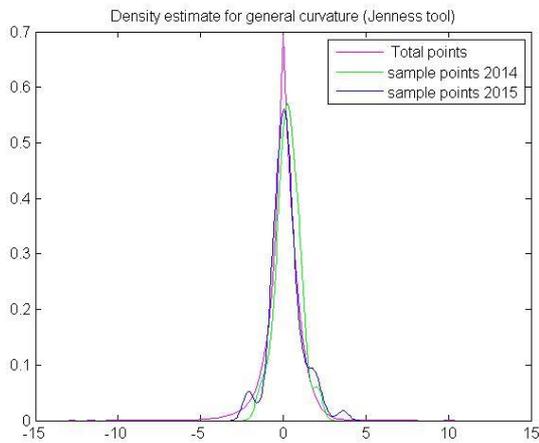


A

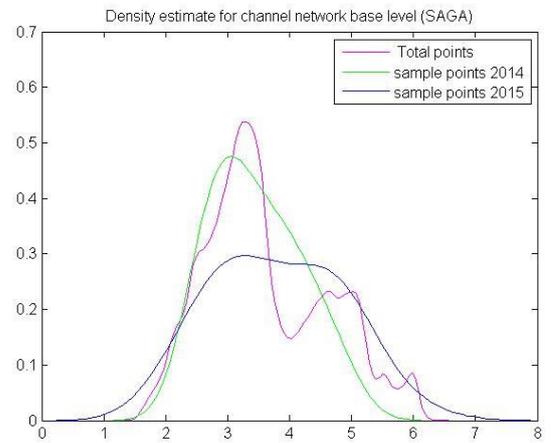


B

**Fig. 4.17** Confronto della densità di probabilità empirica relativa ai punti di campionamento eseguito con il metodo *free survey* (sample poits 2014) con i punti relativi al campionamento effettuato secondo l'approccio proposto (sample poits 2015) e con la densità di probabilità empirica relativa all'intera area di studio (Total points) per la variabile *longitudinal curvature* (Jenness tool) (A) e per la variabile *convergence index* (SAGA) (B).

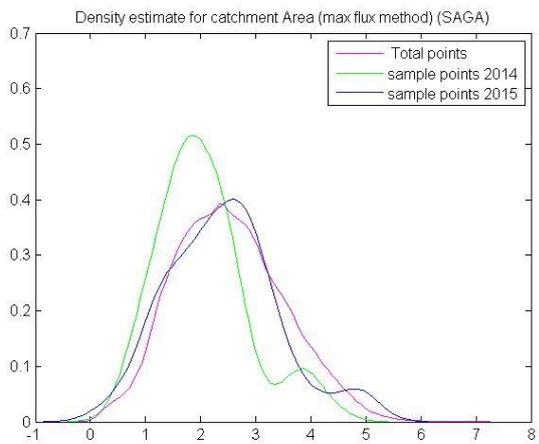


A

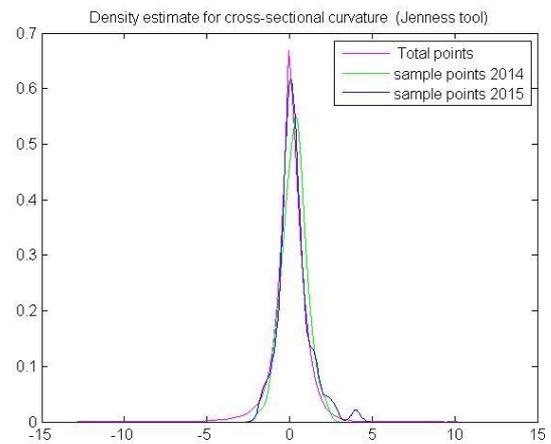


B

**Fig. 4.18** Confronto della densità di probabilità empirica relativa ai punti di campionamento eseguito con il metodo *free survey* (sample poits 2014) con i punti relativi al campionamento effettuato secondo l'approccio proposto (sample poits 2015) e con la densità di probabilità empirica relativa all'intera area di studio (Total points) per la variabile *general curvature* (Jenness tool) (A) e per la variabile *channel network base level* (SAGA) (B).

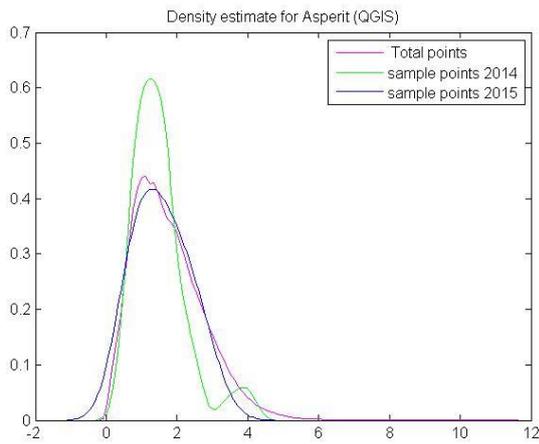


A

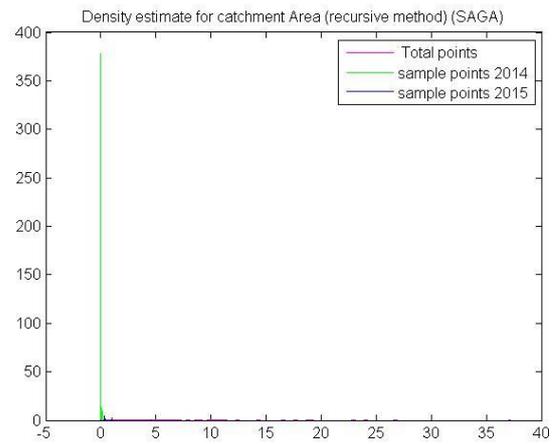


B

**Fig. 4.19** Confronto della densità di probabilità empirica relativa ai punti di campionamento eseguito con il metodo *free survey* (sample poits 2014) con i punti relativi al campionamento effettuato secondo l'approccio proposto (sample poits 2015) e con la densità di probabilità empirica relativa all'intera area di studio (Total points) per la variabile *catchment Area (max flux method)* (SAGA) (A) e per la variabile *cross-sectional curvature* (Jenness tool) (B).

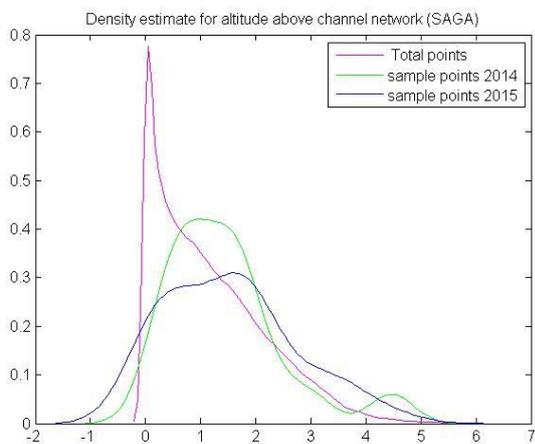


A

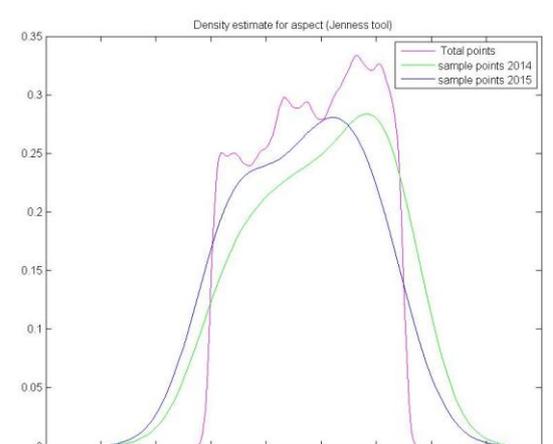


B

**Fig. 4.20** Confronto della densità di probabilità empirica relativa ai punti di campionamento eseguito con il metodo *free survey* (sample points 2014) con i punti relativi al campionamento effettuato secondo l'approccio proposto (sample points 2015) e con la densità di probabilità empirica relativa all'intera area di studio (Total points) per la variabile *Asperit* (QGIS) (A) e per la variabile *catchment Area (recursive method)* (SAGA) (B).



A



B

**Fig. 4.21** Confronto della densità di probabilità empirica relativa ai punti di campionamento eseguito con il metodo *free survey* (sample points 2014) con i punti relativi al campionamento effettuato secondo l'approccio proposto (sample points 2015) e con la densità di probabilità empirica relativa all'intera area di studio (Total points) per la variabile *altitude above channel network* (SAGA) (A) e per la variabile *aspect* (Jenness tool) (B).

La verifica statistica circa la similitudine tra le distribuzioni campionarie e quella dell'intero set di dati è stata effettuata con il test *Two-sample Kolmogorov-Smirnov*, il cui esito è riportato nella tabella 4.6. Esso è un *test* non parametrico basato sull'ipotesi nulla "i due vettori messi a confronto hanno la stessa distribuzione", contro l'ipotesi alternativa "i due vettori differiscono nella loro distribuzione continua". Il risultato è 1 se il *test* rigetta l'ipotesi nulla ad un livello di significatività del 5%, 0 nell'altro caso.

Dai risultati emerge che per il campionamento effettuato seguendo l'approccio proposto, per il 77% delle variabili considerate il *set* dei punti di campionamento riesce a coprire la variabilità dell'intero territorio. Viceversa, seguendo un approccio *free survey*, solo nel 15% dei casi si ottiene l'obiettivo preposto.

**Tabella 4.6** Esito del *test* non parametrico *Two-sample Kolmogorov-Smirnov* quando il risultato è 1 è rigettata l'ipotesi nulla ("i due vettori hanno la stessa distribuzione").

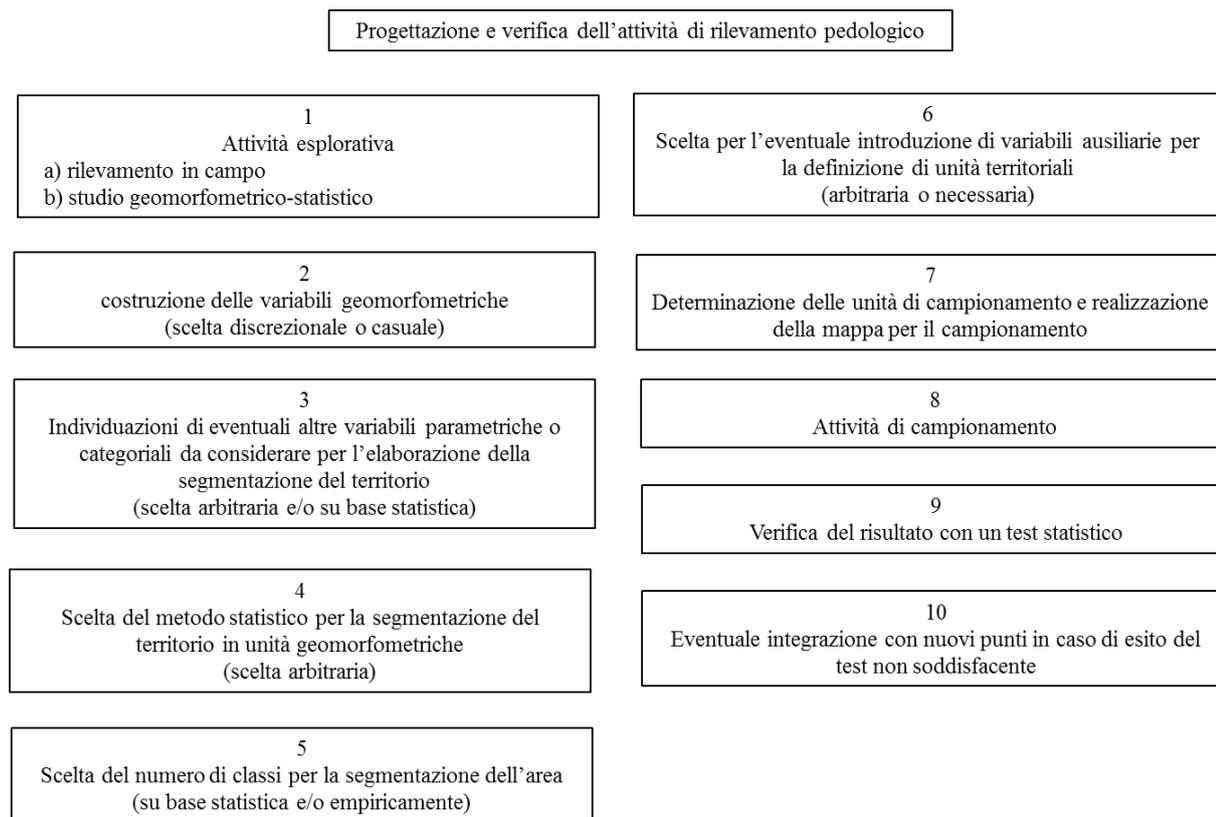
Two-sample Kolmogorov-Smirnov test 5% significance level	elevation	Terrain Ruggedness Index QGIS	Topographic position index QGIS	Asperit QGIS	total Curvature Jenness	tangential Curvature Jenness	profile Curvature Jenness	plan Curvature Jenness	longitudinal Curvature Jenness	general Curvature Jenness	cross-sectional Curvature Jenness	aspect Jenness	slope Jenness	wetness index SAGA	topographic wetness index SAGA
TOT vs point	1	1	1	1	0	1	0	1	0	1	1	0	1	1	1
TOT vs point2015	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0
TOT vs point2014	0	1	1	1	0	1	1	1	1	1	1	1	1	1	1

**Tabella 4.6**

Two-sample Kolmogorov-Smirnov test 5% significance level	Compound Topographic Index GRASS	catchment Area (recursive method) SAGA	mass balance index SAGA	LS factor SAGA	convergence index SAGA	channel network base level SAGA	catchment Area (max flux method) SAGA	altitude above channel network SAGA	flow direction ArcGIS	flow drop ArcGIS	diffuse insolation SAGA	direct insolation SAGA	flow accumulation SAGA	Topographic position index SAGA	Vector Ruggedness Measure SAGA
TOT vs point	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1
TOT vs point2015	0	0	0	0	0	1	0	1	1	0	1	0	0	1	0
TOT vs point2014	1	1	1	1	1	1	1	1	0	1	0	0	1	1	1

## 4.4 Discussione

Al fine di utilizzare in modo efficiente le risorse, per creare mappe tematiche relative al suolo l'attività di campionamento dovrebbe essere preceduta da un lavoro propedeutico, finalizzato all'individuazione delle aree da sottoporre a rilevamento. La progettazione dell'attività di campionamento proposta può essere suddivisa nei seguenti *step*:



**Fig. 4.22** Fasi della progettazione dell'attività di campionamento.

L'attività di rilevamento in campo della fase esplorativa (punto 1 fig. 4.22), di per sé, non dovrebbe essere necessaria per un approccio che procede con una programmazione teorica del campionamento. Occorre però evidenziare che tale attività può essere utile per le aree delle quali

non si conosce nulla riguardo l'oggetto da cartografare, per acquisire elementi di base e meglio progettare l'attività di rilevamento.

Per il metodo presentato, due *steps* risultano particolarmente importanti: 1) la scelta del numero di classi geomorfometriche con cui suddividere il territorio (punto 5 Fig. 4.22); 2) l'eventuale introduzione di tematismi ambientali (variabili ausiliarie) (punto 6 Fig. 4.22). Per evitare una scelta arbitraria, il numero di classi da adottare dovrebbe essere determinato con metodi statistici, come contemplato parzialmente in questo lavoro (es: valutazione della variazione dell'areale massimo e minimo al variare del numero di classi) oppure potrebbe essere stabilito sulla base di conoscenze acquisite nel corso dell'attività esplorativa. Per quanto concerne il presente lavoro, è stata svolta una attività esplorativa che ha portato al rilevamento di 63 profili pedologici. Dette informazioni non sono state considerate per la scelta del numero di classi geomorfometriche, ma hanno indotto l'introduzione del tematismo "geologia"; tuttavia si sottolinea (come indicato nel paragrafo 4.2.1.) che l'introduzione di detta variabile categoriale è avvenuta nello step 6, ovvero non è stato introdotto nel calcolo delle classi geomorfometriche allo step 3.

Riguardo la scelta del numero di classi geomorfometriche da intersecare con la mappa geologica, si evidenzia che la sovrapposizione di detto *layer* è stato effettuato per tutte le 7 mappe create, relative alle *unità geomorfometriche*, costruite con le SOFM (da 4 a 10 classi geomorfometriche). Le 7 nuove mappe derivanti dalla sovrapposizione dei due tematismi forniscono rispettivamente 19, 21, 24, 25, 29, 33, 36 *unità territoriali* potenziali per il campionamento, le quali sono numericamente inferiori al semplice prodotto tra il numero di classi ( $n^\circ$  di classi geomorfometriche  $\times$   $n^\circ$  di classi geologiche): difatti, molte unità geomorfometriche sono specifiche per alcune formazioni geologiche.

Analizzando i dati relativi alle aree delle 7 nuove mappe suddivise in *unità territoriali* per il campionamento, è emerso che: dovendo individuare la mappa di maggior dettaglio (maggior numerosità di *unità territoriali* per il campionamento) con il vincolo di un minimo di 12 poligoni con area maggiore di 400 m<sup>2</sup> per ciascuna unità di campionamento, l'unica scelta possibile ricade sulla mappa con 7 classi geomorfometriche. Questo accade perché, aumentando il numero di classi geomorfometriche, si ha un incremento nel numero di *patches* che non raggiungono la dimensione minima cartografabile.

L'esistenza di discrezionalità in alcune fasi di questo approccio non comporta la perdita di rigorosità, poiché la valutazione del risultato è basata su un *test* statistico.

Secondo lo schema riportato in figura 4.22, le variabili ausiliarie, sia categoriali che numeriche, possono essere introdotte in due differenti *steps* e con due differenti funzioni. Qualora vengano introdotte nel punto 3 diventano delle vere e proprie variabili che partecipano al calcolo per la segmentazione del territorio; altrimenti, servono solo come elemento caratterizzante, fornendo la possibilità di ottenere maggiore informazione sulla variabilità indagata, con un campionamento più scrupoloso.

A tal riguardo occorre tenere in considerazione che: "*In situations where we have doubts on the strength of the correlation between the environmental variables and the soil variable, we may prefer not to use this prior information at the sampling stage of a survey. The use of weakly*

*correlated environmental variables in designing a sample may lead to suboptimal samples. If after the sampling it appears that there is a moderate or strong correlation, then we still can use this prior information at the prediction (estimation) stage of a survey.*” (Brus et al. 2006).

In questo caso, nonostante l'importante relazione tra il substrato litologico e la pedogenesi, non si è ritenuto opportuno introdurre la variabile categoriale “*geologia*” nel calcolo eseguito per la classificazione dell'area in *patches* omogenee. Tale scelta è basata su due considerazioni, riguardanti rispettivamente alcune caratteristiche dell'area di studio e la risoluzione della mappa geologica. La prima considerazione riguarda le dinamiche evolutive dei versanti particolarmente importanti nell'area di studio: ciò comporta che il suolo relativo ad un determinato punto non è detto che abbia come materiale parentale il substrato litologico corrispondente all'unità geologica ivi cartografata. Per quanto riguarda la mappa geologica, si è ritenuto che la risoluzione spaziale del foglio Voghera (1:50.000) è poco appropriata per il livello dettaglio che si vuole ottenere nella realizzazione delle mappe dei suoli (discretizzate a 10 m). Introdurre, dunque, il tematismo “*geologia*” nel calcolo avrebbe potuto determinare un errore, specialmente per le aree vicine al limite delle unità geologiche, forzando l'assegnazione di un determinato punto ad una unità territoriale alla quale, in realtà, può non appartenere. Inoltre, va anche sottolineato che l'introduzione della mappa geologica con le classi originali, come elemento discriminante, avrebbe provocato una moltiplicazione delle aree da campionare, raggiungendo un numero di classi superiore a 180, improponibile dal punto di vista del campionamento pratico, soprattutto considerando le necessarie repliche.

## 4.5 Conclusioni

Nel metodo presentato la segmentazione del territorio crea *patches* che contengono una informazione più significativa rispetto alla *landform segmentation*, utilizzata per il semplice riconoscimento delle forme principali del territorio. Questa informazione è veicolata da *proxy* (le variabili geomorfometriche) che simulano l'attività dei fattori della pedogenesi. Pur non escludendo l'applicazione di variabili ausiliarie (categoriali o numeriche) relative a tematismi ambientali, l'applicazione presentata può anche basarsi su un numero ridotto di variabili geomorfometriche. La bontà dell'attività di campionamento circa la variabilità campionata e, di conseguenza, del metodo, è verificabile con un test statistico. In caso di esito negativo del test, per il raggiungimento dell'obiettivo è possibile incrementare i punti di campionamento, senza dover cambiare la mappa da cui si è partiti. La presenza nella mappa di unità territoriali di campionamento di grandi dimensioni ne permette la facile individuazione in campo. Il numero consistente di unità territoriali consente di procedere agevolmente al campionamento, assicurando nel contempo l'indipendenza e la casualità dei campioni. Dal *test di Kolmogorov-Smirnov*, effettuato per confrontare la distribuzione delle variabili geomorfometriche rispetto a quella delle variabili relative ai punti campionati, è emerso che l'attività di campionamento effettuata con il metodo proposto, a differenza di quella svolta con il metodo tradizionale (*free survey*), è in grado di rilevare la variabilità necessaria per l'approccio statistico del DSM.

## Bibliografia

- Brevik, Eric C.; Calzolari, Costanza; Miller, Bradley A.; Pereira, Paulo; Kabala, Cezary; Baumgarten, Andreas; Jordán, Antonio (2016): Soil mapping, classification, and pedologic modeling. History and future directions. In: *Geoderma*, 264, pp. 256–274. DOI: 10.1016/j.geoderma.2015.05.017.
- Brungard, C. W.; Boettinger, J. L. (2010): Conditioned latin hypercube sampling: Optimal sample size for digital soil mapping of arid rangelands in Utah, USA. In: *Digital soil mapping*: Springer, pp. 67–75.
- Brus, D. J.; Gruijter, J. J. de (1997): Random sampling or geostatistical modelling? Choosing between design-based and model-based sampling strategies for soil (with discussion). In: *Geoderma* n. 1-2, 80, pp. 1–44.
- Brus, D. J.; Gruijter, J. J. de; van Groenigen, J. W. (2006): Designing spatial coverage samples using the k-means clustering algorithm. In: *Developments in Soil Science*, 31, pp. 183–192.
- Dobermann, A.; Simbahan, G. C. (2006): Methodology for using secondary information in sampling optimisation for making fine-resolution maps of soil organic carbon. In: *Developments in Soil Science*, 31, pp. 167–607.
- Grunwald, S. (2009): Multi-criteria characterization of recent digital soil mapping and modeling approaches. In: *Geoderma* n. 3-4, 152, pp. 195–207. DOI: 10.1016/j.geoderma.2009.06.003.
- Hengl, Tomislav; Rossiter, David G.; Stein, Alfred (2003): Soil sampling strategies for spatial prediction by correlation with auxiliary maps. In: *Aust. J. Soil Res.* n. 8, 41, p. 1403. DOI: 10.1071/SR03005.
- Heuvelink, Gerard B. M.; Brus, Dick J.; Gruijter, Jaap J. de (2006): Optimization of sample configurations for digital mapping of soil properties with universal kriging. In: *Developments in Soil Science*, 31, pp. 137–151.
- Lagacherie, Philippe (2008): Digital soil mapping: a state of the art. In: *Digital soil mapping with limited data*: Springer, pp. 3–14.
- Martin, W.K.E.; Timmer, V. R. (2006): Capturing spatial variability of soil and litter properties in a forest stand by landform segmentation procedures. In: *Geoderma* n. 1-2, 132, pp. 169–181. DOI: 10.1016/j.geoderma.2005.05.004.
- McBratney, A.B; Mendonça Santos, M.L; Minasny, B. (2003): On digital soil mapping. In: *Geoderma* n. 1-2, 117, pp. 3–52. DOI: 10.1016/S0016-7061(03)00223-4.
- Minasny, Budiman; McBratney, Alex B. (2006): A conditioned Latin hypercube method for sampling in the presence of ancillary information. In: *Computers & Geosciences* n. 9, 32, pp. 1378–1388. DOI: 10.1016/j.cageo.2005.12.009.
- P.L. Vercesi, P. Faletti, C. Pasquini, L. Papani, C. Parotti, G. Tucci. (2015) - Note Illustrative della carta Geologica d'Italia in scala 1:50.000, foglio178 "Voghera". Servizio Geologico d'Italia - ISPRA, 166 pp.

Pennock, D. J. (2003): Terrain attributes, landform segmentation, and soil redistribution. In: *Soil and Tillage Research* n. 1, 69, pp. 15–26.

Pennock, D.J; Corre, M.D (2001): Development and application of landform segmentation procedures. In: *Soil and Tillage Research* n. 3-4, 58, pp. 151–162. DOI: 10.1016/S0167-1987(00)00165-3.

Royle, J. Andrew; Nychka, Doug (1998): An algorithm for the construction of spatial coverage designs with implementation in SPLUS. In: *Computers & Geosciences* n. 5, 24, pp. 479–488.

Stepinski, T. F.; Bagaria, C. (2009): Segmentation-Based Unsupervised Terrain Classification for Generation of Physiographic Maps. In: *IEEE Geosci. Remote Sensing Lett.* n. 4, 6, pp. 733–737. DOI: 10.1109/LGRS.2009.2024333.

Webster, Richard; Oliver, Margaret A. (2001): *Geostatistics for environmental scientists (Statistics in Practice)*.

White, Robert E. (2013): *Principles and practice of soil science: the soil as a natural resource*: John Wiley & Sons.

## CAPITOLO 5

### Spessore del suolo

#### 5.1 Introduzione

Lo spessore e l'erosione del suolo sono due aspetti di interesse per il governo del territorio (Morgan et al. 1998) ed in particolare per la gestione agricola (Montgomery 2007; Lal 1998), sulla quale hanno riflessi economici (Pimentel et al. 1995). In questo lavoro sono state redatte alcune mappe tematiche utili per la comprensione di questi due fattori antagonisti (v. Cap. 6). In questo capitolo viene descritto il processo per la realizzazione di una mappa per la rappresentazione della presenza o assenza di suolo e dello spessore del *solum*.

Lo spessore del suolo è un importante parametro utilizzato nei modelli deterministici fisicamente basati per lo studio della stabilità dei versanti (Lee e Ho 2009; Zizioli et al. 2013; Menezes et al. 2014), nonché è un importante parametro per la modellazione dei flussi idrologici (Tesfa et al. 2009; Rahimy 2012). Esso è anche un indicatore dello sviluppo pedogenetico, ed è stato incluso nel calcolo degli indici sviluppati da Harden (Harden 1982). Detti indici sono utilizzati per misurare quantitativamente il grado di sviluppo di un profilo e per comprendere la genesi di un suolo ed il processo sotteso. Lo spessore e l'erosione del suolo sono utilizzati dai geomorfologi per comprendere l'evoluzione nel tempo del territorio (Renschler e Harbor 2002) e sono controllati dalla pendenza e dalle dinamiche di evoluzione dei versanti (Tromp-van Meerveld e McDonnell 2006; Dietrich et al.). Occorre tuttavia definire cosa si intende con il termine profondità del suolo o spessore del suolo, sinonimi utilizzati con il medesimo significato nell'ambito degli studi sopracitati. Da un punto di vista strettamente pedologico, ai termini "spessore o profondità del suolo" di rado è difficile associare una misura scevra da determinazioni soggettive. Tale determinazione potrebbe risultare ancora più difficoltosa considerando le diverse definizioni di suolo (Bockheim et al. 2005).

In ogni caso, la misura può essere rilevata nel corso dell'attività di campionamento: in tale occasione, il pedologo rilevatore è il "giudice" e, proprio a causa di questo ruolo, è fondamentale tenere in considerazione il suo background culturale. Infatti, il risultato dell'attività di campionamento può essere dipendente dai giudizi soggettivi dei rilevatori. In particolare l'individuazione degli orizzonti, compreso quello più profondo, può essere di difficile esecuzione, a seconda della complessità dell'area studiata, dell'esperienza del pedologo stesso o delle caratteristiche della stazione di rilevamento. Considerando le differenti definizioni di suolo, al fine di eliminare il più possibile gli equivoci, nel momento in cui si vuole considerare lo spessore del suolo in questo lavoro si fa riferimento al concetto di *solum* (Dietrich et al.; Chartin et al. 2011). In sintesi, il *solum* è la parte superiore e più evoluta del profilo pedologico (Previtali, 2001); è quella parte del profilo pedologico che comprende tutti gli orizzonti individuabili al di sopra dell'orizzonte C o dello strato R, i quali costituiscono il *parent material*, concordemente alla definizione di Kellogg (1930) "*parent material that has been unaltered by soil-building forces*" (Schaeztl e Anderson 2005). Secondo la definizione della *Soil Survey Division Staff* (1993) "*the solum is the set of horizons which are related through the same cycle of pedogenic processes (exclude buried soils)*".

La definizione di *solum*, tuttavia, non è priva di incertezze (Douglas et al. 1967), a causa delle interpretazioni delle definizioni introdotte, nel tempo, nella scienza del suolo (Bockheim et al. 2005). In questo lavoro, date le caratteristiche pedologiche dell'area di studio, la misura dello spessore del *solum* è stata determinata considerando la presenza dei substrati litologici ed escludendo lo spessore degli orizzonti organici, i quali, nell'area di studio, sono dati dalla lettiera forestale (orizzonte OL), presente sui suoli rilevati nelle aree boschive. In altre parole lo spessore del *solum* è stato misurato osservando l'altezza del profilo a partire dalla superficie a contatto con l'atmosfera, escludendo gli orizzonti organici, fino all'orizzonte che presenta caratteristiche univocamente riconducibili ai substrati litologici delle formazioni sedimentarie. Infatti, come termine dello spessore del suolo è stato considerato, a seconda dei casi, la profondità del substrato lapideo, definibile come orizzonte R, oppure la profondità dei depositi colluviali o la profondità del substrato roccioso, definibili come orizzonte C, secondo WRB (2014), considerando appartenenti al *solum* gli eventuali orizzonti di transizione (ad es. BC e CB). In altri termini, sono stati considerati tutti gli orizzonti che hanno subito processi di alterazione; le fratturazioni del substrato litologico, dovute a meccaniche tettoniche, non sono state considerate indicatrici di evoluzione pedologica.

Per la spazializzazione dello spessore del suolo, in letteratura sono presenti lavori nei quali sono stati utilizzati diversi approcci di calcolo. Dietrich ha costruito una mappa della profondità del suolo basandosi sul bilancio di massa tra produzione e trasporto diffuso di suolo (Dietrich et al. 1995). Roering e Pelletier hanno dedotto lo spessore del suolo considerando la morfologia del territorio (Roering 2008; Pelletier e Rasmussen 2009). In altri lavori sono stati utilizzati metodi di geostatistica quali *ordinary kriging*, *co-kriging*, *regression-kriging* e *linear regression* (Penížek e Borůvka 2006) oppure la *multi-linear regression* (Hseu et al. 2002) e anche la *regression tree* (McKenzie e Ryan 1999). Tesfa ha modellato la profondità del suolo utilizzando variabili morfometriche e informazioni derivati da immagini satellitari (Tesfa et al. 2009). Kuriakose ha comparato differenti mappe dello spessore del suolo ottenute con diversi metodi statistici, con cinque variabili morfometriche e la mappa di uso del suolo (Kuriakose et al. 2009). Lu ha mappato lo spessore del suolo integrando la statistica *fuzzy* con metodi di classificazione ad albero (Lu et al. 2016).

In questo lavoro per la redazione della mappa relativa allo spessore del suolo sono stati messi a confronto due metodi: le *Artificial Neural Network* (ANN) (Demuth et al. 2014) e la *Partial Least Square Regression* (PLSR). Al *dataset* costituito dagli spessori dei punti di campionamento rilevati è stato aggiunto quello dei punti ove il suolo non è presente. L'introduzione di detti punti è indispensabile in quanto essi costituiscono gli zeri della funzione utilizzata per modellare lo spessore del suolo. Questi punti sono luoghi privi di vegetazione, a causa della costante erosione del suolo che ostacola lo sviluppo pedogenetico, oppure sono aree ove a causa della morfologia e delle dinamiche di versante è presente un costante movimento gravitativo di materiale più o meno pedogenizzato o, in rari casi, roccia affiorante. Occorre precisare che da satellite non è possibile individuare completamente le aree prive o quasi di suolo; infatti, nell'area di studio sono state individuate zone ove il processo di pedogenesi è costantemente disturbato dalla perdita di materiale o da fenomeni di colluviazione, seppur in presenza di vegetazione arborea. Ciò avviene tipicamente nelle aree soggette a dinamica di tipo calanchivo o in quelle poste su pendenze tali da mantenere periodicamente attivo il movimento di materiale. Sono aree ove si può verificare una temporanea

stabilizzazione del versante, con conseguente radicamento di vegetazione pioniera, cui generalmente seguono eventi che determinano la riattivazione dell'erosione.

## 5.2 Materiali e metodi

La base informativa per la redazione della mappa relativa allo spessore del suolo è costituita dalle variabili geomorfometriche primarie e secondarie utilizzate come predittori (v. Cap. 3) e dalla variabile indipendente spessore del *solum*, rilevato nel corso dell'attività di campionamento (v. Cap. 4).

Il *dataset* dello spessore dei profili rilevati è stato suddiviso in due parti una delle quali è costituita dai profili pedologici campionati seguendo la suddivisione del territorio in unità geomorfometriche (70 punti di campionamento), come descritto nel capitolo 4. Questa parte del *dataset* è stata utilizzata come *training set* per l'elaborazione dei modelli. A tale insieme di punti di campionamento sono stati uniti altri elementi puntuali, come di seguito indicato, relativi a luoghi ove lo spessore del suolo è pari a zero. L'altra porzione del *dataset* è costituita dagli spessori dei profili pedologici rilevati con il metodo *free survey* (v. cap. 4) ed è stata utilizzata interamente sia come *validation set* che come *test set* (62 punti di campionamento). Anche a questo set di dati sono stati associati altri elementi, relativi a punti con *solum* pari a zero, come di seguito indicato. Si precisa che, al di là degli indici di valutazione dei singoli modelli, derivanti dai metodi statistici utilizzati, la bontà dei risultati è stata valutata sulla base della capacità predittiva per i 62 punti campionamento (campionati con il metodo *free survey*, v. Cap. 4) utilizzati come *dataset* esterno (v. paragrafo 5.4).

Il *dataset* relativo ai punti con spessore del *solum* pari a zero è stato costituito partendo dai punti privi di suolo rilevati nel corso dell'attività di campionamento il cui numero è stato incrementato mediante l'utilizzo di immagini satellitari nella regione del visibile. Si precisa i punti aggiunti al *dataset* originale sono esclusivamente quelli che corrispondono ad aree completamente prive di vegetazione (roccia affiorante o calanchi attivi) adiacenti al punto di campionamento e pertanto facilmente individuabili con le immagini satellitari utilizzate (basemap di ArcGIS). Di questo *dataset* sono stati assegnati al *training set* un numero di punti paragonabile a quello impiegato con presenza di suolo (63 punti). L'assegnazione è stata effettuata in modo casuale, in base all'ordine progressivo presente nel *dataset*, avendo cura, però, di considerare almeno un punto per ciascuna tipologia di area priva di suolo rilevata in campo. Assegnazione analoga è stata effettuata per il *validation set* (64 punti). La rimanente aliquota di punti con suolo zero (287 punti) è stata impiegata per il *test set*.

In un processo di modellazione, la selezione delle variabili è utilizzata per ridurre la dimensione del set di dati al fine di individuare le variabili che veicolano l'informazione fondamentale. Per la costruzione di modelli di qualità statisticamente accettabile è utile determinare, prima del processo di modellazione, e specialmente nel caso di modelli di regressione, l'eventuale presenza di correlazione e di collinearità tra le variabili. In altre parole è utile separare l'informazione fondamentale da quella parte di inutile informazione (rumore), che può deteriorare la capacità predittiva del modello, mitigando in tal modo l'effetto denominato “*the curse of dimensionality*” (Behrens et al. 2010). Possono essere applicate due modalità di selezione delle variabili, chiamate *filter approach* e *wrapper approach*. Nel primo caso si procede ad una selezione preliminare delle

variabili cercando di individuare quelle che hanno una importanza fondamentale per ottenere buoni risultati in predizione. Nel secondo approccio vengono confrontati diversi risultati ottenuti con il medesimo algoritmo e con differenti *sub-set* di variabili indipendenti (Behrens et al. 2010). In questo lavoro sono stati adottati entrambi gli approcci contestualmente, per indagare l'influenza delle variabili sulla qualità della risposta. Inoltre, considerati i risultati delle analisi statistiche di base (v. Cap. 3), che hanno evidenziato la peculiarità delle variabili geomorfometriche (non normalità, presenza di outlier, collinearità e correlazione), sono stati scelti due approcci di calcolo. Nel primo metodo, utilizzato per testare la capacità predittiva delle variabili geomorfometriche originali, sono state applicate le *Artificial Neural Network* (ANN) a differenti *sub-set* di variabili, determinati adottando due metodi di selezione, basati sui risultati della *principal component analysis* (PCA). Di contro, pensando ad una trasformazione delle variabili, è stata testata la capacità predittiva delle variabili trasformate con il metodo delle componenti principali, utilizzando la *Partial Least Square Regression* (PLSR).

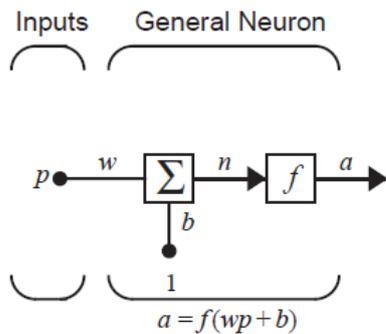
Con le ANN non sono state utilizzate le componenti principali in luogo delle variabili originali, considerando che quando si effettua la trasformazione delle variabili predittive, effettuando la rotazione del sistema di riferimento con la tecnica delle PCA, i versori del nuovo sistema di riferimento sono orientati secondo le direzioni di massima varianza dei predittori (X), senza tenere in considerazione alcuna relazione con la variabile indipendente. In questo modo non è detto che le componenti principali siano la migliore base informativa per effettuare un processo di inferenza (es: una regressione), per la costruzione di un modello che stima la variabile indipendente (Wise et al. 2006). Si precisa che nel processo di modellazione adottato l'obiettivo è stato quello di cercare la migliore approssimazione del risultato finale, delegando la scelta delle variabili da utilizzare ad applicazioni statistiche, senza entrare nel merito del significato fisico delle variabili e quindi della loro eventuale influenza sull'oggetto mappato.

Nell'ambito del DSM, le ANN sono state utilizzate da Zhu per mappare i pedopaesaggi (Zhu 2000), da Behrens per mappare le unità di suolo (Behrens et al. 2005), da Boruvka e Penizek per assegnare profili pedologici a unità di suolo predefinite (Boruvka e Penizek 2006), da Chagas per mappare le classi di suolo (Chagas et al. 2013), da Beucher per mappare i suoli solfato acidi (Beucher et al. 2015).

I principali vantaggi delle ANN sono la capacità di generalizzazione, la possibilità di gestire problemi non lineari e la possibilità di utilizzare grandi quantità di dati. Inoltre, non richiedono particolari tipi di distribuzione dei dati, a differenza dei metodi statistici parametrici tradizionali, che spesso necessitano di una distribuzione normale dei dati. Inoltre, consentono la manipolazione di dati provenienti da fonti diverse, con diversi livelli di precisione e rumore (Chagas et al. 2013).

Le ANN sono tecniche di *machine-learning*, il cui funzionamento mima il meccanismo del sistema neurologico. In questo caso le ANN sono addestrate con un processo *supervised*, in modo che per un particolare segnale in input si ottiene uno specifico output.

Le ANN sono composte da unità elementari dette neuroni (Fig. 5.1), organizzabili in strutture più o meno complesse, il cui funzionamento è sintetizzabile in 3 operazioni fondamentali:



1. il segnale di input  $p$  è moltiplicato per uno scalare (peso)  $w$  per formare lo scalare  $wp$ ;
2. all'input pesato è aggiunto un altro scalare  $b$  detto *bias* o *offset* per formare l'input netto  $n$ ;
3. l'input netto  $n$  viene elaborato tramite la funzione di trasferimento per produrre uno scalare in uscita  $a$ .

**Fig. 5.1** Unità elementare di una ANN.

Una volta inizializzati i pesi, il processo di addestramento (o *training*) di una rete neurale comporta l'aggiustamento, per iterazioni successive, dei pesi e dei *bias* in modo da ottimizzare la rete. L'ottimizzazione avviene osservando la variazione del *mean square error*, tra i dati in output della rete neurale e i rispettivi valori di target. L'addestramento continua fino a che l'errore non decresce per un certo numero di iterazioni, che può essere impostato.

Generalmente, per addestrare le ANN i dati vengono suddivisi in: *training set*, *validation set*, *test set*. Il *training set* è utilizzato per aggiustare pesi e *bias* in accordo con l'errore. Il sottoinsieme di validazione è utilizzato per fermare il processo di addestramento quando la ANN incomincia a perdere la capacità di generalizzazione (*early stopping*). In altre parole il *validation set* è utilizzato per evitare l'*overfitting* che può prodursi nel corso dell'addestramento basato sul *training set*. Nel caso di *overfitting*, l'errore relativo al set di validazione, in genere, inizia a salire. Quando l'errore per la predizione del *validation set* aumenta per un numero specificato di volte (6 in questo caso) il processo di addestramento della rete viene interrotto, e vengono conservati i pesi ed il *bias* relativi alla iterazione per la quale si è verificato il minimo errore in predizione. Il *test set* non viene utilizzato per l'addestramento della rete, ma fornisce indicazioni sulla qualità del processo. Se l'errore sul *test set* raggiunge il minimo ad un numero di iterazioni significativamente diverso da quello dell'errore di validazione, potrebbe esserci una distorta suddivisione del *dataset* tra *training*, *validation* e *test set* (manuale MATLAB).

Il *tool* di MATLAB fornisce gli strumenti per la valutazione dei risultati. In particolare, due sono gli output fondamentali: il plot relativo alla *performance* e gli *scatter plot* relativi ai risultati del processo di addestramento. Il plot relativo alla *performance* rappresenta l'andamento dell'errore prodotto, rispetto al *target*, per ciascuno dei tre set (*training*, *validation* e *test*). Nel caso in cui la curva del *test set* abbia un incremento significativo prima che la curva relativa al set di validazione cresca, allora è possibile che si sia verificato *overfitting*. Il plot relativo ai risultati della regressione fornisce quattro *scatter plots* che mettono in relazione i valori di *target* e i valori calcolati dalla rete, sia per quanto riguarda il *dataset* di *training* che per quelli di validazione e di test, oltre ad un totale. E' stata utilizzata una *multilayer feedforward neural network* con una *log-sigmoid transfer function* per l'*hidden layer* e *linear transfer functions* per l'*output layer* ed è stato trattenuto il miglior risultato di differenti 30 processi di training.

Dei diversi algoritmi di calcolo disponibili con il *tool* di MATLAB, sono stati testati in via preliminare i seguenti metodi: *Levenberg-Marquardt*, *BFGS Quasi-Newton*, *Resilient Backpropagation*, *Scaled Conjugate Gradient*, *Conjugate Gradient with Powell/Beale Restart*, *Fletcher-Powell Conjugate Gradient*, *Polak-Ribière Conjugate Gradient*, *One Step Secant* (Demuth

et al. 2014). Successivamente è stato scelto il metodo *Polak-Ribière Conjugate Gradient* che nelle prove preliminari ha fornito le migliori prestazioni.

Il metodo PLSR condivide caratteristiche con altre tecniche di regressione, ed è utilizzato per la realizzazione di modelli predittivi quando i predittori (X) sono correlati tra loro. E' simile alle *Principal Component Regression* (PCR), in quanto effettua una regressione utilizzando una trasformazione dei dati originali. È affine alla *regressione ridge* per il fatto che è usato con predittori correlati. È simile alla regressione *stepwise*, in quanto l'algoritmo procede per selezionare un *sub-set* di variabili per costruire il modello. PLSR differisce da questi metodi in quanto si basa sulla regressione ai minimi quadrati, preceduta da una trasformazione dei dati originali in nuove variabili dette componenti principali (PCs) o variabili latenti (LVs), riducendo la dimensionalità del *dataset*. Detta trasformazione è una rotazione del sistema di riferimento delle variabili originali, come avviene con la PCA e con le PCR. Nel metodo delle PLS sia la matrice dei predittori (X) che la matrice della variabile indipendente (Y) vengono decomposte e le variabili latenti dei predittori vengono ruotate nella direzione di massima correlazione rispetto a quelle della variabile indipendente. Il metodo utilizza iterativamente coppie di variabili latenti su X e su Y ricercando la massima correlazione, pertanto la massimizzazione della varianza cercata tiene in considerazione la dipendenza tra variabile indipendente e variabile dipendente.

Come già specificato, per il caso in studio le PLS sono più adatte della PCR perché le variabili latenti dei predittori sono scelte massimizzando la correlazione con le variabili latenti della matrice Y (spessore suolo). Nell'impiego della PCR, la rotazione del sistema di riferimento, ovvero la determinazione delle componenti principali, avviene nelle direzioni di massima varianza dei predittori (X); queste direzioni non dipendono dalla matrice Y, quindi non è detto che siano significative per la costruzione di un modello che stima la Y (Wise et al. 2006). Per la PLSR, è stato utilizzato il medesimo *training set* utilizzato per le reti neurali.

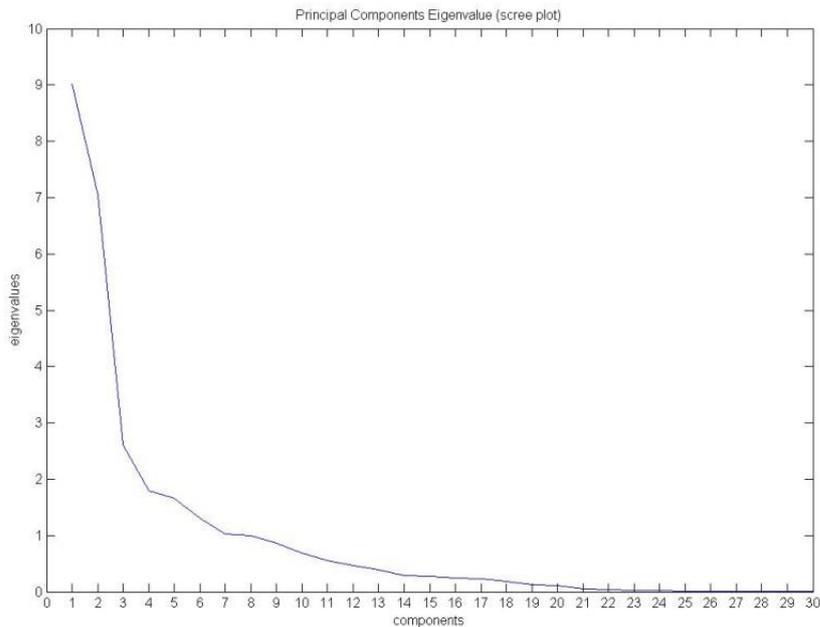
### 5.3 Analisi e risultati

Per evitare scelte discrezionali nella determinazione dei *sub-set* di variabili indipendenti sulle quali applicare le ANN, sono stati adottati due metodi di selezione delle variabili basati sui risultati della PCA (Tab. 5.1), secondo lo schema riportato in Figura 5.3.

In pratica, l'approccio consente di verificare quanto le variabili considerate influenzino la qualità della risposta, in relazione al loro numero ed alla loro presenza o assenza. Inoltre, i differenti *sub-sets* di variabili sono stati testati utilizzando diverse topologie di rete neurale e differenti algoritmi di calcolo.

Nell'analisi delle componenti principali generalmente la *Rank Analysis* è impiegata per individuare il numero di componenti da ritenere valido per le elaborazioni statistiche. Tale operazione, utilizzata per separare l'informazione fondamentale dal rumore, può essere eseguita con diverse tecniche (Zwick e Velicer 1986; Franklin et al. 1995). In questo caso la scelta per individuare le componenti significative è stata effettuata considerando: le caratteristiche dello *scree-plot* (Fig. 5.2); l'autovalore derivante dalla diagonalizzazione della matrice di covarianza, ossia la quota di varianza spiegata da ciascuna componente; ed i valori dei *loadings* che le singole variabili assumono nelle diverse componenti. I *loadings* hanno valori compresi tra -1 e +1 e forniscono un'indicazione di quanto una variabile originale è importante per rappresentare ciascuna componente principale.

Qualora una variabile abbia un *loading* prossimo ad 1 in valore assoluto nel nuovo sistema di riferimento, significa che essa è rappresentata da detta componente (Todeschini 1998).



**Fig. 5.2** Scree plot relativo alla *principal component analysis* eseguito sul set di 30 variabili geomorfometriche.

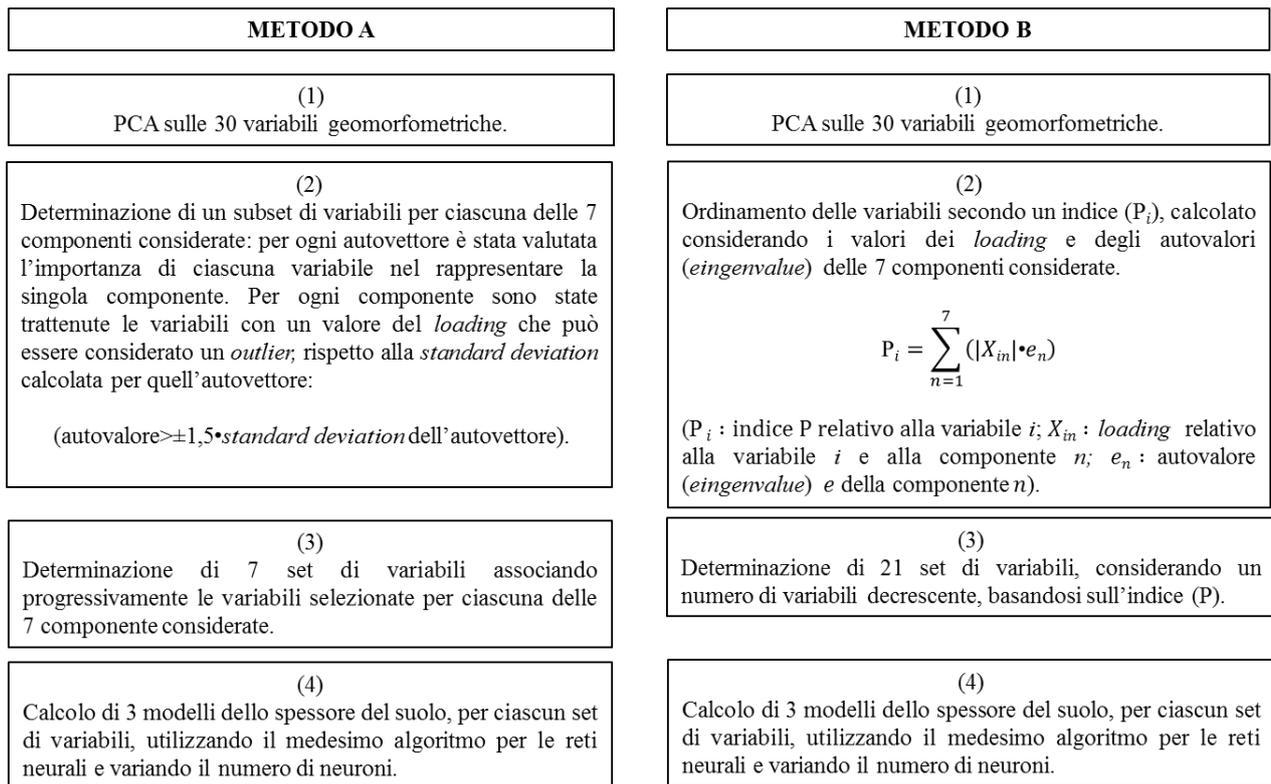
Senza considerare analisi più sofisticate (Crawford et al. 2010; Ledesma et al. 2015), dato l'ordinamento decrescente degli autovalori, l'analisi dello *scree plot* permette di individuare qualitativamente, in corrispondenza di ogni variazione di pendenza, l'importanza delle PCs. Quando in corrispondenza di un brusco cambio di pendenza vi è una differenza trascurabile tra la varianza spiegata dalla componente rispetto a quella successiva ai fini dello scorporo dell'informazione fondamentale veicolata dalle componenti, è inutile considerare la componente successiva al punto ove si verifica il cambio di pendenza. Osservando il grafico di Figura 5.2 è possibile individuare quattro variazioni di pendenza: in corrispondenza della terza, quarta, settima e quattordicesima componente. Considerando le percentuali di varianza spiegata, appare legittimo considerare un numero di componenti superiore a sette. Infatti, in sintesi, secondo i cambi di pendenza dello *scree plot*: le prime tre componenti spiegano il 62,2 % della varianza totale, le prime quattro componenti spiegano il 68,17 % della varianza totale, le prime sette spiegano il 81,45%, fino ad arrivare al 95,61 % nel caso di 14 componenti (v. Tab. 5.1). Occorre osservare che secondo la regola di Kaiser vengono considerate solo le PCs il cui autovalore è maggiore di uno (Zwick e Velicer 1986). Detta discriminazione è basata su due considerazioni: 1) l'autovalore di una PC rappresenta la varianza associata a ciascun autovettore o componente principale (colonna della matrice dei *loadings*); 2) le componenti principali sono calcolate con variabili standardizzate che hanno varianza unitaria; pertanto, si trattengono solo quelle componenti in grado di spiegare una quota di varianza maggiore di quella di una sola variabile. Secondo questa regola, nel caso in esame dovrebbero essere considerate solo le prime sette componenti principali (Tab. 5.1 e Fig. 5.2).

**Tab. 5.1** Valori dei *loadings* delle variabili geomorfometriche, distinti secondo l'importanza (intensità del colore) ed il segno (verdi negativi, rossi positivi). Per ciascuna componente è riportato l'autovalore (*eigenvalue*) e la percentuale di varianza spiegata (*percent\_explained*).

		PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12	PC13	PC14	PC15
percent_explained		<b>30,056</b>	<b>23,452</b>	<b>8,688</b>	<b>5,978</b>	<b>5,528</b>	<b>4,335</b>	<b>3,414</b>	<b>3,311</b>	<b>2,879</b>	<b>2,284</b>	<b>1,867</b>	<b>1,544</b>	<b>1,327</b>	<b>0,948</b>	<b>0,920</b>
eigenvalue		9,017	7,036	2,606	1,793	1,658	1,301	1,024	0,993	0,864	0,685	0,560	0,463	0,398	0,284	0,276
1	elevation	0,110	0,115	0,528	0,014	-0,027	0,191	-0,035	0,088	-0,103	-0,057	-0,008	-0,009	-0,006	0,010	-0,014
2	Terrain Ruggedness Index (QGIS)	0,141	0,336	-0,050	-0,042	-0,011	0,028	-0,013	0,006	-0,007	0,013	-0,008	-0,080	-0,033	0,017	-0,006
3	Topographic position index (QGIS)	0,285	-0,156	-0,070	-0,150	0,020	0,088	-0,037	0,081	0,032	-0,087	-0,058	-0,013	0,083	-0,033	0,015
4	Asperity (QGIS)	0,142	0,330	-0,051	-0,040	0,004	0,021	-0,014	0,009	-0,030	0,017	-0,020	-0,114	-0,043	0,098	0,008
5	total curvature (Jenness tool)	-0,094	0,123	0,103	0,083	-0,027	0,181	-0,018	-0,154	0,881	0,018	0,165	-0,141	-0,014	-0,041	0,042
6	tangential curvature (Jenness tool)	0,249	-0,126	-0,133	0,260	-0,025	0,206	-0,085	0,142	0,060	0,008	-0,291	0,062	0,101	0,040	-0,062
7	profile curvature (Jenness tool)	-0,213	0,129	-0,028	0,474	-0,056	0,077	-0,029	0,018	0,016	0,105	-0,222	0,073	-0,009	0,056	-0,070
8	plan curvature (Jenness tool)	0,170	-0,088	-0,074	0,336	-0,031	0,056	-0,103	0,266	-0,114	0,266	0,680	-0,172	0,090	0,071	-0,349
9	longitudinal curvature (Jenness tool)	-0,228	0,135	-0,011	0,441	-0,048	0,048	-0,021	0,013	0,003	0,112	-0,183	0,069	-0,039	0,055	0,017
10	general curvature (Jenness tool)	0,286	-0,157	-0,066	-0,152	0,021	0,086	-0,036	0,076	0,037	-0,068	-0,059	-0,007	0,086	0,003	-0,022
11	cross-sectional curvature (Jenness tool)	0,261	-0,133	-0,125	0,189	-0,013	0,199	-0,084	0,145	0,067	-0,003	-0,290	0,059	0,109	0,062	-0,021
12	aspect (Jenness tool)	-0,010	-0,010	-0,030	-0,031	-0,283	0,111	0,839	0,304	0,039	0,033	-0,019	-0,068	-0,042	0,283	0,093
13	slope (Jenness tool)	0,150	0,330	-0,059	-0,040	-0,008	0,014	-0,013	0,016	-0,067	0,013	-0,014	-0,081	-0,039	0,026	-0,007
14	wetness index (SAGA)	-0,267	-0,069	-0,047	-0,200	0,024	0,276	-0,060	0,073	0,017	0,027	-0,025	-0,001	0,296	0,127	-0,282
15	topographic wetness index (SAGA)	-0,246	-0,113	-0,119	-0,110	0,010	0,367	-0,052	0,012	-0,076	-0,028	0,101	0,003	0,200	0,018	-0,025
16	Compound Topographic Index (GRASS)	-0,252	-0,012	0,024	-0,135	0,044	-0,121	-0,105	0,363	0,053	-0,035	0,052	-0,058	0,489	0,129	0,293
17	catchment Area (recursive method) (SAGA)	-0,082	-0,059	-0,124	-0,089	-0,019	0,655	0,029	-0,391	-0,234	0,174	0,096	-0,094	-0,164	-0,021	0,236
18	mass balance index (SAGA)	0,267	-0,135	-0,035	-0,113	0,014	0,134	-0,036	0,031	0,268	-0,113	-0,021	-0,040	0,075	-0,040	0,044
19	LS factor (SAGA)	0,031	0,353	-0,036	-0,073	0,004	0,049	-0,042	0,084	0,027	0,029	0,131	-0,177	0,069	0,019	0,034
20	convergence index (SAGA)	0,242	-0,128	-0,069	0,263	-0,016	0,025	-0,047	0,141	-0,023	0,027	0,175	-0,004	-0,075	-0,160	0,617
21	channel network base level (SAGA)	0,075	0,118	0,503	0,027	-0,035	0,249	-0,073	0,177	-0,102	-0,264	0,017	-0,010	-0,144	0,037	-0,073
22	catchment Area (max flux method) (SAGA)	0,042	0,308	-0,102	0,015	-0,005	0,148	-0,055	0,125	-0,074	-0,011	-0,284	-0,161	0,288	-0,166	-0,076
23	altitude above channel network (SAGA)	0,162	0,042	0,322	-0,033	0,014	-0,097	0,106	-0,248	-0,048	0,643	-0,085	-0,002	0,441	-0,082	0,185
24	flow direction (ArcGIS)	0,014	0,006	0,029	0,076	0,716	0,017	-0,016	-0,021	-0,001	-0,012	-0,020	-0,087	-0,064	0,614	0,178
25	flow drop (ArcGIS)	0,192	0,293	-0,052	-0,132	0,003	0,032	-0,011	0,000	-0,024	-0,010	0,030	-0,103	-0,026	0,027	-0,013
26	diffuse insolation (SAGA)	-0,046	-0,214	0,486	0,062	-0,010	0,082	-0,010	0,060	-0,006	-0,055	-0,019	0,034	0,034	-0,036	0,013
27	direct insolation (SAGA)	-0,010	-0,076	0,015	-0,066	-0,615	-0,108	-0,405	-0,111	-0,028	0,008	-0,062	-0,212	-0,043	0,544	0,175
28	flow accumulation (SAGA)	-0,145	-0,001	-0,007	-0,307	0,054	0,079	-0,227	0,526	0,105	0,519	-0,140	0,115	-0,457	-0,092	0,056
29	Topographic position index (SAGA)	0,246	-0,145	0,062	-0,094	0,001	-0,059	0,065	-0,175	0,110	0,289	-0,112	0,079	-0,092	0,269	-0,367
30	Vector Ruggedness Measure (SAGA)	0,096	0,257	-0,049	-0,057	-0,107	0,100	-0,049	-0,025	0,050	-0,037	0,232	0,865	0,124	0,195	0,082

Tab. 5.1 (segue).

		PC16	PC17	PC18	PC19	PC20	PC21	PC22	PC23	PC24	PC25	PC26	PC27	PC28	PC29	PC30
percent explained		<b>0,818</b>	<b>0,744</b>	<b>0,628</b>	<b>0,409</b>	<b>0,371</b>	<b>0,185</b>	<b>0,104</b>	<b>0,085</b>	<b>0,053</b>	<b>0,032</b>	<b>0,027</b>	<b>0,009</b>	<b>0,006</b>	<b>0,000</b>	<b>0,000</b>
eigenvalue		0,245	0,223	0,188	0,123	0,111	0,055	0,031	0,025	0,016	0,010	0,008	0,003	0,002	1,71E-09	8,39E-14
1	elevation	0,003	-0,067	0,047	-0,008	0,053	-0,098	-0,057	-0,115	-0,044	-0,212	-0,093	0,025	-0,021	0,737	0,000
2	Terrain Ruggedness Index (QGIS)	0,066	-0,106	0,041	-0,038	0,052	-0,117	-0,177	-0,188	-0,210	0,256	0,626	0,080	-0,494	0,000	0,000
3	Topographic position index (QGIS)	-0,022	0,001	-0,095	-0,108	0,059	-0,018	-0,064	-0,014	0,773	-0,284	0,348	-0,053	-0,013	0,000	0,000
4	Asperity (QGIS)	0,047	-0,107	0,013	-0,032	0,107	-0,247	0,709	0,482	0,012	-0,106	0,014	0,017	-0,074	0,000	0,000
5	total curvature (Jenness tool)	-0,038	0,116	0,033	-0,222	0,039	-0,056	-0,011	-0,017	0,013	-0,028	-0,007	-0,016	0,041	0,000	0,000
6	tangential curvature (Jenness tool)	0,081	-0,104	-0,102	-0,209	0,001	0,120	0,284	-0,358	-0,097	0,113	-0,160	-0,538	-0,161	0,000	0,000
7	profile curvature (Jenness tool)	0,095	-0,106	-0,041	0,154	0,075	0,014	0,223	-0,337	0,284	0,125	0,027	0,547	0,135	0,000	0,000
8	plan curvature (Jenness tool)	-0,005	0,142	-0,145	0,019	0,043	-0,057	-0,023	0,023	-0,007	-0,002	0,016	0,003	-0,001	0,000	0,000
9	longitudinal curvature (Jenness tool)	0,065	-0,118	0,015	0,165	-0,011	-0,072	-0,337	0,397	0,099	-0,143	0,057	-0,319	-0,094	0,000	0,457
10	general curvature (Jenness tool)	0,005	0,007	-0,048	-0,228	-0,012	0,049	0,025	0,004	-0,187	0,052	-0,036	0,365	0,105	0,000	0,770
11	cross-sectional curvature (Jenness tool)	0,075	-0,109	-0,067	-0,225	-0,032	0,011	-0,303	0,415	-0,221	-0,056	-0,004	0,303	0,085	0,000	-0,446
12	aspect (Jenness tool)	-0,082	0,039	-0,048	0,003	0,008	0,013	-0,013	-0,018	0,000	0,001	0,002	0,000	0,000	0,000	0,000
13	slope (Jenness tool)	0,069	-0,111	0,025	-0,027	0,071	-0,092	-0,103	-0,100	-0,070	0,152	0,237	-0,251	0,806	0,000	0,000
14	wetness index (SAGA)	-0,226	-0,245	0,135	0,027	-0,590	-0,347	0,043	-0,033	0,037	0,072	0,047	-0,018	0,033	0,000	0,000
15	topographic wetness index (SAGA)	-0,277	-0,257	0,332	0,000	0,634	0,201	0,002	0,061	-0,025	0,061	0,030	-0,021	0,000	0,000	0,000
16	Compound Topographic Index (GRASS)	0,614	0,087	0,060	-0,009	0,091	-0,097	-0,006	-0,027	0,002	0,008	-0,017	-0,015	-0,004	0,000	0,000
17	catchment Area (recursive method) (SAGA)	0,354	0,194	-0,147	-0,002	-0,107	-0,015	0,002	-0,007	0,001	-0,011	-0,001	0,005	-0,001	0,000	0,000
18	mass balance index (SAGA)	0,020	-0,144	-0,254	0,820	0,074	-0,049	0,006	0,004	-0,113	0,028	-0,015	-0,006	0,009	0,000	0,000
19	LS factor (SAGA)	0,064	-0,281	0,030	0,051	-0,337	0,752	0,012	0,035	0,045	-0,167	-0,043	0,042	-0,031	0,000	0,000
20	convergence index (SAGA)	-0,168	-0,052	0,522	0,076	-0,226	-0,115	0,056	-0,035	0,030	0,067	0,006	0,031	0,019	0,000	0,000
21	channel network base level (SAGA)	0,055	-0,032	0,098	0,002	0,038	-0,099	-0,059	-0,117	-0,043	-0,219	-0,096	0,026	-0,022	-0,646	0,000
22	catchment Area (max flux method) (SAGA)	-0,265	0,706	0,142	0,159	-0,013	0,073	-0,015	0,006	-0,018	-0,041	-0,034	0,002	-0,010	0,000	0,000
23	altitude above channel network (SAGA)	-0,167	-0,144	-0,143	-0,035	0,071	-0,043	-0,020	-0,049	-0,022	-0,074	-0,031	0,010	-0,005	-0,200	0,000
24	flow direction (ArcGIS)	-0,191	0,102	-0,081	0,013	0,010	0,033	-0,035	-0,049	0,000	0,001	0,007	-0,001	0,001	0,000	0,000
25	flow drop (ArcGIS)	0,030	-0,112	0,008	-0,041	0,070	-0,146	-0,270	0,066	0,361	0,515	-0,553	0,025	-0,146	0,000	0,000
26	diffuse insolation (SAGA)	-0,014	0,148	-0,085	0,005	-0,101	0,274	0,164	0,321	0,122	0,597	0,262	-0,074	0,062	0,000	0,000
27	direct insolation (SAGA)	-0,180	0,073	-0,069	0,013	0,011	0,025	-0,026	-0,036	-0,002	-0,001	0,005	0,000	0,000	0,000	0,000
28	flow accumulation (SAGA)	-0,068	0,035	-0,080	0,014	0,043	0,011	0,004	-0,009	-0,003	0,004	-0,005	0,004	0,000	0,000	0,000
29	Topographic position index (SAGA)	0,336	0,117	0,610	0,164	0,002	0,097	0,003	0,017	0,048	-0,004	0,011	-0,013	-0,002	0,000	0,000
30	Vector Ruggedness Measure (SAGA)	-0,056	0,137	-0,061	0,019	-0,028	0,048	0,016	0,015	0,011	0,003	0,001	-0,001	0,001	0,000	0,000



**Fig. 5.3** Sequenza di operazioni dei METODI A e B per individuare i sub-set di variabili utilizzati per modellare lo spessore del suolo.

Date queste premesse, l'approccio adottato è basato sullo schema di Figura 5.3. Il primo metodo (METODO A) seleziona le variabili in base alla loro importanza nel rappresentare ciascuna delle sette componenti considerate, secondo un criterio restrittivo. Si ha così l'effetto di considerare solo le variabili che hanno una particolare importanza nel rappresentare una componente e, d'altro canto, di verificare l'effetto sul risultato, considerando un numero di variabili via via crescente. I sei *sub-sets* sono stati ottenuti assemblando progressivamente le variabili trattenute per ciascuna componente, come indicato in tabella 5.2, e sono costituiti rispettivamente da 4, 10, 14, 18, 21 e 23 variabili. I *sub-sets* di variabili adottati sono sei anziché sette perché il sesto *sub-set* è composto dalle stesse variabili del settimo (v. Tab. 5.2). Per ciascuno di essi sono stati calcolati 3 modelli utilizzando il metodo *Conjugate Gradient with Powell/Beale Restarts*, variando la topologia delle reti neurali, incrementando il numero di neuroni (10, 20 e 30). Il secondo metodo (METODO B) considera la variazione delle prestazioni sulla risposta, diminuendo il numero di variabili. Queste ultime sono state disposte in ordine crescente secondo un indice, che rappresenta quanto ciascuna di esse contribuisce a spiegare la varianza, in tutte le sette componenti considerate. L'indice è dato dalla sommatoria, sulle sette componenti, dei pesi che ciascuna variabile assume per ogni componente presa in esame. Il peso è calcolato come il prodotto tra l'autovalore, relativo a ciascuna di esse, e il modulo del valore del *loading*, che ciascuna variabile assume nella componente associata all'autovalore (v. Tab. 5.3). Sono stati determinati 21 set di variabili, il primo escludendo le 2 variabili con peso inferiore a 2, il secondo escludendo le 7 variabili con peso inferiore a 3 e i successivi escludendo di volta in volta una variabile, sino a calcolare un risultato con quattro variabili; è stato, inoltre, determinato un risultato con tutte le 30 variabili geomorfometriche.

**Tab. 5.2** Risultati dell'applicazione del METODO A. Sono evidenziati i valori dei *loadings* per le prime sette componenti (PC1-PC7), gli *outliers* per ciascuna componente (valori >±1,5 deviazione standard) e i sub-set di variabili selezionate, evidenziate in giallo.

		PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC1var4	PC2var10	PC3var14	PC4var18	PC5var21	PC6var23	PC7var23
eigenvalue		9,017	7,036	2,606	1,793	1,658	1,301	1,024	outlier													
1	elevation	Zstd	0,110	0,115	0,528	0,014	-0,027	0,191	-0,035							Zstd						
2	Terrain Ruggedness Index (QGIS)	TRI	0,141	0,336	-0,050	-0,042	-0,011	0,028	-0,013							TRI						
3	Topographic position index (QGIS)	TPI	0,285	-0,156	-0,070	-0,150	0,020	0,088	-0,037	0,285						TPI						
4	Asperit (QGIS)	asperit	0,142	0,330	-0,051	-0,040	0,004	0,021	-0,014							asperit						
5	total curvature (Jenness tool)	ctot	-0,094	0,123	0,103	0,083	-0,027	0,181	-0,018							ctot						
6	tangential curvature (Jenness tool)	ctang	0,249	-0,126	-0,133	0,260	-0,025	0,206	-0,085							ctang						
7	profile curvature (Jenness tool)	cprof	-0,213	0,129	-0,028	0,474	-0,056	0,077	-0,029			0,474				cprof						
8	Plan curvature (Jenness tool)	cplan	0,170	-0,088	-0,074	0,336	-0,031	0,056	-0,103				0,336			cplan						
9	longitudinal curvature (Jenness tool)	clong	-0,228	0,135	-0,011	0,441	-0,048	0,048	-0,021				0,441			clong						
10	general curvature (Jenness tool)	cgen	0,286	-0,157	-0,066	-0,152	0,021	0,086	-0,036	0,286						cgen						
11	cross-sectional curvature (Jenness tool)	ccross	0,261	-0,133	-0,125	0,189	-0,013	0,199	-0,084							ccross						
12	aspect (Jenness tool)	aspsharp	-0,010	-0,010	-0,030	-0,031	-0,283	0,111	0,839				-0,283		0,839	aspsharp						
13	slope (Jenness tool)	slsharp	0,150	0,330	-0,059	-0,040	-0,008	0,014	-0,013				0,330			slsharp						
14	wetness index (SAGA)	WI	-0,267	-0,069	-0,047	-0,200	0,024	0,276	-0,060	-0,267				0,276		WI						
15	topographic wetness index (SAGA)	TWI	-0,246	-0,113	-0,119	-0,110	0,010	0,367	-0,052					0,367		TWI						
16	Compound Topographic Index (GRASS)	CTIgrass	-0,252	-0,012	0,024	-0,135	0,044	-0,121	-0,105							CTIgrass						
17	catchment Area (recursive method) (SAGA)	catarea	-0,082	-0,059	-0,124	-0,089	-0,019	0,655	0,029					0,655		catarea						
18	mass balance index (SAGA)	massbal	0,267	-0,135	-0,035	-0,113	0,014	0,134	-0,036	0,267						massbal						
19	LS factor (SAGA)	LSfactor	0,031	0,353	-0,036	-0,073	0,004	0,049	-0,042						0,353	LSfactor						
20	convergence index (SAGA)	ConvInd	0,242	-0,128	-0,069	0,263	-0,016	0,025	-0,047							ConvInd						
21	channel network base level (SAGA)	channelNB	0,075	0,118	0,503	0,027	-0,035	0,249	-0,073						0,503	channelNB						
22	catchment Area (max flux method) (SAGA)	catchslope	0,042	0,308	-0,102	0,015	-0,005	0,148	-0,055						0,308	catchslope						
23	altitude above channel network (SAGA)	altichan	0,162	0,042	0,322	-0,033	0,014	-0,097	0,106						0,322	altichan						
24	flow direction (ArcGIS)	flowdir	0,014	0,006	0,029	0,076	0,716	0,017	-0,016				0,716			flowdir						
25	flow drop (ArcGIS)	flowdrop	0,192	0,293	-0,052	-0,132	0,003	0,032	-0,011						0,293	flowdrop						
26	diffuse insolation (SAGA)	difsol	-0,046	-0,214	0,486	0,062	-0,010	0,082	-0,010							0,486	difsol	difsol	difsol	difsol	difsol	difsol
27	direct insolation (SAGA)	dirsol	-0,010	-0,076	0,015	-0,066	-0,615	-0,108	-0,405				-0,615		-0,405	dirsol						
28	flow accumulation (SAGA)	flowaccS	-0,145	-0,001	-0,007	-0,307	0,054	0,079	-0,227				-0,307			flowaccS						
29	Topographic position index (SAGA)	topPosiS	0,246	-0,145	0,062	-0,094	0,001	-0,059	0,065							topPosiS						
30	Vector Ruggedness Measure (SAGA)	VRM	0,096	0,257	-0,049	-0,057	-0,107	0,100	-0,049							VRM						

**Tab. 5.3** Risultati dell'applicazione del METODO B. Nella tabella le variabili sono disposte secondo l'ordinamento decrescente dell'indice P.

			PC1	PC2	PC3	PC4	PC5	PC6	PC7								
eigenvalue			9,017	7,036	2,606	1,793	1,658	1,301	1,024	∏ PC1	∏ PC2	∏ PC3	∏ PC4	∏ PC5	∏ PC6	∏ PC7	∑ P
6	tangential curvature (Jenness tool)	ctang	0,249	-0,126	-0,133	0,260	-0,025	0,206	-0,085	2,248	0,883	0,348	0,467	0,042	0,268	0,087	4,342
11	cross-sectional curvature (Jenness tool)	ccross	0,261	-0,133	-0,125	0,189	-0,013	0,199	-0,084	2,352	0,934	0,326	0,339	0,022	0,258	0,086	4,317
10	general curvature (Jenness tool)	cgen	0,286	-0,157	-0,066	-0,152	0,021	0,086	-0,036	2,581	1,105	0,172	0,273	0,034	0,112	0,037	4,314
3	Topographic position index (QGIS)	TPI	0,285	-0,156	-0,070	-0,150	0,020	0,088	-0,037	2,568	1,095	0,182	0,270	0,033	0,115	0,038	4,301
25	flow drop (ArcGIS)	flowdrop	0,192	0,293	-0,052	-0,132	0,003	0,032	-0,011	1,736	2,063	0,137	0,237	0,005	0,042	0,011	4,231
15	topographic wetness index (SAGA)	TWI	-0,246	-0,113	-0,119	-0,110	0,010	0,367	-0,052	2,215	0,797	0,309	0,198	0,017	0,478	0,053	4,065
9	longitudinal curvature (Jenness tool)	clong	-0,228	0,135	-0,011	0,441	-0,048	0,048	-0,021	2,054	0,951	0,028	0,791	0,079	0,063	0,022	3,988
7	profile curvature (Jenness tool)	cprof	-0,213	0,129	-0,028	0,474	-0,056	0,077	-0,029	1,920	0,908	0,073	0,850	0,092	0,100	0,030	3,972
13	slope (Jenness tool)	slsharp	0,150	0,330	-0,059	-0,040	-0,008	0,014	-0,013	1,355	2,325	0,154	0,072	0,014	0,019	0,013	3,951
2	Terrain Ruggedness Index (QGIS)	TRI	0,141	0,336	-0,050	-0,042	-0,011	0,028	-0,013	1,273	2,363	0,130	0,076	0,018	0,037	0,014	3,910
18	mass balance index (SAGA)	massbal	0,267	-0,135	-0,035	-0,113	0,014	0,134	-0,036	2,411	0,947	0,091	0,203	0,023	0,174	0,036	3,887
4	Asperit (QGIS)	asperit	0,142	0,330	-0,051	-0,040	0,004	0,021	-0,014	1,283	2,322	0,133	0,072	0,006	0,028	0,014	3,859
20	convergence index (SAGA)	ConvInd	0,242	-0,128	-0,069	0,263	-0,016	0,025	-0,047	2,180	0,898	0,180	0,472	0,026	0,032	0,048	3,835
14	wetness index (SAGA)	WI	-0,267	-0,069	-0,047	-0,200	0,024	0,276	-0,060	2,407	0,485	0,121	0,359	0,039	0,359	0,061	3,831
29	Topographic position index (SAGA)	topPosiS	0,246	-0,145	0,062	-0,094	0,001	-0,059	0,065	2,214	1,019	0,162	0,168	0,002	0,076	0,066	3,708
1	elevation	Zstd	0,110	0,115	0,528	0,014	-0,027	0,191	-0,035	0,989	0,806	1,376	0,026	0,044	0,249	0,036	3,527
26	diffuse insolation (SAGA)	difsol	-0,046	-0,214	0,486	0,062	-0,010	0,082	-0,010	0,418	1,509	1,266	0,111	0,016	0,106	0,011	3,436
21	channel network base level (SAGA)	channelNB	0,075	0,118	0,503	0,027	-0,035	0,249	-0,073	0,677	0,830	1,311	0,048	0,058	0,323	0,075	3,322
30	Vector Ruggedness Measure (SAGA)	VRM	0,096	0,257	-0,049	-0,057	-0,107	0,100	-0,049	0,869	1,809	0,127	0,102	0,177	0,130	0,051	3,265
8	plan curvature (Jenness tool)	cplan	0,170	-0,088	-0,074	0,336	-0,031	0,056	-0,103	1,536	0,621	0,192	0,603	0,051	0,073	0,105	3,182
19	LS factor (SAGA)	LSfactor	0,031	0,353	-0,036	-0,073	0,004	0,049	-0,042	0,281	2,482	0,094	0,131	0,006	0,064	0,043	3,100
22	catchment Area (max flux method) (SAGA)	catchslope	0,042	0,308	-0,102	0,015	-0,005	0,148	-0,055	0,379	2,170	0,265	0,026	0,008	0,192	0,056	3,097
16	Compound Topographic Index (GRASS)	CTIgrass	-0,252	-0,012	0,024	-0,135	0,044	-0,121	-0,105	2,275	0,084	0,063	0,242	0,073	0,157	0,107	3,003
23	altitude above channel network (SAGA)	altichan	0,162	0,042	0,322	-0,033	0,014	-0,097	0,106	1,460	0,292	0,839	0,059	0,023	0,126	0,108	2,908
17	catchment Area (recursive method) (SAGA)	catarea	-0,082	-0,059	-0,124	-0,089	-0,019	0,655	0,029	0,744	0,415	0,322	0,159	0,031	0,852	0,030	2,554
5	total curvature (Jenness tool)	ctot	-0,094	0,123	0,103	0,083	-0,027	0,181	-0,018	0,847	0,865	0,268	0,149	0,046	0,235	0,018	2,428
27	direct insolation (SAGA)	dirsol	-0,010	-0,076	0,015	-0,066	-0,615	-0,108	-0,405	0,094	0,535	0,039	0,119	1,020	0,140	0,415	2,362
28	flow accumulation (SAGA)	flowaccS	-0,145	-0,001	-0,007	-0,307	0,054	0,079	-0,227	1,311	0,005	0,018	0,550	0,090	0,103	0,233	2,309
12	aspect (Jenness tool)	aspsharp	-0,010	-0,010	-0,030	-0,031	-0,283	0,111	0,839	0,089	0,072	0,079	0,055	0,470	0,144	0,859	1,768
24	flow direction (ArcGIS)	flowdir	0,014	0,006	0,029	0,076	0,716	0,017	-0,016	0,122	0,046	0,076	0,137	1,187	0,022	0,016	1,606

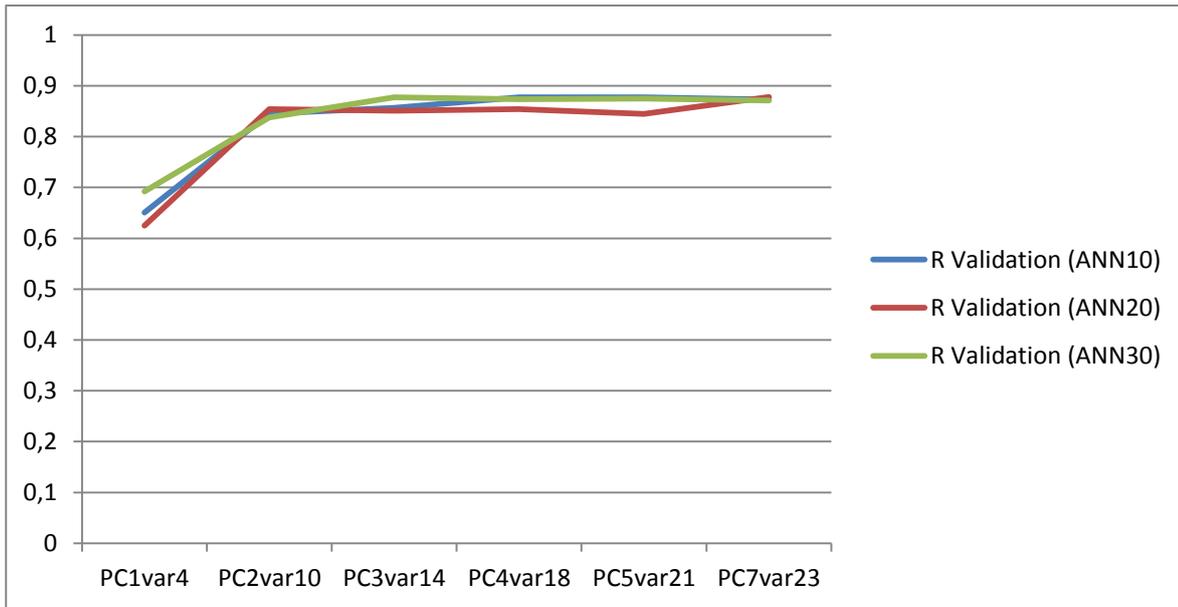
Nelle tabelle che seguono (Tab. 5.4 e Tab. 5.5) sono riportati i coefficienti di correlazione relativi ai migliori risultati ottenuti dall'applicazioni delle ANN ai differenti sub-set di variabili selezionate. Per ciascuna elaborazione, relativa ad una topologia di rete ed a un sub-set di variabili, è stato considerato il miglior risultato di 30 differenti processi di training (30 ripetizioni iniziando la ANN con la funzione *init*). Nella tabella 5.4 sono riportati i risultati relativi al METODO A, mentre nelle tabelle 5.5 sono riportati quelli relativi al METODO B.

**Tab. 5.4** Coefficienti di correlazione ottenuti con i sub-set di variabili selezionati con il METODO A, secondo le differenti topologie di ANN (ANN10: rete neurale con 10 neuroni, ANN20: rete neurale con 20 neuroni, ANN30: rete neurale con 30 neuroni).

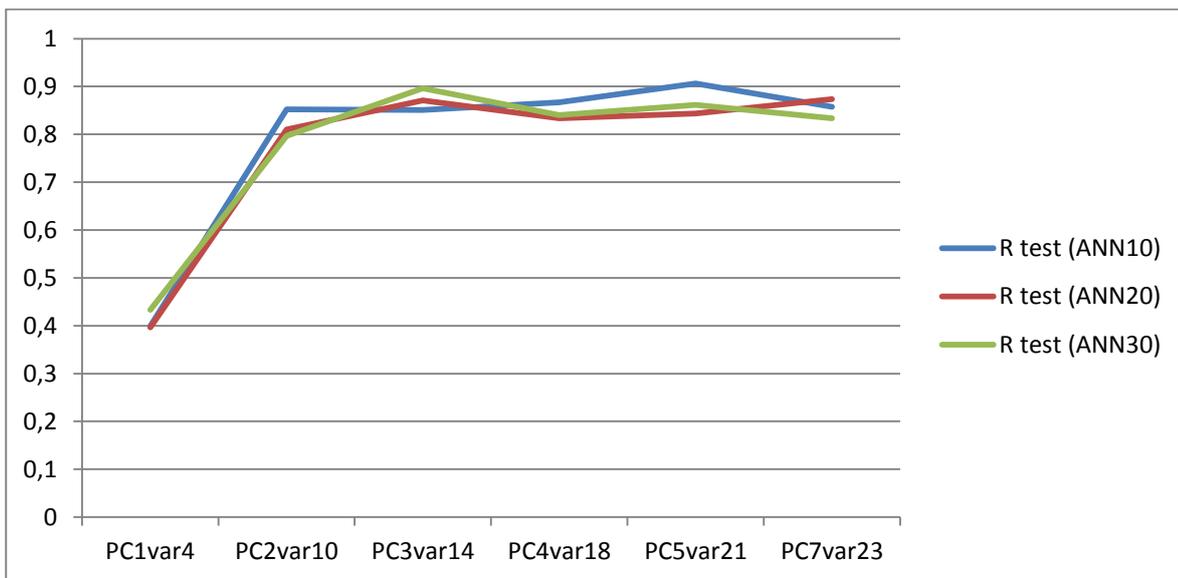
	PC1var4	PC2var10	PC3var14	PC4var18	PC5var21	PC7var23
R Training ANN10	0,71	0,90	0,88	0,87	0,87	0,92
R Validation ANN10	0,65	0,84	0,86	0,88	0,88	0,87
R Test ANN10	0,40	0,85	0,85	0,87	0,91	0,86
R All ANN10	0,58	0,88	0,88	0,88	0,89	0,90
R Training ANN20	0,69	0,82	0,87	0,89	0,90	0,90
R Validation ANN20	0,62	0,85	0,85	0,85	0,84	0,88
R Test ANN20	0,40	0,81	0,87	0,83	0,84	0,87
R All ANN20	0,57	0,84	0,88	0,87	0,88	0,90
R Training ANN30	0,73	0,85	0,89	0,90	0,93	0,93
R Validation ANN30	0,69	0,84	0,88	0,87	0,87	0,87
R Test ANN30	0,43	0,80	0,90	0,84	0,86	0,83
R All ANN30	0,61	0,83	0,90	0,88	0,90	0,88

**Tab. 5.5** Coefficienti di correlazione ottenuti con i sub-set di variabili selezionati con il METODO B, secondo le differenti topologie di ANN (ANN10: rete neurale con 10 neuroni, ANN20: rete neurale con 20 neuroni, ANN30: rete neurale con 30 neuroni).

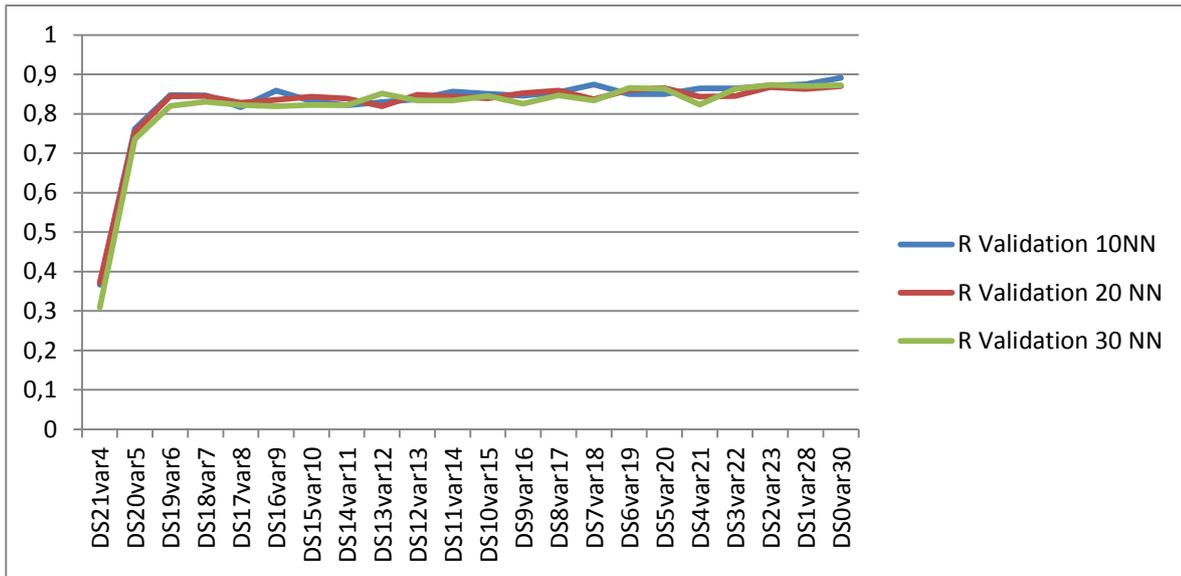
	DS0var30	DS1var28	DS2var23	DS3var22	DS4var21	DS5var20	DS6var19	DS7var18	DS8var17	DS9var16	DS10var15
R Training ANN10	0,92	0,91	0,87	0,90	0,89	0,89	0,85	0,91	0,87	0,87	0,85
R Validation ANN10	0,89	0,88	0,87	0,86	0,86	0,85	0,85	0,87	0,85	0,85	0,85
R Test 10 ANN10	0,88	0,87	0,80	0,74	0,76	0,79	0,80	0,86	0,84	0,78	0,75
R All ANN10	0,91	0,90	0,86	0,84	0,84	0,85	0,84	0,90	0,87	0,84	0,82
R Training ANN20	0,91	0,90	0,89	0,88	0,87	0,88	0,88	0,91	0,87	0,87	0,85
R Validation ANN20	0,87	0,86	0,87	0,85	0,84	0,87	0,86	0,84	0,86	0,85	0,84
R Test ANN20	0,83	0,83	0,75	0,73	0,74	0,77	0,73	0,83	0,82	0,80	0,78
R All ANN20	0,88	0,87	0,84	0,82	0,82	0,84	0,83	0,88	0,86	0,85	0,83
R Training ANN30	0,93	0,91	0,89	0,90	0,89	0,91	0,91	0,92	0,87	0,87	0,86
R Validation ANN30	0,87	0,87	0,87	0,86	0,82	0,86	0,87	0,83	0,85	0,83	0,85
R Test ANN30	0,79	0,82	0,77	0,71	0,79	0,74	0,78	0,81	0,82	0,76	0,75
R All ANN30	0,87	0,87	0,85	0,87	0,85	0,84	0,86	0,87	0,86	0,83	0,82
	DS11var14	DS12var13	DS13var12	DS14var11	DS15var10	DS16var9	DS17var8	DS18var7	DS19var6	DS20var5	DS21var4
R Training ANN10	0,84	0,84	0,81	0,81	0,80	0,80	0,80	0,79	0,79	0,75	0,38
R Validation ANN10	0,86	0,84	0,83	0,82	0,83	0,86	0,82	0,85	0,85	0,76	0,37
R Test ANN10	0,76	0,80	0,77	0,77	0,72	0,77	0,69	0,76	0,76	0,74	0,14
R All ANN10	0,82	0,83	0,81	0,80	0,78	0,81	0,76	0,80	0,80	0,76	0,23
R Training ANN20	0,84	0,84	0,81	0,80	0,80	0,79	0,81	0,79	0,79	0,76	0,42
R Validation ANN20	0,84	0,85	0,82	0,84	0,84	0,84	0,83	0,85	0,84	0,75	0,37
R Test ANN20	0,76	0,73	0,73	0,74	0,73	0,72	0,70	0,73	0,74	0,70	0,22
R All ANN20	0,82	0,80	0,79	0,80	0,79	0,78	0,78	0,79	0,79	0,75	0,31
R Training ANN30	0,85	0,83	0,82	0,81	0,80	0,80	0,78	0,79	0,79	0,76	0,49
R Validation ANN30	0,83	0,83	0,85	0,82	0,82	0,82	0,82	0,83	0,82	0,74	0,31
R Test ANN30	0,73	0,71	0,76	0,71	0,72	0,69	0,76	0,74	0,74	0,69	0,08
R All ANN30	0,81	0,79	0,81	0,78	0,78	0,77	0,79	0,79	0,79	0,73	0,22



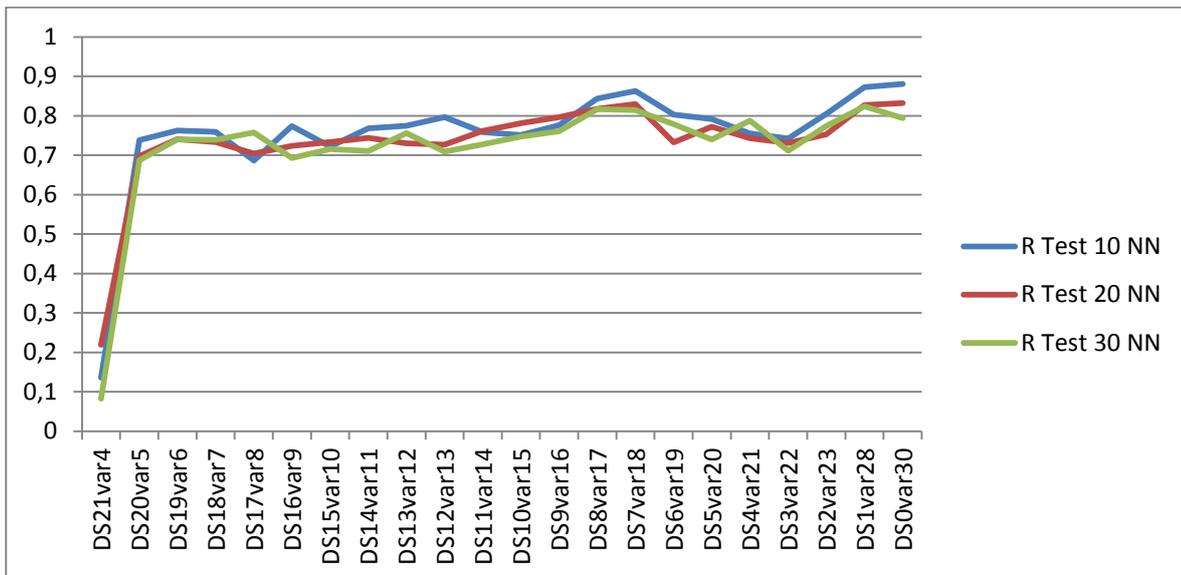
**Fig. 5.4** Variazione del coefficiente di correlazione per il *validation set* rispetto ai differenti sub-set di variabili selezionati con il METODO A.



**Fig. 5.5** Variazione del coefficiente di correlazione per il *test set* rispetto ai differenti sub-set di variabili selezionati con il METODO A.

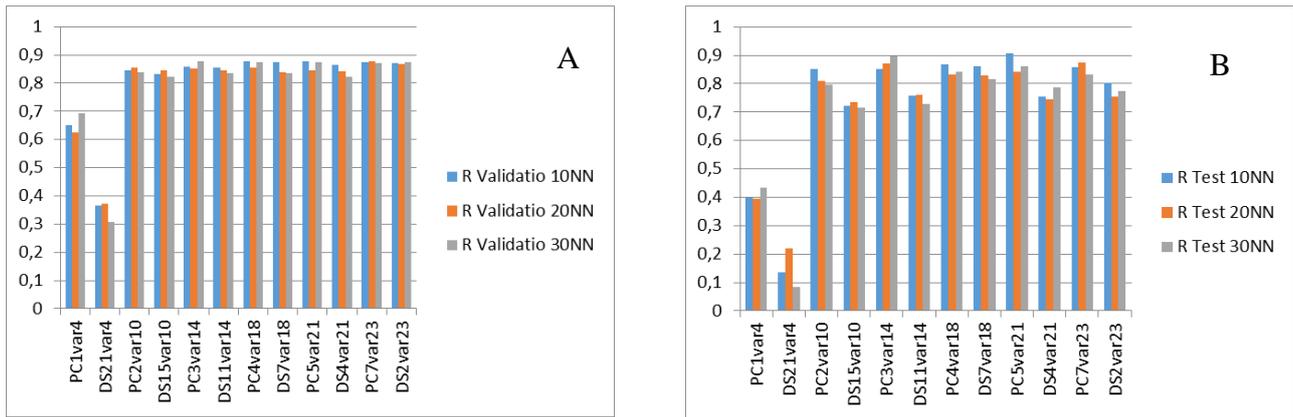


**Fig. 5.6** Variazione del coefficiente di correlazione per il *validation set* rispetto ai differenti sub-set di variabili selezionati con il METODO B.



**Fig. 5.7** Variazione del coefficiente di correlazione per il *test set* rispetto ai differenti sub-set di variabili selezionati con il METODO B.

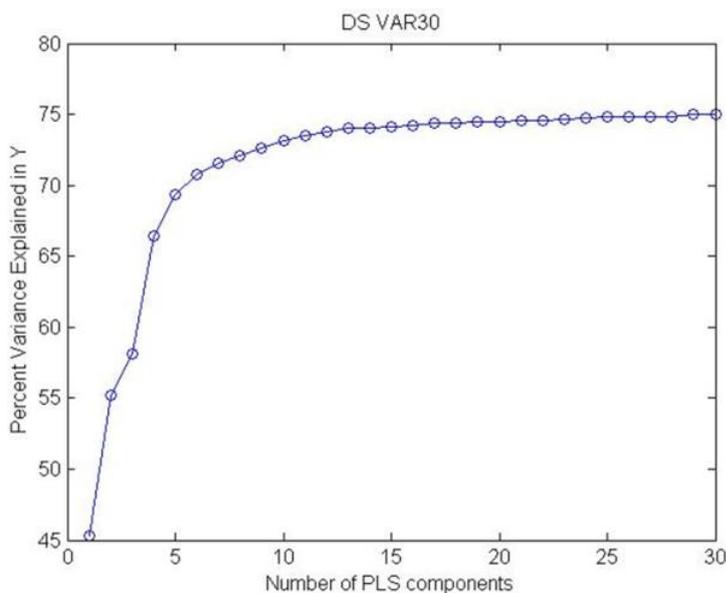
Osservando le variazioni dei valori del coefficiente di correlazione  $R$ , relativamente al *validation set* (Fig. 5.4 e Fig. 5.6) ed al *test set* (Fig. 5.5 e Fig. 5.7), emerge che le prestazioni delle ANN migliorano tendenzialmente all'aumentare del numero di variabili considerate nel calcolo, per entrambi i metodi di selezione delle variabili. Emerge, inoltre, che le reti neurali con una topologia costituita da dieci neuroni hanno generalmente prestazioni superiori a quelle con un numero di neuroni superiore.



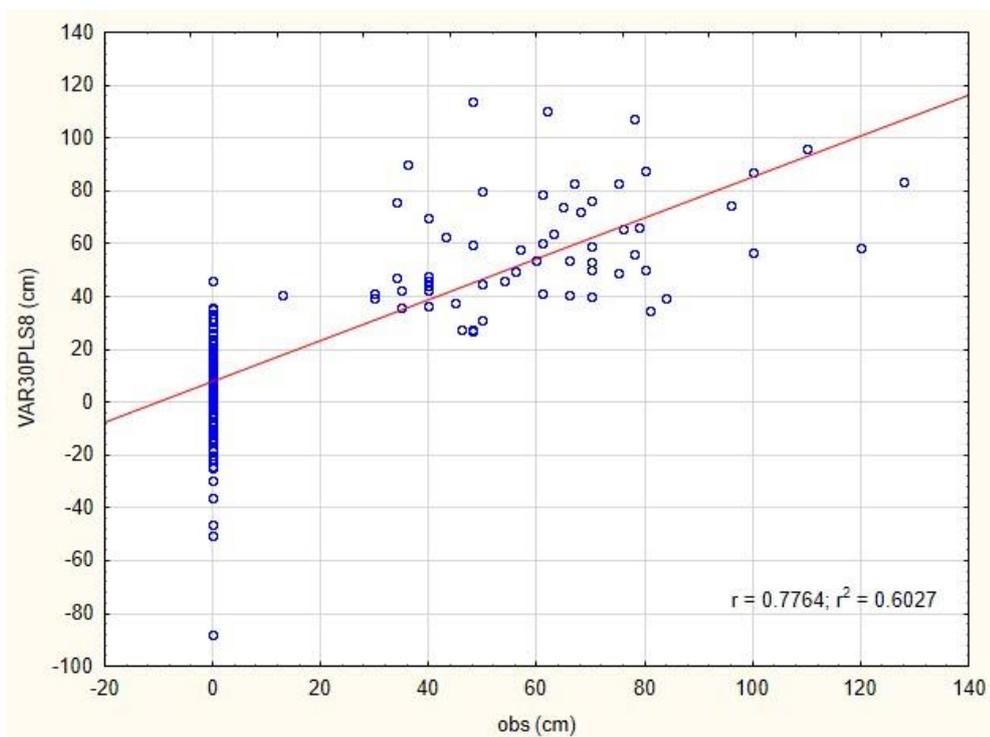
**Fig. 5.8** Confronto tra i valori dei coefficienti di correlazione a parità di numero di variabili per i sub-set determinati con entrambi i metodi di selezione (sub-sets METODO A: PC1var4, PC2var10, PC3var14, PC4var18, PC5var21, PC7var23; sub-sets METODO B: DS1VAR4, DS15var10, DS11var14, DS7var18, DS4var21, DS2var23) relativamente a: *Validation set* (A) e *test set* (B) per le diverse configurazioni di ANN (10 neuroni, 20 neuroni, 30 neuroni).

A parità di numero di variabili si ottengono i migliori risultati utilizzando le variabili selezionate con il METODO A (basato sui valori dei *loadings* che presentano valori estremi) anche se le prestazioni risultano paragonabili (v. Fig. 5.8. A e B).

Analogamente a quanto svolto con le ANN, con il metodo di PLSR sono stati calcolati diversi modelli per verificare quanto il numero delle nuove variabili (le componenti principali) influisca sulla qualità della risposta. Tipicamente, anche per questo metodo è possibile adottare una scelta sul numero di componenti da utilizzare per la regressione, osservando il plot relativo alla varianza spiegata rispetto al numero di componenti (Fig. 5.9). Dall'esame del grafico emerge che a partire dalla quinta componente la percentuale di varianza spiegata ha bassi incrementi all'aumentare del numero di componenti trattenute, sino ad assumere valori quasi asintotici intorno alla ventesima componente. Sono stati comunque testati tutti i set di componenti a partire dalla quinta componente sino alla ventesima e sono stati calcolati i modelli relativamente a 25 componenti e a tutte le componenti trattenute.



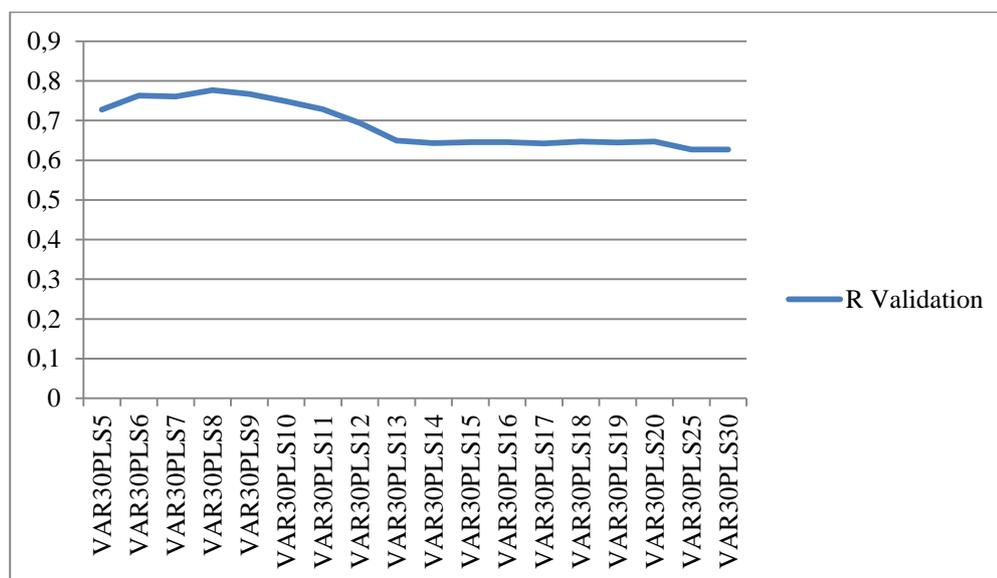
**Fig. 5.9** Percentuale di varianza spiegata rispetto al numero di componenti per il metodo PLSR.



**Fig. 5.10** Scatter plot relativo ai valori calcolati (VAR30PLS8) rispetto ai valori osservati (obs), del miglior risultato ottenuto con il metodo PLSR, calcolato con 8 componenti.

**Tab. 5.6** Valori dei coefficiente di correlazione R ottenuti, con il metodo PLSR, variando il numero di componenti trattenute.

	VAR30PLS5	VAR30PLS6	VAR30PLS7	VAR30PLS8	VAR30PLS9	VAR30PLS10	VAR30PLS11	VAR30PLS12	VAR30PLS13
R Validation	0,727	0,763	0,761	<b>0,776</b>	0,767	0,748	0,729	0,694	0,650
	VAR30PLS14	VAR30PLS15	VAR30PLS16	VAR30PLS17	VAR30PLS18	VAR30PLS19	VAR30PLS20	VAR30PLS25	VAR30PLS30
R Validation	0,643	0,645	0,645	0,642	0,647	0,645	0,647	0,627	0,627



**Fig. 5.11** Grafico relativo alla variazione del coefficiente di correlazione R per il metodo PLSR, al variare del numero di componenti trattenute.

Dai risultati emerge che, nonostante il graduale incremento della varianza spiegata all'aumentare del numero di componenti, il coefficiente di correlazione peggiora utilizzando più di otto componenti. Emerge, inoltre, che: utilizzando i *sub-set* costituiti da 6,7,8 e 9 componenti si ottengono prestazioni paragonabili; il miglior risultato è quello ottenuto trattenendo 8 componenti (R per il *validation set* 0,77) (v.Tab. 5.6 e Fig. 5.10, 5.11).

Considerando i risultati ottenuti, applicando sia le ANN che le PLSR (cfr. Tab. 5.4, 5.5, 5.6), emerge che le prestazioni che si ottengono con il metodo PLSR sono tendenzialmente inferiori a quello delle reti neurali. Per quanto riguarda i risultati ottenuti con l'applicazione delle reti neurali considerando i valori dei coefficienti di correlazione per il *validation set* ed il *test set*, il modello con la migliore prestazione è quello prodotto con il *set* DS0VAR30 (*set* completo di variabili), utilizzando una ANN con una topologia costituita da 10 neuroni (v. Fig. 5.12 A), la cui *performance* (v. Fig. 5.12 B) non esibisce sintomi di *overfitting*. La mappa è stata prodotta portando a zero i valori stimati inferiori allo zero e impostando come limite massimo 2 m di spessore (cfr. WRB - *World Reference Base for Soil Resources* (IUSS Working Group 2014)) ed è riportata in Figura 5.13 (A e B).

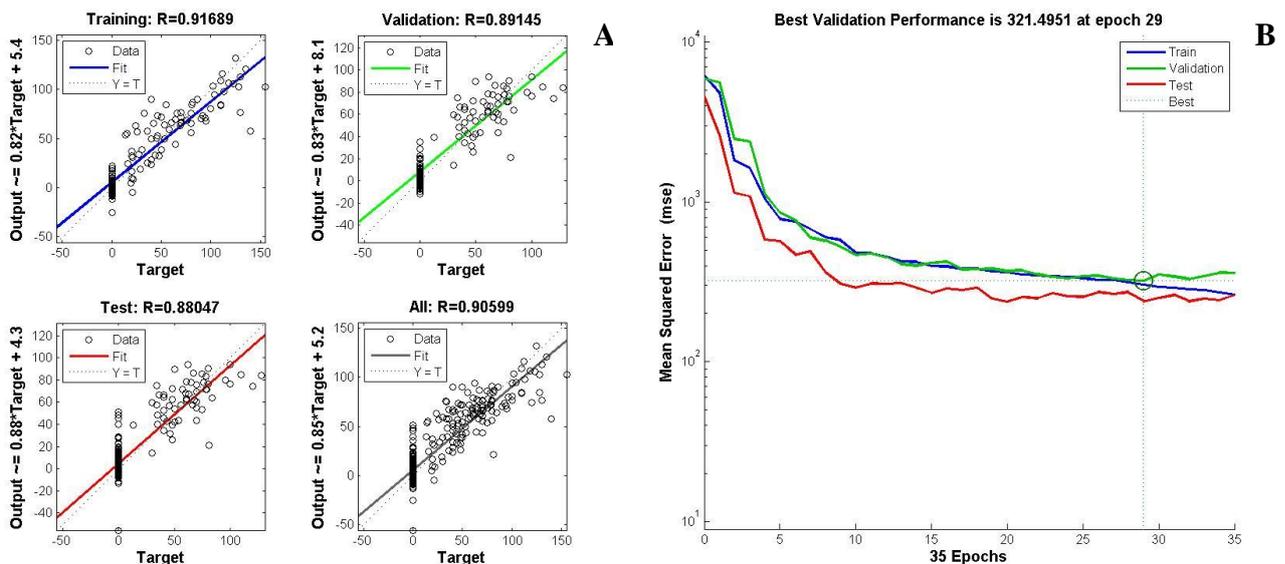
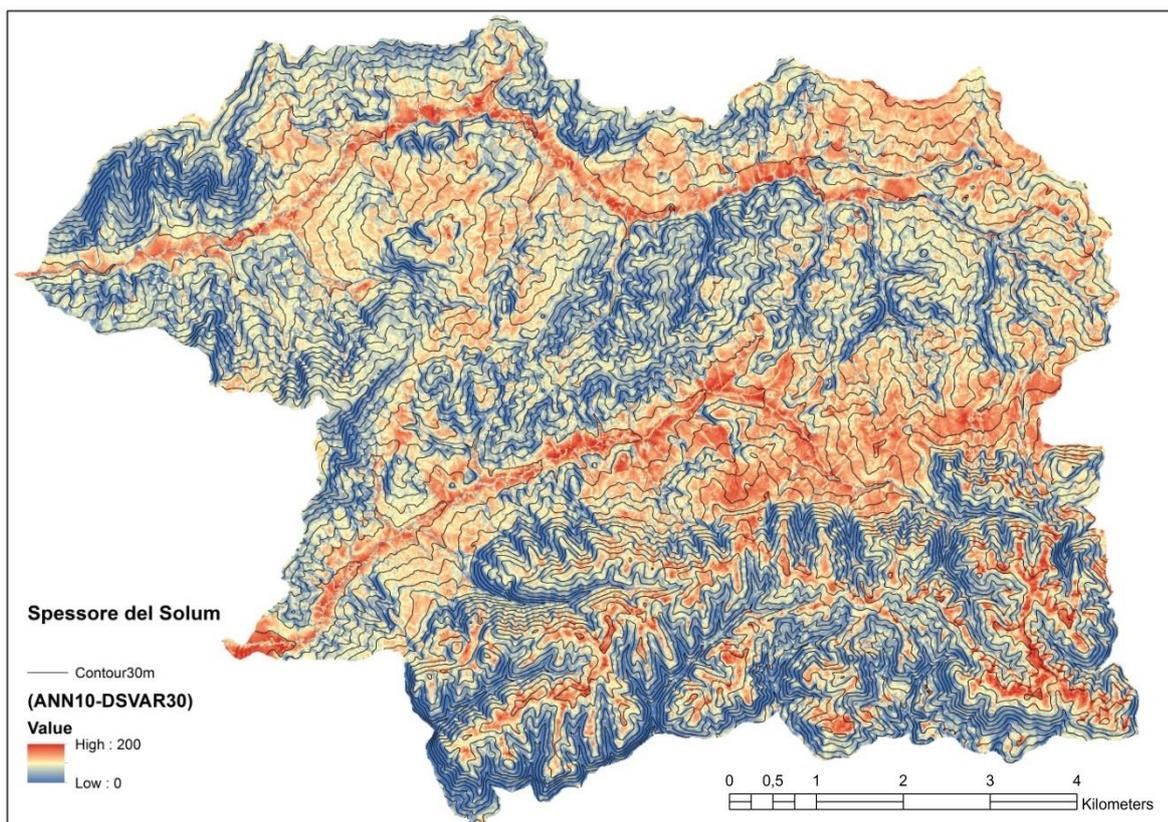


Fig 5.12 *Scatter plots* relativi ai valori calcolati rispetto ai valori predetti, per il set di variabili DS0VAR30 a cui è stata applicata una ANN costituita da 10 neuroni (A) e curva della sua *performance*. Essa rappresenta l'andamento dell'errore prodotto (in termini di errore quadratico medio) per il valore calcolato rispetto al *target* osservato, per ciascuno dei tre set (*training*, *validation* e *test*).

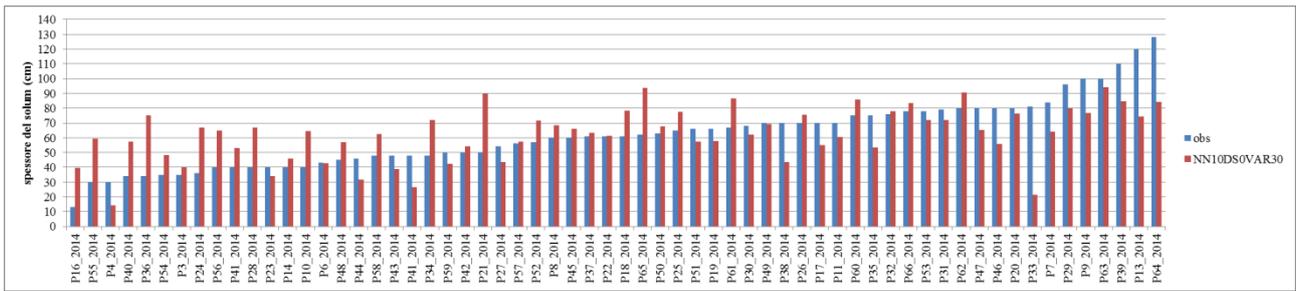


**Fig. 5.13** Immagine relativa alla mappa dello spessore del *solum*, calcolata con il dataset DS0VAR30 e una ANN costituita da 10 neuroni, rappresentata unitamente alle isoipse a 30 m.

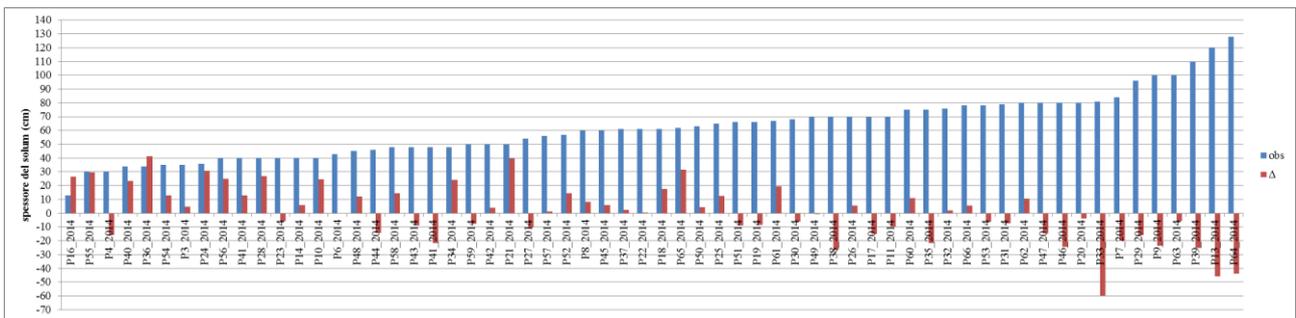
## 5.4 Discussione e conclusioni

Per meglio comprendere la capacità predittiva di questi metodi è stata effettuata una ulteriore analisi sui dati del *validation set*. Sono stati analizzati i residui, calcolati come differenza tra il valore osservato ed il valore calcolato, per il set utilizzato per la validazione. Questa analisi non fornisce indicazioni sulla qualità generale del modello, ma è indicativa di quanto gli errori determinati dal modello siano compatibili con la rappresentazione della realtà.

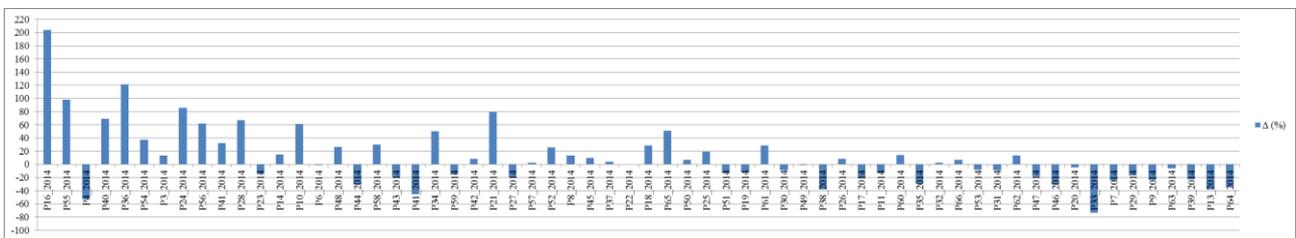
Per la valutazione dei residui sono stati creati due istogrammi relativi allo spessore del *solum* dei profili di suolo campionati: nel primo sono rappresentati i valori osservati e quelli predetti (v. Fig. 5.14), nel secondo i valori osservati e la differenza tra l'osservato ed il predetto, cioè la sovrastima ( $\Delta > 0$ ) o la sottostima ( $\Delta < 0$ ) (v. Fig. 5.15). Un ulteriore istogramma riguarda i punti zero, sul quale è rappresentata la differenza tra valore osservato e valore predetto. Per quanto riguarda i valori predetti sui profili di suolo, i risultati mostrano che nel 18% dei casi viene predetto il valore esatto, nel 23% dei casi l'errore in valore assoluto è entro i 10 cm, nel 25% dei casi l'errore è tra 10 e 20 cm, infine per una quota pari al 22% dei casi l'errore è tra 20 e 30 cm. Al di là delle valutazioni statistiche, per meglio comprendere se il risultato sia in grado di rappresentare la realtà in modo accettabile, occorre osservare i casi estremi relativi alle differenze tra osservato e predetto. Tali casi si verificano per i seguenti profili pedologici: P16\_2014, P55\_2014, P36\_2014, P40\_2014, P24\_2014, P21\_2014, P28\_2014, P33\_2014, P13\_2014, P64\_2014.



**Fig. 5.14** Istogramma relativo allo spessore del *solum* dei profili utilizzati per il *validation set* (barre blu) rispetto ai relativi valori calcolati con il metodo NN10DS0VAR30 (barre rosse).

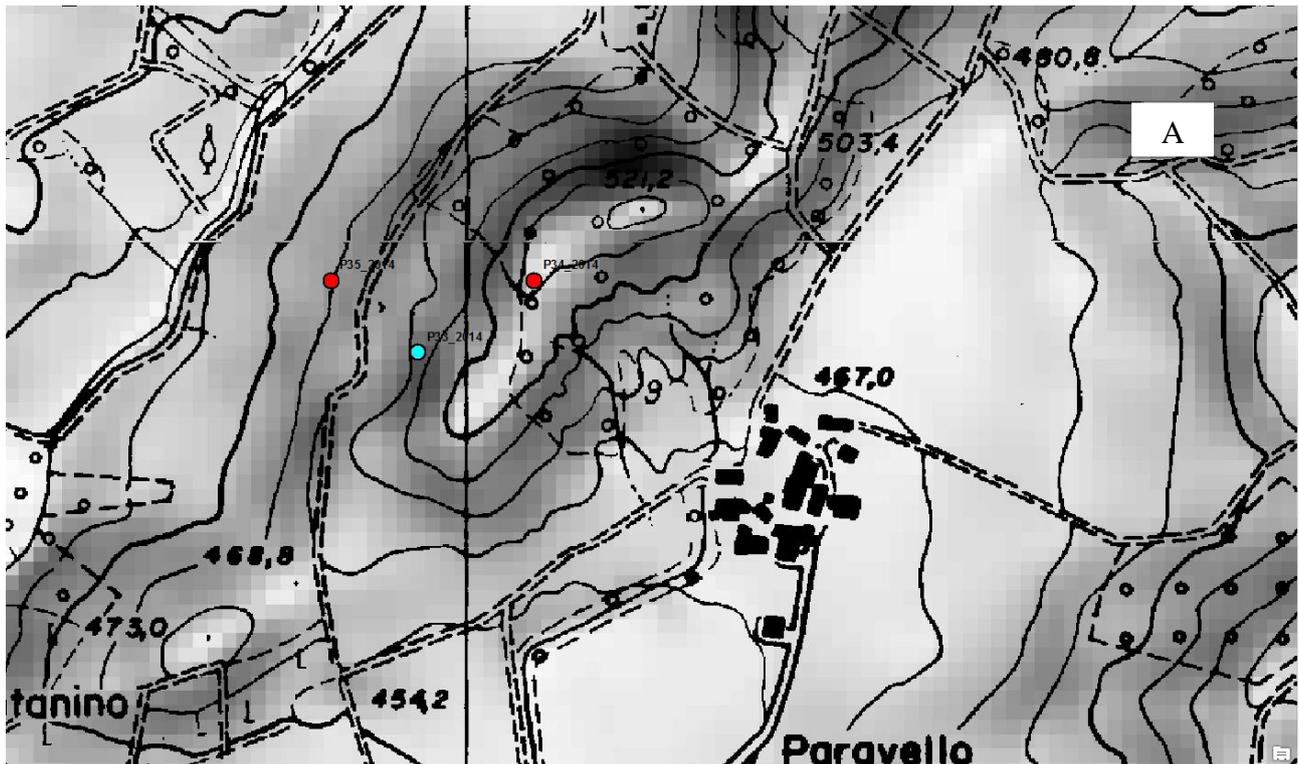


**Fig. 5.15** Istogramma relativo allo spessore del *solum* dei profili utilizzati per il *validation set* (barre blu) rispetto alla differenza ( $\Delta$ ) tra il valore calcolato ed il valore osservato (barre rosse): la sovrastima ( $\Delta > 0$ ) ha un valore di ordinata positivo, opposto alla sottostima ( $\Delta < 0$ ).



**Fig. 5.15** Istogramma relativo alla differenza tra il valore calcolato ed il valore osservato ( $\Delta$ ), espressa in percentuale rispetto all'osservato.

Per valutare l'importanza (anche pratica) di tali differenze viene considerata la scala adottata da ERSAF (Regione Lombardia) per discretizzare lo spessore del suolo nella cartografia di semidettaglio. Lo spessore del suolo viene distinto in: molto sottile (<25 cm), sottile (25-50 cm), moderatamente profondo (50-100), profondo (100-150 cm), molto profondo (>150 cm). Osservando le classi osservate rispetto a quelle assegnate (Tab. 5.7), emerge che nel caso delle sovrastime si ha un passaggio alla classe successiva per il valore predetto, mentre nel caso di sottostima, del profilo P33\_2014, si ha un salto di due classi. Detto profilo è posizionato su un versante di una piccola collina, in località Pavarello, ove la vegetazione mostra segni di stress associabili ad una instabilità del versante (Fig. 5.16 B). Osservando il DEM o le curve di livello (Fig. 5.16 A) questo profilo risulta ubicato su un brusco cambio di pendenza, mentre i profili P34\_2014 ed il P35\_2014, campionati a poca distanza (affetti da un errore di circa 20 cm) e sono posti in posizione sub-pianeggiante. Non si ritiene tuttavia che vi siano evidenti anomalie che possano generare tale errore.

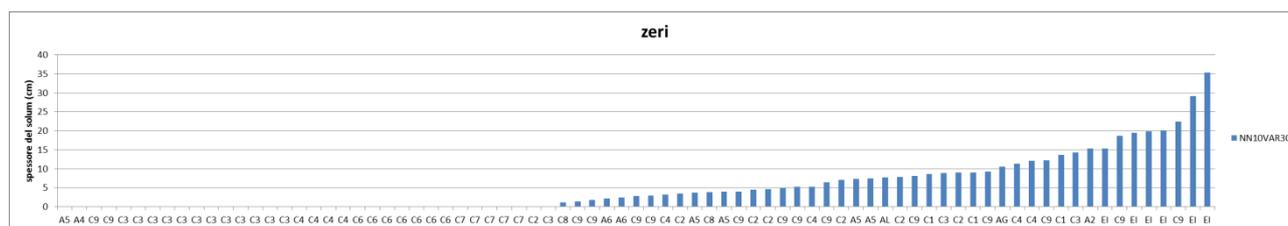


**Fig. 5.16** Area ove è stato campionato il profilo P33\_2014, con l'indicazione dell'ubicazione del profilo (punto evidenziato in azzurro) e dei profili P34 e 53\_2014, rilevati in prossimità; il raster sotteso è il DEM discretizzato a 10 m (A) immagine del versante sul quale è stato campionato il P33\_2014, ove sono visibili le deformazione dei fusti delle piante dovute a stress di natura meccanica (B).

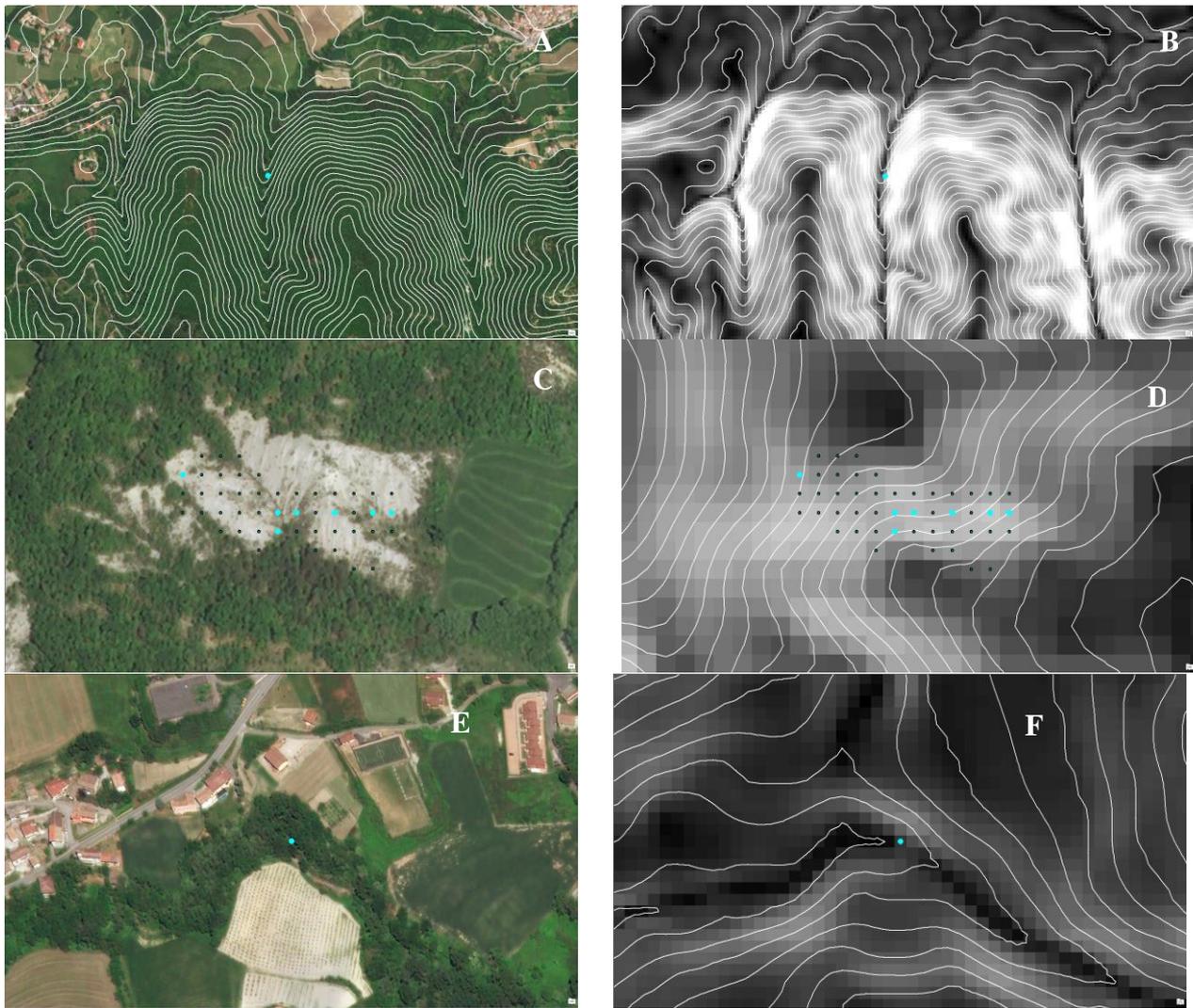
**Tab. 5.7** Assegnazione della classe di spessore di suolo secondo la scala adottata da ERSAF (Regione Lombardia), per i casi che presentano l'errore maggiore.

prof.	obs	classe vera	NN10DS0VAR30	classe calcolata	$\Delta$ spessore	$\Delta$ classe
<b>P16_2014</b>	13,00	<25	39,50	25-50	-26,50	1
<b>P55_2014</b>	30,00	25-50	59,52	50-100	-29,52	1
<b>P40_2014</b>	34,00	25-50	57,50	50-100	-23,50	1
<b>P36_2014</b>	34,00	25-50	75,19	50-100	-41,19	1
<b>P24_2014</b>	36,00	25-50	66,92	50-100	-30,92	1
<b>P28_2014</b>	40,00	25-50	66,85	50-100	-26,85	1
<b>P21_2014</b>	50,00	25-50	89,74	50-100	-39,74	1
<b>P33_2014</b>	81,00	50-100	21,37	<25	59,63	2
<b>P13_2014</b>	120,00	100-150	74,33	50-100	45,67	1

Nel caso dei punti con valore zero di spessore del solum, nel 56% dei casi il valore predetto è corretto o comunque entro i 5 cm di errore, a cui si aggiunge una quota del 22% con un errore entro 10 cm, ed una quota pari al 16% con un errore tra 10 e 20 cm (v. Fig. 5.17). La rimanente parte con gli errori maggiori, per un massimo di 35 cm, sono punti principalmente assegnati agli impluvi, soggetti ad erosione da flusso idrico concentrato o a calanchi. In particolare dei tre punti che calcolato son errori superiori a 20 cm vi è un punto mal modellato relativo al calanco C9 indicato nella figura 5.18 C e D. Come indicato nel Cap. 6, sui calanchi vi possono essere zone ove l'accumulo di materiale colluviale crea la situazione favorevole al radicamento della vegetazione: parte di tali punti sembrerebbe corrispondere ad aree di questo tipo. L'esame a posteriori porta a considerare anche l'ipotesi che alcuni di essi non avrebbero dovuto essere inclusi nel set per i punti con *solum* pari a zero. Ugualmente, per quanto concerne gli impluvi, in un caso si può osservare che il punto appartenente alla classe EI con spessore predetto di 28 cm, secondo il DEM (Fig. 5.18 A e B), non corrisponde all'impluvio ma al versante. Analogamente, per il punto con l'errore più elevato, osservando la pendenza relativa al suo corrispondente pixel (Fig. 5.18 E e F), si desume che esso non rappresenta adeguatamente il letto del canale.

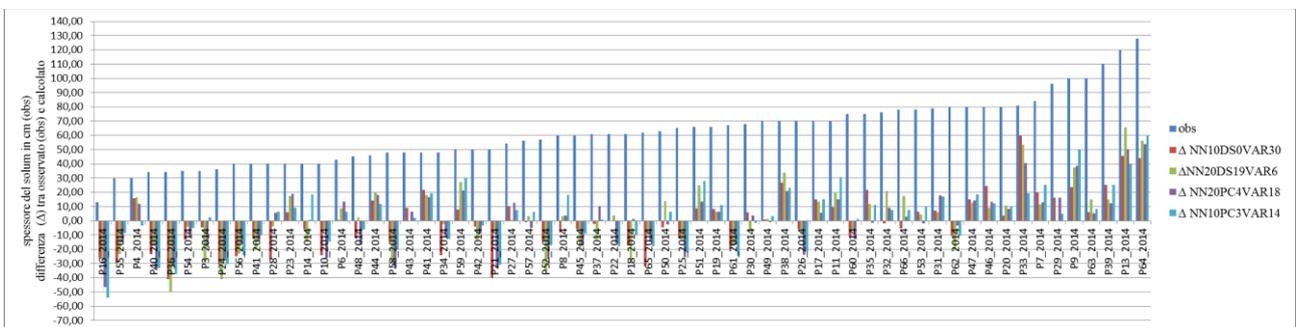


**Fig. 5.17** Istogramma relativo alla differenza tra valore osservato e valore predetto per i punti con suolo zero. I punti indicati con EI appartengono alle zone degli impluvi soggetti a erosione da flusso idrico concentrato. I punti indicati con la lettera C appartengono a calanchi. I punti indicati con la lettera A appartengono a roccia affiorante.



**Fig. 5.18** Ubicazione dei punti con suolo zero con le peggiori predizioni (A, C, E) e indicazione della relativa situazione morfologica (DEM con isoipse a 5 m: B, D, F).

Confrontando i valori dei residui, relativi ai migliori modelli (Fig. 5.19), emerge una concordanza circa l'errore della stima; ciascun modello è contraddistinto da una propria capacità predittiva, ma la dimensione dell'errore (valore e segno) per ciascun punto utilizzato per il *validation set* è simile per tutti i modelli.



**Fig. 5.19** Confronto, per ciascuna misura di profilo osservato (obs), dei residui (osservato - calcolato) relativi ai migliori risultati ottenuti, dal quale emerge una sostanziale concordanza circa l'errore della stima per i modelli calcolati.

Date queste considerazioni, occorre tener presente che, nell' eseguire questo tipo di elaborazione, alcune imprecisioni sono insite nel metodo stesso, a cominciare dall' esecuzione dell' attività di rilevamento. Infatti, l' imprecisione del GPS utilizzato per rilevare le coordinate può provocare un errore nella georeferenziazione (soprattutto nel caso di GPS non differenziale); i profili possono essere rilevati in particolari posizioni topografiche, per le quali il modello digitale del terreno è inadeguato per rappresentare la reale variabilità del territorio e quindi la relazione tra gli attributi geomorfometrici ed il tipo di suolo. D' altro canto, al di là di queste valutazioni, occorre considerare che nel rilevamento in campo spesso non è semplice la determinazione dello spessore del *solum*. A volte, l' individuazione esatta del punto dove l' alterazione del substrato litologico termina è particolarmente difficile, soprattutto nel caso delle formazioni sedimentarie presenti nell' area di studio. Tale individuazione è resa ancor più complicata per i profili profondi: infatti, nel caso di scavo effettuato a mano, sovente non si ha la possibilità di osservare tutte le parti del profilo da una prospettiva favorevole. Oltre a ciò, occorre considerare che anche se il DEM a 10 m è considerato un ottimo compromesso tra risoluzione spaziale e capacità di rappresentare la morfologia del territorio (Maynard e Johnson 2014), per aree particolarmente articolate può non essere sufficientemente accurato. Alla luce delle analisi effettuate si ritiene che il metodo presentato fornisca risultati accettabili; il giudizio considera che: è stato utilizzato un numero esiguo di punti di campionamento; la qualità del modello è stata valutata con un *set* di punti non utilizzato per la costruzione del modello; il risultato può essere migliorato con l' introduzione di nuovi punti di campionamento. Un' altra considerazione può essere avanzata riguardo l' importanza che le diverse variabili hanno sulla risposta, ovvero sul coefficiente di correlazione per il *dataset* di validazione. Confrontando i risultati dei due metodi, proposti per la selezione delle variabili, è emersa una diversità per il tipo di variabili incluse nei diversi sub-set, ma anche una vicinanza dei risultati a parità di numero di variabili. In generale, sono stati ottenuti migliori risultati utilizzando i sub-set selezionati con il metodo A; tuttavia, per costruzione detto metodo utilizza un massimo di 23 variabili. Riguardo l' importanza assegnata alle diverse variabili, solo tre (*general curvature*, *topographic position index* e *flow drop*) sono rilevanti sia per il metodo A (che seleziona le variabili secondo l' importanza che ognuna di esse assume in ciascuna delle componenti considerate) che per il metodo B (che considera l' importanza delle variabili complessivamente su tutte le sette componenti); tale concordanza non emerge nel caso delle variabili che potenzialmente hanno scarsa influenza. Infine, occorre sottolineare che, considerando il *trend* dei coefficienti di correlazione al variare del numero di variabili incluse nelle elaborazioni, è stato calcolato anche un modello con l' intero set di variabili il quale, di fatto, ha fornito complessivamente il migliore risultato.

## **Bibliografia**

Behrens, Thorsten; Förster, Helga; Scholten, Thomas; Steinrücken, Ulrich; Spies, Ernst-Dieter; Goldschmitt, Michael (2005): Digital soil mapping using artificial neural networks. In: *Z. Pflanzenernähr. Bodenk.* n. 1, 168, pp. 21–33. DOI: 10.1002/jpln.200421414.

Behrens, Thorsten; Zhu, A-Xing; Schmidt, Karsten; Scholten, Thomas (2010): Multi-scale digital terrain analysis and feature selection for digital soil mapping. In: *Geoderma* n. 3-4, 155, pp. 175–185. DOI: 10.1016/j.geoderma.2009.07.010.

Beucher, A.; Siemssen, R.; Fröjdö, S.; Österholm, P.; Martinkauppi, A.; Edén, P. (2015): Artificial neural network for mapping and characterization of acid sulfate soils. Application to Sirppujoki

River catchment, southwestern Finland. In: *Geoderma*, 247-248, pp. 38–50. DOI: 10.1016/j.geoderma.2014.11.031.

Bockheim, J. G.; Gennadiyev, A. N.; Hammer, R. D.; Tandarich, J. P. (2005): Historical development of key concepts in pedology. In: *Geoderma* n. 1, 124, pp. 23–36.

Boruvka, L.; Penizek, V. (2006): A test of an artificial neural network allocation procedure using the Czech soil survey of agricultural land data. In: *Developments in Soil Science*, 31, pp. 415–424.

Chagas, César da Silva; Vieira, Carlos Antônio Oliveira; Fernandes Filho, Elpídio Inácio (2013): Comparison between artificial neural networks and maximum likelihood classification in digital soil mapping. In: *Revista Brasileira de Ciência do Solo* n. 2, 37, pp. 339–351.

Chartin, C.; Bourennane, H.; Salvador-Blanes, S.; Hirschberger, F.; Macaire, J.-J. (2011): Classification and mapping of anthropogenic landforms on cultivated hillslopes using DEMs and soil thickness data — Example from the SW Parisian Basin, France. In: *Geomorphology* n. 1-2, 135, pp. 8–20. DOI: 10.1016/j.geomorph.2011.07.020.

Crawford, Aaron V.; Green, Samuel B.; Levy, Roy; Lo, Wen-Juo; Scott, Lietta; Svetina, Dubravka; Thompson, Marilyn S. (2010): Evaluation of parallel analysis methods for determining the number of factors. In: *Educational and Psychological Measurement*.

Demuth, Howard; Beale, Mark; Hagan, Martin (2008): Neural network toolbox™ 6. In: *User's guide*, pp. 37–55.

Demuth, Howard B.; Beale, Mark H.; Jess, Orlando de; Hagan, Martin T. (2014): Neural network design: Martin Hagan.

Dietrich; W. E.; Reiss; R.; Hsu; M. L. et al.: A process-based model for colluvial soil depth and shallow landsliding using digital elevation data.

Dietrich, William E.; Reiss, Robert; Hsu, Mei-Ling; Montgomery, David R. (1995): A process-based model for colluvial soil depth and shallow landsliding using digital elevation data. In: *Hydrological processes* n. 3-4, 9, pp. 383–400.

Douglas, C. L.; Fehrenbacher, J. B.; Ray, B. W. (1967): The lower boundary of selected mollisols. In: *Soil Science Society of America Journal* n. 6, 31, pp. 795–800.

Franklin, Scott B.; Gibson, David J.; Robertson, Philip A.; Pohlmann, John T.; Fralish, James S. (1995): Parallel analysis: a method for determining significant principal components. In: *Journal of Vegetation Science* n. 1, 6, pp. 99–106.

Harden, Jennifer W. (1982): A quantitative index of soil development from field descriptions: Examples from a chronosequence in central California. In: *Geoderma* n. 1, 28, pp. 1–28.

Hseu, Zeng-Yei; Chen, Zueng-Sang; Tsai, Chen-Chi; Tsui, Chun-Chih; Cheng, Shuang-Fu; Liu, Chyan-Lan; Lin, Haw-Tarn (2002): Digestion methods for total heavy metals in sediments and soils. In: *Water, Air, & Soil Pollution* n. 1, 141, pp. 189–205.

IUSS Working Group (2014): World reference base for soil resources 2014 international soil classification system for naming soils and creating legends for soil maps. In: *FAO, Rome*.

- Kuriakose, Sekhar L.; Devkota, Sanjaya; Rossiter, D. G.; Jetten, V. G. (2009): Prediction of soil depth using environmental variables in an anthropogenic landscape, a case study in the Western Ghats of Kerala, India. In: *CATENA* n. 1, 79, pp. 27–38.
- Lal, Rattan (1998): Soil erosion impact on agronomic productivity and environment quality. In: *Critical reviews in plant sciences* n. 4, 17, pp. 319–464.
- Ledesma, Rubén Daniel; Valero-Mora, Pedro; Macbeth, Guillermo (2015): The scree test and the number of factors: a dynamic graphics approach. In: *The Spanish journal of psychology*, 18, E11.
- Lee, Kwan Tun; Ho, Jui-Yi (2009): Prediction of landslide occurrence based on slope-instability analysis and hydrological model simulation. In: *Journal of Hydrology* n. 3-4, 375, pp. 489–497. DOI: 10.1016/j.jhydrol.2009.06.053.
- Lu, Yuanyuan; Zhang, Ganlin; Zhao, Yuguo; Li, Decheng; Yang, Jinling; Liu, Feng (2016): Mapping Soil Thickness by Integrating Fuzzy C-Means with Decision Tree Approaches in a Complex Landscape Environment. In: *Digital Soil Mapping Across Paradigms, Scales and Boundaries*: Springer, pp. 63–76.
- Maynard, J. J.; Johnson, M. G. (2014): Scale-dependency of LiDAR derived terrain attributes in quantitative soil-landscape modeling. Effects of grid resolution vs. neighborhood extent. In: *Geoderma*, 230-231, pp. 29–40. DOI: 10.1016/j.geoderma.2014.03.021.
- McKenzie, Neil J.; Ryan, Philip J. (1999): Spatial prediction of soil properties using environmental correlation. In: *Geoderma* n. 1-2, 89, pp. 67–94. DOI: 10.1016/S0016-7061(98)00137-2.
- Menezes, Michele Duarte de; Silva, Sérgio Henrique Godinho; Mello, Carlos Rogério de; Owens, Phillip Ray; Curi, Nilton (2014): Solum depth spatial prediction comparing conventional with knowledge-based digital soil mapping approaches. In: *Scientia Agricola* n. 4, 71, pp. 316–323.
- Montgomery, David R. (2007): Soil erosion and agricultural sustainability. In: *Proceedings of the National Academy of Sciences* n. 33, 104, pp. 13268–13272.
- Morgan, R. P.C.; Quinton, J. N.; Smith, R. E.; Govers, Gerard; Poesen, J. W.A.; Auerswald, K. et al. (1998): The European Soil Erosion Model (EUROSEM): a dynamic approach for predicting sediment transport from fields and small catchments. In: *Earth surface processes and landforms* n. 6, 23, pp. 527–544.
- Pelletier, Jon D.; Rasmussen, Craig (2009): Geomorphically based predictive mapping of soil thickness in upland watersheds. In: *Water Resour. Res.* n. 9, 45. DOI: 10.1029/2008WR007319.
- Penížek, V.; Borůvka, L. (2006): Soil depth prediction supported by primary terrain attributes: a comparison of methods. In: *Plant, Soil and Environment* n. 9, 52, pp. 424–430.
- Pimentel, David; Harvey, Celia; Resosudarmo, Pradnja; Sinclair, K. (1995): Environmental and economic costs of soil erosion and conservation benefits. In: *Science* n. 5201, 267, p. 1117.
- Rahimy, Pooyan (2012): Effects of soil depth spatial variation on runoff simulation, using the Limburg Soil Erosion Model (LISEM), a case Study in Faucon Catchment, France. In: *Soil & Water Research* n. 2, 7.

- Renschler, Chris S.; Harbor, Jon (2002): Soil erosion assessment tools from point to regional scales—the role of geomorphologists in land management research and implementation. In: *Geomorphology* n. 2-4, 47, pp. 189–209. DOI: 10.1016/S0169-555X(02)00082-X.
- Roering, Joshua J. (2008): How well can hillslope evolution models “explain” topography? Simulating soil transport and production with high-resolution topographic data. In: *Geological Society of America Bulletin* n. 9-10, 120, pp. 1248–1262.
- Schaetzl, R. J.; Anderson, S. (2005): *Soil Genesis and Morphology*: Cambridge University Press, New York, NY.
- Tesfa, Teklu K.; Tarboton, David G.; Chandler, David G.; McNamara, James P. (2009): Modeling soil depth from topographic and land cover attributes. In: *Water Resour. Res.* n. 10, 45, n/a-n/a. DOI: 10.1029/2008WR007474.
- Todeschini, Roberto (1998): Introduzione alla chemiometria. In: *EdiSES, Napoli*, p. 321.
- Tromp-van Meerveld, H. J.; McDonnell, J. J. (2006): On the interrelations between topography, soil depth, soil moisture, transpiration rates and species distribution at the hillslope scale. In: *Advances in Water Resources* n. 2, 29, pp. 293–310. DOI: 10.1016/j.advwatres.2005.02.016.
- Wise, Barry M.; Gallagher, Neal B.; Bro, Rasmus; Shaver, Jeremy M.; Windig, Willem; Koch, R. Scott (2006): Chemometrics tutorial for PLS\_Toolbox and Solo. In: *Eigenvector Research, Inc*, 3905.
- Zhu, A-Xing (2000): Mapping soil landscape as spatial continua: the neural network approach. In: *Water Resources Research* n. 3, 36, pp. 663–677.
- Zizioli, D.; Meisina, C.; Valentino, R.; Montrasio, L. (2013): Comparison between different approaches to modeling shallow landslide susceptibility. A case history in Oltrepo Pavese, Northern Italy. In: *Nat. Hazards Earth Syst. Sci.* n. 3, 13, pp. 559–573. DOI: 10.5194/nhess-13-559-2013.
- Zwick, William R.; Velicer, Wayne F. (1986): Comparison of five rules for determining the number of components to retain. In: *Psychological bulletin* n. 3, 99, p. 432.

## CAPITOLO 6

### Erosione del suolo

#### 6.1 Introduzione

Date le differenze mineralogiche e, di conseguenza, di composizione chimica e delle caratteristiche meccanica dei diversi substrati geologici presenti nell'area (vgs. Cap. 2), nonché la conseguente differenza morfologica del territorio, in quest'area è presente una notevole geodiversità e differenti tipologie di fenomeni di erosione del suolo. Queste ultimi sono quasi del tutto specifiche a seconda delle formazioni geologiche e macroscopicamente sono accorpabili in due gruppi. Il primo racchiude quei fenomeni erosivi che si verificano sui substrati costituiti dalle formazioni sedimentarie appartenenti alle Unità Epiliguri, dell'Eocene medio e Miocene superiore. Per questi le maggiori evidenze di erosione sono quelle relative alle aree soggette a dinamica erosiva di tipo calanchivo. Infatti, nelle due valli di piccola dimensione (92 Km<sup>2</sup> circa) dell'area di studio, sono presenti, nelle loro differenti forme, le strutture geomorfologiche che in letteratura sono chiamate calanchi e biancane. Inoltre, in queste aree sono presenti processi quali la suffosione probabilmente associata a fenomeni di *piping*, in sostituzione a frane superficiali e a volte a fenomeni di *rill*. In alcune aree in corrispondenza delle Breccie di Baiso è stata rilevata la presenza di soliflusso. Questi fenomeni erosivi si esplicano con differente importanza a seconda del substrato litologico e, a volte, anche in combinazione tra di loro. Il secondo gruppo di fenomeni erosivi riguarda quasi esclusivamente la "placca di Pizzocorno-Pietragravina", costituita dalle Arenarie di Monte Vallassa del Miocene superiore. Data la totale differenza delle caratteristiche geotecniche di questa formazione e la peculiare morfologia locale, in quest'area si verificano fenomeni di erosione con dinamiche differenti rispetto a quelle precedentemente descritte. In questa area sono l'erosione differenziale degli strati più o meno competenti del substrato roccioso e la pendenza dei versanti i due fattori, più importanti, che guidano i processi erosivi e la pedogenesi.

#### 6.2 Descrizione dei fenomeni di erosione del suolo

Come indicato nel capitolo 2, secondo una lettura semplificata della legenda della carta d'uso del suolo, l'area di studio comprende sostanzialmente due sole tipologie: a) aree adibite ad uso agricolo; b) aree ove vi è una crescita di vegetazione arborea o arbustiva principalmente spontanea (le aree antropizzate sono solo una parte residuale, meno del 5% dell'area). Uno dei fattori principali che determina questa dicotomia è la differenza delle caratteristiche chimico-fisiche e strutturali, delle formazioni geologiche presenti. Questa differenza si manifesta determinando una morfologia del territorio formazione-specifica; inoltre, la singolarità delle diverse formazioni geologiche influenza il processo di formazione e di erosione del suolo, e di conseguenza la possibilità di sfruttamento da parte dell'uomo. Infatti, per quanto concerne i fenomeni di erosione del suolo ed i loro effetti è consolidata l'idea che sia influente il fattore antropico. Tipicamente alcune pratiche agricole, che tendono a ridurre la copertura vegetale, sono state ritenute responsabili per l'innescio di processi erosivi. (Ruiz-Flan et al. 1992; Clarke e Rendell 2000; Romero Díaz et al. 2007; Torri et al. 2013). Confrontando la ripartizione del territorio relativa all'anno 2012 con quella desumibile dalla carta d'uso del suolo del 1954, emerge che nel corso del tempo si è verificata una diminuzione delle pratiche agricole che ha determinato un incremento delle aree con vegetazione

arbustiva o arborea. Tale variazione è pari a circa il 15% dell'intera superficie considerata (v. Cap. 2).

Tenendo conto delle relazioni tra substrati litologici, fenomeni di erosione e tipi di suolo, emerge una suddivisione tra le aree occupate dalle formazioni sedimentarie marine, appartenenti alle Unità Epiliguri del pre-Miocene medio e l'areale delle Arenarie del Monte Vallassa, del Miocene superiore. Le prime sono formazioni marnose o con tessiture miste (peliti-siltiti, conglomerati), ove è sempre importante la componente dei minerali argillosi: questo determina la netta differenziazione rispetto alla formazione delle Arenarie del Monte Vallassa.

L'areale costituito dalla formazione delle Arenarie del Monte Vallassa (Fig. 6.1) corrisponde quasi totalmente alla presenza della placca di Pizzocorno-Pietragravina, sebbene sul Foglio Voghera siano stati anche cartografati due piccoli areali, presenti anche nella Val Ardivestra (Vercesi et al. 2015).



**Fig. 6.1** Formazione delle Arenarie del Monte Vallassa.

Data la specificità delle caratteristiche mineralogiche e tessiture di questa formazione, in quest'area si verificano fenomeni di erosione con dinamiche completamente differenti rispetto a quelle, più articolate, delle formazioni per le quali la componente pelitica e siltitica è preponderante. In quest'area sono l'erosione differenziale del substrato roccioso e la pendenza dei versanti i fattori più importanti che guidano i processi erosivi e la pedogenesi. L'evoluzione, che ha portato all'orogenesi, ha determinato la formazione di due sinclinali e di incisioni profonde, determinando una differenziazione nella giacitura degli strati rocciosi. La formazione, completamente fratturata, è principalmente organizzata in strati più competenti, costituiti in prevalenza da biocalcareni e arenarie bioclastiche, alternati a strati meno competenti caratterizzati da banchi sabbiosi. Questi fattori determinano un modello erosionale caratteristico, ove nelle parti sommitali dei versanti sono presenti le

aree più stabili, costituiti da plateau. Un esempio di suolo rilevabile sui plateau è quello relativo al profilo P10\_2015 (Fig. 6.2 A), rilevato al margine di un'area agricola, in località Panzini. Questo profilo è caratterizzato da uno spessore del solum pari 155 cm e da una tessitura franco sabbiosa argillosa, sostanzialmente costante lungo tutto lo spessore (la percentuale di sabbia è intorno al 30%, quella di limo di poco superiore al 40% e la percentuale di argilla varia dal 20% al 23%). Presenta un basso contenuto di carbonato di calcio (inferiore al 7%). La struttura degli aggregati, negli orizzonti sottosuperficiali, è moderatamente sviluppata, con forma poliedrica sub-angolare e dimensione grande, mentre l'orizzonte Ap presenta una struttura fortemente sviluppata. Lo scheletro è sostanzialmente assente lungo tutto il profilo sino a raggiungere l'orizzonte C, costituito da Arenaria fratturata. Il suolo è stato classificato secondo la tassonomia WRB come Calcaric Cambisol (Loamic). Viceversa, nelle porzioni costituite dalle creste sono presenti suoli con un basso grado di sviluppo. Un esempio di questo tipo di suolo è dato dal profilo P3\_2014 (Fig. 6.2 B),

caratterizzato da uno spessore del solum pari 35 cm, da una tessitura sabbioso franca, con un contenuto di carbonato pari al 20%. La struttura degli aggregati è moderatamente sviluppata con forma granulare nell'orizzonte A e poliedrica angolare di dimensione media nell'orizzonte AC. La presenza di scheletro è scarsa e l'orizzonte C è costituito dallo strato sabbioso della formazione. Questo suolo è stato classificato secondo la tassonomia WRB come Leptic Calcaric Regosol.



**Fig. 6.2** Profilo rilevato in località Panzini con le caratteristiche tipiche dei suoli presenti sui plateau della Placca di Pizzocorno-Pietragravina. (A). Profilo rilevato su una costa della Placca di Pizzocorno-Pietragravina. (B)

Ai lati di queste aree è visibile ciò che rimane dalla demolizione dei plateau: infatti, in queste parti del versante, a volte costituite da strutture a terrazzo, sono presenti i blocchi derivanti dalle frane di crollo dei soprastanti strati rocciosi fratturati. Queste sono le aree di transizione a quelle porzioni del versante che completano il pendio e che portano agli impluvi. Nelle zone di accumulo, i blocchi sono diventati il nuovo substrato litologico per la pedogenesi. Un esempio di suolo rilevabile in queste aree è quello relativo al profilo P39\_2015 (Fig. 6.3 A), rilevato in un'area boschiva. È un suolo, con uno spessore del solum di 14 cm circa, con un orizzonte AC avente una tessitura franco sabbiosa (sabbia 65%, limo 20%, argilla 15%), sviluppatosi al di sotto di un orizzonte OL. È caratterizzato, inoltre, da un basso contenuto di carbonato di calcio (4%) e presenta una struttura granulare media, debolmente sviluppata. L'orizzonte C è costituito da blocchi di arenaria di dimensione grande o molto grande (25-50 cm). Il suolo è stato classificato secondo la tassonomia WRB come Leptic Skeletic Calcaric Regosol.

Sui versanti ai lati dei plateau, invece, i suoli si formano sul materiale derivante dagli strati meno competenti della formazione, a volte di origine colluviale. Questi suoli, a seconda delle condizioni topografiche, presentano differenti gradi di sviluppo pedogenetico. Su questi pendii, in dipendenza dal processo preponderante (rimozione o apporto di materiale), i suoli hanno caratteristiche e uno spessore fortemente influenzato dalla dinamica del versante e la matrice, più o meno pedogenizzata,

può inglobare frammenti di roccia. Un esempio di suoli rilevati in queste aree è dato dal profilo P36\_2015 (Fig. 6.3 B). Questo suolo, caratterizzato da uno spessore del solum di 53 cm, presenta una tessitura sabbioso franca ed un quantitativo di carbonato di calcio intorno al 15%. La struttura è debolmente sviluppata con forma poliedrica sub-angolare di dimensione grande, mentre quella dell'orizzonte superficiale è granulare. Nella matrice dell'orizzonte Bw sono inglobati clasti di arenarie e l'orizzonte C è costituito da blocchi di arenaria alterata, che poggiano sullo strato sabbioso della formazione. Il suolo è stato classificato secondo la tassonomia WRB come Leptic Calcaric Cambisol (Loamic).



**Fig. 6.3** Profilo rilevato in una area di accumulo dei blocchi derivanti dalla demolizione della formazione delle Arenarie del Monte Vallassa (A). Profilo con frammenti di roccia della formazione delle Arenarie del Monte Vallassa (B).

Nelle porzioni più ripide del versante, che portano agli impluvi, sono attivi processi di erosione che



**Fig. 6.4** Profilo rilevato in prossimità di un impluvio, con blocchi di grandi dimensioni.

determinano una asportazione laminare del materiale, a causa del dilavamento delle acque meteoriche; detta erosione determina l'affioramento dell'apparato radicale delle piante, o del substrato litologico (Fig. 6.5). In alternativa vi sono zone più stabili ove nella matrice del materiale, più o meno pedogenizzato, possono essere presenti parti di roccia poco degradate di grandi dimensioni. Un esempio di profilo rilevabile in queste aree è P47\_2015 (Fig. 6.4): è un suolo, con uno spessore del *solum* di 50 cm, rilevato in un'area boschiva, su una pendenza di 50°, sviluppatosi al di sotto di un orizzonte OL. Presenta una tessitura franco sabbiosa lungo tutto il profilo, sebbene l'orizzonte A si differenzi per avere un contenuto di argilla inferiore (7,5%) rispetto agli orizzonti sottostanti (16%). Presenta un contenuto di carbonato di calcio che varia da 6 al 13% ed una struttura moderatamente o debolmente sviluppata. A 50 cm circa di profondità si incontrano blocchi di dimensione molto grande (50 cm) insieme a materiale colluviale. Il suolo è stato classificato secondo la tassonomia WRB come

Leptic Skeletic Calcaric Regosol. Negli impluvi sono attivi i processi di incisione dovuta al deflusso idrico concentrato.



**Fig. 6.5** Erosione di tipo laminare sui versanti più ripidi in prossimità degli impluvi (A) e (B).

Per quanto riguarda l'area costituita dalle formazioni sedimentarie appartenenti alle Unità Epiliguri, del pre-Miocene medio, per comprendere i fenomeni di erosione del suolo che ivi si verificano occorre considerare, prima di tutto, gli elementi geomorfologici soggetti a dinamica di tipo calanchivo: i calanchi e le biancane.

I calanchi e le biancane sono stati ampiamente studiati in Italia, tanto che tali denominazioni sono utilizzati anche dai ricercatori non italiani, negli studi riguardanti le *badlands*. Per poter classificare e comprendere le forme calanchive dell'area di studio è opportuno riconsiderare gli studi già effettuati, con particolare riguardo a quelli svolti in Italia.

Nonostante l'abbondante letteratura riguardante detti elementi geomorfologici, va sottolineato che non sono stati svolti specifici studi né nell'area di studio né sull'appennino vogherese. Va segnalato che, sulle mappe create per indicare le aree soggette a dinamica calanchiva, seppur non di dettaglio, non risulta segnalata l'area corrispondente alla parte appenninica dell'Oltrepo Pavese. Per esempio come risulta nell'articolo pubblicato da Phillips (Phillips 1998), oppure come pubblicato da D. E. Alexander (Alexander 1980) (Fig. 6.5).

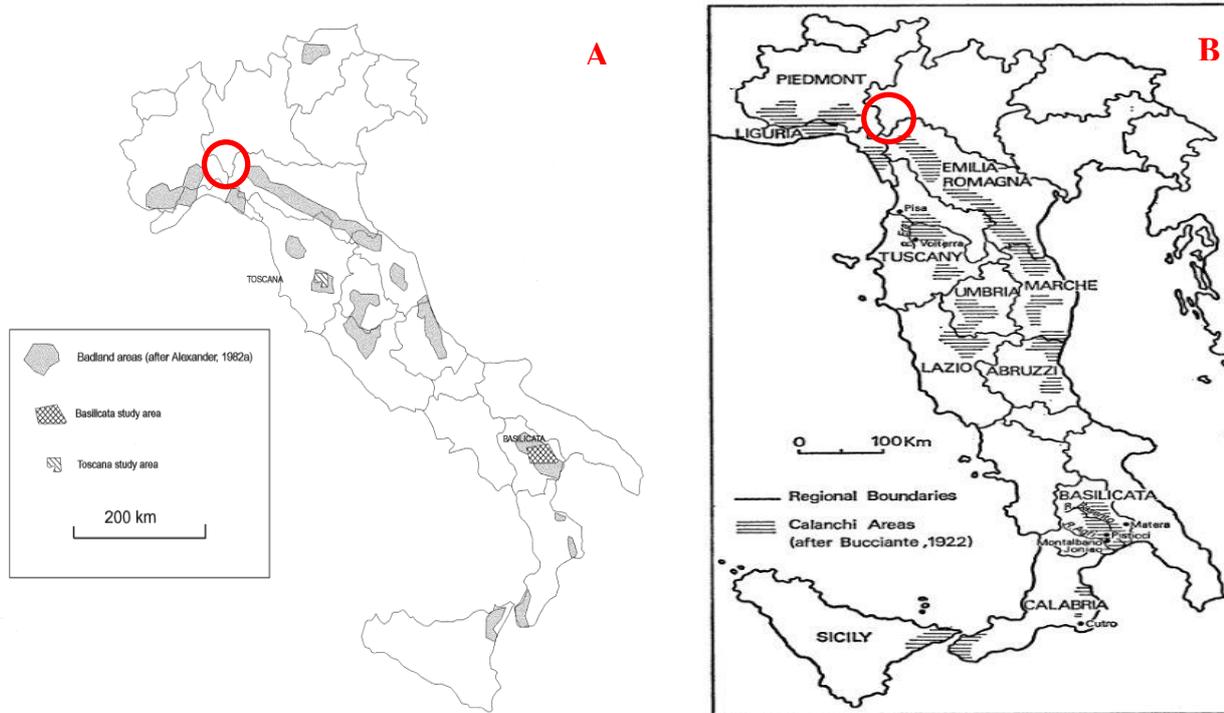


Figure 3 Location map of the study areas

Fig. 1—Italy: location map with distribution of calanchi areas.

**Fig. 6.6** mappa degli areali delle aree soggette a dinamica di tipo calanchivo pubblicata da Phillips (Phillips 1998) (A) e da D. E. Alexander (Alexander 1980) (B).

Tale rappresentazione incompleta si ritrova anche in pubblicazioni recenti, riportanti mappe generali sull'erosione accelerata del suolo (Cocco et al. 2015). Questo è probabilmente dovuto ai precedenti presenti in letteratura, che tipicamente sono stati sviluppati su casi del centro e sud Italia. In particolare, esiste una importante letteratura per casi studiati nelle seguenti regioni: Toscana (Torri et al. 1994; Guasparri 1993; Phillips 1998; Calzolari e Ungaro 1998; Battaglia et al. 2003; Della Seta et al. 2007; Ciccacci et al. 2008; Della Seta et al. 2009; Torri et al. 2013; Buccolini e Coco 2013;), Marche (Buccolini et al. 2007; Buccolini e Coco 2010, 2013; Cocco et al. 2015),

Lazio (Della Seta et al. 2007; Della Seta et al. 2009), Basilicata (Rendell 1982; Farifteh e Soeters 1999; Farifteh e Soeters 2006; Clarke e Rendell 2000; Summa e Luigia 2013; Piccarreta, et al. 2006), Abruzzo (Moretti e Rodolfi 2000), Calabria (Sdao et al. 1984), Sicilia (Cappadonia et al. 2011; Buccolini et al. 2012; Cappadonia et al. 2015).

Gli elementi geomorfologici presenti in Val di Nizza e Val Ardivestra non sempre mostrano caratteristiche tali da renderli facilmente classificabili, secondo le definizioni accettate nella letteratura di settore. Va evidenziato che in questa piccola area (92 km<sup>2</sup> circa) è presente una geodiversità tale da contenere, oltre alle forme calanchive e ai fenomeni ad essi associati, descritti in letteratura, elementi che, seppur riconducibili alle macrocategorie dei calanchi e delle biancane, posseggono una loro peculiarità, con caratteristiche intermedie.

Nel corso del tempo, nel mondo scientifico sono state avanzate diverse ipotesi circa la genesi dei calanchi e delle biancane. Sembrerebbe tuttavia che gli studi, effettuati con diversi approcci, e il ricorso a definizioni non abbiano esaurito il vivace dibattito riguardo la loro identificazione e, ancor più, la loro genesi.

A tal riguardo, un elenco della letteratura meno recente è riportato in Crocetti e Cavaglià nella pubblicazione intitolata: I CALANCHI Caratteristiche delle morfologie calanchive (Crocetti e Cavaglià s.d.). Le differenti teorie, riportate nella pubblicazione, si sono evolute nel tempo e sembrano essere state influenzate dalle condizioni stazionali delle aree di studio, a cominciare da Castiglioni (Castiglioni 1933) e Passerini: il primo per la formazione delle morfologie calanchive ha dato importanza alla giacitura degli strati, il secondo riteneva che fosse il tipo di esposizione il fattore determinante. Altri ricercatori ritenevano necessario il vincolo strutturale. Altri ancora hanno messo in relazione la composizione granulometria dei materiali alla genesi delle morfologie calanchive: tale ipotesi è stata sostenuta anche da altri autori, fino ad associare la genesi delle forme calanchive alla presenza di terreni argillosi dotati di scheletro silto-sabbioso. Successivamente è stata presa in considerazione l'ipotesi che i fattori fondamentali siano il macroclima e l'azione antropica.

Le definizioni di calanco e biancana e le descrizioni della loro genesi, ricorrenti nella letteratura recente, sono state formalizzate sin dal 1980. Come riportato in tutte le basilari pubblicazioni (Torri et al. 1994; Torri e Brya 1997; Calzolari e Ungaro 1998), anche nel fondamentale articolo di Moretti e Rodolfi (Moretti e Rodolfi 2000) vengono considerate come riferimento le pubblicazioni di Alexander (Alexander 1980, 1982), il quale fornisce una definizione basata sulla descrizione morfologica ed ipotizza l'esistenza dell'associazione tra calanchi e biancane, nonché la loro differenza.

Moretti e Rodolfi nella loro pubblicazione (Moretti e Rodolfi 2000) aggiungono la distinzione tra calanchi di tipo "A" e calanchi di tipo "B", di seguito riportata.

*"Type "A" develops because of the action of concentrated water runoff on clayey substrata having a very high silt and sand content; it gives rise to sharp and dissected landforms, referable to the classic 'knife edged' type \_sensu Scheidegger et al., 1968. with a very dense drainage pattern with channels having a deep V-shaped cross profile;"*

*"Type "B" is mainly due to recurrent superficial slides of soil or regolith on the unweathered substratum, usually containing a considerable amount of expendable clay; slopes become gentler, the divides get blunt and the drainage pattern less dense; slides may often evolve like mudflows."*

Inoltre, gli autori ammettono l'esistenza di tipologie intermedie tra i due tipi fondamentali, che si generano in situazioni ove i fattori modellanti (erosione idrica e mobilitazione del materiale) diventano più o meno importanti a seconda dei parametri geotecnici del substrato. Nel 2008 altri autori (Ciccacci et al. 2008) pubblicano l'esistenza di un terzo tipo di calanchi, indicati come tipo C: *“is characterised by a higher frequency of mass movements with numerous small scattered landslides, which almost completely destroy the calanchi ridge and fill up the bottom of the small valleys with flow materials. At the foot of calanchi badland slopes, superficial running waters deposit a large amount of colluvium that show characteristic mud-cracks up to 30 cm deep and which can be seen even just a few days after rainfall.”*

Nel 2006 ne viene individuata una ulteriore tipologia denominata “calanco mammellone” come una forma di transizione intermedia tra calanco e biancana dovuta alle caratteristiche chimico-fisiche e morfologiche intermedie (Piccarreta et al. 2006). Inoltre, in detto lavoro è stato osservato il coinvolgimento di *landslide* nella genesi degli stessi.

Per la definizione e descrizione delle biancane e della loro genesi, oltre agli autori già citati, occorre far riferimento alle pubblicazioni di Torri (Torri et al. 1994; Torri e Brya 1997) e alla pubblicazione di Costanza Calzolari e Fabrizio Ungaro (Calzolari e Ungaro 1998). Secondo le prime definizioni, le biancane sono alte meno di 20 metri, hanno una forma a cupola, sono a volte isolate e a volte in gruppo; hanno generalmente il versante esposto a sud più pendente e privo di vegetazione rispetto a quello esposto a nord, ricoperto da specie erbacee e il contatto con il basamento è determinato da un cambio abrupto della pendenza. Viene, inoltre, indicata la differenza tra la litologia sulla quale si formano i calanchi rispetto e quella ove si formano le biancane; in particolare, queste ultime, in Italia, si formano solo sulle argille marine del Pliocene (Calzolari e Ungaro 1998). Secondo un approccio morfologico per lo studio delle biancane vengono descritte cinque tipologie (*typical domes; bean forms; cone forms; souffle-like forms; scarp forms and sliding slopes*) ma nella pubblicazione viene anche precisato che esistono altre forme e che comunque tutte le categorie raramente possono essere rappresentate da un centrotipo, in quanto, di fatto, ne esistono forme intermedie (Calzolari e Ungaro 1998). Tuttavia, tutti i tipi di biancana sono accomunati dallo stesso modello concettuale di evoluzione e l'esistenza della varietà morfologica è dipendente dal peso dei concomitanti fattori modellanti (Calzolari e Ungaro 1998).

Riguardo la genesi, Torri e Brya concludevano il loro lavoro affermando che il movimento idrologico e l'erosione di tipo *small pipes* e *micropipes* sono i fattori che principalmente determinano l'evoluzione delle biancane, mentre l'erosione da *rills* gioca un ruolo secondario (Torri e Brya 1997). La Calzolari, basandosi sugli articoli di Torri, forniva una descrizione della genesi più articolata:

*“originally gently sloping surface is dissected by a system of retreating gullies, mainly affected by fluvial dynamics. In the first stage of landform evolution, small domes are isolated on which small mass movements can start. When the vegetation turf is broken, surface erosion processes, lead to the formation of the typical biancane. Through the combined effects of surface and subsurface erosion processes, the domes evolve towards small cone forms, which slowly wear down, or, if subsurface processes are predominant, form the soufflé-like biancane. The rate of such a process depends on the morphological position and processes involved, such as vegetation dynamics, anthropic disturbance, grazing load and the respective rates of erosion and soil forming processes”* (Calzolari e Ungaro 1998).

Nel 1999 Farifteh e Soeters (Farifteh e Soeters 1999) individuavano quattro fondamentali tipi di piping che possono verificarsi sulle biancane:

- *Pipes mainly associated with desiccation cracks (micro pipes)*
- *Pipes associated with desiccation cracks and tectonic joints (small pipes)*
- *Pipes associated with tectonic joints (moderate to large pipes)*
- *Pipes associated with lithological contacts (moderate to large pipes).*

Clarke e Rendell (Clarke e Rendell 2000) hanno evidenziato una differenza nei processi di formazione delle biancane in Basilicata, ove predominano fenomeni di *rill*, rispetto a quelle presenti in Toscana studiate da Torri. A riguardo Clarke ha ipotizzato che la causa finale di questa diversità fosse imputabile alla differenze climatiche.

Battaglia ha analizzato le proprietà del parent material dei calanchi e delle biancane concludendo che “*biancane tend to develop in very fine sediments, such as silty clays with a very high clay content (65–70%); calanchi tend to occur in relatively coarser sediments, such as clayey silts and sandy clayey silts having a notable sand fraction (6–18%)*” (Battaglia et al. 2003). Tuttavia l’autore ha affermato che con gli studi effettuati non è stata trovata una chiara correlazione tra granulometria e forme calanchive, in quanto esistono biancane e calanchi con granulometrie intermedie. Come pure, non esiste una tipicità nella mineralogia delle argille nelle due forme calanchive. Le analisi chimiche hanno portato a misurare sulle biancane alte concentrazioni di sodio, a differenza di quanto rilevato sui calanchi.

Altri autori hanno dimostrato che le proprietà chimico-fisiche, in particolare dei minerali argillosi sodici, influenzano i processi erosionali e hanno concluso che il *piping* gioca un ruolo essenziale per la formazione dei calanchi mammellonari che si possono evolvere in biancane (Piccarreta, et al. 2006).

Farifteh e Soeters (Farifteh e Soeters 2006), a seguito degli studi condotti in Basilicata, ha proposto un preciso modello di formazione dei calanchi e delle biancane. Secondo Farifteh e Soeters le biancane sono associate ad un sistema di fratture lungo il quale si forma un sistema di *rills*. Successivamente, durante una fase di forte erosione, una rete interconnessa di *gully* e *pipes* collassati genererebbe blocchi separati, con la parte sommitale vegetata ed i fianchi con suolo nudo. Durante le fasi di debole erosione, *piping* e *rill* di piccole proporzioni, e crolli rimodellerebbero il blocco ormai isolato, verso forme riconoscibili come biancane. Invece, i Calanchi si formano su ripidi pendii, dove il tasso di incisione è più alto rispetto a quello di denudazione.

Inoltre, ha dedotto che le biancane non sono il prodotto dell’evoluzione dei calanchi, mentre un calanco, occasionalmente, può generarsi dall’erosione di una biancana. Nelle conclusioni l’autore pone dubbi circa la relazione tra tipo di materiale e forme, attribuendo quindi maggior peso ai processi rispetto ai fattori.

Nel 2007 Marta Della Seta (Della Seta et al. 2007) ha osservato un modello distributivo dell’associazione tra calanchi e biancane. Le biancane tipicamente sono localizzate su versanti con deboli pendenze, sia ai piedi che nella parte sommitale di aree interessate da fenomeni di erosione, che portano alla formazione di calanchi. Nella sua area di studio (bacini del fiume Tevere e dell’Ombrone) ha registrato anche altri fenomeni di erosione superficiale quali *mudflows*, *soil creep* and *solifluction*. Nel 2009 Della Seta, in accurati studi sulla quantificazione spazio temporale dell’erosione su calanchi e biancane, in sintesi osservava che: “*at calanchi slopes we observed deeper piping and larger gravitational movements. Deep piping is favored by steep hydraulic*

*gradients and develops at the boundary between mobilized clays (mainly deriving from abandoned cropland and mud flowing) and bedrock clays. Positive feedback processes are likely to favor the re-mobilization of the clay mass above the deep pipes through following collapses and mudflows. On the other hand, at biancane sites shallow micro-piping contributes to denudation without triggering gravitational movements, which occur mainly along major rills and gullies” (Della Seta et al. 2009).*

Il breve e non esaustivo excursus è necessario per poter mettere in relazione la conoscenza pregressa con quanto rilevato nell’area di studio. Per poter inquadrare le geoforme osservate e le loro dinamiche erosive è importante considerare le differenze tra i substrati litologici, sapendo che per la loro evoluzione giocano un ruolo fondamentale i minerali delle argille. Tali differenze possono essere evidenziate considerando la tessitura e la composizione mineralogica dei suoli.

Come è risultato dalle analisi tessiturali sui suoli e dagli esami effettuati con la diffrazione ai raggi X, in corrispondenza delle Breccie di Baiso i suoli campionati, hanno una componente di argilla più importante rispetto ai suoli campionati sulle altre formazioni (da 30 a oltre 50%), ed il minerale delle argille più presente è la smectite. Gli altri suoli o i materiali sciolti campionati hanno tessiture con la componente siltosa generalmente più importante della componente argillosa, che può avere quote inferiori al 20% e al massimo arrivare al 40%. La composizione mineralogica delle argille ha una maggiore varietà, e comprende oltre alla smectite anche caolinite, clorite, illite, e le loro fasi intermedie, con percentuali variabili (Mazuerk & Previtali, 1981).

Per queste proprietà le forme di erosione che si verificano in corrispondenza delle Breccie di Baiso hanno caratteristiche che le differenziano da quelle che si evolvono sulle formazioni delle Marne di Antognola, e sulle formazioni di Contignaco e di Ranzano, nelle loro diverse *facies*.

Nel corso dell’attività di rilevamento dei suoli sono state individuati, oltre a numerosi calanchi attivi di semplice classificazione, anche numerosi elementi morfologici che possono essere considerati casi intermedi tra i due prototipi di calanchi e biancane, assimilabili ai calanchi mammellonari (Piccarreta, et al. 2006). Inoltre, sono stati rilevati casi singoli di tipologie di erosione tipicamente presenti o associate ai fenomeni erosivi di tipo calanchivo, quali *soil creep*, soliflusso e *gully* (Della Seta et al. 2007).

Per quanto riguarda la distribuzione e l’associazione tra le morfologie, va sottolineato che nonostante emergano in modo evidente nel territorio numerose forme soggette a dinamiche di tipo calanchivo, osservando il territorio anche tramite immagini satellitari nel visibile non è possibile percepire la reale presenza e distribuzione di dette tipologie di erosione. Questo è dovuto al fatto che le aree soggette a tali dinamiche sono a volte coperte totalmente o parzialmente dalla *canopy* della vegetazione arborea (Fig. 6.7 A,B,C,D).



**Fig. 6.7** Aree soggette a dinamica di tipo calanchivo parzialmente occupate da vegetazione pioniera (A) (B) (C) (D).

Infatti, in entrambe le valli, nella maggior parte delle aree soggette a dinamica di tipo calanchivo si può verificare una temporanea stabilizzazione del versante, con conseguente radicamento di vegetazione pioniera, cui generalmente seguono eventi che determinano la riattivazione dell'erosione. In altri casi la vegetazione arborea trova punti favorevoli nelle zone di accumulo del materiale colluviale (Fig. 6.9 B). L'evidenza più importante di questa dinamica è data dal portamento (Fig. 6.8 B e 6.9. A) e dalla distribuzione della vegetazione arborea: in particolare, le radici esposte lignificate e il portamento strisciante sono il segno che si è verificata una riattivazione dell'erosione.



**Fig. 6.8** Il denudamento dell'apparato radicale indica la riattivazione del processo erosivo, successivamente ad un periodo di stabilità (A) e (B).



**Fig. 6.9** Il particolare portamento della Roverella indica che la presenza di una dinamica erosiva lenta ma continua che costringe la pianta ad uno sviluppo orizzontale (A). Area all'interno di un calanco ove l'accumulo di materiale colluviale ha creato le condizioni per la crescita di vegetazione arborea (B).

Per evidenziare la complessità delle forme e dei processi attivi nell'area, a titolo di esempio, si riportano i seguenti casi. Sulla formazione delle Breccie di Baiso, ai piedi della località di Costa Croce, vi è un versante con il fianco rivolto a ovest ove è presente un tipico calanco classificabile di tipo A. Sul versante esposto a nord, a circa 200 m, sono presenti una piccola frana ed evidenze di *rill* (Fig. 6.11 A). Nella parte soprastante a quest'ultimo è presente un'area agricola incolta esposta a nord-ovest, soggetta a erosione da soliflusso e poco distante sullo stesso versante si è generato un *gully* (Fig. 6.11 B). Nell'area sono stati rilevati 3 profili pedologici: il profilo P15\_2015 (Fig. 6.10. A) sull'area soggetta a *soil creep*, il profilo P16\_2015 (Fig. 6.10. B) in un'area nelle immediate adiacenze (a 30 m circa), apparentemente più stabile, ed il profilo P14\_2015 (Fig. 6.12 A) sulla parete esposta dall'erosione da *gully* (Fig. 6.12 B).



**Fig. 6.10** Profilo rilevato su un'area agricola incolta in una zona soggetta a soliflusso (A) Profilo rilevato a 30 m circa dal P15\_2015 su un'area apparentemente più stabile (B).

I profili P15\_2015 e P16\_2015<sup>1</sup> sono particolari tipologie di suolo rinvenute solo sulle Breccie di Baiso nelle aree soggette a erosione da *soil creep* ubicate nei pressi di aree con dinamiche di tipo calanchivo. In entrambi i casi, il profilo è costituito da un orizzonte Ap mediamente di 20 cm di spessore sovrapposto ad un orizzonte avente un basso grado di alterazione (CB) di circa 30 cm, la cui tessitura è caratterizzata dalla preponderante presenza di argilla. Infatti il quantitativo di quest'ultima supera il 50% sino ad arrivare al 56%. I quantitativi di limo sono in un *range* intorno al 30% e la percentuale di sabbia ha una quota residuale (10-20%): la classe tessiturale è tipicamente argillosa. Il carbonato di calcio, presente con quote comprese tra il 20 e il 30%, è a volte visibile in accumuli. La struttura degli aggregati, di norma, è fortemente sviluppata con forma poliedrica angolare o sub-angolare di dimensione media o grande; l'aggregato sottoposto a pressione ha un evidente comportamento plastico. Lo scheletro, di piccole dimensioni (20-75 mm), è presente in quantità frequente e in forma non alterata. I suoli appartenenti a questa categoria sono stati classificati secondo la tassonomia WRB come Leptic Vertisol (Calcaric). Il profilo P14\_2015, con uno spessore del solum di 125 cm, è stato rinvenuto sulla formazione delle Marne di Antognola. La tessitura, costante lungo tutto il profilo, è franca o franco argillosa, con un leggero aumento di argilla nell'orizzonte più profondo: la percentuale di sabbia è intorno al 30%, quella di limo di poco superiore al 40% mentre l'argilla varia tra il 26 e il 30%. Il contenuto di carbonato di calcio, che pervade indifferentemente l'intero profilo, è di poco inferiore al 30%. La struttura degli aggregati è moderatamente sviluppata con forma poliedrica angolare o sub-angolare di dimensione media o grande. Lo scheletro è praticamente assente lungo tutto il profilo sino a raggiungere l'orizzonte C. I suoli appartenenti a questa categoria sono stati classificati secondo la tassonomia WRB come Calcaric Cambisol (Loamic).

<sup>1</sup> Per questi profili è stata misurata la concentrazione di sodio (FAAS): OP15/2016 1-Ap: 0,33 (Na+ cmol<sub>(+)</sub>/kg), 2-C1: 0,43 (Na+ cmol<sub>(+)</sub>/kg); OP16/2016 1-Ap: 0,42 (Na+ cmol<sub>(+)</sub>/kg), 2-C: 0,58 (Na+ cmol<sub>(+)</sub>/kg).



**Fig. 6.11** Fenomeni di erosione in corrispondenza della formazione delle Brecce di Baiso in località di Costa Croce. Sul versante con il fianco rivolto a ovest è presente un tipico calanco classificabile di tipo A, sul versante esposto a nord, a circa 200 m, è presente erosione di tipo laminare ed evidenze di *rill*, è inoltre presente una piccola frana (A). Nella parte soprastante, esposta a nord-ovest, in un'area agricola incolta, è presente erosione da soliflusso a poca distanza si è generato un *gully*.



**Fig. 6.12** Profilo rilevato all'interno del gully (P14\_2015) (A) Gully (B).

In corrispondenza della medesima formazione, a nord-ovest di Sant'Albano, è presente un versante esposto ad ovest con un'associazione di fenomeni erosivi: un'area, interessata da *soil creep*, a cui è



**Fig. 6.13** Profilo rilevato Su un versante soggetto a erosione di tipo *soil creep*, in corrispondenza della Formazione delle Breccie di Baiso.

saldata ad una biancana con copertura a “pop-corn”<sup>2</sup>, del tipo descritto da Torri e Howard (Torri et al. 1994; Torri e Brya 1997; Howard 1994) (v. Fig. 6.14 B). Adiacente a questo sito e sul medesimo versante sono presenti forme classificabili come calanchi di tipo B, associate a fenomeni di *rill*. Nel medesimo contesto ed a soli 400 metri di distanza sono presenti calanchi di tipo A (Moretti e Rodolfi 2000) (v. Fig. 6.15) e, in prossimità del contatto tra le Marne di Monte Piano e le Arenarie di Ranzano, è presente un altro calanco sul quale è visibile una frana, del tipo descritto da Ciccacci per quanto riguarda i calanchi di tipo C (Ciccacci et al. 2008). Sul versante soggetto a erosione di tipo *soil creep* è stato rilevato il profilo P62\_2015 (Fig. 6.13). Detto profilo presenta caratteristiche, dal punto di vista morfologico e chimico-fisico, molto simili ai profili P15\_2015 e P16\_2015.

<sup>2</sup> Su un campione di pop-corn è stata misurata la seguente concentrazioni di sodio (FAAS): 16,93 Na<sup>+</sup> cmol<sub>(+)</sub>/kg.



**Fig. 6.14** Area a nord-ovest di Sant’Albano (A) Biancana a “pop-corn” (B).



**Fig. 6.15** Area a nord-ovest di Sant’Albano, con l’ubicazione della Biancana a “pop-corn” della Fig. 6.14 (B) e del profilo P62\_2015.

Sul fondovalle della Val di Nizza, nei pressi della località Risaia, in corrispondenza del contatto tra le Marne di Antognola e la formazione di Ranzano (*litofacies* RAN 3b), è presente una struttura che ha una forma assimilabile a quella di una biancana, anche se dimensioni superiori a quelle descritte dalla Calzolari (Calzolari e Ungaro 1998) (v. Fig. 6.17). Detta struttura ha una forma tronco conica, è alta oltre 30 metri ed ha un diametro alla base di oltre 100 metri. Su detta struttura è stato rilevato il profilo P70\_2015 (Fig. 6.16). E’ un suolo con uno spessore del *solum* di 38 cm, la cui tessitura, costante lungo tutto il profilo, è franco limosa. La percentuale di sabbia è intorno al 16%, quella di limo intorno al 50% e di argilla intorno al 24%. Il contenuto di carbonato di calcio è pari al 15%. La struttura degli aggregati è granulare di dimensioni medie, debolmente sviluppata. Lo scheletro è presente in forme di scaglie di peliti di piccole dimensioni con abbondanza crescente verso il substrato litologico. L’orizzonte CB, compreso tra 20 e 38 cm, è composto quasi completamente da scaglie di pelite; per questo profilo sono state misurate le seguenti concentrazioni di sodio (FAAS): 1A1: 0,09 (Na+ cmol<sub>(+)</sub>/kg), 2A2: 0,12 (Na+ cmol<sub>(+)</sub>/kg), 3CB: 0,20 (Na+ cmol<sub>(+)</sub>/kg). Questo suolo è stato classificato secondo la tassonomia WRB come Leptic Calcaric Regosol. Profili con analoghe caratteristiche sono stati rilevati anche sulle strutture adiacenti a forma di duomo, poste nelle vicinanze e su alcune strutture con caratteristiche assimilabili ai calanchi mammellonari descritti da Piccarreta.



**Fig. 6.16** Profilo rilevato sul versante della struttura assimilabile a una biancana (A) Profilo rilevato sulla parte sommitale della struttura assimilabile a una biancana (B).



**Fig. 6.17** Struttura assimilabile a una biancana ubicata sul fondovalle della Val di Nizza, in località Risaia.

In particolare, è stato osservato che strutture in stato di erosione a forma di duomo, saldate al retrostante versante, sono presenti sia nella parte alta dei versanti associate a calanchi, sia sul fondovalle, ove ugualmente sono presenti fenomeni di erosione di tipo calanchivo. Occorre precisare che nelle aree ove sono diffusi questi fenomeni esistono anche versanti la cui forma piana

e la pendenza determinano un'erosione di tipo laminare, eventualmente associata a fenomeni di *rill* che non danno origine a forme paragonabili ai calanchi (v. Fig. 6.18 B). Pur considerando il modello distributivo rappresentato da Della Seta (Della Seta et al. 2007) (v. Fig. 6.19 A) risulta difficile classificare come biancane le suddette strutture. In particolare, nella complessa struttura collinare, costituita dalle Arenarie di Ranzano nelle sue diverse *facies*, vi sono elementi che presentano una parte del versante a forma di duomo con suolo nudo o parzialmente ricoperte da vegetazione e una parte del versante con un calanco attivo. Alcuni di questi elementi del paesaggio mostrano ancora evidenze di passate lavorazioni agricole, ma sono ora colonizzati da vegetazione pioniera. Riguardo queste strutture occorre considerare che, in passato, l'agricoltura era praticata su un'area più estesa rispetto all'attuale (v. Cap. 2) e che per il suo svolgimento sono stati effettuati interventi che hanno modificato il territorio, tanto da obliterare parzialmente o totalmente alcuni connotati tipici di queste forme, eliminando i risultati dell'erosione. Tuttavia, riguardo dette strutture pare invece maggiormente applicabile la classificazione di calanco mammellonare secondo la definizione di Piccarreta, il quale ha definito queste forme come una transizione tra calanchi e biancane e ne ha spiegato la genesi (Piccarreta et al. 2006) (v. Fig. 6.18 A e 6.19 B).



**Fig. 6.18** Calanco mammellonare (A). Erosione di tipo laminare associata a fenomeni di *rill* (B).

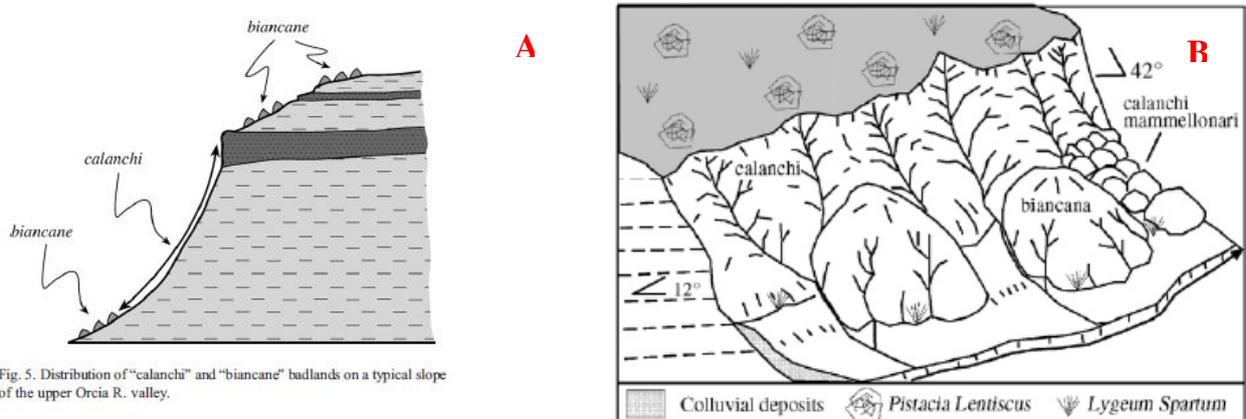


Fig. 5. Distribution of "calanchi" and "biancane" badlands on a typical slope of the upper Orcia R. valley.

**Fig. 6.19** Modello distributivo dei calanchi e delle biancane (Della Seta et al. 2007)(A). Modello esplicativo della formazione delle biancane e dei calanchi mammellonari (Piccarreta et al. 2006) (B).

Oltre alle strutture con le caratteristiche tipiche delle forme calanchive, un altro fenomeno di erosione del suolo è la suffosione. Tale tipo di erosione è l'evidenza più effimera presente nell'area. E' diffusa sui pendii a quote intermedie, nelle aree adibite a coltivazioni erbacee. La presenza di tale fenomeno erosivo non dipende dalla tipologia del substrato litologico: infatti, la suffosione è riscontrabile in diverse aree delle due valli (Sant'Albano, Lumello, ecc.) (v. Fig. 6.20 e 6.21). Essa è facilmente individuabile da immagini satellitari nel visibile, dal momento che l'area sottoposta a tale fenomeno presenta chiazze di forma circolare ed a volte ellittica. Queste, a seconda del periodo stagionale e del ciclo di lavorazione colturale a cui sono sottoposti i terreni, sono riconoscibili o come deboli sprofondamenti della superficie o come aree ove le colture manifestano forme di stress.



**Fig. 6.20** Suffosione in località Sant'Albano.



**Fig. 6.21** Suffosione in località Lumello (A) e (B).

Nell'area di studio l'altro fenomeno di erosione del suolo che si verifica, indipendentemente dal tipo di substrato litologico, è quello delle frane superficiali. Le tipologie di frane superficiali, in quest'area dell'Oltrepo Pavese, possono essere distinte secondo la classificazione di Campus e Cruden e Varnes della pubblicazione di Zizioli (Zizioli et al. 2013). Nell'area si sono verificati i seguenti tipi di frane superficiali, come documentato anche nell'Inventario dei Fenomeni Franosi in Italia (IFFI):

- 1) Frana traslazionale incipiente (*incipient translational slide*): sono visibili fratture nel terreno ma il movimento di massa è limitato e vi è poca deformazione interna;
- 2) Frana traslazionale del suolo (*translational soil slide*);
- 3) Frana roto-traslazionale (*rotational slide-translational slide*): tipicamente si sviluppano sui versanti ai bordi delle strade;
- 4) Frane complesse (*complex landslide*): la frana inizia con movimenti roto-traslazionali ed evolve, durante piogge continue, in flussi di terra che possono estendersi per notevoli distanze lungo il pendio (es: 250 m);
- 5) *Soil slip*: nel corso delle piogge, l'aumento del contenuto idrico determina un flusso di massa per il quale è difficile identificare la zona di accumulo.

La maggior parte delle frane che si verificano nelle due valli, salvo eventi eccezionali, sono le frane superficiali indicate al punto 2 (v. Fig 6.22 A e B). Queste sono state registrate principalmente in corrispondenza delle Formazioni di Antognola e delle Breccie di Baiso e interessano superfici mediamente di 1600 m<sup>2</sup>, ma possono raggiungere dimensioni notevoli (9000 m<sup>2</sup>). Questo tipo di frane presenta una superficie di scivolamento, lungo la discontinuità che si forma tra il *solum* ed il substrato costituente il materiale parentale. L'ipotesi riguardo il meccanismo di innesco considera che l'aumento della pressione dell'acqua interstiziale, durante le piogge più intense, può diminuire la resistenza al taglio del suolo e che la differenza di permeabilità del substrato litologico, rispetto al soprastante *solum*, può creare una superficie di scorrimento preferenziale del flusso idrologico, con lo sviluppo di "falda acquifera sospesa" (Bordoni et al. 2013). Nell'area di studio, i comportamenti idromeccanici dei suoli possono essere anche più complessi, a causa della concomitanza di differenti processi fisico-chimici, come nel caso dei suoli con un alto contenuto di argille espandibili, che causano cicli ripetuti di contrazione e rigonfiamento del *solum* e che possono favorire l'instabilità del versante (Meisina 2004). Lo spostamento della massa terrosa determina un distacco di pochi metri tra nicchia di frana e materiale traslato. Queste sono le frane che periodicamente si verificano sulle aree adibite a coltivazioni erbacee, a seguito di eventi piovosi estremi. A riguardo sembrerebbe che per le pratiche agricole, sia stato compreso questo problema, sostituendo le coltivazione arboree e quelle di pregio con coltivazioni erbacee (Erba medica). Il ripristino del terreno, successivamente all'evento, consiste nel ricollocamento della massa traslata nella nicchia di distacco, seguito da un livellamento.



Fig. 6.22 Frane superficiali in Val Ardivestra (A) (B).

### 6.3 Mappa dell'erosione del suolo e delle aree soggette a dinamica di tipo calanchivo

Per la rappresentazione dei processi attivi di erosione del suolo sono state create tre mappe tematiche. In queste elaborazioni non sono stati inclusi i processi per i quali sono stati rilevati casi singoli (soliflusso, *soil creep* e biancana a “popcorn”), su superfici aventi dimensioni troppo ridotte, tali da non poter essere utilizzate nell'elaborazione statistica. Sono state create una mappa generale per le tipologie di erosione del suolo individuate e due per singole tematiche: una relativa alle aree soggette a dinamica erosiva di tipo calanchivo ed una per individuare le aree con la propensione alle frane superficiali. Nella mappa generale dell'erosione del suolo sono state rappresentate le seguenti tipologie di erosione:

- *calanchi* (C);
- *calanchi mammellonari* (CM);
- *erosione da flusso idrico concentrato* (EIC);
- *roccia affiorante* (R);
- *erosione laminare* (EL);
- *erosione/deposizione* (ED);
- *suffosione* (S);
- *aree stabili* (AS).

La categoria *calanchi* (C) si riferisce agli areali ove i processi erosivi generano le geoforme classificabili come calanchi di tipo A, B o C; la categoria *calanchi mammellonari* (CM) comprende quelle strutture con una dinamica erosiva che genera una morfologia intermedia tra calanchi e biancane. La categoria *erosione da flusso idrico concentrato* (EIC) si riferisce all'erosione causata dalla canalizzazione del flusso idrologico superficiale, mentre la classe *roccia* (R) è per quelle aree ove emerge il substrato roccioso. Per *erosione laminare* (EL) si intende quel processo di erosione del suolo che si verifica in aree particolarmente acclivi (tipicamente nella parte terminale dei versanti che porta agli impluvi), ove la perdita di materiale pedogenizzato, che si accumula ai piedi del versante, determina l'affioramento dell'apparato radicale delle piante o del substrato litologico. La categoria *erosione/deposizione* (ED) contempla quelle dinamiche di disturbo della pedogenesi che si verificano generalmente su versanti con pendenze meno accentuate rispetto alla precedente categoria, ove il processo evolutivo dei suoli è stato o è condizionato da fasi di mobilitazione di materiale (che può comprendere blocchi di roccia) alternate a fasi deposizionali; la categoria *suffosione* (S) è specifica per il fenomeno. Infine, la classe *aree stabili* (AS) indica quelle porzioni di territorio che attualmente non sono interessate da nessuno dei fenomeni di erosione del suolo.

Come evidenziato, le diverse forme di tipo calanchivo, presenti nell'area di studio, spesso presentano morfologie intermedie che non facilitano la loro categorizzazione. Pertanto è stato scelto di elaborare una mappa specifica per verificare quanto queste forme possano essere distinte, dal punto di vista morfologico, utilizzando le variabili geomorfometriche, stante il fatto che i diversi processi danno origine a caratteristiche topografiche ben differenti. Nel caso dei calanchi che possono essere distinti in tipo A, B, C, i processi erosivi tendono a suddividere il versante in creste e piccoli valli; la loro genesi, la loro evoluzione e la forma attuale dipendono dalla composizione

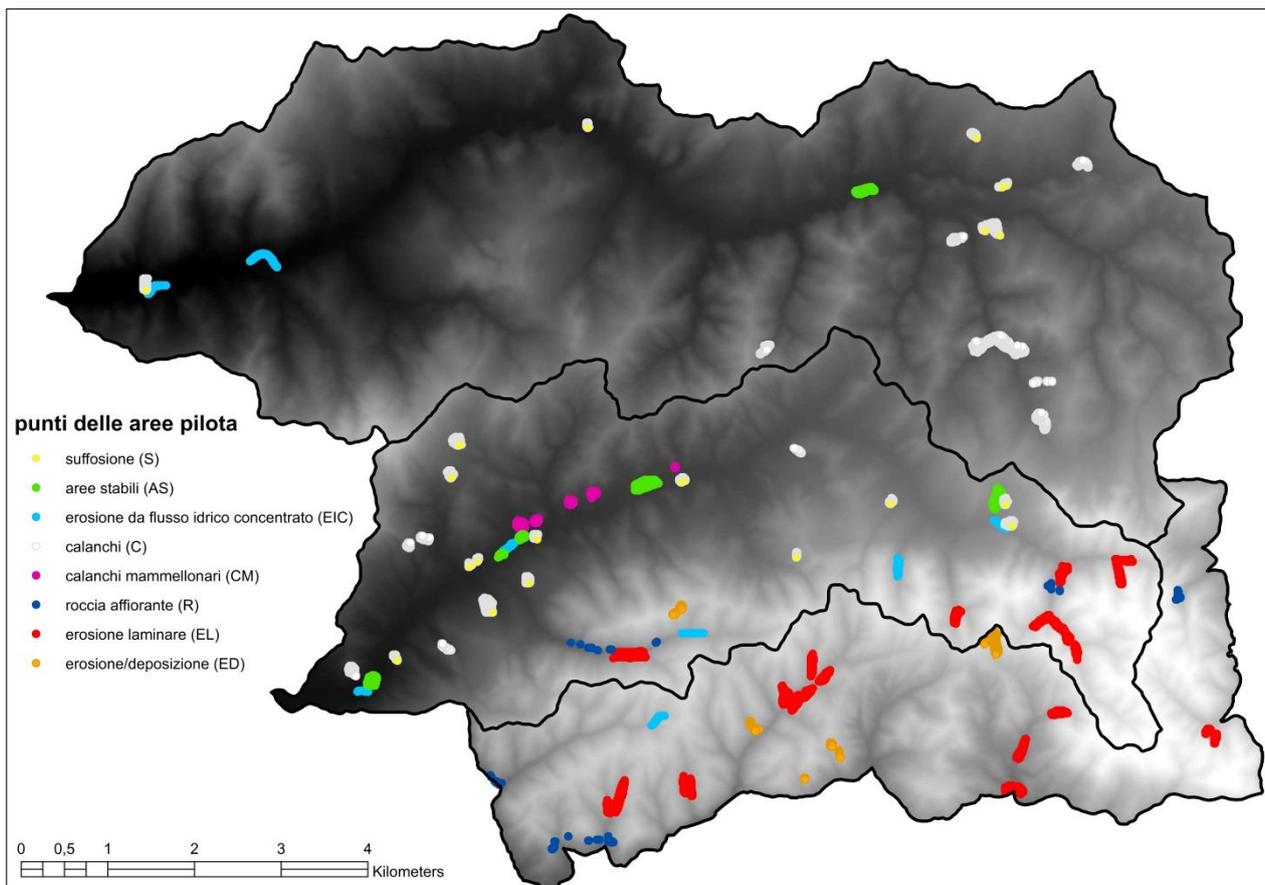
mineralogica del substrato e dal processo erosivo genetico dominante. Considerando il tipo di DEM a disposizione (pixel 10 m) e le caratteristiche di queste forme, si ritiene impraticabile la loro distinzione. Infatti, l'alternanza tra creste e valli può avvenire su brevi distanze (qualche metro) e la forma delle creste più o meno accentuata non è rilevabile con detta discretizzazione del territorio. Nel caso dei calanchi mammellonari, il processo di erosione tende a sviluppare una prominenzza rispetto al versante. Pertanto, sia nella mappa generale dell'erosione del suolo che in quella specifica sono stati rappresentati due tipologie: la prima comprende i calanchi A, B o C, la seconda i calanchi mammellonari. La differenza tra le due mappe sta nel fatto che in quella generale le due categorie sono in competizione con altre sei, mentre in quella specifica hanno una sola altra categoria antagonista (comprendente tutti gli altri casi): in questo modo il grado di confusione è misurato in modo più specifico.

La scelta di redigere una mappa per le frane superficiali è stata indotta dall'evidenza che questo tipo di erosione si verifica in aree ove è presente anche il fenomeno di suffosione, ed è probabile che vi siano aree per le quali i due processi non sono mutuamente esclusivi. Pertanto, includere questa categoria nella mappa generale dell'erosione del suolo avrebbe portato ad una rappresentazione incerta. Si evidenzia che nell'area si verificano differenti fenomeni franosi (Carelli et al. 2006) ma in questo lavoro e nella redazione della mappa specifica è stata presa in considerazione una sola tipologia di frane: quelle superficiali, la cui dinamica è stata precedentemente descritta (v. paragrafo 6.2).

### 6.3.1 Materiali e metodi

La mappa relativa all'erosione del suolo e quella delle tipologie calanchive sono stata realizzate utilizzando il metodo della *Discriminant Analysis* (DA), impiegando l'intero set di variabili geomorfometriche (v. Cap. 3) e partendo da un dataset di punti creato nel corso dell'attività di rilevamento in campo. Infatti, durante il campionamento, oltre al rilevamento pedologico, sono state registrate le coordinate di un punto valido per ciascun areale ove è stato osservato un processo di erosione attivo. Per poter applicare la DA è stato necessario incrementarne il numero, ottenendo così un set di punti sufficiente per l'elaborazione statistica. Tale artificio è stato realizzato digitalizzando 5392 punti mediante con l'ausilio di immagini satellitari (*Basemap* di ArcGIS), includendo più punti per ciascuna area pilota individuata (Fig. 6.23):

- 860 punti per la categoria AS;
- 213 punti per la categoria CM;
- 1060 punti, rispettivamente per la categoria C;
- 262 punti per la categoria EIC;
- 73 punti per la R
- 1754 punti per la EL;
- 254 punti per la ED;
- 916 punti per la AS.



**Fig. 6.23** Punti delle aree pilota, individuate nel corso dell'attività di campionamento.

Tale dataset è stato costituito in più fasi, in quanto le prime elaborazioni hanno fornito risultati con elevato grado di confusione tra le categorie considerate. A riguardo, si evidenzia che sono stati inclusi nella categoria *aree stabili* (S) (contenente i punti di campionamento pedologico non rilevati su nicchia di frana) alcune aree del fondovalle che nel corso delle prime elaborazioni sono state assegnate ad una delle categorie di erosione precedentemente indicate. Per la realizzazione di entrambe le mappe è stato utilizzato il medesimo *training set*, costituito dall'80% dei punti digitalizzati. Per la mappa relativa all'erosione calanchiva, il *training set* di punti non appartenenti alle classi C o CM è stato accorpato nella classe N. La scelta dei punti è stata effettuata casualmente, avendo però cura di includere un certo numero di punti per ciascuna delle aree pilota individuate. Per la valutazione dei risultati sono state prodotte le matrici di confusione per l'intero set di dati.

La DA è un metodo bayesiano, che per il calcolo della densità di probabilità utilizza un algoritmo basato su una funzione normale multivariata: infatti, l'ipotesi sottesa all'applicazione di questo metodo di classificazione è che le variabili abbiano una distribuzione normale. Per le variabili utilizzate nel presente lavoro, questo assunto è solo parzialmente verificato (v. Cap. 3). Tuttavia, si sottolinea che il metodo fornisce buoni risultati anche in presenza di scostamenti non piccoli dalle condizioni di normalità (Todeschini 1998). Si precisa che non si è proceduto ad alcun tipo di trasformazione delle variabili per ottenere una distribuzione più aderente alla gaussiana, per le argomentazioni indicate nel capitolo 3 e che, prima di introdurre nel calcolo, le variabili sono state normalizzate rispetto alle corrispondenti deviazione standard.

Nell'analisi discriminante un oggetto  $x_i$  viene classificato nella classe  $g$  se:

$$p(g|x_i) > p(k|x_i) \text{ per } \forall i \ k \neq g \text{ e } k = 1 \dots, G$$

Dove  $p(g|x_i)$  è la probabilità a posteriori che l'oggetto  $x_i$  appartenga alla classe  $g$  e  $G$  è il numero totale delle classi. Questa probabilità viene calcolata secondo la regola di Bayes:

$$p(g|x_i) = \frac{P_g f(x_i|g)}{\sum_k P_k f(x_i|k)}$$

Dove  $P_g$  è la probabilità a priori della classe  $g$  e  $f(x_i|g)$  è la funzione di densità di probabilità che deve essere stimata dal *training set*.

Un oggetto viene classificato nella classe  $g$  se è minima l'espressione:

$$D_g(x_i) = (x_i - \bar{x}_g)^T S_g^{-1} (x_i - \bar{x}_g) + \ln | S_g | - 2 \ln P_g$$

Dove  $D_g(x_i)$  è detto *discriminant score*,  $S_g^{-1}$  è l'inversa della matrice di covarianza della classe  $g$ ,  $\bar{x}_g$  è il corrispondente centroide di classe.

In pratica è stato utilizzato il modello di analisi discriminante già implementato in MATLAB, il quale opera in modo differente a seconda che l'algoritmo voglia ottenere una superficie di separazione degli elementi in modo lineare oppure mediante funzioni quadratiche. Nel primo caso il modello ha la stessa matrice di covarianza per ogni classe e varia soltanto la media; nel secondo caso, invece, variano ambedue i termini in base alla classe.

Una volta che il classificatore è stato addestrato, è possibile predire la classe di ogni oggetto secondo l'espressione:

$$\hat{y} = \arg \min \sum_{k=1}^K \hat{P}(k|x) C(y|k)$$

Dove

- $\hat{y}$  è la classe predetta;
- $k$  è il numero di classi;
- $\hat{P}(k|x)$  è la probabilità a posteriori dell'appartenenza alla classe  $k$  per l'osservazione  $x$ ;
- $C(y|k)$  è il costo per classificare una osservazione come  $\hat{y}$  quando la classe vera è  $k$ .

Con *Statistics Toolbox* di MATLAB (Jones 1997) è possibile settare i valori della *Prior Probability* e del *True Misclassification Cost* per un errore di classificazione, oltre a poter scegliere il tipo di metodo discriminante.

La *Prior Probability* di una classe può essere settata in modo:

'*uniform*' – ove la probabilità a priori della classe  $k$  è  $1/(\text{numero totale di classi})$ ;

'*empirical*' – ove la probabilità a priori di classe  $k$  è il numero di campioni appartenenti alla classe  $k$  diviso per il numero totale di campioni del training set.

Il *True Misclassification Cost* ( $i,j$ ) è il peso che a priori si ritiene opportuno assegnare se una osservazione è assegnata alla classe  $j$  quando in realtà la sua vera classe è  $i$ .

E' possibile, inoltre, scegliere il metodo di calcolo secondo i seguenti tipi di *Discriminant Analysis* (DA):

- *linear*: stima una matrice di covarianza per tutte le classi,
- *quadratic*: stima una matrice di covarianza per ciascuna classe;
- *diagLinear*: utilizza la matrice di covarianza diagonalizzata del metodo *linear*, e la sua pseudoinversa se necessario;
- *diagQuadratic*: utilizza le matrici di covarianza diagonalizzate del metodo *quadratic*, e le loro pseudoinverse se necessario;
- *pseudolinear*: utilizza la matrice pseudoinversa della matrice di covarianza del metodo *linear*;
- *pseudoQuadratic*: utilizza le matrici di covarianza pseudoinverse del metodo *quadratic*.

Per le mappe realizzate con il metodo della DA (erosione del suolo - aree calanchive), non conoscendo a priori la distribuzione sul territorio delle diverse classi, riguardo la probabilità a priori è stata adottata l'opzione *uniform*. A tal proposito si evidenzia che, imputando una probabilità a priori uguale per ogni classe, si può inserire un *bias*. In tal modo si assume che ogni classe abbia le stesse probabilità di essere presente nell'area di studio; d'altro canto, assegnando una probabilità a priori calcolata sulla base dell'abbondanza dei campioni (*empirical*) per ciascuna classe, si introduce un'altra distorsione, secondo la quale si considera che la numerosità dei campioni per ciascuna classe è proporzionale alla presenza di quest'ultima sull'intero territorio di interesse (occorrenza non verificabile a priori). Per quanto riguarda la matrice dei costi (*True Misclassification Cost*), è stata adottata l'impostazione: costo ( $i, j$ )=1 se  $i \neq j$ , e il costo ( $i, j$ )=0 se  $i=j$ , in parole non si sono volute considerare differenze d'importanza fra le diverse classi. Sono stati calcolati i risultati con i seguenti tipi di DA: *linear*, *diagLinear*, *diagQuadratic*, *pseudolineare*, *pseudoQuadratic*.

Gli output del metodo sono la *Posterior Probability* e la *Expected Misclassification Cost per Observation*. La *Posterior Probability* che un punto  $x$  appartenga alla classe  $k$  è il prodotto della probabilità a priori per il valore assunto dalla funzione densità normale multivariata. La funzione di densità normale multivariata con media  $\mu_k$  e covarianza  $\Sigma_k$ , in un punto  $x$  è così determinata:

$$P(x | k) = \frac{1}{(2\pi|\Sigma_k|)^{1/2}} \exp\left(-\frac{1}{2}(x - \mu_k)^T \Sigma_k^{-1} (x - \mu_k)\right) \quad \text{dove } |\Sigma_k| \text{ è il determinante di } \Sigma_k, \text{ e } \Sigma_k^{-1} \text{ è la matrice inversa.}$$

Se  $P(k)$  rappresenta la probabilità a priori di classe  $k$ , allora la probabilità a posteriori che una osservazione  $x$  appartenga alla classe  $k$  è:

$$\hat{P}(k | x) = \frac{P(x | k)P(k)}{P(x)}$$

dove  $P(x)$  è una costante di normalizzazione, cioè, la somma su  $k$  di  $P(x|k)P(k)$ .

Il *Expected Misclassification Cost per Observation* è il costo atteso (medio) dovuto all'applicazione del classificatore, per ciascuna osservazione relativamente a ciascuna classe. Il calcolo viene effettuato mediante la seguente espressione:

$$\sum_{k=1}^K \hat{P}(i | x)C(k | i)$$

- $k$  è il numero di classi.
- $\hat{P}(i | x)$  è la probabilità a posteriori di classe  $i$  per l'osservazione da predire  $x$ .
- $C(k | i)$  è il costo per classificare una osservazione come  $k$  quando la sua vera classe è  $i$ .

Per la valutazione dei risultati e la validazione di ciascun metodo di analisi discriminante, utilizzando le matrici di confusione sono stati calcolati: la percentuale di classificazioni corrette (*Non-Error-Rate*) e il rischio di errore di classificazione sia per l'intera matrice di confusione (*Misclassification risk*) sia relativamente a ciascuna classe. Sono state, inoltre, calcolate la sensibilità (*Sensibility*) e la specificità (*Specificity*), relativamente a ciascuna classe. La prima indica la capacità di una classe di rappresentare gli oggetti di tale classe, la seconda indica la capacità di una classe di isolarne gli oggetti (Todeschini 1998). La matrice di confusione riporta la classe vera sulle righe e quella assegnata sulle colonne. I casi assegnati correttamente si trovano sulla diagonale, mentre gli altri elementi della matrice sono: le assegnazioni degli oggetti di una classe che il metodo determina ( $n_g$ ), leggendo lungo le ( $g$ ) righe; gli oggetti che il metodo assegna ad una determinata classe ( $n_{g'}$ ), leggendo lungo le ( $g'$ ) colonne.

$$\text{Non-Error-Rate (NER\%)} = \frac{\sum_g c_{gg}}{n} 100 \quad (\text{calcolato sulla matrice di confusione})$$

$$\text{Error-Rate (ER\%)} = \frac{n - \sum_g c_{gg}}{n} 100 = 100 - \text{NER\%} \quad (\text{calcolato sulla matrice di confusione})$$

$$\text{Misclassification risk (MR\%)} = \sum_g \frac{\left( \sum_{g'} L_{g'g} C_{gg'} \right) P_g}{n_g} 100 \quad (\text{calcolato sulla matrice di confusione})$$

$$\text{Misclassification risk (MR}_g\%) = \frac{\left( \sum_{g'} L_{g'g} C_{gg'} \right) P_g}{n_g} 100 \quad (\text{calcolato sulla matrice di confusione})$$

$$\text{Sensitivity (Sn}_g) = \frac{c_{gg}}{n_g} 100 \quad (\text{calcolato per ogni singola classe della matrice di confusione})$$

$$\text{Specificity (Sp}_g) = \frac{c_{gg'}}{n_{g'}} 100 \quad (\text{calcolato per ogni singola classe della matrice di confusione})$$

(Dove  $L$  è la matrice dei costi;  $c$  è un elemento della matrice di confusione;  $g$  è l'indice relativo alle classi vere;  $g'$  è l'indice relativo alle classi calcolate dal modello;  $P_g$  è la probabilità a priori di una classe).

### 6.3.2 Analisi e risultati

Sia per la mappa *erosione del suolo* sia per quella di Aree soggette *erosione di tipo calanchivo* sono stati calcolati cinque modelli (con i metodi *linear*, *diagLinear*, *diagQuadratic*, *pseudolineare*, *pseudoQuadratic*), in quanto non è stato possibile effettuare il calcolo con il metodo *Quadratic* (il quale stima una matrice di covarianza per ogni classe). Questo accade perché il determinante della matrice di covarianza di una o più classi tende a zero in quanto due o più righe del *training set* sono o tendono a essere in combinazione lineare tra loro (collinearità).

Per quanto riguarda la mappa dell'erosione del suolo, esaminando i risultati delle matrici di confusione, la migliore prestazione è stata ottenuta con il metodo *pseudoQuadratic* (*Non-Error-Rate* pari a 81%) (v. Fig. 6.24). Anche gli altri parametri di valutazione indicano una elevata performance del metodo, infatti risulta basso il valore del *Misclassification risk*, sia complessivamente (MR pari a 12,7%) che relativamente a ciascuna classe (MR<sub>g</sub>). Ugualmente, i valori relativi a *Sensitivity* e a *Specificity* sono mediamente elevati. Per quanto riguarda i valori di *Sensitivity* la classe *calanchi* è la peggiore (50,5%). In particolare tale classe perde quasi la metà dei suoi elementi a favore delle classi R (*roccia affiorante*), EL (*erosione laminare*) e AS (*aree stabili*), tuttavia mantiene una alta *Specificity*, infatti vengono assegnati a questa classe pochi oggetti appartenenti ad altre classi. L'altra classe con una bassa *Sensitivity* è la AS (*aree stabili*) che perde elementi a favore della classe S (*suffosione*).

**Tab. 6.1 A** Matrice di confusione per le classi di erosione del suolo, relativa al metodo *pseudoQuadratic* (Non-Error-Rate (NER) Error-Rate (ER) e Misclassification risk (MR)).

		pseudoQuadratic									
		ED'	EL'	R'	CM'	C'	EIC'	AS'	S'	ng	
	ED	252	0	0	0	0	0	2	0	254	
NER (%)	81,03	2	1726	0	0	10	14	1	0	1753	
	EL	1	25	45	0	1	1	0	0	73	
ER (%)	18,97	0	0	0	209	0	0	2	2	213	
	R	0	0	0	61	535	103	108	9	1060	
MR (%)	12,70	0	12	0	0	0	218	27	5	262	
	CM	6	238	0	0	4	22	536	262	861	
	C	0	1	0	3	0	12	52	848	916	
	EIC	10	19	0	8	4	22	536	262	861	
	AS	0	1	0	3	0	12	52	848	916	
	S	0	1	0	3	0	12	52	848	916	
	ng'	271	2021	45	281	550	370	728	1126	5392	

**Tab. 6.1 B** Valori della Sensibility (Sn), Specificity (Sp) e del Misclassification risk (Mr), per ciascuna classe di erosione del suolo, relativi al metodo *pseudoQuadratic*.

SnED (%)	99,2	SpED (%)	93,0	MrED (%)	0,16
SnEL (%)	98,5	SpEL (%)	85,4	MrEL (%)	0,04
SnR (%)	61,6	SpR (%)	100,0	MrR (%)	4,50
SnCM (%)	98,1	SpCM (%)	74,4	MrCM (%)	0,38
SnC (%)	50,5	SpC (%)	97,3	MrC (%)	3,87
SnEIC (%)	83,2	SpEIC (%)	58,9	MrEIC (%)	1,41
SnAS (%)	62,3	SpAS (%)	73,6	MrAS (%)	0,42
SnS (%)	92,6	SpS (%)	75,3	MrS (%)	2,07

**Tab. 6.2 A** Matrice di confusione per le classi di erosione del suolo, relativa al metodo *diagQuadratic* (Non-Error-Rate (NER) Error-Rate (ER) e Misclassification risk (MR)).

		diag Quadratic									
		ED'	EL'	R'	CM'	C'	EIC'	AS'	S'	ng	
	ED	210	24	12	0	0	2	6	0	254	
NER (%)	77,36	14	1639	57	0	9	34	0	0	1753	
	EL	2	17	53	0	1	0	0	0	73	
ER (%)	22,64	0	0	0	186	1	9	2	15	213	
	R	0	0	0	104	618	114	38	26	1060	
MR (%)	17,71	0	19	0	0	0	171	71	1	262	
	CM	38	95	27	0	11	22	548	212	861	
	C	0	19	0	0	0	135	746	916		
	EIC	29	19	1	19	11	22	548	212	861	
	AS	16	0	0	13	0	6	135	746	916	
	S	0	0	0	3	0	10	103	800	916	
	ng'	309	1813	150	322	640	358	800	1000	5392	

**Tab. 6.2 B** Valori della Sensibility (Sn), Specificity (Sp) e del Misclassification risk (Mr), per ciascuna classe di erosione del suolo, relativi al metodo *diagQuadratic*.

SnED (%)	82,7	SpED (%)	68,0	MrED (%)	3,48
SnEL (%)	93,5	SpEL (%)	90,4	MrEL (%)	0,18
SnR (%)	72,6	SpR (%)	35,3	MrR (%)	3,21
SnCM (%)	87,3	SpCM (%)	57,8	MrCM (%)	2,55
SnC (%)	58,3	SpC (%)	96,6	MrC (%)	3,26
SnEIC (%)	65,3	SpEIC (%)	47,8	MrEIC (%)	2,91
SnAS (%)	63,6	SpAS (%)	68,5	MrAS (%)	0,41
SnS (%)	81,4	SpS (%)	74,6	MrS (%)	5,18

**Tab. 6.3 A** Matrice di confusione per le classi di erosione del suolo, relativa al metodo *Linear* (Non-Error-Rate (NER) Error-Rate (ER) e Misclassification risk (MR)).

		Linear									
		ED'	EL'	R'	CM'	C'	EIC'	AS'	S'	ng	
	ED	217	3	5	0	23	0	2	4	254	
NER (%)	78,99	10	1709	0	1	20	11	2	0	1753	
	EL	4	20	45	0	4	0	0	0	73	
ER (%)	21,01	0	0	0	168	10	1	0	34	213	
	R	0	0	0	157	504	50	40	107	1060	
MR (%)	22,00	1	25	0	1	0	190	33	12	262	
	CM	40	18	0	29	3	20	626	125	861	
	C	0	0	0	3	0	10	103	800	916	
	EIC	1	25	0	1	0	190	33	12	262	
	AS	40	18	0	29	3	20	626	125	861	
	S	0	0	0	3	0	10	103	800	916	
	ng'	273	1969	57	359	564	282	806	1082	5392	

**Tab. 6.3 B** Valori della Sensibility (Sn), Specificity (Sp) e del Misclassification risk (Mr), per ciascuna classe di erosione del suolo, relativi al metodo *Linear*.

SnED (%)	85,4	SpED (%)	79,5	MrED (%)	2,93
SnEL (%)	97,5	SpEL (%)	86,8	MrEL (%)	0,07
SnR (%)	61,6	SpR (%)	78,9	MrR (%)	4,50
SnCM (%)	78,9	SpCM (%)	46,8	MrCM (%)	4,25
SnC (%)	47,5	SpC (%)	89,4	MrC (%)	4,10
SnEIC (%)	72,5	SpEIC (%)	67,4	MrEIC (%)	2,30
SnAS (%)	72,7	SpAS (%)	77,7	MrAS (%)	0,30
SnS (%)	87,3	SpS (%)	73,9	MrS (%)	3,54

**Tab. 6.4 A** Matrice di confusione per le classi di erosione del suolo, relativa al metodo *diagLinear* (Non-Error-Rate (NER) Error-Rate (ER) e Misclassification risk (MR)).

		diag Linear									
		ED'	EL'	R'	CM'	C'	EIC'	AS'	S'	ng	
	ED	207	16	9	0	10	1	0	11	254	
NER (%)	75,43	31	1628	26	6	22	40	0	0	1753	
	EL	3	15	52	0	3	0	0	0	73	
ER (%)	24,57	0	0	0	157	14	9	0	33	213	
	R	0	0	0	255	511	90	6	64	1060	
MR (%)	21,39	1	18	0	0	0	182	55	6	262	
	CM	34	87	13	0	4	23	597	132	861	
	C	0	0	0	13	0	8	157	733	916	
	EIC	48	10	0	47	4	23	597	132	861	
	AS	5	0	0	13	0	8	157	733	916	
	S	0	0	0	3	0	10	103	800	916	
	ng'	329	1774	100	478	564	353	815	979	5392	

**Tab. 6.4 B** Valori della Sensibility (Sn), Specificity (Sp) e del Misclassification risk (Mr), per ciascuna classe di erosione del suolo, relativi al metodo *diagLinear*.

SnED (%)	81,5	SpED (%)	62,9	MrED (%)	3,72
SnEL (%)	92,9	SpEL (%)	91,8	MrEL (%)	0,20
SnR (%)	71,2	SpR (%)	52,0	MrR (%)	3,37
SnCM (%)	73,7	SpCM (%)	32,8	MrCM (%)	5,29
SnC (%)	48,2	SpC (%)	90,6	MrC (%)	4,05
SnEIC (%)	69,5	SpEIC (%)	51,6	MrEIC (%)	2,56
SnAS (%)	69,3	SpAS (%)	73,3	MrAS (%)	0,34
SnS (%)	80,0	SpS (%)	74,9	MrS (%)	5,58

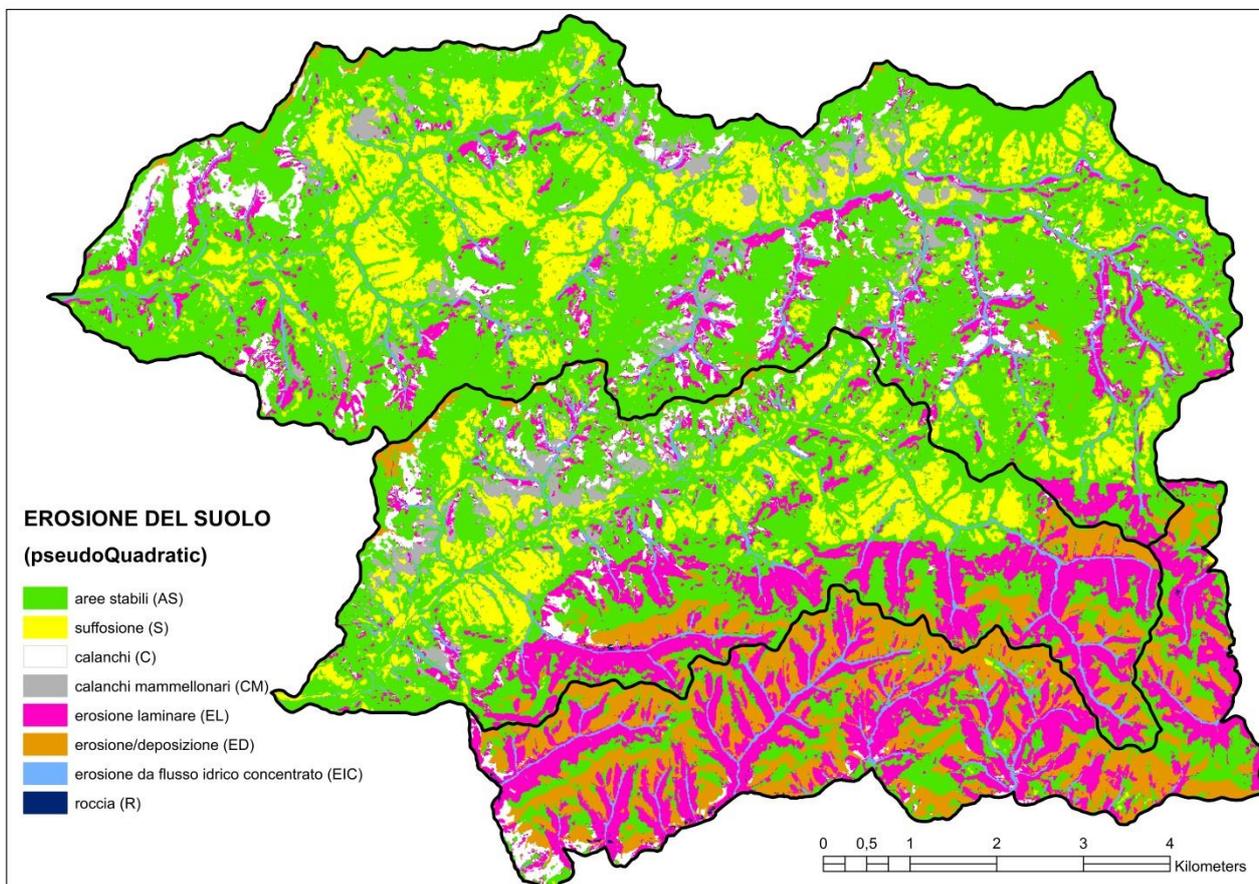
**Tab. 6.5 A** Matrice di confusione per le classi di erosione del suolo, relativa al metodo *pseudoLinear* (*Non-Error-Rate* (NER) *Error-Rate* (ER) e *Misclassification risk* (MR))..

		pseudo Linear								
		ED'	EL'	R'	CM'	C'	EIC'	AS'	S'	ng
	ED	217	3	5	0	23	0	2	4	254
NER (%)	78,99	10	1709	0	1	20	11	2	0	1753
ER (%)	21,01	4	20	45	0	4	0	0	0	73
MR (%)	19,07	0	0	0	168	10	1	0	34	213
	C	1	194	7	157	504	50	40	107	1060
	EIC	1	25	0	1	0	190	33	12	262
	AS	40	18	0	29	3	20	626	125	861
	S	0	0	0	3	0	10	103	800	916
	ng'	273	1969	57	359	564	282	806	1082	5392

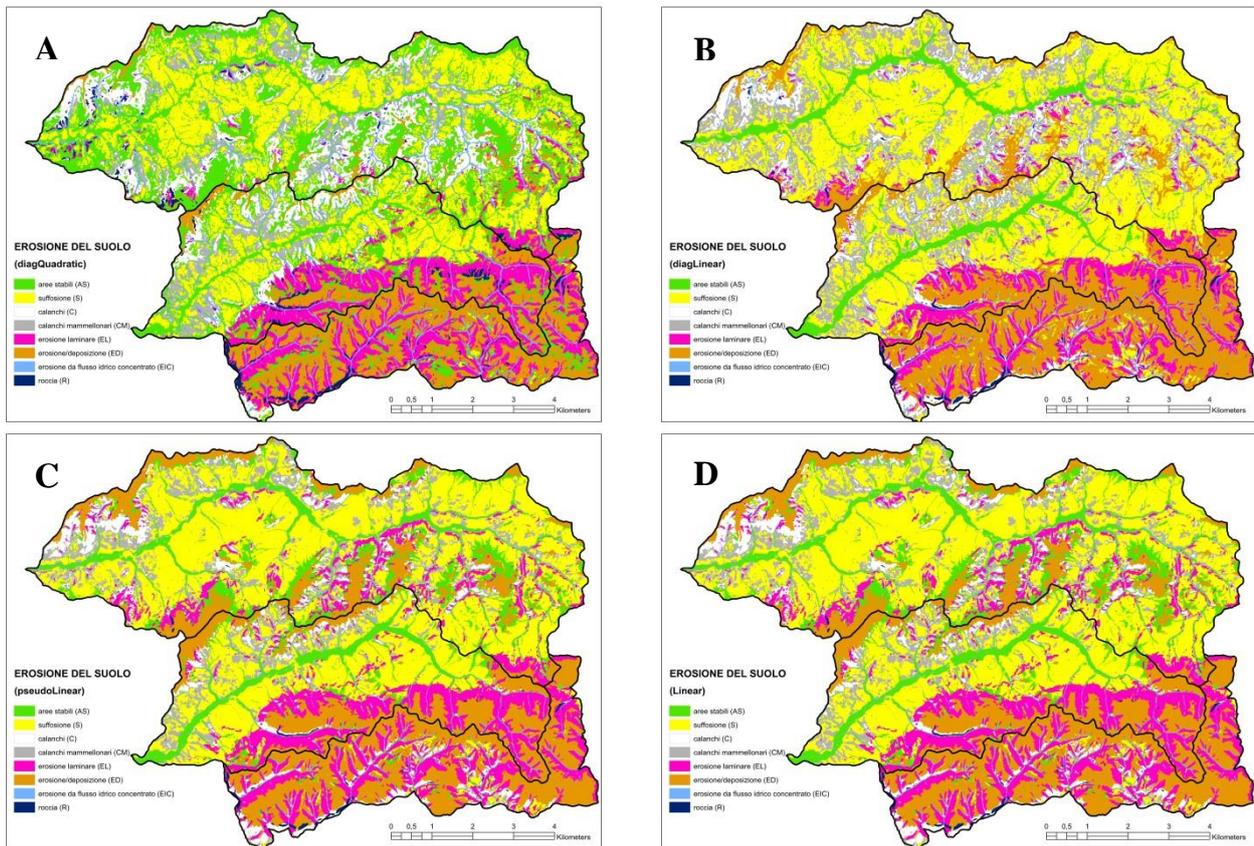
**Tab. 6.5 B** Valori della *Sensibility* (Sn), *Specificity* (Sp) e del *Misclassification risk* (Mr), per ciascuna classe di erosione, relativi al metodo *pseudoLinear*.

SnED (%)	85,4	SpED (%)	79,5	MrED (%)	2,93
SnEL (%)	97,5	SpEL (%)	86,8	MrEL (%)	0,07
SnR (%)	61,6	SpR (%)	78,9	MrR (%)	4,50
SnCM (%)	78,9	SpCM (%)	46,8	MrCM (%)	4,25
SnC (%)	47,5	SpC (%)	89,4	MrC (%)	4,10
SnEIC (%)	72,5	SpEIC (%)	67,4	MrEIC (%)	2,30
SnAS (%)	72,7	SpAS (%)	77,7	MrAS (%)	0,30
SnS (%)	87,3	SpS (%)	73,9	MrS (%)	3,54

Osservando i risultati si nota che tutte le mappe calcolate hanno alti valori degli indici di valutazione, tuttavia vi sono differenze nel pattern di distribuzione delle diverse classi. Infatti vi è una differenza sostanziale tra i metodi lineari (che utilizzano una matrice di covarianza per tutte le classi) e i metodi quadratici (che utilizzano una matrice di covarianza per ciascuna delle classi) (v. Fig. 6.25 A,B,C,D). In particolare le classi *suffosione* ed *erosione/deposizione* sono quelle che determinano le maggiori differenze. La classe *suffosione*, a seconda del metodo utilizzato, è più o meno estesa e tende ad occupare maggiori porzioni dei versanti, altrimenti classificati come *aree stabili* (AS). Analogamente la classe *erosione/deposizione* quando ha una maggiore assegnazione di punti occupa aree che in altre mappe sono attribuite a erosione di tipo calanchivo; anche la classe erosione laminare concorre con la classe calanchi con la quale è in associazione.



**Fig. 6.24** Mappa dell'erosione del suolo prodotta con il metodo *pseudoQuadratic*, con il perimetro delle due valli.



**Fig. 6.25** Confronto tra la distribuzione spaziale delle classi, per le mappe dell'erosione del suolo, ottenute con i metodi *diagQuadratic* (A), *diagLinear*(B), *pseudoLinear* (C), *Linear* (D), con il perimetro delle due valli.

Relativamente alla mappa delle aree soggette a erosione di tipo calanchivo, esaminando i parametri calcolati con la matrici di confusione, il migliore risultato, anche in questo caso, è stata ottenuto con il metodo *pseudoQuadratic* (*Non-Error-Rate* pari a 87,6 %) (v. Fig. 6.26). Un risultato ugualmente elevato è stato ottenuto con il metodo *diagQuadratic* (v. Fig. 6.27). Particolarmente bassi sono i valori del *Misclassification risk*. In questa elaborazione, diversamente dal caso con otto categorie, la classe CM (*calanchi mammellonari*) mantiene alti i valori per la *Sensibility* (perde pochi oggetti), ma ha un basso valore per la *Specificity*, in quanto gli vengono assegnati casi appartenenti alle altre classi, mentre la classe C (*calanchi*) si comporta come nel caso precedente (bassa *Sensibility* e alta *Specificity*), cedendo oggetti a entrambe le classi.

**Tab. 6.6 A** Matrice di confusione relativa alle classi utilizzate per la redazione della mappa dell'erosione calanchiva, relativamente al metodo *pseudoQuadratic* (*Non-Error-Rate* (NER) *Error-Rate* (ER) e *Misclassification risk* complessivo MR)).

		pseudo Quadratic				
		N'	CM'	C'	ng	
NER(%)	87,59	N	3989	104	26	4119
ER(%)	12,41	CM	3	210	0	213
MR(%)	5,09	C	470	66	524	1060
		ng'	4462	380	550	5392

**Tab. 6.6 B** Valori della *Sensibility* (Sn), *Specificity* (Sp) e del *Misclassification risk* (Mr), per ciascuna classe utilizzata per la redazione della mappa dell'erosione calanchiva, relativi al metodo *pseudoQuadratic*.

SnN(%)	96,84	SpN(%)	89,40	MrC (%)	2,93
SnCM(%)	98,59	SpCM(%)	55,26	MrF (%)	0,04
SnC(%)	49,43	SpC(%)	95,27	MrP (%)	2,12

**Tab. 6.7 A** Matrice di confusione relativa alle classi utilizzate per la redazione della mappa dell'erosione calanchiva, relativamente al *metodo diagQuadratic* (Non-Error-Rate (NER) Error-Rate (ER) e Misclassification risk complessivo MR)).

		diag Quadratic				
		N'	CM'	C'	ng	
NER(%)	87,20	N	3943	130	46	4119
ER(%)	12,80	CM	16	196	1	213
MR(%)	6,18	C	444	53	563	1060
		ng'	4403	379	610	5392

**Tab. 6.7 B** Valori della *Sensibility* (Sn), *Specificity* (Sp) e del *Misclassification risk* (Mr), per ciascuna classe utilizzata per la redazione della mappa dell'erosione calanchiva, relativi al metodo *diagQuadratic*.

SnN(%)	95,73	SpN(%)	89,55	MrC (%)	3,96
SnCM(%)	92,02	SpCM(%)	51,72	MrF (%)	0,25
SnC(%)	53,11	SpC(%)	92,30	MrP (%)	1,97

**Tab. 6.8 A** Matrice di confusione relativa alle classi utilizzate per la redazione della mappa dell'erosione calanchiva, relativamente al *metodo Linear* (Non-Error-Rate (NER) Error-Rate (ER) e Misclassification risk complessivo MR)).

		Linear				
		N'	CM'	C'	ng	
NER(%)	86,24	N	3994	46	79	4119
ER(%)	13,76	CM	31	167	15	213
MR(%)	5,75	C	458	113	489	1060
		ng'	4483	326	583	5392

**Tab. 6.8 B** Valori della *Sensibility* (Sn), *Specificity* (Sp) e del *Misclassification risk* (Mr), per ciascuna classe utilizzata per la redazione della mappa dell'erosione calanchiva, relativi al metodo *Linear*.

SnN(%)	96,97	SpN(%)	89,09	MrC (%)	2,81
SnCM(%)	78,40	SpCM(%)	51,23	MrF (%)	0,67
SnC(%)	46,13	SpC(%)	83,88	MrP (%)	2,26

**Tab. 6.9 A** Matrice di confusione relativa alle classi utilizzate per la redazione della mappa dell'erosione calanchiva, relativamente al *metodo diagLinear* (Non-Error-Rate (NER) Error-Rate (ER) e Misclassification risk complessivo MR)).

		diag Linear				
		N'	CM'	C'	ng	
NER(%)	73,16	N	3548	184	387	4119
ER(%)	26,84	CM	64	145	4	213
MR(%)	17,04	C	642	166	252	1060
		ng'	4254	495	643	5392

**Tab. 6.9 B** Valori della *Sensibility* (Sn), *Specificity* (Sp) e del *Misclassification risk* (Mr), per ciascuna classe utilizzata per la redazione della mappa dell'erosione calanchiva, relativi al metodo *diagLinear*.

SnN(%)	86,14	SpN(%)	83,40	MrC (%)	12,85
SnCM(%)	68,08	SpCM(%)	29,29	MrF (%)	0,99
SnC(%)	23,77	SpC(%)	39,19	MrP (%)	3,20

**Tab. 6.10 A** Matrice di confusione relativa alle classi utilizzate per la redazione della mappa dell'erosione calanchiva, relativamente al *metodo pseudoLinear* (Non-Error-Rate (NER) Error-Rate (ER) e Misclassification risk complessivo MR)).

		pseudo Linear				
		N'	CM'	C'	ng	
NER(%)	86,24	N	3994	46	79	4119
ER(%)	13,76	CM	31	167	15	213
MR(%)	5,75	C	458	113	489	1060
		ng'	4483	326	583	5392

**Tab. 6.10 B** Valori della *Sensibility* (Sn), *Specificity* (Sp) e del *Misclassification risk* (Mr), per ciascuna classe utilizzata per la redazione della mappa dell'erosione calanchiva, relativi al metodo *pseudoLinear*.

SnN(%)	96,97	SpN(%)	89,09	MrC (%)	2,81
SnCM(%)	78,40	SpCM(%)	51,23	MrF (%)	0,67
SnC(%)	46,13	SpC(%)	83,88	MrP (%)	2,26

Nelle figure 6.26 e 6.27 vengono riportati i risultati delle elaborazioni solo per le mappe derivanti dall'applicazione dei metodi *pseudoQuadratic* e *diagQuadratic*, elaborate associando alla classe di ciascun pixel il valore della *Posterior Probability*. Si può osservare che il *pattern* relativo alla distribuzione di queste categorie è sostanzialmente comparabile con quello ottenuto con più categorie in competizione. Tuttavia, a parità di metodo, mentre la parte relativa ai calanchi mammellonari è praticamente la stessa, si ha una contrazione degli areali relativi ai calanchi. Detta differenza si ripercuote in modo proporzionale a seconda dell'algoritmo di calcolo che, come sopra evidenziato, dà maggiore enfasi ad un tipo di erosione piuttosto che ad un altro.

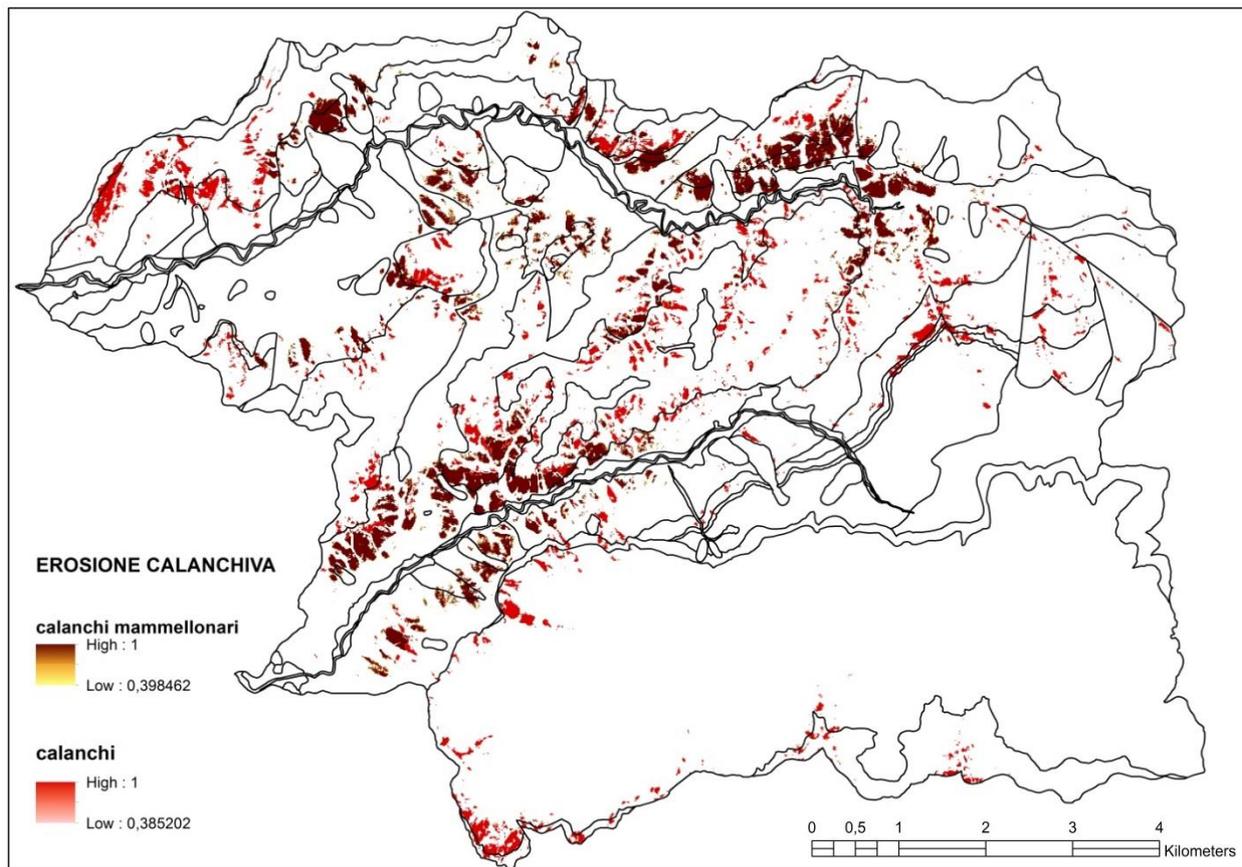


Fig. 6.26 Aree soggette a dinamica calanchiva elaborate con il metodo *pseudoQuadratic*, con i lineamenti delle formazioni geologiche (cfr. fig. 2.6).

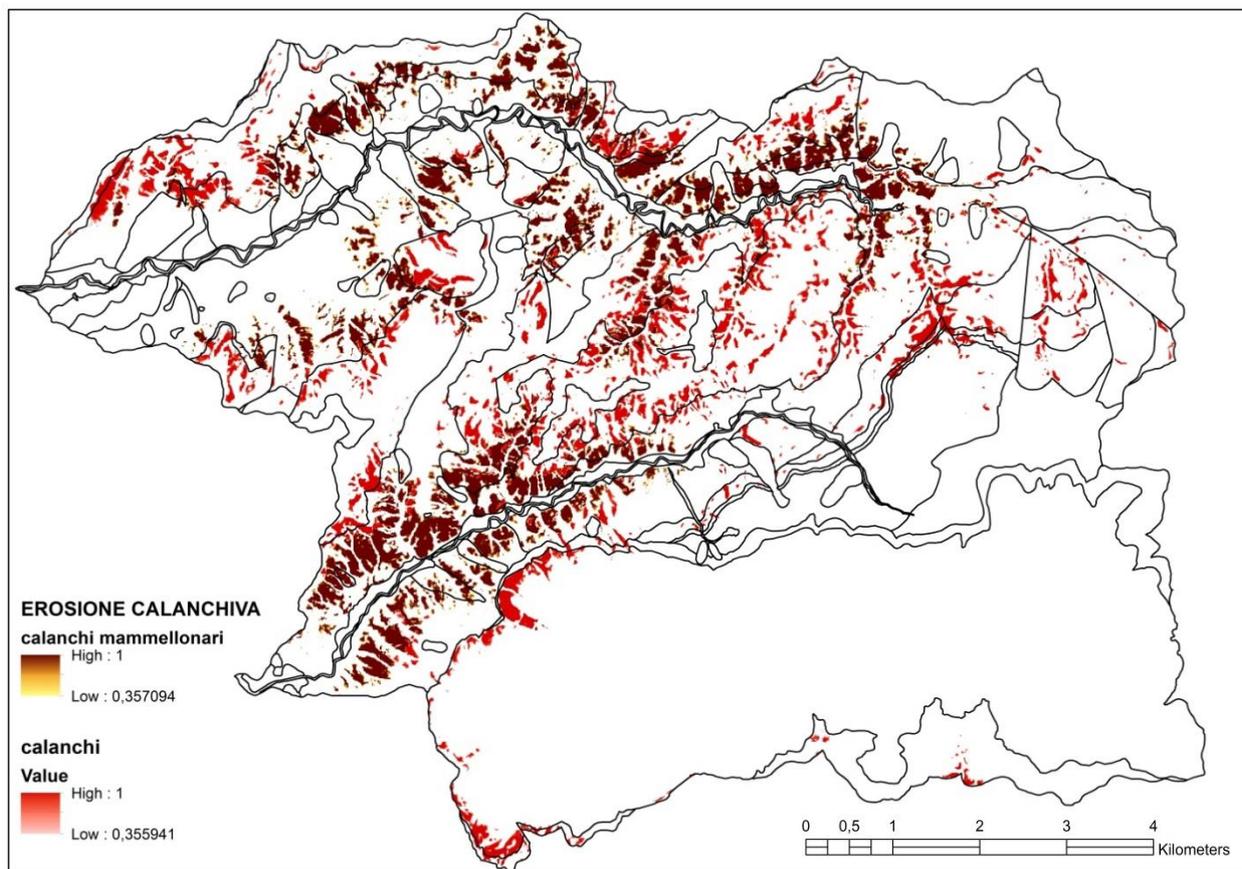


Fig. 6.27 Aree soggette a dinamica calanchiva elaborate con il metodo *diagQuadratic*, con i lineamenti delle formazioni geologiche (cfr. fig. 2.6).

## 6.4 Mappa della propensione alle frane superficiali

### 6.4.1 Materiali e metodi

Per la costruzione della mappe di suscettibilità alle frane, in letteratura esistono differenti approcci sia deterministici che probabilistici o di *machine learning*, e risulta difficile stabilire a priori quale sia il metodo più appropriato (Guzzetti et al. 2006; Lee e Ho 2009, Melchiorre et al. 2011): questo perché i metodi ottimali, usati per la redazioni di mappe tematiche, spesso differiscono a seconda dall'area geografica considerata. Poiché una mappa relativa alla suscettibilità alle frane dovrebbe essere utilizzati per la pianificazione territoriale, in letteratura viene data particolare enfasi alla validazione dei modelli (Melchiorre et al. 2008; Melchiorre et al. 2011). A riguardo F. Guzzetti (Guzzetti et al. 2006) ha individuato quattro aspetti fondamentali da considerare per effettuare una validazione di una mappa relativa alla suscettibilità da frana:

- 1) *“investigating the role of the thematic information in the production of the susceptibility model;*
- 2) *determining the model sensitivity and robustness to variations in the input data;*
- 3) *determining the error associated with the susceptibility prediction obtained for each mapping unit;*
- 4) *testing the model prediction against independent landslide information.”*

Questo tipo di modellazione generalmente non viene effettuata solo sulla base dei parametri geomorfometrici, ma prende in considerazioni strati informativi, tipicamente considerati importanti per spiegare questo tipo di fenomeno (es: dati geologici strutturali o geomeccanici).

Occorre evidenziare che l'obiettivo di questo lavoro non è quello di ottenere un modello di suscettibilità alla frane: si è tuttavia cercato di verificare se l'impiego delle sole variabili geomorfometriche potesse evidenziare le aree potenzialmente soggette al fenomeno, nonché ricercare l'eventuale relazione tra quest'ultimo e le caratteristiche dei suoli.

La base informativa utilizzata per la redazione di questa mappa tematica è costituita dalle variabili geomorfometriche e da un database costituito da 112 punti di innesco (v. fig. 6.34), relativi alle frane superficiali avvenute negli anni 2009 e 2014, messo a disposizione da ricercatori dell'Università di Pavia (Zizioli et al. 2013). La classe antagonista è costituita da tutti i punti appartenenti alle diverse categorie di erosione individuate nell'area di studio.

Indipendentemente dal metodo utilizzato per la redazione della mappa, mentre per la fase di *training* del procedimento di classificazione dicotomica è relativamente semplice fornire i casi positivi certi (presenza dell'evento), in quanto si fa riferimento agli eventi registrati, può risultare difficile fornire un numero congruo di casi negativi certi (assenza dell'evento). Per non incorrere nel problema di assegnare a priori casi falsi negativi (Melchiorre et al. 2008) (casi per i quali a priori è esclusa la verifica dell'evento ove invece è possibile), non sono stati inclusi nel *training set* i punti relativi al campionamento pedologico effettuato su nicchie di frana (P7\_2014, P17\_2014, P18\_2014, P19\_2014, P25\_2014, P32\_2014, P48\_2014, P53\_2014, P63\_2014) e quelli delle aree ove è stata verificata la presenza della suffosione. Come già esposto, l'areale relativo a questo tipo di erosione del suolo è potenzialmente interessato dalle frane superficiali, pertanto l'esclusione di questi punti è dovuta all'impossibilità di stabilire a priori quali parti del territorio siano stabili, e per

le quali sia certo che non si possa verificare l'evento frana superficiale. Il dataset dei punti di innesco (112 punti) è stato ripartito in modo casuale nel modo seguente: il 70% dei punti è andato a costituire il *training set*, mentre il rimanente 30% è stato suddiviso equamente tra *validation set* e *test set*. Per quanto riguarda i punti ove non si verifica l'evento, ai tre dataset dei punti di innesco è stato aggiunto un pari numero di punti, scelti casualmente, relativi alle categorie: *calanchi di tipo A, B, C* (C), *calanchi mammellonari* (CM), *erosione da flusso idrico concentrato* (EIC), *roccia affiorante* (R), *erosione laminare* (EL), *erosione/deposizione* (ED), *aree stabili* (A), accorpati in unica categoria. Mantenendo costante il *set* di punti relativi a queste categorie, sono state utilizzate alcune differenti combinazioni di punti di innesco, relativamente al *training set*, al *test set* ed al *validation set*.

La mappa è stata costruita utilizzando una *Artificial neural network for pattern recognition* (ANN) del *toolbox* di MATLAB. E' stata utilizzata una rete neurale standard per *pattern recognition*, è una *feed-forward network* a doppio strato con una funzione *tangent-sigmoid* sia per i neuroni dell'*hidden layer* che per i neuroni in output (Demuth et al. 2008). Sono stati ottenuti differenti risultati modificando la topologia della rete, aumentando il numero di neuroni nell'*hidden layer* (10, 20 e 30). Sono stati trattenuti i migliori risultati ottenuti reiterando il *training* 30 volte; per la valutazione dei risultati sono stati considerati gli output del *tool*, in particolare la matrice di confusione, la *Receive Operating Characteristic* (ROC) e il plot relativo alla *performance*.

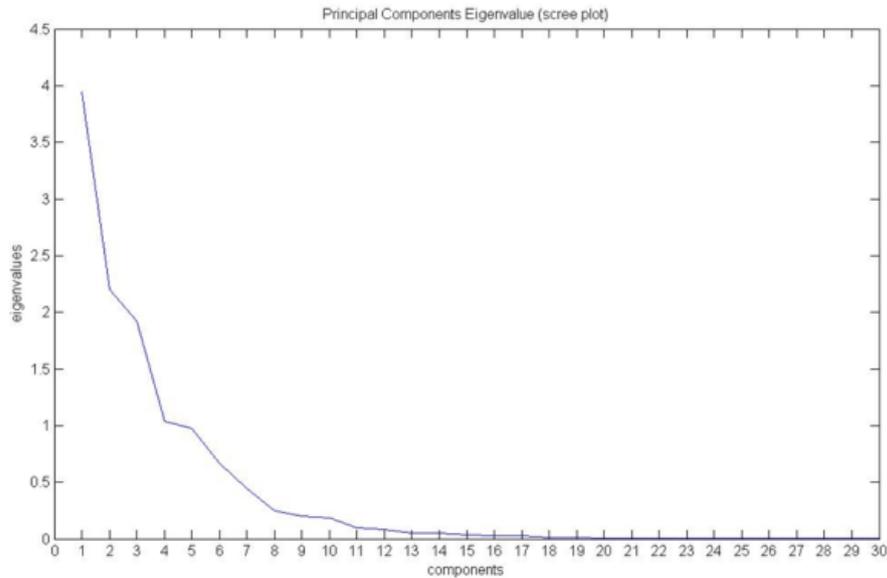
## 6.4.2 Analisi e risultati

Analogamente a quanto svolto nel capitolo 5 la selezione delle variabili da utilizzare è stata basata sui risultati della *Principal Component Analysis*, effettuata considerando solo i record relativi ai punti di innesco delle frane. Quest'ultimi sono stati individuati determinando il punto con maggiore elevazione di ogni areale di frana mappato.

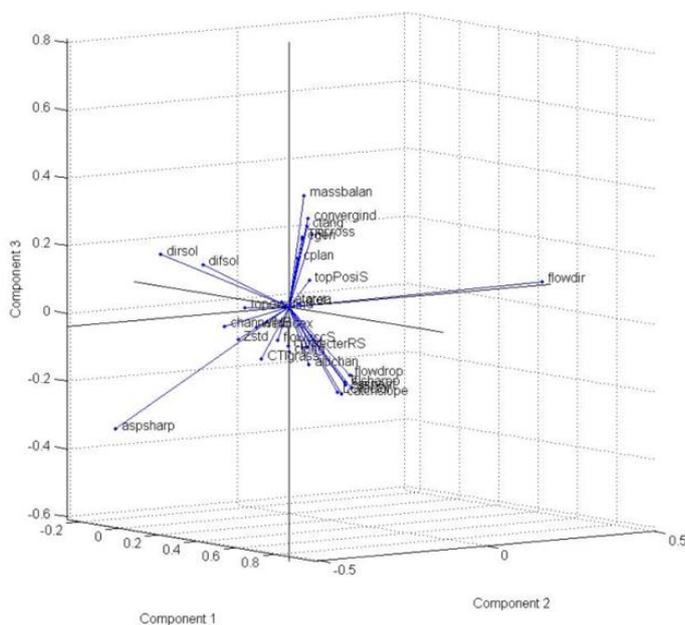
Oltre alle considerazioni già indicate nel capitolo 5, è stato scelto di non utilizzare le componenti principali come variabili indipendenti e di mantenere le variabili originarie per le seguenti motivazioni:

- 1) la determinazione delle componenti principali, basata sul set di dati relativi ai punti di innesco, fornisce un risultato completamente differente rispetto alla trasformazione delle variabili relative all'intero set di dati, sui quali occorre operare la predizione. Pertanto non avrebbe senso addestrare le reti neurali con le componenti principali derivanti dalla matrice dei punti di innesco, per poi utilizzare la rete addestrata per effettuare le predizioni sull'intero set di punti dell'area;
- 2) si vuole verificare se vi siano variabili geomorfometriche il cui eventuale significato fisico è in grado di fornire informazione sul meccanismo di innesco delle frane.

Anche in questo caso (v. Cap. 5), considerando i soli valori dei *loadings* (Tab. 6.11 e Fig. 6.29), in assenza di criteri oggettivi, possono essere considerati differenti *sub-sets* di variabili indipendenti, a seconda del numero di componenti che si vuole trattenere e di quello di variabili che si vuole considerare per ciascuna componente.



**Fig. 6.28** Scree plot relativo alla *principal component analysis* eseguito sul set di 30 variabili geomorfometriche.



**Fig. 6.29** Loadings plot relativo alla *principal component analysis* eseguito sul set di 30 variabili geomorfometriche.

Le prime otto componenti spiegano il 93,36%, fino ad arrivare al 97,38% nel caso di 11 componenti (v. Tab. 6.11). Secondo la regola di Kaiser in questo caso dovrebbero essere considerate solo le prime quattro componenti principali (v. Cap. 5), per le quali l'autovalore è maggiore di zero (v. Tab. 6.11 e Fig. 6.28).

Pertanto la costruzione dei *sub-sets* è stata effettuata basandosi sulla *rank analysis*, considerando le caratteristiche dello *scree-plot* (Fig. 6.28) e tenendo in considerazione i valori dei *loadings* che le singole variabili assumono nelle diverse componenti. Osservando il grafico (Fig. 6.28), per il caso in esame, è possibile individuare quattro variazioni di pendenza, in corrispondenza della seconda, quarta, ottava e undicesima componente.

La percentuale di varianza spiegata, secondo i cambi di pendenza dello *scree-plot*, è per le prime due componenti pari al 50,15% della varianza totale, per le prime quattro componenti pari 74,31%, mentre le

Date queste premesse, tenuto in considerazione che si vuole cercare una combinazione di variabili che possa fornire predizioni accettabili, è stato scelto di testare quattro differenti *sub-sets* di variabili, considerando quelle che nelle prime quattro componenti forniscono un maggior apporto. Per ciascuna di esse è stato calcolato l'apporto delle singole variabili, moltiplicando il valore del *loadings* per l'autovalore della componente. Sono stati evidenziati i contributi più importanti secondo la regola:  $\text{apporto} > \pm \text{deviazione standard}$  (calcolato per ciascuna componente). I quattro *sub-sets* sono stati determinati considerando le variabili evidenziate per ciascuna componente, accorrandole secondo il numero di componenti trattenute. Considerando solo la prima componente, sono state utilizzate 17 variabili (*dataset* DSPC1) incrementate a 19, 21, 24 trattenendo rispettivamente due (*dataset* DSPC2), poi tre (*dataset* DSPC3), ed infine quattro (*dataset* DSPC4) componenti (v. Tab. 6.12).

**Tab. 6.11** Valori dei *loadings* dei punti di innesco delle frane superficiali, distinti secondo l'importanza (intensità del colore) ed il segno (verdi negativi, rossi positivi). Per ciascuna componente è riportato l'autovalore (*eigenvalue*) e la percentuale di varianza spiegata (*percent\_explained*).

		PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12	PC13	PC14	PC15	
<b>percent_explained</b>		<b>32,220</b>	<b>17,929</b>	<b>15,695</b>	<b>8,464</b>	<b>7,976</b>	<b>5,464</b>	<b>3,612</b>	<b>2,005</b>	<b>1,672</b>	<b>1,545</b>	<b>0,797</b>	<b>0,647</b>	<b>0,443</b>	<b>0,433</b>	<b>0,321</b>	
<b>eigenvalue</b>		<b>3,947</b>	<b>2,196</b>	<b>1,923</b>	<b>1,037</b>	<b>0,977</b>	<b>0,669</b>	<b>0,443</b>	<b>0,246</b>	<b>0,205</b>	<b>0,189</b>	<b>0,098</b>	<b>0,079</b>	<b>0,054</b>	<b>0,053</b>	<b>0,039</b>	
1	elevation	<b>Zstd</b>	0,013	-0,159	-0,081	<b>0,493</b>	0,009	-0,250	0,089	-0,048	-0,095	0,110	-0,054	-0,036	-0,015	-0,035	-0,014
2	Terrain Ruggedness Index (QGIS)	<b>tri</b>	<b>0,270</b>	0,009	-0,198	0,026	-0,096	0,053	0,041	-0,065	-0,059	-0,021	-0,042	-0,110	-0,122	-0,074	-0,031
3	Topographic position index (QGIS)	<b>TPI</b>	<b>0,185</b>	-0,069	<b>0,229</b>	-0,063	-0,055	-0,158	0,051	0,060	0,035	-0,026	0,064	0,096	-0,072	-0,049	-0,116
4	Asperit (QGIS)	<b>asperit</b>	<b>0,273</b>	0,027	-0,216	0,005	-0,148	0,039	0,015	-0,068	-0,162	-0,010	0,087	-0,159	0,078	-0,148	-0,193
5	total Curvature (Jenness tool)	<b>ctot</b>	0,063	-0,003	0,017	-0,003	-0,034	-0,025	-0,003	0,085	0,018	0,021	0,102	0,172	0,015	0,127	<b>-0,329</b>
6	tangential Curvature (Jenness tool)	<b>ctang</b>	0,150	-0,033	<b>0,256</b>	0,016	0,111	0,132	<b>0,329</b>	0,039	-0,060	-0,068	0,009	0,120	-0,029	-0,059	<b>-0,231</b>
7	profile Curvature (Jenness tool)	<b>cprof</b>	<b>-0,144</b>	0,083	-0,137	0,113	0,184	<b>0,380</b>	<b>0,253</b>	-0,040	-0,130	-0,008	-0,139	-0,048	0,072	-0,059	0,080
8	plan Curvature (Jenness tool)	<b>cplan</b>	0,068	-0,013	0,151	0,005	0,061	0,090	0,183	-0,015	-0,061	-0,019	-0,001	-0,020	-0,005	-0,021	-0,014
9	longitudinal Curvature (Jenness tool)	<b>clong</b>	<b>-0,155</b>	0,092	-0,155	0,117	0,159	<b>0,341</b>	<b>0,245</b>	-0,053	-0,157	-0,030	-0,144	-0,081	0,041	0,044	0,090
10	general Curvature (Jenness tool)	<b>cgen</b>	0,192	-0,073	<b>0,227</b>	-0,062	-0,055	-0,159	0,037	0,064	0,026	-0,028	0,082	0,116	-0,029	-0,023	<b>-0,209</b>
11	cross-sectional Curvature (Jenness tool)	<b>ccross</b>	0,172	-0,032	<b>0,233</b>	0,014	0,068	0,075	<b>0,315</b>	0,056	-0,116	-0,079	-0,006	0,117	-0,009	0,005	<b>-0,269</b>
12	aspect (Jenness tool)	<b>aspsharp</b>	0,060	<b>-0,551</b>	<b>-0,307</b>	<b>-0,337</b>	<b>0,573</b>	-0,192	0,011	-0,130	-0,204	-0,135	0,182	0,059	0,011	0,037	0,055
13	slope (Jenness tool)	<b>slsharp</b>	<b>0,272</b>	0,010	-0,204	0,028	-0,102	0,055	0,029	-0,064	-0,070	-0,005	-0,040	-0,120	-0,060	-0,100	-0,050
14	wetness index (SAGA)	<b>wetindex</b>	<b>-0,205</b>	0,025	-0,081	-0,133	-0,160	<b>-0,231</b>	0,189	0,087	-0,050	-0,109	-0,050	0,070	-0,019	0,010	0,105
15	topographic wetness index (SAGA)	<b>topowetind</b>	<b>-0,217</b>	-0,003	-0,023	<b>-0,277</b>	-0,098	-0,214	0,212	0,324	<b>-0,333</b>	<b>0,600</b>	-0,038	<b>-0,205</b>	-0,013	-0,044	-0,073
16	Compound Topographic Index (GRASS)	<b>CTIgrass</b>	<b>-0,239</b>	0,059	-0,183	-0,115	-0,221	<b>-0,245</b>	<b>0,352</b>	0,018	0,004	<b>-0,391</b>	-0,107	0,220	0,257	<b>-0,480</b>	-0,035
17	catchment Area (recursive method) (SAGA)	<b>catarea</b>	<b>-0,019</b>	-0,009	0,013	-0,073	0,013	-0,018	0,028	0,109	-0,110	<b>0,286</b>	0,019	-0,097	0,001	-0,023	-0,026
18	mass balance index (SAGA)	<b>massbalan</b>	0,244	-0,098	<b>0,360</b>	-0,088	-0,045	-0,204	0,114	-0,068	0,069	-0,052	-0,165	<b>-0,242</b>	-0,053	-0,174	<b>0,632</b>
19	LS factor (SAGA)	<b>LSfactor</b>	0,184	0,038	<b>-0,237</b>	-0,024	-0,184	-0,043	0,121	-0,081	-0,063	-0,140	-0,072	<b>-0,214</b>	<b>-0,296</b>	0,119	-0,011
20	convergence index (SAGA)	<b>convergind</b>	0,110	-0,007	<b>0,272</b>	0,006	0,110	0,129	<b>0,336</b>	-0,012	<b>-0,207</b>	-0,028	0,005	-0,053	-0,082	0,157	0,134
21	channel network base level (SAGA)	<b>channelNB</b>	<b>-0,034</b>	-0,171	-0,046	<b>0,454</b>	-0,028	<b>-0,281</b>	0,123	<b>-0,305</b>	-0,081	0,186	-0,099	-0,050	-0,002	-0,020	-0,078
22	catchment Area (max flux method) (SAGA)	<b>catchslope</b>	0,198	0,041	<b>-0,242</b>	0,014	-0,017	0,020	<b>0,426</b>	-0,006	<b>0,535</b>	0,314	0,411	0,113	0,283	0,141	0,205
23	altitude above channel network (SAGA)	<b>altichan</b>	0,158	-0,033	-0,153	<b>0,353</b>	0,123	-0,013	<b>-0,070</b>	<b>0,808</b>	-0,091	-0,194	0,121	0,032	-0,046	-0,064	0,200
24	flow direction (ArcGIS)	<b>flowdir</b>	0,079	<b>0,711</b>	0,021	0,032	0,290	<b>-0,337</b>	-0,038	-0,172	<b>-0,328</b>	-0,070	0,344	0,028	0,053	0,045	0,090
25	flow drop (ArcGIS)	<b>flowdrop</b>	<b>0,317</b>	-0,005	-0,170	0,002	-0,135	-0,039	-0,008	-0,031	-0,061	-0,005	0,031	-0,141	-0,004	-0,096	-0,081
26	diffuse insolation (SAGA)	<b>difsol</b>	<b>-0,234</b>	-0,119	0,113	<b>0,371</b>	0,108	-0,212	0,065	-0,061	0,078	-0,019	0,064	0,073	-0,085	0,092	-0,038
27	direct insolation (SAGA)	<b>dirsol</b>	<b>-0,204</b>	-0,262	0,158	0,082	<b>-0,409</b>	0,213	-0,043	-0,096	<b>-0,324</b>	<b>-0,152</b>	<b>0,643</b>	<b>-0,168</b>	0,156	-0,001	0,114
28	flow accumulation (SAGA)	<b>flowaccS</b>	<b>-0,067</b>	0,007	-0,106	-0,085	-0,217	-0,182	0,191	0,085	-0,013	-0,291	-0,119	-0,102	-0,040	<b>0,734</b>	-0,010
29	Topographic position index (SAGA)	<b>topPosiS</b>	0,200	-0,056	0,102	0,045	0,023	-0,058	<b>-0,150</b>	0,041	-0,144	-0,023	<b>-0,285</b>	<b>-0,122</b>	<b>0,823</b>	0,181	-0,038
30	Vector Ruggedness Measure (SAGA)	<b>vectorRS</b>	0,182	-0,052	-0,097	0,010	-0,258	0,093	<b>-0,082</b>	-0,096	-0,358	0,201	-0,132	<b>0,739</b>	-0,044	0,123	0,309

Tab. 6.11 (segue).

		PC16	PC17	PC18	PC19	PC20	PC21	PC22	PC23	PC24	PC25	PC26	PC27	PC28	PC29	PC30	
percent explained		<b>0,202</b>	<b>0,196</b>	<b>0,113</b>	<b>0,080</b>	<b>0,057</b>	<b>0,049</b>	<b>0,027</b>	<b>0,018</b>	<b>0,016</b>	<b>0,010</b>	<b>0,007</b>	<b>0,002</b>	<b>0,001</b>	<b>0,000</b>	<b>0,000</b>	
eigenvalue		0,025	0,024	0,014	0,010	0,007	0,006	0,003	0,002	0,002	0,001	0,001	0,000	0,000	1,82E-12	5,81E-14	
1	elevation	Zstd	0,032	0,040	-0,034	-0,136	-0,030	-0,101	-0,132	-0,059	-0,118	0,061	0,085	-0,041	-0,030	0,737	0,001
2	Terrain Ruggedness Index (QGIS)	tri	-0,132	-0,058	0,103	-0,279	-0,176	-0,116	-0,054	0,134	0,571	-0,219	-0,220	-0,476	0,084	0,000	0,000
3	Topographic position index (QGIS)	TPI	-0,096	0,019	0,051	-0,050	0,043	-0,046	-0,024	0,673	-0,227	-0,408	0,379	-0,012	-0,010	0,000	0,000
4	Asperit (QGIS)	asperit	0,298	0,262	-0,495	0,414	0,339	-0,026	-0,093	0,104	0,061	-0,046	-0,083	-0,058	-0,018	0,000	0,000
5	total Curvature (Jenness tool)	ctot	0,018	-0,123	0,197	-0,278	0,650	-0,421	0,241	-0,124	0,038	0,024	-0,104	0,061	-0,010	0,000	0,000
6	tangential Curvature (Jenness tool)	ctang	-0,223	0,238	0,022	0,142	-0,241	-0,177	0,132	0,049	-0,092	0,238	-0,266	-0,101	-0,555	0,000	0,000
7	profile Curvature (Jenness tool)	cprof	-0,155	0,190	0,000	0,034	-0,044	-0,326	0,182	0,214	-0,225	0,037	-0,103	0,040	0,575	0,000	0,000
8	plan Curvature (Jenness tool)	cplan	0,134	-0,128	-0,044	0,265	-0,220	-0,396	0,169	-0,401	0,251	-0,291	0,518	0,064	0,000	0,000	0,000
9	longitudinal Curvature (Jenness tool)	clong	-0,082	0,069	0,060	-0,146	0,320	0,325	-0,004	-0,003	0,082	-0,182	0,246	-0,028	-0,326	-0,001	0,457
10	general Curvature (Jenness tool)	cgen	-0,019	0,076	0,000	0,027	-0,116	0,065	-0,114	-0,134	0,015	0,112	-0,075	0,036	0,367	-0,001	0,770
11	cross-sectional Curvature (Jenness tool)	ccross	-0,117	0,201	0,062	-0,104	0,128	0,447	-0,201	-0,235	0,110	0,007	0,123	0,033	0,299	0,001	-0,446
12	aspect (Jenness tool)	aspsharp	-0,031	0,007	0,007	-0,002	0,013	-0,016	-0,004	-0,009	-0,005	-0,004	0,005	0,005	-0,002	0,000	0,000
13	slope (Jenness tool)	slsharpp	-0,078	-0,025	0,067	-0,152	-0,149	-0,021	-0,048	0,107	0,189	-0,016	-0,093	0,839	-0,091	0,000	0,000
14	wetness index (SAGA)	wetindex	0,514	0,512	0,350	-0,079	-0,067	-0,026	0,146	0,136	0,208	0,152	0,097	0,021	-0,007	0,000	0,000
15	topographic wetness index (SAGA)	topowetind	-0,104	-0,015	0,022	0,037	-0,028	0,022	0,001	-0,104	-0,086	-0,277	-0,195	0,022	-0,018	0,000	0,000
16	Compound Topographic Index (GRASS)	CTIgrass	-0,122	-0,341	-0,121	-0,011	0,026	-0,003	-0,043	-0,017	-0,017	0,022	-0,017	-0,010	-0,008	0,000	0,000
17	catchment Area (recursive method) (SAGA)	catarea	-0,185	-0,199	-0,115	0,055	0,074	0,007	0,076	0,279	0,325	0,658	0,392	-0,037	0,046	0,000	0,000
18	mass balance index (SAGA)	massbalan	-0,195	0,149	-0,066	-0,036	0,321	-0,082	0,098	-0,105	0,039	0,023	-0,083	0,031	0,011	0,000	0,000
19	LS factor (SAGA)	LSfactor	-0,100	-0,188	0,553	0,455	0,112	0,003	-0,159	-0,050	-0,224	0,117	0,053	-0,093	0,025	0,000	0,000
20	convergence index (SAGA)	convergind	0,558	-0,499	-0,070	-0,118	-0,024	0,050	-0,079	0,172	-0,067	0,073	-0,220	0,005	0,020	0,000	0,000
21	channel network base level (SAGA)	channeINB	0,035	0,047	-0,037	-0,160	-0,033	-0,108	-0,136	-0,058	-0,128	0,068	0,091	-0,045	-0,030	-0,646	-0,001
22	catchment Area (max flux method) (SAGA)	catchslope	0,001	0,000	0,046	0,039	-0,014	0,031	-0,055	-0,004	-0,011	-0,016	0,002	-0,010	0,003	0,000	0,000
23	altitude above channel network (SAGA)	altichan	0,005	-0,005	-0,007	0,015	-0,005	-0,024	-0,049	-0,030	-0,022	0,007	0,018	-0,008	-0,015	-0,200	0,000
24	flow direction (ArcGIS)	flowdir	-0,096	0,031	0,025	-0,023	-0,021	-0,028	-0,004	-0,012	-0,001	-0,001	0,008	0,006	0,000	0,000	0,000
25	flow drop (ArcGIS)	flowdrop	0,067	-0,067	-0,052	-0,189	-0,143	0,316	0,724	-0,117	-0,281	0,040	0,098	-0,108	0,039	0,000	0,000
26	diffuse insolation (SAGA)	difsol	-0,084	-0,111	0,050	0,416	0,090	0,249	0,395	0,176	0,341	-0,172	-0,251	0,119	0,088	0,000	0,000
27	direct insolation (SAGA)	dirsol	-0,114	0,026	0,093	-0,057	-0,024	-0,016	-0,016	-0,008	-0,003	0,004	0,005	0,001	0,001	0,000	0,000
28	flow accumulation (SAGA)	flowaccS	-0,205	0,047	-0,374	-0,087	-0,071	-0,065	0,029	0,002	0,002	-0,007	0,026	0,031	0,007	0,000	0,000
29	Topographic position index (SAGA)	topPosiS	0,010	-0,045	0,249	0,113	-0,039	0,016	-0,003	0,069	0,046	-0,020	-0,035	-0,022	-0,010	0,000	0,000
30	Vector Ruggedness Measure (SAGA)	vecterRS	-0,059	-0,008	-0,030	0,109	0,005	0,029	0,039	-0,015	0,001	-0,014	-0,004	-0,002	-0,003	0,000	0,000

**Tab. 6.12** Determinazione dei sub-sets di variabili. Sono evidenziati i valori dei *loadings* per le prime quattro componenti (PC1-PC4), gli *outliers* per ciascuna componente (valori > ± deviazione standard dell'autovettore) e i sub-set di variabili selezionate evidenziate in giallo.

			PC1	PC2	PC3	PC4								
percent explained			32,220	17,929	15,695	8,464	PC1	PC2	PC3	PC4	DSPC1	DSPC2	DSPC3	DSPC4
eigenvalue			3,947	2,196	1,923	1,037	outlier	outlier	outlier	outlier				
1	elevation	Zstd	0,013	-0,159	-0,081	0,493				0,493	Zstd	Zstd	Zstd	Zstd
2	Terrain Ruggedness Index (QGIS)	tri	0,270	0,009	-0,198	0,026	0,270		-0,198		tri	tri	tri	tri
3	Topographic position index (QGIS)	TPI	0,185	-0,069	0,229	-0,063	0,185		0,229		TPI	TPI	TPI	TPI
4	Asperit (QGIS)	asperit	0,273	0,027	-0,216	0,005	0,273		-0,216		asperit	asperit	asperit	asperit
5	total Curvature (Jenness tool)	ctot	0,063	-0,003	0,017	-0,003					ctot	ctot	ctot	ctot
6	tangential Curvature (Jenness tool)	ctang	0,150	-0,033	0,256	0,016			0,256		ctang	ctang	ctang	ctang
7	profile Curvature (Jenness tool)	cprof	-0,144	0,083	-0,137	0,113					cprof	cprof	cprof	cprof
8	plan Curvature (Jenness tool)	cplan	0,068	-0,013	0,151	0,005					cplan	cplan	cplan	cplan
9	longitudinal Curvature (Jenness tool)	clong	-0,155	0,092	-0,155	0,117					clong	clong	clong	clong
10	general Curvature (Jenness tool)	cgen	0,192	-0,073	0,227	-0,062	0,192		0,227		cgen	cgen	cgen	cgen
11	cross-sectional Curvature (Jenness tool)	ccross	0,172	-0,032	0,233	0,014	0,172		0,233		ccross	ccross	ccross	ccross
12	aspect (Jenness tool)	aspsharp	0,060	-0,551	-0,307	-0,337		-0,551	-0,307	-0,337	aspsharp	aspsharp	aspsharp	aspsharp
13	slope (Jenness tool)	slsharp	0,272	0,010	-0,204	0,028	0,272		-0,204		slsharp	slsharp	slsharp	slsharp
14	wetness index (SAGA)	wetindex	-0,205	0,025	-0,081	-0,133	-0,205				wetindex	wetindex	wetindex	wetindex
15	topographic wetness index (SAGA)	topowetind	-0,217	-0,003	-0,023	-0,277	-0,217			-0,277	topowetind	topowetind	topowetind	topowetind
16	Compound Topographic Index (GRASS)	CTIgrass	-0,239	0,059	-0,183	-0,115	-0,239		-0,183		CTIgrass	CTIgrass	CTIgrass	CTIgrass
17	catchment Area (recursive method) (SAGA)	catarea	-0,019	-0,009	0,013	-0,073					catarea	catarea	catarea	catarea
18	mass balance index (SAGA)	massbalan	0,244	-0,098	0,360	-0,088	0,244		0,360		massbalan	massbalan	massbalan	massbalan
19	LS factor (SAGA)	LSfactor	0,184	0,038	-0,237	-0,024	0,184		-0,237		LSfactor	LSfactor	LSfactor	LSfactor
20	convergence index (SAGA)	convergind	0,110	-0,007	0,272	0,006			0,272		convergind	convergind	convergind	convergind
21	channel network base level (SAGA)	channelNB	-0,034	-0,171	-0,046	0,454			0,454		channelNB	channelNB	channelNB	channelNB
22	catchment Area (max flux method) (SAGA)	catchslope	0,198	0,041	-0,242	0,014	0,198		-0,242		catchslope	catchslope	catchslope	catchslope
23	altitude above channel network (SAGA)	altichan	0,158	-0,033	-0,153	0,353			0,353		altichan	altichan	altichan	altichan
24	flow direction (ArcGIS)	flowdir	0,079	0,711	0,021	0,032		0,711			flowdir	flowdir	flowdir	flowdir
25	flow drop (ArcGIS)	flowdrop	0,317	-0,005	-0,170	0,002	0,317				flowdrop	flowdrop	flowdrop	flowdrop
26	diffuse insolation (SAGA)	difsol	-0,234	-0,119	0,113	0,371	-0,234			0,371	difsol	difsol	difsol	difsol
27	direct insolation (SAGA)	dirsol	-0,204	-0,262	0,158	0,082	-0,204		-0,262		dirsol	dirsol	dirsol	dirsol
28	flow accumulation (SAGA)	flowaccS	-0,067	0,007	-0,106	-0,085					flowaccS	flowaccS	flowaccS	flowaccS
29	Topographic position index (SAGA)	topPosiS	0,200	-0,056	0,102	0,045	0,200				topPosiS	topPosiS	topPosiS	topPosiS
30	Vector Ruggedness Measure (SAGA)	vecterRS	0,182	-0,052	-0,097	0,010	0,182				vecterRS	vecterRS	vecterRS	vecterRS

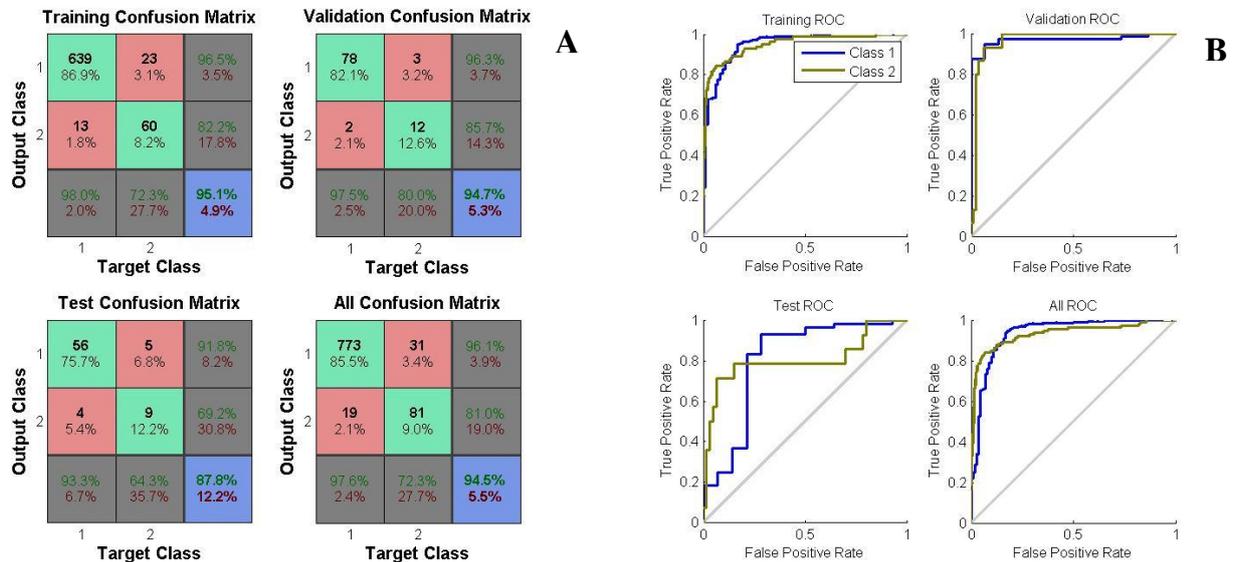
Osservando la percentuale dei punti di innesco classificati correttamente, derivanti dai risultati delle matrici di confusione (Tab. 6.13), risulta che le differenti configurazioni delle ANN, applicate ai diversi sets di variabili, hanno ottenuto prestazioni comparabili. Confrontando la percentuale di punti classificati correttamente, relativamente al *validation set* e al *test set*, emerge che il miglior risultato è stato ottenuto utilizzando 19 variabili (dataset DSPC2), relative alle prime due componenti. Per questo set di variabili sono state ottenute prestazioni analoghe sia utilizzando una ANN con topologia costituita da 10 neuroni, sia con quella con 30 neuroni (v. Tab. 6.13). Per tutte le elaborazioni effettuate è stata ottenuta una bassa percentuale di punti di innesco classificati correttamente relativamente al *test set*, che comunque raggiunge il 64% per il set citato.

**Tab. 6.13** Percentuale dei punti di innesco classificati correttamente per ciascun sub-set di variabili, derivanti dai risultati delle matrici di confusione, secondo le differenti topologie di ANN (ANN10: rete neurale con 10 neuroni, ANN20: rete neurale con 20 neuroni, ANN30: rete neurale con 30 neuroni).

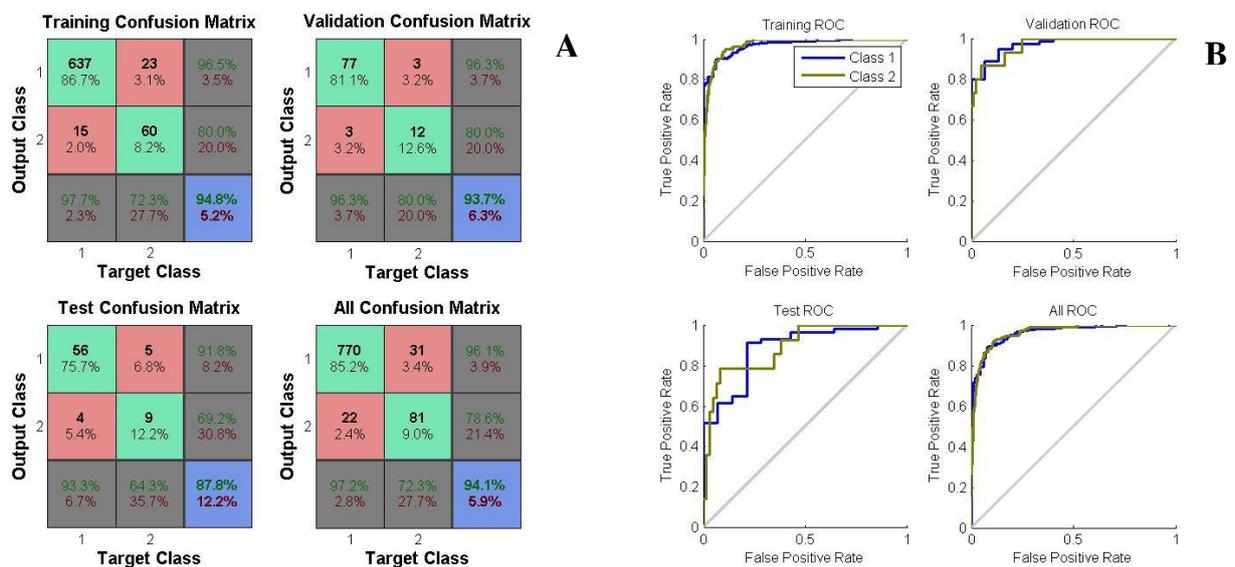
	DSPC1	DSPC2	DSPC3	DSPC4
training set ANN10	71,1%	72,3%	75,9%	77,1%
validation set ANN10	73,3%	80,0%	73,0%	80,0%
test set ANN10	57,1%	64,3%	50,0%	50,0%
training set ANN20	65,1%	66,3%	75,9%	78,3%
validation set ANN20	73,3%	73,3%	80,0%	80,0%
test set ANN20	57,0%	50,0%	50,0%	57,1%
training set ANN30	62,7%	72,3%	73,5%	79,5%
validation set ANN30	73,3%	80,0%	73,3%	73,3%
test set ANN30	57,1%	64,3%	50,0%	50,0%

Nelle seguenti figure sono rappresentati gli output delle ANN applicate al set di variabili DSPC2 per la ANN con dieci (Fig. 6.30 A e B) e con 30 neuroni (Fig. 6.31 A e B). Nelle matrici di confusione è riportato il numero di punti classificati correttamente e la relativa percentuale (riquadri verdi), viceversa con il colore rosso è indicata la quota mal classificata. La prima riga (classe 1 nel plot della ROC - *Receive Operating Characteristic*) è relativa al caso “evento assente” mentre la seconda si riferisce ai punti di innesco (classe 2 nel plot della ROC). La ROC (Fig. 6.30 B 6.31 B) permette di fare una ulteriore valutazione: osservando i grafici a destra, si può confrontare la quota di veri positivi (sensibilità) verso i falsi positivi (1-specificità), facendo variare il valore di soglia critica. Nel caso di perfetta classificazione la curva coinciderebbe con l’angolo in alto a sinistra: per il caso in esame emerge chiaramente la differenza di prestazioni riguardanti il *test set*. Anche se quest’ultimo non viene utilizzato per l’addestramento della rete, esso può fornire una indicazioni sulla qualità del processo di addestramento e quindi sulla capacità di generalizzazione della ANN. Informazioni a riguardo possono essere ottenute osservando la curva relativa alla *performance* (Fig. 6.32 A e B). Questa, riprodotta su scala logaritmica, indica la variazione del *mean squared error*, tra il valore calcolato ed il valore *target*, per ogni step (*epoch*) del processo di addestramento. Se in questa rappresentazione l’errore sul *test set* raggiunge il minimo ad un numero di iterazioni significativamente diverso da quello dell’errore di validazione, vi potrebbe essere una distorta suddivisione del dataset e nel caso in cui la curva del *test set* abbia un incremento significativo prima che la curva relativa al *validation set* cresca, allora è possibile che si sia verificato un *overfitting* (Beale et al. 1992). Osservando le curve relative alla *performance*, per i risultati

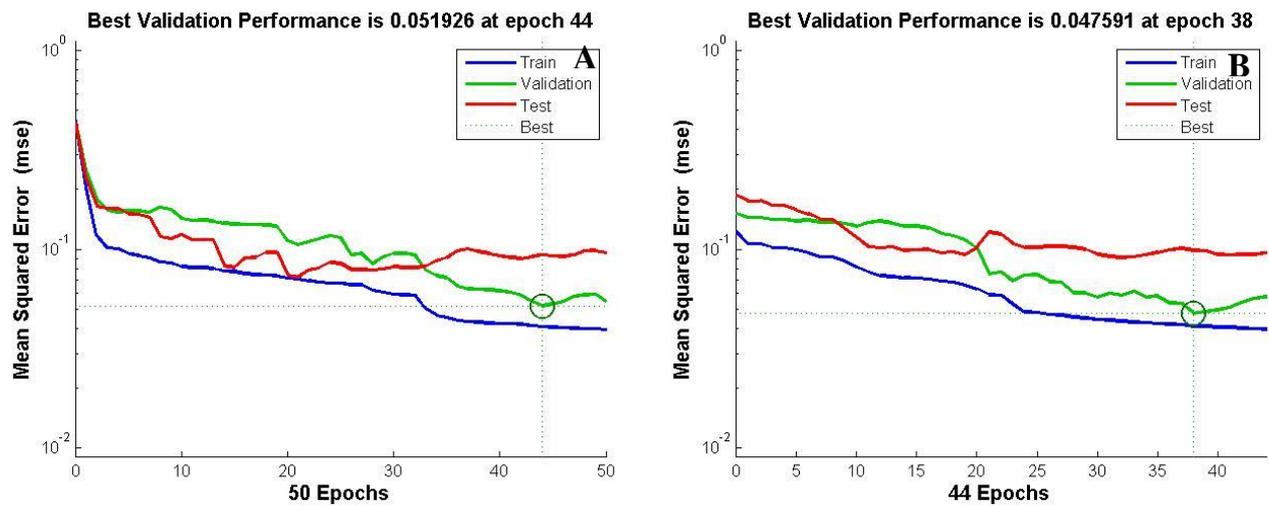
considerati, emerge che per entrambe le ANN il *test set* raggiunge il minimo del *mean squared error* con un numero di iterazioni significativamente diverso rispetto a quello del *validation set*, mentre la ANN con 30 neuroni potrebbe essere affetta anche da *overfitting* (v. Fig. 6.32 B). Pertanto è stata prodotta la mappa ottenuta con la ANN con 10 neuroni, assegnando a ciascun pixel il valore di probabilità calcolato dalla rete neurale (Fig. 6.33 e Fig. 6.34).



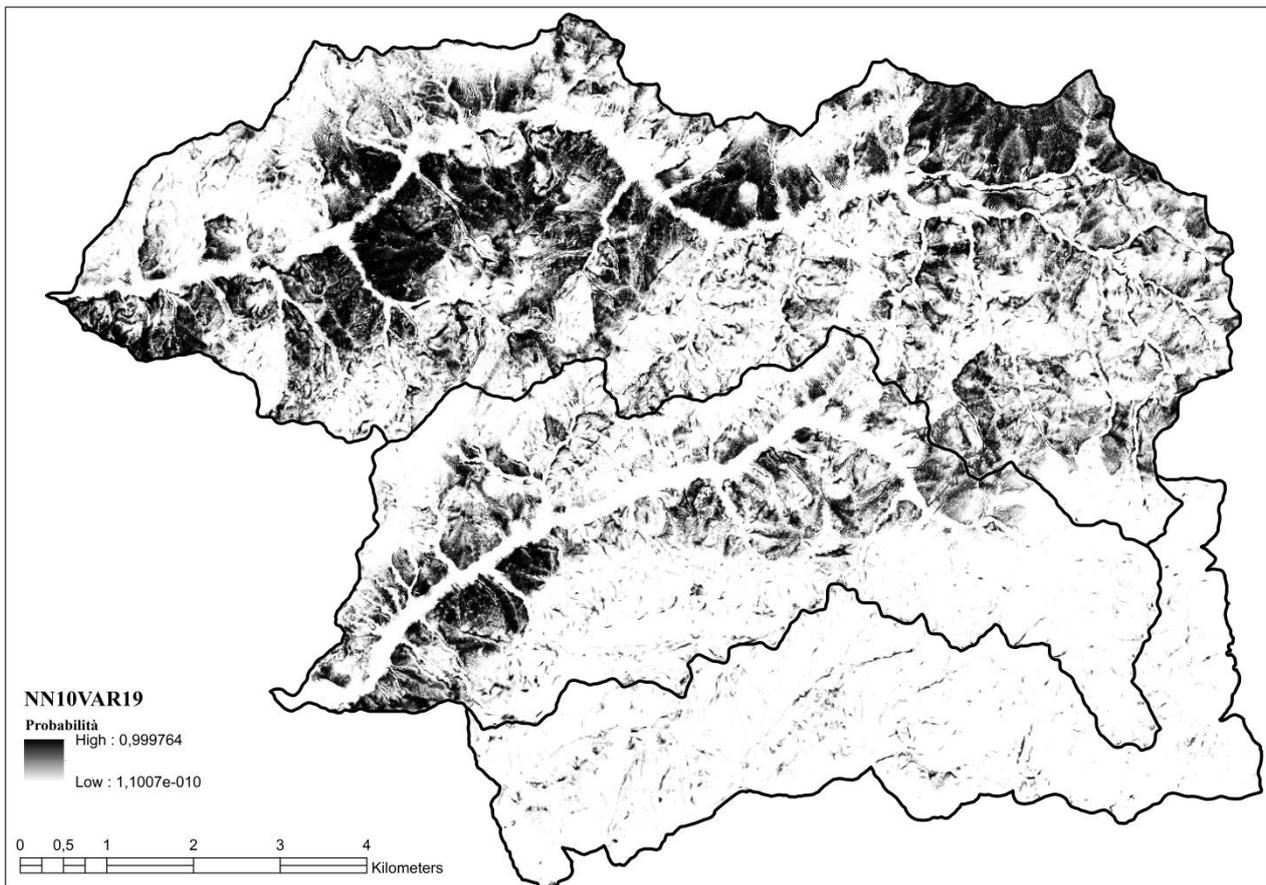
**Fig. 6.30** Output della ANN con 10 neuroni, applicata al sub-set di variabili DSPC2: matrice di confusione (A) curva ROC (classe 1 “evento assente” classe 2 “ punto di innesco” (B).



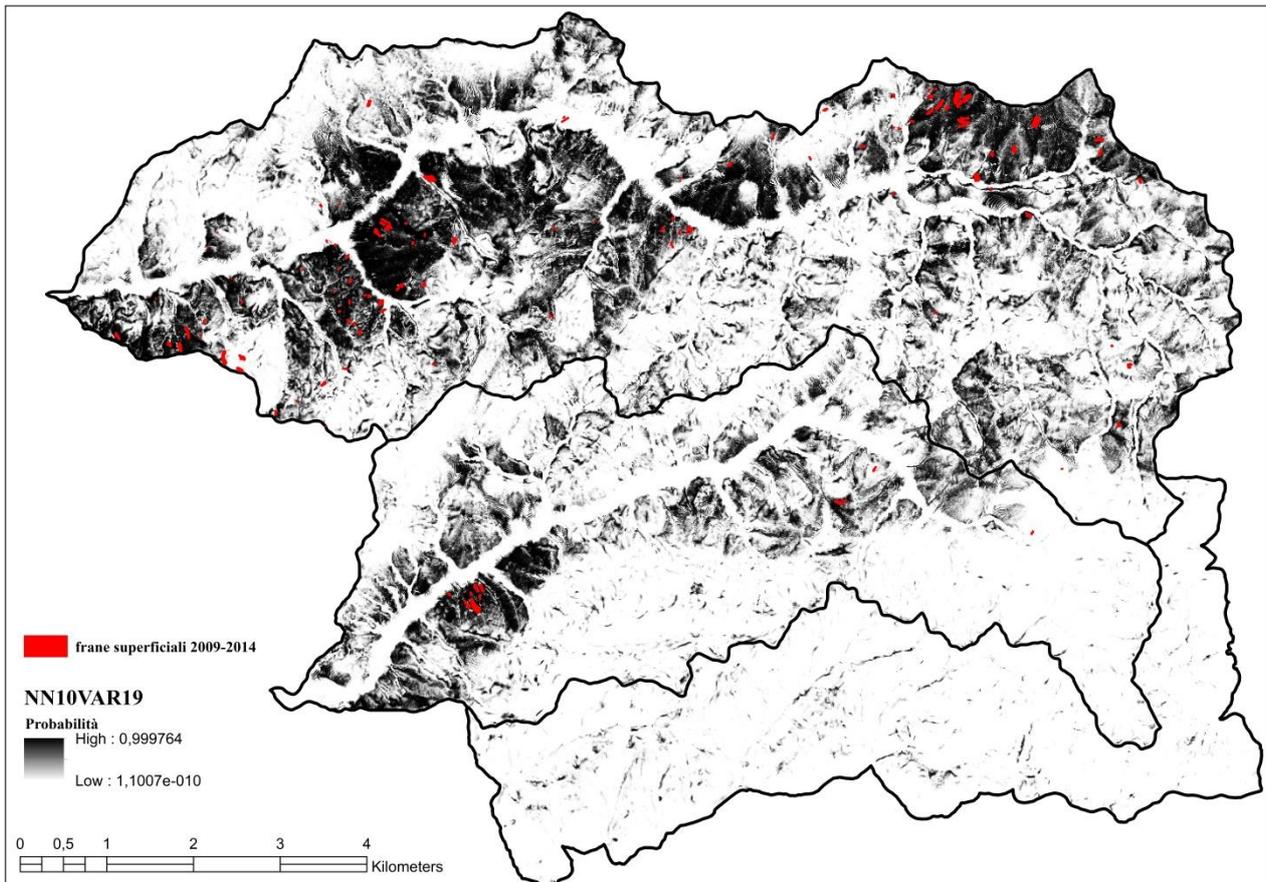
**Fig. 6.31** Output della ANN con 30 neuroni, applicata al sub-set di variabili DSPC2: matrice di confusione (A) curva ROC (classe 1 “evento assente” classe 2 “ punto di innesco” (B).



**Fig. 6.32** Performance della ANN applicata al sub-set di variabili DSPC2: ANN con 10 neuroni (A) ANN con 30 neuroni (B).



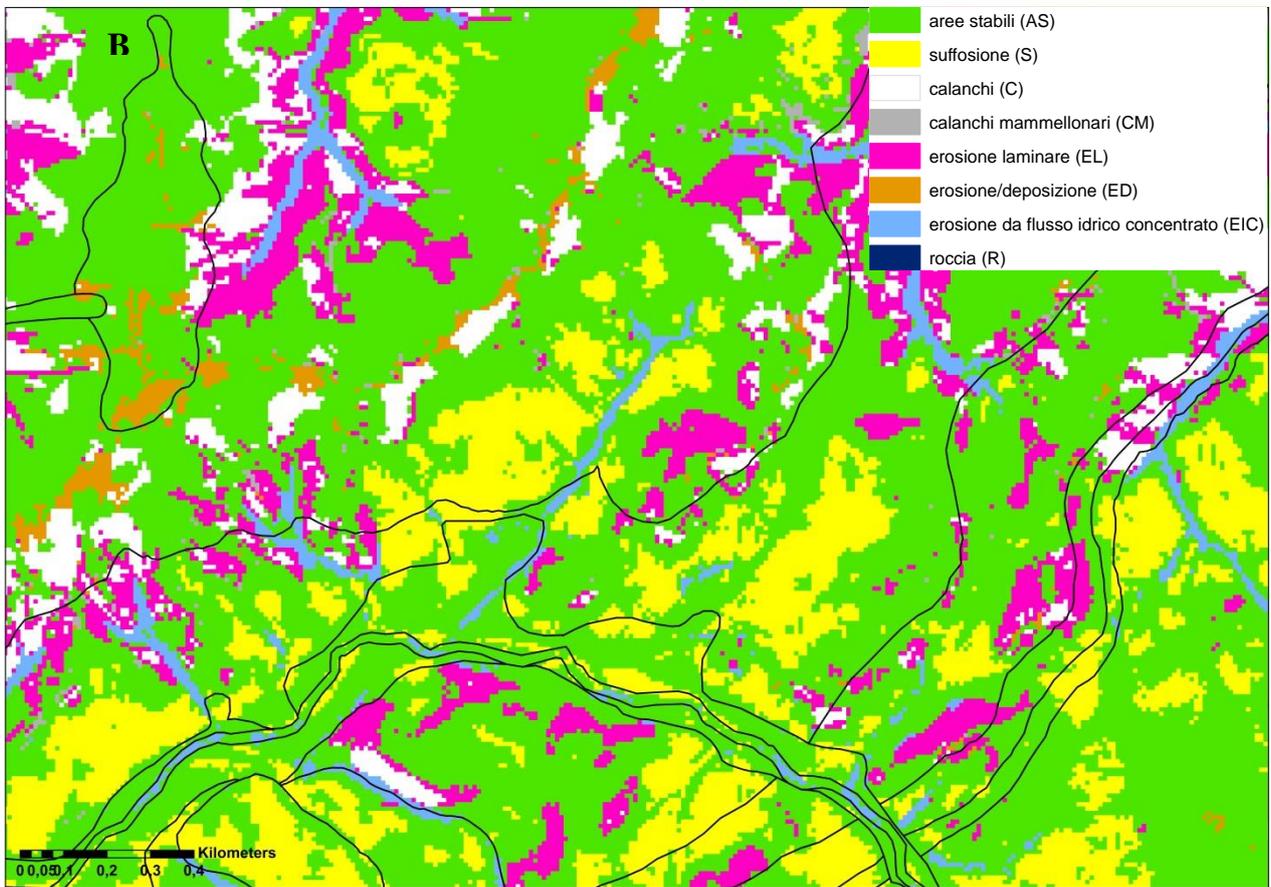
**Fig. 6.33** Mappa della PROPENSIONE ALLE FRANE SUPERFICIALI realizzata con il sub-set di variabili DSPC2 e con la ANN con 10 neuroni.



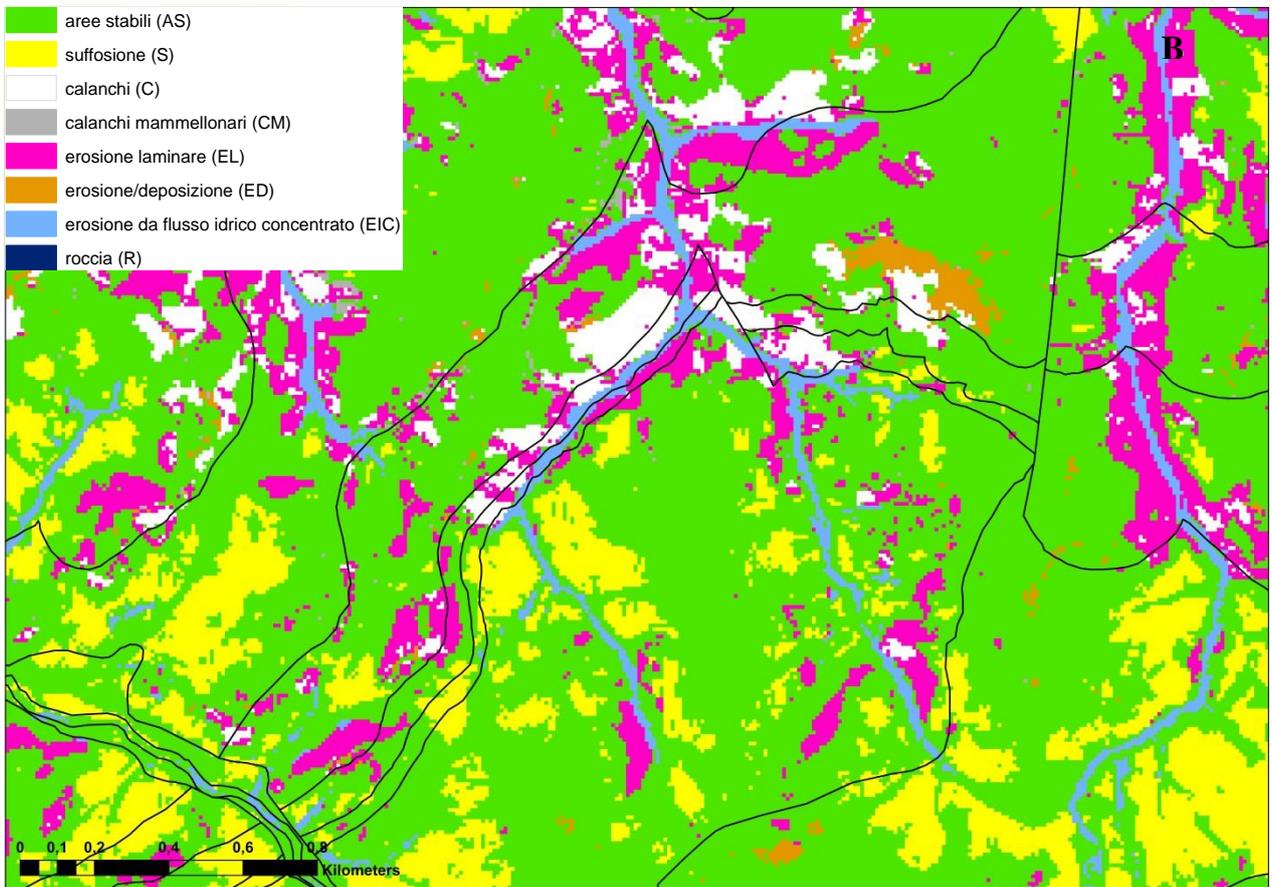
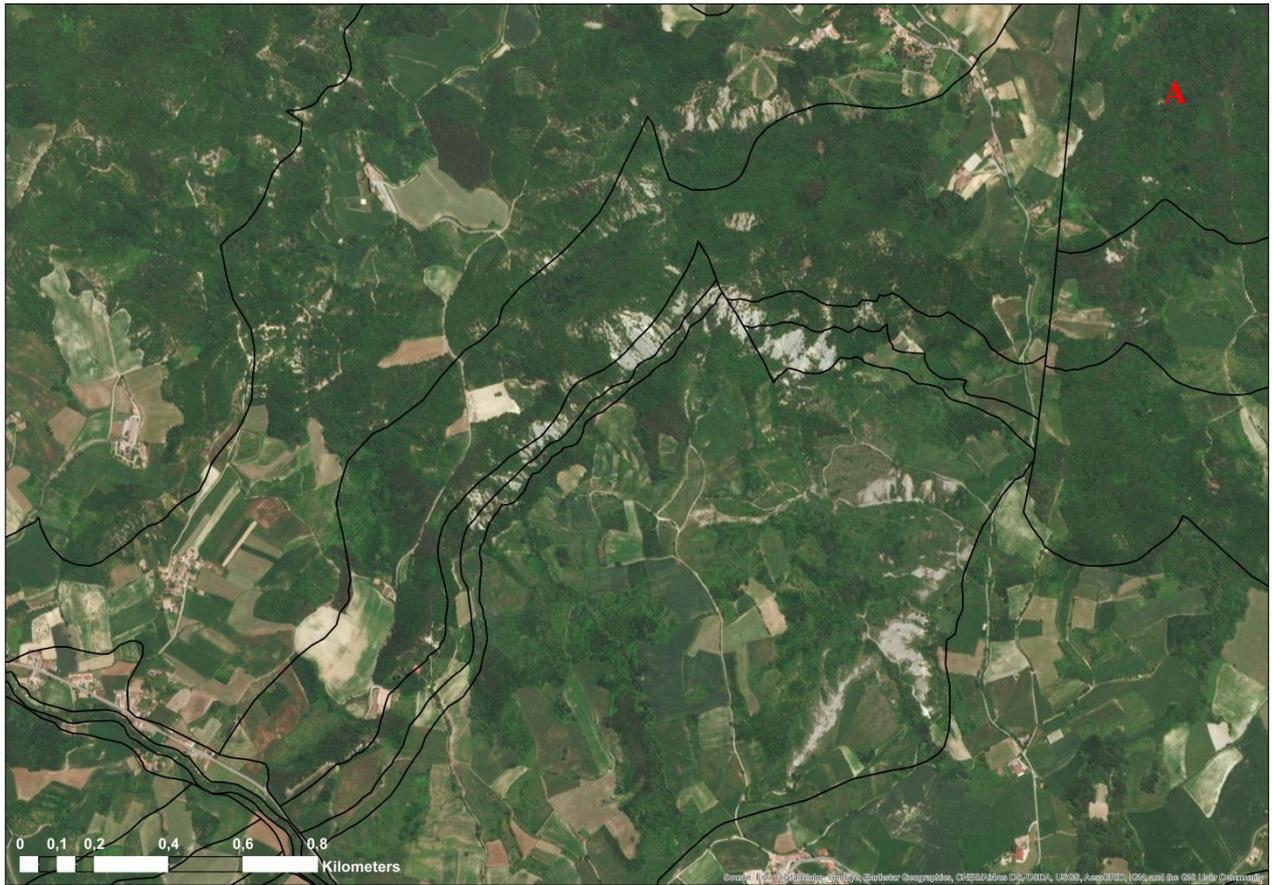
**Fig. 6.34** Mappa della PROPENSIONE ALLE FRANE SUPERFICIALI realizzata con il sub-set di variabili DSPC2 e con la ANN con 10 neuroni, con l'ubicazione delle frane superficiali impiegate per le elaborazioni.

## 6.5 Discussione e Conclusioni

Le mappe ottenute con la DA, anche se apparentemente mostrano differenze importanti per la dimensione della unità cartografiche, sono in realtà paragonabili tra loro. Infatti, le classi che si confondono o si sostituiscono sono compatibili tra loro. Come più volte evidenziato le forme calanchive presenti nell'area raramente sono associabili alle tipologie comunemente descritte nell'ambito scientifico (Calzolari e Ungaro 1998) e spesso negli areali soggetti ad erosione sono attivi più processi. Per esempio, spesso i calanchi hanno quote di areali planari ove la perdita di materiali non dà origine all'alternanza di creste e valli. Nelle diverse mappe calcolate dell'erosione del suolo, la maggior differenza emerge riguardo la distribuzione della soffusione. Per quanto riguarda questo fenomeno non si conosce la sua reale distribuzione: infatti, come evidenziato, questo tipo di erosione è stato individuato in corrispondenza di differenti formazioni geologiche, ed è stato osservato solo su aree adibite a coltivazioni agricole; l'evidenza di questo fenomeno svanisce con l'applicazione delle lavorazioni del terreno (aratura e fresatura). Con il *remote sensing* il fenomeno è osservabile se l'immagine satellitare è relativa al periodo stagionale in cui il fenomeno si manifesta (tipicamente al termine del periodo piovoso della primavera) e a seconda di quando è stata effettuata la lavorazione del terreno e la semina. Due esempi relativi alla complessità dei processi erosivi sono rappresentati nelle Figure 6.35 A e 6.36 A, unitamente alle relative predizioni, ottenute con il metodo *pseudoQuadratic* (Figure 6.35 B e 6.36 B).



**Fig. 6.35** Area soggetta ad erosione del suolo (A) e relativa predizioni ottenuta con il metodo *pseudoQuadratic* (B).

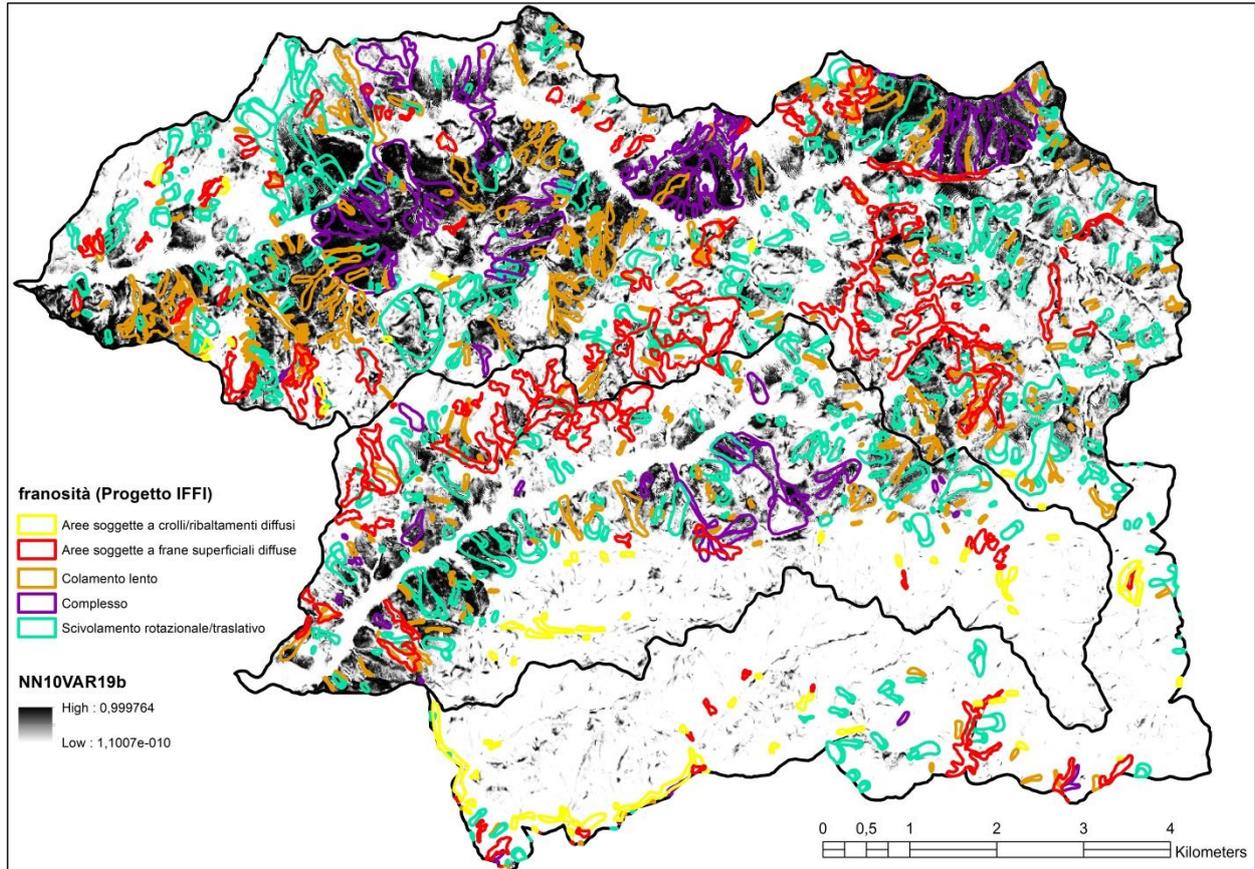


**Fig. 6.36** Area soggetta ad erosione del suolo (A) e relativa predizioni ottenuta con il metodo *pseudoQuadratic* (B).

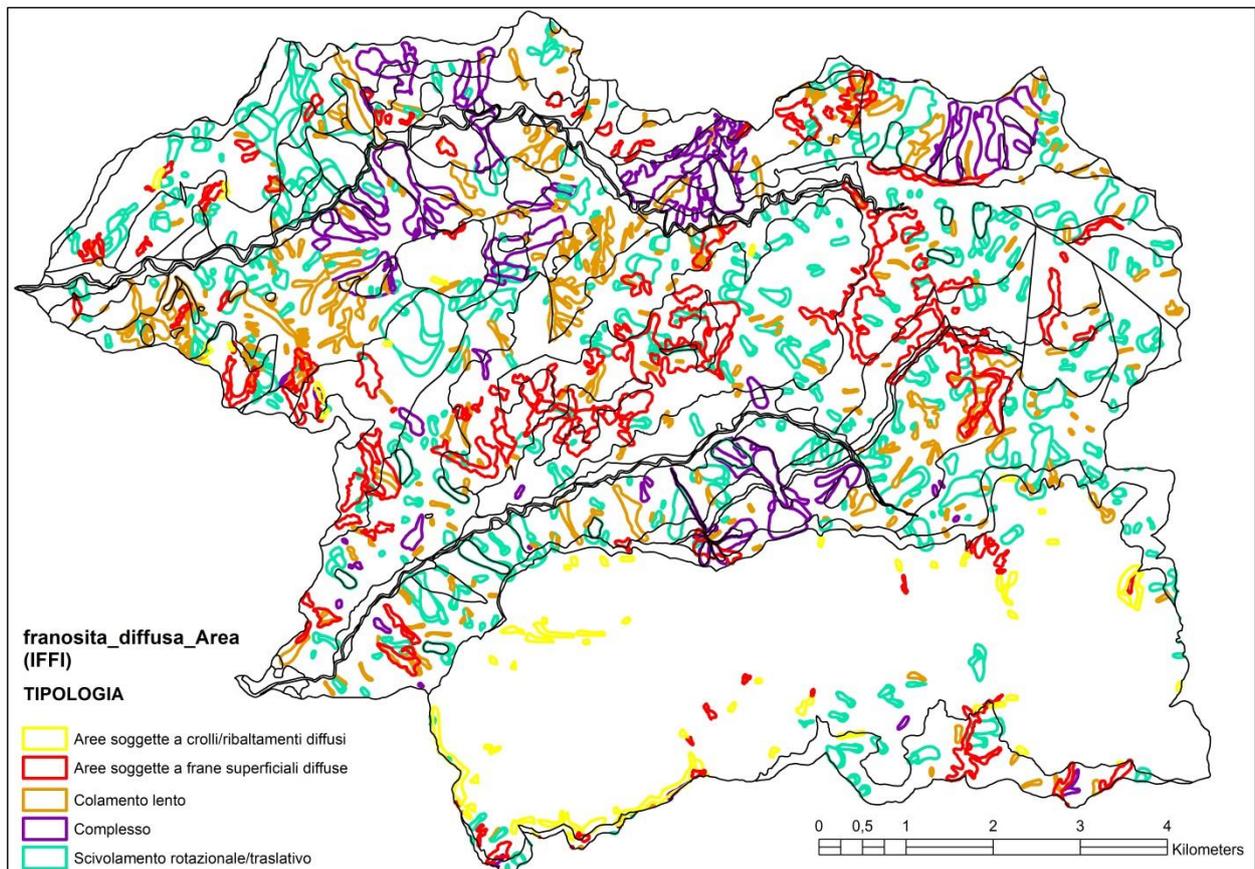
Riguardo la costruzione e la validazione della mappa relativa alla propensione alle frane superficiali, si evidenzia che non è stata effettuata una indagine approfondita, sia per il processo di calcolo, sia per quanto riguarda la validazione. In particolare, non è stata effettuata una analisi per indagare l'influenza delle singole variabili sull'output (*sensitivity analysis*); tuttavia il risultato accettato è relativo ad uno dei quattro differenti *sub-sets* di variabili utilizzate. Inoltre, non è stata effettuata una *cross-validation*, ottenendo più risultati con differenti distribuzioni di punti di innesco nei tre set (*training, validation, test*). Tuttavia a riguardo si evidenzia che una prima suddivisione casuale del set di punti di innesco ha costantemente fornito scarsi risultati. E' stato deciso di non effettuare ulteriori analisi perché: 1) la validazione ottenuta è ritenuta sufficiente per gli scopi di questo lavoro, 2) le indicazioni fornite sono da ritenersi complementari alla mappa dell'erosione del suolo; 3) le evidenze dimostrano che il risultato ottenuto sostanzialmente concorda con le indicazioni derivanti dal catalogo del progetto IFFI (Inventario dei Fenomeni Franosi in Italia) (v. Fig. 6.37 e 6.38).

Confrontando la mappa dell'erosione del suolo e quella della propensione alle frane superficiali, l'evidenza di maggior interesse riguarda il fatto che, come ipotizzato, vi sono aree soggette sia a soffusione che a frane superficiali. Questa evidenza induce a supporre che possa esserci un legame tra i due fenomeni di erosione: i processi di soffusione e la creazione di percorsi di scorrimento preferenziale delle acque sub-superficiali potrebbero favorire l'innesco delle frane superficiali. Detta congruenza avviene in particolare in corrispondenza delle formazioni delle Marne di Antognola e Brecce di Baiso. Confrontando gli areali di dette tipologie di erosione con le mappe create per la distribuzione spaziale dei tipi di suolo (v. Cap. 7) emerge che tali fenomeni tendenzialmente si sviluppano sui suoli con i maggiori contenuti di argilla quali i Leptic Vertisol (Calcaric) e i Leptic Calcaric Cambisol (Clayic).

Riguardo l'elaborazione della mappa relativa alla PROPENSIONE ALLE FRANE SUPERFICIALI, risulta di interesse la rilevanza che alcune variabili geomorfometriche hanno potenzialmente sul risultato. Detta importanza è emersa adottando il metodo di selezioni delle variabili proposto, il quale, secondo l'approccio utilizzato in questo lavoro, evita scelte soggettive e delega la selezione ad una analisi statistica. Infatti, considerando i sub-set di variabili (v. Tab. 6.12), ed in particolare il sub-set DSPC2 che ha fornito il migliore risultato, risulta che il maggior contributo è fornito dalle variabili *flow drop* (0,317) (v. fig. 3.21) e *flow direction* (0,711) (v. fig. 3.28), rispettivamente nella prima e nella seconda componente. Questa evidenza confermerebbe l'influenza che la circolazione idrica ha sui fenomeni franosi, presi in considerazione. Le altre variabili che, secondo il calcolo adottato per la selezione, potenzialmente forniscono i contributi più alti, sono: due relative alla morfologia generale dell'area di studio (*Terrain Ruggedness Index* e l'*Asperit*), insieme allo *slope*, per la prima componente e all'*aspect*, per la seconda componente. Questo secondo gruppo di variabili probabilmente è quello che concorre a determinare la differenza del risultato per le due valli, rispetto alla porzione appartenente alla placca di Pizzocorno-Pietragravina, che secondo la mappa è scevra da propensione alle frane superficiali, e per la quale il dataset di *training* non presentava questa tipologia di fenomeno. Nella realtà tale disparità è imputabile alle differenze nella composizione mineralogica delle formazioni ed alla loro disposizione nella successione stratigrafica che hanno determinato un differente risultato morfologico nelle evoluzioni oloceniche (v. paragrafo 6.1).



**Fig. 6.37** Mappa della PROPENSIONE ALLE FRANE SUPERFICIALI con indicazione delle aree franose dell'inventario IFFI.



**Fig. 6.38** Indicazione delle aree franose riportate nel catalogo IFFI.

## Bibliografia

- Alexander, D. E. (1980): I calanchi-accelerated erosion in Italy. In: *Geography*, pp. 95–100.
- Alexander, D. E. (1982): Difference between " calanchi" and" biancane" badlands in Italy. In: Geobooks.
- Battaglia, S.; Leoni, L.; Sartori, F. (2003): Mineralogical and grain size composition of clays developing calanchi and biancane erosional landforms. In: *Geomorphology* n. 1, 49, pp. 153–170.
- Beale, Mark Hudson; Hagan, Martin T.; Demuth, Howard B. (1992): Neural Network Toolbox™ User's Guide. In: *MathWorks Inc.*
- Bordoni, Massimiliano; Meisina, Claudia; Valentino, Roberto; Bittelli, Marco; Chersich, Silvia; Musetti, Marco; Persichillo, Maria Giuseppina (2017): Analysis of hydro-meteorological monitoring data collected in different contexts prone to shallow landslides of the Oltrepò Pavese (northern Italy). 4<sup>th</sup> World Landslide Forum Ljubijana (Slovenia).
- Buccolini, Marcello; Coco, Laura (2010): The role of the hillside in determining the morphometric characteristics of "calanchi". The example of Adriatic central Italy. In: *Geomorphology* n. 3-4, 123, pp. 200–210. DOI: 10.1016/j.geomorph.2010.06.003.
- Buccolini, Marcello; Coco, Laura (2013): MSI (morphometric slope index) for analyzing activation and evolution of calanchi in Italy. In: *Geomorphology*, 191, pp. 142–149. DOI: 10.1016/j.geomorph.2013.02.025.
- Buccolini, Marcello; Coco, Laura; Cappadonia, Chiara; Rotigliano, Edoardo (2012): Relationships between a new slope morphometric index and calanchi erosion in northern Sicily, Italy. In: *Geomorphology*, 149, pp. 41–48.
- Buccolini, Marcello; Gentili, Bernardino; Materazzi, Marco; Aringoli, Domenico; Pambianchi, Gilberto; Piacentini, Tommaso (2007): Human impact and slope dynamics evolutionary trends in the monoclinial relief of Adriatic area of central Italy. In: *CATENA* n. 1, 71, pp. 96–109. DOI: 10.1016/j.catena.2006.07.010.
- Calzolari, Costanza; Ungaro, Fabrizio (1998): Geomorphic features of a badland (biancane) area (Central Italy): characterisation, distribution and quantitative spatial analysis. In: *CATENA* n. 4, 31, pp. 237–256.
- Cappadonia, Chiara; Coco, Laura; Buccolini, Marcello; Rotigliano, Edoardo (2015): From slope morphometry to morphogenetic processes: an integrated approach of field survey, geographic information system morphometric analysis and statistics in Italian Badlands. In: *Land Degradation & Development*.
- Cappadonia, Chiara; Conoscenti, Christian; Rotigliano, Edoardo (2011): Monitoring of erosion on two calanchi fronts-Northern Sicily (Italy). In: *Landform analysis*, 17, pp. 21–25.
- Carelli, M.; Ceriani, M.; Valbuzzi, E. (2006): Inventario dei Fenomeni Franosi in Lombardia: Regione Lombardia.
- Castiglioni, BRUNO (1933): Osservazioni sui calanchi appenninici. In: *Bollettino della Società Geologica Italiana* n. 2, 2, pp. 357–360.

- Ciccacci, Sirio; Galiano, Mariachiara; Roma, Maria Antonietta; Salvatore, Maria Cristina (2008): Morphological analysis and erosion rate evaluation in badlands of Radicofani area (Southern Tuscany — Italy). In: *CATENA* n. 2, 74, pp. 87–97. DOI: 10.1016/j.catena.2008.03.012.
- Clarke, Michèle L.; Rendell, Helen M. (2000): The impact of the farming practice of remodelling hillslope topography on badland morphology and soil erosion processes. In: *CATENA* n. 2, 40, pp. 229–250. DOI: 10.1016/S0341-8162(99)00047-8.
- Cocco, Stefania; Brecciaroli, Giorgia; Agnelli, Alberto; Weindorf, David; Corti, Giuseppe (2015): Soil genesis and evolution on calanchi (badland-like landform) of central Italy. In: *Geomorphology*, 248, pp. 33–46. DOI: 10.1016/j.geomorph.2015.07.031.
- Crocetti, Cesare; Cavaglià, Adriana: I CALANCHI Caratteristiche delle morfologie calanchive.
- Della Seta, M.; Del Monte, M.; Fredi, P.; Lupia Palmieri, E. (2009): Space–time variability of denudation rates at the catchment and hillslope scales on the Tyrrhenian side of Central Italy. In: *Geomorphology* n. 3-4, 107, pp. 161–177. DOI: 10.1016/j.geomorph.2008.12.004.
- Della Seta, Marta; Del Monte, Maurizio; Fredi, Paola; Palmieri, Elvidio Lupia (2007): Direct and indirect evaluation of denudation rates in Central Italy. In: *CATENA* n. 1, 71, pp. 21–30. DOI: 10.1016/j.catena.2006.06.008.
- Demuth, Howard; Beale, Mark; Hagan, Martin (2008): Neural network toolbox™ 6. In: *User's guide*, pp. 37–55.
- Farifteh, Jamshid; Soeters, Rob (2006): Origin of biancane and calanchi in East Aliano, southern Italy. In: *Geomorphology* n. 1-2, 77, pp. 142–152. DOI: 10.1016/j.geomorph.2005.12.012.
- Farifteh, Jamshid; Soeters, Robert (1999): Factors underlying piping in the Basilicata region, southern Italy. In: *Geomorphology* n. 4, 26, pp. 239–251. DOI: 10.1016/S0169-555X(98)00070-1.
- Guasparri, G. (1993): I lineamenti geomorfologici dei terreni argillosi pliocenici. In: *La storia naturale della Toscana Meridionale*, pp. 89–106.
- Guzzetti, Fausto; Reichenbach, Paola; Ardizzone, Francesca; Cardinali, Mauro; Galli, Mirco (2006): Estimating the quality of landslide susceptibility models. In: *Geomorphology* n. 1, 81, pp. 166–184.
- Howard, Alan D. (1994): Badlands. In: *Geomorphology of desert environments*: Springer, pp. 213–242.
- Jones, Bradley (1997): MATLAB: Statistics Toolbox; User's Guide: MathWorks.
- Lee, Kwan Tun; Ho, Jui-Yi (2009): Prediction of landslide occurrence based on slope-instability analysis and hydrological model simulation. In: *Journal of Hydrology* n. 3, 375, pp. 489–497.
- Meisina, Claudia (2004): Swelling-shrinking properties of weathered clayey soils associated with shallow landslides. In: *Quarterly Journal of Engineering Geology and Hydrogeology*. 37(2): 77-94.
- Melchiorre, C.; Matteucci, M.; Azzoni, A.; Zanchi, A. (2008): Artificial neural networks and cluster analysis in landslide susceptibility zonation. In: *Geomorphology* n. 3, 94, pp. 379–400.

- Melchiorre, Caterina; Abella, E. CastellanosA; van Westen, Cees J.; Matteucci, Matteo (2011): Evaluation of prediction capability, robustness, and sensitivity in non-linear landslide susceptibility models, Guantánamo, Cuba. In: *Computers & Geosciences* n. 4, 37, pp. 410–425.
- Moretti, S.; Rodolfi, G. (2000): A typical “calanchi” landscape on the Eastern Apennine margin (Atri, Central Italy): geomorphological features and evolution. In: *CATENA* n. 2, 40, pp. 217–228.
- Phillips (1998): PII: S0143-6228(98)00005-8. In: *Mathematical Programming*, 82, p. 199. DOI: 10.1016/S0025-5610(97)00093-2.
- Piccarreta, M.; Capolongo, D.; Boenzi, F.; Bentivenga, M. (2006): Implications of decadal changes in precipitation and land use policy to soil erosion in Basilicata, Italy. In: *CATENA* n. 2, 65, pp. 138–151.
- Piccarreta, Marco; Faulkner, Hazel; Bentivenga, Mario; Capolongo, Domenico (2006a): The influence of physico-chemical material properties on erosion processes in the badlands of Basilicata, Southern Italy. In: *Geomorphology* n. 3-4, 81, pp. 235–251. DOI: 10.1016/j.geomorph.2006.04.010.
- Rendell, Helen (1982): Clay hillslope erosion rates in the Basento Valley, S. Italy. In: *Geografiska Annaler. Series A. Physical Geography*, pp. 141–147.
- Romero Díaz, A.; Marín Sanleandro, P.; Sánchez Soriano, A.; Belmonte Serrato, F.; Faulkner, H. (2007): The causes of piping in a set of abandoned agricultural terraces in southeast Spain. In: *CATENA* n. 3, 69, pp. 282–293. DOI: 10.1016/j.catena.2006.07.008.
- Ruiz-Flan, P.; Garci, J. M.; Ortigosa, Luis (1992): Geomorphological evolution of abandoned fields. A case study in the Central Pyrenees. In: *CATENA* n. 3-4, 19, pp. 301–308.
- Sdao, GIUSEPPE; Simone, ANTONIO; Vittorini, SEBASTIANO (1984): Osservazioni geomorfologiche su calanchi e biancane in Calabria. In: *Geografia Fisica e Dinamica Quaternaria*, 7, pp. 10–16.
- Summa, Vito; Luigia, Maria (2013): The Role of Soil Mineralogy, Geochemistry and Grain Size in the Development of Mediterranean Badlands: A Review. In: Maria C. Hernandez Soriano (a cura di): *Soil Processes and Current Trends in Quality Assessment: InTech*.
- Todeschini, Roberto (1998): Introduzione alla chemiometria. In: *EdiSES, Napoli*, p. 321.
- Torri, D.; Santi, E.; Marignani, M.; Rossi, M.; Borselli, L.; Maccherini, S. (2013): The recurring cycles of biancana badlands. Erosion, vegetation and human impact. In: *CATENA*, 106, pp. 22–30. DOI: 10.1016/j.catena.2012.07.001.
- Torri, Dino; Brya, Rorke (1997): Micropiping processes and biancana evolution in southeast Tuscany, Italy. In: *Geomorphology* n. 3-4, 20, pp. 219–235. DOI: 10.1016/S0169-555X(97)00025-1.
- Torri, Dino; Colica, Antonella; Rockwell, David (1994): Preliminary study of the erosion mechanisms in a biancana badland (Tuscany, Italy). In: *CATENA* n. 3-4, 23, pp. 281–294.
- Zizioli, D.; Meisina, C.; Valentino, R.; Montrasio, L. (2013): Comparison between different approaches to modeling shallow landslide susceptibility: a case history in Oltrepo Pavese, Northern Italy. In: *Natural Hazards and Earth System Sciences* n. 3, 13, p. 559.

## CAPITOLO 7

### Tipologie di suolo

#### 7.1 Introduzione

Considerando i notevoli cambiamenti nell'uso del territorio, il consumo di suolo e il suo degrado, è fondamentale conoscere le caratteristiche dei suoli e la loro distribuzione (Baruck et al. 2016).

La mappatura delle classi di suolo è stata importante per il progresso delle conoscenze nella scienza del suolo, sin dalla nascita della pedologia (Brevik et al. 2016). Una mappa dovrebbe essere lo strumento più semplice per trasmettere l'informazione, sia nel caso di rappresentazione di valori parametrici che per caratteristiche categoriali. Nel caso delle categorie tassonomiche dei suoli le mappe racchiudono una considerevole quantità di dati, che derivano in parte dall'osservazione effettuata nel corso del rilevamento pedologico ed in parte determinati con le analisi chimiche e fisiche. Altre mappe categoriali per comprendere la distribuzione di una o più caratteristiche del suolo possono essere create *ad hoc*, adottando una specifica classificazione. Nell'ambito del DSM sono stati utilizzati diversi metodi per mappare le variabili categoriali del suolo, che sono state distinte in *Pure classification techniques*, *Pure regression techniques*, *Pure geostatistical techniques Hybrid*, *statistical/geostatistical approaches* ed *Expert systems* (Hengl et al. 2007). Hengl ha comparato i risultati derivanti da: *supervised classification using maximum likelihoods*; *multinomial logistic regression*; *regression-kriging on memberships*; *classification of taxonomic distances*. Una sezione dedicata alla spazializzazione di attributi del suolo è contenuta nella pubblicazione di Lagacherie, McBratney e Voltz (Lagacherie et al. 2008).

Chagas ha utilizzato le *artificial neural networks* e la *maximum likelihood classification* (Chagas et al. 2013), W. Li ha utilizzato la *Bayesian Markov Chain Cosimulation* (Li et al. 2013), Taghizadeh ha utilizzato la *decision trees* (Taghizadeh-Mehrjardi et al. 2012), mentre Hitziger ha comparato i risultati ottenuti con la *linear regression*, con la *random forest classification*, e la *stochastic gradient boosting of regression trees* (Hitziger e Ließ 2014).

Considerando le caratteristiche dell'area di studio e le argomentazioni riguardo gli strati informativi, eventualmente utilizzabili nelle elaborazioni statistiche per la redazione delle mappe, in questo lavoro è stato utilizzato l'approccio teorico descritto nel capitolo 1. In particolare è stato considerato che:

- in molti casi, la topografia fornisce la migliore indicazione per comprendere la distribuzione spaziale delle proprietà del suolo nei luoghi dove la variazione degli altri fattori ambientali (ad esempio, clima, litologia, uso del suolo) è relativamente piccola (Maynard e Johnson 2014);
- che il suolo normalmente si sviluppa in relazione al movimento dell'acqua e in seguito all'alterazione del materiale litologico del substrato (Maynard e Johnson 2014; Florinsky 2012; Florinsky 2016);
- che l'analisi geomorfometrica del terreno è fondamentale nei territori dove la topografia è fortemente legata ai processi di formazione del suolo (Maynard e Johnson 2014; Florinsky 2012; Florinsky 2016).

In questo lavoro sono state elaborate due mappe, una utilizzando la classificazione tassonomica WRB - World Reference Base for Soil Resources (IUSS Working Group 2014), l'altra impiegando una classificazione empirica, che utilizza come caratteri discriminanti la tessitura e lo spessore del *solum*, che è stata pensata per applicazioni in campo agronomico.

Come informazione di base sono state utilizzate le variabili geomorfometriche, descritte nel capitolo capitolo 2. e le informazioni derivanti dall'attività di campionamento (v. Cap. 4); il metodo adottato per la derivazione delle mappe è quello dell'analisi discriminante (v. Cap. 6).

## 7.2 Caratteristiche generali dell'area di studio

Nell'area di studio (v. Cap. 2) i fattori che maggiormente influenzano i processi pedogenetici sono legati alle caratteristiche chimico-fisiche del *parent material* e alle dinamiche evolutive dei versanti. Le differenze strutturali e mineralogiche tra le formazioni geologiche determinano una caratteristica morfologia del territorio e come risultato fenomeni di erosione formazione-specifica, a cui è legata l'intensità di sviluppo dei suoli. In un contesto di così forte erosione (v. Cap. 6), l'equilibrio tra formazione del suolo ed asportazione del materiale pedogenizzato è spostato verso quest'ultimo, con la conseguenza di un rallentamento o interruzione dei processi di pedogenesi, che determinano una elevata presenza di suoli poco sviluppati.

Considerando l'influenza del *parent material* sulla pedogenesi, i suoli presenti nell'area di studio possono essere suddivisi in tre macrocategorie:

- suoli che si sviluppano sulle Arenarie del Monte Vallassa (arenarie bioclastiche e biocalcareni, intercalate da marne sabbiose grigio azzurre), caratterizzati da una tessitura con una importante componente sabbiosa;
- suoli che si sviluppano sulle Breccie di Baiso (corpi rocciosi a tessitura caotica costituiti da breccie poligeniche a matrice argillosa grigia a vari colori), con tessitura a forte contenuto di argilla, in prevalenza smectitica;
- suoli che si sviluppano sulla Formazione di Antognola (marne e marne argillose), sulla Formazione di Contignaco (marne siltose grigie, marne calcaree chiare, intercalate nella parte inferiore da strati arenaceo-pelitici), sulla Formazione di Ranzano con i suoi diversi membri nelle loro differenti *Litofacies*: (v. Cap 2), caratterizzati da una tessitura con poca sabbia e preponderanza di limo e argilla.

Considerando il sistema tassonomico WRB al primo livello della classificazione (Reference Soil Groups), nell'area di studio sono presenti quattro categorie: Cambisols, Regosols, Vertisols, Luvisols. I processi di pedogenesi attivi in quest'area sono la brunificazione, la lisciviazione e la vertisolizzazione. Nel processo di brunificazione la liberazione del ferro per l'alterazione del *parent material* porta alla formazione di complessi argilla-Fe-humus che strutturano il suolo e danno origine ad un orizzonte cambico (Previtali 2001). Nel processo di lisciviazione si ha il traslocamento, operato dall'acqua di percolazione, delle argille fini e grossolane in soluzione o in sospensione, attraverso i pori e le fessure; durante la discesa si ha la distruzione dei composti organici, per azione batterica e i suoli sono caratterizzati da una reazione acida o sub-acida (Previtali 2001). Il processo di vertisolizzazione è determinato dalla presenza di argille espandibili, che nel caso dei vertisuoli dell'area di studio sono ereditate dal *parent material*. La presenza di detto materiale determina la formazione di fessure verticali nella stagione secca e la loro chiusura

nella stagione umida. Queste variazioni volumetriche inducono movimenti detti “vertici”, che tendono a omogeneizzare il profilo (Previtali 2001).

### 7.3 Materiali e metodi

Per entrambe le mappe la base informativa per l’elaborazione statistica è costituita dalle variabili geomorfometriche primarie e secondarie utilizzate come predittori (v. Cap.2), mentre le variabili indipendenti sono in un caso le categorie tassonomiche secondo il sistema tassonomico WRB, e nell’altro le categorie tipologiche, stabilite secondo una chiave dicotomica basata sulla quota della frazione argillosa e sabbiosa e sullo spessore del *solum*. Dette assegnazioni sono state effettuate sulla base delle caratteristiche rilevate nel corso dell’attività di rilevamento pedologico e sulla base delle analisi chimico-fisiche svolte sui campioni di suolo.

L’attività di rilevamento pedologico dell’area è stata realizzata nel corso di due campagne, la prima delle quali è stata svolta nel periodo estivo-autunnale del 2014 basandosi sull’approccio fisiografico tipicamente utilizzato nella cartografia tradizionale (*free survey* (White 2013)), mediante osservazione diretta delle caratteristiche territoriali e con l’ausilio di una carta geologica (1:10.000) disponibile per la sola Val di Nizza (De Wrachien et al.1981) (V. Cap. 2). La seconda campagna di rilevamento (svolta dalla primavera all’autunno del 2015) è stata effettuata basandosi sull’apposita mappa elaborata su base statistica, come descritto nel capitolo 4.

I punti di campionamento scelti per il rilevamento pedologico sono stati georeferenziati e per ogni stazione di rilevamento è stata redatta una descrizione dell’uso del suolo, della fisiografia e della vegetazione. E’ stata eseguita l’apertura manuale dei profili pedologici (uno per ogni stazione) fino al raggiungimento del substrato litologico. L’esame dei profili è stato effettuato mediante l’individuazione degli orizzonti e l’annotazione su un’apposita scheda delle caratteristiche morfologiche principali, in particolare: profondità e andamento dei limiti degli orizzonti, colore a umido (*Munsell Soil Color Charts*), abbondanza, dimensione, e grado di alterazione dello scheletro, tipo di struttura, quantità e dimensione delle radici, effervescenza all’HCl. Da ciascun orizzonte è stato prelevato un campione di suolo per le successive analisi chimico-fisiche. Nel corso della prima campagna di rilevamento sono stati rilevati 62 profili pedologici, dai quali sono stati raccolti 233 campioni di suolo; nella seconda parte del rilevamento sono stati campionati 70 profili pedologici, dai quali sono stati raccolti 235 campioni di suolo.

Sui campioni di suolo prelevati in campo sono state eseguite analisi fisiche e chimiche di laboratorio per determinare i parametri necessari per la caratterizzazione dei suoli e la classificazione tassonomica.

Prima di effettuare le analisi di laboratorio, i campioni di suolo sono stati essiccati all’aria e (dopo frantumazione degli aggregati) setacciati a 2 mm.

Sulla terra fine di tutti i campioni sono state eseguite le seguenti analisi chimiche e fisiche: tessitura apparente a 5 frazioni, carbonati totali, pH in H<sub>2</sub>O. Il carbonio organico è stato analizzato solo per gli orizzonti genetici A e B (in totale 254 campioni). Su 120 campioni di suolo è stata determinata la capacità di scambio cationico. Inoltre, è stato determinato il quantitativo di ferro libero e quello di ferro totale su 55 campioni di suolo. Sono stati impiegati i metodi ufficiali di analisi del suolo descritti nel DM 185 del 1999 (Ministero Politiche Agricole e Forestali 2000).

La tessitura apparente è stata determinata per setacciatura e sedimentazione mediante il *Metodo della pipetta*. Il materiale è stato trattato in soluzione di acqua deionizzata e disperdente di sodio

esametafosfato e agitato per circa 8 ore. Sono state determinate 5 frazioni granulometriche: sabbia grossa (0,1-2 mm), sabbia fine (0,05-0,1 mm), limo grosso (0,02-0,05 mm), limo fine (0,002-0,02 mm), argilla (<0,002 mm). La sabbia grossa è stata separata mediante setacciatura, mentre per la determinazione delle altre frazioni è stato utilizzato il levigatore di Esenwein. Con le misure ricavate sono determinate le classi tessiturali (triangolo USDA (United States Department of Agriculture)).

La determinazione del calcare totale è stata eseguita con il metodo gasvolumetrico, mediante trattamento con acido cloridrico e impiegando il calcimetro di Dietrich-Fruehling.

Il carbonio organico è stato determinato mediante titolazione previa ossidazione della sostanza organica con bicromato di potassio in ambiente acido, secondo il metodo *Walkley & Black*. La misura del pH in H<sub>2</sub>O è stata effettuata per via potenziometrica, su una sospensione suolo-acqua in rapporto 1:2,5.

La CSC è stata determinata mediante titolazione con EDTA (acido etilendiamminotetraacetico) dopo trattamento del campione con cloruro di bario e trietanolammina a pH 8.1.

La determinazione del ferro totale è stata effettuata mediante digestione del campione di suolo in ambiente acido (HNO<sub>3</sub> al 65%, HCl al 37%, HF al 40%), secondo il metodo stabilito dalla US Environmental Protection Agency (EPA).

La determinazione del ferro libero è stata effettuata mediante il metodo Holmgren, con una soluzione di sodio ditionito e sodio citrato e successiva determinazione per spettrofotometria in assorbimento atomico (Flame Atomic Absorption Spectroscopy (FASS)).

Utilizzando le chiavi dei gruppi pedologici di riferimento (primo livello di classificazione) del WRB, nonché i prefissi (*principal qualifiers*) e i suffissi (*supplementary qualifiers*) (secondo livello di classificazione), i profili di suolo sono stati assegnati a 13 categorie tassonomiche (v. Tab 1).

**Tab. 7.1** Categorie tassonomiche assegnate ai profili campionati e sigla utilizzata nella mappa dei suoli.

<b>denominazione WRB</b>	<b>sigla</b>
<b>Leptic Vertisol (Calcaric)</b>	<b>LVC</b>
<b>Abruptic Luvisol (Loamic, Differentic)</b>	<b>ALLD</b>
<b>Abruptic Leptic LUVISOL (Loamic, Differentic)</b>	<b>ALLLD</b>
<b>Calcaric Cambisol (Loamic)</b>	<b>CCL</b>
<b>Leptic Calcaric Cambisol (Loamic)</b>	<b>LCCL</b>
<b>Leptic Calcaric Regosol</b>	<b>LCR</b>
<b>Dystric Cambisol (Loamic)</b>	<b>DCL</b>
<b>Leptic Skeletic Calcaric Regosol</b>	<b>LSCR</b>
<b>Leptic Dystric Cambisol (Loamic)</b>	<b>LDCM</b>
<b>Leptic Dystric Regosol</b>	<b>LDR</b>
<b>Leptic Calcaric Regosol (Loamic)</b>	<b>LCRL</b>
<b>Leptic, Calcaric Cambisol (Clayic)</b>	<b>LCCC</b>

Per quanto riguarda le categorie tipologiche, l'attribuzione è stata effettuata utilizzando tre chiavi discriminanti: le prime due contemplano le componenti sabbia e argilla, la terza tiene conto dello spessore del solum. Il criterio utilizzato dalla chiave primaria per suddividere gli oggetti è basato sul contenuto medio di argilla lungo il profilo: <25% (Ab), 25-45% (Am), >45% (Aa). La seconda

suddivisione è dicotomica ed è basata sul contenuto medio di sabbia: >50% (H), <50% (L). L'ulteriore suddivisione è basata sullo spessore del *solum*: <35 cm (suoli sottili - Ss), 35-60 cm (suoli moderatamente sottili - Sms), 61-100 cm (suoli moderatamente profondi - Smp), >100 cm (suoli profondi - Sp). Sono state riscontrate 12 categorie tipologiche (Tab. 2).

**Tab. 7.2** Categorie tipologiche di suolo assegnate secondo le caratteristiche tessiturali e dello spessore.

caratterisitche	sigla
<b>argilla&gt;45% (Aa), sabbia&lt;50% (L), solum 61-100 cm (Smp)</b>	<b>AaLSmp</b>
<b>argilla&gt;45% (Aa), sabbia&lt;50% (L), solum 35-60 cm (Sms)</b>	<b>AaLSms</b>
<b>argilla&lt;25% (Ab), sabbia&lt;50% (L), solum &gt;100 cm (Sp)</b>	<b>AbLSp</b>
<b>argilla&lt;25% (Ab), sabbia&gt;50% (H), solum&lt;35 cm (Ss)</b>	<b>AbHSs</b>
<b>argilla&lt;25% (Ab), sabbia&gt;50% (H), solum 35-60 cm (Sms)</b>	<b>AbHSms</b>
<b>argilla 25-45% (Am), sabbia &gt;50% (H), solum 61-100 (Smp)</b>	<b>AmHSmp</b>
<b>argilla 25-45% (Am), sabbia &gt;50% (H), solum &gt;100 cm (Sp)</b>	<b>AmHSp</b>
<b>argilla&lt;25% (Ab), sabbia&gt;50% (H), solum &gt;100 cm (Sp)</b>	<b>AbHSp</b>
<b>argilla 35-45% (Am), sabbia&lt;50% (L), solum &gt;100 cm (Sp)</b>	<b>AmLSp</b>
<b>argilla 25-45% (Am), sabbia&lt;50% (L), solum&lt;35 cm (Ss)</b>	<b>AmLSs</b>
<b>argilla 25-45% (Am), sabbia&lt;50% (L), solum 35-60 cm (Sms)</b>	<b>AmLSms</b>
<b>argilla 25-45% (Am), sabbia&lt;50% (L), solum 61-100 cm (Smp)</b>	<b>AmLSmp</b>

Per l'elaborazione delle mappe è stata utilizzata la DA; nella matrice dei costi sono stati mantenuti gli elementi fuori dalla diagonale tutti uguali a uno, mentre per la probabilità a priori è stata adottata l'opzione *uniform* (v. Cap. 6). La valutazione dei risultati è stata basata sui risultati della matrice di confusione (v. Cap. 6). La mappa relativa alle tipologie tassonomiche è stata realizzata utilizzando 83 punti del dataset, mentre la matrice di confusione è stata calcolata sui 127 considerati. Analogamente, la mappa relativa alle tipologie di suolo è stata prodotta utilizzando 82 punti e la matrice di confusione è stata calcolata sui 129 punti considerati. I punti utilizzati nel *training set* sono quelli relativi alla campagna di rilevamento effettuata secondo il criterio descritto nel capitolo 4 (profili del 2015), integrati con alcuni del 2014 per i casi in cui gli oggetti di una categoria non raggiungessero il numero minimo di quattro. Le elaborazioni hanno riguardato l'intera area di studio, con esclusione delle aree con spessore del *solum* inferiore a 5 cm, individuate sulla base della mappa dello spessore del *solum* (V. Cap. 5). Per la mappa relativa alle classi tassonomiche WRB sono state escluse le tipologie di suolo per le quali sono stati rilevati singoli casi: Dystric Cambisol (Loamic) - P28\_2015; Eutric Cambisol (Loamic) - P11\_2015; Leptic Dystric Regosol - P49\_2015; Leptic Dystric Cambisol (Loamic) - P35\_2015. Di conseguenza, nella mappa relativa alle classi tassonomiche secondo il WRB il territorio è suddiviso in sole nove classi.

### 7.3.1 Tipologie di suolo: principali caratteristiche e classificazione



**Fig. 7.1** Profilo di suolo rappresentativo della classe AaLSmp.

orizzonte Ap di spessore 30-35 cm, sovrapposto a uno o più orizzonti ove la presenza di slickensides e di crepacciature è più o meno evidente a seconda del contenuto di umidità; il venir meno di queste evidenze è indicativo dell'inizio dell'orizzonte C. In questi suoli può essere presente un orizzonte di transizione tra quello di alterazione (Bwss) e il sottostante C, nei quali è possibile trovare accumuli di carbonati. La struttura degli aggregati è fortemente sviluppata con forma poliedrica angolare o sub-angolare di dimensione media o grande. Lo scheletro è praticamente assente lungo tutto il profilo sino al raggiungimento dell'orizzonte C. I suoli appartenenti a questa categoria sono stati classificati secondo il WRB come Leptic Vertisol (Calcaric) (P1\_2015, P13\_2015, P26\_2015, P29\_2015, P30\_2015, P55\_2015, P56\_2015, P67\_2015, P7\_2014, P17\_2014, P18\_2014, P19\_2014, 29\_2014, 30\_2014, 31\_2014, 39\_2014, P61\_2014, P62\_2014, P63\_2014).

La classe **AaLSmp** (argilla >45% (Aa), sabbia <50% (L), solum 61-100 cm (Smp)) comprende suoli, tipicamente rilevati sulla formazione delle Breccie di Baiso, con uno spessore del solum compreso tra 70 e 100 cm (media dei profili campionati pari a 76 cm). Sono suoli che nell'area di studio sono generalmente occupati da colture erbacee, hanno una tessitura praticamente costante lungo il profilo e sono caratterizzati dalla preponderante presenza di argilla. Infatti, il quantitativo minimo di argilla è sempre maggiore del 45% sino a raggiungere quote che superano abbondantemente il 50%. I quantitativi di limo sono sempre importanti (maggiori del 25%) e la percentuale di sabbia ha una quota residuale variabile, che raggiunge al massimo il 25%. La tessitura è tipicamente argillosa e a volte franco-argillosa. Nei profili campionati è costante la presenza di carbonati lungo il profilo, con percentuali variabili tra 10% e 40%. Sono suoli tipicamente costituiti da un



**Fig. 7.2** Profilo di suolo rappresentativo della classe AaLSms.

WRB come Leptic Vertisol (Calcaric) (P15\_2015, P16\_2015, P\_62\_2015, P22\_2014, P25\_2014, P\_26\_2014, P\_27\_2014, P28\_2014).



**Fig. 7.3** Profilo di suolo rappresentativo della classe AbLSp.

La classe **AaLSms** (argilla > 45% (Aa), sabbia < 50% (L), solum 35-60 cm (Sms)) comprende suoli, tipicamente rilevati sulla formazione delle Breccie di Baiso, generalmente su aree agricole abbandonate, occupate da vegetazione erbacea. La tessitura è caratterizzata dalla dominante presenza di argilla, che può raggiungere quote pari al 56%. Il quantitativo di limo varia tra il 28 ed il 38%, la sabbia ha una quota residuale variabile con valori superiori al 10% sino a raggiungere il 23%; la tessitura è tipicamente argillosa. Il carbonato di calcio è presente con quote comprese tra il 20 e il 30% e a volte sono presenti arricchimenti localizzati. La struttura degli aggregati è, normalmente, fortemente sviluppata con forma poliedrica angolare o sub-angolare di dimensione media o grande. Lo scheletro, se presente, è di piccole dimensioni (20-75 mm) può essere a volte frequente e non è alterato. I suoli appartenenti a questa categoria sono stati classificati secondo il

WRB come Leptic Vertisol (Calcaric) (P15\_2015, P16\_2015, P\_62\_2015, P22\_2014, P25\_2014, P\_26\_2014, P\_27\_2014, P28\_2014).

La classe **AbLSp** (argilla < 25% (Ab), sabbia < 50% (L), solum > 100 cm (Sp)) comprende suoli rilevati sul fondovalle (sull'Unità di Nizza e Unità dell'Ardivestra) in aree occupate da colture erbacee. La tessitura tipicamente franca è costante lungo il profilo, con sabbia di poco superiore al 30%, limo variabile tra il 35 e il 50% e argilla intorno al 20%. La presenza di carbonato di calcio è costante lungo il profilo, con quote maggiori del 10%, sino a un massimo che supera di poco il 20%. Presentano un orizzonte superficiale Ap il cui spessore varia tra i 35 e i 45 cm, sovrapposto a uno orizzonte di alterazione Bw con successivi orizzonti di transizione al sottostante C. Lo scheletro è praticamente assente lungo tutto il profilo. I suoli appartenenti a questa categoria sono stati classificati secondo il WRB come Calcaric Cambisol (Loamic); appartengono a questa categoria i profili P37\_2015, P38\_2015, P68\_2015, P69\_2015.



**Fig. 7.4** Profilo di suolo rappresentativo della classe AbHSs.

Regosol (P23\_2015, P24\_2015, P65\_2015, P3\_2014), Leptic Skeletic Calcaric Regosol (P34\_2015, P39\_2015, P45\_2015, P46\_2015) e Leptic Dystric Regosol (P49\_2015).

I suoli della classe **AbHSs** (argilla < 25% (Ab), sabbia > 50% (H), solum < 35 cm (Ss)), tipicamente rilevati sulla Formazione delle Arenarie del Monte Vallassa, sono tra i meno sviluppati dell'area di studio. Presentano una tessitura da franco-sabbiosa a sabbioso-franca, con un quantitativo di sabbia sempre superiore al 60%, mentre l'argilla raramente supera il 10%. Sono presenti in aree boschive e sono tipicamente costituiti da un orizzonte A di pochi cm, sovrapposto a un AC o CA. La struttura degli aggregati è debolmente o moderatamente sviluppata. I suoli appartenenti a questa categoria sono stati classificati secondo il WRB come Leptic Calcaric



**Fig. 7.5** Profilo di suolo rappresentativo della classe AbHSms.

La classe **AbHSms** (argilla < 25% (Ab), sabbia > 50% (H), solum 35-60 cm (Sms)), tipicamente rilevati sulla Formazione delle Arenarie del Monte Vallassa, comprende suoli poco sviluppati che sono stati rilevati in aree boschive. Sono tipicamente costituiti da un sottile orizzonte A (mediante di 10 cm), sovrapposto a un orizzonte che può presentare segni di alterazione a volte evidenti (Bw) e che può inglobare ciottoli. La struttura degli aggregati è debolmente o al limite moderatamente sviluppata. Il carbonato di calcio è presente in quote comprese tra meno del 10% e il 20%. La tessitura, costante lungo il profilo, è tipicamente franco-sabbiosa (sabbia media 66%, limo 22%, argilla 12%). I suoli appartenenti a questa categoria sono stati classificati secondo il WRB come Leptic Calcaric Regosol (P64\_2015, P6\_2014), Leptic Calcaric Cambisol (Loamic) (P33\_2015, P36\_2015, P21\_2014) e Leptic Skeletic Calcaric Regosol (P47\_2015, P48\_2015).



**Fig. 7.6** Profilo di suolo rappresentativo della classe AmHSmp.

La classe **AmHSmp** (argilla 25-35% (Am), sabbia >50% (H), solum 61-100 (Smp)), tipicamente rilevati sulla Formazione delle Arenarie del Monte Vallassa, comprende suoli rilevati in aree boschive. Sono caratterizzati dall'assenza di carbonato di calcio e da una tessitura da franco-sabbiosa a franco-sabbioso-argillosa. In particolare, questi profili possiedono un orizzonte profondo che presenta un significativo incremento di argilla, rispetto agli orizzonti soprastanti, con una quota generalmente intorno al 30% e a volte superiore. Lo scheletro è assente lungo tutto il profilo. I suoli appartenenti a questa categoria sono stati classificati secondo il WRB come: Abruptic Leptic LUVISOL (Loamic, Differentic) P25\_2015, P60\_2015, P8\_2014, P9\_2014, P65\_2014, P66\_2014) e Leptic Dystric Cambisol (Loamic) (P35\_2015).



**Fig. 7.7** Profilo di suolo rappresentativo della classe AmHSp.

La classe **AmHSp** (argilla 25-35% (Am), sabbia >50% (H), solum >100 cm (Sp)), tipicamente rilevati sulla Formazione delle Arenarie del Monte Vallassa, comprende suoli con caratteristiche analoghe alla classe AmHSmp, ma uno spessore del solum maggiore di 100 cm. I suoli appartenenti a questa categoria sono stati classificati secondo il WRB come: Abruptic Luvisol (Loamic, Differentic) (P2\_2015, P5\_2015, P59\_2015, P64\_2014).



**Fig. 7.8** Profilo di suolo rappresentativo della classe AbHSp.

La classe **AbHSp** (argilla <25% (Ab), sabbia >50% (H), solum >100 cm (Sp)) comprende suoli che sono stati campionati in aree agricole a volte abbandonate. Sono caratterizzati da un contenuto di carbonato di calcio fra 3 e 20% e da una tessitura da franco-sabbiosa sino a franco-sabbioso-argillosa. I suoli appartenenti a questa categoria sono stati classificati secondo il WRB come: Calcaric Cambisol (Loamic) (P10\_2015, P44\_2015, P53\_2015, P12\_2014, P13\_2014) ed Eutric Cambisol (Loamic) (P11\_2015).



**Fig. 7.9** Profilo di suolo rappresentativo della classe AmLSp.

La classe **AmLSp** (argilla 25-45% (Am), sabbia <50% (L), solum >100 cm (Sp)) comprende suoli con uno spessore del solum compreso tra 110 e 140 cm, una tessitura costante franco-argillosa o franco-limoso-argillosa. Il contenuto di sabbia va dal 14 al 30%, quello di limo dal 35% al 50%, quello di argilla da poco meno del 30% fino a oltre il 40%. Il contenuto di carbonato di calcio lungo l'intero profilo varia da poco meno del 30% fino a oltre il 40%. La struttura degli aggregati è moderatamente sviluppata, con forma poliedrica angolare o sub-angolare e dimensione media o grande. Lo scheletro è praticamente assente lungo tutto il profilo. I suoli appartenenti a questa categoria sono stati rilevati in corrispondenza delle formazioni di Antognola e Contignaco e sono stati classificati secondo il WRB come: Calcaric Cambisol (Loamic) (P3\_2015, P4\_2015, P12\_2015, P14\_2015, P40\_2015, P52\_2015, P54\_2015).



**Fig. 7.10** Profilo di suolo rappresentativo della classe AmLSs.

meno del 10% a oltre il 40%. La struttura degli aggregati è debolmente o moderatamente sviluppata; la forma varia da poliedrica granulare media o grande a poliedrica sub-angolare media o grande. La presenza dello scheletro può variare, a seconda della formazione litologica, da comune ad abbondante e la dimensione da piccola (20-75 mm) a molto piccola (2-20 mm). I suoli appartenenti a questa categoria sono stati rilevati sulle formazioni di Contignaco, Antognola e Ranzano, sono stati classificati secondo il WRB come: Leptic Calcaric Regosol (P8\_2015, P20\_2015, P21\_2015, P22\_2015, P31\_2015, P42\_2015, P50\_2015, P57\_2015, P70\_2015, P4\_2014, P23\_2014, P24\_2014) e Leptic Calcaric Regosol (Loamic) (P55\_2014, P54\_2014).



**Fig. 7.11** Profilo di suolo rappresentativo della classe AmLSms.

La classe **AmLSs** (argilla 25-45% (Am), sabbia < 50% (L), solum < 35 cm (Ss)) comprende suoli sottili, rilevati in aree boschive ma anche su aree agricole e incolte. La tessitura è differente e correlata al diverso *parent material*. I quantitativi di sabbia possono essere di poco inferiori al 20% e giungere a superare di poco il 40%; il limo ha quote che variano tra il 20 ed il 50%, mentre l'argilla può raggiungere valori di poco superiori al 40%. La tessitura è tipicamente franco-argillosa, ma anche franco-limoso-argillosa o franca. E' sempre presente il carbonato di calcio, da

La classe **AmLSms** (argilla 25-45% (Am), sabbia < 50% (L), spessore solum 35-60 cm (Sms)) comprende suoli rilevati in aree boschive ma anche in terreni agricoli a coltivazioni erbacee. La tessitura è costante lungo il profilo ed è caratterizzata dalla preponderante presenza di limo (da meno del 40% a oltre il 60%); l'argilla è sempre maggiore del 30% e la sabbia ha una quota residuale, generalmente superiore al 10% (talvolta può raggiungere quote del 30%). La tessitura è tipicamente franco-argillosa ma anche franco-limoso-argillosa. Costante è la presenza di carbonati lungo il profilo, con percentuali tra il 20 e oltre il 40%. La struttura degli aggregati, moderatamente sviluppata nella parte più superficiale del profilo, diventa debolmente sviluppata all'approssimarsi dell'orizzonte C, con forma poliedrica sub-angolare media o grande. Lo scheletro è di dimensioni piccole (20-75 mm) o molto piccole (2-20 mm), in quantità scarsa, con

tendenza ad aumentare verso la parte più profonda del profilo ove può diventare comune. I suoli appartenenti a questa categoria sono stati rilevati sulle formazioni di Contignaco, Antognola e Ranzano, sono stati classificati secondo il WRB come: Leptic, Calcaric Cambisol (Loamic) (P7\_2015, P9\_2015, P17\_2015, P18\_2015, P41\_2015, P43\_2015, P51\_2015, P10\_2014, P14\_2014, P34\_2014, P36\_2014, P41\_3°\_2014, P44\_2014, P56\_2014), Leptic Calcaric Regosol (Loamic) (P19\_2015, P37\_2014, P40\_2014, P41\_2°\_2014, P42\_2014, P43\_2014, P58\_2014) e Leptic, Calcaric Cambisol (Clayic) (P45\_2014, P48\_2014, P52\_2014).



**Fig. 7.12** Profilo di suolo rappresentativo della classe AmLSmp.

La classe **AmLSmp** (argilla 25-45% (Am), sabbia < 50% (L), spessore solum 61-100 cm (Smp)) comprende suoli rilevati in aree boschive ma anche in aree agricole utilizzate e incolte. La tessitura è caratterizzata un quantitativo minimo di sabbia intorno al 30%, di limo variabile tra il 38 e il 46% e di argilla variabile tra il 20 ed il 35%. La tessitura è tipicamente franca, ma anche franco-argilloso-limoso, franco-argilloso o franco-limoso. Nei profili campionati è sempre presente carbonato di calcio (dal 10% fino al 30%). La struttura degli aggregati è moderatamente o fortemente sviluppata, con forma poliedrica sub-angolare media o grande. Lo scheletro è generalmente assente nella parte superiore del profilo, ma può essere comune o abbondante di dimensioni piccole (20-75 mm) o molto piccole (2-20 mm) nella parte inferiore. I suoli appartenenti a questa categoria sono stati rilevati sulle formazioni di Contignaco, Antognola e Ranzano, sono stati classificati

secondo il WRB come: Leptic Calcaric Cambisol (Loamic) (P6\_2015, P17\_2015, P27\_2015, P32\_2015, P11\_2014, P20\_2014, P32\_2014, P33\_2014, P35\_2014, P38\_2014, P46\_2014, P59\_2014, P60\_2014), Dystric Cambisol (Loamic) (P28\_2015), Leptic Calcaric Cambisol (Clayic) (P47\_2014, P49\_2014, P50\_2014, P51\_2014, P53\_2014) e Leptic Calcaric Regosol (Loamic) (P37\_2014).

## 7.4 Analisi e risultati

Come per le altre applicazioni della DA (v. Cap. 6) sono stati calcolati cinque modelli (con i metodi *linear*, *diagLinear*, *diagQuadratic*, *pseudolinear*, *pseudoQuadratic*) sia per la mappa *classi di suolo WRB* sia per quella di relativa alle *tipologie tessiturali e spessore*.

Per quanto riguarda la mappa delle categorie tassonomiche, esaminando i risultati delle matrici di confusione (Tab. 7.3, 7.5, 7.7, 7.9, 7.11), la migliore prestazione è stata ottenuta con il metodo

*diagQuadratic* (Non-Error-Rate pari a 58%) (Tab. 7.7), il cui risultato è rappresentato in figura 7.13. Questo ha i valori di *Misclassification risk*, relativamente a ciascuna classe (Tab. 7.8), generalmente contenuti (MR complessivo pari a 41%), dai quali emerge la classe LCCL (MR<sub>g</sub> pari a 14,96%) . I valori relativi a *Sensibility* e a *Specificity* sono mediamente elevati. Per quanto riguarda i valori di *Sensibility* la classe LCCL (Leptic, Calcaric Cambisol (Loamic)) è la peggiore (36%). In particolare, tale classe perde 2/3 dei suoi oggetti, tuttavia mantiene una alta *Specificity*, infatti vengono assegnati a questa classe pochi oggetti appartenenti ad altre classi. L'altra classe con una bassa *Sensibility* è la ALLLD (Abruptic Leptic Luvisol (Loamic, Differentic)) che perde due dei quattro elementi a favore della classe ALLD (Abruptic Luvisol (Loamic, Differentic)) e LCCC (Leptic, Calcaric CAMBISOL (Clayic)).

**Tab. 7.3** Matrice di confusione per le classi tassonomiche, relativa al metodo *Linear* (Non-Error-Rate (NER) Error-Rate (ER) e *Misclassification risk* (MR)).

		Linear										
		LVC'	LSCR'	LCRL'	LCR'	LCCL'	LCCC'	CCL'	ALLLD'	ALLD'	ng	
		LVC	19	0	0	2	1	3	2	0	0	27
NER (%)	54,33	LSCR	0	6	0	1	0	0	0	0	0	7
ER (%)	45,67	LCRL	3	0	2	2	1	1	0	0	0	9
MR (%)	45,67	LCR	1	2	1	12	2	1	1	0	0	20
		LCCL	9	2	1	2	8	3	4	0	1	30
		LCCC	4	0	0	0	0	5	0	0	0	9
		CCL	4	0	0	0	0	0	10	1	1	16
		ALLLD	0	0	0	0	1	0	0	2	1	4
		ALLD	0	0	0	0	0	0	0	0	5	5
		ng'	40	10	4	19	13	13	17	3	8	127

**Tab. 7.4** Valori della *Sensibility* (Sn), *Specificity* (Sp) e del *Misclassification risk* (Mr), per ciascuna classe tassonomica, relativi al metodo *Linear*.

Sn LVC (%)	70,4	Sp LVC (%)	47,5	Mr LVC (%)	6,30
Sn LSCR (%)	85,7	Sp LSCR (%)	60,0	Mr LSCR (%)	0,79
Sn LCRL (%)	22,2	Sp LCRL (%)	50,0	Mr LCRL (%)	5,51
Sn LCR (%)	60,0	Sp LCR (%)	63,2	Mr LCR (%)	6,30
Sn LCCL (%)	26,7	Sp LCCL (%)	61,5	Mr LCCL (%)	17,32
Sn LCCC (%)	55,6	Sp LCCC (%)	38,5	Mr LCCC (%)	3,15
Sn CCL (%)	62,5	Sp CCL (%)	58,8	Mr CCL (%)	4,72
Sn ALLLD (%)	50,0	Sp ALLLD (%)	66,7	Mr ALLLD (%)	1,57
Sn ALLD (%)	100,0	Sp ALLD (%)	62,5	Mr ALLD (%)	0,00

**Tab. 7.5** Matrice di confusione per le classi tassonomiche, relativa al metodo *diagLinear* (Non-Error-Rate (NER) Error-Rate (ER) e *Misclassification risk* (MR)).

		diag Linear										
		LVC'	LSCR'	LCRL'	LCR'	LCCL'	LCCC'	CCL'	ALLLD'	ALLD'	ng	
		LVC	12	0	1	2	1	5	6	0	0	27
NER (%)	48,03	LSCR	0	6	0	1	0	0	0	0	0	7
ER (%)	51,97	LCRL	3	0	1	4	1	0	0	0	0	9
MR (%)	51,97	LCR	2	4	0	10	2	1	1	0	0	20
		LCCL	5	1	1	6	13	2	1	0	1	30
		LCCC	3	0	0	0	2	4	0	0	0	9
		CCL	1	0	0	0	4	0	9	0	2	16
		ALLLD	0	0	0	0	1	0	1	2	0	4
		ALLD	0	1	0	0	0	0	0	0	4	5
		ng'	26	12	3	23	24	12	18	2	7	127

**Tab. 7.6** Valori della *Sensibility* (Sn), *Specificity* (Sp) e del *Misclassification risk* (Mr), per ciascuna classe tassonomica, relativi al metodo *diagLinear*.

Sn LVC (%)	44,4	Sp LVC (%)	46,2	Mr LVC (%)	11,81
Sn LSCR (%)	85,7	Sp LSCR (%)	50,0	Mr LSCR (%)	0,79
Sn LCRL (%)	11,1	Sp LCRL (%)	33,3	Mr LCRL (%)	6,30
Sn LCR (%)	50,0	Sp LCR (%)	43,5	Mr LCR (%)	7,87
Sn LCCL (%)	43,3	Sp LCCL (%)	54,2	Mr LCCL (%)	13,39
Sn LCCC (%)	44,4	Sp LCCC (%)	33,3	Mr LCCC (%)	3,94
Sn CCL (%)	56,3	Sp CCL (%)	50,0	Mr CCL (%)	5,51
Sn ALLLD (%)	50,0	Sp ALLLD (%)	100,0	Mr ALLLD (%)	1,57
Sn ALLD (%)	80,0	Sp ALLD (%)	57,1	Mr ALLD (%)	0,79

**Tab. 7.7** Matrice di confusione per le classi tassonomiche, relativa al metodo *diagQuadratic* (*Non-Error-Rate* (NER) *Error-Rate* (ER) e *Misclassification risk* (MR)).

		diag Quadratic									
		LVC'	LSCR'	LCRL'	LCR'	LCCL'	LCCC'	CCL'	ALLLD'	ALLD'	ng
	LVC	18	0	0	3	0	4	2	0	0	27
NER (%)	58,27	LSCR	0	7	0	0	0	0	0	0	7
ER (%)	41,73	LCRL	1	0	2	3	1	2	0	0	9
MR (%)	41,73	LCR	2	2	0	11	0	1	2	1	20
	LCCL	6	2	0	4	11	2	3	1	1	30
	LCCC	0	0	0	0	0	9	0	0	0	9
	CCL	2	0	0	0	1	0	10	0	3	16
	ALLLD	0	0	0	0	0	1	0	2	1	4
	ALLD	0	0	0	0	0	0	1	0	4	5
	ng'	29	11	2	21	13	19	18	4	10	127

**Tab. 7.8** Valori della *Sensibility* (Sn), *Specificity* (Sp) e del *Misclassification risk* (Mr), per ciascuna classe tassonomica, relativi al metodo *diagQuadratic*.

Sn LVC (%)	66,7	Sp LVC (%)	62,1	Mr LVC (%)	7,09
Sn LSCR (%)	100,0	Sp LSCR (%)	63,6	Mr LSCR (%)	0,00
Sn LCRL (%)	22,2	Sp LCRL (%)	100,0	Mr LCRL (%)	5,51
Sn LCR (%)	55,0	Sp LCR (%)	52,4	Mr LCR (%)	7,09
Sn LCCL (%)	36,7	Sp LCCL (%)	84,6	Mr LCCL (%)	14,96
Sn LCCC (%)	100,0	Sp LCCC (%)	47,4	Mr LCCC (%)	0,00
Sn CCL (%)	62,5	Sp CCL (%)	55,6	Mr CCL (%)	4,72
Sn ALLLD (%)	50,0	Sp ALLLD (%)	50,0	Mr ALLLD (%)	1,57
Sn ALLD (%)	80,0	Sp ALLD (%)	40,0	Mr ALLD (%)	0,79

**Tab. 7.9** Matrice di confusione per le classi tassonomiche, relativa al metodo *pseudoLinear* (*Non-Error-Rate* (NER) *Error-Rate* (ER) e *Misclassification risk* (MR)).

		pseudo Linear									
		LVC'	LSCR'	LCRL'	LCR'	LCCL'	LCCC'	CCL'	ALLLD'	ALLD'	ng
	LVC	19	0	0	2	1	3	2	0	0	27
NER (%)	54,33	LSCR	0	6	0	1	0	0	0	0	7
ER (%)	45,67	LCRL	3	0	2	2	1	1	0	0	9
MR (%)	45,67	LCR	1	2	1	12	2	1	1	0	20
	LCCL	9	2	1	2	8	3	4	0	1	30
	LCCC	4	0	0	0	0	5	0	0	0	9
	CCL	4	0	0	0	0	0	10	1	1	16
	ALLLD	0	0	0	0	1	0	0	2	1	4
	ALLD	0	0	0	0	0	0	0	0	5	5
	ng'	40	10	4	19	13	13	17	3	8	127

**Tab. 7.10** Valori della *Sensibility* (Sn), *Specificity* (Sp) e del *Misclassification risk* (Mr), per ciascuna classe tassonomica, relativi al metodo *pseudoLinear*.

Sn LVC (%)	70,4	Sp LVC (%)	47,5	Mr LVC (%)	6,30
Sn LSCR (%)	85,7	Sp LSCR (%)	60,0	Mr LSCR (%)	0,79
Sn LCRL (%)	22,2	Sp LCRL (%)	50,0	Mr LCRL (%)	5,51
Sn LCR (%)	60,0	Sp LCR (%)	63,2	Mr LCR (%)	6,30
Sn LCCL (%)	26,7	Sp LCCL (%)	61,5	Mr LCCL (%)	17,32
Sn LCCC (%)	55,6	Sp LCCC (%)	38,5	Mr LCCC (%)	3,15
Sn CCL (%)	62,5	Sp CCL (%)	58,8	Mr CCL (%)	4,72
Sn ALLLD (%)	50,0	Sp ALLLD (%)	66,7	Mr ALLLD (%)	1,57
Sn ALLD (%)	100,0	Sp ALLD (%)	62,5	Mr ALLD (%)	0,00

**Tab. 7.11** Matrice di confusione per le classi tassonomiche, relativa al metodo *pseudoQuadratic* (*Non-Error-Rate* (NER) *Error-Rate* (ER) e *Misclassification risk* (MR)).

		pseudo Quadratic									
		LVC'	LSCR'	LCRL'	LCR'	LCCL'	LCCC'	CCL'	ALLD'	ALLD'	ng
	LVC	19	0	2	0	0	0	1	0	5	27
NER (%)	37,80	LSCR	0	7	0	0	0	0	0	0	7
ER (%)	62,20	LCRL	2	0	5	0	0	0	0	2	9
MR (%)	62,20	LCR	0	1	16	0	0	0	0	3	20
	LCCL	0	0	13	1	0	0	1	0	15	30
	LCCC	2	0	1	0	0	6	0	0	0	9
	CCL	1	0	1	0	0	0	10	0	4	16
	ALLLD	0	0	1	0	0	0	0	0	3	4
	ALLD	0	0	4	0	0	0	0	0	1	5
	ng'	24	8	43	1	0	6	12	0	33	127

**Tab. 7.12** Valori della *Sensibility* (Sn), *Specificity* (Sp) e del *Misclassification risk* (Mr), per ciascuna classe tassonomica, relativi al metodo *pseudoQuadratic*.

Sn LVC (%)	70,4	Sp LVC (%)	79,2	Mr LVC (%)	6,30
Sn LSCR (%)	100,0	Sp LSCR (%)	87,5	Mr LSCR (%)	0,00
Sn LCRL (%)	55,6	Sp LCRL (%)	11,6	Mr LCRL (%)	3,15
Sn LCR (%)	0,0	Sp LCR (%)	0,0	Mr LCR (%)	15,75
Sn LCCL (%)	0,0	Sp LCCL (%)	#DIV/0!	Mr LCCL (%)	23,62
Sn LCCC (%)	66,7	Sp LCCC (%)	100,0	Mr LCCC (%)	2,36
Sn CCL (%)	62,5	Sp CCL (%)	83,3	Mr CCL (%)	4,72
Sn ALLLD (%)	0,0	Sp ALLLD (%)	#DIV/0!	Mr ALLLD (%)	3,15
Sn ALLD (%)	20,0	Sp ALLD (%)	3,0	Mr ALLD (%)	3,15

Per quanto riguarda la mappa delle tipologie di suolo (basata sulla tessitura e lo spessore del *solum*), secondo i parametri calcolati con le matrici di confusione, il migliore risultato, anche in questo caso, è stato ottenuto con il metodo *diagQuadratic* (*Non-Error-Rate* 55%) (Tab. 7.13) il cui risultato è rappresentato in figura 7.14. I suoi valori di *Misclassification risk* relativi a ciascuna classe risultano spesso alti (Tab. 7.14), in particolare per le classi AbHSms (71%) e AmbLSmp (61%) e AmLSmp (53%). Queste sono le classi che presentano i tassi di *Sensibility* più bassi (rispettivamente 28, 35, 32%). Per quanto riguarda la *Sensibility* i valori sono mediamente alti ed il caso peggiore è la classe AmLSp (24%).

**Tab. 7.13** Matrice di confusione per le tipologie di suolo, relativa al metodo *diagQuadratic* (Non-Error-Rate (NER) Error-Rate (ER) e Misclassification risk (MR)).

		diag Quadratic												
		AbHSs'	AmLSs'	AbHSms'	AmLSmp'	AmLSms'	AmHSmp'	AbHSp'	AmHSp'	AmLSp'	AbLSp'	AaLSms'	AaLSmp'	ng
		AbHSs	9	0	0	0	0	0	0	0	0	0	0	9
NER (%)	55,04	AmLSs	1	8	0	2	1	1	0	2	0	2	0	17
ER (%)	44,96	AbHSms	2	0	2	0	0	0	3	0	0	0	0	7
MR (%)	405,75	AmLSmp	0	0	1	7	3	0	0	5	0	3	1	20
		AmLSms	0	2	1	5	7	1	0	1	0	1	4	22
		AmHSmp	0	0	1	0	1	4	0	1	0	0	0	7
		AbHSp	0	0	1	0	0	4	0	0	0	0	0	5
		AmHSp	0	0	0	0	0	0	3	1	0	0	0	4
		AmLSp	0	0	0	0	1	0	0	4	0	0	2	7
		AbLSp	0	0	0	0	0	0	0	0	4	0	0	4
		AaLSms	0	0	0	1	0	0	0	2	0	5	0	8
		AaLSmp	0	0	0	3	0	1	0	1	0	0	14	19
		ng'	12	10	6	15	16	6	5	6	17	4	11	129

**Tab. 7.14** Valori della Sensibility (Sn), Specificity (Sp) e del Misclassification risk (Mr), per le tipologie di suolo, relativi al metodo *diagQuadratic*.

Sn AbHSs (%)	100,0	Sp AbHSs (%)	80,0	Mr AbHSs (%)	0,00
Sn AmLSs (%)	47,1	Sp AmLSs (%)	66,7	Mr AmLSs (%)	52,94
Sn AbHSms (%)	28,6	Sp AbHSms (%)	40,0	Mr AbHSms (%)	71,43
Sn AmLSmp (%)	35,0	Sp AmLSmp (%)	44,4	Mr AmLSmp (%)	60,78
Sn AmLSms (%)	31,8	Sp AmLSms (%)	44,4	Mr AmLSms (%)	53,10
Sn AmHSmp (%)	57,1	Sp AmHSmp (%)	66,7	Mr AmHSmp (%)	42,86
Sn AbHSp (%)	80,0	Sp AbHSp (%)	57,1	Mr AbHSp (%)	20,00
Sn AmHSp (%)	75,0	Sp AmHSp (%)	42,9	Mr AmHSp (%)	25,00
Sn AmLSp (%)	57,1	Sp AmLSp (%)	18,8	Mr AmLSp (%)	15,84
Sn AbLSp (%)	100,0	Sp AbLSp (%)	100,0	Mr AbLSp (%)	0,00
Sn AaLSms (%)	62,5	Sp AaLSms (%)	30,8	Mr AaLSms (%)	37,50
Sn AaLSmp (%)	73,7	Sp AaLSmp (%)	63,6	Mr AaLSmp (%)	26,32

**Tab. 7.15** Matrice di confusione per le tipologie di suolo, relativa al metodo *Linear* (Non-Error-Rate (NER) Error-Rate (ER) e Misclassification risk (MR)).

		Linear												
		AbHSs'	AmLSs'	AbHSms'	AmLSmp'	AmLSms'	AmHSmp'	AbHSp'	AmHSp'	AmLSp'	AbLSp'	AaLSms'	AaLSmp'	ng
		AbHSs	8	0	1	0	0	0	0	0	0	0	0	9
NER (%)	51,16	AmLSs	1	8	0	0	2	0	0	2	0	3	1	17
ER (%)	48,84	AbHSms	1	0	2	0	0	1	0	3	0	0	0	7
MR (%)	447,64	AmLSmp	0	1	1	8	2	0	1	0	4	0	3	20
		AmLSms	0	3	0	4	4	1	0	2	0	3	5	22
		AmHSmp	0	0	1	2	0	4	0	0	0	0	0	7
		AbHSp	0	0	0	0	0	0	4	0	1	0	0	5
		AmHSp	0	0	0	0	0	0	3	1	0	0	0	4
		AmLSp	0	0	0	1	0	0	1	3	0	0	2	7
		AbLSp	0	0	0	0	0	0	0	0	4	0	0	4
		AaLSms	0	0	0	1	0	0	0	3	0	4	0	8
		AaLSmp	0	0	0	2	1	0	2	0	0	0	14	19
		ng'	10	12	5	18	9	6	7	7	16	4	13	129

**Tab. 7.16** Valori della Sensibility (Sn), Specificity (Sp) e del Misclassification risk (Mr), per le tipologie di suolo, relativi al metodo *Linear*.

Sn AbHSs (%)	88,9	Sp AbHSs (%)	80,0	Mr AbHSs (%)	11,11
Sn AmLSs (%)	47,1	Sp AmLSs (%)	66,7	Mr AmLSs (%)	47,83
Sn AbHSms (%)	28,6	Sp AbHSms (%)	40,0	Mr AbHSms (%)	71,43
Sn AmLSmp (%)	40,0	Sp AmLSmp (%)	44,4	Mr AmLSmp (%)	60,00
Sn AmLSms (%)	18,2	Sp AmLSms (%)	44,4	Mr AmLSms (%)	62,97
Sn AmHSmp (%)	57,1	Sp AmHSmp (%)	66,7	Mr AmHSmp (%)	42,86
Sn AbHSp (%)	80,0	Sp AbHSp (%)	57,1	Mr AbHSp (%)	20,00
Sn AmHSp (%)	75,0	Sp AmHSp (%)	42,9	Mr AmHSp (%)	25,00
Sn AmLSp (%)	42,9	Sp AmLSp (%)	18,8	Mr AmLSp (%)	30,12
Sn AbLSp (%)	100,0	Sp AbLSp (%)	100,0	Mr AbLSp (%)	0,00
Sn AaLSms (%)	50,0	Sp AaLSms (%)	30,8	Mr AaLSms (%)	50,00
Sn AaLSmp (%)	73,7	Sp AaLSmp (%)	63,6	Mr AaLSmp (%)	26,32

**Tab. 7.17** Matrice di confusione per le tipologie di suolo, relativa al metodo *diagLinear* (Non-Error-Rate (NER) Error-Rate (ER) e *Misclassification risk* (MR)).

		diag Linear												
		AbHSs'	AmLSs'	AbHSms'	AmLSmp'	AmLSms'	AmHSmp'	AbHSp'	AmHSp'	AmLSp'	AbLSp'	AaLSms'	AaLSmp'	ng
		AbHSs	6	0	0	0	0	0	1	0	0	0	0	9
NER (%)	49,61	AmLSs	1	9	1	0	2	0	0	1	1	1	1	17
ER (%)	50,39	AbHSms	1	0	4	0	0	1	0	1	0	0	0	7
MR (%)	501,42	AmLSmp	0	3	1	6	2	0	0	3	0	3	2	20
		AmLSms	0	4	0	4	6	0	0	1	0	3	4	22
		AmHSmp	1	0	1	1	0	3	0	1	0	0	0	7
		AbHSp	0	0	0	0	0	4	0	1	0	0	0	5
		AmHSp	0	0	1	0	0	0	1	2	0	0	0	4
		AmLSp	0	0	0	2	0	0	1	4	0	0	0	7
		AbLSp	0	0	0	0	0	0	0	0	4	0	0	4
		AaLSms	0	1	0	0	0	0	0	4	0	3	0	8
		AaLSmp	0	0	0	2	0	0	1	0	2	0	13	19
		ng'	9	17	8	15	10	6	6	5	17	5	20	129

**Tab. 7.18** Valori della *Sensibility* (Sn), *Specificity* (Sp) e del *Misclassification risk* (Mr), per le tipologie di suolo, relativi al metodo *diagLinear*.

Sn AbHSs (%)	66,7	Sp AbHSs (%)	80,0	Mr AbHSs (%)	33,33
Sn AmLSs (%)	52,9	Sp AmLSs (%)	66,7	Mr AmLSs (%)	41,95
Sn AbHSms (%)	57,1	Sp AbHSms (%)	40,0	Mr AbHSms (%)	42,86
Sn AmLSmp (%)	30,0	Sp AmLSmp (%)	44,4	Mr AmLSmp (%)	61,55
Sn AmLSms (%)	27,3	Sp AmLSms (%)	44,4	Mr AmLSms (%)	57,65
Sn AmHSmp (%)	42,9	Sp AmHSmp (%)	66,7	Mr AmHSmp (%)	57,14
Sn AbHSp (%)	80,0	Sp AbHSp (%)	57,1	Mr AbHSp (%)	20,00
Sn AmHSp (%)	50,0	Sp AmHSp (%)	42,9	Mr AmHSp (%)	50,00
Sn AmLSp (%)	57,1	Sp AmLSp (%)	18,8	Mr AmLSp (%)	42,86
Sn AbLSp (%)	100,0	Sp AbLSp (%)	100,0	Mr AbLSp (%)	0,00
Sn AaLSms (%)	37,5	Sp AaLSms (%)	30,8	Mr AaLSms (%)	62,50
Sn AaLSmp (%)	68,4	Sp AaLSmp (%)	63,6	Mr AaLSmp (%)	31,58

**Tab. 7.19** Matrice di confusione per le tipologie di suolo, relativa al metodo *pseudoLinear* (Non-Error-Rate (NER) Error-Rate (ER) e *Misclassification risk* (MR)).

		pseudo Linear												
		AbHSs'	AmLSs'	AbHSms'	AmLSmp'	AmLSms'	AmHSmp'	AbHSp'	AmHSp'	AmLSp'	AbLSp'	AaLSms'	AaLSmp'	ng
		AbHSs	8	0	1	0	0	0	0	0	0	0	0	9
NER (%)	51,16	AmLSs	1	8	0	0	2	0	0	2	0	3	1	17
ER (%)	48,84	AbHSms	1	0	2	0	0	1	0	3	0	0	0	7
MR (%)	447,64	AmLSmp	0	1	1	8	2	0	1	0	4	0	3	20
		AmLSms	0	3	0	4	4	1	0	0	2	0	3	22
		AmHSmp	0	0	1	2	0	4	0	0	0	0	0	7
		AbHSp	0	0	0	0	0	4	0	1	0	0	0	5
		AmHSp	0	0	0	0	0	0	3	1	0	0	0	4
		AmLSp	0	0	0	1	0	0	1	3	0	0	2	7
		AbLSp	0	0	0	0	0	0	0	0	4	0	0	4
		AaLSms	0	0	0	1	0	0	0	3	0	4	0	8
		AaLSmp	0	0	0	2	1	0	2	0	0	0	14	19
		ng'	10	12	5	18	9	6	7	7	16	4	22	129

**Tab. 7.20** Valori della *Sensibility* (Sn), *Specificity* (Sp) e del *Misclassification risk* (Mr), per le tipologie di suolo, relativi al metodo *pseudoLinear*.

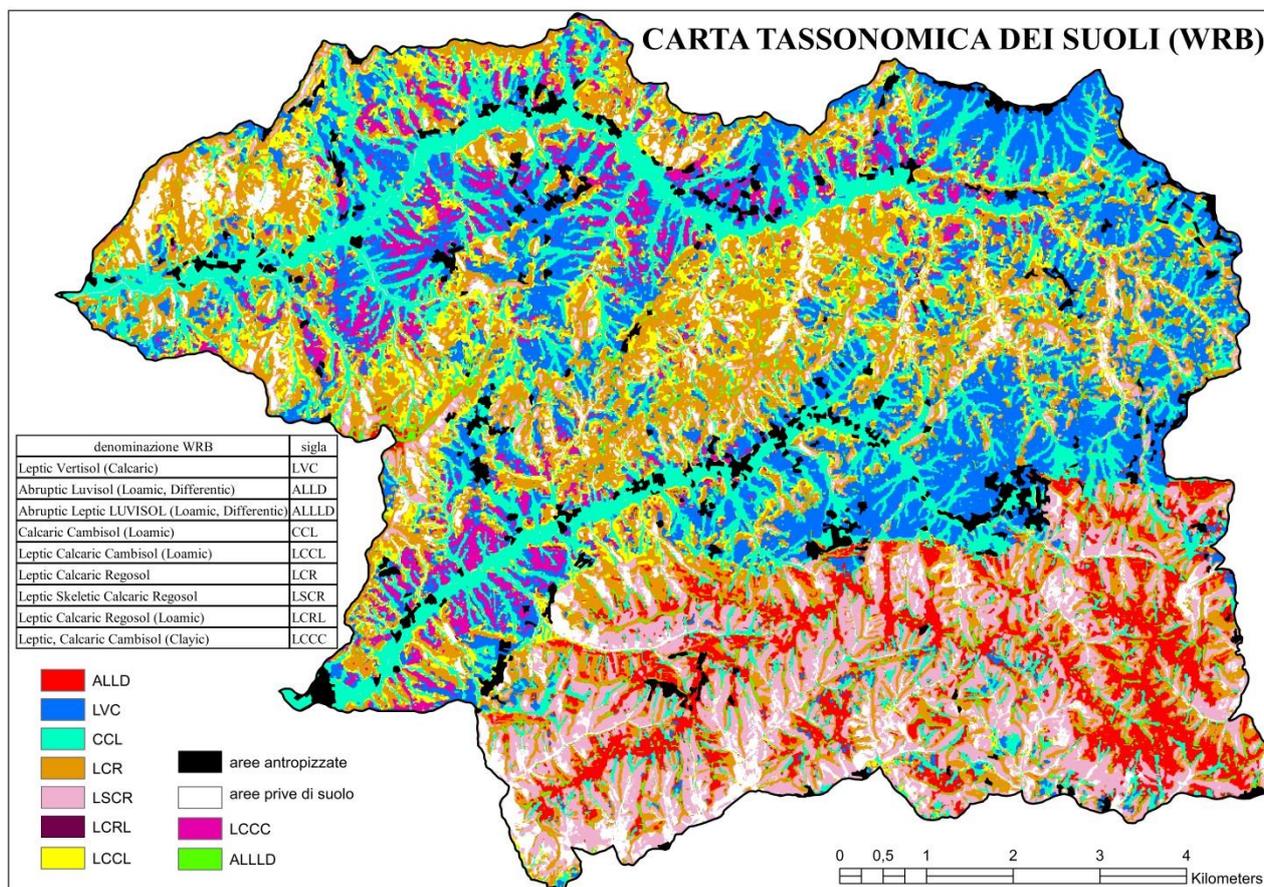
Sn AbHSs (%)	88,9	Sp AbHSs (%)	80,0	Mr AbHSs (%)	11,11
Sn AmLSs (%)	47,1	Sp AmLSs (%)	66,7	Mr AmLSs (%)	47,83
Sn AbHSms (%)	28,6	Sp AbHSms (%)	40,0	Mr AbHSms (%)	71,43
Sn AmLSmp (%)	40,0	Sp AmLSmp (%)	44,4	Mr AmLSmp (%)	60,00
Sn AmLSms (%)	18,2	Sp AmLSms (%)	44,4	Mr AmLSms (%)	62,97
Sn AmHSmp (%)	57,1	Sp AmHSmp (%)	66,7	Mr AmHSmp (%)	42,86
Sn AbHSp (%)	80,0	Sp AbHSp (%)	57,1	Mr AbHSp (%)	20,00
Sn AmHSp (%)	75,0	Sp AmHSp (%)	42,9	Mr AmHSp (%)	25,00
Sn AmLSp (%)	42,9	Sp AmLSp (%)	18,8	Mr AmLSp (%)	30,12
Sn AbLSp (%)	100,0	Sp AbLSp (%)	100,0	Mr AbLSp (%)	0,00
Sn AaLSms (%)	50,0	Sp AaLSms (%)	30,8	Mr AaLSms (%)	50,00
Sn AaLSmp (%)	73,7	Sp AaLSmp (%)	63,6	Mr AaLSmp (%)	26,32

**Tab. 7.21** Matrice di confusione per le tipologie di suolo, relativa al metodo *pseudoQuadratic* (Non-Error-Rate (NER) Error-Rate (ER) e Misclassification risk (MR)).

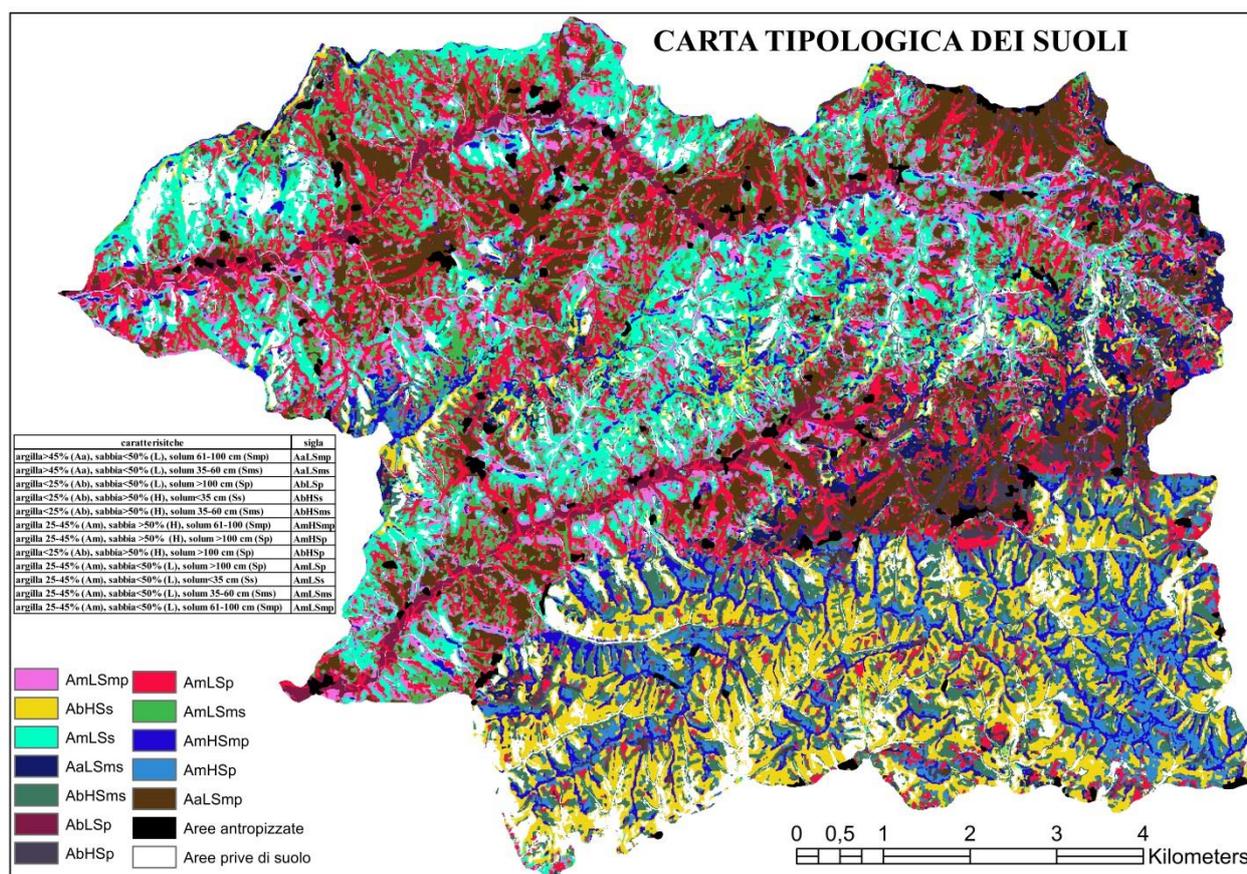
		pseudo Quadratic												
		AbHSs'	AmLSs'	AbHSms'	AmLSmp'	AmLSms'	AmHSmp'	AbHSp'	AmHSp'	AmLSp'	AbLSp'	AaLSms'	AaLSmp'	ng
	AbHSs	5	0	0	0	0	0	1	0	0	0	3	0	9
NER (%)	37,98	AmLSs	0	0	0	1	0	5	2	1	0	8	0	17
ER (%)	62,02	AbHSms	1	0	0	0	0	1	3	0	0	2	0	7
MR (%)	659,10	AmLSmp	0	0	0	9	0	4	1	1	0	5	0	20
	AmLSms	0	0	0	0	8	0	2	1	0	0	8	3	22
	AmHSmp	0	0	0	0	0	0	5	1	0	0	1	0	7
	AbHSp	0	0	0	0	0	0	4	0	0	0	1	0	5
	AmHSp	0	0	0	0	0	0	2	1	0	0	1	0	4
	AmLSp	0	0	0	0	0	0	2	0	1	0	2	2	7
	AbLSp	0	0	0	0	0	0	0	0	0	4	0	0	4
	AaLSms	0	0	0	0	0	0	2	0	0	0	3	3	8
	AaLSmp	0	0	0	0	0	0	4	0	0	0	1	14	19
	ng'	6	0	0	9	9	0	32	9	3	4	35	22	129

**Tab. 7.22** Valori della Sensibility (Sn), Specificity (Sp) e del Misclassification risk (Mr), per le tipologie di suolo, relativi al metodo *pseudoQuadratic*.

Sn AbHSs (%)	55,6	Sp AbHSs (%)	80,0	Mr AbHSs (%)	44,44
Sn AmLSs (%)	0,0	Sp AmLSs (%)	66,7	Mr AmLSs (%)	100,00
Sn AbHSms (%)	0,0	Sp AbHSms (%)	40,0	Mr AbHSms (%)	100,00
Sn AmLSmp (%)	45,0	Sp AmLSmp (%)	44,4	Mr AmLSmp (%)	55,00
Sn AmLSms (%)	36,4	Sp AmLSms (%)	44,4	Mr AmLSms (%)	52,33
Sn AmHSmp (%)	0,0	Sp AmHSmp (%)	66,7	Mr AmHSmp (%)	100,00
Sn AbHSp (%)	80,0	Sp AbHSp (%)	57,1	Mr AbHSp (%)	20,00
Sn AmHSp (%)	25,0	Sp AmHSp (%)	42,9	Mr AmHSp (%)	75,00
Sn AmLSp (%)	14,3	Sp AmLSp (%)	18,8	Mr AmLSp (%)	58,69
Sn AbLSp (%)	100,0	Sp AbLSp (%)	100,0	Mr AbLSp (%)	0,00
Sn AaLSms (%)	37,5	Sp AaLSms (%)	30,8	Mr AaLSms (%)	27,33
Sn AaLSmp (%)	73,7	Sp AaLSmp (%)	63,6	Mr AaLSmp (%)	26,32



**Fig. 7.13** Mappa dei suoli secondo la tassonomia WRB, ottenuta con il metodo *diagQuadratic*.



**Fig. 7.14** Mappa delle tipologie di suolo (basata sulla tessitura e lo spessore del *solum*), ottenuta con il metodo *diagQuadratic*.

## 7.5 Discussione e conclusioni

Indipendentemente dagli indici di valutazione calcolati, riguardo le elaborazioni presentate occorre tenere in considerazione alcuni importanti aspetti, utili per comprendere la loro affidabilità e di conseguenza la capacità di rappresentare la realtà.

Le mappe presentate sono state prodotte utilizzando un numero di punti di campionamento che è limitato per le elaborazioni statistiche eseguite (82 punti per la mappa WRB e 83 per quella delle tipologie di suolo). Tale argomentazione assume un particolare peso se si considera il ridotto numero di oggetti a disposizione per alcune categorie, sia tassonomiche che tipologiche. Questa limitazione ha costretto a modellare le mappe utilizzando tutti i punti a disposizione per le categorie costituite da solo quattro oggetti (es: la classe Abruptic Luvisol (Loamic, Differentic) corrispondente alla classe tipologica AmHSp). Modellare con quattro oggetti significa avere una bassa rappresentatività statistica e, di fatto, l'impossibilità di procedere con una cross validazione. Anche se questa è una situazione che si può verificare con le campagne di rilevamento pedologico su vaste aree, sarebbe auspicabile per i processi di inferenza procedere con un numero di oggetti superiore a quattro per ciascuna delle classi che si vogliono rappresentare in mappa.

Inoltre, occorre considerare che la suddivisione degli oggetti in categorie rigide costringe ad assegnare ciascun oggetto ad una classe anche quando gli oggetti presentano caratteristiche

intermedie tra due classi come può capitare nel caso dei suoli,. In questo caso, se non si introduce una ulteriore categoria si è costretti a forzare l'assegnazione a una delle classi prescelte. Sia per il caso delle categorie tassonomiche che per le tipologie di suolo tali assegnazioni spesso vengono effettuate sulla base di valori al limite delle categorie (es: un Leptic Calcaric Regosol (Loamic) può differire da un Leptic Calcaric Regosol per una differenza di qualche centimetro dello spessore di un orizzonte). Riguardo gli errori calcolati occorre evidenziare che mentre alcuni sono trascurabili in quanto la confusione avviene tra categorie simili (es: la categoria ALLLD cede un oggetto alla categoria ALLD, oppure la classe LCCL cede tre oggetti alla classe CCL: per questi casi ciò che distingue i due oggetti può essere la differente posizione dell'orizzonte diagnostico rispetto al substrato, che può variare anche solo di pochi centimetri), altri hanno un peso differente da un punto di vista pedologico (es: la classe LVC cede oggetti alla classe LCR e dalla stesse ne riceve, oltre ad avere diversi oggetti da più classi: in questi casi le tipologie tassonomiche coinvolte non sono fra loro sovrapponibili, neppure in parte). Analogamente, riguardo la mappa delle tipologie dei suoli, alla classe AmLSp costituita da 7 oggetti vengono assegnati 17 altri oggetti, alcuni dei quali totalmente incompatibili AmLSs (suoli profondi - suoli sottili). Le due mappe prodotte possono dunque essere considerata una prima approssimazione utile per l'interpretazione del territorio; esse forniscono sia una informazione che permette una interpretazione genetico-evolutiva dei suoli presenti nell'area sia una indicazione di carattere tecnico-applicativo per esempio in ambito agronomico.

## Bibliografia

Baruck, Jasmin; Nestroy, Othmar; Sartori, Giacomo; Baize, Denis; Traidl, Robert; Vrščaj, Borut et al. (2016): Soil classification and mapping in the Alps. The current state and future challenges. In: *Geoderma*, 264, pp. 312–331. DOI: 10.1016/j.geoderma.2015.08.005.

Brevik, Eric C.; Calzolari, Costanza; Miller, Bradley A.; Pereira, Paulo; Kabala, Cezary; Baumgarten, Andreas; Jordán, Antonio (2016): Soil mapping, classification, and pedologic modeling. History and future directions. In: *Geoderma*, 264, pp. 256–274. DOI: 10.1016/j.geoderma.2015.05.017.

Chagas, César da Silva; Vieira, Carlos Antônio Oliveira; Fernandes Filho, Elpídio Inácio (2013): Comparison between artificial neural networks and maximum likelihood classification in digital soil mapping. In: *Revista Brasileira de Ciência do Solo* n. 2, 37, pp. 339–351.

De Wrachien D., Previtali F, Calzi L., Semenza C. (1981) Lineamenti Geologici-Agrari del bacino del torrente Nizza (Oltrepo Pavese).

Florinsky, Igor (2016): Digital terrain analysis in soil science and geology: Academic Press.

Florinsky, Igor V. (2012): Digital Terrain Analysis in Soil Science and Geology: Elsevier.

Hengl, Tomislav; Toomanian, Norair; Reuter, Hannes I.; Malakouti, Mohammad J. (2007): Methods to interpolate soil categorical variables from profile observations. Lessons from Iran. In: *Geoderma* n. 4, 140, pp. 417–427. DOI: 10.1016/j.geoderma.2007.04.022.

Hitziger, Martin; Ließ, Mareike (2014): Comparison of Three Supervised Learning Methods for Digital Soil Mapping. Application to a Complex Terrain in the Ecuadorian Andes. In: *Applied and Environmental Soil Science* n. 1, 2014, pp. 1–12. DOI: 10.1155/2014/809495.

- Italiana, Repubblica (1999): Ministero delle politiche agricole alimentari e forestali. Dm 185 del 13 settembre 1999 Approvazione dei “Metodi ufficiali di analisi chimica del suolo”. In: Gazzetta Ufficiale della Repubblica Italiana–suppl. ord. no. 248 del 21 ottobre 1999–Serie generale.
- IUSS Working Group (2014): World reference base for soil resources 2014 international soil classification system for naming soils and creating legends for soil maps. In: *FAO, Rome*.
- Lagacherie, Philippe (2008): Digital soil mapping: a state of the art. In: Digital soil mapping with limited data: Springer, pp. 3–14.
- Li, Weidong; Zhang, Chuanrong; Dey, Dipak K.; Willig, Michael R. (2013): Updating categorical soil maps using limited survey data by Bayesian Markov chain cosimulation. In: *TheScientificWorldJournal*, 2013, p. 587284. DOI: 10.1155/2013/587284.
- Maynard, J. J.; Johnson, M. G. (2014): Scale-dependency of LiDAR derived terrain attributes in quantitative soil-landscape modeling. Effects of grid resolution vs. neighborhood extent. In: *Geoderma*, 230-231, pp. 29–40. DOI: 10.1016/j.geoderma.2014.03.021.
- Ministero Politiche Agricole e Forestali (2000): Metodi di Analisi chimiche del suolo Franco Angeli Editore Milano.
- Taghizadeh-Mehrjardi, R.; Minasny, B.; McBratney, A. B.; Triantafyllis, J.; Sarmadian, F.; Toomanian, N. (2012): Digital soil mapping of soil classes using decision trees in central Iran.
- White, Robert E. (2013): Principles and practice of soil science: the soil as a natural resource: John Wiley & Sons.

## Capitolo 8

### Conclusioni generali

L'obiettivo principale del progetto di dottorato è stato quello di produrre una base informativa cartografica, relativa agli aspetti principali del suolo, per un'area appenninica. L'attività si è svolta nelle condizioni che tipicamente si verificano nell'ambito della mappatura dei suoli su aree vaste, per la quale il fattore economico è quello che maggiormente condiziona la qualità del risultato finale. Pertanto, il lavoro svolto è stato incentrato sull'applicazione di metodologie che potessero fornire prodotti cartografici con un alto livello di dettaglio e un basso costo di esecuzione, nonché sulla verifica circa la loro attendibilità.

L'attività svolta per l'indagine pedologica, e in particolare la creazione della mappa propedeutica al rilevamento, ha dimostrato che una programmazione preventiva del campionamento ha maggiori probabilità di cogliere la variabilità dei suoli e nel contempo di minimizzare il numero di punti di campionamento. Inoltre, tale mappa si è dimostrata di utilità pratica nel corso del lavoro, poiché la Carta Tecnica Regionale, sovrapposta alle unità cartografiche di maggiori dimensioni, determinate con un metodo oggettivo, ha facilitato l'individuazione sul campo degli areali da considerare per l'indagine.

E' stato considerato, inoltre, che per la redazione di mappe tematiche relative al suolo spesso si opera in aree per le quali non sono disponibili gli strati informativi, relativi ai fattori pedogenetici, tipicamente utilizzati nella cartografia tradizionale: si è dunque scelto di procedere alla redazione delle mappe utilizzando metodi basati soltanto sulla geomorfometria. Il lavoro svolto ha dimostrato che è possibile procedere alla modellazione di mappe tematiche, relative al suolo, utilizzando la sola informazione derivante dal DEM, e prescindendo dall'informazione ausiliaria tipicamente utilizzata per la redazione di mappe pedologiche.

Infatti, i modelli distributivi, ottenuti con differenti approcci statistici, che hanno portato alla redazione di mappe tematiche sia parametriche che categoriali, sono ritenuti attendibili, almeno in prima approssimazione.

Tale giudizio tiene in considerazione i seguenti aspetti:

- per la modellazione delle mappe è stato utilizzato un numero esiguo di punti rispetto a quelli normalmente richiesti dai metodi statistici utilizzati. Questo si è verificato in particolare nel caso delle mappe relative alle classi di suolo, alcune delle quali sono state modellate utilizzando solo quattro casi;
- la valutazione dei modelli è stata effettuata almeno parzialmente con dati non utilizzati nel processo di modellazione;
- gli indici di valutazione dei modelli hanno valori accettabili;
- vi è congruenza tra i diversi prodotti cartografici.

Al di là degli indici statistici derivanti dai metodi di validazione, può essere formulato un giudizio di tipo qualitativo. Nelle carte dell'erosione del suolo sono stati mappati oggetti con vere e proprie morfologie individuabili sul territorio (es: calanchi e calanchi mammellonari), insieme con oggetti che non hanno una tale specificità morfologica (es: frane superficiali e suffosione), tuttavia esiste coerenza fra i diversi risultati prodotti. Riguardo i tipi di suolo (tassonomia WRB e tipologie) sono state prodotte due mappe con un numero di classi differente che esibiscono una parziale sovrapposizione, per cui si può affermare che esiste fra loro una apprezzabile conformità. Ancor più interessante è il confronto tra la mappa dello spessore del suolo, che rappresenta un carattere parametrico, e le altre mappe, tutte con variabili categoriali. Vi è ad esempio una buona corrispondenza tra i suoli sottili (o quelli potenti), presenti nella mappa delle tipologie, e la stima numerica dello spessore. Esiste parimenti una buona corrispondenza fra assenza di suolo (mappa dello spessore) e dinamiche calanchive, oppure fra spessore del suolo e suscettibilità a frane superficiali e suffosione.

Le mappe elaborate sono state concepite per ottenere una conoscenza utile a comprendere la dinamica evolutiva dell'area: esse, infatti, forniscono indicazioni per lo studio dei fenomeni erosivi in corso e per la gestione territoriale; possono, inoltre, essere utili per comprendere le dinamiche del recente passato e aiutare a creare ipotesi circa l'evoluzione futura. Occorre tuttavia tener presente che la cartografia è stata ottenuta attraverso un approccio di tipo statistico, partendo dalle informazioni derivanti dal rilevamento di campo, le quali comportano valutazioni soggettive. Esiste inoltre, una serie di fattori che possono influenzare la validità e l'accuratezza della cartografia realizzata, i quali sono imputabili a imprecisioni introdotte dall'operatore, oppure di origine computazionale.

La precisione con la quale vengono effettuate le osservazioni di campo, il livello di dettaglio con le quali vengono descritte, il numero di orizzonti pedologici campionati e l'uniformità delle modalità descrittive, influenzano in modo rilevante la qualità del dataset iniziale. Inoltre, nel caso dei suoli dell'area di studio sono da considerare le difficoltà intrinseche incontrate nel corso del rilevamento, a cominciare dalla valutazione dello spessore del solum con precisione centimetrica: questo può risultare difficile, sia per il tipo di formazioni soggiacenti ai suoli, sia per la difficoltà nell'esaminare i profili scavati manualmente (in alcuni casi il limite inferiore del *solum* è stato rilevato a profondità non visionabili direttamente, come nel caso di suoli su bassa pendenza). Inoltre, alcune caratteristiche dei suoli sono di difficile individuazioni a causa del periodo stagionale in cui viene effettuato il campionamento: nel caso dei suoli vertici, ad esempio, alcune caratteristiche non sono evidenti tutto l'anno, mentre gran parte del campionamento è stata eseguita nel periodo più piovoso.

Un altro fattore importante è l'accuratezza e la precisione con la quale vengono condotte le analisi di laboratorio, le quali influenzano in modo decisivo il processo di classificazione e quindi le successive fasi per la derivazione delle mappe e la loro validazione.

Inoltre, la qualità intrinseca del DEM si riflette nelle successive cartografie derivate e nelle cartografie tematiche realizzate. La dimensione del pixel condiziona la scala virtuale alla quale

vengono realizzate le cartografie tematiche: quello utilizzato, seppure di dettaglio, potrebbe non essere idoneo a rappresentare la variabilità topografica di zone particolarmente articolate.

Un altro fattore soggettivo riguarda la scelta delle stazioni in cui aprire i profili pedologici: questo, oltre a condizionare la possibilità di descrivere in modo completo ed esaustivo le diverse situazioni presenti nell'area di studio, assume un'importanza particolare nel caso dell'approccio impiegato, poiché il punto campionato con determinate coordinate geografiche è anche un punto nell'iperspazio determinato dalle variabili impiegate per la modellazione. A questo riguardo, alcuni suoli sono stati campionati in punti ove risultava più semplice mettere in evidenza il profilo (ad esempio, su scarpata): tuttavia, dal punto di vista del DEM e delle variabili derivate, questi punti potrebbero essere fuorvianti per le elaborazioni.

Altri aspetti che condizionano il prodotto finale riguardano la fase di calcolo, in particolare la scelta dei metodi statistici da impiegare o anche la scelta di procedere modellando l'area in modalità unitaria oppure suddivisa. Nel caso studiato, data la morfologia profondamente differente delle Arenarie del Monte Vallassa rispetto al resto dell'area, si sarebbero anche potute scindere le modellazioni.

Malgrado questi aspetti problematici, la metodologia adottata possiede, rispetto a quella cartografica tradizionale, alcuni peculiari vantaggi:

- l'impiego di un approccio statistico consente di ridurre il numero di osservazioni di campo, in modo da abbattere i costi di realizzazione della cartografia pedologica;
- la carta dei suoli, ottenuta con un limitato impiego di risorse, costituisce un valido punto di partenza per successivi studi di dettaglio;
- la qualità delle mappe prodotte, nel caso di aree modellate male, può essere facilmente incrementata con l'aggiunta di altri punti e successivo ricalcolo delle mappe utilizzando i modelli statistici migliorati.