

SPERIMENTAZIONE MEDIANTE LOGICA FUZZY PER LA CARATTERIZZAZIONE DI SEDIMENTI MARINI

Nicoletta GAZZEA (*), Emiliana VALENTINI (*), Maria Elena PICCIONE (*),
Andrea TARAMELLI (*) (**)

(*) ISPRA Istituto Superiore per la Protezione e Ricerca Ambientale
(ex ICRAM Istituto Centrale per la Ricerca Scientifica e Tecnologica Applicata al Mare)
Via di Casalotti, 300 – 00166 – Roma Tel. +390661570508, Fax. +390661561906.

Indirizzo e-mail n.gazzea@icram.org

(**) Lamont Doherty Earth Observatory of Columbia University, New York, USA

Riassunto

Il presente lavoro intende contribuire allo studio delle tecniche più recenti di inferenza in ambiti territoriali. La metodologia utilizzata si basa su un approccio probabilistico non parametrico che sfrutta la logica fuzzy per prevedere potenziali aree contaminate. Il modello impiegato è basato su criteri diagnostici definiti dallo spessore di substrato, dalla batimetria del fondale e dalla presenza e quindi influenza degli scarichi, civili ed industriali, localizzati nell'area di indagine. Il risultato preliminare è uno scenario di potenziali zone sulle quali focalizzare i campionamenti al fine di caratterizzare al meglio i sedimenti marini.

Abstract

In this paper, we examine an application of sediment characterization using an ontological approach in GIS, which helps to define a specific scenario grounded on the abilities from an existing spatial dataset.

A comparison between the fuzzy-based scenario and the existing spatial dataset carried out in the study area shows areas suitable and unsuitable for potentially contaminated sediments. A sensitivity analysis applied to the area illustrates a significant increase of object suitable for classification when bathymetry and distance from the contaminating sources are decreased within the lower limit while the sediments thickness is increased. So that inaccuracies in attributes influence the determination of the objects. The sensitivity of the model shows that artificial boundaries for the contaminated areas are not reasonable and consequently should be treated as vague objects.

Introduzione

Per una mirata caratterizzazione dei sedimenti marini, è necessario determinare nell'area di studio delle sottoaree omogenee sia dal punto di vista fisico-morfologico che per livello di contaminazione. A tale scopo si è utilizzata la logica fuzzy o teoria degli insiemi sfumati (Burrough, 1989), per la quale le aree di studio sono caratterizzate da completa vaghezza o incertezza. A differenza della probabilità che si dissolve quando i dati noti diventano numerosi, la natura fuzzy dei fenomeni emerge all'aumentare delle informazioni disponibili. Più dati quindi aiutano a fissare la sfumatura mediante la quale gli insiemi si sovrappongono, eliminando i confini che segnano dove una realtà oggettiva cessa di essere quella realtà. A ciascun elemento vengono attribuiti infiniti valori, compresi tra 0 e 1, che corrispondono ai gradi di appartenenza ad uno specifico insieme fuzzy (Zadeh, 1965). Tale metodologia si basa sul giudizio ed esperienza del ricercatore sia nel selezionare le variabili da utilizzare che nell'associare ad esse i pesi più corretti.

Area di Studio

L'area oggetto dello studio è il bacino portuale della Rada di Augusta, comunicante con il mare attraverso due imboccature ed originata dalla chiusura di buona parte dell'insenatura naturale per mezzo di dighe foranee (Fig. 1). Tale area, già oggetto di studi ambientali pregressi, è caratterizzata dalla presenza di numerosi insediamenti industriali e civili (ICRAM, 2008).

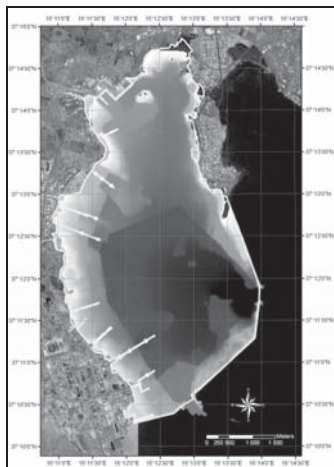
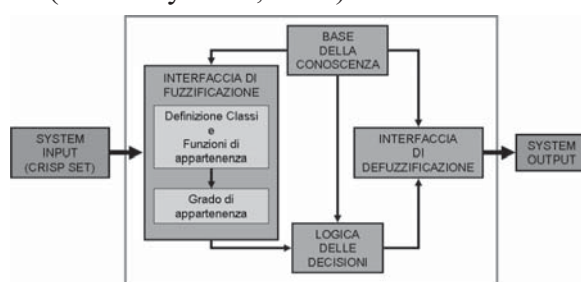


Fig. 1- La Rada di Augusta.

Metodologia di lavoro

La metodologia utilizzata, rappresentata nello schema è stata sviluppata mediante l'utilizzo, in ambiente ArcGis®9.2, del *toolbox* Spatial Data Modeller (SDM), un insieme di *tools* che permettono il *geoprocessing* per la modellazione di dati spaziali (Sawatzky *et al.*, 2008).



System Input

Il database utilizzato in questo lavoro è stato realizzato da ICRAM (2008) su dati georiferiti nel datum WGS 84 – UTM fuso 33, con una risoluzione spaziale dei grid pari a 15x15 m:

- *feature* puntuale dei risultati della pregressa caratterizzazione dei sedimenti marini, utilizzata per verificare la coerenza dei risultati ottenuti;
- *feature* puntuale degli scarichi industriali e civili presenti lungo la costa, dalla quale è stato calcolato, in ambiente ArcGis, il grid (15x15 m) relativo alla distanza euclidea dagli scarichi nell'area di indagine;
- *feature* poligonale dell'area di indagine;
- *grid* della batimetria, per il quale è stata necessaria una trasformazione in valori positivi, al fine di utilizzarlo come dato di input nel tool Fuzzy Evidence di SDM;
- *grid* dello spessore del substrato.

Le variabili selezionate e utilizzate, come *system input*, nella metodologia fuzzy sono:

- il *grid* relativo alla batimetria (bati_pos) per considerare le discontinuità morfologiche, i fenomeni di risospensione e i trascorsi dragaggi portuali;
- il *grid* della distanza dagli scarichi industriali e civili (extract_dist) per considerare le diverse e possibili fonti di inquinamento;
- il *grid* del substrato (substrato) per considerare le eventuali rocce affioranti o aree nelle quali risulterebbe comunque difficile effettuare campionamenti.

Interfaccia di Fuzzificazione

Primo step di un modello *fuzzy logic* è la fuzzificazione del set di dati originali (*crisp set*), ossia la loro “conversione” in valori di appartenenza fuzzy (*fuzzy set*) compresi nell'intervallo [0,1], selezionando la funzione di appartenenza (*membership function*) più idonea a rappresentare il set di dati. Le funzioni di appartenenza disponibili nel tool SDM sono: Lineare, Large, Small, Near, Gaussiana, MS Large e MS Small (Sawatzky *et al.*, 2008 e relative citazioni).

Per ognuna delle tre variabili considerate ed in funzione del loro *crisp set*, è stata selezionata quella lineare che permette di convertire con una relazione diretta l'insieme dei dati di input in valori del

grado di appartenenza. Per ogni *crisp set* è stato utilizzato l'intero intervallo e non dei sottoinsiemi di esso, per non tralasciare alcuna informazione inerente l'area di studio.

Al grid *bati_pos* è stata applicata la *membership function* lineare con i seguenti parametri: min = 40.78 m e max = 0 m. Il grado di appartenenza maggiore è stato attribuito ai bassi fondali per considerare i fenomeni di risospensione. La funzione utilizzata per determinare il grado di appartenenza *li_bati* è il sistema (I). Al grid *extract_dist* è stata applicata la *membership function* lineare con i seguenti parametri: min = 3187.99 m e max = 0 m. Il grado di appartenenza maggiore è stato attribuito alle distanze inferiori dagli scarichi, sia industriali che civili, che potrebbero rappresentare eventuali fonti di inquinamento. La funzione utilizzata per determinare il grado di appartenenza *li_scarichi* è il sistema (II). Al grid *substrato* è stata applicata la *membership function* lineare con i seguenti parametri: min = 0 m e max = 6.34 m tali da conferire un grado di appartenenza superiore alle aree dove, la maggiore presenza di substrato, permette l'esecuzione di campionamenti tramite carotiere garantendo il prelievo di sedimento in profondità e pertanto una caratterizzazione più completa. La funzione corrispondente per determinare il grado di appartenenza *li_sub* è il sistema (III).

$$\begin{aligned}
 \text{(I)} \quad & \begin{cases} 0 & \left(\frac{\text{float}(\text{bati_pos}) - 40.78}{0 - 40.78} \right) < 0 \\ \left(\frac{\text{float}(\text{bati_pos}) - 40.78}{0 - 40.78} \right) & 0 < \left(\frac{\text{float}(\text{bati_pos}) - 40.78}{0 - 40.78} \right) < 1 \\ 1 & \left(\frac{\text{float}(\text{bati_pos}) - 40.78}{0 - 40.78} \right) > 1 \end{cases} \\
 \text{(II)} \quad & \begin{cases} 0 & \left(\frac{\text{float}(\text{extract_dist}) - 3187.99}{0 - 3187.99} \right) < 0 \\ \left(\frac{\text{float}(\text{extract_dist}) - 3187.99}{0 - 3187.99} \right) & 0 < \left(\frac{\text{float}(\text{extract_dist}) - 3187.99}{0 - 3187.99} \right) < 1 \\ 1 & \left(\frac{\text{float}(\text{extract_dist}) - 3187.99}{0 - 3187.99} \right) > 1 \end{cases} \\
 \text{(III)} \quad & \begin{cases} 0 & \left(\frac{\text{float}(\text{substrato}) - 0}{6.34 - 0} \right) < 0 \\ \left(\frac{\text{float}(\text{substrato}) - 0}{6.34 - 0} \right) & 0 < \left(\frac{\text{float}(\text{substrato}) - 0}{6.34 - 0} \right) < 1 \\ 1 & \left(\frac{\text{float}(\text{substrato}) - 0}{6.34 - 0} \right) > 1 \end{cases}
 \end{aligned}$$

Sono stati determinati i gradi di appartenenza rappresentati dalla legenda, suddivisa in 20 intervalli, (Sawatzky et al., 2008), dopo aver selezionato ed applicato, ad ogni variabile considerata, le funzioni di appartenenza. La legenda, riportata in tabella, è stata tematizzata in una scala di grigi, dove il colore bianco indica la non appartenenza, mentre il nero la completa appartenenza.

In Fig. 2 sono mostrati i grafici, ottenuti mediante l'impiego di Statistica[®] 7.1, e le equazioni delle funzioni di appartenenza utilizzate. Le pendenze delle rette dipendono strettamente dai vincoli imposti (intervallo dei dati, massimo e minimo da associare ad una completa o nulla appartenenza, ect..) e nello specifico caso (A) e (B) hanno pendenza negativa, mentre (C) positiva. I valori degli intervalli riportati in tabella sono stati calcolati tramite le equazioni delle funzioni di appartenenza.

Grado di Appartenenza		Crisp Set		
$\mu(x)$		Substrato (m)	Batimetria (m)	Distanza dagli Scarichi (m)
0 - 0.05	Nonmember	0.00-0.31	40.81-38.77	3187.99-3166.66
0.05 - 0.1		0.31-0.63	38.77-36.73	3166.66-3000.00
0.1 - 0.15		0.63-0.94	36.73-34.69	3000.00-2833.33
0.15 - 0.2		0.94-1.26	34.69-32.65	2833.33-2666.66
0.2 - 0.25		1.26-1.58	32.65-30.61	2666.66-2500.00
0.25 - 0.3		1.58-1.90	30.61-28.57	2500.00-2333.33
0.3 - 0.35		1.90-2.21	28.57-26.53	2333.33-2166.66
0.35 - 0.4		2.21-2.53	26.53-24.49	2166.66-2000.00
0.4 - 0.45		2.53-2.85	24.49-22.45	2000.00-1833.33
0.45 - 0.5		2.85-3.16	22.45-20.41	1833.33-1666.66
0.5 - 0.55		3.16-3.48	20.41-18.37	1666.66-1500.00
0.55 - 0.6		3.48-3.80	18.37-16.33	1500.00-1333.33
0.6 - 0.65		3.80-4.12	16.33-14.29	1333.33-1166.66
0.65 - 0.7		4.12-4.43	14.29-12.24	1166.66-1000.00
0.7 - 0.75		4.43-4.75	12.24-10.20	1000.00-833.33
0.75 - 0.8		4.75-5.07	10.20-8.16	833.33-666.66
0.8 - 0.85		5.07-5.38	8.16-6.12	666.66-500.00
0.85 - 0.9		5.38-5.70	6.12-4.08	500.00-333.33
0.9 - 0.95		5.70-6.02	4.08-2.04	333.33-166.66
0.95 - 1	Full Member	6.02-6.34	2.04-0.00	166.66-0.00

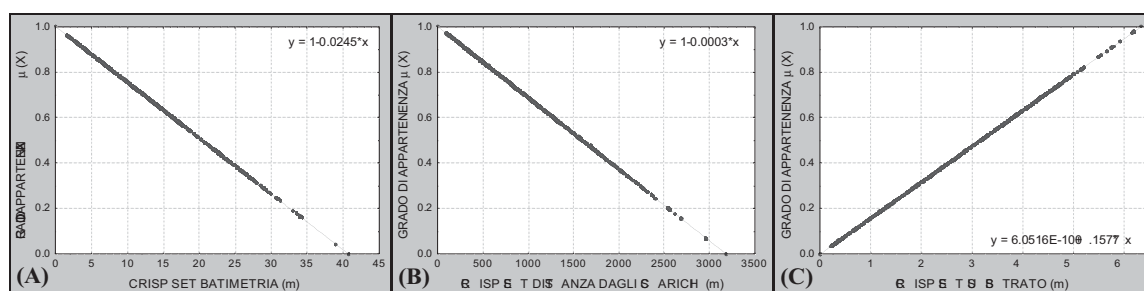


Fig. 2 – Fuzzificazione Crisp Set Batimetria (A), Distanza dagli scarichi (B) e Substrato (C). In ascissa sono visualizzati i valori del crisp set e in ordinata il grado di appartenenza corrispondente ai valori minimi e massimi imposti nella funzione di appartenenza.

Si riportano in Fig. 3 le mappe di output della fuzzificazione relative al grado di appartenenza di ciascuna variabile.

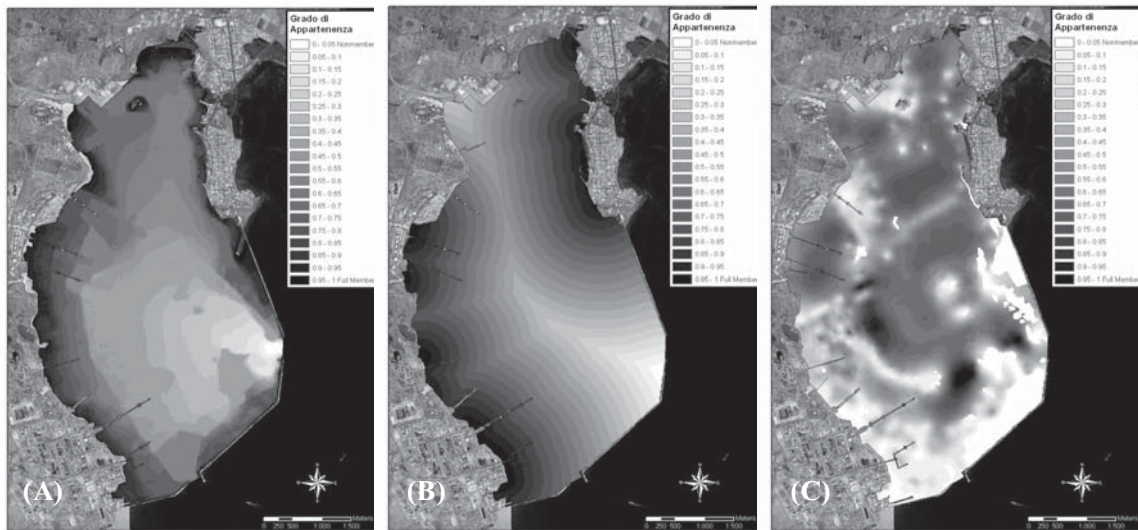


Fig. 3 – Mappa del Grado di Appartenenza della Batimetria (A), Distanza dagli scarichi (B) e Substrato (C).

Logica delle Decisioni

Lo step successivo riguarda l'applicazione delle regole di combinazione tra gli insiemi fuzzy (*fuzzy rules*) nella forma *IF x IS A THEN y IS B* per sintetizzare la conoscenza dell'esperto. In letteratura sono molteplici le formulazioni linguistiche e le applicazioni di esse (Klir e Yuan, 1995).

Dopo la fuzzificazione dei tre *crisp set*, si è applicata una regola che fornisse in output una suddivisione del sito in aree prioritarie da campionare. Da uno studio approfondito dei principali operatori fuzzy (Bonham-Carter, 1994), si è determinato che l'operatore più conforme al raggiungimento dell'obiettivo, fosse la *gamma operation*, in grado di restituire in output un risultato, al variare del parametro gamma, compreso nell'intervallo di input. L'equazione di riferimento è :

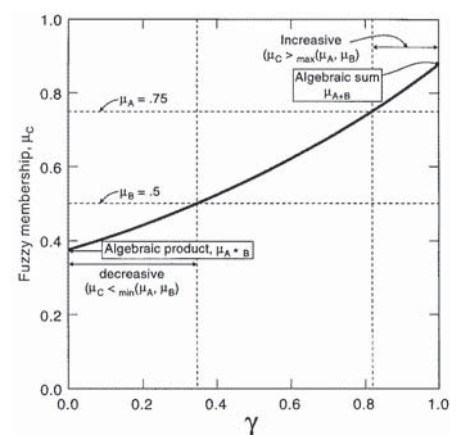
$$\mu_{combination} = (Fuzzy\ algebraic\ sum)^{\gamma} * (Fuzzy\ algebraic\ product)^{1-\gamma}$$

Dove γ è compreso tra 0 e 1 (Zimmermann e Zysno, 1980).

Per selezionare il parametro gamma si è utilizzato il grafico (Bonham-Carter, 1994) nel quale, al variare di gamma, è indicato l'andamento del grado di appartenenza μ_C in funzione della combinazione di due gradi di appartenenza μ_A e μ_B forniti in input. Per $\gamma=1$ si ottiene lo stesso μ_C della *Fuzzy Algebraic Sum*, per $\gamma=0$ quello della *Fuzzy Algebraic Product*, per valori di γ compresi tra 0.8 e 1 si ottengono μ_C superiori a quelli di ingresso, per γ compresi tra 0 e 0.35 inferiori mentre per γ compresi tra 0.35 e 0.8 valori inclusi nell'intervallo in input.

E' stata implementata la *gamma operation* per valori di γ compresi nell'intervallo tra 0.35 e 0.8. Il valore di gamma, più idoneo a rappresentare l'obiettivo prefissato, è risultato $\gamma=0.75$

(Fig. 4). La mappa ottenuta è rappresentata da sole 18 classi e non 20; questo deriva dal fatto di combinare i gradi di appartenenza delle tre variabili e soprattutto di utilizzare la *gamma operation* che è funzione dell'*algebraic product*, che restituisce valori molto più bassi di quelli forniti in ingresso.



Interfaccia di Defuzzificazione e output finale

L'output del processo può essere utilizzato sia nella forma di grado di appartenenza, per restituire un'informazione qualitativa o può essere defuzzificato e quindi riportato, mediante una scala di conversione, a numero reale.

In tabella è riportato l'output finale, dove ad ogni classe di appartenenza sono associati gli intervalli dei valori reali delle tre variabili considerate.

Classe	Grado di Appartenenza $\mu(x)$	SUBSTRATO min	SUBSTRATO medio	SUBSTRATO max	BATIMETRIA min	BATIMETRIA medio	BATIMETRIA max	DISTANZA SCARICHI min	DISTANZA SCARICHI medio	DISTANZA SCARICHI max
1	0 - 0.05 Nonmember	0.00	1.25	2.50	0.88	20.83	40.78	96.05	1642.02	3187.99
2	0.05 - 0.1	0.10	1.28	2.46	24.17	32.45	40.73	2623.46	2655.80	3088.14
3	0.1 - 0.15	0.01	1.63	3.24	2.96	21.51	40.06	1448.10	2307.62	3167.13
4	0.15 - 0.2	0.01	1.98	3.94	2.10	20.77	39.45	906.60	2037.95	3189.30
5	0.2 - 0.25	0.02	2.31	4.60	3.32	21.00	38.67	764.85	1944.82	3124.79
6	0.25 - 0.3	0.06	2.56	5.06	2.81	20.39	37.97	657.44	1861.45	3065.47
7	0.3 - 0.35	0.09	2.81	5.53	2.51	19.67	36.84	141.51	1589.84	3038.16
8	0.35 - 0.4	0.12	3.15	6.17	2.46	19.19	35.91	75.00	1526.97	2978.33
9	0.4 - 0.45	0.20	3.22	6.23	1.32	18.33	35.35	15.00	1284.22	2553.44
10	0.45 - 0.5	0.28	2.93	5.58	1.29	15.22	29.15	15.00	1155.43	2301.85
11	0.5 - 0.55	0.44	3.43	6.21	1.57	15.52	29.06	30.00	1134.31	2228.62
12	0.55 - 0.6	0.65	3.49	6.33	1.32	15.03	28.75	33.54	1137.71	2241.99
13	0.6 - 0.65	0.94	3.59	6.22	1.01	14.30	27.60	45.00	1149.25	2253.50
14	0.65 - 0.7	1.26	3.70	6.14	0.92	12.09	23.26	33.54	1117.61	2201.68
15	0.7 - 0.75	1.67	3.94	6.20	0.95	11.94	22.93	0.00	981.93	1963.95
16	0.75 - 0.8	2.28	3.62	4.96	1.00	9.70	18.40	47.43	534.87	1022.31
17	0.8 - 0.85	2.97	4.03	5.09	1.02	7.21	13.39	45.00	477.29	908.57
18	0.85 - 0.9	3.63	4.26	4.89	0.69	4.05	7.42	45.00	366.68	688.37
19	0.9 - 0.95	n.p.	n.p.	n.p.	n.p.	n.p.	n.p.	n.p.	n.p.	n.p.
20	0.95 - 1 Full Member	n.p.	n.p.	n.p.	n.p.	n.p.	n.p.	n.p.	n.p.	n.p.

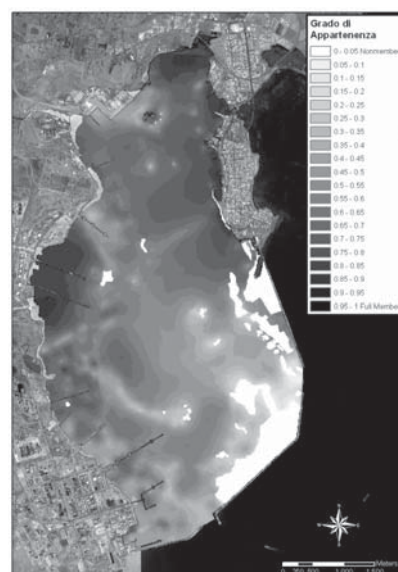


Fig. 4 – Mappa del grado di appartenenza determinato dalla combinazione delle tre variabili di input.

In Fig. 5, sono visualizzati i grafici dei risultati ottenuti, nei quali la classificazione dei gradi di appartenenza, derivati dalla combinazione delle tre variabili fuzzificate, è associata agli intervalli dei valori reali di Batimetria, Distanza dagli scarichi e Substrato. Il grafico di ciascuna variabile mostra il medesimo trend osservato nei grafici di Fig. 2, rappresentativi delle funzioni di appartenenza utilizzate nella fuzzificazione del *crisp set*.

La principale singolarità è relativa alla classe 0-0.005. Si è verificato che l'intervallo così ampio deriva esclusivamente dall'utilizzo della *gamma operation* e dai vincoli imposti (sistema I, II e III) durante la fuzzificazione delle variabili considerate. Infatti per uno spessore di substrato nullo, il prodotto restituisce, per qualsiasi valore di batimetria o distanza dagli scarichi, la classe di appartenenza 0-0.005 nella quale ricadono anche batimetrie profonde o elevate distanze dagli scarichi.

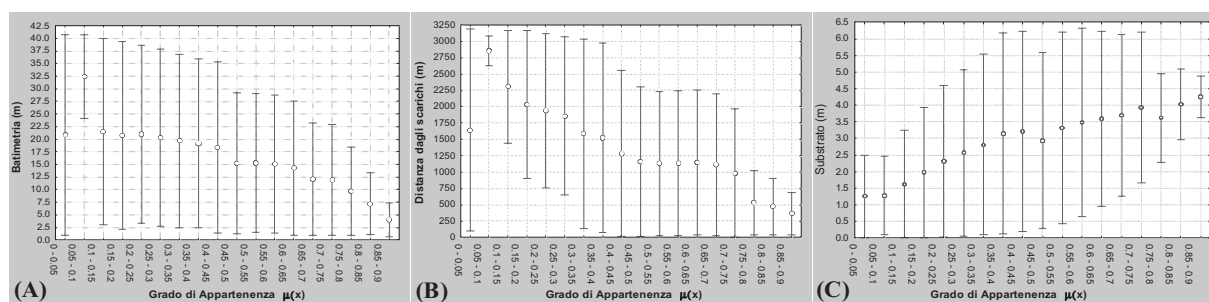


Fig. 5 – Grafici relativi agli intervalli di valori reali corrispondenti alle classi dei gradi di appartenenza per le tre variabili: Batimetria (A), Distanza dagli scarichi (B) e Substrato (C).

Visualizzando i grafici è possibile associare ad ogni classe del grado di appartenenza i diversi intervalli dei *crisp set*. Ad esempio, alla classe con grado di appartenenza 0.85-0.9 è associata batimetria di basso fondale, scarsa distanza dagli scarichi industriali e civili e spessore di substrato profondo. Alla classe 0.45-0.5 è associata batimetria di medio fondale, media distanza dagli scarichi industriali e civili e spessore di substrato medio. La classe 0-0.005, secondo quanto discusso precedentemente, è da considerare singolare per quanto riguarda batimetria e distanza dagli scarichi, poiché è influenzata maggiormente dal valore nullo dello spessore del substrato.

Risultati e Conclusioni

Il risultato preliminare ottenuto, mostra una rappresentazione spaziale discreta delle variabili considerate, che individua sottoinsiemi di aree, all'interno dei quali esiste un'omogeneità morfologica e sedimentologica, utile a guidare la distribuzione delle stazioni di campionamento.

A titolo di confronto, si riporta la mappa dell'output finale alla quale sono state sovrapposte le concentrazioni del Mercurio, campionate in una pregressa campagna di caratterizzazione e tematizzate secondo un grado di contaminazione individuato applicando la legislazione vigente (Fig. 6). Si evidenzia che l'area etichettata con la lettera A non mostra coerenza con i risultati della pregressa caratterizzazione in quanto, pur essendo un'area omogenea dal punto di vista morfologico e sedimentologico, è caratterizzata dalla presenza di attività industriali non connesse a possibili sversamenti di Mercurio. Una debole coerenza si osserva per l'area B, nella quale insistono diversi scarichi civili, ai quali, in questo studio, è stato associato lo stesso peso di quelli industriali. L'area che ha mostrato la maggiore coerenza con i risultati ottenuti è la C, nella quale le tre variabili selezionate risultano sufficienti a descrivere la potenziale contaminazione.

I risultati della pregressa caratterizzazione evidenziano una contaminazione non riconducibile ad aree discrete, a sottolineare la necessità di applicare metodologie che permettano di rendere la logica di indagine più aderente alle sfumature della variabilità reale.

E' emerso che per ottenere una risposta coerente con il reale scenario di riferimento è necessario valutare attentamente il numero e la tipologia delle variabili da includere nel modello. Ad esempio l'aver associato uguale peso sia agli scarichi industriali che civili ha reso omogenee aree esposte a contaminazioni differenti. Inoltre è opportuno introdurre variabili che influenzano la risospensione, dispersione e diffusione dei contaminanti, come ad esempio il contributo delle correnti generate dal vento e dalla marea.

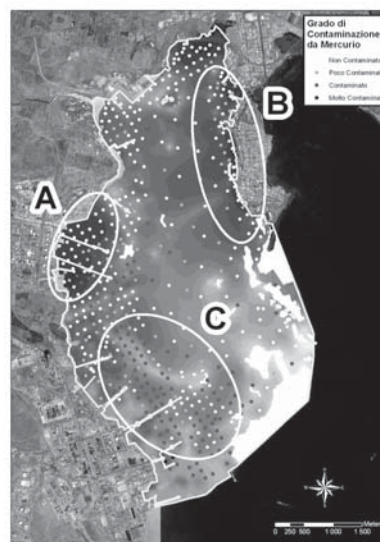


Fig. 6 – Mappa del grado di appartenenza relativo alla combinazione delle tre variabili di input con sovrapposto il grado di contaminazione da Mercurio.

Bibliografia

- Bonham-Carter, G. F., (1994), Geographic information systems for geoscientists, modelling with GIS. Ontario, Pergamon, 398p.
- Burrough, P. A., (1989), Fuzzy mathematical methods for soil survey and land evaluation. Journal of Soil Science, 40: 477-492
- KLIR, G. J. and YUAN, B., (1995), Fuzzy Sets and Fuzzy Logic: Theory and Applications. Prentice-Hall, Upper Saddle River, NJ
- Sawatzky, D.L., Raines, G.L., Bonham-Carter, G.F., and Looney, C.G., (2008), Spatial Data Modeller (SDM): ArcMAP 9.2 geoprocessing tools for spatial data modelling using weights of evidence, logistic regression, fuzzy logic and neural networks. <http://arcscrips.esri.com/details.asp?dbid=15341> data ultimo accesso 01/09/2008
- Zadeh, L.A., (1965), Fuzzy sets. Inst. Electr. Elettron. Eng., Information and Control. v.8, p. 338-353
- Zimmermann, H.J. and Zysno, P., (1980), Latent connectives in human decision making: Fuzzy Sets and System, v.4, p.37-51
- ICRAM (2008), BoI-Pr-SI-PR-Rada di Augusta-03.22