

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

SCUOLA DI SCIENZE

Corso di laurea magistrale in *Biologia Marina*

**IMPLEMENTAZIONE DI UN SISTEMA
PER LA CLASSIFICAZIONE DI BANCHI DI PESCE
DA DATI DI ACUSTICA**

Tesi di laurea in

MODELLI DI VALUTAZIONE E GESTIONE DELLE RISORSE ALIEUTICHE

Relatore

Prof. Victor Ugo Ceccherelli

Presentata da

Irene Giliberto

Correlatori

Dott. Fabio Bozzeda

Dott. Angelo Bonanno

III sessione

Anno Accademico 2012/2013

INDICE

	Pag.
<i>PREMESSA</i>	7
1. INTRODUZIONE	9
1.1. Importanza della valutazione degli <i>stock</i> ittici	11
1.2. Variabili ambientali per la valutazione degli <i>stock</i> ittici	14
1.3. Acustica per i banchi di pesce: stato dell'arte	15
1.3.1. Accuratezza di metodi acustici in <i>fish stock assessment surveys</i>	19
1.4. Capo Granitola	24
2. MATERIALI E METODI	27
2.1. <i>Sonar</i> ed <i>echosounder</i>	27
2.2. Trattamento di dati a singola frequenza per l'estrazione di banchi di pesce e delle aggregazioni planctoniche	29
2.3. Trattamento dei dati a multifrequenza per la caratterizzazione e l'identificazione dei banchi di pesci	31
2.4. Trattamento dei dati acustici per l'estrazione delle informazioni sui banchi di pesce da dati a multifrequenza	32
2.5. L'approccio <i>fuzzy naïve Bayes</i> per la classificazione dei segnali da acustica	33
2.5.1. <i>Fuzzy Logic</i>	33
2.5.2. Compilatore <i>naïve Bayes</i>	34
2.5.3. Approccio <i>fuzzy naïve Bayes</i>	36
2.5.4. Considerazioni sul metodo	38
2.6. Approccio metodologico: campagne oceanografiche	39
3. ELABORAZIONI	45
3.1 Criteri adottati per l'elaborazione	45
3.2. Scelta delle variabili	45
3.3. Correlazioni tra le variabili e le percentuali di pesci catturati	51

3.4. Metodo <i>fuzzy naïve Bayes</i> : fuzzyficazione	54
4. RISULTATI	55
5. DISCUSSIONE E CONCLUSIONI	59
<i>BIBLIOGRAFIA</i>	65
<i>SITOGRAFIA</i>	79
<i>APPENDICE</i>	81

PREMESSA

Il mare copre circa il 71% della superficie della Terra e rappresenta da sempre una fonte di sostentamento fondamentale per l'uomo. Fino a qualche decennio fa si riteneva, in modo del tutto errato, che le risorse ittiche da esso offerte alla pesca fossero inesauribili e che riuscissero a rigenerarsi indipendentemente dal prelievo effettuato.

Non sono emersi problemi finché l'attività di pesca era artigianale e locale, contribuendo a sostenere cultura, economia, occupazione e approvvigionamento alimentare delle comunità costiere (Gabriel *et al.*, 2005; World Bank, 2009). A partire dal 1880, con l'avvento della pesca industriale e dei primi pescherecci a vapore in Inghilterra, lo sforzo di pesca applicato agli *stock* ittici di interesse commerciale ha cominciato a crescere. Le Guerre Mondiali hanno temporaneamente sospeso tale tendenza allo sfruttamento, dato che gli *stock* ittici in quei periodi storici non subivano la pressione del prelievo grazie alla riduzione dello sforzo di pesca applicato; è anche vero che la tecnologia sviluppata per scopi bellici ha dato modo alle attività alieutiche, una volta riprese, di aumentare l'efficacia dello sforzo di pesca (Anticamera *et al.*, 2011). L'impatto antropico che ne è derivato si configurava, e si configura, essenzialmente come un disequilibrio dovuto al prelievo smodato: a causa delle tecnologie sempre più evolute, le catture effettuate dall'uomo avvengono in tempi troppo brevi affinché la biomassa ittica riesca, attraverso i naturali tempi di riproduzione, a ripristinare la risorsa prelevata. Tale disequilibrio intacca le popolazioni ittiche non solo riducendone le dimensioni, ma anche alterandone le strutture d'età.

La crescente globalizzazione e la domanda sempre più insistente di risorse ittiche da parte dei paesi sviluppati (Swartz *et al.*, 2010) non fanno altro che accrescere lo sforzo di pesca su scala mondiale continuando a causare pressioni sulla risorsa. Secondo quanto riferito nel 2009 dalla FAO (*Food and Agriculture Organization of the United Nations*), l'espansione dello sforzo di pesca avvenuta dal 1970 in poi ha comportato che nel 2007 il 52% degli *stock* ittici mondiali è stato completamente sfruttato, il 28% è stato sovrasfruttato o impoverito, il 20% moderatamente sfruttato e solo l'1% ha mostrato segni di ripresa. Se si considera lo sforzo di pesca in termini di potenza dei motori tradotta in kilowatt per giorni di pesca (KW/giorni), la Figura 1 mostra che il totale dei KW/giorni globali stimati per tutte le nazioni è rimasto quasi invariato tra il 1950 e il 1960, mentre dal 1961 si è verificata una tendenza all'aumento dello sforzo di pesca

(1,1% di incremento annuo), fino al valore più alto (4,4 miliardi di KW/giorni) tra il 1950 e il 2010 con un aumento del 54 % (Anticamara *et al*, 2011).

J.A. Anticamara *et al.* / *Fisheries Research* 107 (2011) 131–136

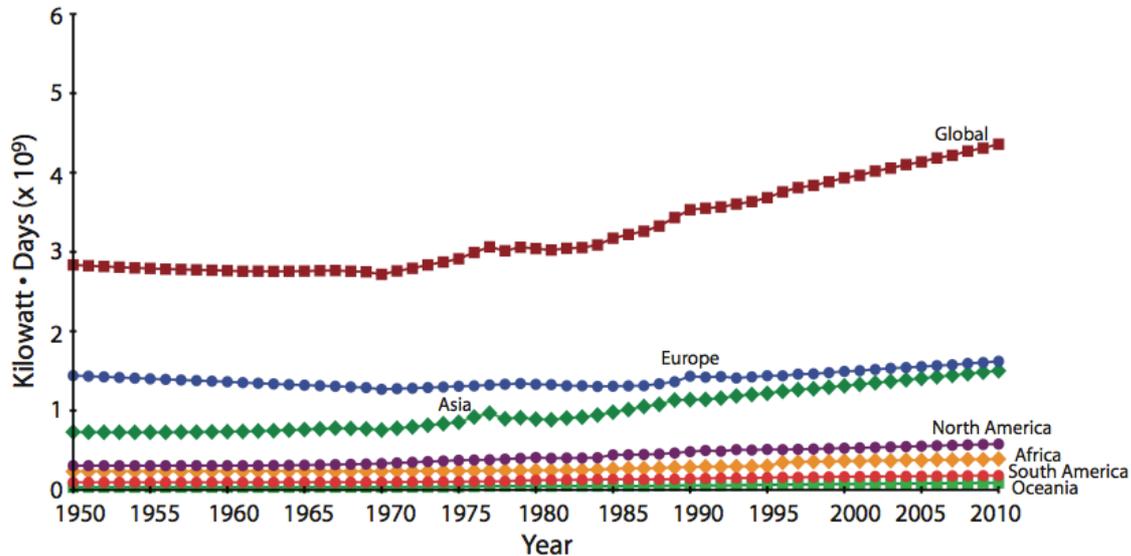


Figura 1. Serie temporali di stime annuali dello sforzo totale di pesca esercitato sia a livello globale, sia a livello dei singoli continenti (1950-2010).

Ad oggi l'attività di pesca è in continuo accrescimento e l'*overfishing* sta portando sempre più al depauperamento sia delle risorse ittiche che dei loro ambienti naturali. Risulta quindi necessario, al fine di tenere sotto controllo lo stato della risorsa, effettuare una valutazione preliminare degli *stock* ittici, che in genere fino ad ora è stata sempre più lenta e a rimorchio dell'attività di pesca; ciò al fine di arrivare ad avere per tempo informazioni atte ad attuare strumenti di gestione idonei ad una pesca razionale, consapevole e gestita senza intaccare la naturale capacità di ripristino delle popolazioni ittiche.

L'obiettivo del mio lavoro di tesi pertanto è stato quello di analizzare i dati acquisiti in tre zone prospicienti la costa della Sicilia meridionale, nel tentativo di mettere a punto tecniche per l'acquisizione dati meno invasive rispetto ai metodi tradizionali e potenzialmente adatte ad aumentare la conoscenza della risorsa sia nello spazio che nel tempo.

1. INTRODUZIONE

Nel mar Mediterraneo la valutazione dello stato delle risorse marine e la gestione dell'attività di pesca non sono spesso correlate, come sarebbe invece auspicabile. La prima rappresenta un'attività di fondamentale importanza svolta dalla comunità scientifica la quale, attraverso specifici progetti di valutazione, ha il compito di acquisire dati e monitorare lo stato degli *stock* ittici e delle relative popolazioni (Leonard & Maynou, 2003).

Paesi mediterranei, specialmente l'Italia (Ardizzone & Corsi, 1997), ma anche altri Paesi europei che si affacciano sul Mediterraneo, come dimostrato dal progetto MEDITS dal 1994 (Bertrand *et al.*, 1998; Abella *et al.*, 1999; Abellò *et al.*, 2002), utilizzano come principale strumento di valutazione delle attività di pesca i “*trawl surveys*”; l'impiego di cale con rete a strascico è quindi un mezzo molto utile nella valutazione degli *stock* ittici e per la gestione delle risorse marine, ma determina anche inevitabili impatti antropici. Diverse specie ittiche pelagiche di interesse commerciale ed ecologico sono monitorate da tempo attraverso *survey* acustici, che ne stimano la biomassa (per “biomassa” si intende, in generale, la massa totale di tutti gli organismi presenti in un dato tempo in un volume d'acqua o di terra) e la distribuzione degli *stock* (Simmonds and MacLennan, 2005). Sebbene lo strumento acustico impiegato (*echosounder*) sia poco invasivo, nello svolgimento dei *survey* acustici vengono effettuate delle catture con rete pelagica proprio al fine di identificare le specie ittiche insonificate dallo strumento e di determinare la composizione per classi di taglia degli *stock*.

Alcuni rilevanti studi di valutazione per mezzo di *echosurvey* riguardano ad esempio piccoli pelagici quali acciughe (*Engraulis encrasicolus*) e sardine (*Sardina pilchardus*) (Abad *et al.*, 1996; Patti *et al.*, 2000; Guennegan *et al.*, 2000; Garcia Lafuente *et al.*, 2002; Patti *et al.*, 2004; Bonanno *et al.*, 2005, Giannoulachi *et al.*, 2005) nel mar Mediterraneo (ad es. nel Canale di Sicilia e nel Mar Egeo) e nella Baia di Biscaglia (Massé *et al.*, 1996). Altri coinvolgono il Mar del Nord per i *survey* acustici sulle aringhe (*Clupea harengus*) (Simmond, 2003) e sui *sandeel* (*Ammodytes marinus*) (Greenstreet *et al.*, 1998); l'Atlantico nord-orientale sul *blu whiting* (*Micromesistius poutassou*) (ICES 2004g). Nel nord Europa ci sono più di 20 *stock* di pesce soggetti a valutazioni acustiche (ICES 2000). Altri studi di valutazione, che vedono l'applicazione delle tecniche acustiche nel resto del mondo, coinvolgono ad esempio il merluzzo del Newfoundland

(*Gadus morhua*) (Rose, 1995) e l'aringa del Pacifico in Alaska (*Clupea pallasii*) (Stokesbury *et al.*, 2002).

Inoltre, in virtù di un necessario approccio ecosistemico da tenere presente per una gestione sostenibile delle risorse marine, anche altri pesci e organismi marini non sfruttati commercialmente sono stati interessati da stime di abbondanza e distribuzione (Sainsbury *et al.*, 2000).

Sebbene le tecniche acustiche ai fini della stima della biomassa di pesce siano giunte a livelli molto sofisticati se si pensa a multi-frequenza, banda larga, *multibeam*, *echosounding* verticale e orizzontale (Georgakarakos *et al.*, 2011), esse si imbattono nel rilevante problema della identificazione della specie (D'Elia, 2012). I banchi di pesce sono aggregazioni di individui, presenti nella colonna d'acqua, in grado di restituire un'eco di ritorno del segnale acustico emesso dall'*echosounder*: nei risultanti ecogrammi (rappresentazioni bidimensionali) si vengono a formare delle ecotracce (aggregazioni coerenti) che contengono varie informazioni, come posizione, morfologia ed energia, e che risultano importanti per le stime di biomassa e densità (D'Elia, 2012). Come sopra accennato, per l'identificazione delle specie e per la definizione delle classi di taglia per specie è necessario effettuare campagne sperimentali con raccolta di diversi campioni. Tale metodo però, oltre a essere invasivo, risulta deficitario di accuratezza e obiettività, producendo stime che possono portare con sé un errore di classificazione delle ecotracce anche dell'80% (Simmonds & MacLennan, 2005). Poiché non è possibile raccogliere un campione per ogni banco visualizzato, ci si affida da una parte alle caratteristiche di similarità (per forma o profondità) tra tutti i banchi registrati nel corso del *survey* acustico e i soli banchi soggetti alla cala (Reid *et al.*, 2000), dall'altra all'esperienza di coloro che analizzano gli ecogrammi. L'opportunità di identificare le specie direttamente sulla base dei dati acustici apporta maggiore obiettività e accuratezza alle stime di biomassa e rende possibili studi più dettagliati su distribuzione e comportamento dei banchi. Al contempo, potrebbe fornire ai pescatori la possibilità di svolgere un'attività alieutica più selettiva che riesca ad individuare i banchi delle specie target, senza esercitare inutili pressioni su altre specie e sul cosiddetto "scarto" (*by-catch*), quindi sull'intero ecosistema e sulla biodiversità (D'Elia, 2012).

Il presente lavoro di tesi è volto a migliorare l'accuratezza, la precisione e la capacità discriminante del dato acustico; inoltre fornisce un esempio di ricerca per l'identificazione delle specie, la quale è definita come "la grande sfida dell'acustica applicata alla pesca" (MacLennan and Holliday, 1996) o come il "Santo Graal" per chi

lavora in questo ambito (Horne, 2000). Cercando di trovare un metodo basato sulle tecniche acustiche, efficace per l'identificazione delle specie in aree in cui sono presenti banchi relativi a più specie ittiche, esso si lega in modo diretto e complementare ad una ricerca che forniva invece strumenti atti all'identificazione della specie ittiche a partire da catture sperimentali monospecifiche (D'Elia, 2012) e dalla quale è stato implementato un classificatore *decision tree* (D'Elia *et al.*, 2014).

Poiché nel Mediterraneo è molto difficile individuare aree di pesca popolate da una sola specie, sarebbe utile considerare questo aspetto nella valutazione, tramite metodi specifici di raccolta dati e approcci multi-specie di valutazione (Lleonard and Maynou, 2003), e sviluppare delle misure di gestione integrate che tengano conto di tale caratteristica (Caddy, 1993).

1.1. Importanza della valutazione degli *stock* ittici

La valutazione delle dimensioni e dello stato degli *stock* ittici sfruttati dalla pesca rappresenta una delle sfide più importanti nella moderna gestione delle risorse alieutiche, importante anche per fornire idonee misure gestionali da rivolgere a *stakeholders*, industrie della pesca e pescatori. A tal fine risulta fondamentale l'osservazione diretta delle popolazioni ittiche, utilizzando strumenti elettronici specializzati di *remote sensing* come ecoscandagli e integratori eco. Inoltre, a partire da dati relativi alle catture, i ricercatori hanno sviluppato tecniche per valutare la natura, l'abbondanza, la distribuzione, la struttura e le dinamiche delle popolazioni ittiche. A seconda dei dati disponibili, gli *stock* talvolta vengono valutati utilizzando, contemporaneamente, due approcci alternativi: i metodi sintetici che utilizzano il rapporto teorico tra il livello di intensità della pesca e il totale delle catture; i metodi analitici che utilizzano la relazione tra reclutamento, crescita e mortalità. Entrambi gli approcci di valutazione dello stato di uno *stock* ittico partono dall'acquisizione di dati per determinare le dimensioni della cattura totale (FAO, 2014).

La pesca riduce l'abbondanza dello *stock* vergine, la durata della vita, la dimensione media e l'età della popolazione, con conseguente diminuzione delle catture per unità di sforzo (da qui in poi: $CPUE = Catch Per Unit of Effort$).

Attualmente sono notevoli i progressi fatti nell'ambito della valutazione delle risorse ittiche, sebbene nella maggior parte dei Paesi essi non siano ancora pienamente operativi.

Una nota della FAO cercava già nel 1983 di migliorare la reciproca comprensione tra gli amministratori e gli scienziati in merito alla valutazione degli *stock* ittici.

Per “valutazione degli *stock* ittici” si intende qualsiasi studio scientifico volto a determinare la produttività di una risorsa alieutica, gli effetti della pesca su tale risorsa, l’impatto sulla risorsa e sulle catture. Una buona valutazione degli *stock* dovrebbe portare alla formulazione di suggerimenti che conducano verso decisioni integrate di gestione.

Secondo la nota della FAO (Gulland, 1983), in relazione al grado di sfruttamento delle risorse ittiche, è necessario uno specifico modello di valutazione dello *stock*. I metodi di valutazione si basano solitamente su manuali pubblicati dalla FAO o da altri (ad esempio, Gulland, 1969; Ricker, 1958; Saville, 1977).

Esistono diversi metodi, come quello della stima del potenziale, che possono dare un’idea della probabilità di riscontrare determinati tipi di specie ittiche: attraverso studi si confrontano zone simili al fine di stimarne il potenziale totale.

Le specie e le quantità di pesce disponibili per la cattura dipendono dalla produzione biologica di base e quindi dalle caratteristiche dell’area (quali correnti, disponibilità di sostanze nutritive, ecc). Regioni simili avranno probabilmente una produttività biologica simile e simile potenziale di raccolta di pesce per unità di superficie. Osservazioni abbastanza semplici di una data area, combinate con informazioni sul potenziale di altre zone simili ma meglio conosciute, possono quindi dare informazioni importanti per il potenziale dell’area in questione. Questo metodo è stato più ampiamente utilizzato nei laghi (Henderson e Welcomme, 1974) ma è stato applicato anche ai fiumi (Welcomme, 1979) e agli ecosistemi marini, ad esempio ad aree di barriera corallina nei Caraibi (Munro, 1982). Se i metodi di pesca utilizzati sono abbastanza simili, questa tecnica comparativa può anche essere applicata per stimare il numero di pescatori o di imbarcazioni standard da impiegare nella pesca. La Figura 2 dimostra che non è vero che pescando tanto e in tanti le catture aumentano. Infatti, se anche lo sforzo di pesca aumenta, superato un determinato valore soglia le catture non aumentano più, anzi vanno diminuendo e lo *stock* ittico va verso il collasso compromettendo l’utilizzo futuro della risorsa. Ciò dimostra, da una parte, la necessità di una gestione volta ad ottimizzare l’utilizzo della risorsa di pesca abbattendone però i possibili impatti negativi e, dall’altra, la conseguente importanza di effettuare una valutazione degli *stock* ittici prima della stagione di pesca.

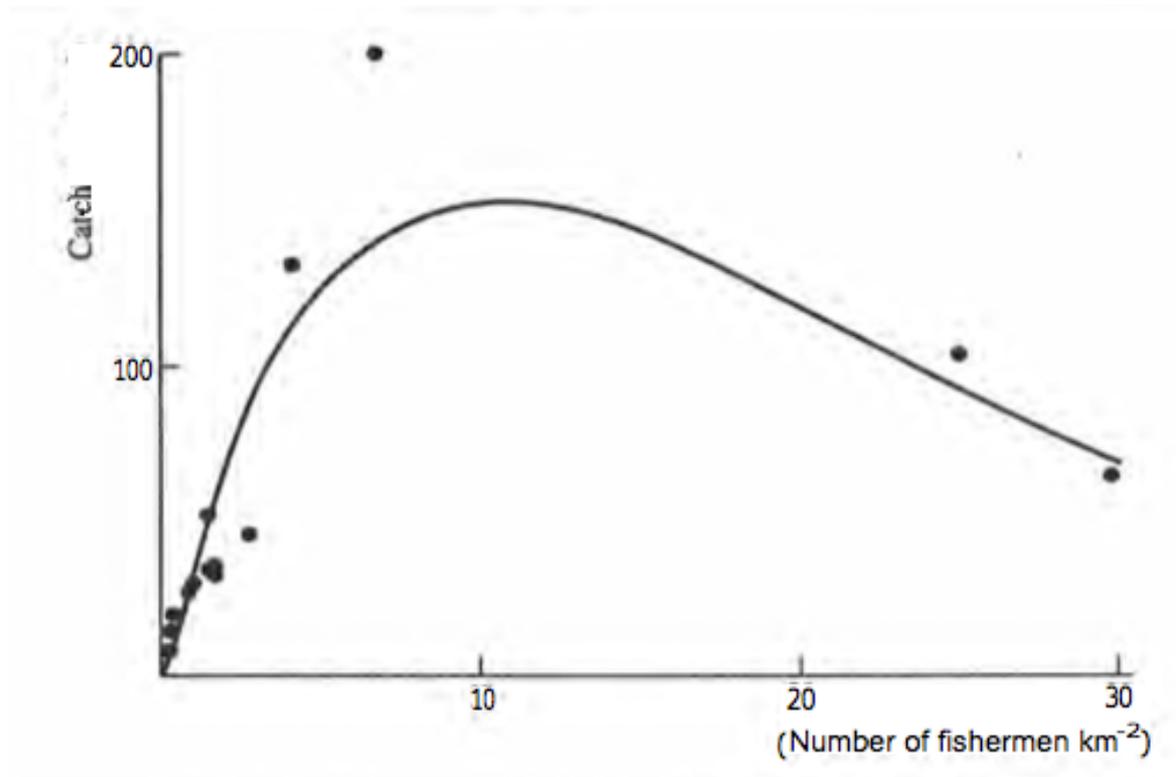


Figura 2. Per un certo numero di fiumi africani, la relazione tra la resa per unità di superficie della pianura alluvionale e il numero di pescatori per unità di superficie (Welcomme, 1979).

Con modelli di produzione si intende, invece, la classe dei metodi di analisi che considerano la relazione tra l'abbondanza di uno *stock* ittico, la quantità di attività di pesca (*effort*, cioè lo sforzo di pesca) e il totale delle catture. Questi modelli considerano lo *stock* come una massa uniforme e non tentano di affrontare la composizione dello *stock* (ad esempio, la percentuale di pesci vecchi e/o grandi).

I modelli analitici riconoscono che uno *stock* è costituito da un gran numero di individui: gli eventi nello *stock* possono quindi essere studiati considerando ciò che accade a un gruppo di pesci dal momento in cui sono abbastanza grandi per essere catturati (reclute), fino all'ultimo individuo, morto per cattura o per cause naturali.

La *multi-species fisheries* necessita di studi specialistici piuttosto complessi.

Appurata la necessità di effettuare la valutazione degli *stock* ittici, sarebbe opportuno che essa avvenisse passando attraverso dei metodi meno invasivi e di minore impatto nei confronti della risorsa. I metodi suggeriti dal presente lavoro di tesi hanno come obiettivi quello di causare meno impatti antropici riducendo il numero di peschate (cale) a scopo scientifico e quello di consentire l'acquisizione di una maggiore mole di dati, risparmiando tempo e mezzi.

1.2. Variabili ambientali per la valutazione degli *stock* ittici

I *Survey* di valutazione degli *stock* effettuati utilizzando esclusivamente attrezzi da pesca presuppongono che le catture per unità di sforzo (CPUE) siano proporzionali al grado di ricchezza della popolazione (Stoner, 2004). Tuttavia Punsly & Nakano (1992) sostengono che il comportamento delle specie *target* in risposta alle variabili ambientali potrebbe avere effetti maggiori sulla CPUE rispetto all'abbondanza di pesce. La variazione di CPUE è influenzata da due fattori: la disponibilità di pesce per gli attrezzi da pesca (vale a dire la distribuzione nel tempo e nello spazio) e la vulnerabilità a tali attrezzi (cioè la catturabilità) (Engås & Løkkeborg, 1994; Arreguín-Sánchez, 1996).

Poiché qualsiasi variabile ambientale influenza in generale il comportamento dei pesci, ed in particolare la loro locomozione e l'ecologia della loro alimentazione, e poiché proprio dal comportamento della specie *target* dipende l'efficacia degli attrezzi da pesca e quindi la catturabilità del pesce (in termini di CPUE), è importante prendere in considerazione anche queste variabili – oltre ai dati direttamente legati alle osservazioni dei banchi di pesce – per rendere più complete le stime di abbondanza e le valutazioni degli *stock ittici*.

Lo studio di Stoner (2004) si riferisce a metodologie di pesca che fanno uso di esche per il pesce, motivo per cui risultano di fondamentale importanza le variabili ambientali che possono influenzare in modo determinante la catturabilità del pesce. In generale, ciò che è descritto riguardo alla catturabilità e alle variabili ambientali può valere anche per metodologie di cattura che fanno uso di reti pelagiche – come quelle utilizzate per la presente tesi –, in quanto la presenza di una data specie di pesce in una data area può essere fortemente determinata dalle variabili ambientali.

Il sopra citato studio ha avuto lo scopo in prima istanza di identificare, valutare e classificare le variabili ambientali per quanto riguarda la loro probabile influenza sulla vulnerabilità dei pesci agli attrezzi da pesca (e quindi sulla loro catturabilità), in seconda istanza di determinare se possono essere realizzati o meno miglioramenti per le valutazioni degli *stock* attraverso la correzione della variazione ambientale. Stoner (2004) osserva che la variazione del comportamento alimentare legata all'ambiente può agire attraverso quattro diversi meccanismi: processi metabolici, limiti sensoriali, interazioni sociali e impatti diretti. La temperatura dell'acqua, il livello di luce, la velocità della corrente e la densità di prede naturali nell'ambiente possono avere effetti marcati sulla catturabilità del pesce, con potenziali ripercussioni sulla variazione in CPUE di un fattore dieci (Stoner, 2004). Poiché le specie *target* possono occupare nel tempo e nello spazio

vaste gamme di condizioni ambientali, la variazione spaziale e temporale per l'alimentazione potrebbe avere maggiori implicazioni sulla CPUE rispetto ai modelli di abbondanza. E' necessario comprendere meglio come le variabili ambientali influenzano le prestazioni di alimentazione correlate agli attrezzi da pesca. Combinando osservazioni sul campo ed esperimenti di laboratorio, si possono parametrizzare i modelli di valutazione degli *stock* sulla base di variazioni nel comportamento dei pesci. Lo studio considera come "spazio attivo" l'area campionata nella quale le esche sono presenti in concentrazioni superiori alla soglia di risposta per il pesce. La velocità di rilascio e trasporto di segnali chimici attraverso l'acqua e la soglia chemio-sensoriale per un individuo in un momento specifico (Løkkeborg *et al.*, 1995) sono altamente variabili e risultano molto importanti perché, anche in base ad essi, solo una certa percentuale di pesce di una determinata zona risponde agli attrezzi da pesca e solo alcuni pesci sono motivati a carpire le esche o ad entrare nelle trappole (Furevik, 1994; Løkkeborg, 1994; Kaimmer, 1999; Stoner, 2003). Questo studio sottolinea le più gravi potenziali limitazioni nella valutazioni degli standard per gli *stock* esca-dipendenti in modo da supportare la ricerca di base necessaria per migliorare i modelli di valutazione degli *stock*. Sempre Stoner (2004) afferma infine che, per una migliore gestione della pesca, i modelli di valutazione degli *stock* dovranno considerare le variazioni dei coefficienti di catturabilità causati dalla variazione spaziale e temporale delle variabili ambientali. In caso contrario, le CPUE serviranno solo per stimare l'abbondanza delle specie bersaglio, che può risultare molto confusa a causa della variabilità spazio-temporale nell'ecologia dell'alimentazione di quella specie e dei suoi competitori.

1.3. Acustica per i banchi di pesce: stato dell'arte

A partire dagli anni '80, la classificazione delle ecotracce per l'identificazione della specie ha iniziato ad essere oggetto di numerosi studi. L'approccio iniziale era basato su informazioni riguardanti il comportamento aggregativo, quali forma, taglia e distribuzione spaziale, finché non si è capita la necessità di adottare un secondo tipo di approccio, che tenesse conto del fatto che per ogni specie il comportamento aggregativo può variare (D'Elia, 2012), ad esempio a seconda della presenza o meno di altre specie nella medesima area geografica (Massè *et al.*, 1996) o a seconda della stagione o durante l'arco della giornata.

La prima tipologia di approccio ha portato per primi Rose & Leggett (1988) a sviluppare le prime classificazioni di banchi di pesce, sulla base di ecogrammi a singola frequenza, tramite identificazione delle tracce acustiche e utilizzando informazioni di forma e taglia delle singole aggregazioni. Altri tentativi di classificazione sono stati effettuati nel 1991 da Massè & Rouxel che hanno combinato informazioni relative a catture pelagiche con l'uso del software MOVIES-B (Weill *et al.*, 1993). Tra le cale monospecifiche di alcuni *survey* effettuati tra il 1989 e il 1993, alcune ecotracce sono state selezionate e studiate da Scalabrin *et al.* (1996) e più di 7000 banchi sono stati caratterizzati da una ventina di parametri descrittivi di morfologia, batimetria ed energia: questi permettevano la distinzione di *Engraulis encrasicolus*, *Sardina pilchardus* e *Micromesistius poutassou*, ma non di *Trachurus trachurus*. Il secondo approccio ha aperto le porte ad un'altra tipologia di studi che integrano descrittori dell'ambiente e descrittori del banco: tali studi pongono la distribuzione dei banchi di pesce in relazione ad esempio alla profondità del termocline (Swartzmann *et al.*, 1994), alla temperatura e alla salinità superficiale (Walsh *et al.*, 1995; Castillo *et al.*, 1996), al tipo di fondale (D'Elia *et al.*, 2009), alla densità di cibo (Nøttestad *et al.*, 1996; Patti *et al.*, 2011). Volendo descrivere questo approccio un po' più nel dettaglio, si può far riferimento allo studio condotto da Tugores *et al.* (2011) i quali, per primi, hanno tentato di costruire dei modelli di differenti stadi vitali (in questo caso, uova e adulti di sardine) e di ampliare le conoscenze riguardo all'influenza delle variabili ambientali sulla loro distribuzione spaziale. Associando i dati acustici relativi alle sardine - reperiti nell'arco di 5 anni (2003-2008) presso quattro aree del mar Mediterraneo - a dati di variazione ambientali (soprattutto la SST: *Sea Surface Temperature*) e batimetrici, sono stati realizzati dei modelli di habitat potenziale di sardine durante alcune stagioni dell'anno; questi modelli sono risultati utili in termini predittivi di presenza-assenza e per la costruzione di mappe annuali (tramite *software arc-GIS*) di adeguatezza di habitat per la presenza di sardine, da tenere presenti per efficaci azioni di gestione (*habitat suitability maps*).

L'uso di tecniche statistiche per analizzare le relazioni tra comportamento aggregativo e distribuzione spaziale ha cominciato a prendere piede a partire dal progetto "CLUSTER (FAIR-CT-96.1799)" approvato tra il 1997 e il 1999, il quale ha avuto come oggetto tutti gli *stock* ittici europei; esso si è basato su serie storiche di ecotracce e ha esaminato le variazioni interannuali in relazione a più parametri. Inoltre va considerata la contestuale ricerca di nuove procedure di analisi degli ecogrammi e la stesura del protocollo per la definizione di parametri caratterizzanti un banco (Reid *et al.*, 2000). Le tecniche

statistiche di cui sopra – finalizzate all'identificazione della specie ed allo studio della distribuzione spaziale dei banchi – sono essenzialmente tecniche di *clustering*, analisi discriminanti (Scalabrin *et al.*, 1996; Lawson *et al.*, 2001), *neural network* (Simmonds *et al.*, 1996); per analizzare invece le relazioni tra aggregazioni e fattori esterni si fa riferimento ai GAMs (*Generalized Additive Models*) (Agenbag *et al.* 2003; Swartzmann *et al.*, 2008). Questi studi hanno identificato alcune regole generali valide annualmente e per differenti *stock*: il numero di individui del banco è direttamente proporzionale alla dimensione geometrica dello stesso, mentre il numero di individui del banco è inversamente proporzionale al numero di banchi vicini; l'orario influenza il numero di banchi e le loro caratteristiche; esiste una relazione tra numero totale di banchi e organizzazione spaziale delle aggregazioni, cioè più elevato è il numero dei banchi più essi sono piccoli e organizzati in più *cluster* (D'Elia, 2012).

A seconda che le ecotracce siano state acquisite da un *echosounder* a singola frequenza (Scalabrin *et al.*, 1996) o a più frequenze (Madureira *et al.*, 1993; Kloser *et al.*, 2002) o da *broad band echosounders* (Simmonds *et al.*, 1996; Lundgren & Nielsen, 2008), esse vengono classificate ed identificate in base alle informazioni acquisite (D'Elia, 2012). Le ecotracce a singola frequenza sono adatte più che altro ad ambienti monospecifici (MacLennan & Holliday, 1996), mentre quelle multifrequenza forniscono più informazioni acustiche per ogni ecotraccia e quindi maggiore accuratezza delle informazioni biologiche acquisite (Madureira *et al.*, 1993). Il terzo è un metodo basato su una banda larga che permette di discriminare meglio nel segnale di ritorno; il *broad band echosounder* equivale, infatti, all'utilizzo contemporaneo di un centinaio di sonar a singola frequenza (Øvergaard & Skjøsberg, 2012).

Si ritiene che le tecniche che si basano sulle differenze nelle proprietà di retrodiffusione (*backscattering* = eco di ritorno; vedi paragrafo 2.1) delle varie frequenze siano fondamentali per l'identificazione della specie (Kang *et al.*, 2002; Kloser *et al.*, 2002; Watkins & Brierley, 2002) anche perché rendono meno soggettiva l'interpretazione dell'ecogramma.

Anche alla classificazione delle ecotracce isolate sono state applicate tecniche statistiche, quali analisi delle componenti principali e analisi discriminante, *artificial neurale networks*, *nearest-neighbour analysis* (Richards *et al.*, 1991), *k-mean clustering* (Tegoswski *et al.*, 2003) o un insieme di questi modelli (Fleischman & Burwen, 2003): molti sono buoni classificatori ma difficilmente applicabili, mentre uno dei metodi più utili parrebbe essere quello del *classification tree* (Fernandes, 2009).

L'identificazione del segnale acustico, proveniente da diversi gruppi di taglia dello zooplancton - alla base dei modelli di *scattering* usati per la ricerca nell'acustica della pesca e del plancton - ha fatto nascere una nuova idea per la discriminazione in specie a seconda del tipo di dati acustici a multifrequenza (le frequenze impiegate per la pesca sono 18, 38, 120, 200 kHz): secondo tali modelli alcuni *target*, come il plancton, si trovano in una regione in cui il *backscattering* varia significativamente con la frequenza (regione di *Rayleigh*), mentre altri (come i pesci dotati di vescica natatoria) nello stesso *range* di frequenze danno un *backscattering* geometrico cioè costante al variare della frequenza (Lavery *et al.*, 2007). Per i pesci il modello più usato è quello che descrive il comportamento della vescica natatoria, anche se alcuni modelli più recenti (Fassler *et al.*, 2007) combinano l'effetto di questa con quello del corpo: a certe profondità, il corpo del pesce determina un comportamento risonante che, alle varie frequenze, dipende dalla taglia del pesce. Ad esempio, ad una profondità di 112 m la compressione della vescica natatoria porta il corpo di un'aringa atlantica di 20-25 cm ad avere una energia totale di *backscatter* predominante su quello della vescica natatoria (D'Elia, 2012). Per esempio, in Figura 3, la linea continua spessa riporta i risultati della forza dell'eco di ritorno del modello derivato dalla combinazione dell'effetto della vescica natatoria con l'effetto del corpo; la linea tratteggiata raffigura i risultati del modello usato per rappresentare l'effetto della vescica natatoria; la linea continua sottile raffigura quelli del modello usato per rappresentare l'effetto del corpo del pesce (Fässler *et al.*, 2007).

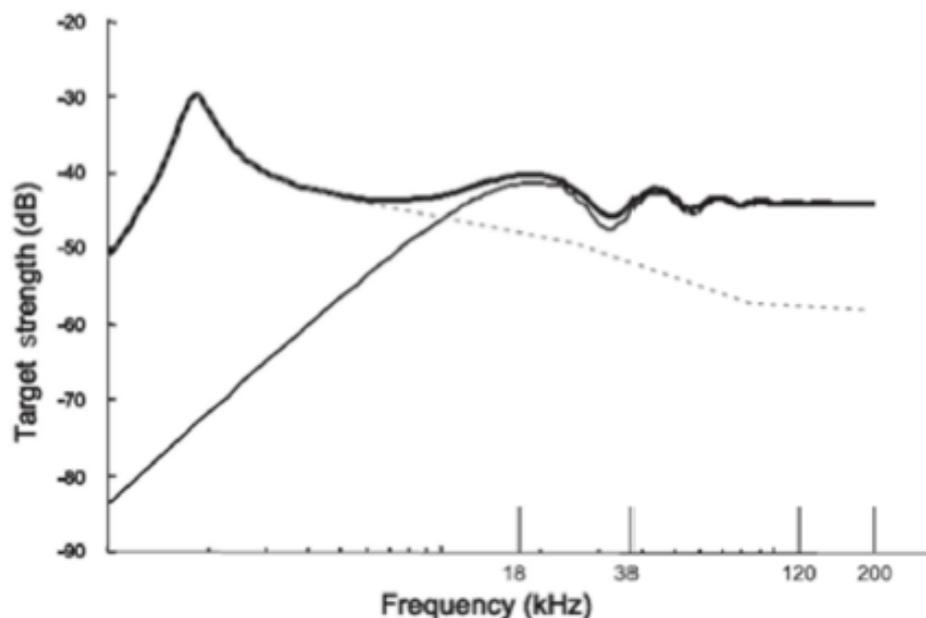


Figura 3. Modello predittivo della *backscatter strength* di un'aringa Atlantica di 20-25 cm di taglia ad una profondità di 112 m. (Fässler *et al.*, 2007).

1.3.1. Accuratezza di metodi acustici in *fish stock assessment surveys*

Secondo Bertrand (2003) e Koslow (2009), le tecniche acustiche per la stima della biomassa di pesce risultano di rilevante importanza poiché accrescono le conoscenze, non solo riguardo agli *stock* ittici in sé, ma anche riguardo alla struttura dell'ecosistema.

I metodi acustici adottati per la valutazione degli *stock* di pesce pelagico durante i *survey* sono infatti risultati relativamente più affidabili di altri (Simmonds, 2003): essi sono caratterizzati da una buona accuratezza nel definire l'abbondanza del banco e dalla procedura di “*intrinsic error analysis*” la quale fa una buona analisi d'errore (stima l'errore totale, le fonti di errore, la natura *random* o sistematica e la grandezza di ciascun errore) e applica un approccio di *de-biasing* (cioè minimizza l'errore, ove possibile).

L'accuratezza nel presente contesto viene trattata in modo ampio e interessante ai nostri fini, nello studio “*Accuracy of Acoustic Methods in Fish Stock Assessment Surveys*” (Georgakarakos *et al.*, 2011), motivo per cui si è scelto di entrare un po' più nel dettaglio. Lo studio in questione esamina alcune delle più importanti fonti di errore e il loro impatto sulla stima acustica della biomassa, focalizzandosi sulle specie pelagiche. Capire quanto errore ci si può aspettare in un *survey* acustico standard è possibile grazie agli indicatori proposti da Simmond & MacLennan (2005).

Lo studio distingue gli errori in due gruppi: quelli legati alla strumentazione e quelli dovuti alla complessità del comportamento dei pesci.

Le tre più importanti fonti di errore messe a fuoco sono:

- 1) la stima del TS (*Target Strength*), che è una misura del coefficiente di riflessione del target, misurato in decibel;
- 2) l'identificazione di specie e l'allocazione della biomassa per ogni specie;
- 3) le reazioni di “evitamento” (*avoiding*) da parte del pesce nei confronti delle navi da *survey*.

1) STIMA DELLA FORZA TARGET (*Target Strength*)

Nello studio in esame si afferma che ogni approccio di monitoraggio di abbondanza degli *stock* di pesce deve avere, come prerequisito, la stima priva di *bias* del TS, poiché tale stima è un parametro chiave che non fa altro che permettere di trasformare il *backscattering* acustico in biomassa.

Ciascun pesce presenta un particolare TS che dipende dalla frequenza del suono e da alcune caratteristiche del pesce come taglia, anatomia, fisiologia, comportamento nel nuoto e in generale alcuni adattamenti morfologici causati da fattori ambientali.

Da ciò deriva la necessità, sottolineata dagli autori dello studio, di attuare da un lato *surveys* di stima della biomassa, senza però trascurare dall'altro una elaborazione delle relazioni multivariate tra TS e i suoi possibili parametri di regressione.

Ad oggi l'efficienza del campionamento biologico, cioè quella attraverso le cale, è ancora maggiore rispetto a quella fornita dalle tecniche acustiche. Queste ultime stanno però acquisendo sempre maggior rilevanza, grazie agli *echosounder* moderni che determinano all'interno del fascio acustico posizione e taglia dei *target* singoli: gli algoritmi “*single target detector*” riescono a filtrare echi multipli dall'eco dell'intera popolazione. Essi funzionano bene, però, solo in condizioni di densità molto basse di pesce (Ehrenberg & Torkelson, 1996; Simmonds & MacLennan, 2005).

Per i piccoli pesci pelagici il rapporto segnale/rumore è più basso (ciò chiaramente implica che il rumore sia maggiore), per cui il *bias* della stima del TS a causa della presenza di echi multipli di ritorno può essere molto alto. Anche il plancton che si trova in mezzo ai pesci di un banco produce degli echi e per ridurlo viene utilizzata una procedura che tiene conto di una soglia minima di *backscattering* (*thresholding*). Purtroppo questa tecnica introduce un *bias* alla distribuzione del TS, soprattutto per pesci di piccole dimensioni come appunto i piccoli pelagici (Simmonds & MacLennan, 2005).

Gli algoritmi di monitoraggio isolano gli echi del singolo pesce e forniscono stime di punti di TS, nonché un'ulteriore informazione riguardante il pattern di nuoto. Infatti, in base al tipo di nuoto, un pesce produce una particolare energia di *backscattering*: essa non è altro che una quantità dinamica che dipende molto dal comportamento del pesce, specialmente dall'angolo di inclinazione del nuoto, cioè dal cosiddetto “*swim tilt angle*” (Huse & Ona, 1996). Per stimare un pattern medio di nuoto è necessario tenere presente che, a seconda del comportamento nel nuoto di un pesce, si ha una particolare inclinazione del corpo e quindi una particolare eco di ritorno: ecco perché le stime del TS vanno fatte su pesci selvatici e non su pesci confinati in un'area.

Ad esempio, per il progetto di ricerca “FISHCAL”, finalizzato alla stima della TS delle 3 specie commerciali più importanti (*Sardina pilchardus*, *Engraulis encrasicolus* e *Trachurus trachurus*), sono stati registrati i dati acustici nel Thermaikos Gulf (nord-Egeo) e si è impiegata una rete pelagica francese modificata per l'identificazione di specie e le

misurazioni di lunghezza del pesce. Per le misurazioni idroacustiche, è stato impiegato un *echosounder* con dei trasduttori circolari.

Gli algoritmi che rilevano la singola eco (SED, *Single-Echo-Detection*) non sono capaci di restituire tutti i *target* multipli, per cui risultavano necessari alcuni procedimenti di *post-processing* tramite l'implementazione di filtri di rilevamento.

Quando si attua un monitoraggio che fa uso di tecniche acustiche, è necessario tenere presenti le limitazioni alle quali si va incontro: si prediligono i *target* più grandi poiché maggiore è la dimensione del pesce, più alto è il rapporto segnale/rumore; si ha cioè minore rumore per un maggiore volume di campionamento, specialmente nelle acque più profonde. Purtroppo invece, gli echi più bassi dei *target* più piccoli hanno meno possibilità di essere rilevati in *ping* successivi (per *ping* si intende il segnale ad alta energia acustica emesso con cadenza regolare dall'*echosounder*) e sono sottostimati dal monitoraggio.

Il TS di un pesce con vescica natatoria è legato alla sezione trasversale della medesima ed è profondità dipendente, secondo la legge di Boyle (Edwards & Armstrong, 1983), quindi bisogna prestare una buona dose di attenzione nel confrontare echi provenienti da differenti profondità. Non è chiaro se responsabile della riduzione del TS sia la sola compressione della vescica natatoria o se, in più, il comportamento di nuoto del pesce cambi con la profondità.

2) IDENTIFICAZIONE DI SPECIE

Tradizionalmente nei *survey* acustici per la stima della biomassa ittica, come già precedentemente evidenziato, il *backscattering* acustico viene usato in affiancamento a campioni di cala; ma questa procedura impiega molto del tempo disponibile in nave, ha un notevole aspetto di soggettività e non è utilizzabile nella riduzione del *bias* del segnale.

Negli anni '90 sono stati sviluppati dei *software* specifici, come ad esempio gli *Artificial Neural Networks* (ANNs) che rendono più obiettivo l'approccio per l'identificazione della specie, attraverso un'estrazione automatica dei parametri di un banco. Inoltre, sono stati sviluppati algoritmi più veloci di scrutinamento di un banco (Reid *et al.*, 2000; Fernandes *et al.*, 2002), descrittori multifrequenza (Korneliussen & Ona, 2002; Korneliussen *et al.*, 2009) o multidimensionali (Paramo *et al.*, 2007; Trygonis *et al.*, 2009) e ANNs avanzati (Cabreira *et al.*, 2009).

In particolare, in questo lavoro, vengono confrontate le tecniche sopra elencate con tecniche avanzate di classificazione, basate su *Bayesian Neural Network*. I dati sono stati

acquisiti in due aree geografiche: Thermaikos Gulf (nord-eggeo) e Mar del Nord (Shetland Island e Norvegia). Gli ecogrammi del banco hanno subito validazione incrociata con campioni di cala e le immagini dei banchi sono state caratterizzate con un attributo di identificazione di una specie.

Nel *layer* di output ogni unità corrisponde ad una certa specie, assumendo il valore di 1 se il banco appartiene a questa categoria e 0 altrimenti.

Per quanto riguarda i risultati dello studio, le specie *target* sono state previste applicando ANNs e metodi di *Discriminant Function Analysis* (DFA).

Per entrambi i *data-set*, i descrittori batimetrici sembravano essere più importanti di quelli energetici, mentre questi ultimi erano più importanti dei descrittori morfometrici.

I *network* Bayesiani hanno incrementato la loro performance nel riconoscimento dei banchi delle stesse specie, ma riscontrate in differenti aree di *survey*. Questo è importante, considerato che i banchi sono stati insonificati a diverse profondità e che i loro descrittori morfologici ed energetici sono stati distorti.

Le classificazioni sul set di test erano sopra il 97%; la performance era perfetta (100%) per la classificazione di aringa e merluzzo norvegese, cioè le 2 specie che avevano abbastanza rappresentanti nel *training test* (45,8% e 34,3% dei casi rispettivamente).

In un secondo set sperimentale, questo *network* già formato e validato con dati raccolti dalle Shetland Islands è stato applicato a differenti data-set di banchi raccolti in Norvegia. Per cui questo test ha funzionato come misura di “*generalisation power*” del *Network* Bayesiano, cioè per prevedere i banchi di aringa sconosciuti in Norvegia.

Quando è stato utilizzato un ANN non-Bayesiano, la percentuale di banchi di aringa riconosciuti correttamente scendeva a circa il 60%. L’alto tasso di previsione (88,3%) del *Bayesian Neural Network* riguardante i banchi di aringa di un’altra area di *survey* sottolinea le sue buone proprietà di generalizzazione. Una importante differenza tra i dati raccolti in 2 diverse aree era la profondità media del banco. Portando i banchi Norvegesi a profondità simili a quelle delle Shetlands, il punteggio di previsione delle aringhe è migliorato dall’88,3 al 97,8%, quindi i banchi individuati più in profondità possono essere meglio riconosciuti da *network* neurali.

In conclusione, la metodologia ha dimostrato che è possibile identificare la specie in un banco, grazie a descrittori specifici di estrazione da ecogrammi, generati da trasduttori *single-beam* (singolo fascio) e a banda stretta.

3) REAZIONI DI EVITAMENTO DA PARTE DEL PESCE NEI CONFRONTI DELLE NAVI DA *SURVEY*

Studi abbastanza recenti (Olsen, 1990; Aglen, 1994; Misund, 1997; Mitson & Knudsen, 2003) hanno dimostrato che non è assolutamente vera l'assunzione sulla quale si basavano i metodi acustici di *stock fish assessment*, secondo la quale le navi da *survey* non turbano il comportamento del pesce.

E' invece vero che il rumore indotto dalle navi e dagli attrezzi da pesca (*fishing-gear*) determina una reazione nei pesci (Ona & Godo, 1990; Gerlotto & Frèon, 1992), che dipende dalla specie in questione, dal suo comportamento alimentare, dalla migrazione, da stimoli visivi e auditivi, dalle condizioni idrografiche, e dalla superficie dell'acqua e interazioni con il fondo (Frèon *et al.*, 1993; Levenez *et al.*, 1990; Misund 1990).

Il comportamento del pesce causato dalle navi da *survey* è noto in acustica come "*fish avoidance effect*": si ha cioè un iniziale movimento verticale del pesce verso il fondo ed un successivo trasferimento orizzontale perpendicolare al corso delle navi. Ciò comporta l'alterazione del TS a causa della compressione della vescica natatoria (legge di Boyle), cambiamenti nell'angolo di inclinazione, riduzione della densità nel *beam* acustico.

I sonar *multibeam* (multi-fascio) lavorano su ampi volumi, per cui risultano potenti strumenti per studiare le reazioni di un banco di pesce a navi da ricerca, a strumenti di pesca e/o a predatori.

Diner e Masse (1987) hanno usato sistemi sonar per registrare le reazioni dei banchi di clupeidi indotte dalla presenza di una nave da ricerca e Hafsteinsson e Misund (1995) hanno riportato differenze specie-specifiche: quasi il 20% dei banchi di aringhe riscontrati nel Mar del Nord hanno mostrato qualche grado di reazione alla nave in avvicinamento, ma questo comportamento non è stato osservato nella migrazione del "capelin" (*Mallotus villosus*) nel Barent Sea. Inoltre, una modificazione scientifica dei sonar commerciali ha fornito misure quantitative di struttura interna (Gerlotto *et al.* 2000), tipologia (Paramo *et al.* 2007) e comportamento spaziale di un banco di pesce.

I dati acustici sono stati registrati durante il *survey* AVITIS-98 in Grecia (Thermaikos Gulf, aprile-maggio 1998) usando un sonar *multibeam* a 450 kHz.

I banchi incontrati sono stati esaminati in successive sezioni trasversali 2D verticali e i coefficienti di *backscattering* del volume acquisito sono stati registrati e analizzati *off-line* (cioè in laboratorio e non contemporaneamente alla fase di acquisizione). Veniva usato un *software* interno per stimare i descrittori dei banchi morfologici ed energetici in 2D e 3D. La maggior parte dei banchi osservati mostrava il primo bordo anteriore più vicino alla

nave rispetto a quelli successivi e al bordo finale. Questa distanza che cresce indicava un movimento medio del banco di 0.2 m/s perpendicolare alla nave. Si è anche osservato un simile incremento di profondità. Usando insieme *echosounder* e osservazioni di sonar, Soria *et al.* (2003) hanno formulato un modello di comportamento per i banchi di acciughe e sardine nel Mediterraneo e hanno dedotto che le variazioni in lunghezza del banco nei dati da *echosounder* sono inizialmente legati alla reazione di “evitamento” di fronte alla nave da ricerca, mentre i fattori ambientali come la profondità del termoclino e dell'alocline possono avere anche un effetto successivo.

In conclusione, sviluppi tecnologici recenti come *hardware* di sonar ad alta risoluzione con output calibrati (Trenkel *et al.*, 2008) o sistemi di *software* operazionali che possono quantitativamente processare una buona parte di dati grezzi *multibeam* (Balabanian *et al.*, 2007; Trygonis *et al.*, 2009) assicurano evidenze nell'effetto di evitamento del pesce e misurazioni *in situ* di dinamiche di aggregazione di specie pelagiche.

1.4. Capo Granitola

La valutazione dello stato delle risorse ittiche pelagiche, volta a stimare la loro distribuzione e abbondanza nello Stretto di Sicilia, è oggetto dal 1998 di campagne di ricerca tramite *survey* acustici (Patti *et al.*, 2004) condotti dall'IAMC-CNR (*Istituto per l'Ambiente Marino Costiero - Consiglio Nazionale delle Ricerche*, presso il quale sono stata ospitata per il lavoro di tesi) di Capo Granitola. Tra le specie pelagiche dello Stretto di Sicilia, acciughe (*Engraulis encrasicolus*) e sardine (*Sardina pilchardus*) rivestono un ruolo di centrale importanza, dal momento che sono le due specie di piccoli pesci pelagici più rilevanti dal punto di vista economico e più abbondanti negli sbarchi di Sciacca, il maggior porto per questa pesca nello Stretto di Sicilia (Mazzola *et al.* 2002). Bisogna anche tenere presente che queste due specie sono fondamentali sia per i pescatori locali - costituendo più del 90% dello sbarcato di piccoli pelagici (Patti *et al.*, 2004) - sia nell'ambito dell'industria della pesca e di quella conserviera nazionali (Alvares *et al.*, 2003): le acciughe ricoprono il 26 % della pesca marina complessiva del Mediterraneo, mentre le sardine il 9% (dati IREPA 2010). Secondo i dati IREPA 2009, le catture di piccoli pelagici hanno subito importanti riduzioni: la variabilità ambientale da un lato e l'aumento dello sforzo di pesca dall'altro danneggiano gli *stock* di piccoli pelagici, riducendone la biomassa e causandone un pericolo di collasso (D'Elia, 2012). E' dunque

necessario continuare a svolgere *survey* acustici per la valutazione della biomassa e dello stato degli *stock* di piccoli pelagici, ai fini di un'attenta pratica di gestione nazionale e regionale che tenga ben presenti le indicazioni suggerite dagli esperti di “*stock assessment*” e dalle linee guida dell’“*Ecosystem approach to fisheries*” (Fao, 2003).

L'interpretazione degli ecogrammi acquisiti nel Canale di Sicilia, principalmente basata sull'esperienza di analisi, associata alle informazioni biologiche delle catture sperimentali, introducono un certo grado di soggettività alle stime di biomassa di acciughe e sardine.

Un sistema capace di apportare maggiore obiettività sarebbe determinante nell'accrescere l'accuratezza delle stime di biomassa e la comprensione delle dinamiche di popolazione di queste specie, ai fini di una gestione sostenibile della risorsa. Il lavoro di Marta D'Elia (2012) e quello sviluppato nel presente lavoro di tesi, sulla scia del primo, sono proprio finalizzati al conseguimento di una maggiore obiettività nella classificazione dei banchi di pesci.

2. MATERIALI E METODI

2.1. *Sonar ed echosounder*

Poiché le strumentazioni utilizzate ai fini della raccolta dei dati per il presente studio sono stati principalmente *sonar* ed *echosounder* acustici, appare necessario descriverne a grandi linee il funzionamento.

Il termine **SONAR**, acronimo di *SOund Navigation And Ranging* che sta a significare “navigazione e misurazione di distanza per mezzo del suono”, è una tecnica di propagazione del suono molto usata in mare. Si tratta di un localizzatore a ultrasuoni che veniva utilizzato durante la seconda guerra mondiale dalla marina militare americana per individuare i sommergibili nemici; il termine ha poi assunto un significato più generico che indica i sistemi di ricerca subacquea basati sull’emissione di suoni o ultrasuoni e sulla captazione di echi riflessi dagli oggetti incontrati, come fondo marino e corpi immersi.

Si distinguono due categorie di *sonar*: quelli passivi e quelli attivi. I *sonar* attivi (Figura 4) si basano sulla ricetrasmisione di impulsi ultrasonori, mentre i *sonar* passivi sono tali in quanto non trasmettono impulsi ma si basano solo sulla ricezione. Un *sonar* passivo non è altro che un sistema idrofonico costituito da uno o più sensori che captano i suoni trasmessi dagli oggetti sommersi, permettendo di localizzarne la direzione e la fonte nonché di definirne la distanza. I *sonar* attivi funzionano invece secondo il principio dei *RADAR* (*Radio Detection And Ranging* ossia “individuazione e misurazione di distanza via radio”), infatti permettono di localizzare i corpi sommersi tramite la rivelazione degli echi restituiti dagli oggetti che hanno ricevuto impulsi sonori o ultrasonori emessi da un trasduttore. Quest’ultimo emette infatti un segnale ad alta energia acustica detto *ping*, mentre gli oggetti investiti dagli impulsi sonori riflettono il suono, per cui l’onda riflessa (eco) viene raccolta dai sensori. La misurazione del tempo che intercorre tra l’emissione del segnale acustico e il raggiungimento dell’emettitore da parte dell’onda permette di misurare la distanza (www.treccani.it).

Per quanto riguarda i sistemi **echosounder**, funzionano tutti più o meno allo stesso modo e si tratta di apparecchi che permettono di individuare oggetti sotto la superficie delle acque, fornendo informazioni circa profondità e forma del fondo e di eventuali pesci presenti. Un ecoscandaglio è sostanzialmente un ricetrasmittitore che genera impulsi elettrici e li invia al trasduttore, il quale converte gli impulsi in onde sonore. Queste, rimbalzando contro gli oggetti sotto la barca, producono degli echi che tornano al

trasduttore, il quale converte l'energia del suono dell'eco in impulso elettrico che viene restituito all'ecoscandaglio. L'ecoscandaglio quindi misura il tempo intercorso tra la trasmissione di un impulso sonoro e la ricezione della sua eco, per calcolare la distanza dell'oggetto. A questo punto il microprocessore legge queste informazioni, le elabora e le mostra sullo schermo, in modo che vengano rappresentati su di esso gli oggetti e/o la posizione dei pesci. Si può quindi parlare di sistema-ecoscandaglio, il cui cuore è il trasduttore perché è proprio questo che trasforma gli impulsi elettrici in sonori e viceversa; è cioè il dispositivo che trasmette le onde sonore e riceve gli echi, permettendo allo scandaglio di interpretare, di leggere ciò che si trova sott'acqua. Il sistema-ecoscandaglio interpreta dunque il risultato e mostra la profondità del fondo. E' importante tenere presente che gli echi di alcune onde sonore sono più veloci di altri e che esiste una drastica differenza di densità tra l'aria e il resto del corpo di un pesce: questo concetto spiega come sia possibile che il trasduttore veda un pesce grazie all'aria presente nella sua vescica natatoria. Questa differenza di densità fa sì che le onde sonore emesse dall'ecoscandaglio rimbalzino (*backscattering*) in modo differente: il trasduttore riceve gli echi, mentre l'ecoscandaglio riconosce queste differenze e visualizza tutto questo come pesce (www.airmartechology.com).

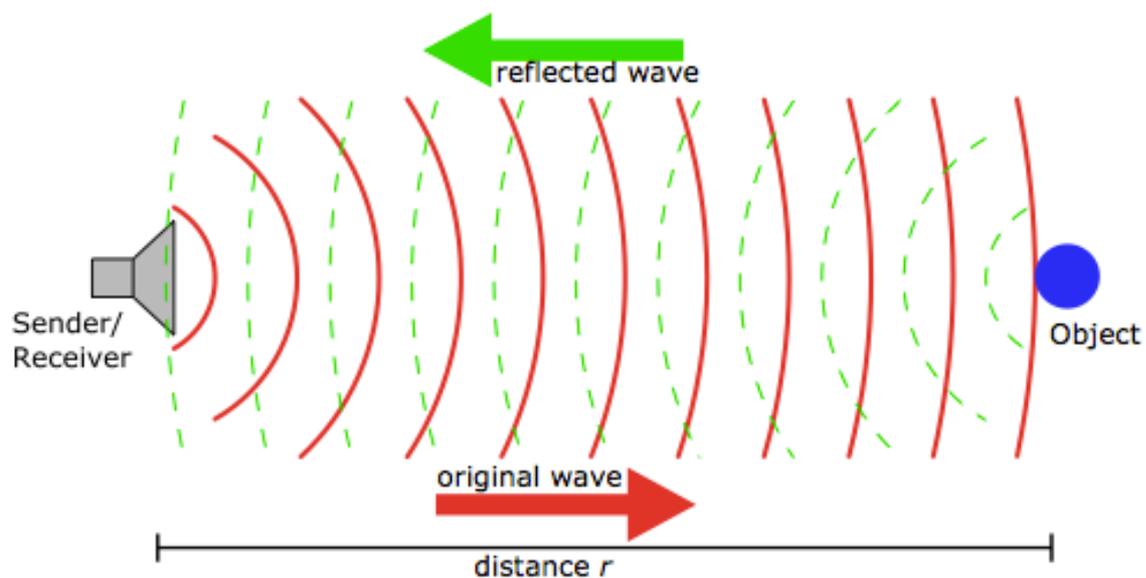


Figura 4. Principio di funzionamento del sonar attivo (Bozma & Kuc, 1991).

2.2. Trattamento di dati a singola frequenza per l'estrazione di banchi di pesce e delle aggregazioni planctoniche

Si ritiene necessario spiegare nel dettaglio sia le procedure di acquisizione, sia quelle di elaborazione dei dati impiegate nel lavoro di Marta D'Elia (2012), in quanto entrambe rappresentano il substrato metodologico indispensabile sul quale si è sviluppato il presente lavoro di tesi.

L'area di studio, rappresentata dalla porzione del Canale di Sicilia prospiciente la costa meridionale della Sicilia, è stata suddivisa in tre zone (Figura 5). Nell'ambito di essa, durante le campagne oceanografiche effettuate con la nave "Dallaporta" (Figura 6) dell'IAMC – CNR, è stato eseguito un tracciato acustico lungo i transetti perpendicolari alla costa.

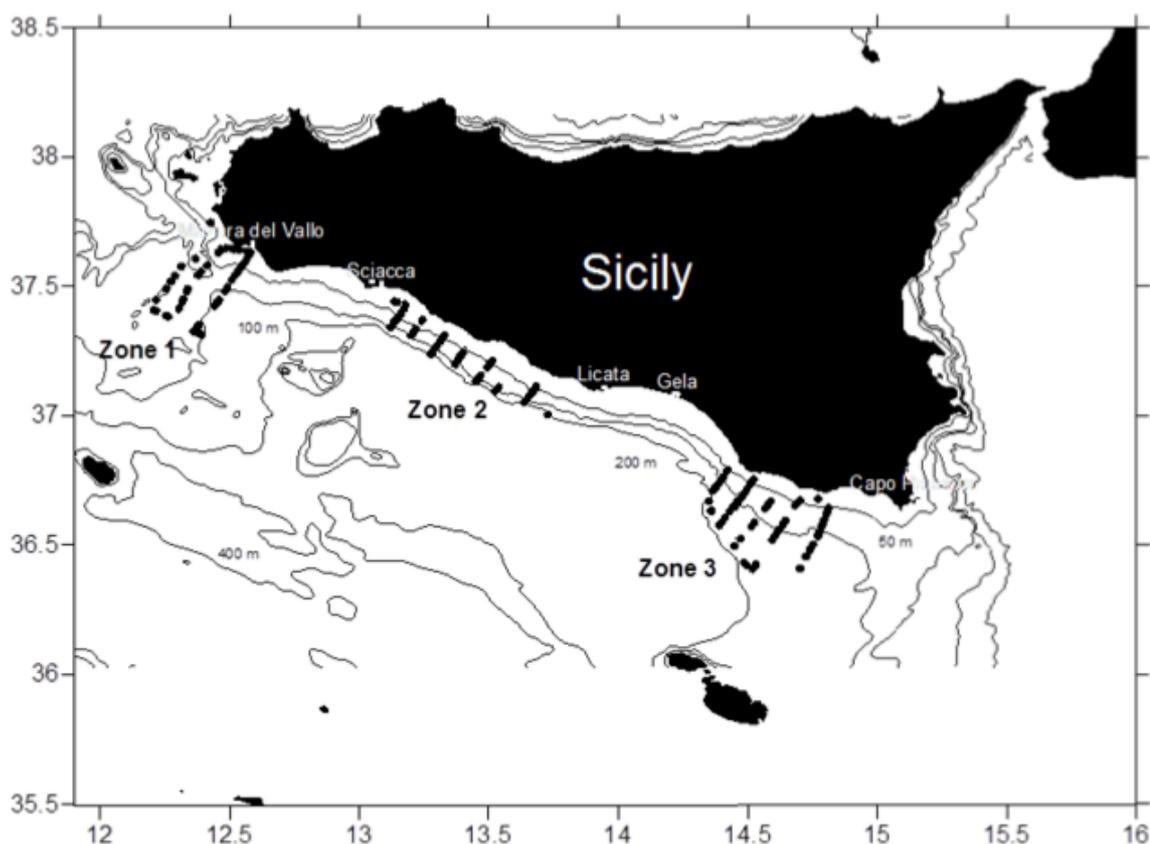


Figura 5. Area di studio. Mostra i siti, lungo il tracciato acustico giornaliero del survey, in cui banchi di pesce e plancton sono compresenti. E' riportata anche la separazione dell'area di studio in tre zone, indicate come Zona 1, Zona 2, Zona 3 (Patti et al., 2011).



Figura 6. Nave oceanografica *Dallaporta* (a, b).

Sulla base dei dati acustici acquisiti alle frequenze di 38 e 120 kHz, grazie all'utilizzo dell'*echosounder* scientifico della SIMRAD EK500, sono state ricavate contemporaneamente informazioni sui banchi di pesce e sulle aggregazioni planctoniche. Di tutti i transetti del tracciato acustico giornaliero del *survey* sono stati selezionati solo quelli diurni, in quanto durante il giorno le specie di piccoli pesci pelagici formano dei banchi.

I dati acustici, estratti tramite il *software Echoview V.3* e convertiti in formato ASCII, sono stati sottoposti ad algoritmi di analisi di immagine in Matlab: in pratica, dagli ecogrammi di entrambe le frequenze e sulla base dell'energia di *backscattering*, è stato possibile individuare e delimitare *patch* di plancton e banchi di pesce.

Per l'identificazione di *patch* acustiche identificabili come banchi di pesce ed esclusivamente agli ecogrammi a 38 kHz è stata applicata la procedura riportata in Patti *et al.* (2011): essa utilizza come soglia dell'area minima i 4 m² e un filtro nell'intervallo [-58, -36 dB]; tramite il filtro, tutte le aggregazioni poco dense interpretate come informazione di banchi presenti nella metà superiore della colonna d'acqua sono state eliminate dalla totalità delle informazioni raccolte su tutti i banchi.

Invece, per l'identificazione delle *patch* acustiche identificabili come plancton, dapprima gli ecogrammi a 38 e 120 kHz sono stati filtrati includendo nell'analisi valori di *backscattering* compresi solo nell'intervallo [-70, -50 dB]; è stata poi utilizzata una maschera binaria che poneva a 1 i valori di ogni *pixel* all'interno di questo intervallo, a 0 i valori posti al di fuori. La differenza tra gli ecogrammi binari delle due frequenze ha dato

luogo a un ulteriore ecogramma risultante che ha dato luogo a una maschera binaria: sono stati posti uguali a 0 i *pixel* con valore minore o uguale a 5 dB; uguali a 1 invece i *pixel* con valore superiore a 5 dB. In ragione della tipica distribuzione in strati orizzontali del plancton, questa seconda maschera binaria è stata sottoposta a un filtro morfologico che utilizza un elemento strutturale di dimensione 3x3. L'ecogramma binario risultante moltiplicato per l'ecogramma originario a 120 kHz conclude la procedura (Patti *et al.*, 2011). Dalle *patch* individuate, precisamente dalla sola frequenza a 38 kHz, vengono estratti i parametri descrittivi della struttura della *patch*, cioè quelli di *posizione* (come profondità e distanza dal fondo), di *dimensione* (come area, lunghezza, altezza, elongazione), di *energia* sia in NASC (indice di biomassa) sia in NASC/AREA (indice di densità).

Ogni miglio nautico i valori medi di NASC sono stati mediati e sono state escluse dall'analisi quelle aree in cui non erano compresenti banchi di pesce e plancton, al fine di analizzare la relazione tra biomassa o densità dei banchi di pesce e biomassa o densità delle *patch* di plancton. Per investigare tale relazione e in particolare per individuare eventuali relazioni non lineari, si è scelto di fare uso dei Modelli Additivi Generalizzati (GAMs). Per proseguire l'analisi e analizzare le relazioni per ognuna delle tre Zone, il plancton che si trovava tra 1 e 20 m di profondità è stato poi separato da quello che si trovava a distanza dal fondo inferiore ai 5 m, cioè il plancton superficiale è stato separato da quello profondo. I modelli sono stati applicati grazie ad una distribuzione Gaussiana dell'errore con l'identità come funzione di link. E' stata poi applicata sui dati una trasformazione logaritmica dei predittori perché fosse rispettata la normalità nelle distribuzioni. Sono state selezionate le funzioni di *smoothing* secondo la metodologia di Hastie & Tibshirani (1990) usando la libreria "mgcv" del software statistico R (R. Development Core Team, 2005).

2.3. Trattamento dei dati a multifrequenza per la caratterizzazione e l'identificazione dei banchi di pesci

I due *echosounder* EK500 ed EK60 utilizzati per raccogliere i dati acustici presentano due o tre trasduttori *split-beam* i quali, calibrati secondo il metodo raccomandato da Foote *et al.* (1987), operano alle frequenze di 38, 120 e 200 kHz.

I tre trasduttori emettevano *ping* contemporaneamente ogni secondo, con la differenza che l'impulso prodotto dall'*echosounder* EK500 per le frequenze di 38, 120 e 200 kHz

dura rispettivamente 1.0, 1.0, 0.6 ms, mentre l'impulso prodotto dall'*echosounder* EK60 per tutte e tre le frequenze dura 1.024 ms. Grazie a tali strumentazioni, sono stati raccolti i dati acustici lungo transetti costa-largo (Figura 7) effettuati presso la piattaforma continentale della Sicilia meridionale, con una velocità di navigazione pari a 10 nodi. Le operazioni di cattura sono state svolte, invece, ad una velocità di 4 nodi, utilizzando una rete pelagica caratterizzata da apertura orizzontale di 13-15 metri, apertura verticale della bocca di 6-8 metri, dimensione delle maglie terminali di 10 mm.

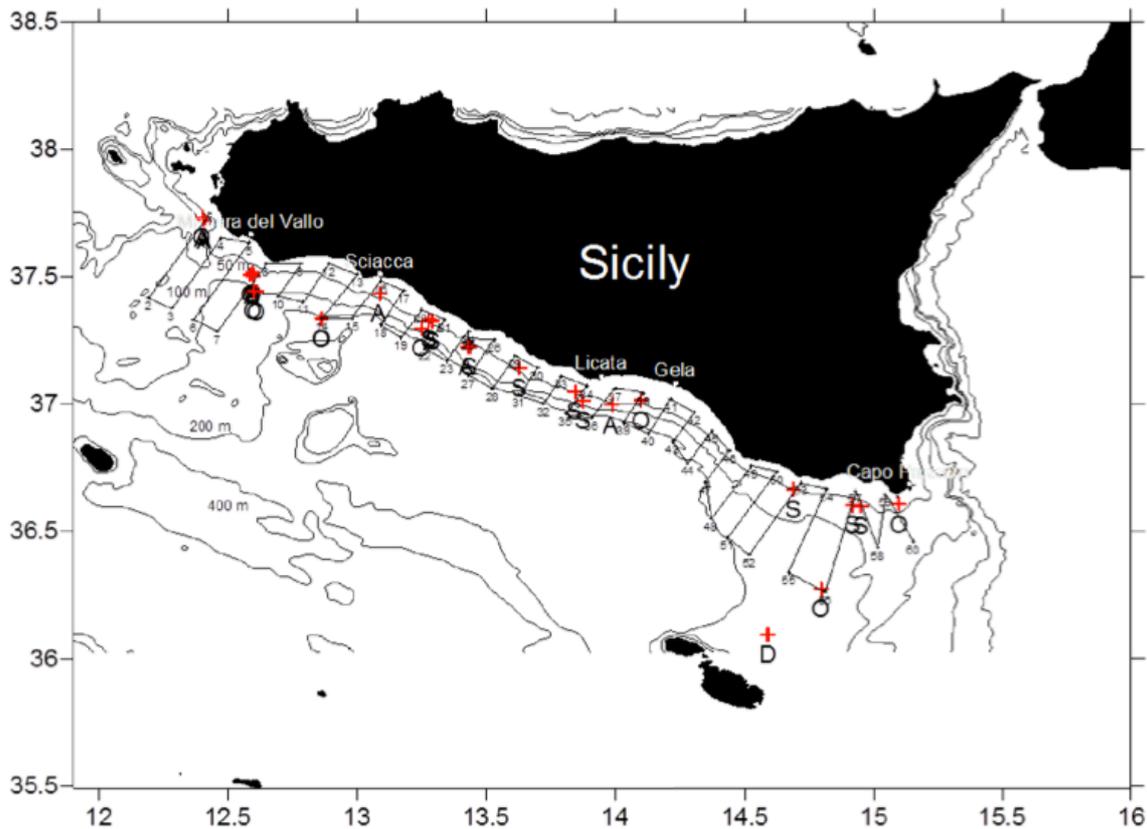


Figura 7. Area di studio. In rosso sono visualizzate le catture monospecifiche e i transetti costa-largo effettuati durante il survey ANCHEVA2009. La lettera indica la specie predominante nella cala: A=acciughe, O=altre specie pelagiche, S=sardine, D=demersali (D'Elia, 2012).

2.4. Trattamento dei dati acustici per l'estrazione delle informazioni sui banchi di pesce da dati a multifrequenza

Dall'ecogramma sono state isolate le ecotracce riferibili a potenziali banchi di pesce. In una prima fase si è applicato agli ecogrammi un "*fish school filter*", cioè il filtro sviluppato da SINFAMI (Fernandes *et al.*, 2006), modificato, adattato ai banchi e quindi implementato come algoritmo. Tale algoritmo applica dei filtri al fine di isolare i banchi di pesce con vescica natatoria e rende spazialmente comparabili dati acquisiti a differenti

frequenze; esso ha permesso di ottenere un ecogramma “più pulito” da analizzare per identificare le caratteristiche dei banchi di pesce.

Per la seconda fase si è proceduto ad utilizzare un altro algoritmo: esso, basandosi su alcuni parametri di banco, è servito per individuare e isolare le ecotracce corrispondenti a potenziali banchi presenti nell’area campionata dalla rete. Le ecotracce isolate in questo modo vengono selezionate come oggetti da utilizzare per la successiva classificazione.

Le operazioni effettuate durante la terza fase hanno permesso di estrarre i parametri energetici, morfologici e batimetrici.

2.5. L’approccio *fuzzy naïve Bayes* per la classificazione dei segnali da acustica

La metodologia proposta in questo lavoro di tesi nasce dalla fusione di logica *fuzzy* e teorema di *Bayes* per dar luogo ad un approccio modellistico consistente in un compilatore *naïve Bayes* operante in ambiente *fuzzy* (Tang e Xu, 2005). Si tratta, in sintesi, di un compilatore *naïve bayes* in cui le probabilità in input sono definite a partire dalla *fuzzy*ficazione delle variabili di sistema; l’output rende un valore di probabilità associato ad un *fuzzy set* appartenente alla partizione *fuzzy* del dominio della variabile target di interesse, nel caso specifico la composizione percentuale in specie di pesce di quanto rilevato col sonar durante ogni singola cala (Bozzeda, 2013).

2.5.1. *Fuzzy Logic*

La *fuzzy logic* si basa sul concetto di *fuzzy set* (Zadeh, 1965) ossia di un insieme al quale un elemento appartiene con un grado di verità compreso nell’intervallo [0,1] (Figura 8).

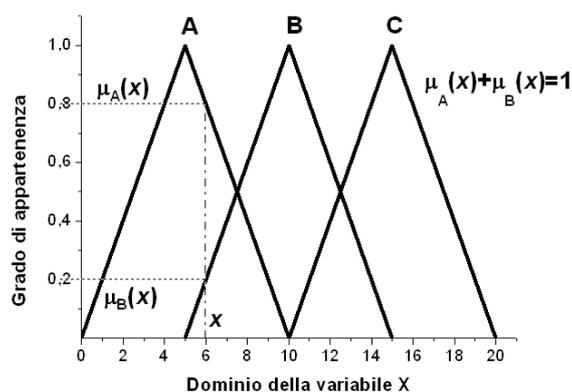


Figura 8. *Fuzzy set*: partizione del dominio di una variabile X in tre *fuzzy sets* che si compenetrano nel punto medio e che rispondono alla funzione d’appartenenza a “triangolo isoscele”. $\mu_{A/B}(x)$ =coefficiente d’appartenenza di x ad un *fuzzy set*; la relazione $\mu_A(x)+\mu_B(x)=1$ esprime la relazione che esiste fra i coefficienti d’appartenenza di un elemento x conosciuto con incertezza nulla (Bozzeda, 2013).

Assunta l'impossibilità di ottenere misure di grandezze o fenomeni reali con un errore nullo o con l'assoluta sicurezza della costanza dell'errore, la logica *fuzzy* inquadra il grado d'incertezza come una "qualità" della grandezza misurata. Un qualunque dato può rispondere ad una determinata caratteristica secondo un valore compreso fra 0 e 1.

Il grado di veridicità di un certo valore rispetto ad una determinata caratteristica è chiamato "coefficiente d'appartenenza" ed esprime l'appartenenza di un valore ad un *fuzzy set* che ne definisce una caratteristica. E' dimostrato che i coefficienti di appartenenza esprimono valori di possibilità; rimane invece aperta la discussione sul considerarli o meno valori di probabilità. Formalmente i coefficienti d'appartenenza verificano i tre assiomi di Kolmogorov (Zadeh, 1968) ed il problema del complementare per cui se ne può dimostrare la loro valenza probabilistica.

Attorno alla teoria dei *fuzzy set* sono state sviluppate in diversi campi applicazioni in grado di trattare l'incertezza di stato di sistemi di variabili.

Ritornando al significato formale dei "coefficienti di appartenenza", essi indicano un valore di possibilità. Per definire il significato di "valore di possibilità" si può ricorrere ad un esempio utilizzato da Zadeh, il padre della logica *fuzzy*. Si consideri l'enunciato: "è possibile che Giacomo questa mattina si mangi 3 uova, ma non l'ha mai fatto, è quindi improbabile che succeda".

L'esempio mostra come il termine *possibilità* convogli in sé un significato di fattibilità/veridicità ed un significato di consistenza con l'informazione disponibile. Un evento è oggettivamente possibile se esiste una fattibilità fisica per il suo realizzarsi. Dal punto di vista "frequentista" una possibilità può essere considerata come limite superiore di una distribuzione e quindi un'estensione del concetto più restrittivo (ma più robusto in virtù del numero più elevato di implicazioni) di probabilità.

2.5.2. Compilatore *naïve Bayes*

Un compilatore bayesiano è una metodologia di tipo probabilistico che, avvalendosi di una serie di dati etichettati, stima i parametri del modello generativo e classifica le nuove istanze osservate utilizzando il teorema di Bayes (John e Langley, 1995) e selezionando la classe o categoria o intervallo di valori in un dominio continuo che ha la probabilità maggiore di aver generato l'istanza. Questo metodo di classificazione è considerato importante per diversi motivi, fra cui la semplicità di implementazione e di calcolo dei parametri che non necessita di complicati schemi di calcolo.

Teorema di Bayes :

$$P(A_i | E) = \frac{P(E | A_i)P(A_i)}{P(E)} = \frac{P(E | A_i)P(A_i)}{\sum_{j=1}^n P(E | A_j)P(A_j)}$$

Il *naïve Bayes* è uno dei compilatori più semplici e versatili poiché si basa sull'unica assunzione semplificativa che tutti gli attributi che descrivono un'istanza siano fra loro condizionatamente indipendenti (Pazzani, 1995; Kohavi, 1996). Nonostante la sua semplicità e le restrizioni dell'assunzione semplificativa, in pratica mai rispettata per la maggior parte di pattern complessi, dalla letteratura emerge come il compilatore funzioni sorprendentemente bene e come le sue prestazioni complessive siano le migliori fra i classificatori bayesiani conosciuti. L'assunzione di "indipendenza condizionale" viene detta "assunzione *naïve Bayes*" e, in base a questa, il classificatore esegue una classificazione di tipo MAP (*Maximum A posteriori Probability*). L'ambito applicativo del *naïve Bayes* riguarda la classificazione di istanze descrivibili attraverso un insieme di attributi di cardinalità anche molto elevate. A differenza di altri algoritmi di apprendimento, il *naïve Bayes* non ricerca la soluzione nello spazio delle possibili soluzioni (aprioristiche), ma ne stima la veridicità sulla base delle frequenze osservate nelle istanze di apprendimento.

Sia $X = \{x_1, x_2, \dots, x_n\}$ un insieme di variabili che descrivono gli attributi delle istanze da classificare e sia $Y = \{y_1, y_2, \dots, y_m\}$ una variabile i cui stati rappresentano le categorie di appartenenza possibili delle istanze (Figura 9). Il compilatore stima i parametri di una funzione che fornisce la classe di appartenenza dell'istanza stessa.

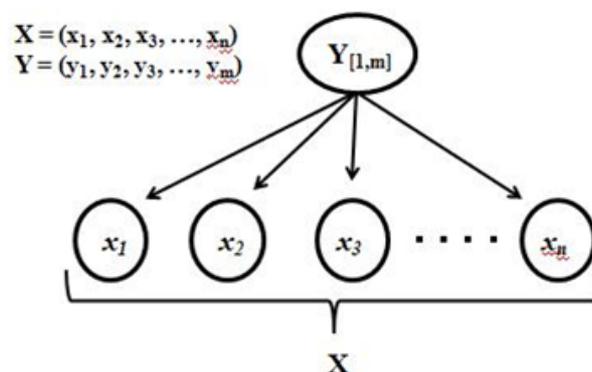


Figura 9: Schema concettuale del funzionamento di un classificatore *naïve Bayes* in cui i parametri sono equivalenti alle probabilità condizionate di ogni n-esima "features"(x) data ogni classe m-esima(y) (Bozzeda, 2013).

L'assunzione d'indipendenza condizionale per cui le componenti delle istanze sono indipendenti all'interno di ciascuna classe fa in modo che:

$$f(x|y) = \prod_{j=1}^n f(x_j|y)$$

da cui si desume che le distribuzioni marginali di ogni attributo delle classi sono stimabili separatamente. In conseguenza di questo, un problema n-multivariato risulta risolvibile come un problema n-univariato semplificando la parte computazionale della classificazione.

La relazione *naïve Bayes* fra attributi delle istanze e classe di attribuzione è una relazione non lineare che non richiede monotonicità in x_j per cui risulta potenzialmente in grado di replicare qualunque dinamica di classificazione.

L'applicazione del *naïve Bayes* proposta in questo lavoro si basa principalmente su queste due ultime caratteristiche del classificatore che permettono rispettivamente il trattamento di *data set* multivariati ad elevata dimensionalità e l'individuazione nel sistema delle variabili di punti soglia che il modello identifica come punti di "treshold" che descrivono l'assegnazione di un banco ad una specie piuttosto che ad un'altra.

2.5.3. Approccio *fuzzy naïve Bayes*

L'alta complessità e l'alta incertezza (quindi l'elevata dimensionalità) associate ai segnali sonar registrati possono essere trattate attraverso l'utilizzo congiunto di logica *fuzzy* e compilatore *naïve Bayes*. Il sistema derivante è un compilatore *fuzzy naïve Bayes* che consiste in un compilatore *naïve Bayes* operante in ambiente *fuzzy*.

Il compilatore, come un normale *naïve Bayes*, elabora i dati per la stima dei parametri sotto forma di istanze.

L'algoritmo del compilatore è composto da 5 step successivi (Widyantoro e Yen, 2000):

1. fuzzyficazione delle variabili;
2. fase di apprendimento;
3. inferenza *fuzzy naïve Bayes* sulle nuove istanze osservate;
4. definizione del risultato utilizzando la regola del centroide all'interno dell'intervallo di incertezza risultante dal compilatore;
5. defuzzyficazione dell'output.

- 1 – La fuzzyficazione delle variabili consiste nel descrivere le variabili considerate secondo formalismo *fuzzy*. Viene creata una partizione *fuzzy* all'interno del dominio di valori di ciascuna variabile per cui ogni dominio di sistema viene descritto da una successione di *fuzzy set* secondo la quale verranno fuzzyficate le osservazioni che compongono il *set* di *training* in base al quale il compilatore effettuerà la fase di *learning*.
- 2 – Durante la fase di *learning* il sistema stima i parametri del classificatore dati dalle probabilità marginali di ogni *fuzzy set*

$$P(x_i) = \int_{e \in L} \mu_{x_i}^e dx$$

Dove:

L è il *set* di apprendimento

x_i indica la generica i -esima variabile

e indica una generica istanza del *set* di apprendimento L ,

e dalle probabilità condizionate di ogni *fuzzy set* dato il valore della variabile *target*

$$P(x_i|y) = \frac{\int_{e \in L} \mu_{x_i}^e \mu_y^e dy}{P(y)}$$

Nella fase di *learning* i valori di probabilità congiunta e condizionata sono calcolati secondo la definizione di spazio di probabilità di Lebesgue e sono riferiti all'evento *fuzzy* cioè alla possibilità che una data variabile assuma valori inclusi in un particolare *fuzzy set* (Zadeh 1968).

3 – Inferenza *fuzzy naïve Bayes*: data una nuova istanza osservata, prevista o simulata il compilatore calcola per ogni *fuzzy set* della partizione della variabile *target* la probabilità condizionata del *fuzzy set target* rispetto alla nuova istanza osservata.

$$P(y|e) = P(y) \left(\sum_{x_1 \in X_1} \frac{P(x_1|y)}{P(x_1)} \mu_{x_1}^e \right) \dots \left(\sum_{x_n \in X_n} \frac{P(x_n|y)}{P(x_n)} \mu_{x_n}^e \right) = \mu(y)$$

L'output finale (della variabile *target*) è composto da una serie di probabilità

condizionate associate ai *fuzzy set* di cui si compone la partizione della variabile *target*.

4 – I valori previsti sono ottenuti secondo criterio MAP (*Maximum a posteriori probability*) attraverso il metodo del centroide, per cui si avranno uno o più valori secondo il carattere estremo delle nuove istanze a cui il compilatore associa un campo di incertezza calcolato in base alla somiglianza statistica con le istanze di *training*.

L'utilizzo congiunto di logica *fuzzy* e compilatore *naïve Bayes* rende l'output di classificazione simile per accuratezza all'output di un metodo predittivo. Il dominio della variabile output (variabile *target*) sottoposto a fuzzyficazione non è infatti più rappresentato da classi discrete ma da classi sfumate.

2.5.4. Considerazioni sul metodo

L'approccio *fuzzy naïve Bayes* si presenta robusto rispetto al rumore inevitabilmente presente in un set complesso di dati. I set di output dovuti ad una o più particolari forzanti mostrano una diminuzione generale della variabilità in quanto la classificazione è influenzata dall'intero set di *training*. Una misura di variabilità reale degli output è calcolabile in riferimento ai diversi *range* d'incertezza dei risultati.

Il rumore presente nei dati viene "splittato" nelle varianze associate alle diverse variabili per cui in fase di elaborazione si osserva una diminuzione del rumore totale in ragione, appunto, della distribuzione del rumore. Quest'ultima caratteristica rende l'approccio potenzialmente adatto all'analisi e all'interpretazione di ecogrammi mono-multispecie di banchi di pesce. La robustezza dell'approccio rispetto a pattern d'apprendimento ad elevata variabilità è legata all'elevato numero di parametri che vengono stimati durante la fase di apprendimento. Se si considera n come il numero di variabili considerate e k il numero di *fuzzy set* della partizione, durante la fase di apprendimento vengono stimati $k^2(n-1) + k(n-1) + 1$ parametri.

Approcci analoghi sono stati utilizzati per la descrizione, l'analisi e la modellizzazione di pattern di dati complessi in diversi ambiti; esistono infatti applicazioni *fuzzy naïve Bayes* in oceanografia (Randon *et al.*, 2008), robotica (Bustamante *et al.*, 2007), classificazione del testo (Storr, 2002) ed analisi di comunità bentoniche (Bozzeda, 2013).

Ogni nuova istanza osservata, inoltre, può essere utilizzata in un procedimento di autocalibrazione del modello, aggiornando la fase di *learning*.

2.6. Approccio metodologico: campagne oceanografiche

Come già specificato, per bypassare la tempistica che intercorre fra l'acquisizione dati, la loro "pulizia" e la successiva elaborazione, le osservazioni utilizzate per questo lavoro di tesi provengono da campagne oceanografiche effettuate tra il 2002 e il 2011 dall'IAMC-CNR di Capo Granitola.

Oltre ai motivi sopra riportati, l'applicazione di un algoritmo di classificazione allo stesso set di dati alternativo a quello normalmente utilizzato (*classification tree*) ne agevolerà il futuro confronto particolareggiato e l'eventuale integrazione dei due metodi.

I tempi di lavoro in barca sono scanditi da una turnazione.

Le attività da svolgere a bordo sono diverse; quelle eseguite nell'ambito delle campagne cui ho partecipato sono le seguenti:

- attività relative ad *echosurvey* (acquisizione dati acustici con *echosounder* Simrad EK60);
- attività relative alle cale con rete pelagica (operazioni di prelievo del pesce; pesatura; smistamento per specie e taglia; eventuali analisi da svolgere a bordo; eventuale congelamento o conservazione in formalina per analisi da svolgere successivamente a terra);
- rilevamento dei profili delle variabili oceanografiche (tramite sonda multiparametrica *CTD 9/11 plus* della Seabird);
- campionamento di acqua (tramite bottiglie *Niskin*);
- campionamenti planctonici (tramite retino WP2);
- campionamenti ittioplanttonici e zooplanctonici con retini BONGO40;
- calibrazione dell'*echosounder*, attraverso il metodo della sfera standard.

Appare utile descrivere più nel dettaglio la parte che riguarda le tecniche acustiche e le strumentazioni adoperate.

Nello scafo della nave (a 3.5 m dalla superficie) sono installati tre trasduttori che inviano l'impulso acustico verso il fondo, ognuno con una diversa frequenza (38 kHz, 120 kHz e 200 kHz), e che ricevono l'eco di risposta.

I trasduttori sono collegati a tre *GPT* (*General Purpose Transceiver*) mostrati in figura 10. Uno per ogni trasduttore, essi fungono da interfaccia tra trasduttore e computer. I tre *GPT*, sono connessi in rete attraverso un concentratore detto *switch* (posto al di sopra dei tre *GPT*, in figura 10) per cui viene creata una "rete locale". Allo *switch* è collegato il

computer in modo tale che tutte le informazioni vengano trasferite in un senso e nell'altro (dallo *switch* al computer e viceversa).

L'applicativo software ER60 della Simrad programma e gestisce i tre GPT nel loro funzionamento.



Figura 10. Dai trasduttori agli ecogrammi.

A seconda dell'oggetto che il segnale acustico incontra, si hanno diversi tipi di echi di ritorno, perché, in base alla lunghezza d'onda, le frequenze inviate interferiscono o meno con certi oggetti. Sebbene gli echi da soli non diano informazioni complete sulle specie di pesce insonificate, l'analisi degli ecogrammi permette di eliminare il segnale del plancton (rappresentato da echi di ampiezza inferiore rispetto a quelli dei pesci). Questi ultimi due elementi sono risultati molto utili agli obiettivi delle ricerche di Marta D'Elia (2012), che è riuscita a classificare banchi monospecifici di pesce e che considerava come "rumore" l'eco acustica del plancton. La presenza di quest'ultimo può invece tornare utile, ad esempio, per la caratterizzazione della forma di un banco di pesce, nel caso ad esempio si tratti di una specie planctivora.

Ad ognuno dei 3 trasduttori corrisponde un ecogramma (Figura 11) e ogni riga verticale sull'ecogramma rappresenta un impulso trasmesso e un'eco ricevuta. Considerato che viene emesso un impulso al secondo e che la velocità del suono è di circa 1500 m/s, è possibile scandagliare in un secondo i primi 750 m di colonna d'acqua, sebbene per le specie che ci interessa studiare basterebbe acquisire dati fino ad una profondità decisamente inferiore.

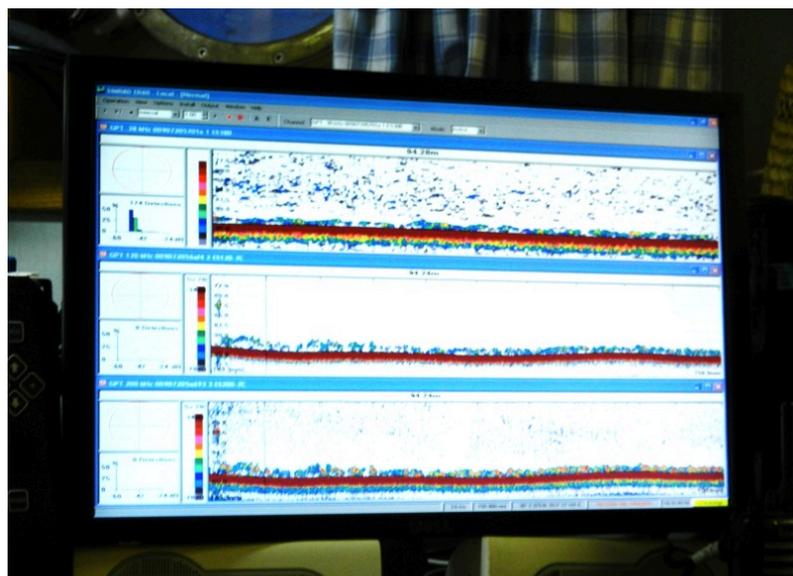


Figura 11. Schermata dell'echosounder in cui si nota un ecogramma per ogni trasduttore (38, 120, 200 kHz).

Tutto quello che il segnale incontra nella colonna d'acqua dà, dunque, riflessione in relazione alla differenza di impedenza acustica dell'oggetto riflettente rispetto all'acqua. Il segnale più forte è quello restituito dal fondo ed è molto importante, infatti l'ispezione di cala - oltre che per vedere se ci sono banchi di pesce - si esegue per vedere se il fondo presenta o meno repentine variazioni di pendenza che potrebbero danneggiare la rete nella fase del campionamento biologico.

Ogni ecogramma (Figura 12) mostra tipicamente 1 miglio nautico e presenta in alto una zona nascosta: i primi 2 metri dal trasduttore rappresentano infatti il cosiddetto "campo vicino" che è ininterpretabile.

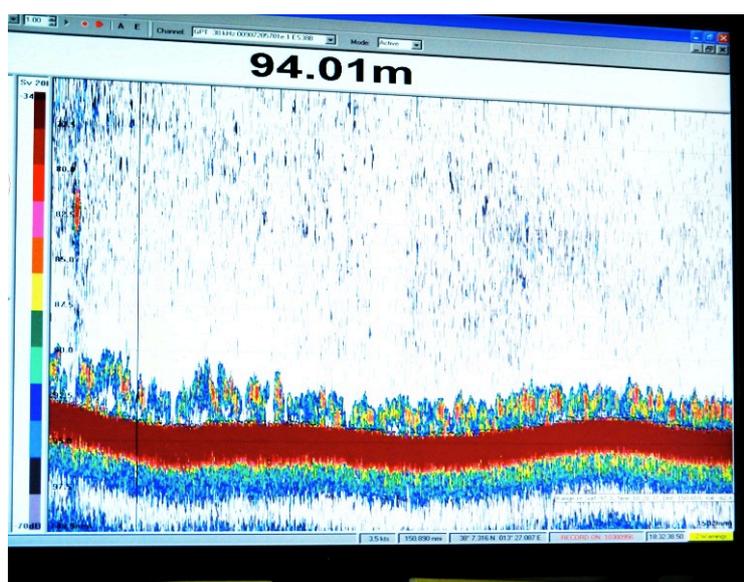


Figura 12. Esempio di Ecogramma

Ogni trasduttore è circolare ed è diviso in 4 settori, ciascuno dei quali trasmette contemporaneamente lo stesso impulso. Nel campo vicino l'interferenza tra le quattro onde di pressione generate dai quattro settori non dà luogo ad una "interferenza costruttiva" (che si ha quando l'onda diventa piana) perché ciascuna di esse è ancora circolare.

Il sistema sopra descritto acquisisce informazioni con la nave sempre in movimento e a velocità intorno ai 9-10 nodi; in caso di condizioni meteo-marine avverse, la nave deve rallentare il moto per ridurre il rumore prodotto dalle bolle d'aria mescolate all'acqua nello strato più superficiale, ma comunque è buona norma non ridurre la velocità al di sotto dei 5 nodi.

Tutti i dati sono georeferenziati, tramite sistema *GPS*, che associa le coordinate al singolo *ping* per poter ricostruire il tracciato.

I dati acquisiti vengono registrati su *hard disk* in modo continuo. Occasionalmente, mentre si percorreva un certo transetto si calava la rete (cala) per la pesca dei banchi di pesce su cui campionare a bordo gli esemplari per le analisi biologiche. La cala percorreva un tratto di mare di circa 4 miglia nautiche e catturava tutti i banchi che incontrava. I dati acustici da abbinare al periodo e al tratto di mare in cui si era svolta la cala, una volta contati i pesci, venivano poi rintracciati attraverso il riscontro delle coordinate.

Durante le campagne, ci si ferma in punti ben precisi per rilevare i parametri chimico-fisici dell'acqua: in corrispondenza dei vertici dei transetti (di solito perpendicolari alla costa e paralleli tra loro) vengono fissate le stazioni per il rilevamento dei parametri fisici con CTD (sonda multiparametrica che rileva informazioni su conduttività, temperatura e profondità).

Superati i 5 nodi, viene attivato il sistema per la registrazione dei dati, in formato *RAWDATA*, dati grezzi: essi conservano il segnale acquisito dai trasduttori in formato digitale. I parametri di calibrazione permettono in un qualunque momento di risalire al livello di pressione acustica di *backscattering* ricevuto per ogni *ping* emesso.

La calibrazione dei trasduttori è un procedimento, che viene svolto almeno una volta nel corso del survey, di fondamentale importanza perché permette di compensare eventuali derive dei componenti che costituiscono gli *echosounder* e consente di tenere conto di eventuali fattori esterni (fattori biotici e abiotici). Per le procedure di *POSTPROCESSING*, naturalmente è necessario combinare i dati acquisiti a bordo nel corso del *survey* con i dati di calibrazione.

La calibrazione è necessaria perché permette che i dati, interpolati tra di loro, vengano tradotti in quantità di biomassa di organismi: se essa non viene effettuata, non è quindi possibile analizzare il dato e produrre informazioni quantitative sulla biomassa ittica investigata.

Il principale obiettivo del *survey* era quello di ottenere la distribuzione spaziale di acciughe e sardine unitamente ad altre specie di pesce pelagici presenti.

Il sistema ER60 acquisisce 3 tipi di file, il primo dei quali è un *filename.raw*, che riporta la data e l'ora di registrazione di Greenwich e contiene informazioni sia sulla colonna d'acqua che sul fondale. Esso è accompagnato da altri due file: *filename.bot*, un file più piccolo che dà informazioni su quello che il sistema ha interpretato come fondo, e *filename.idx* che contiene le posizioni GPS.

3. ELABORAZIONI

3.1 Criteri adottati per l'elaborazione

L'enorme mole di dati estratti a seguito dell'elaborazione dei dati acustici - si tratta in particolare di 435 osservazioni relative ad ogni banco rilevato, per cui erano state rilevate 87 variabili (D'Elia, 2012) - ha reso possibile un'oculata cernita di quelli più adeguati per le elaborazioni da attuare secondo il nuovo approccio modellistico.

3.2. Scelta delle variabili

Una volta eliminate tutte quelle variabili non necessarie alla classificazione, quali etichette, date, orari, coordinate geografiche (latitudine e longitudine) ecc., la matrice *Excel* originaria è stata ripulita anche delle osservazioni di tutte quelle variabili i cui dati non erano disponibili (NA) per tutti i banchi. In tal modo si è ottenuta una matrice costituita da 403 osservazioni per le seguenti 40 variabili, ciascuna delle quali rappresenta un parametro morfologico, batimetrico o energetico e per ciascuna delle quali è riportato tra parentesi il codice identificativo:

- 1) valore medio del volume di *backscattering strenght* alla frequenza di 38 kHz per il banco analizzato (Sv_mean_38);
- 2) *Nautical Area Scattering Coefficient* alla frequenza di 38 kHz (NASC_38);
- 3) valore massimo del volume di *backscattering strenght* alla frequenza di 38 kHz per il banco analizzato (Sv_max_38);
- 4) valore minimo del volume di *backscattering strenght* alla frequenza di 38 kHz per il banco analizzato (Sv_min_38);
- 5) altezza media del banco (Height_mean);
- 6) profondità media del banco (Depth_mean);
- 7) valore della lunghezza (estensione orizzontale) del banco in assenza di correzione (Uncorrected_length);
- 8) valore dello spessore (estensione verticale) del banco in assenza di correzione (Uncorrected_thickness);
- 9) valore del perimetro del banco in assenza di correzione (Uncorrected_perimeter);
- 10) valore dell'area del banco in assenza di correzione (Uncorrected_area);
- 11) deviazione standard dei valori di Sv compresi nel banco (Standard_deviation_38);
- 12) *skewness* dei valori di Sv compresi nel banco (Skewness_38);

- 13) valore corretto della lunghezza (estensione orizzontale) del banco (Corrected_length);
- 14) valore corretto dello spessore (estensione verticale) del banco (Corrected_thickness);
- 15) valore corretto del perimetro del banco (Corrected_perimeter);
- 16) valore corretto dell'area del banco (Corrected_area);
- 17) parametro statistico utilizzato per misurare la dispersione di energia a 38 KHz all'interno di un banco lungo la direzione orizzontale (Horizontal_roughness_38);
- 18) parametro statistico utilizzato per misurare la dispersione di energia a 38 KHz all'interno di un banco lungo la direzione verticale (Vertical_roughness_38);
- 19) area calcolata assumendo che la forma del banco sia cilindrica (X3D_school_area);
- 20) volume calcolato assumendo che la forma del banco sia cilindrica (X3D_school_volume)
- 21) misura dell'area di *backscattering* alla frequenza di 38 kHz su scala lineare (ABC_38)
- 22) misura dell'area di *backscattering* alla frequenza di 38 kHz su scala logaritmica (Area_Backscatter_Strength_38)
- 23) valore dello spessore (estensione verticale) del banco corretto per la geometria del *beam* in accordo al sistema del 1998 di Diner (Thickness_mean);
- 24) per traduttori verticali questo parametro coincide con la Depth mean (profondità media del banco calcolata nel centro geometrico) (Range_mean);
- 25) volume campionato (Beam_volume_sum);
- 26) valore medio del volume di *backscattering strenght* alla frequenza di 120 kHz per il banco analizzato (Sv_mean_120);
- 27) *Nautical Area Scattering Coefficient* alla frequenza di 120 kHz (NASC_120);
- 28) valore massimo del volume di *backscattering strenght* alla frequenza di 120 kHz per il banco analizzato (Sv_max_120);
- 29) valore minimo del volume di *backscattering strenght* alla frequenza di 120 kHz per il banco analizzato (Sv_min_120);
- 30) deviazione standard dei valori di Sv compresi nel banco per la frequenza 120 kHz (Standard_deviation_120);
- 31) *skewness* dei valori di Sv compresi nel banco per la frequenza 120 kHz (Skewness_120);

- 32) parametro statistico utilizzato per misurare la dispersione di energia a 120 KHz all'interno di un banco lungo la direzione orizzontale (Horizontal_roughness_120);
- 33) parametro statistico utilizzato per misurare la dispersione di energia a 120 KHz all'interno di un banco lungo la direzione verticale (Vertical_roughness_120);
- 34) misura dell'area di *backscattering* alla frequenza di 120 kHz su scala lineare (ABC_120);
- 35) misura dell'area di *backscattering* alla frequenza di 120 kHz su scala logaritmica (Area_Backscatter_Strength_120);
- 36) variabilità della superficie del banco; si tratta di un indice pari al rapporto tra il perimetro del banco e il perimetro del rettangolo al contorno del banco, avente una certa altezza e lunghezza (Unevenness2);
- 37) dimensione frattale; si tratta di un indice della complessità della forma del banco (FractalDim);
- 38) distanza media del banco dal fondo (Mean_dist_bottom);
- 39) numero di banchi per cala (N° banchi per cala);
- 40) distanza del banco dalla costa in metri (mDistanceToCoast).

La matrice dei dati bruti costituita dalle 403 osservazioni per ciascuna delle 40 variabili è stata utilizzata per costruire il grafico della PCA (*Principal Component Analysis*) (Figura 13) con la sua tabella esplicativa (Tabella 1); poiché l'analisi risultava poco chiara e poco esplicativa, è stato necessario mantenere nella sua interezza la variabilità di tutte le variabili per costruire il classificatore *fuzzy naïve Bayes*.

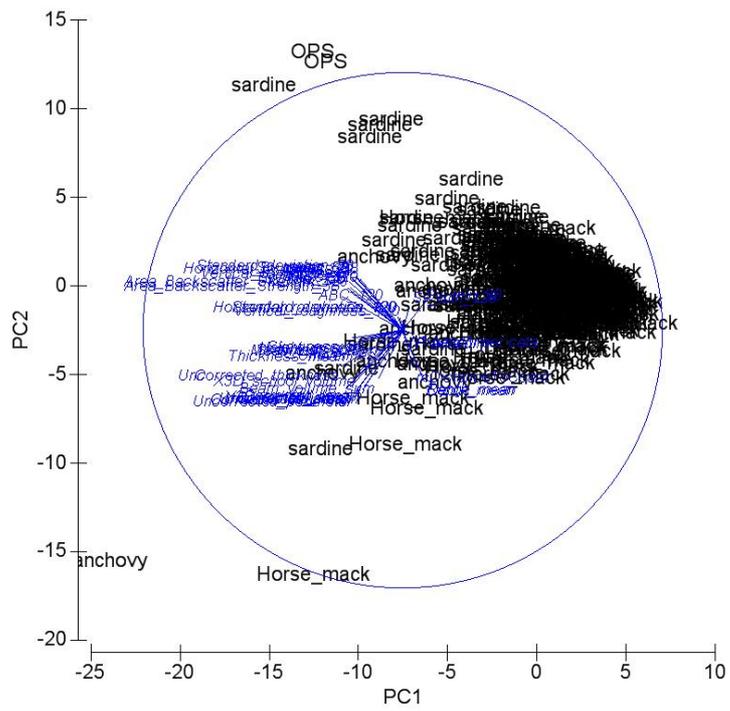


Figura 13. PCA totale.

Tabella 1. Tabella esplicativa della PCA in figura 13.

PCA
Principal Component Analysis

Data worksheet
Name: Data1
Data type: Environmental
Sample selection: All
Variable selection: All

Eigenvalues

PC	Eigenvalues	%Variation	Cum.%Variation
1	11.5	28.7	28.7
2	7.12	17.8	46.5
3	3.91	9.8	56.2
4	2.78	6.9	63.2
5	2.43	6.1	69.3

Eigenvectors
(Coefficients in the linear combinations of variables making up PC's)

Variable	PC1	PC2	PC3	PC4	PC5
Sv_mean_38	-0.196	0.204	0.036	0.119	0.079
NASC_38	-0.195	0.197	0.037	-0.160	-0.193
Sv_max_38	-0.210	0.162	-0.007	0.135	0.070
Sv_min_38	0.076	0.097	0.025	-0.112	0.048
Height_mean	-0.195	-0.030	0.082	-0.222	0.217
Depth_mean	0.082	-0.198	-0.040	-0.351	-0.208
Uncorrected_length	-0.155	-0.231	-0.080	0.199	-0.219
Uncorrected_thickness	-0.236	-0.140	0.060	-0.093	0.160
Uncorrected_perimeter	-0.181	-0.239	-0.060	0.178	-0.159
Uncorrected_area	-0.221	-0.227	0.005	-0.043	0.034
Standard_deviation_38	-0.180	0.209	0.019	-0.089	-0.310
Skewness_38	-0.137	-0.028	-0.200	0.135	-0.082
Corrected_length	-0.147	-0.219	-0.081	0.239	-0.227
Corrected_thickness	-0.236	-0.140	0.060	-0.093	0.160
Corrected_perimeter	-0.178	-0.231	-0.059	0.206	-0.163
Corrected_area	-0.216	-0.225	0.001	0.009	0.027
Horizontal_roughness_38	-0.163	0.198	0.014	-0.082	-0.302
Vertical_roughness_38	-0.160	0.182	-0.006	-0.056	-0.352
X3D_school_area	-0.201	-0.220	-0.008	0.079	0.007
X3D_school_volume	-0.169	-0.161	0.030	-0.110	0.169
ABC_38	-0.204	0.186	0.033	-0.172	-0.216
Area_Backscatter_Strength_38	-0.234	0.131	0.094	0.031	0.059
Thickness_mean	-0.210	-0.065	0.097	-0.248	0.230
Range_mean	0.081	-0.195	-0.023	-0.342	-0.185
Beam_volume_sum	-0.094	-0.188	0.004	-0.222	-0.013
Sv_mean_120	-0.178	0.203	0.079	0.075	0.108
NASC_120	-0.164	0.167	0.012	-0.159	-0.117
Sv_max_120	-0.202	0.165	0.064	0.089	0.115
Sv_min_120	0.025	0.086	0.035	-0.099	0.029
Standard_deviation_120	-0.011	0.048	-0.485	-0.074	0.093
Skewness_120	-0.152	-0.039	-0.110	0.064	0.060
Horizontal_roughness_120	-0.018	0.050	-0.445	-0.065	0.082
Vertical_roughness_120	-0.002	0.036	-0.484	-0.060	0.102
ABC_120	-0.067	0.096	-0.417	-0.133	0.040
Area_Backscatter_Strength_120	-0.202	0.149	-0.139	-0.035	0.169
Unevenness2	-0.062	-0.060	0.036	-0.053	0.075
FractalDim	0.040	-0.014	-0.019	-0.048	-0.034
Mean_dist_bottom	-0.088	-0.042	0.086	-0.151	0.151
N° banchi per cala	0.015	-0.009	-0.081	0.120	-0.114
mDistanceToCoast	0.040	-0.145	-0.009	-0.358	-0.171

Outputs
Plot: Graph1
Worksheet: Data3

Quindi, ancora dalla matrice dei dati bruti è stata calcolata una matrice triangolare di similarità a coppie tra variabili utilizzando il Coefficiente di Correlazione di Pearson; infine, mediante *software Primer*, è stato costruito un grafico a *cluster* riportato in Appendice 1.

Con l'obiettivo di ottenere pattern di informazione più univoci possibile, sono state eliminate le variabili maggiormente correlate che avrebbero potuto portare ridondanza informativa al momento della stima dei parametri di classificazione: sulla base del grafico a *cluster* e soprattutto della sua tabella esplicativa (Appendice 2) è stato possibile ridurre la mole di variabili da utilizzare per le successive elaborazioni. E' stato pertanto stabilito di mantenere tutte le variabili meno correlate tra loro scegliendo quelle con un valore del Coefficiente di Pearson inferiore a 0.9 (90%) (risalendo, quando occorreva, lungo i *cluster* di ordine superiore fino ad individuare le variabili originali che per prime formavano gruppo). Alcune variabili con valore di correlazione maggiore o uguale a 0.9 sono state ulteriormente trattenute secondo i seguenti criteri:

- nel caso di variabili come una media correlata ad un massimo o a un minimo, è stata mantenuta solo la variabile della media;
- nel caso di un minimo correlato a un massimo, sono state mantenute entrambe le variabili;
- nel caso di variabili corrette e non corrette secondo la correzione di Diner sono state mantenute quelle non corrette (poiché le dimensioni orizzontali e verticali dei banchi soffrono di alcune distorsioni legate alla lunghezza dell'impulso e alla larghezza del *beam*, Diner nel 2001 ha suggerito una procedura di correzione da applicare ai parametri morfologici quando si usa un *echosounder* verticale; essa tuttavia non è sempre applicabile).

Le 17 variabili selezionate sono riportate nel seguente elenco:

- 1) valore medio del volume di *backscattering strenght* alla frequenza di 38 kHz per il banco analizzato (Sv_mean_38);
- 3) valore massimo del volume di *backscattering strenght* alla frequenza di 38 kHz per il banco analizzato (Sv_max_38);
- 4) valore minimo del volume di *backscattering strenght* alla frequenza di 38 kHz per il banco analizzato (Sv_min_38);
- 6) profondità media del banco (Depth_mean);
- 7) valore della lunghezza (estensione orizzontale) del banco in assenza di correzione (Uncorrected_length);

- 8) valore dello spessore (estensione verticale) del banco in assenza di correzione (Uncorrected_thickness);
- 9) valore del perimetro del banco in assenza di correzione (Uncorrected_perimeter);
- 10) valore dell'area del banco in assenza di correzione (Uncorrected_area);
- 11) deviazione standard dei valori di Sv compresi nel banco (Standard_deviation_38);
- 17) parametro statistico utilizzato per misurare la dispersione di energia a 38 KHz all'interno di un banco lungo la direzione orizzontale (Horizontal_roughness_38);
- 18) parametro statistico utilizzato per misurare la dispersione di energia a 38 KHz all'interno di un banco lungo la direzione verticale (Vertical_roughness_38);
- 19) area calcolata assumendo che la forma del banco sia cilindrica (X3D_school_area);
- 21) misura dell'area di *backscattering* alla frequenza di 38 kHz su scala lineare (ABC_38);
- 26) valore medio del volume di *backscattering strength* alla frequenza di 120 kHz per il banco analizzato (Sv_mean_120);
- 30) deviazione standard dei valori di Sv compresi nel banco per la frequenza 120 kHz (Standard_deviation_120);
- 32) parametro statistico utilizzato per misurare la dispersione di energia a 120 KHz all'interno di un banco lungo la direzione orizzontale (Horizontal_roughness_120);
- 33) parametro statistico utilizzato per misurare la dispersione di energia a 120 KHz all'interno di un banco lungo la direzione verticale (Vertical_roughness_120).

3.3. Correlazioni tra le variabili e le percentuali di pesci catturati

E' stato studiato l'insieme dei dati relativi alle osservazioni di ciascuna delle 17 variabili illustrate nell'elenco sopra riportato, correlando ciascuna delle 17 variabili con ciascuno dei 5 *output*, dove per *output* si intende: % *anchovy* (*Engraulis encrasicolus*), % *sardine* (*Sardina pilchardus*), % *Horse mack* (*Trachurus trachurus*), % *OPS* (Other Pelagic Species), % *Demersal* (specie demersali). Costruendo 85 grafici di dispersione con indicatori, la linea di tendenza che veniva fuori e i bassi valori di R^2 (min: 4.4E-0.5 e max: 0.473) hanno dimostrato che nessuna delle 17 variabili si correla con le 5 variabili di *output* in modo sufficiente da poter essere utilizzata come unico indicatore delle specie presenti nei banchi (si riporta qualche esempio dei nostri grafici di correlazione nelle Figure 14-15-16-17-18).

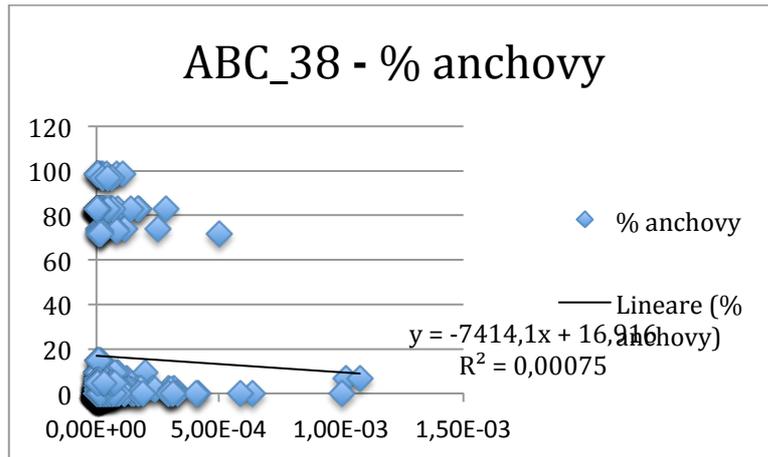


Figura 14. Esempio di grafico di dispersione tra la variabile di input “ABC_38” e la variabile di output “% di *Engraulis encrasicolus*”.

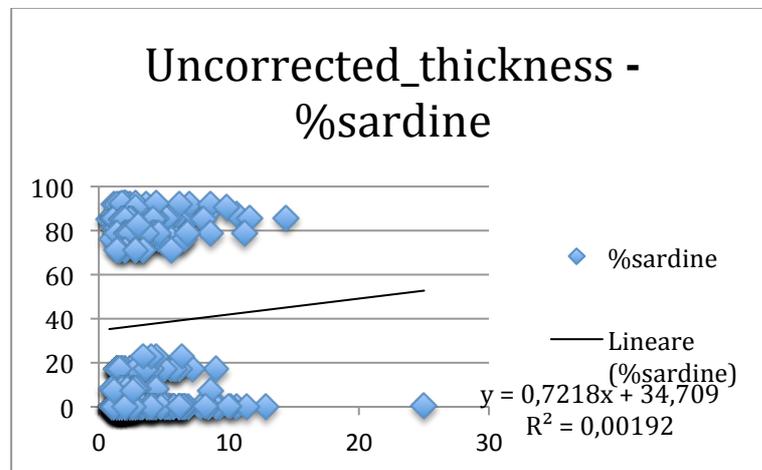


Figura 15. Esempio di grafico di dispersione tra la variabile di input “Uncorrected_thickness” e la variabile di output “% di *Sardina pilchardus*”.

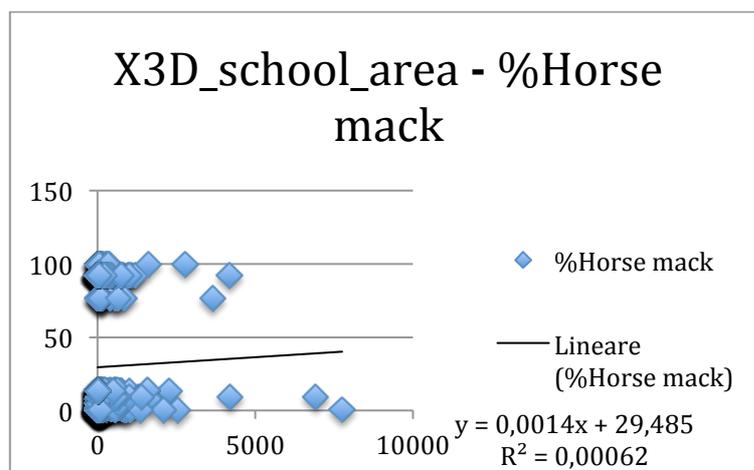


Figura 16. Esempio di grafico di dispersione tra la variabile di input “X3D_school_area” e la variabile di output “% di *Trachurus trachurus*”.

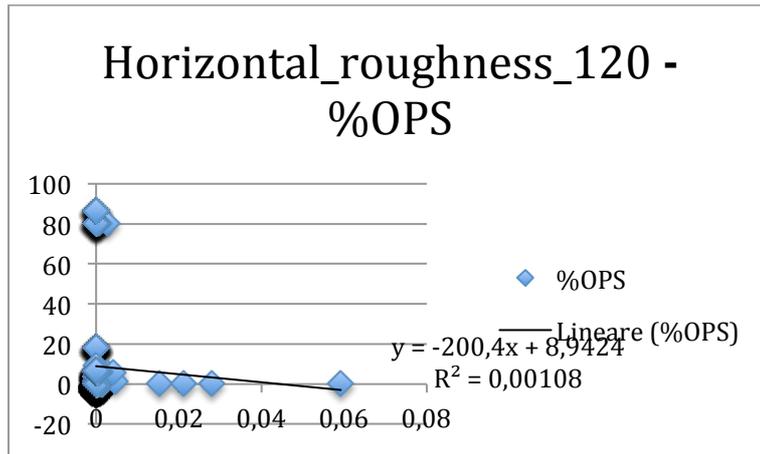


Figura 17. Esempio di grafico di dispersione tra la variabile di input “Horizontal_roughness_120” e la variabile di output “% di Other Pelagic Species”.

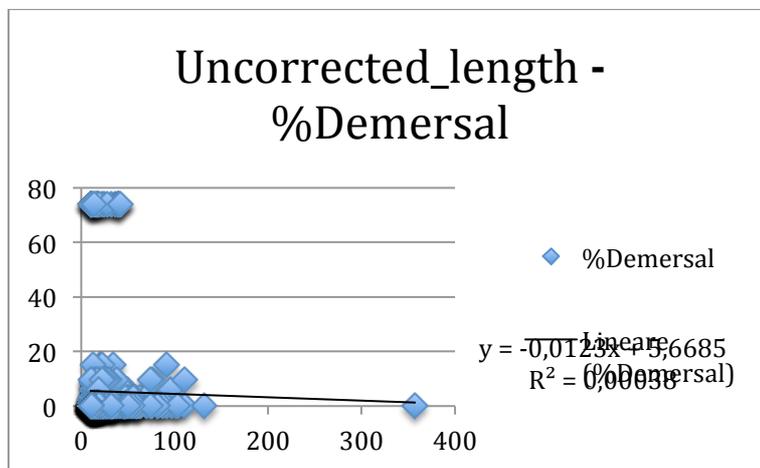


Figura 18. Esempio di grafico di dispersione tra la variabile di input “Uncorrected_length” e la variabile di output “% di specie demersali”.

3.4. Metodo *fuzzy naïve Bayes*: fuzzyficazione

Infine si è proceduto alla fuzzyficazione di tutti i valori della matrice “ripulita”, allo scopo di costruire il classificatore *fuzzy naïve Bayes* che rappresenta il nostro risultato complessivo. Le formule per la fuzzyficazione sono le due seguenti:

$$(n - \text{Min})/20 \quad (1)$$

$$(\text{Max} - n)/20 \quad (2)$$

dove :

“n” rappresenta il valore standardizzato;

“Min” è il minimo del fuzzysset;

“Max” è il massimo del fuzzysset;

“20” è la metà di ciascun fuzzysset.

4. RISULTATI

Sulla base delle elaborazioni svolte, è stato costruito il “classificatore”, entro il quale sono state effettuate delle simulazioni. Per procedere sono stati utilizzati i dati relativi a 31 cale, cioè quei tratti dei transetti in cui, oltre che essere rilevati i dati acustici, erano state effettuate anche le catture, con rete pelagica, di banchi di pesce su cui stimare le percentuali reali delle specie presenti.

Il “classificatore” è stato, innanzi tutto, “istruito” durante la fase di *training*; ciò significa che è stato costruito un *learning sample* inserendo nel classificatore alcune righe – scelte in modo rigorosamente *random* – della matrice “ripulita” e fuzzyficata. Nella fattispecie, per la fase di *training*, sono stati inseriti nel simulatore, assieme ai relativi *output* (cioè le percentuali di acciughe, le percentuali di sardine, le percentuali di suri, le percentuali di altre specie pelagiche e le percentuali di demersali), i dati delle osservazioni da *echosounder* relative alle 11 righe della matrice di *training*. Avendo così istruito il simulatore, si è proceduto alla sua validazione, classificando i dati relativi alle rimanenti 20 righe – ovviamente sprovvisti dei dati relativi agli *output* – che costituiscono il *test set* (Tabella 2).

Una volta ottenuta la classificazione, i risultati (cioè le percentuali simulate di ciascun *output*) sono stati messi a confronto con gli *output* reali provenienti dalle tecniche tradizionali di campionamento (cioè quelle tramite prelievo con rete pelagica) (Appendice 3). In pratica è stato verificato il buon accoppiamento (*matching*) dei dati forniti dal simulatore (attesi) e le percentuali reali (osservate) di ciascun *output*, cioè sia tra le percentuali simulate e quelle osservate relative a ciascuna tipologia di pesce catturata (Figure 19, 20, 21, 22, 23), sia sull’insieme di tutte le osservazioni (Figura 24).

Gli altissimi valori di R^2 ottenuti (0,992 – 0,997 – 0,993 – 0,995 – 0,913 – 0,988, riportati rispettivamente nelle Figure 19, 20, 21, 22, 23, 24) dimostrano la validità del nostro simulatore, poiché esso riesce a classificare in maniera ottimale acciughe, sardine, suri, OPS e demersali all’interno di aree con banchi appartenenti a più specie.

Tabella 2. Tabella riassuntiva delle 11 cale utilizzate per il *training set* e delle 20 cale utilizzate per il *test set*.

	DATA	N°CALA
TRAINING SET	13/07/02	22
	26/06/05	7
	10/07/07	2
	13/07/07	4
	16/07/07	16
	14/07/09	10
	17/07/09	20
	29/06/11	8
	09/07/11	10
	11/07/11	18
	11/07/11	19
TEST SET	24/06/05	1
	25/06/05	3
	27/06/05	11
	28/06/05	15
	05/07/05	29
	20/06/06	3
	21/06/06	8
	26/06/06	22
	14/07/07	8
	05/08/08	2
	06/08/08	5
	12/07/09	2
	14/07/09	9
	15/07/09	11
	24/08/09	18
	25/08/09	21
	31/08/09	31
	27/05/11	21
	07/06/11	34
	29/06/11	7

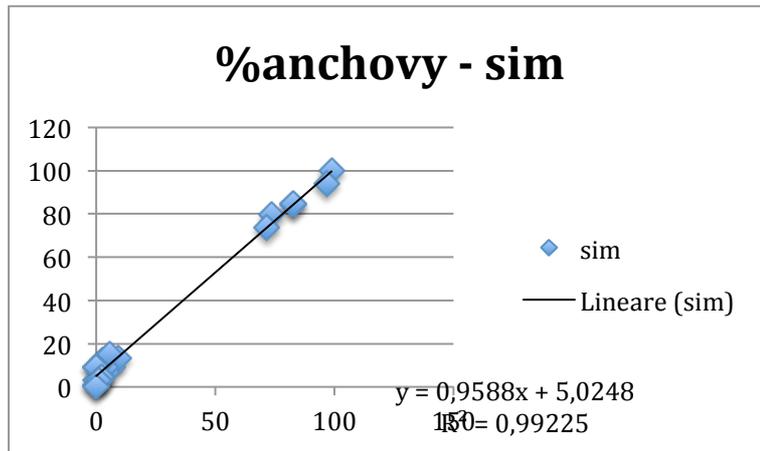


Figura 19. Grafico di correlazione tra % di *Engraulis encrasicolus* e % simulate.

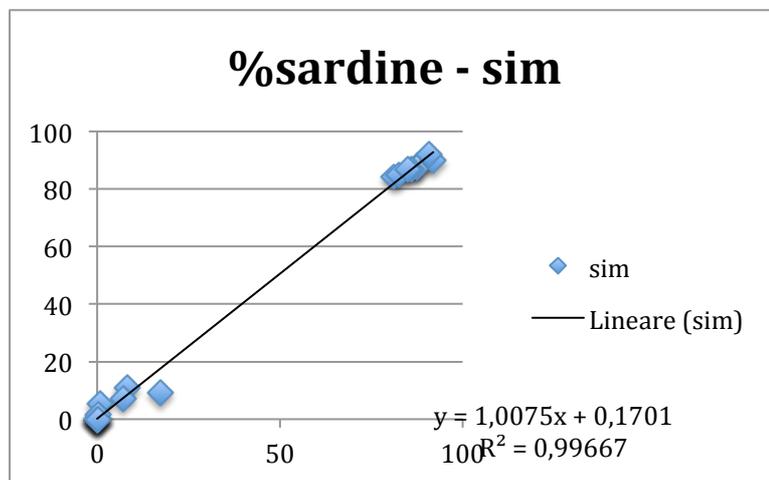


Figura 20. Grafico di correlazione tra %di *Sardina pilchardus* e % simulate.

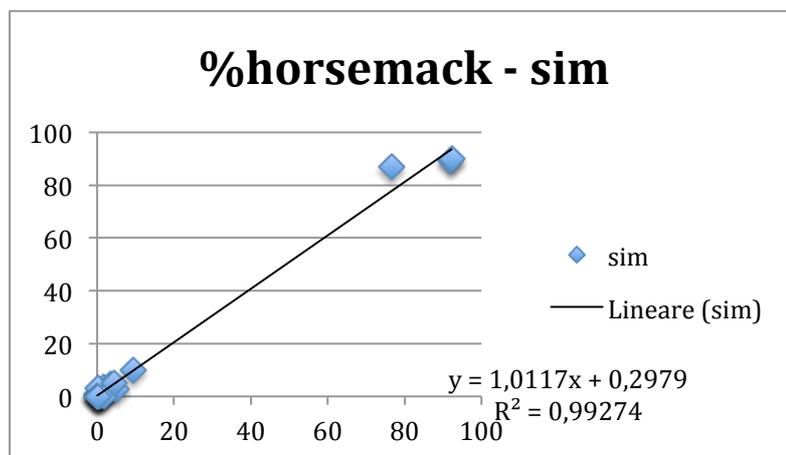


Figura 21. Grafico di correlazione tra % di *Trachurus trachurus* e % simulate.

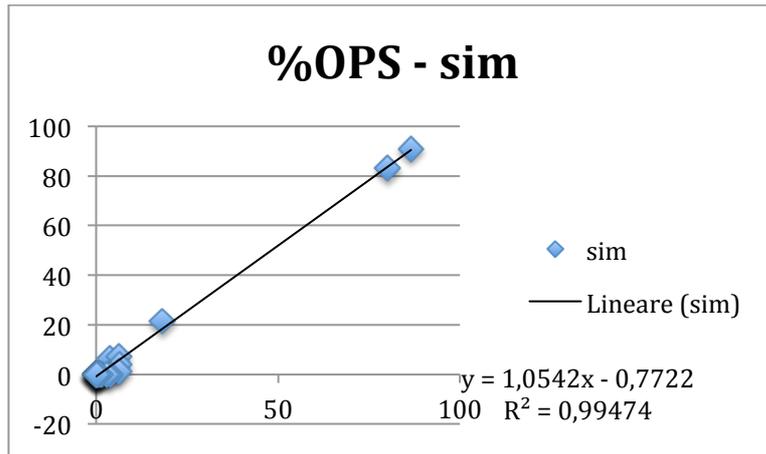


Figura 22. Grafico di correlazione tra % di “other pelagic species” e % simulate.

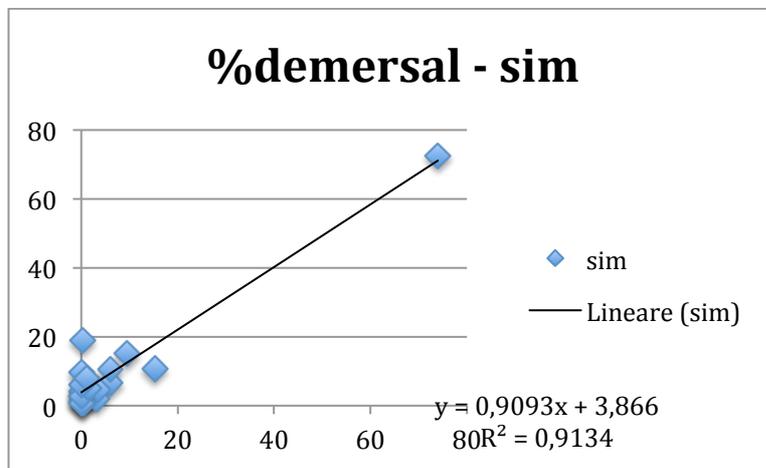


Figura 23. Grafico di correlazione tra % di pesci demersali e % simulate.

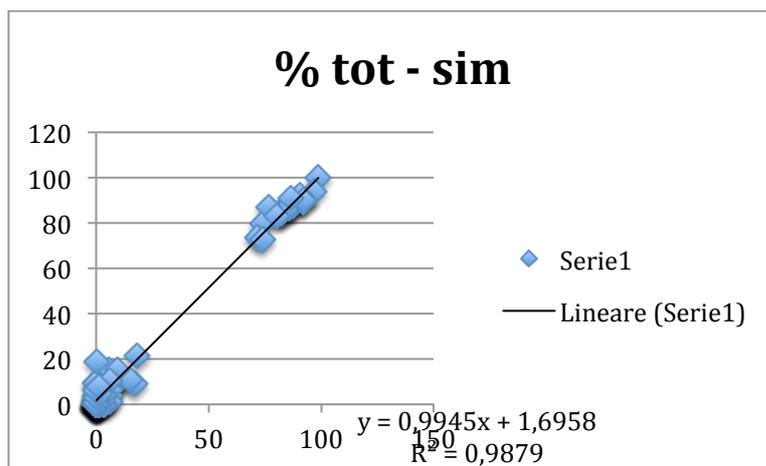


Figura 24. Grafico di correlazione tra % totali e % simulate.

5. DISCUSSIONE E CONCLUSIONI

I ripetuti riferimenti allo studio di Marta D'Elia (2012) mostrano già di per sé l'importanza che esso assume per lo sviluppo di questa tesi. In sede di discussione, appare opportuno illustrare in sintesi quanto fatto precedentemente durante la messa a punto del *classification tree*, in modo da chiarire il confronto col classificatore *fuzzy naïve Bayes* da me utilizzato.

Quando sono state raccolte informazioni abiotico-ambientali relative alla colonna d'acqua (ricavate nelle "stazioni CTD") e al fondo (ricavate con tecniche acustiche), assieme a parametri di tipo morfologico, batimetrico ed energetico dei banchi con segnali di più frequenze, l'obiettivo era quello di caratterizzare e identificare banchi di specie pelagiche *target* (acciughe e sardine) nel Canale di Sicilia. Il lavoro di caratterizzazione è stato sviluppato secondo due fasi: la prima utilizzava informazioni estratte da una sola frequenza a 38 kHz e la loro relazione con densità e biomassa del plancton (interpretata come indice della disponibilità di cibo) per caratterizzare i banchi e fare ipotesi sulla composizione in specie; con la seconda si volevano individuare dei parametri utili per discriminare le due specie, o con i descrittori del banco o con i parametri ambientali.

La prima fase si è articolata in ulteriori tre punti:

- 1- estrazione delle caratteristiche strutturali di un banco e individuazione del plancton tramite un algoritmo a singola frequenza;
- 2- analisi delle relazioni tra biomassa/densità dei banchi e plancton tramite GAMS, suddividendo il Canale di Sicilia in aree e distinguendo tra plancton superficiale e profondo;
- 3- caratterizzazione per aree di banchi e aggregazioni planctoniche, tramite analisi statistica descrittiva.

La seconda fase si è svolta secondo i seguenti 8 punti:

- 1- creazione e implementazione di un algoritmo a multifrequenza, per identificare banchi di pesce con vescica natatoria ed estrazione delle caratteristiche strutturali a più frequenze. L'algoritmo "*Fish school filter*" del progetto SINFAMI (Fernandes *et al.* 2006) è stato modificato e adattato ai banchi del Canale di Sicilia (cioè alle loro caratteristiche acustiche) ed è stato implementato nel modulo "*Virtual Variable*" del software *Sonardata Echoview* v.4.9, per rendere dati acustici a diverse frequenze comparabili nello spazio, cioè per confrontare le informazioni energetiche di vari banchi a varie

frequenze; il nuovo algoritmo, inoltre, elimina parte del segnale acustico di altre componenti presenti nella colonna d'acqua (come plancton o sedimento in sospensione) che nelle analisi rappresentano una fonte di rumore e che, nel caso si usi una sola frequenza, possono dar luogo ad erronee interpretazioni, ad esempio segnalarli come banchi di pesce; mantiene invece reale la struttura del banco, diversamente dal metodo tradizionale;

- 2- implementazione di un metodo acustico per estrarre le proprietà di durezza del fondo;
- 3- creazione di un *learning sample* per la classificazione, cioè creazione di un *database* di dati biologici (da cale monospecifiche) e dati acustici (rilevati in corrispondenza delle stesse cale);
- 4- analisi statistica descrittiva per la caratterizzazione batimetrica e morfologica dei banchi, tramite descrizione delle caratteristiche medie di ogni specie (distribuzione batimetrica e caratteristiche morfologiche estratte dalla singola frequenza a 38kHz);
- 5- analisi delle differenze nelle proprietà energetiche tra le due specie *target*; cioè i parametri energetici ricavati a differenti frequenze sono stati analizzati per valutarne l'efficacia nel riconoscere/classificare le due specie. Parametri energetici a più frequenze servono per identificare la specie in quanto legati alla presenza o meno della vescica natatoria, forma e taglia del pesce (Fernandes *et al.* 2006). E' sempre stato difficile distinguere tra acciughe e sardine, in quanto entrambe le specie presentano caratteristiche simili dal punto di vista acustico (ad esempio, entrambe hanno vescica natatoria ed entrambe formano banchi nelle ore diurne dalla struttura molto simile). Tuttavia le caratteristiche acustiche tra le due specie possono differire in quanto di solito le sardine hanno taglia maggiore rispetto alle acciughe, soprattutto in autunno quando le acciughe sono in fase di reclutamento.
- 6- analisi delle differenze nelle proprietà energetiche per *range* di profondità diversi. Poiché acciughe e sardine sono pesci *fisostomi* (presentano vescica natatoria aperta, ossia comunicante con l'apparato digerente per mezzo del dotto pneumatico), non presentano meccanismi per aggiustare attivamente il volume della vescica natatoria in risposta a cambiamenti di profondità e quindi di pressione: ecco perché queste specie possono mostrare diminuzione del volume

della vescica natatoria all'aumentare della pressione, secondo la legge di Boyle (si pensi a quanto osservato per le aringhe da Fässler *et al.*, 2009).

- 7- costruzione di modelli decisionali ad albero sulla base dei parametri strutturali del banco, per individuare i parametri meglio discriminanti delle specie e valutarne le capacità di classificazione;
- 8- costruzione di modelli decisionali ad albero sulla base di parametri ambientali e strutturali del banco, per valutarne le capacità di classificazione.

Il *classification tree* si è dimostrato molto efficiente con i banchi individuati nel corso di catture monospecifiche, mentre sono stati riscontrati problemi nell'identificazione di specie nel caso di cale plurispecifiche, vale a dire cale in cui non si osserva una netta predominanza di una specie di pesce sulle altre.

Questo “deficit” classificatorio può essere imputabile alla filosofia stessa dei *classification tree*, il cui obiettivo è operare assegnazioni utilizzando caratteristiche discriminanti univoche nella ricerca di una buona classificazione con la minor spesa computazionale possibile.

A differenza del *classification tree*, il compilatore *fuzzy naïve Bayes* non ricerca in maniera prioritaria il risparmio computazionale ma include in sé un maggior numero di variabili in modo che, in fase di classificazione, entrino in gioco quelle dinamiche complesse che definiscono il sistema in esame.

Un altro punto a favore della classificazione *fuzzy bayesiana* è la continuità con cui l'informazione viene *splittata* nella diverse categorie che, essendo *fuzzy* e non discrete, richiedono un minor grado di approssimazione nell'assegnare un valore ad una classe piuttosto che ad un'altra.

Le differenze riscontrate fra i funzionamenti dei due approcci mostrano come il *classification tree* sia più affidabile in casi di banchi monospecifici, mentre sia completamente inaffidabile in caso di banchi dalla composizione maggiormente eterogenea.

Pensando ad un utilizzo effettivo di tali algoritmi per le valutazioni degli *stock* in fase di gestione della pesca, è evidente come la resa migliore potrebbe essere data da un'integrazione dei due approcci in modo da far emergere le peculiarità di ognuno ed allo stesso tempo *bypassare* le problematiche dei due algoritmi presi singolarmente.

Tornando alla discussione dei risultati ottenuti nel contesto di questo lavoro di tesi, non si può fare a meno di notare come questi siano stati “troppo buoni”. Si è proceduto infatti ad una calibrazione e validazione del classificatore su un grande numero di osservazioni in

zone molto simili per condizioni ambientali al contorno. Per cui ci si può aspettare, al momento dell'utilizzo reale del classificatore, una resa meno buona rispetto alle osservazioni in quanto non è pensabile per ogni sito una fase di calibrazione/validazione estesa come quella effettuata in fase sperimentale.

Inoltre va verificata l'applicabilità a specie diverse e, soprattutto, in condizioni al contorno differenti per variabilità intrinseca.

La soluzione migliore che si pensa di portare avanti nel futuro prossimo è quella di integrare i due approcci utilizzando la sicurezza del *classification tree* come base di classificazione in cui il *fuzzy naïve Bayes* potrebbe fungere da filtro, calibrando i parametri decisionali dell'albero *bayesiano* in funzione delle condizioni ambientali al contorno e dell'incertezza ad esse associata.

BIBLIOGRAFIA:

Abad R., Miquel J., Iglesias M.. 1996. Campañas de evaluación por métodos acústicos de sardina, boquerón y ochavo en el Mediterráneo Occidental. *FAO Fish Rep.*, 537: 191-193.

Abella A., Belluscio A., Bertrand J., Carbonara P.L., Giordano D., Sbrana M., Zamboni A.. 1999. Use of MEDITS trawl survey data and commercial fleet information for the assessment of some Mediterranean demersal resources. *Aquat. Living Resour.*, 12(3): 155-166.

Abellò P., Bertrand J. A., Gil de Sola L., Papaconstantinou C., Relini G., Souplet A. (Eds.). 2002. Mediterranean Marine Demersal Resources: The MEDITS international Trawl Survey (1994-1999) *Sci. Mar.*, 66 (Suppl. 2).

Agenbag J.J., Richardson A.J, Demarcq H., Frèon P., Weeks S., Shillington F.A.. 2003. Estimating environmental preferences of South African pelagic fish species using catch size- and remote sensing data. *Progress in Oceanography* 59(2-3): 275-300.

Aglen A.. 1994. Sources of error in acoustics estimation of fish abundance. In: *Marine Fish Behaviour in Capture and Abundance Estimation*, A. Fernø, & S. Olsen (Eds.), 107-133, Fishing News Books, Oxford.

Alvares F., Cingolati N., Kallianiotis A.. 2003. Piccoli pelagici: gestione futura delle risorse condivise in Mediterraneo. *Il Pesce* N° 5, 57-65.

Anticamara J.A., Watsona R., Gelchu A., Pauly D.. 2011. Global fishing effort (1950–2010): Trends, gaps, and implications. *Fisheries Research* 107 (2011) 131–136.

Ardizzone G. D. & Corsi F. (Eds.). 1997. Atlas of Italian demersal fishery resources. Trawl surveys 1985-1987. *Biol. Mar. Medit.*, 4(2): 1-568.

Arreguín-Sánchez F.. 1996. Catchability: a key parameter for fish stock assessment. *Reviews in Fish Biology and Fisheries* 6, 221–242.

Balabanian J.-P., Viola I., Ona E., Patel R., Gröller M.E.. 2007. Sonar Explorer: A new tool for visualization of fish schools from 3D sonar data. *Proceedings of Data Visualization - EuroVis 2007*, ISBN: 9783905673456, Norköpping, Sweden, May 23-25 2007.

Bertrand J., Gil de Sola L., Papaconstantinou C., Relini G., Souplet A.. 1998. An international bottom trawl survey in the Mediterranean: the MEDITS programme. In: J.A. Bertrand and G. Relini (co-ordinators). *Demersal Resources in the Mediterranean. Actes de Colloques IFREMER n° 26*: 76-93.

Bertrand A., Josse E., Bach P. & Dagorn L.. 2003. Acoustics for ecosystem research: lessons and perspectives from a scientific programme focusing on tuna-environment relationships. *Aquatic Living Resources* 16, pp. 197-203.

Bonanno A., Patti B., Goncharov S., Basilone G., Cuttitta A., Buscaino G., García Lafuente J., García A., Palumbo V., Cancemi M., Mazzola S.. 2005. Fluctuation of sardine and anchovy abundance in the Strait of Sicily investigated by acoustic surveys. *MedSudMed Technical Documents*, 5: 39-47.

Bozma Ö. & Kuc R.. 1991. Building a sonar map in a specular environment using a single mobile sensor. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 13(12), 1260-1269.

Bozzeda F.. 2013. Sviluppo di un approccio fuzzy-bayesiano per l'analisi e la modellizzazione degli ecosistemi: applicazione ad ecosistemi marini costieri. Tesi di dottorato in Scienze Ambientali. Università degli studi di Bologna.

Bustamante C., Garrido L. and Soto R.. 2007. Fuzzy naive Bayesian classification in RoboSoccer 3D: A Hybrid Approach to Decision Making. *RoboCup 2006: Robot Soccer World Cup X*. Pp. 507-515.

Cabreira A. G., Tripode M., Madirolas A.. 2009. Artificial neural networks for fish-species identification. *ICES Journal of Marine Science*, 66: 1119-1129.

Caddy J.F.. 1993. Some future perspectives for assessment and management of Mediterranean fisheries. In: J. Lleonard (Ed.) *Northwestern Mediterranean fisheries*. *Sci. Mar.*, 57 (2-3): 121-130.

Castillo J., Barbieri M. A., Gonzalez A.. 1996. Relationships between sea surface temperature, salinity, and pelagic fish distribution off northern Chile. – *ICES Journal of Marine Science*, 53: 139–146.

Cressie N., Calder C.A., Clark J.S., Ver Hoef J.M. and Wikle C.K.. 2009. Accounting for uncertainty in ecological analysis: the strengths and limitations of hierarchical statistical modelling. *Ecological Applications* 19: 553-570.

D’Elia M., Patti B., Sulli A., Tranchida G., Bonanno A., Basilone G., Giacalone G., Fontana I., Genovese S., Guisande C., and Mazzola S. 2009. Distribution and spatial structure of pelagic fish schools in relation to the nature of the seabed in the Sicily Channel (Central Mediterranean), *Mar. Ecol.* 30, pp. 151–160.

D’Elia M. 2012. Caratterizzazione e identificazione di banchi di specie pelagiche mediante l’uso di descrittori strutturali e parametri ambientali. Tesi di dottorato in Scienze Ambientali. Ca’ Foscari, Venezia.

D’Elia M., Patti B., Bonanno A., Fontana I., Giacalone G., Basilone G., Fernandes P.G.. 2014. Analysis of backscatter properties and application of classification procedures for the identification of small pelagic fish species in the Central Mediterranean. *Fisheries Research* 149: 33–42.

Diner N., Masse J.. 1987. Fish school behaviour during echo survey observed by acoustic devices. *Proceedings of the International Symposium on Fisheries Acoustics*, Seattle, USA, June 22-26, 1987.

Diner N.. 2001. Correction on school geometry and density: approach based on acoustic image simulation. *Aquatic Living Resources*, 14: 211-222.

Edwards J. I., Armstrong F..

1983. Measurement of the target strength of live herring and mackerel. *FAO Fish. Rep.* 300, 69-77.

Ehrenberg J. E., Torkelson T. C.. 1996. Application of dual-beam and split-beam target tracking in fisheries acoustics. *ICES Journal of Marine Science*, 53: 329-334.

Engås A. & Løkkeborg S.. 1994. Abundance estimation using bottom gillnet and longline – the role of fish behaviour. In *Marine Fish Behaviour in Capture and Abundance Estimation* (Fernö A. & Olsen S., eds), pp. 134–165. Oxford: Fishing News Books.

FAO. 2009. *State of the World's Fisheries and Aquaculture 2008*. Food and Agriculture Organization of the United Nations, Rome.

Fässler S.M.M., Santos R., Garcia-Núñez N., Fernandes P.G.. 2007. Multifrequency backscattering properties of Atlantic herring (*Clupea harengus*) and Norway pout (*Trisopterus esmarkii*). *Can. J. Fish. Aquat. Sci.* 64, 362–374.

Fässler S.M.M., Brierley A. S., Fernandes P.G.. 2009. A Bayesian approach to estimating target strength. *ICES Journal of Marine Science*, 66: 1197-1204.

Fernandes P.G., Gerlotto G., Holliday D.V., Nakken O., Simmonds E.J.. 2002. Acoustic applications in fisheries science: the ICES contribution. *ICES Mar. Sci. Symp.* 215, 483-492.

Fernandes P. G., Korneliussen R. J., Lebourges-Dhaussy A., Masse J., Iglesias M., Diner N., Ona E., *et al.*. 2006. The SIMFAMI Project: Species Identification Methods from Acoustic Multifrequency Information. Final Report to the EC, Q5RS-2001–02054.

Fernandes P. G. 2009. Classification trees for species identification of fish-school echotraces. *ICES Journal of Marine Science*, 66: 1073–1080.

Fleischman S.J., Burwen D.L.. 2003. Mixture models for the species apportionment of hydroacoustic data, with echo-envelope length as the discriminatory variable. *ICES Journal of Marine Science*, 60: 592–598.

Fréon P., Soria M., Mullon C., Gerlotto F.. 1993. Diurnal variation in fish density estimate during acoustic surveys in relation to spatial distribution and avoidance reaction. *Aquatic Living Resources*, Vol.6, N°3, pp. 221-234.

Furevik D. M.. 1994. Behaviour of fish in relation to pots. In *Marine Fish Behaviour in Capture and Abundance Estimation* (Fernö A. & Olsen S., eds), pp. 28–44. Oxford: Fishing News Books.

Gabriel O., Lange K., Dahm E., Wendt T. (Eds.). 2005. *Von Brandt's Fish Catching Method of the World*. Blackwell Publishing Ltd., Oxford, UK.

García Lafuente J., García A., Mazzola S., Quintanilla L., Delgado J., Cuttitta A., Patti, B.. 2002. Hydrographic phenomena influencing early life stages of the Sicilian Channel anchovy. *Fisheries Oceanography*, 11(1), 31–44.

Garcia S.M., Zerbi A., Aliaume C., Do Chi T., Lasserre G.. 2003. The ecosystem approach to fisheries. Issues, terminology, principles, institutional foundations, implementation and outlook. *FAO Fisheries Technical Paper*. No. 443. Rome, FAO. 2003. 71 p..

Georgakarakos S.. 2005. Investigation of the beam effect on fish school parameters. A simulation approach. In *Proceedings of the 1st International Conference on Underwater Acoustic Measurements: Technologies and Results, 28 June – 1 July 2005*. J.S. Papadakis & L. Bjørnø (Ed), FORTH/IACM, Heraklion, Greece. 6 pp. ISBN 960-88702-08.

Georgakarakos S., Trygonis V., Haralabous J.. 2011. Accuracy of Acoustic Methods in Fish Stock Assessment Surveys. In: *Sonar Systems*, InTech, ISBN: 978-953-307-345-3.

Gerlotto F., Fréon P.. 1992. Some elements on vertical avoidance of fish schools to a vessel during acoustic surveys. *Fisheries Research*, Vol.14, N°4, pp. 251-259.

Gerlotto F., Georgakarakos S., Eriksen P.K.. 2000. The application of multibeam sonar technology for quantitative estimates of fish density in shallow water acoustic surveys. *Aquatic Living Resources*, Vol.13, No.5, pp. 385-393.

Giannoulaki M., Machias A., Somarakis S., Tsimenides N.. 2005. The spatial distribution of anchovy and sardine in the northern Aegean Sea in relation to hydrographic regimes. *Belgian Journal of Zoology*, 135: 151e156.

Greenstreet S.P.R., McMillan J.A., Armstrong F.. 1998. Seasonal variation in the importance of pelagic fish in the diet of piscivorous fish in the Moray Firth, NE Scotland: a response to variation in prey abundance?. *ICES J. Mar. Sci.* 55: 121-133.

Guennegan Y., Liorzou B., Bigot J.L.. 2000. Exploitation des petits pelagiques dans le Golf du Lion et suivi de l'évolution des stocks par echo-integration de 1999 a 2000. Paper presented at WG on small Pelagics. Fuengirola (Spain) 1-3 March 2000: 27p.

Gulland J. A.. 1983. Stock assessment: why?. *FAO Fisheries Circular N°759*, pp.1-21.

Gulland J.A.. 1969. Manual of methods for fish stock assessment. Part 1. Fish population analysis. *FAO Man.Fish.Sci.* (4):154 p.

Hafsteinsson M.T., Misund O.A.. 1995. Recording the migration behaviour of fish schools by multi-beam sonar during conventional acoustic surveys. *ICES Journal of Marine Science*, Vol.52, N°6, pp. 915-924.

Hastie T. and Tibshirani R.. 1990. *Generalised Additive Models*, Chapman and Hall, London.

Henderson H.F. and Welcomme R.L.. 1974. The relationship of yield to morpho-edaphic index and numbers of fishermen in African inland waters. *Relation entre la production,*

l'indice morpho- edaphique et le nombre de pecheurs des pecheries des eaux continentales. CIFA. Occas.Pap./Doc.Occas.CPCA, (1):19 p.

Horne J.. 2000. Acoustic approaches to remote species identification: a review. *Fisheries Oceanography*, 9: 356–371.

Huse I., Ona E.. 1996. Tilt angle distribution and swimming speed of overwintering Norwegian spring spawning herring. *ICES Journal of Marine Science*, 53: 863-873.

ICES 2000. Report of the Working Group on Fisheries Acoustics Science and Technology (WG-FAST), Ijmuiden/Haarlem, Netherlands, 10-14 April 2000. International Council for the Exploration of the Sea, Copenhagen (Denmark). ICES CM 2000/B:04, 35 pp.

ICES 2004g. Report of the planning group on Northeast Atlantic Pelagic Ecosystem Surveys (PGNAPES). ICES CM 2004g/D:07, 57 pp.

IREPA, 2010. Osservatorio economico sulle strutture produttive della pesca marittima in Italia 2008. Napoli: Edizioni Scientifiche Italiane, pp. 184.

John G.H. and Langley P.. 1995. Estimating continuous distributions in Bayesian classifiers. Pages 338-345 in P. Besnard and S. Hanks editors. *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence*, Morgan Kaufmann Publishers, San Mateo.

Kaimmer S. M.. 1999. Direct observations on the hooking behavior of Pacific halibut, *Hippoglossus stenolepis*. *Fishery Bulletin* 97, 873–883.

Kang M., Furusawa M., Miyashita K.. 2002. Effective and accurate use of difference in mean volume backscattering strength to identify fish and plankton. *ICES Journal of Marine Science*, 59: 794-804.

Kloser R., Ryan T., Sakov P., Williams A., Koslow J.A.. 2002. Species identification in deep water using multiple acoustic frequencies. *Can. J. Fish. Aquat. Sci.* 59: 1065-1077.

Koslow J.A.. 2009. The role of acoustics in ecosystem-based fishery management. *ICES Journal of Marine Science*, 66: 966-973.

Korneliussen R. J., Ona E.. 2002. An operational system for processing and visualizing multi-frequency acoustic data. *ICES Journal of Marine Science*, 159: 293-313.

Korneliussen R. J., Heggelund Y., Eliassen I. K., Johansen G. O.. 2009. Acoustic species identification of schooling fish. *ICES Journal of Marine Science*, 66: 1111-1118.

Lavery A.C., Wiebe P.H., Stanton T.K., Lawson G.L., Benfield M.C., and Copley N.. 2007. Determining dominant scatterers of sound in mixed zooplankton populations. *Journal of the Acoustical Society of America*, 122: 3304–3326.

Lawson G.L., Barange M. and Frèon. 2001. Species identification of pelagic fish schools on the South African continental shelf using acoustic descriptors and ancillary information. *ICES J. Mar. Sci.* 58: 275-287.

Levenez J.-J., Gerlotto F., Petit D.. 1990. Reactions of tropical coastal pelagic species to artificial lighting and implications for the assessment of abundance by echo integration. *Rapports et Procès-Verbaux des Réunions du Conseil International pour l'Exploration de la Mer*, Vol. 189, pp. 128-314

Lleonard J. & Maynou F.. 2003. Fish stock assessments in the Mediterranean: state of art. *Sci. Mar.*, 67 (Suppl. 1): 37-49.

Løkkeborg S. (1994). Fish behaviour and longlining. In *Marine Fish Behaviour in Capture and Abundance Estimation* (Fernö A. & Olsen S., eds), pp. 9–27. Oxford: Fishing News Books.

Løkkeborg S., Olla B. L., Pearson W. H. & Davis M. W.. 1995. Behavioural responses of sablefish, *Anoplopoma fimbria*, to bait odour. *Journal of Fish Biology* 46, 142–155.

Lundgren B., Nielsen J.R.. 2008. A method for the possible species discrimination of juvenile gadoids by broad-bandwidth backscattering spectra vs. angle of incidence. *ICES Journal of Marine Science*, 65: 581–593.

MacLennan D.N., and Holliday D.V.. 1996. Fisheries and plankton acoustics: past, present, and future. *ICES J. Mar. Sci.* 53: 513-516.

Madureira L.S.P., Everson I. and Murphy E.J.. 1993. Interpretation of acoustic data at two frequencies to discriminate between Antarctic krill (*Euphausia superba* Dana) and other scatterers. *Journal of Plankton Research*, 15(7): 787-802.

Massè J., Rouxel C.. 1991. Improvement in acoustic assessments by discrimination of pelagic shoals with INES/MOVIES system. *ICES CM 1991/B:26*. 8 pp.

Massè J., Koutsikopoulos C., Patty W.. 1996. The structure and spatial distribution of pelagic fish schools in multispecies clusters: an acoustic study. *ICES Journal of Marine Science* 53(2): 155-160.

Mazzola S., Garcia A., Garcia Lafuente J.. 2002. The Sicilian channel anchovy fishery and the underlying oceanographic and biological processes conditioning their inter-annual fluctuations. DG XIV, MED 98-070 Final Report.

Misund O.A.. 1990. Sonar observation of schooling herring: school dimensions, swimming behaviour and avoidance of vessel and purse seine. *Rapports et Procès-Verbaux des Réunions du Conseil International pour l'Exploration de la Mer*, Vol. 189, pp. 135-146.

Misund O.A.. 1997. Underwater acoustics in marine fisheries and fisheries research. *Reviews in Fish Biology and Fisheries*, Vol.7, No.1, pp. 1-34.

Mitson R.B., Knudsen H.P.. 2003. Causes and effects of underwater noise on fish abundance estimation. *Aquatic Living Resources*, Vol.16, N°3, pp. 255-263.

Munro J.L. (ed.). 1982. Caribbean coral reef fishery resources. *ICLARM Stud.Rev.*

Nøttestad L., Aksland M., Beltestad A., Ferno A., Johannessen A., Misund O.A.. 1996. Schooling dynamics of Norwegian spring spawning herring (*Clupea harengus* L.) in a coastal spawning area. *Sarsia* 80, 277–284.

Olsen K.. 1990. Fish behaviour and acoustic sampling. *Rapports et Procés-Verbaux des Réunions du Conseil International pour l'Exploration de la Mer*, Vol.189, pp. 147-158.

Ona E., Georgakarakos S., Simmonds E.J.. 1998. Acoustic properties of fish and their exploitation in classification schemes for fish stock assessment. In (ed.), *Third european marine science and technology conference MAST conference*, Lisbon, 23-27 May 1998: *Projectsynopses Vol.5: Fisheries and Aquaculture AIR: 1990-94 Selected projects from the research programme for Agriculture and Agro Industry including Fisheries*, 241-244.

Ona E., Godø O.R.. 1990. Fish reaction to trawling noise: the significance for trawl sampling. *Rapports et Procés-Verbaux des Réunions du Conseil International pour l'Exploration de la Mer*, Vol.189, pp. 159-166.

Øvergaard S. and Skjøsberg E.. 2012. Next-generation echo sounder soon ready. The Research Council of Norway.

Paramo J., Bertrand S., Villalobos H., Gerlotto F.. 2007. A three-dimensional approach to school typology using vertical scanning multibeam sonar. *Fisheries Research*, 84: 171-179.

Patti B., Mazzola S., Calise L., Bonanno A., Buscaino G., Cosimi G.. 2000. Echosurveys estimates and distribution of small pelagics Fish concentration in the Strait of Sicily during June 1998. Paper presented at WG on small Pelagics. Fuengirola (Spain) 1-3 March 2000: 11p.

Patti B., Bonanno A., Basilone G., Goncharov S., Mazzola S., Buscaino G., Cuttitta A., García Lafuente J., García A., Palumbo V., Cosimi G.. 2004. Interannual fluctuations in acoustic biomass estimates and in landings of small pelagic fish populations in relation to hydrology in the Strait of Sicily. *Chem. Ecol.* 20: 365–375.

Patti B., Bonanno A., Mazzola S., Basilone G., Buscaino G., Cuttitta A., Gontcharov S., Aronica S., Giacalone G., Fontana I., Tegolo D.. 2006. Distribution of fish schools and plankton patches along the southern coast of Sicily in relation to hydrology. 4-8 Settembre 2006 41° European Marine Biology Symposium-EMBS, Cork.

Patti B., Bonanno A., D'Elia M., Quinci E., Giacalone G., Fontana I., Aronica S., Basilone G. and Mazzola S.. 2011. Daytime pelagic schooling behaviour and relationships with plankton patch distribution in the Sicily Strait (Mediterranean Sea)', *Advances in Oceanography and Limnology*, 2: 1, 79-92.

Punsly, R. G. & Nakano, H.. 1992. Analysis of variance and standardization of longline hook rates of bigeye (*Thunnus obesus*) and yellowfin (*Thunnus albacares*) in the eastern Pacific Ocean during 1975–1987. *Inter-American Tropical Tuna Commission Bulletin* 20, 165–184.

Randon N.J., Lawry J., Horsburgh K. and Cluckie I.D.. 2008. Fuzzy Bayesian modelling of sea-level along the east coast of Britain. *IEEE Transactions on Fuzzy Systems* 16: 725-737.

Reid D., Scalabrin C., Petitgas P., Massè J., Aukland R., Carrera P., and Georgakarakos S.. 2000. Standard protocols for the analysis of school based data from echo sounder surveys. *Fisheries Research*, 47: 125–136.

Richards L. J., Kieser R., Mulligan T. J., Candy, J. R.. 1991. Classification of fish assemblages based on echo integration surveys. *Canadian Journal of Fisheries and Aquatic Sciences*, 48: 1264–1272.

Ricker W.E.. 1958. Handbook of computations for biological statistics of fish populations. *J.Fish.Res.Board Can.*, (119): 300 p.

Rose G.A.. 1995. Cod spawning on a migration highway in the north-west Atlantic. *Nature* 366: 458-461.

Rose G.A., Leggett W.C.. 1988. Hydroacoustical signal classification of fish schools by species. *Can. J. Fish. Aquat. Sci* 45: 597-607.

Sainsbury K.J., Punt A.E., Smith A.D.M.. 2000. Design of operational management strategies for achieving fishery ecosystem objectives. *ICES J. Mar. Sci.* 57: 731-741.

Saville A.. 1977. Survey methods of appraising fishery resources. *FAO Fish.Tech.Pap.*, (171):76 p.

Scalabrin C., Diner N., Weill A., Hillion A. and Mouchot M.C.. 1996. Narrow band acoustic identification of monospecific fish shoals. *ICES J. Mar. Sci.* 53: 181-188.

Simmonds E.J., Armstrong F., Copland P.J.. 1996. Species identification using wideband backscatter with neural network and discriminant analysis. *ICES J. Mar. Sci.* 53: 189-195.

Simmonds E.J.. 2003. Weighting of acoustic- and trawl-survey indices for the assessment of North Sea herring. *ICES J. Mar. Sci.* 60: 463-471.

Simmonds J., MacLennan D.. 2005. *Fisheries acoustics: theory and practice*, 2nd edn. Blackwell Publishing, Oxford.

Soria M., Bahri T., Gerlotto F.. 2003. Effect of external factors (environment and survey vessel) on fish school characteristics observed by echosounder and multibeam sonar in the Mediterranean Sea. *Aquatic Living Resources*, Vol.16, No.3, pp. 145-157.

Stokesbury K.D.E., Kirsch J., Patrick E.V., Norcross B.L.. 2002. Natural mortality estimates of juvenile Pacific herring (*Clupea pallasii*) in Prince William Sound, Alaska. *Can. J. Fish. Aquat. Sci.* 59(3): 416-423.

Stoner A. W.. 2003. Hunger and light level alter response to bait by Pacific halibut: laboratory analysis of detection, location and attack. *Journal of Fish Biology* 62, 1176–1193.

Stoner A. W.. 2004. Effects of environmental variables on fish feeding ecology: implications for the performance of baited fishing gear and stock assessment. *Journal of Fish Biology* (2004) 65, 1445–1471.

Störr H.P.. 2002. A compact fuzzy extension of the Naive Bayesian classification algorithm. Pages 172-177 in *Proceedings InTech/VJFuzzy*.

Swartz W., Rashid Sumaila U., Watson R., Pauly D.. 2010. Sourcing seafood for the three major markets: the EU, Japan and the USA. *Mar. Policy* 34, 1366–1373.

Swartzman G., Stuetzle W., Kulman K., Powojowski M.. 1994. Relating the distribution of pollock schools in the Bering Sea to environmental factors. *ICES Journal of Marine Science*, 51: 481–492.

Swartzman G., Bertrand A., Gutiérrez M., Bertrand S., Vasquez L.. 2008. The relationship of anchovy and sardine to water masses in the Peruvian Humboldt Current System from 1983–2005. *Progress in Oceanography* 79: 228–237.

Tang Y. & Xu Y.. 2005. Application of fuzzy Naive Bayes and a real-valued genetic algorithm in identification of fuzzy model. *Information Science*. Volume 169, Issues 3-4, Pages 205–226.

Tegowski J., Gorska N., Klusek Z.. 2003. Statistical analysis of acoustic echoes from underwater meadows in the eutrophic Puck Bay (southern Baltic Sea). *Aquatic Living Resources*, 16: 215–221.

Trenkel V.M., Mazauric V., Berger L.. 2008. The new fisheries multibeam echosounder ME70: description and expected contribution to fisheries research. *ICES Journal of Marine Science*, Vol.65, No.4, pp. 645-655.

Trygonis V., Georgakarakos S., Simmonds E.J.. 2009. An operational system for automatic school identification on multibeam sonar echoes. *ICES Journal of Marine Science*, 66: 935-949.

Tugores M.P., Giannoulaki M., Iglesias M., Bonanno A., Tičina V., Leonori I., Machias A., Tsagarakis K., Diaz N., Giráldez A., Patti B., De Felice A., Basilone G., Valavanis V.. 2011. Habitat suitability modelling for sardine *Sardina pilchardus* in a highly diverse ecosystem: the Mediterranean Sea. Mar. Ecol. Prog. Ser. 443: 181-205.

Viertl R. and D. Hareter. 2004. Fuzzy information and imprecise probability. ZAMM. Zeitschrift für Angewandte Mathematik und Mechanik 84: 731-739.

Walsh M., Reid D. G., Turrell W. R.. 1995. Understanding mackerel migration off Scotland: tracking with echosounders and commercial data, and including environmental correlates and behaviour. ICES Journal of Marine Science, 52: 925–939.

Watkins J.L., Brierley A.S.. 2002. Verification of the acoustic techniques used to identify Antarctic krill. ICES J. Mar. Sci. 59: 1326-1336.

Weill A., Scalabrin C., Diner N.. 1993. MOVIES-B: An acoustic detection description software. Application to shoal species' classification. Journal of Aquatic Living Resources 6(3): 255-267.

Welcomme R.L.. 1979. Fisheries ecology of floodplain fisheries. London, Longman, 317 p..

Widyantoro D. H. and Yen J.. 2000. A fuzzy similarity approach in text classification task. In Proceedings of the IEEE International Conference on Fuzzy Systems, FUZZ-IEEE.

World Bank. 2009. The Sunken Billions: The Economic Justification for Fisheries Reform. The International Bank for Reconstruction and Development/The World Bank, Washington, DC.

Zadeh L. A.. 1965. Fuzzy sets. Information and Control 8: 338-353.

Zadeh L. A.. 1968. Probability measures of fuzzy events. Journal of mathematical analysis and applications 23: 421-427.

SITOGRAFIA:

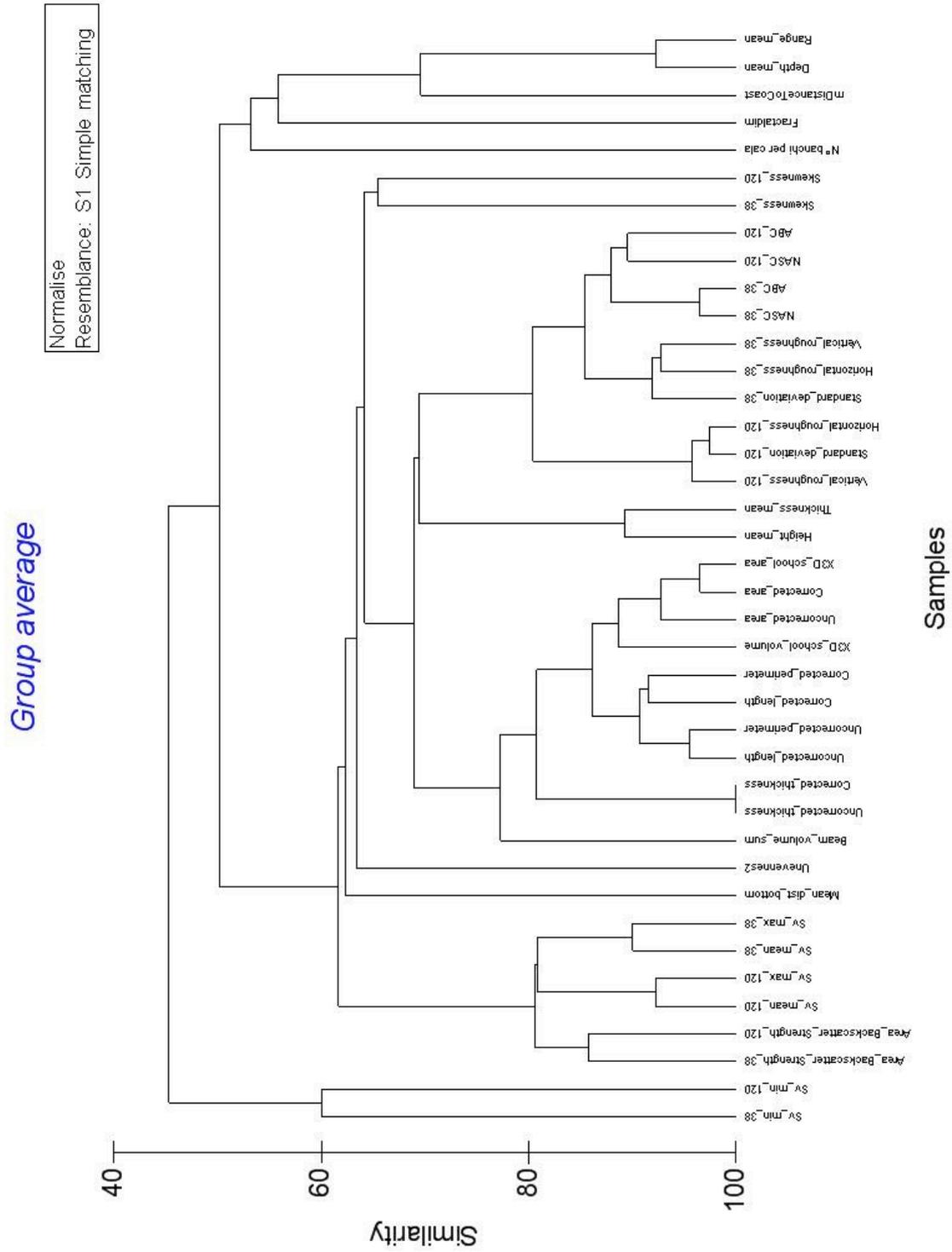
Airmar technology corporation: www.airmartechology.com

Food and Agriculture Organization of the United Nations (FAO): www.fao.org

Treccani: www.treccani.it

APPENDICE:

Appendice 1. Cluster graph delle 40 variabili



Appendice 2. Tabella esplicativa del *Cluster Graph* (Appendice1)

		Parameters
CLUSTER		Cluster mode: Group average
Hierarchical Cluster analysis		Combining
Resemblance worksheet		8+14 -> 41 at 100
Name: Resem4		30+32 -> 42 at 97,52
Data type: Similarity		16+19 -> 43 at 96,53
Selection: All		2+21 -> 44 at 96,53
		33+42 -> 45 at 95,78
		7+9 -> 46 at 95,53
Samples		17+18 -> 47 at 92,8
1 Sv_mean_38		10+43 -> 48 at 92,8
2 NASC_38		6+24 -> 49 at 92,31
3 Sv_max_38		26+28 -> 50 at 92,31
4 Sv_min_38		11+47 -> 51 at 91,94
5 Height_mean		13+15 -> 52 at 91,56
6 Depth_mean		46+52 -> 53 at 90,69
7 Uncorrected_length		1+3 -> 54 at 90,07
8 Uncorrected_thickness		27+34 -> 55 at 89,58
9 Uncorrected_perimeter		5+23 -> 56 at 89,33
10 Uncorrected_area		20+48 -> 57 at 88,67
11 Standard_deviation_38		44+55 -> 58 at 87,97
12 Skewness_38		53+57 -> 59 at 86,14
13 Corrected_length		22+35 -> 60 at 85,86
14 Corrected_thickness		51+58 -> 61 at 85,44
15 Corrected_perimeter		50+54 -> 62 at 80,89
16 Corrected_area		41+59 -> 63 at 80,8
17 Horizontal_roughness_38		60+62 -> 64 at 80,65
18 Vertical_roughness_38		45+61 -> 65 at 80,44
19 X3D_school_area		25+63 -> 66 at 77,3
20 X3D_school_volume		40+49 -> 67 at 69,6
21 ABC_38		56+65 -> 68 at 69,48
22 Area_Backscatter_Strength_38		66+68 -> 69 at 68,96
23 Thickness_mean		12+31 -> 70 at 65,51
24 Range_mean		69+70 -> 71 at 64,19
25 Beam_volume_sum		36+71 -> 72 at 63,43
26 Sv_mean_120		38+72 -> 73 at 62,34
27 NASC_120		64+73 -> 74 at 61,66
28 Sv_max_120		4+29 -> 75 at 60,05
29 Sv_min_120		37+67 -> 76 at 55,83
30 Standard_deviation_120		39+76 -> 77 at 53,23
31 Skewness_120		74+77 -> 78 at 50,22
32 Horizontal_roughness_120		75+78 -> 79 at 45,35
33 Vertical_roughness_120		
34 ABC_120		Outputs
35 Area_Backscatter_Strength_120		Plot: Graph3
36 <u>Unevennes2</u>		
37 <u>FractalDim</u>		
38 <u>Mean_dist_bottom</u>		
39 N [*] banchi per cala		
40 mDistanceToCoast		

Appendice 3. Percentuali reali e percentuali simulate dal classificatore *fuzzy naïve*

Bayes

	A	B	C	D	E	F	G	H	I	J
1	% anchovy	sim	%sardine	sim	%Horse maci	sim	%OPS	sim	%Demersal	sim
2	98,75	100	0	0	0,83	1	0	0	0	1
3	82,57	84,153333	17,2	9,18	3	3,7633333	0	0,5	0	2,7366667
4	6,78	9,11	8,06	10,89	0,86	0	80,2	83	1,14	3
5	9,17	13,27	0	0	76,53	87	3,74	6,35	6,04	6,62
6	0	3,1	0	0	5	2,75	18,17	21,5	73,97	72,65
7	3,65	8,815	0	0	92,04	89,185	0	0	3,23	2
8	2,13	11,46	87,68	88,31	1,69	2,56	6,18	7,3	0	9,63
9	0	9,33	0	0,12	1,65	3,5	86,55	91	0	3,95
10	2,73	12,06	87,61	87,65	0	0,8	6,35	4	3,31	4,51
11	5,79	15,12	81,14	84,21	0,42	0	0	0	0,46	0,67
12	0	9,33	91,94	90	3,49	4,8	1,65	0	0,34	4,13
13	73,71	79,65	0,62	5,35	0,42	3,2	6,15	1,15	15,22	10,65
14	0	1,0843333	0,37	1,4366667	92,23	90,2	0,29	0	0,11	0,68
15	1,81	3,115	85,92	86,885	9,45	10,1	0,18	0	0,17	2,77
16	82,73	84,84	0,05	0	0,02	0	0,15	0	9,54	15,16
17	2,44	4,55	82,6	85	0,08	0	3,87	0	6,05	10,45
18	1,01	3,12	90,72	92	0	0	3,08	0	1,42	4,88
19	71,54	73,65	7,11	7,35	1,8	0	0,58	0	0,33	19
20	96,67	93,87	0,4	0	0,37	0	1,15	0	0	6,13
21	0	0,038	85,11	87	4,58	5	0,12	0	1,02	7,962

Grazie!

Al Prof. Victor Ugo Ceccherelli che, con i suoi vasti insegnamenti e il suo modo di porgerli, ha suscitato in me un grande interesse per la materia. Grazie perché mi ha offerto il suo aiuto pratico e professionale, senza dimenticare gli aspetti umani.

A Fabione (Dott. Fabio Bozzeda) che mi ha fatto diventare matta (!) ma che mi ha seguita in modo preciso e competente, facendomi impadronire di tutti gli strumenti necessari ai miei fini.

Ad Angelo (Dott. Angelo Bonanno) che, sia durante le campagne oceanografiche, sia in fase di stesura della tesi, mi ha fornito il suo aiuto e supporto, anche nonostante le distanze.

All'intero IAMC-CNR che mi ha permesso di partecipare alle campagne oceanografiche e di lavorare sui dati raccolti in seno ai suoi progetti.

A Marta (Dott.ssa Marta D'Elia) che mi ha permesso di elaborare i miei studi sulla base dei suoi.

A Ignazio (Dott. Ignazio Fontana) che mi ha aiutata in modo molto efficiente, anche da lontano.

A tutti i ragazzi con cui ho condiviso la bellissima e professionalizzante esperienza delle campagne oceanografiche, in particolar modo ai capo missione Angelo (Dott. Angelo Bonanno) e Walter (Dott. Gualtiero Basilone), e all'intero equipaggio della N/O Dallaporta.

A tutti i miei compagni dell'avventura "Biologia Marina a Ravenna", in particolare ad Ale che ha sempre creduto in me.

A tutti i miei amici di Catania che non mi hanno mai fatto mancare il loro affetto.

A Kurry che mi ha sopportata e supportata costantemente, restandomi sempre vicino e dandomi coraggio.

Alla mia famiglia che mi ha sostenuta da lontano e da vicino.

Al mare che, se non ci fosse, nuoterei nell'aria!