Valutazione dello stato ecologico dei sistemi lotici mediante analisi dei popolamenti ittici ai sensi della Direttiva 2000/60/CE: una proposta basata sulla ricostruzione del giudizio esperto con tecniche di Intelligenza Artificiale.

Evaluating the ecological status of streams and rivers by means of fish assemblages according to 2000/60/CE Directive: a method based on the reconstruction of expert judgment using Artificial Intelligence techniques.

Michele Scardi
Stefano Cataudella
Eleonora Ciccotti
Daniele Ciuffa
Dipartimento di Biologia, Università di Roma "Tor Vergata" - Roma

Paolo Tito Colombari
ARSIAL - Stabilimento Ittiogenico - Roma

Paola Di Dato
Dipartimento di Biologia, Università di Roma "Tor Vergata" - Roma

Giancarlo Gibertini
Dipartimento di Biologia - Università "Roma Tre" - Roma

Stefano Larsen
Laura Mancini
Dipartimento di Ambiente e Connessa Prevenzione Primaria, Istituto Superiore di Sanità - Roma

Giuseppe Moccia
Provincia di Roma - Servizio 3 - Caccia e Pesca - Roma

Michele Munafò
Servizio Gestione Modulo Nazionale SINAnet - APAT - Roma

Monika Grygielewicz
Dipartimento di Biologia - Università di Roma "Tor Vergata" - Roma

Giorgio Pace
Dipartimento di Ambiente e Connessa Prevenzione Primaria, Istituto Superiore di Sanità - Roma

Tommaso Russo
Dipartimento di Biologia, Università di Roma "Tor Vergata" - Roma

Massimiliano Scalici
Dipartimento di Biologia - Università "Roma Tre" - Roma

Lorenzo Tancioni
Dipartimento di Biologia - Università di Roma "Tor Vergata" - Roma

Key words
Fish assemblages, Water Framework Directive, neural networks, ecological status, decision support systems, expert judgment.

Summary

The implementation of the Water Framework Directive (2000/60/EC) raises a number of complex scientific and technical problems, and the evaluation of the ecological status of streams and rivers based on fish fauna is probably the most challenging among these problems. In fact, no widely accepted methods are available and developing new ones is not a trivial task. Several biotic indices have been proposed for European as well as Italian streams and rivers so far, but none gained enough momentum to be used on a routine basis as a monitoring tool. Multimetric biotic indices are the most popular tools in this field, but, although certainly useful, they are far from perfect, because the assumptions upon which they are based may be either false or too simplistic. Moreover, no procedure – including biotic indices - can be entirely objective in evaluating ecological status, because this very concept is inherently subjective. We therefore tried to address this problem by means of a Decision Support System (DSS) based on Artificial Intelligence techniques, namely on an artificial neural network. In our approach the unavoidable subjective elements, namely the expert judgment about ecological status, only play a role in the earliest step of development, while the subsequent optimization of the assessment procedure is as objective as possible. A working implementation of the DSS, based on a user-friendly interface, has been developed for the Latium (Central Italy) streams and rivers. The accuracy of the classification relative to a test data set is quite high (66.7%, n=69), while errors in classification never exceed a single quality class.

Riassunto

L'implementazione della Direttiva Quadro sulle Acque 2000/60/CE pone una serie di complessi problemi tecnici e scientifici, fra i quali la valutazione dello stato ecologico dei corsi d'acqua sulla base della fauna ittica è probabilmente il più arduo. In effetti, non esistono metodi ampiamente accettati e lo sviluppo di nuovi approcci non è un compito banale. Diversi indici biotici sono stati proposti per i fiumi europei ed italiani a tutt'oggi, ma nessuno ha acquisito una base di consenso tale da poter essere utilizzato su base routinaria come strumento di monitoraggio. Gli indici più popolari in questo campo sono di tipo multimetrico, ma, per quanto utili, essi non sono certo perfetti, perché le assunzioni su cui sono basati possono essere false o troppo semplicistiche. Inoltre, nessuna procedura – compresi gli indici biotici – può essere completamente oggettiva nel valutare lo stato ecologico, perché questo stesso concetto è intrinsecamente soggettivo. Per questo motivo si è cercato di affrontare il problema mediante un Sistema di Supporto Decisionale (DSS) basato su tecniche di Intelligenza Artificiale, ed in particolare su una rete neurale artificiale. Nel nostro approccio gli ineludibili elementi soggettivi, ovvero i giudizi esperti sullo stato ecologico, giocano un ruolo esclusivamente nella fase iniziale dello sviluppo, mentre la successiva ottimizzazione della procedura di valutazione è il più possibile oggettiva. Una implementazione funzionante del DSS, basata su un'interfaccia utente user-friendly, è stata sviluppata per i corsi d'acqua del Lazio. L'accuratezza della classificazione riferita ad un insieme di dati di test è decisamente elevata (66.7%), con gli errori di classificazione che non eccedono mai una singola classe di qualità.

Introduzione

L'attuazione della Direttiva Quadro sulle Acque 2000/60/CE (European Union, 2000) richiede la messa a punto di metodi di valutazione dello *stato ecologico* delle acque superficiali basati su *elementi di qualità* biotici. Fra questi ultimi, i popolamenti ittici possono giocare un ruolo molto importante, sia per le loro peculiarità biologiche ed ecologiche (es. facilità di riconoscimento, integrazione delle risposte biotiche su scale spaziali più ampie di altri taxa, cicli vitali anche molto lunghi), sia per il loro valore iconico, che li associa alla qualità ambientale anche nella percezione di un pubblico non tecnico (Tancioni *et al.*, 2005; Scardi *et al.*, 2006; Tancioni *et al.*, 2006).

Anche se non è affatto indicato dalla Direttiva che la valutazione dello stato ecologico dei sistemi acquatici debba essere basata su indici biotici (il termine *indice* non è mai menzionato), quasi tutti i metodi sviluppati, sia in Europa che altrove, sono basati proprio su tali strumenti. La ragione di ciò non è del tutto ovvia, ma è probabilmente legata al fatto che la maggior parte degli ecologi non ha, per formazione, una reale vocazione per il trattamento statistico e matematico dei dati ed è quindi portata a premiare la semplicità di un approccio basato su indici biotici, soprattutto di tipo multimetrico, piuttosto che l'efficienza di metodi più performanti da un punto di vista computazionale, ma percepiti come troppo complessi o lontani dalla propria esperienza personale.

Negli ultimi anni sono stati proposti diversi indici biotici basati sui pesci, sia in Italia (Zerunian, 2004; Forneris *et al.*, 2004, 2005) sia in altri paesi dell'UE. Tuttavia, l'approccio basato sugli indici può non rappresentare la risposta migliore al problema posto dalla Direttiva Quadro sulle Acque. Lasciando da parte le soluzioni che hanno un valore conservazionistico e che quindi rispondono alle istanze di altre Direttive (es. Direttiva Habitat), tutti gli approcci mirati alla valutazione della funzionalità degli ecosistemi, come gli indici derivati dall'IBI (Indice di Integrità Biotica) (Karr, 1981; Karr *et al.*, 1986; Plafkin *et al.*, 1989; Oberdoff *et al.*, 1992; Maio *et al.*, 1996; Karr e Chu, 1999; La Porta *et al.*, 2001) o il recente EFI (*European Fish Index*) (FAME Consortium, 2004; Pont *et al.*, 2006), sono basati su un approccio multimetrico. Quest'ultimo, per quanto concettualmente attraente, da un punto di vista pratico presenta dei limiti, sia per le proprietà delle singole metriche che lo sottendono, che non necessariamente sono legate alla qualità ambientale da relazioni monotone o lineari (es. ricchezza specifica totale o di singoli gruppi), sia per la selezione delle metriche stesse, che può introdurre ridondanze dovute alla loro non indipendenza (es. specie litofile e specie intolleranti), sia per la mancanza di qualsiasi forma di ottimizzazione computazionale (chi definirebbe i parametri di una retta di

regressione senza usare l'algoritmo dei minimi quadrati?), sia per numerosi altri aspetti.

In sostanza, l'uso di indici biotici multimetrici, pur se nella percezione di molti addetti ai lavori è considerato concettualmente solido ed oggettivo, porta inevitabilmente alla sovrasemplificazione di relazioni complesse e contemporaneamente comporta un elevato grado di soggettività, come ad esempio nella selezione delle metriche o nella discretizzazione delle scale dei punteggi (scelta dei livelli soglia tra le classi di qualità), che non possono che essere governati dal giudizio esperto. Non è certamente un caso, da questo punto di vista, il fatto che in nessun settore applicativo in cui si debbano effettuare valutazioni o diagnosi critiche il ricorso a indici o altri metodi di valutazione automatica abbia un ruolo prevalente e tale da relegare il giudizio esperto in secondo piano. L'esempio della diagnostica medica è ben noto a tutti, ma anche in altri campi l'uso di indici è del tutto subordinato ed ancillare rispetto al giudizio esperto.

Nel caso della valutazione dello stato ecologico ai fini della Direttiva Quadro sulle Acque, inoltre, la soggettività del giudizio esperto non può essere in alcun modo aggirata, poiché il concetto stesso di stato ecologico rappresenta un'interpretazione umana della fenomenologia naturale piuttosto che una proprietà emergente della stessa. Se si accetta questo principio, allora può essere vantaggioso porre proprio l'interpretazione umana, sotto forma di giudizio esperto espresso da più fonti, al centro dell'attenzione. L'uso ottimale di questa fonte di informazione richiede certamente un approccio di nuova generazione ed in particolare l'uso di tecniche di Intelligenza Artificiale. Da questo punto di vista, l'uso di questi strumenti, ed in particolare delle reti neurali artificiali, non è più una novità in Ecologia, poiché questo tipo di applicazioni si sono andate moltiplicando nel corso dell'ultimo decennio, soprattutto per quanto riguarda le applicazioni modellistiche, l'analisi di gradiente ed il riconoscimento di patterns (Scardi, 1996; Lek e Guegan, 1999; Scardi e Harding, 1999; Scardi 2001). Anche nel campo specifico dello studio della fauna ittica degli ambienti lotici sono state realizzate numerose applicazioni, prevalentemente mirate alla stima della fauna potenziale in base a descrittori ambientali (es. Scardi *et al.*, 2004, 2005).

In questo quadro di riferimento, il nostro approccio si propone di associare i dati relativi ad un insieme di variabili ambientali ed alla composizione osservata del popolamento ittico ai giudizi esperti sullo stato ecologico dei siti di campionamento formulati da un gruppo di esperti. Ciò consente di generalizzare il razionale del giudizio esperto e di ricostruire, sulla base dell'informazione ambientale e di quella faunistica relativa all'ittiofauna, un giudizio sullo stato ecologico di qualsiasi nuovo sito. In questo modo si può ottenere il miglior giudizio di consenso basato sull'informazione disponibile, riproducendo il comportamento di un gruppo di esperti chiamati a consulto.

Materiali e metodi

Nelle applicazioni fino ad oggi disponibili (es. Scardi et al., 2004, 2005) i descrittori ambientali venivano utilizzati come elementi predittivi rispetto alla composizione del popolamento ittico, che veniva stimata dalla rete neurale. Nell'applicazione presentata in questa sede, invece, sia i dati relativi ai descrittori ambientali, sia quelli relativi al popolamento ittico vengono considerati come elementi predittivi, mentre la rete neurale ricostruisce in base ad essi una stima del giudizio esperto.

Poiché una rete neurale del tipo di quella utilizzata in questo caso, ovvero un perceptron multistrato, apprende attraverso un adeguato insieme di esempi, sia i dati relativi ai descrittori predittivi, sia quelli relativi ai descrittori da prevedere devono essere raccolti in un numero sufficiente di casi, da utilizzare poi per l'addestramento e la validazione della rete neurale.

Da un punto di vista pratico, ciò ha richiesto che al momento del campionamento, oltre a raccogliere dati biotici ed abiotici, venisse anche formulato un giudizio esperto relativo allo stato ecologico complessivo del sito. È importante sottolineare come, in quest'ottica, il giudizio debba essere riferito allo stato ecologico generale, e non a quello specifico dell'ittiofauna, che non rappresenta l'oggetto dello studio, ma piuttosto una chiave di lettura delle proprietà ecologiche complessive del biotopo analizzato. Per facilitare la formulazione del giudizio nei casi dubbi e per permettere l'uso di più valutazioni indipendenti per lo stesso sito, il giudizio esperto è stato codificato in termini di probabilità di assegnazione a ciascuna delle cinque classi di qualità, piuttosto che come un singolo valore. Ad esempio, se il giudizio esperto è intermedio fra le classi Buono e Sufficiente, vuoi per uno stato ecologico effettivamente intermedio fra le due classi, vuoi per il differente giudizio espresso da due esperti, si deve assegnare un valore del 50% a ciascuna delle due classi.

Per dimostrare le potenzialità di questo approccio è stato utilizzato un insieme di dati relativi ai fiumi del Lazio, il quale includeva 62 campionamenti, effettuati nel corso dell'estate 2005. Alle osservazioni di campo, complete di giudizio esperto, sono state affiancate altre osservazioni, che possono essere considerate come virtuali. In pratica, si è chiesto agli esperti impegnati nell'esercizio di redigere delle nuove schede di campionamento per ogni sito studiato, modificando i dati relativi ai descrittori ambientali e/o quelli relativi alla fauna ittica e variando conseguentemente il loro giudizio. Si consideri, a titolo di esempio, il caso di un sito in cui non fossero stati rinvenuti stadi giovanili di particolari specie nel campionamento reale: l'eventuale

presenza di questi ultimi avrebbe certamente indotto l'esperto a formulare un giudizio più positivo e tutto ciò può essere tradotto in un'osservazione virtuale, ma utilissima ai fini della esemplificazione del razionale che sottende la formulazione del giudizio esperto e, di conseguenza, utilissima ai fini dell'addestramento di una rete neurale capace di generalizzare efficacemente quanto appreso.

È ovvio che il ruolo di tali osservazioni virtuali è tanto più importante quanto più limitata è la base di dati di campo di cui si può disporre, ma non è subordinato a quest'ultima, poiché trasferisce alla rete neurale elementi di conoscenza essenziali per poter ricostruire le modalità di formulazione del giudizio esperto. In totale, le 62 osservazioni di campo sono state affiancate da 78 osservazioni virtuali, per un totale di 140 osservazioni.

I giudizi esperti formulati per questo insieme di osservazioni non erano distribuiti in maniera omogenea fra le 5 classi di stato ecologico, poiché quelle intermedie risultavano preponderanti rispetto a quelle estreme. Per evitare uno squilibrio nella risposta della rete neurale indotto dalla maggior frequenza delle condizioni intermedie, sono state generate altre osservazioni utilizzando come base le osservazioni reali relative alle classi meno rappresentate ed applicando una modesta perturbazione aleatoria ($\pm 10\%$) ai valori relativi alle sole variabili ambientali quantitative. L'obiettivo era quello di fare in modo che tutte le 5 classi di stato ecologico fossero rappresentate da non meno di 40 osservazioni e per ottenere ciò sono state generate 79 nuove osservazioni virtuali.

Un sottoinsieme dei dati disponibili, sia reali che virtuali, è stato utilizzato per l'addestramento della rete neurale (150 osservazioni), mentre un altro sottoinsieme (69 osservazioni) non è stato utilizzato ed è stato invece riservato per un test *ex post* del sistema. Quest'ultimo è fondamentale per scongiurare effetti indesiderati, primo fra tutti la mancanza di capacità di generalizzazione, che trasforma una rete neurale in una memoria, capace di associare con precisione i casi noti ai relativi giudizi e al tempo stesso incapace di valutare correttamente casi anche solo leggermente differenti.

Prima di dividere le osservazioni nei due sottoinsiemi appena menzionati, l'intero insieme dei dati è stato stratificato per classe di qualità assegnata in base al giudizio esperto ed in base all'altitudine, per evitare che la composizione dei due sottoinsiemi fosse sbilanciata per uno di questi due importanti aspetti.

Per ciò che riguarda i descrittori predittivi (ovvero diagnostici, nell'ottica della previsione del giudizio esperto), la Tabella 1 mostra quelli abiotici, mentre la Tabella 2 mostra quelli biotici. Per ciò che invece riguarda le specie ittiche, la lista utilizzata include 30 taxa sui 43 presenti nell'insieme delle osservazioni disponibili, poiché si è scelto di escludere le specie meno frequenti, presenti in meno del 5% delle osservazioni, non informative ai fini della formulazione del giudizio su una base strettamente numerica. È evidente che alcune delle specie escluse possono avere un significato ecologicamente rilevante, ma tale significato deve essere supportato da un'adeguata quantità di esempi (reali o virtuali) per poter essere trasferito efficacemente alla rete neurale. Le specie escluse da una menzione esplicita, tuttavia, concorrono alla ricchezza specifica, sia in termini totali, sia per ciò che riguarda i soli stadi giovanili.

La rete neurale utilizzata per questa applicazione aveva un singolo strato nascosto ed un'architettura 59-25-5, ovvero 59 nodi di input (27 descrittori ambientali, 30 taxa, 2 ricchezze specifiche), 25 nodi nello strato nascosto e 5 nodi di output (le probabilità di assegnazione a ciascuna delle 5 classi di stato ecologico). Tutti i dati sono stati normalizzati in modo da ricadere nell'intervallo [0,1] prima dell'addestramento. I nodi nascosti erano dotati di una funzione di attivazione sigmoide del tipo $f(x)=1/(1-e^{-x})$, mentre i nodi di output erano dotati di una funzione di attivazione *softmax* (Bridle, 1990). Tale funzione fa sì che la somma degli output della rete neurale sia unitaria, coerentemente con la loro interpretazione probabilistica. L'architettura ottimale è stata determinata sperimentalmente, facendo variare il numero dei nodi dello strato nascosto fra 10 e 50 e selezionando la soluzione ottimale. Va segnalato tuttavia il fatto che questa caratteristica non è critica, poiché reti con architetture diverse nello strato intermedio hanno mostrato risposte molto simili in termini di accuratezza, non solo rispetto ai dati usati per l'addestramento, ma anche rispetto al sottoinsieme indipendente di dati di test. L'addestramento della rete è stato effettuato utilizzando il più diffuso algoritmo, ovvero quello di error *back-propagation* (Rumelhart et al., 1986), con un tasso di apprendimento costante pari a 0.9 ed un momento pari a 0.1. Durante l'addestramento, per favorire una corretta generalizzazione, ad ogni valore di input è stato aggiunto una rumore gaussiano compreso nell'intervallo [-0.01,+0.01].

Inoltre, per valutare la conformità del giudizio fornito dalla rete neurale con il giudizio esperto originale si è fatto ricorso alla statistica *kappa* pesata (Cohen, 1960; Fleiss *et al.*, 1969) per l'analisi della risposta discreta, ovvero della classificazione in 5 classi di stato ecologico, ed alla correlazione lineare di Pearson (1896) per l'analisi della risposta continua, ottenuta come media ponderata delle probabilità di appartenenza a ciascuna delle 5 classi.

Infine, per facilitare l'applicazione del metodo è stata realizzata un'interfaccia utente grafica (GUI) in ambiente Microsoft Windows (figura 1). Questa GUI è stata dotata di sliders per consentire di variare in maniera interattiva il valore di tutte le variabili predittive ad esclusione di quelle tassonomiche, che invece hanno una codifica strettamente binaria. In questo modo, facendo scorrere gli sliders, è possibile rilevare in tempo reale la variazione indotta nella clas-

sificazione dello stato ecologico e comprendere meglio il ruolo dei diversi descrittori e la sensibilità del sistema alle loro variazioni.

Variabile	Unità
Altitudine	m
Profondità	m
Correntini	superficie %
Pozze	superficie %
Raschi	superficie %
Flusso indistinto alla superficie	superficie %
Presenza zone umide connesse	0=no, 1=si
Barre di meandro e puntiformi o isole	0=no, 1=si
Massi	superficie %
Sassi e ciottoli	superficie %
Ghiaia	superficie %
Sabbia	superficie %
Limo e argilla	superficie %
Velocità del flusso	punteggio 0-5
Copertura vegetale in alveo	superficie %
Ombreggiamento	superficie %
Disturbo antropico	punteggio 0-4
Sbarramento a monte	se si, distanza in km (max=100), se no 100
Sbarramento a valle (0/1)	0=no, 1 =si
Lago a monte	se si, distanza in km (max=50), se no 50
Temperatura dell'acqua	°C
Torbidità	NTU
pH	
Conducibilità specifica	$\mu S\ cm^{-1}$
Saturazione O_2	%
Bacino versante	km (come radice quadrata della superficie)
Distanza sorgente	km

Tab. 1 - Le variabili ambientali utilizzate per la valutazione dello stato ecologico.
Tab. 1 - The environmental variables used for the ecological status evaluation.

Abramis brama	*Leuciscus cephalus*
Alburnus alburnus alborella	*Leuciscus lucumonis*
Alosa fallax	*Leuciscus souffia muticellus*
Anguilla anguilla	*Liza ramada*
Barbus plebejus/tyberinus	*Mugil cephalus*
Carassius carassius	*Petromyzon marinus*
Chondrostoma genei	*Pseudorasbora parva*
Cobitis taenia bilineata	*Rutilus rubilio*
Cyprinus carpio	*Rutilus rutilus*
Dicentrarchus labrax	*Salaria fluviatilis*
Esox lucius	*Salmo trutta*
Gambusia holbrooki	*Sander lucioperca*
Gasterosteus aculeatus	*Scardinius erythrophthalmus*
Gobius nigricans	*Tinca tinca*
Lampetra fluviatilis	Ricchezza specifica totale
Lampetra planeri	Ricchezza specifica degli stadi giovanili

Tab. 2 - Specie e variabili faunistiche utilizzate per la valutazione dello stato ecologico. Le specie più rare (presenti in meno del 5% delle osservazioni) non sono considerate esplicitamente, ma concorrono a determinare i valori di ricchezza specifica, sia totale che riferita ai giovanili.

Tab. 2 - Species and faunistic variables which were used for the evaluation of ecological status. Rarest species (i.e. present in less than 5% of records) were not explicitly taken into account, while they contributed to overall and juveniles species richness.

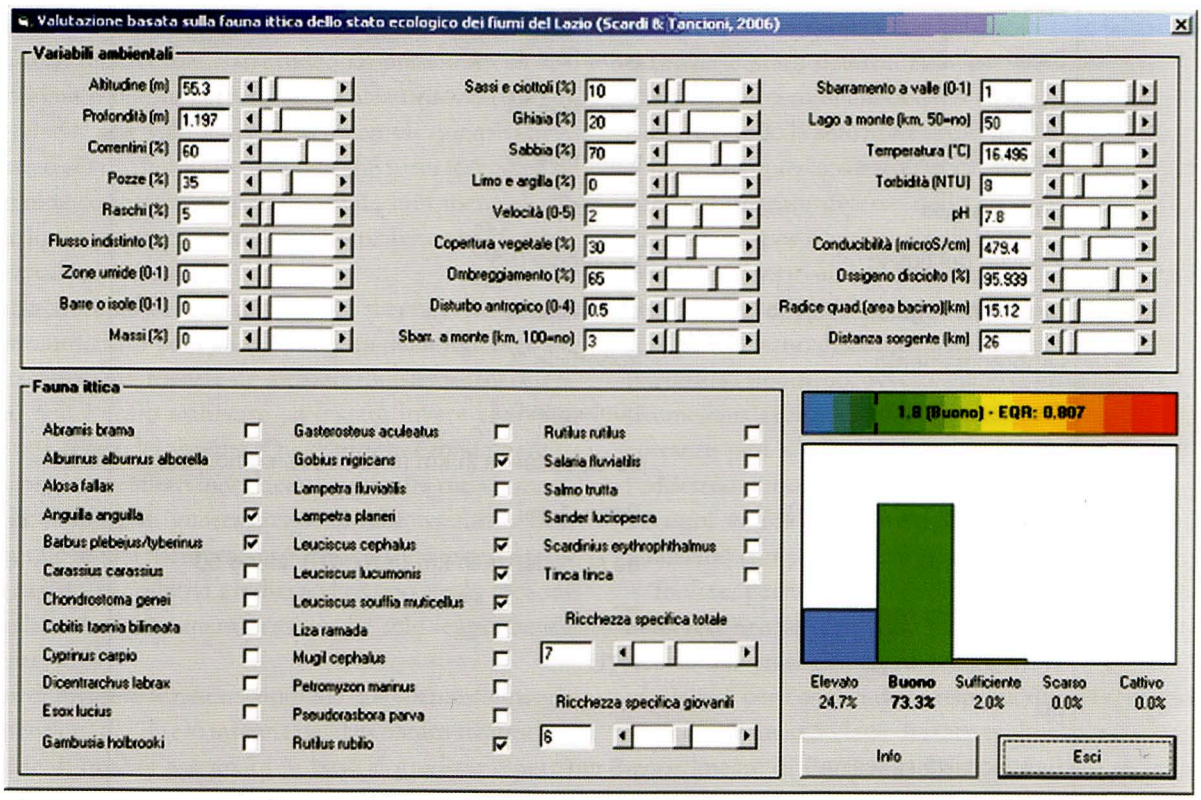

Fig. 1 - L'interfaccia utente grafica (GUI) del sistema di supporto decisionale (DSS). La classificazione viene aggiornata in tempo reale ad ogni variazione dei dati di input e quindi agendo sugli sliders è possibile esplorare la risposta del sistema alla variazione di ogni singola variabile. La classificazione è fornita in due forme: discreta, come probabilità di appartenenza a ciascuna classe di stato ecologico, e continua, sulla base di una media ponderata di tali probabilità. Quest'ultima esprime sia un punteggio da 1 a 5 (analogo alla classificazione discreta nella sua interpretazione), sia un Environmental Quality Ratio (EQR).

Fig. 1 - The Graphical User Interface (GUI) of the Decision Support System (DSS). The classification is updated in real time according to changes in input data and therefore responses to variations in each variable can be easily gauged by moving the corresponding slider. The classification is presented in two ways: discrete, as the membership probability for each ecological status class, and continuous, on the basis of a weighted average of such probabilities. The continuous classification is showed both as a score ranging from 1 to 5 (that can be interpreted exactly like the discrete classification) and as an Environmental Quality Ratio (EQR).

Risultati

Il risultato della procedura di valutazione dello stato ecologico è rappresentato da un vettore di 5 elementi, che rappresentano le probabilità di appartenenza del caso in esame a ciascuna delle 5 classi di qualità in cui viene classificato lo stato ecologico, mostrate in forma grafica mediante un istogramma. Lo stesso risultato è espresso anche in forma continua: (1) come un punteggio compreso fra 1 (stato Elevato al 100%) e 5 (stato Cattivo al 100%), che viene calcolato come una media ponderata delle probabilità di appartenenza alle 5 classi di stato ecologico; (2) come Environmental Quality Ratio (EQR), con un valore compreso fra 0, corrispondente al massimo grado di alterazione, ed 1, che invece indica uno stato conforme alle condizioni di riferimento. E' ovvio, data l'impostazione del metodo, che queste ultime sono quelle basate implicitamente sull'insieme dei giudizi esperti su cui il metodo di valutazione è stato calibrato. È importante, a questo proposito, ricordare che la WFD prevede la definizione delle condizioni di riferimento sia su base modellistica sia mediante il giudizio esperto e che l'approccio da noi proposto realizza una sintesi dei due criteri.

Nel caso della classificazione discreta, in realtà esiste una doppia possibilità di attribuzione, perché è possibile sia assegnare la classe che corrisponde al livello di probabilità più elevato (strategia *winner takes all*), sia arrotondare il punteggio continuo all'intero più vicino. Quest'ultima soluzione, però, sfavorisce le classi estreme a vantaggio di quelle intermedie ed è per questo motivo che in questa sede è stata utilizzata la prima.

La capacità della rete neurale di incorporare in maniera efficace l'informazione fornita nella fase di addestramento può essere desunta dalle tavole di confusione che mettono a confronto le classi di stato ecologico assegnate in base al giudizio esperto con quelle fornite dalla rete neurale stessa. Nella tabella 3 è mostrata la tavola di confusione relativa all'insieme di osservazioni utilizzate per l'addestramento, mentre nella tabella 4 è riportata quella relativa alle osservazioni utilizzate esclusivamente come test. Mentre la prima tavola di confusione rende conto di quanto la rete neurale abbia appreso durante l'addestramento, la seconda fornisce una stima della capacità di generalizzazione dell'apprendimento.

Nel primo caso le istanze correttamente classificate (CCI) sono il 73.5% del totale, mentre nel secondo sono il 66.7%, ovvero 2 su 3. Questa percentuale è certamente soddisfa-

		Rete neurale					
		Elevato	Buono	Sufficiente	Scarso	Cattivo	
Giudizio esperto	Elevato	**19**	21				*40*
	Buono	2	**40**	4			*46*
	Sufficiente		14	**35**	2		*51*
	Scarso			9	**31**		*40*
	Cattivo				6	**36**	*42*
		21	*75*	*48*	*39*	*36*	***219***

Tab. 3 - Matrice di confusione relativa all'insieme dei dati di addestramento. Sono correttamente classificati il 73.5% dei casi, con un margine di errore che, nei rimanenti casi, non eccede una singola classe.

Tab. 3 - Confusion matrix for the training data set. Correctly classified cases are 73.5% and errors never exceed a single quality class.

		Rete neurale					
		Elevato	Buono	Sufficiente	Scarso	Cattivo	
Giudizio esperto	Elevato	**5**	7				*12*
	Buono	2	**15**	1			*18*
	Sufficiente		5	**6**	2		*13*
	Scarso			2	**11**		*13*
	Cattivo				4	**9**	*13*
		7	*27*	*9*	*17*	*9*	***69***

Tab. 4 - Matrice di confusione relativa all'insieme dei dati di test. I casi correttamente classificati sono il 66.7% e l'errore non supera una classe di qualità. La statistica kappa pesata è risultata pari a 0.775 ed altamente significativa (p>99.9%), a testimonianza dell'accuratezza della riproduzione del giudizio esperto.

Tab. 4 - Confusion matrix for the test data set. Correctly classified cases are 66.7% and errors never exceed a single quality class. Weighted kappa statistics is highly significant (k=0.775, p>99.9%), thus showing that expert judgment was correctly reproduced.

cente, data soprattutto la complessità del problema, ma è da considerarsi in una luce ancor più favorevole per il fatto che in nessun caso, né nell'insieme di osservazioni di addestramento, né in quelle di test, sono stati osservati scarti fra la classificazione basata sul giudizio esperto e quella fornita dalla rete neurale che fossero più grandi di una singola classe di stato ecologico.

Le tavole di confusione e le CCI forniscono un quadro utile da un punto di vista descrittivo, ma la valutazione del livello di accordo fra giudizio esperto reale e giudizio esperto stimato dalla rete neurale richiede una verifica più formale. A questo fine è stato applicato il test *kappa* (Cohen, 1960; Fleiss *et al.*, 1969) per confronti fra due criteri di classificazione basati su categorie ordinali multiple. In particolare, è stata utilizzata la forma pesata della statistica *kappa*, che attribuisce un peso maggiore agli accordi relativi alle medesime categorie, seguiti da quelli relativi a categorie adiacenti nella tavola di confusione e così via. Il valore della statistica *kappa* per la tavola di confusione basata sull'insieme delle osservazioni di test è risultato pari a 0.775, che, secondo una scala empirica dei valori di *kappa* può essere considerato come indicativo di un accordo "sostanziale". Ciò che più conta, tuttavia, è che è stato possibile rigettare l'ipotesi nulla di identità fra la statistica *kappa* osservata e quella corrispondente ad una classificazione causale delle osservazioni, ovvero l'ipotesi di una statistica *kappa* nulla ad un livello di probabilità superiore al 99.9%. Se si aggregano i dati relativi alle prime due classi, considerate come accettabili ai fini della Direttiva Quadro sulle Acque, e quelli relativi alle successive tre, insufficienti ai fini della stessa Direttiva, la statistica *kappa* sale fino a 0.826, per un livello di probabilità anche in questo caso superiore al 99.9%.

L'analisi delle tavole di confusione implica l'assegnazione di una specifica classe di stato ecologico a ciascuna osservazione, ma in realtà la formulazione del giudizio esperto originale e la sua ricostruzione mediante la rete neurale è più articolata e si presta anche ad essere espressa come media ponderata delle probabilità di appartenenza a ciascuna delle classi di stato ecologico. Se si confrontano i valori di tale media ponderata (che si esprime come un punteggio compreso fra 1 e 5) per ciò che riguarda il giudizio esperto originale e quello ricostruito dalla rete neurale si può notare come l'accordo sia molto buono, sia per le osservazioni utilizzate per l'addestramento, sia per quelle utilizzate per testare *ex post* la rete neurale (Figura 2). In particolare, il valore del coefficiente di correlazione di Pearson è pari a 0.978 nel primo caso ed a 0.932 nel secondo, risultando in entrambi i casi altamente significativo (p>99.9%).

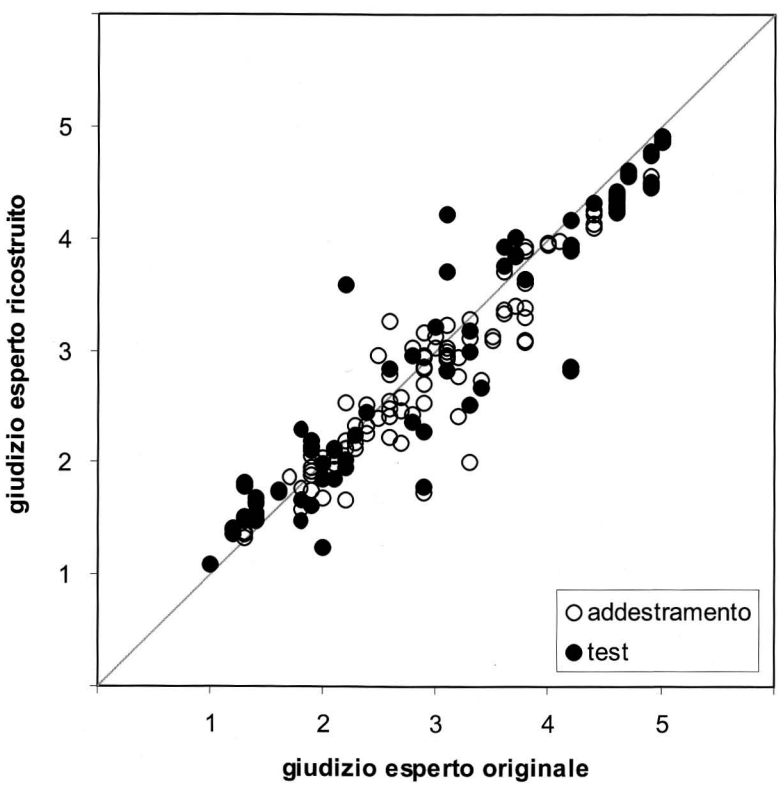

Fig. 2 - Relazione fra giudizi esperti originali e ricostruiti in forma continua, come media ponderata della probabilità di appartenenza a ciascuna classe di stato ecologico. Sono mostrate sia le osservazioni di addestramento (cerchi bianchi, r=0.978**) che quelle di test (cerchi neri, r=0.932**).

*Fig. 2 - Relationship between original and reproduced export judgments, expressed in continuous form, as the weighted average of the membership probability relative to each ecological status class. Both the training (white circles, r=0.978**) and the test cases (black circles, r=0.932**) are shown.*

Discussione e conclusioni

In passato era già stato dimostrato (Scardi *et al.*, 2004, 2005) come fosse possibile utilizzare una rete neurale per stimare la composizione della fauna ittica di un biotopo di cui fossero noti i valori relativi ad un insieme di grandezze fisiche, chimiche e morfologiche. Questo approccio aveva fornito previsioni corrette in una frazione molto elevata dei casi analizzati (oltre il 91%) ed aveva un'ovvia rilevanza non solo ai fini dello studio della distribuzione spaziale della fauna ittica, ma anche in rapporto alla Direttiva Quadro sulle Acque. Infatti, se si confronta la fauna attesa con quella osservata e se, come esplicitamente previsto dalla Direttiva stessa, si considera la prima come espressione di una condizione di riferimento, allora si può facilmente ottenere un valore per il cosiddetto Environmental Quality Ratio (EQR) da utilizzare come base per la valutazione dello stato ecologico. Ad esempio, ciò può essere ottenuto calcolando il valore dell'indice di similarità di Jaccard, che altro non è se non la proporzione di specie comuni nelle due liste. In pratica, questo approccio richiede (1) la scelta di un coefficiente di similarità o distanza fra fauna osservata e fauna attesa fra le decine di coefficienti disponibili, (2) la sua eventuale normalizzazione nell'intervallo [0,1] perché possa esprimere un EQR e (3) la discretizzazione della scala dell'EQR in intervalli corrispondenti alle 5 classi di stato ecologico (Elevato, Buono, Sufficiente, Scarso, Cattivo). E' evidente, tuttavia, che tutte queste scelte sono intrinsecamente soggettive, ma, al tempo stesso, determinanti al fine della formulazione del giudizio finale.

Utilizzando la stessa informazione di base necessaria per questo tipo di approccio, però, è stato possibile associare direttamente i dati rilevati sul campo al giudizio esperto, ponendo l'unico elemento necessariamente soggettivo della procedura di valutazione alla base della strategia di valutazione piuttosto che al termine della stessa. In questo modo, e facendo leva su una strategia di valutazione plurima, basata sui giudizi formulati da più esperti in maniera indipendente per il medesimo sito è possibile ottenere una stima del giudizio esperto "di consenso" sulla base della fauna ittica osservata e dei dati relativi alle variabili ambientali. In pratica, queste ultime forniscono il contesto ambientale entro cui viene valutata l'informazione contenuta nella composizione della fauna ittica, mentre il fatto di disporre di più di un giudizio rende conto dell'inevitabile margine di incertezza insito nella formulazione di un giudizio di qualità.

Il risultato ottenuto in questa prima applicazione ai fiumi del Lazio ha un valore dimostrativo e si presta certamente ad essere migliorato. Già in questa forma, tuttavia, è sufficientemente accurato, con un margine di errore che non supera in nessun caso una singola classe di stato ecologico. In particolare, se ci si concentra sulla separazione, critica ai fini della Direttiva Quadro sulle Acque, fra gli stati Elevato e Buono da una parte e quelli Sufficiente, Scarso e Cattivo dall'altra, si può notare (vedi tabelle 3 e 4) come i casi non correttamente classificati siano soltanto 18 su 219 (8.2%) per le osservazioni di addestramento e 6 su 69 (8.7%) per quelle di test. In entrambi i casi, l'incertezza riguarda soltanto osservazioni attribuite ad uno stato Sufficiente in base al giudizio esperto e valutate come in uno stato Buono dalla rete neurale e viceversa. È ovvio che una base di dati più ampia potrà migliorare ulteriormente l'accuratezza della ricostruzione del giudizio esperto operata dalla rete neurale.

Per ciò che riguarda le variabili utilizzate, alcune di esse, soprattutto fra quelle abiotiche, potrebbero essere escluse da applicazioni future, poiché giocano un ruolo marginale o strettamente contingente rispetto all'insieme dei dati utilizzati (es. la temperatura dell'acqua ha un valore diagnostico nel caso di un campionamento esclusivamente estivo in un'area geografica limitata, ma non è informativo altrimenti). Analogamente, alcuni descrittori potrebbero essere esclusi perché di difficile determinazione, come ad esempio la torbidità dell'acqua. Va anche precisato che un certo livello di approssimazione nella stima delle variabili ambientali è del tutto accettabile (es. nella quantificazione delle granulometrie) e che la rete neurale correttamente addestrata non risponde a piccole variazioni delle stesse.

Viceversa, per rispondere in maniera più efficace a quanto indicato nella stessa Direttiva sulle Acque, potrebbero essere utilizzati altri descrittori della fauna ittica di tipo quantitativo e demografico. In particolare, si prevede di implementare il sistema attraverso l'integrazioni dei descrittori qualitativi di presenza/assenza delle specie ittiche rilevate e la ricchezza specifica dei loro stadi giovanili (0+) con quelli quantitativi di densità espresse, in termini di abbondanza e biomassa.

In conclusione, va ribadita l'importanza non soltanto di un'adeguata base di dati di campo, il cui ruolo è tanto ovvio quanto fondamentale, ma anche di uno sfruttamento ottimale dell'informazione che può essere fornita dal giudizio esperto. In ogni caso in cui sia possibile, ad esempio, il giudizio esperto deve essere espresso in maniera indipendente da più soggetti diversi, sempre in funzione dello stato ecologico complessivo del sito esaminato e non soltanto in base alla composizione della fauna ittica. Sempre sfruttando tutti i soggetti disponibili, inoltre, ulteriore informazione può essere ottenuta mediante la simulazione di osservazioni virtuali, basate su scenari del tutto ipotetici o su una o più modifiche delle condizioni biotiche o abiotiche riscontrate in un sito e del relativo giudizio. Ciò consente, infatti, di trasmettere alla rete neurale la capacità di valutare correttamente anche sfumature relativamente marginali del quadro ecologico. Per questo stesso motivo, è opportuno che, anche in una fase di applicazione routinaria

del metodo, il giudizio esperto continui ad essere formulato da chi opera sul campo, in modo da consentire di riaddestrare periodicamente la rete neurale, migliorandone ulteriormente l'accuratezza.

In ogni caso, quanto proposto si configura come un sistema di supporto decisionale e non come un metodo di valutazione assoluto. Infatti, coerentemente con la convinzione della centralità del giudizio esperto che ha informato il nostro lavoro, riteniamo che nessun metodo di valutazione, sia esso basato su un indice biotico o su una rete neurale, debba sostituire il giudizio esperto. Quest'ultimo, infatti, può essere utilmente affiancato dal giudizio ricostruito dalla rete neurale, che tuttavia deve essere considerato come un riferimento di consenso, utile per meglio focalizzare la valutazione esperta e per favorirne l'omogeneità.

Dal punto di vista applicativo, in considerazione della attuale struttura della rete di monitoraggio nazionale, basata prevalentemente sul "sistema agenziale" (APAT-ARPA-APPA), è plausibile ipotizzare un periodo di verifica/validazione/formazione che veda coinvolti i tecnici deputati al monitoraggio ambientale, coadiuvati da esperti settoriali.

In conclusione, posto che nessun metodo di valutazione dello stato ecologico può essere al tempo stesso semplice, generale ed accurato, è ovvio che gli indici biotici, non potendo prescindere dalla semplicità, possono essere accurati, ma non generali (es. molte versioni bacino-specifiche dell'IBI), o generali, ma non accurati (es. EFI). Una valutazione accurata e generale, quindi, deve rinunciare ad essere semplice, almeno dal punto di vista della sua meccanica interna. L'approccio proposto in questa sede realizza esattamente questa strategia, sia pur rendendola trasparente all'utente finale grazie ad una implementazione software user-friendly, ed al tempo stesso restituisce al giudizio esperto un ruolo centrale ed imprescindibile.

Bibliografia

Bridle J.S., 1990. Probabilistic interpretation of feedforward classification network outputs, with relationships to statistical pattern recognition. In: F. Fogleman Soulie and J.Herault (eds.), Neurocomputing: Algorithms, Architectures and Applications, Berlin: Springer-Verlag, pp. 227-236.

Cohen J., 1960. A coefficient of agreement for nominal scales., Educ. Psychol. Meas., 20: 27-46.

Fleiss J.L., Cohen J., Everitt B.S., 1969. Large sample standard errors of kappa and weighted kappa. Psychol. Bull., 72: 323-327.

European Union, 2000. Parliament and Council Directive 2000/60/EC of 23rd October 2000. Establishing a Framework for Community Action in the Field of Water Policy. Official Journal PE-CONS 3639/1/00 REV 1, 2000. Brussels: European Union.

Fame Consortium, 2004. Manual for the application of the European Fish Index - EFI. A fish-based method to assess the ecological status of European rivers in support of the Water Framework Directive. Version 1.1, January 2005, 92 pp. (http://fame.boku.ac.at).

Forneris G., Merati F., Pascale M., Perosino G.C., 2004. Proposta di un indice ittico (II) per il bacino occidentale del Po.

Forneris G., Merati F., Pascale M., Perosino G.C., 2005. Proposta di un indice ittico (II) per il bacino occidentale del Po e prime applicazioni in Piemonte. Riv. Piem. St. Nat., XXVI: 3-39.

Karr J.R., 1981. Assessment of biotic integrity using fish communities. Fisheries, (6): 21-27.

Karr J.R., Fausch K.D., Angermeier P.L., Yant P.R., Schlosser I.J., 1986. Assessing biological integrity in running waters: a method and its rationale. Illinois Natural History Survey Special Publication, 1986; 5. Champaign, IL.

Karr J.R., Chu E.W, 1999. Restoring Life in Running Waters - Better Biological Monitoring. Island Press Washington, D.C. Covelo California, 206 pp.

La Porta G., Lorenzoni M., Carosi A., Mearelli M., 2001. Definizione di un indice di integrità biologica per il bacino del Fiume Tevere. Atti del XI Congresso Nazionale della SITE – Atti 25 (a cura di M. Falcucci e V. Hull). Parco Nazionale del Circeo, Sabaudia 12-14 settembre 2001.

Lek S., Guégan J.-F., 1999. Artificial neural networks as a tool in ecological modeling, an introduction. Ecol. Modelling, 120, 65-73.

Maio G., Rigatti Luchini S., Castaman D., Mojetta A., Salviati S., Marconato E, 1996. Prima applicazione ed adeguamento dell'Index of Biotic Integrity (IBI) in Provincia di Vicenza. Atti del VI Conv. Naz. A.I.I.A.D., Varese Ligure (SP), 6-8 Giugno 1996.

Oberdoff T., Hughes R. M., 1992. Modification of an index of biotic integrity based on fish assemblages to characterize rivers of the Seine Basin, France. Hydrobiologia, 228: 117-130.

PEARSON K., 1986. Mathematical contributions to the theory of evolution. III. Regression, heredity and panmixia. Philos. Trans. Royal Soc. London Ser. A , 187: 253–318.

PLAFKIN J.L., M.T. BARBOUR, K.D. PORTER, S.K. GROSS, R.M. HUGHES, 1989. Rapid Bioassessment Protocols for Use in Streams and Rivers. Benthic Macroinvertebrates and Fish. U.S. Environmental Protection Agency, EPA 440/4-891001.

PONT D., HUGUENY B., BEIER U., GOFFAUX D., MELCHER A., NOBLE R., ROGERS C., ROSET N., SCHMUTZ S., 2006. Assessing the biotic integrity of rivers at the continental scale: a European approach using fish assemblages. Journ. Applied Ecol., 43: 70-80.

SCARDI M., 1996. Artificial neural networks as empirical models of phytoplankton production. Mar. Ecol. Prog. Ser., 139: 289-299.

SCARDI M., HARDING L.W., 1999. Developing an empirical model of phytoplankton primary production: a neural network case study. Ecol. Model., 120 (2-3): 213-223.

SCARDI M., 2001. Advances in neural network modeling of phytoplankton primary production. Ecol. Model., 146 (1-3): 33-45.

SCARDI M., CATAUDELLA S., CICCOTTI E., DI DATO P., MAIO G., MARCONATO E., SALVIATI S., TANCIONI L., TURIN P., ZANETTI M., 2004. Previsione della composizione della fauna ittica mediante reti neurali artificiali. Biologia Ambientale, 18, 1-8.

SCARDI M., CATAUDELLA S., CICCOTTI E., DI DATO P., MAIO G., MARCONATO E., SALVIATI S., TANCIONI L., TURIN P., ZANETTI M., 2005. Optimisation of artificial neural networks for predicting fish assemblages in rivers. In: Modelling Community Structure in Freshwater Ecosystems, Lek S., Scardi M., Verdonschot P.F., Descy J.P. and Park Y.S. (Eds). Springer-Verlag: Berlin, Germany, pp. 114-129.

SCARDI M., TANCIONI L., CATAUDELLA S., 2006. Monitoring methods based on fish. In: G. Ziglio, M. Siligardi. G. Flaim (eds.), Biological Monitoring of Rivers: Applications and Perspectives. Wiley, London: 135-153.

TANCIONI L., SCARDI M., CATAUDELLA S., 2005. I pesci nella valutazione dello stato ecologico dei sistemi acquatici. Ann. Ist. Super. Sanità 2005; 41(3): 399-402.

TANCIONI L., M. SCARDI, S. CATAUDELLA, 2006. Riverine fish assemblages in temperate rivers. In: G. Ziglio, M. Siligardi., G. Flaim (eds.), Biological Monitoring of Rivers: Applications and Perspectives. Wiley, London: 47-69.

ZERUNIAN S., 2004. Proposta di un indice dello stato ecologico delle comunità ittiche viventi nelle acque interne italiane. Biol. Amb. 18(2): 25-30.