



POLITECNICO DI TORINO

Corso di Laurea Magistrale in Ingegneria per l'Ambiente
e il Territorio

Raccolta, immagazzinamento e analisi di dati biologici e idromorfologici per la modellazione dell'habitat fluviale

Relatore:

Prof. Paolo Vezza

Candidato:

Simone Forte

Co-relatore:

Ing. Giovanni Negro

Dipartimento di Ingegneria dell'Ambiente, del Territorio e
delle Infrastrutture

A.A. 2021/2022 - Sessione di Laurea luglio 2022

Sommario

I fiumi sono spesso soggetti a forti pressioni idro-morfologiche, come l'alterazione del regime di flusso naturale che provoca una diminuzione della disponibilità di habitat acquatico per la comunità biotica esistente.

Per valutare, quantificare e gestire correttamente i cambiamenti nella disponibilità di habitat, sia in Italia che all'estero, sono ampiamente utilizzati gli strumenti di modellazione dell'habitat fluviale. Uno degli approcci più efficaci è rappresentato dalla metodologia MesoHABSIM, uno strumento di modellazione a meso-scala basato sulla valutazione e sull'analisi dei mesohabitat acquatici, che corrisponde per dimensione e posizione alle unità morfologiche (es. pools, riffles, rapids).

Per stimare la quantità di habitat disponibile nei fiumi, vengono utilizzati modelli di distribuzione di specie che distinguono i meso-habitat in non idonei, idonei e ottimali. Questi criteri di idoneità possono essere sviluppati per diversi target ecologici (come singole specie, stadi vitali o comunità di pesci o macroinvertebrati) utilizzando la tecnica di machine learning Random Forest (RF).

L'applicazione della tecnica RF, costituita da una moltitudine di alberi decisionali, permette di definire le caratteristiche dell'habitat che maggiormente influenzano la distribuzione dei target ecologici. Vengono generati due modelli statistici binari di assenza/presenza e presenza/abbondanza per valutare rispettivamente habitat non idonei/idonei e habitat idonei/ottimali. La tecnica RF permette inoltre di rappresentare la relazione tra le caratteristiche di ciascun habitat e la probabilità di presenza o abbondanza attraverso l'utilizzo dei Partial Dependence Plots (PDP). Questi grafici forniscono quindi un'importante interpretazione ecologica dell'influenza dei singoli parametri del modello.

Nel presente lavoro di tesi i dataset, biologici e idro-morfologici raccolti in Italia per l'applicazione della metodologia MesoHABSIM, sono stati formattati e accorpati in un unico database nazionale. A partire dal 2009 sono stati campionati per i pesci più di 1300 mesohabitat. Circa 19000 esemplari, appartenenti a 40 specie

ittiche d'acqua dolce italiane, sono stati inseriti nel database e le loro presenze e abbondanze sono state georeferenziate alla meso-scala e legate alle caratteristiche fisiche dell'habitat.

Inoltre, per la raccolta dei dati biologici in campo, è stato sviluppato un nuovo strumento informatico costituito da un'app mobile, progettata per l'utilizzo su telefoni cellulari, che consente un'acquisizione rapida e sistematica dei dati di distribuzione del pesce che vengono periodicamente salvati nel cloud. L'app aumenta la velocità di acquisizione dei dati rispetto alle tecniche comunemente utilizzate, riducendo al minimo la loro possibilità di perdita. La funzionalità e l'efficacia dello strumento sono state testate e verificate durante due settimane di rilievi, effettuati tra marzo e maggio 2022, in 8 corsi d'acqua ubicati in provincia di Savona e Imperia (regione Liguria).

Utilizzando il nuovo database, facilmente alimentato dalla nuova app mobile, è attualmente possibile estrarre informazioni su singole specie e stadi vitali in formato standardizzato, utile per la generazione di modelli di distribuzione di specie. Come esempio di applicazione, è stato sviluppato un modello RF per lampreda padana (*Lampetra zanandreae*), specie minacciata autoctona del bacino del fiume Po. L'analisi delle prestazioni del modello e l'interpretazione dei PDP hanno portato ad ottenere risultati ecologicamente consistenti, evidenziando che la presenza di lamprede all'interno dei mesohabitat è altamente correlata alla presenza di substrati fini (ghiaia e sabbia), bassa profondità d'acqua (tra 15 e 30 cm) e bassa velocità del flusso (inferiore a 15 cm/s).

Abstract

Rivers are often subject to severe hydro-morphological pressure, such as the alteration of the natural flow regime that causes a decrease in the availability of aquatic habitat for the existing biotic community.

In order to correctly assess, quantify and manage changes in habitat availability, habitat modeling tools are widely used, both in Italy and abroad. One of the most effective approach is represented by the MesoHABSIM methodology, a meso-scale habitat modeling tool based on the assessment and analysis of aquatic mesohabitats, which corresponds in size and location to geomorphic units (e.g., pools, riffles, rapids).

To estimate the amount of available habitat in the rivers, species distribution models are used to distinguish between not suitable, suitable and optimal habitats. Such habitat suitability criteria can be developed for different ecological targets (such as single species, life stages or communities of fish or macroinvertebrates) using a machine learning technique called Random Forest (RF).

The application of the RF technique, consisting of a multitude of decision trees, allowing to define the habitat characteristics that most influence the distribution of the ecological targets. Two binary statistical models of presence / absence and presence / abundance are generated to respectively assess not suitable / suitable habitats and suitable / optimal habitats. RF technique also allows to represent the relationship between each habitat characteristic and the probability of presence or abundance through the use of the so called Partial Dependence Plots (PDP). These plots that graph the probability of presence (or abundance), provide an important ecological interpretation of the influence of single model parameters.

In this thesis, the biological and hydro-morphological datasets collected in Italy for the application of the MesoHABSIM methodology have been formatted and merged in a single nation-wide database. Starting from 2009, more than 1300 mesohabitats (or geomorphic units) have been quantitatively sampled for fish. About

19000 fish, belonging to 40 Italian freshwater fish species, have been included in the database and their presence and abundance have been georeferenced at the mesohabitat scale and linked with physical habitat characteristics.

In addition, for the collection of biological data in the field, a new IT tool has been developed. This tool consists of a mobile app, designed for use on mobile phones, allowing a quick and systematic acquisition of fish distribution data that is periodically saved to the cloud. The app increases the rate of data acquisition compared to commonly used techniques, minimizing, at the same time, the possibility of data loss. The functionality and effectiveness of the IT tool has been tested and verified during two weeks of surveys, carried out between March and May 2022, in 9 river reaches located in the province of Savona and Imperia (Liguria region).

Using the new nation-wide database, easily feeded by the new mobile app, it is currently possible to extract information of single species or life stages in a standardized format, useful for the generation of species distribution models. As an example of application, a RF model has been developed for the endangered freshwater lamprey (*Lampetra zanandreae*), an autochthonous Petromyzontiformes of Po river basin. The analysis of the model performances and the interpretation of the PDPs led to obtain ecologically consistent results, highlighting that lampreys occurrence within mesohabitats is highly correlated to the presence of fine sediments (gravel and sand), shallow water depths (between 15 and 30 cm) and low flow velocity (less than 15 cm/s).

Elenco delle figure

1.1	Principali fasi MesoHABSIM e relativi input e output, da [16].	4
1.2	Layers progetto MapStream.	7
1.3	Stazione selezionata e Layer VERTEX.	7
1.4	Layer HMU.	8
1.5	Raccolta descrittori ambientali.	8
1.6	Layer MEAS.	8
1.7	Esempio di anguilla (a sinistra) e trota mediterranea (a destra) campionate all'interno di 2 riffle, Bevera, $0.15 \text{ m}^3/\text{s}$, 11/03/2022. . .	10
1.8	Mappa di idoneità di habitat per l'Anguilla allo stadio vitale giovanile, Bevera, 11/03/2022, $0.15 \text{ m}^3/\text{s}$	10
1.9	Schema generale Random Forest.	12
1.10	Esempio di curva habitat-portata, per la trota fario adulta e giovane, nel torrente Olen ([16]).	13
1.11	Generazione serie storiche di habitat.	14
1.12	Costruzione della curva UCUT da serie temporale di habitat disponibile, [16].	15
1.13	Schermata SimStream-Web per il caricamento delle generalità del progetto.	17
1.14	Schermata SimStream-Web per il caricamento dei dati dei rilievi. . .	18
1.15	Messaggio di warning durante il caricamento dei rilievi in SimStream-Web.	19
1.16	Geometria shapefile HMU non accettata da SimStream-Web, presenza di buchi (in alto) e auto intersezioni (in basso).	20
1.17	Schermata SimStream-Web di inserimento serie temporale di portata. .	21
1.18	Schermata SimStream-Web di selezione specie, stadio vitale e modello biologico.	21
1.19	Dettagli modello biologico scazzone adulto (<i>Cottus gobio</i>).	22
1.20	Schermata SimStream-Web di selezione output.	22
1.21	Output Simstream-Web. Shapefile unità idro-morfologiche, Malacqua, $0.652 \text{ m}^3/\text{s}$	24

1.22	Output Simstream-Web. Distribuzioni di frequenza, Malaqua, 0.652 m^3/s	24
1.23	Output SimStream-Web. Shapefile mappa idoneità di habitat per vairone adulto (<i>Telestes muticellus</i>), Malacqua, 0.652 m^3/s	25
1.24	Output SimStream-Web. Shapefile mappa risultato modello biologico di Assenza/Presenza per vairone adulto (<i>Telestes muticellus</i>), Malacqua, 0.652 m^3/s , 23/01/2010.	25
1.25	Schermata SimStream-Web per download output.	26
2.1	Elaborazioni su tabella attributi dello shapefile HMU.	29
2.2	Elaborazioni su poligono dello shapefile HMU.	30
2.3	Elaborazioni su file di testo misure puntuali.	31
2.4	Dato biologico registrato su supporto cartaceo.	32
2.5	Dato biologico registrato su supporto cartaceo e trascritto su file Excel.	33
2.6	Check box per selezione output ("Hydro Bio Tool")	38
2.7	Associazione cromatica HMU CODE dato biologico e idro-morfologico.	40
2.8	Codice identificativo dell'unità "HMU_CODE".	42
2.9	Codice identificativo del sito di campionamento "CODE".	42
2.10	Sheets cartella di lavoro "MesoHABSIM DataBase".	43
2.11	Strumenti per la gestione del Database MesoHABSIM.	45
2.12	Siti analizzati per la generazione del modello statistico di distribuzione della Lampetra zanandreai allo stadio vitale giovanile. I fiumi analizzati sono i seguenti: 1 - Belbo, 2/3 - Ghiandone, 4 - Orco, 5 - Pellice.	47
2.13	Valori classe assenza/presenza pre oversampling.	48
3.1	Schermata app per inserimento generalità sito di campionamento.	51
3.2	Visualizzazione stato rete e codice per la sua gestione.	51
3.3	Dettaglio comandi di inserimento caratteristiche generali del sito di campionamento.	52
3.4	Dettaglio codice per operazioni avviate da barra di avanzamento.	53
3.5	Schermata di registrazione dato biologico.	54
3.6	Dettaglio comandi di registrazione dato biologico.	55
3.7	Dettaglio opzioni e procedura di registrazione dato biologico.	56
3.8	Dettaglio pulsante ed evoluzione post inserimento.	57
3.9	Esempio raccolta _BIO_SAMPLING.	58
3.10	Cartella di salvataggio e file di backup .csv (Q in m^3/s).	58
3.11	Flusso Power Automate per il salvataggio dei dati.	59
3.12	Sequenza comandi per modifica esemplare.	60
3.13	Pulsanti per eliminazione esemplari inseriti.	60

3.14	Flusso di operazioni per salvataggio file .csv.	62
3.15	Schermata di inserimento generalità del rilievo e bounding box HMU.	63
3.16	Validazione caratteristiche generali del rilievo inserite.	64
3.17	Schermata di inserimento del dato grezzo biologico (a) ed idro- morfologico, (b).	65
3.18	Schermata di selezione dato di input grezzo idro-morfologico.	66
3.19	Schermata di selezione dato di input grezzo biologico.	67
3.20	Messaggio di errore, attenzione e informazione per l’inserimento dei dati biologici.	69
3.21	Errata nomenclatura intestazioni e inversione ordine colonne.	70
3.22	Convertitori campi dato di input grezzo biologico.	72
3.23	Esempio di dato biologico con valori non conformi allo standard prescelto e loro modifica via convertitore.	73
3.24	Schermata di selezione output.	74
3.25	Confronto CODE dati biologici e idro-morfologici.	75
3.26	Interfaccia tool "Delete Relief".	77
3.27	Grafico siti rilievi ottenuto tramite tool "Relief Map".	78
3.28	Strumento "Extract Biological Data" per l'estrazione di dati biologici.	79
3.29	Selezione specie (sinistra) e stadio vitale (destra) da "Extract Bio- logical Data".	80
4.1	Mappa dei siti campionati in Liguria nei mesi di marzo e maggio 2022.	83
4.2	Campionamento HMU con elettrostorditore spallabile.	84
4.3	Dettaglio retino elettrostorditore spallabile.	84
4.4	Utilizzo di retino da pesca come ausilio durante le operazioni di campionamento biologico.	85
4.5	Raccolta esemplari in apposito secchio.	85
4.6	Raccolta esemplari in apposita nassa.	85
4.7	Misurazione esemplare di anguilla con ittiometro.	86
4.8	Misurazione esemplare di anguilla con bindella metrica.	86
4.9	Test app sviluppata e registrazione pesce.	87
4.10	Esemplare di anguilla appena liberato nella medesima unità di ri- trovamento (foto subacquea).	88
4.11	Operazione di calibrazione telemetro laser.	89
4.12	Distinzione meso-habitat all'interno di un tratto fluviale.	90
4.13	Utilizzo MapStream da computer da campo e mappatura meso- habitat tramite telemetro laser.	91
4.14	Telemetro laser montato su treppiede.	91
4.15	Antenna e asta per la misura del punto stazione.	92
4.16	Distribuzione punti e misurazione di profondità, velocità e substrato.	93

4.17	Shapefile rilievo in campo, necessaria traslazione verso i punti Antenna.	96
4.18	Shapefile HMU georeferenziato rispetto ai punti Antenna.	96
5.1	Schermata di registrazione dato biologico (prima versione app).	98
5.2	Elenco di selezione specie (prima versione app).	98
5.3	Confronto interfaccia di inserimento dati tra prima (sinistra) e seconda (destra) versione app.	99
5.4	Confronto dato biologico campionato con apposita app (sinistra) e file Excel (destra).	100
5.5	Numero di esemplari campionati per specie e stadio vitale (N° esemplari tra 50 e 500).	105
5.6	Numero di esemplari campionati per specie e stadio vitale (N° esemplari maggiore di 500).	105
5.7	Numero di esemplari campionati per specie, per ciascuna tipologia di unità (N° HMU maggiori di 50, N° esemplari tra 50 e 500).	106
5.8	Numero di esemplari campionati per specie, per ciascuna tipologia di unità (N° HMU maggiori di 50, N° esemplari maggiori di 500).	106
5.9	Regioni con campionamenti di trota fario (<i>Salmo trutta</i>) adulta, giovane e totale per regioni campionate.	107
5.10	Importanza variabili da algoritmo BORUTA.	109
5.11	Stabilizzazione dell' E_{OOB} in funzione del numero di alberi, valore di stabilizzazione: 240 alberi.	109
5.12	Partial Dependence Plots per il modello parsimonioso oversampled per ammocete di lampreda padana.	110

Elenco delle tabelle

1.1	Caratteristiche UMI da rilevare secondo la metodologia MesoHAB-SIM ([16]).	6
1.2	Classi di integrità dell'habitat secondo l'Indice IH, [16].	16
2.1	Dato biologico registrato su supporto informatico (foglio Excel). . .	34
2.2	Formato scheda dati di input grezzi idro-morfologici	37
2.3	Esempio di Hydro_Sheet_River_*m3s_ddmmyyyy	39
2.4	Esempio di Bio_Sheet_River_*m3s_ddmmyyyy	39
2.5	Esempio di Resume_AJ_*m3s_ddmmyyyy	41
2.6	Esempio di Resume_AJ_*m3s_ddmmyyyy per rilievo biologico senza soglia A/J.	41
2.7	Esempio di Resume_All_*m3s_ddmmyyyy.	41
2.8	Estratto elenco siti caricati nel database.	44
2.9	Estratto tabella dati idro-morfologici.	44
2.10	Estratto tabella dati biologici.	44
2.11	Tabella estrazione sanguinerola (<i>Phoxinus phoxinus</i>) adulta (estratto).	45
2.12	Tabella estrazione sanguinerola (<i>Phoxinus phoxinus</i>) giovane (estratto).	45
3.1	Valori campi accettati dato biologico.	71
3.2	Estratto tabella CODE univoci elaborata con lo strumento "Site List"	76
3.3	Estrazione dato biologico con "Extract Biological Data" per il varrone (<i>Telestes muticellus</i>) adulto.	81
4.1	Dato biologico raccolto con apposita app, Carpasina, 16/05/2022	94
5.1	Informazioni generali riassuntive per rilievi idro-morfologici e biologici.	102
5.2	<i>a)</i> Condizioni massime e minime di portata rilievi. <i>b)</i> Numero minimo e massimo di unità campionate. <i>c)</i> Numero di campionamenti quantitativi e semi-quantitativi effettuati. <i>d)</i> Numero minimo e massimo di esemplari individuati.	103
5.3	Numero di rilievi e unità campionate per regione.	103

5.4	Analisi esemplari campionati per tipologia di unità.	104
5.5	Analisi esemplari campionati per tipologia di unità.	108
5.6	Performance predittive modello statistico Lampreda padana.	109

Indice

Sommario	I
Abstract	III
Elenco delle figure	V
Elenco delle tabelle	IX
Introduzione	1
1 Modellazione dell’habitat fluviale	3
1.1 Metodologia MesoHABSIM	4
1.1.1 Rilievo idro-morfologico e descrizione habitat fisico	4
1.1.2 Modellazione statistica di distribuzione di specie	9
1.1.3 Disponibilità spazio-temporale di habitat	13
1.2 SimStream Web	17
1.2.1 Input SimStream-Web	17
1.2.2 Output SimStream-Web	22
2 Organizzazione Dati Idro-morfologici e Biologici	27
2.1 Formato pregresso e formattazione dati per applicazione MesoHAB-SIM	28
2.1.1 Dati idro-morfologici	28
2.1.2 Dati biologici	32
2.2 Struttura Database	34
2.2.1 Schema generale ”MesoHABSIM DataBase”	35
2.3 Tipologia dati di input	37
2.3.1 Input dato grezzo idro-morfologico	37
2.3.2 Input dato grezzo biologico	37
2.4 Esportazione output con ”Hydro Bio Tool”	38
2.4.1 Formato schede di output	39
2.5 Codici univoci per il confronto di HMUs	42

2.6	Componenti "MesoHABSIM DataBase"	43
2.6.1	Sheets DataBase	43
2.6.2	Tools aggiuntivi	45
2.7	Modello Random Forest per Lampreda padana (Lampetra zanandreae)	46
2.7.1	Area di studio	46
2.7.2	Configurazione modello RF	46
3	Raccolta, Elaborazione ed Immagazzinamento Dati	50
3.1	App "Bio Sampling" per raccolta dati biologici in campo.	50
3.1.1	Inserimento caratteristiche generali sito di campionamento	50
3.1.2	Registrazione dato biologico	54
3.1.3	Modifica dati inseriti	59
3.1.4	Salvataggio file campionamento finale	61
3.2	"Hydro Bio Tool" per inserimento ed elaborazione dati.	63
3.2.1	Inserimento e Validazione Generalità del Rilievo	63
3.2.2	Inserimento e Validazione Dati Grezzi di Input	64
3.2.3	Selezione dato idro-morfologico	66
3.2.4	Selezione dato biologico	67
3.2.5	Validazione dato biologico	68
3.3	"MesoHABSIM Database" per Immagazzinamento ed Estrazione dati.	75
3.3.1	"Site List" per elenco siti	75
3.3.2	"Delete Relief" per rimozione dati	77
3.3.3	"Relief Map" per distribuzione spaziale rilievi	78
3.3.4	"Extract Biological Data" per estrazioni dati biologici	79
4	Attività di Campo e Test Strumenti	82
4.1	Area di studio	82
4.1.1	Campionamento biologico	83
4.1.2	Rilievo idro-morfologico	89
4.2	Trattamento dati	94
4.2.1	Elaborazione dati biologici	94
4.2.2	Elaborazione dati idro-morfologici	95
5	Risultati	97
5.1	Ottimizzazione App di registrazione dato biologico	97
5.2	Confronto dato grezzo campionato	100
5.3	Elaborazioni dati di campo	101
5.3.1	Output SimStream-Web	101
5.3.2	Output "Hydro Bio Tool"	101
5.4	Analisi "MesoHABSIM DataBase"	102
5.5	Preferenza di habitat per ammocete di Lampetra zanandreae	108

Conclusioni 112

Bibliografia 115

Introduzione

L'ecoidraulica è una scienza che lega discipline come biologia, ecologia, geomorfologia fluviale, idraulica e idrologia.

Tra i suoi campi di ricerca rientra anche la modellazione dell'habitat fluviale che nasce dalla necessità di descrivere le interazioni tra l'ecosistema fluviale e l'idromorfologia locale del corso d'acqua, al fine di valutare l'impatto di eventuali alterazioni sullo stesso.

Tra le metodologie per la modellazione dell'habitat fluviale ricopre un ruolo fondamentale il modello MesoHABSIM (Meso-habitat Simulation Model), con cui è possibile valutare la disponibilità di habitat per la fauna acquatica.

L'applicazione del suddetto approccio necessita l'esecuzione di rilievi idro-morfologici per la mappatura dell'habitat fisico a cui verrà affiancata la quantificazione della distribuzione (in termini di assenza, presenza e abbondanza) delle specie di fauna presenti.

Associando alla descrizione fisica dell'habitat, quello che è il dato biologico campionato in ciascuna unità, sarà possibile generare modelli statistici binari di assenza/presenza e presenza/abbondanza con cui poter classificare ogni unità idromorfologica come non idonea (assenza della specie), idonea (presenza della specie), oppure ottimale (abbondanza della specie).

Dalle attività di campo svolte, dal 2009 ad oggi, per l'implementazione della metodologia, è derivata un'elevatissima quantità di dati idro-morfologici (1303 mesohabitat registrati), e biologici (più di 19000 pesci campionati), descritti in una moltitudine di formati differenti, che con il presente lavoro di tesi sono stati raccolti, formattati e organizzati.

Ad oggi, per la mappatura dell'habitat fisico è prevista una procedura standardizzata tramite il plugin MapStream, operante in ambiente QGIS, attraverso cui è possibile disporre di un dato di campo ben strutturato.

Per i campionamenti biologici non esiste invece alcun software di questo genere, per questo motivo, nel presente lavoro, ci si è affacciati a tale problematica attraverso lo sviluppo di un'apposita App, con cui è possibile registrare il dato biologico in campo in formato standardizzato.

Questi dati sono stati organizzati in un unico database, a scala nazionale ed unico

nel suo genere, al fine di disporre di una banca dati per l'implementazione della metodologia MesoHABSIM, soprattutto grazie alla generazione di nuovi modelli statistici di distribuzione di specie. L'organizzazione dei dati, per un utilizzo così specifico, ha richiesto lo sviluppo di appositi strumenti di elaborazione e gestione del dato.

Disponendo quindi di un dato idro-morfologico, facilmente e univocamente collegabile a quello biologico è stato possibile estrarre importanti informazioni relative allo stato di conservazione della fauna ittica italiana. Si è analizzato il caso della lampreda padana (*Lampetra zanandreae*), per cui è stato costruito, attraverso la tecnica del Random Forest (RF), il modello di assenza/presenza per lo stadio vitale ammocete. Il modello parsimonioso prodotto ha espresso ottime performance predittive evidenziando un'accuratezza totale (total accuracy) del 95,2% e true skill statistics (TSS) del 90,4 %.

Capitolo 1

Modellazione dell'habitat fluviale

I fiumi sono spesso soggetti a forti pressioni idro-morfologiche, come l'alterazione del regime di deflusso naturale idrologico, che provoca una diminuzione della disponibilità di habitat acquatico per la comunità biotica esistente. È necessario quantificare la disponibilità spatio-temporale di habitat attraverso lo studio delle relazioni che intercorrono tra l'idro-morfologia locale del corso d'acqua e l'ecosistema dello stesso.

In funzione della risoluzione spaziale si distinguono modelli a scala del microhabitat, di dimensione media 10 cm ([13]) e quelli a scala del mesohabitat, dove la dimensione può essere compresa tra 10^{-1} e 10^3 m ([9]). Nel presente lavoro di tesi si è fatto riferimento all'approccio a meso-scala, denominato MesoHABSIM ([8], [9]).

L'approccio del meso-habitat offre numerosi vantaggi in contrapposizione alle metodologie a micro-scala (es. PHABSIM, [1]) in quanto:

- La classificazione dei dati morfologici è adattabile a qualsiasi tipologia di corso d'acqua ed è basata su un sistema di classificazione robusto ([13]).
- Permette l'utilizzo di un set di variabili ambientali più ampio che si rispecchia in un'analisi più mirata delle specie e delle comunità target a livello biologico (ad es. [17], [18], [19]).
- Non necessita l'accoppiamento a modelli di simulazione idraulica ed è quindi utilizzabile anche nei corsi d'acqua, come quelli a regime torrentizio, dove venga meno l'ipotesi di acqua bassa su cui i modelli sono basati.

1.1 Metodologia MesoHABSIM

Il Meso-habitat Simulation Model è stato sviluppato, a partire dagli anni 2000, dal Prof. Piotr Parasiewicz ([7], [8]), ed è stata adattata al contesto italiano da Vezza et al.([10], [15], [18], [19]).

La metodologia si articola in 3 fasi principali, rappresentate in fig. 1.1, ciascuna delle quali necessita di un input per poter essere avviata, generando uno specifico output. Le fasi procedurali sono le seguenti:

- Descrizione dell'habitat fisico, tramite rilievo idro-morfologico, per diverse condizioni di deflusso.
- Applicazione dei modelli biologici di idoneità di habitat.
- Analisi variazioni spazio-temporali dell'habitat.

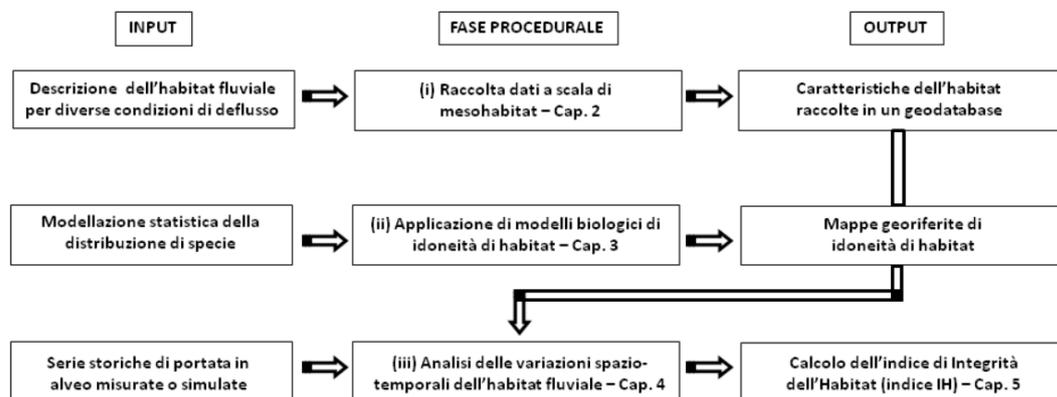


Figura 1.1: Principali fasi MesoHABSIM e relativi input e output, da [16].

Nel seguito sono analizzate le diverse macro fasi.

1.1.1 Rilievo idro-morfologico e descrizione habitat fisico

Di seguito viene descritta la procedura, prevista dalla metodologia MesoHABSIM, per la realizzazione dei rilievi idro-morfologici.

Per la descrizione dell'habitat fisico vengono seguite le indicazioni del manuale SUM (Sistema di rilevamento e classificazione delle Unità Morfologiche dei corsi d'acqua, [13]) secondo cui la descrizione delle unità idro-morfologiche deve avvenire all'interno di un porzione di fiume, denominata sottotratto, rappresentativa della distribuzione delle unità morfologiche tipiche del tratto (scala spaziale tra 10^{-1} e 10^1 km). La procedura di definizione del tratto è dettagliatamente spiegata nel

manuale sopra citato.

La strumentazione necessaria per la realizzazione del rilievo idro-morfologico è la seguente:

- Telemetro laser per la mappatura delle unità idro-morfologiche.
- Treppiede per il sostegno del telemetro.
- Computer o tablet da campo per l'utilizzo dei software necessari (es. QGIS).
- Correntometro per la misura della portata al momento del rilievo, della velocità e della profondità della corrente.
- Bussola per la calibrazione del telemetro.

In primo luogo viene montata la strumentazione necessaria e vengono calibrati i sensori interni del telemetro (inclinometro e bussola elettronica).

Viene poi eseguita la mappatura del mosaico di meso-habitat (denotate nel seguito anche come UMI, unità idro-morfologiche, o HMU, hydromorphologic unit) che prevede, dopo l'acquisizione dell'immagine georeferenziata di background, sul computer da campo e tramite software GIS, l'inserimento del punto stazione di partenza, rappresentativo della posizione del telemetro da cui viene poi mappato il perimetro delle HMU.

Per ciascuna unità sono poi raccolti i descrittori ambientali; per alcuni si misura un valore medio per l'intera UMI, per altri, come la distribuzione di profondità, velocità e substrato, sono necessarie misure puntuali in almeno 15 punti di misura (2 punti m^2 o non meno di 7 per unità di estensione $< 5m^2$) rappresentativi delle caratteristiche idrodinamiche della stessa.

Per la mappatura e la raccolta dei descrittori ambientali viene utilizzato Map-Stream, plugin del software QGIS, descritto al punto 1.1.1.1.

I parametri fisici misurati durante i rilievi in campo sono riportati in tabella 1.1 (reperita dal [16]).

All'inizio e al termine del rilievo, valutando il campo di velocità della corrente in una sezione trasversale, viene misurata la portata defluente in alveo.

Nel capitolo 4 verrà ripresa la descrizione delle operazioni, eseguite in 9 siti liguri, con annessa documentazione fotografica.

Tabella 1.1: Caratteristiche UMI da rilevare secondo la metodologia MesoHABSIM ([16]).

DESCRITTORE	UNITA' DI MISURA	CLASSI	CATEGORIE/DESCRIZIONE
Principali unità UMI	si/no	13	POTHOLE, CASCADE, RAPID, RIFFLE, STEP, GLIDE, POOL, DUNE, AQUAT_VEG, SEC_CHAN, FLOOD_LAKE, WETLAND, ARTIF_ELEM, WATERFALL, PLUNGE_POOL, BACKWATER, ROCK_GLIDE, ISOL_POND
Pendenza	%	1	pendenza media della superficie libera dell'UMI
Zone di rifugio	si/no	10	BOULDER, CANOP_SHAD, OVERHA_VEG, ROOTS, SUBMER_VEG, EMERG_VEG, UNDERC_BAN, WOODY_DEBR, RIPRAP, SHALL_MARG
Substrato	% di misurazioni random stratificate	12	GIGALITHAL, MEGALITHAL, MACROLITHAL, MESOLITHAL, MICROLITHAL, AKAL, PSAMMAL, PELAL, DETRITUS, XYLAL, SAPROPEL, PHYTAL
Profondità della corrente	% misurazioni random stratificate	9	Classi definite da intervalli di 15 cm (fino a ≥ 120 cm)
Velocità della corrente	% misurazioni random stratificate	9	Classi definite da intervalli di 15 cm/s (fino a ≥ 120 cm/s)
Numero di Froude	-	1	media sull'area dell'unità morfologica

1.1.1.1 Mapstream

Dopo aver acquisito l'immagine di background del sotto tratto in analisi, viene avviato MapStream. Creando un nuovo progetto potranno essere inseriti il nome del corso d'acqua, la data del rilievo e il sistema di riferimento desiderato (nel seguente lavoro di tesi si è utilizzato, e si farà riferimento al WGS84/UTM zone 32N). Dopo la creazione del progetto sono generati, in QGIS, i 4 layers osservabili in fig. 1.2.

In VERTEX e in HMU sono contenuti rispettivamente i vertici e i poligoni delle UMI, STATION contiene invece i punti stazione, da cui sono state mappate le unità, infine nel layer MEAS sono contenute le informazioni di profondità, velocità e substrato, campionate per ciascun poligono.

Dopo aver eseguito la connessione Bluetooth tra computer da campo e telemetro sarà possibile procedere alla registrazione del contorno delle unità.

In primo luogo occorrerà selezionare la stazione di riferimento per l'acquisizione dei dati; il punto selezionato diventerà di colore rosso e sarà possibile modificare il layer VERTEX, registrando i punti del poligono dell'unità (come rappresentato in fig. 1.3).

Terminato il rilievo topografico dei vertici, viene generato il poligono rappresentato in fig. 1.4; per ciascun nuovo poligono saranno inseribili la tipologia di unità e le

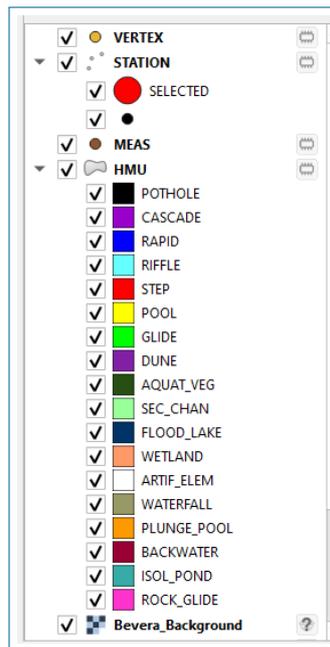


Figura 1.2: Layers progetto MapStream.



Figura 1.3: Stazione selezionata e Layer VERTEX.

covers riconosciute. Selezionando l'unità desiderata, potranno poi essere inserite le misurazioni locali di profondità (cm), velocità (m) e substrato, per ciascun punto MEAS di quella unità; in fig. 1.5 è riportata l'interfaccia di inserimento delle misure.

In figura 1.6 è osservabile l'aspetto dei layer terminata la fase di misurazione per la prima unità. Ripetendo il procedimento di registrazione delle UMI e delle MEAS, è possibile, seguendo una poligonale con i punti stazione, generare il mosaico delle meso-unità per il sottotratto di interesse.

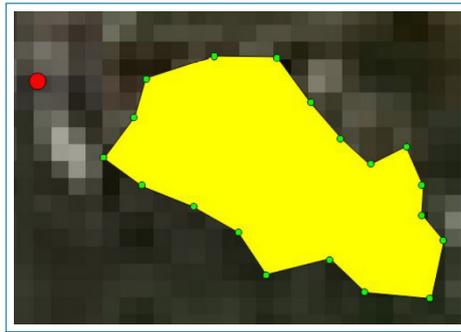


Figura 1.4: Layer HMU.

MEAS - Attributi elemento

HMU NUM.	HMU TYPE	
1	Pool	
SD	AZ	INC
12	345	15
DEPTH [cm]	VELOCITY [m/s]	
23	0.34	
SUBSTRATE	MEAS. NUM.	
Macrolithal (20-40)	1	
COMMENTS	<input type="checkbox"/> ESTIMATED	

OK Cancel

Figura 1.5: Raccolta descrittori ambientali.

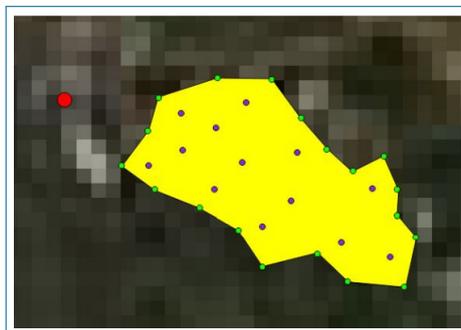


Figura 1.6: Layer MEAS.

1.1.2 Modellazione statistica di distribuzione di specie

Da studi disponibili in letteratura scientifica è noto come la comunità ittica risponda negativamente ad alterazioni, anche contenute, del regime idrologico naturale ([11]). La facilità di identificazione della fauna ittica e di quantificazione della loro risposta agli stress, unita al limitato numero di specie presenti e all'elevata valenza economica assunta da alcune specie (come i salmonidi), l'hanno resa la metrica fondamentale per la valutazione della qualità ambientale del corso([16]). Per stimare la quantità di habitat disponibile nel tratto analizzato è necessaria la realizzazione di modelli di distribuzione di specie che permetteranno di classificare ciascun meso-habitat, su cui sono applicati, come non idoneo, idoneo oppure ottimale.

I modelli che definiscono l'idoneità di habitat, per i diversi stadi vitali delle diverse specie target selezionate, si basano su dati biologici quantitativi raccolti in campo, alla mesoscala, e in assenza di alterazioni idro-morfologiche o artificiali per la comunità ittica).

La quantificazione biologica consiste nell'identificazione, per ciascuna HMU, della presenza e dell'abbondanza delle specie ittiche presenti per ciascun stadio vitale (per la tipologia di rilievi che nel presente lavoro viene indicata come semi-quantitativa, S) o della raccolta degli individui presenti, descritti ciascuno per lunghezza e peso (per la tipologia di rilievi indicata come quantitativa, Q); in figura 1.7 è riportata la fase di misurazione di due esemplari pescati nelle unità mappate. La descrizione dell'esecuzione di un campionamento biologico viene descritta nel capitolo 4.1.1.

Associando il dato biologico di assenza, presenza o abbondanza (per stadio vitale e specie selezionata), al valore dei descrittori fisici caratterizzanti ciascuna unità, tramite la tecnica di machine learning Random Forest, verranno costruiti 2 modelli statistici binari di assenza/presenza e presenza/abbondanza.

L'applicazione dei modelli statistici (disponibili in SimStream-Web), genererà come risultato la probabilità di presenza o abbondanza, che porterà alla classificazione di ciascuna UMI come non idoneo, idoneo, per prob. presenza > 0.5 , o ottimale, con prob. abbondanza > 0.5 ([16]). In figura 1.8 è riportato l'esempio di una mappa di idoneità di habitat per lo stesso tratto rappresentato nella figura precedente (in rosso sono evidenziati i meso-habitat non idonei mentre in giallo quelli idonei, nel caso analizzato non sono presenti habitat ottimali).

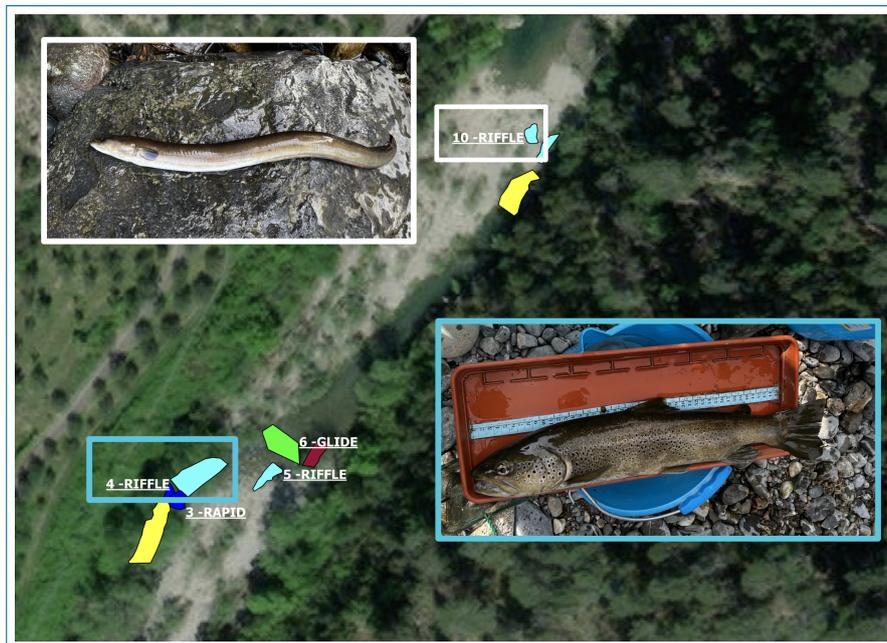


Figura 1.7: Esempio di anguilla (a sinistra) e trota mediterranea (a destra) campionate all'interno di 2 riffle, Bevera, $0.15 \text{ m}^3/\text{s}$, 11/03/2022.



Figura 1.8: Mappa di idoneità di habitat per l'Anguilla allo stadio vitale giovanile, Bevera, 11/03/2022, $0.15 \text{ m}^3/\text{s}$.

1.1.2.1 Random Forest

Il Random Forest è una tecnica ensemble di machine learning basata sulla combinazione di un set di alberi di decisione. Nel RF non viene ottimizzato il miglior predittore possibile, come avviene per i CART (Classification And Regression Tree), ma vengono generati, e successivamente aggregati, una serie di alberi, non necessariamente ottimali, ma che generano migliori performance predittive ([3]).

Nel presente lavoro di tesi, la variabile risposta che si desidera prevedere è categorica (assenza/presenza o presenza/abbondanza di specie), viene quindi trattato un problema di classificazione a due classi. Come osservabile in figura 1.9, l'algoritmo opera suddividendo il dataset di training in t campioni di bootstrap; per ciascuno di questi verrà costruito, tramite partizionamento ricorsivo binario, un albero di classificazione non potato (unpruned). Un campione bootstrap è generato, estraendo randomicamente un numero prefissato di osservazioni, con ripetizione, dal dataset di training. Ad ogni nodo, il dataset originale viene poi suddiviso in 2 nodi figli, sulla base della variabile predittrice che riduce maggiormente l'indice di Gini, in modo che si ottengano nodi risultanti il più puri possibile (cioè contenenti osservazioni classificate dalla stessa etichetta di classe). Le m variabili tra cui selezionare l'attributo per il miglior splitting, vengono estratte in modo casuale e senza ripetizione tra le p variabili totali.

Ciascun albero costituente la foresta viene quindi addestrato su un diverso dataset, selezionando per lo split, ad ogni nodo, un diverso sotto insieme di variabili senza ripetizione. La previsione su un'osservazione verrà effettuata generando la previsione con ciascun albero della foresta e associando quell'etichetta di classe predetta il maggior numero di volte.

Grazie alle perturbazioni introdotte in primis dall'estrazione casuale dei campioni di bootstrap e, successivamente, dall'estrazione random delle variabili in cui ricercare quella per il miglior splitting, si ottengono alberi di classificazione molto diversificati che, assemblati, generano migliori performance predittive; come riportato in [14], il RF accresce la foresta di alberi decisionali non solo manipolando il dataset di training (usando campioni bootstrap), ma anche gli attributi di input (usando sottoinsiemi diversi di attributi ad ogni nodo).

Un grande vantaggio di tale tecnica risiede nel fatto che è già presente al suo interno una stima diretta dell'errore commesso nella predizione, l'Out Of Bag error (E_{OOB}). Questo tipo di cross-validazione evita la necessita di un campione di test. L'idea di base, come riportato in [3], è quella di usare come test le osservazioni non selezionate nel campione di bootstrap (in media il 37% delle osservazioni). L' E_{OOB} viene calcolato con la 1.1:

$$E_{OOB} = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{Y_i \neq \hat{Y}_i} \quad (1.1)$$

Per ognuna delle n osservazione del dataset di addestramento, per predire la risposta associata all'osservazione i -esima X_i , vengono aggregati tutti i t predittori costruiti con campioni bootstrap non contenenti l'osservazione (X_i, Y_i) ; a ciascuna osservazione viene associata la stima \hat{Y}_i con cui sarà possibile calcolare l'errore. Il vantaggio principale è quindi quello di stimare l'errore con osservazioni non utilizzate per l'accrescimento dell'albero.

Le operazioni eseguite per la costruzione del modello statistico di presenza/assenza per la Lampetra zanandreai allo stadio giovanile sono descritte nella sezione 2.7.

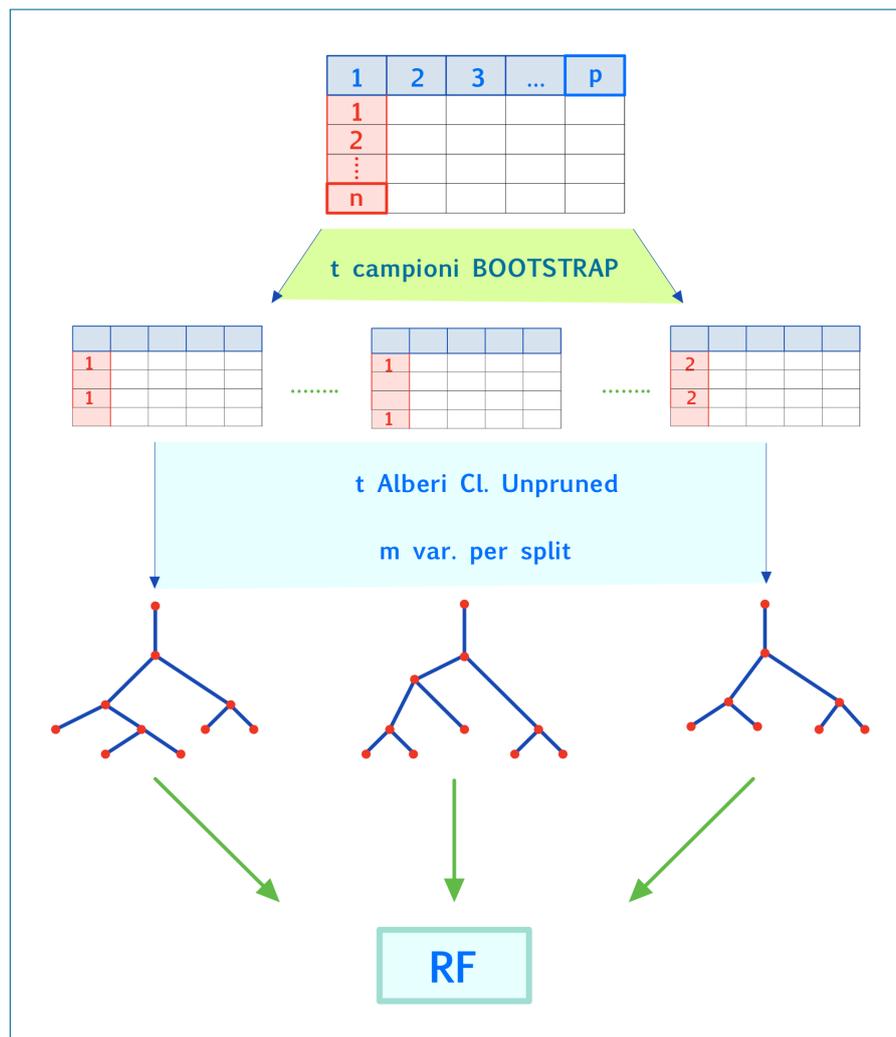


Figura 1.9: Schema generale Random Forest.

1.1.3 Disponibilità spazio-temporale di habitat

Seguendo la procedura riportata in [16], i dati idro-morfologici raccolti vengono applicati i modelli statistici di distribuzione di specie con cui è valutata l' idoneità di ciascun meso-habitat. Successivamente, ai fini della metodologia MesoHABSIM, viene quantificato, con la 1.2, l'habitat totale disponibile (H_d) per il tratto analizzato.

$$H_d = H_I * 0.25 + H_O * 0.75 \quad (1.2)$$

H_I indica la quantità di habitat idoneo mentre H_O quella ottimale, . Rappresentando sull'asse delle ordinate l'habitat totale disponibile in percentuale (rispetto all'area bagnata relativa al valore massimo di portata analizzata) in funzione della portata defluente (in m^3/s), viene definita la curva habitat-portata osservabile in figura 1.10.

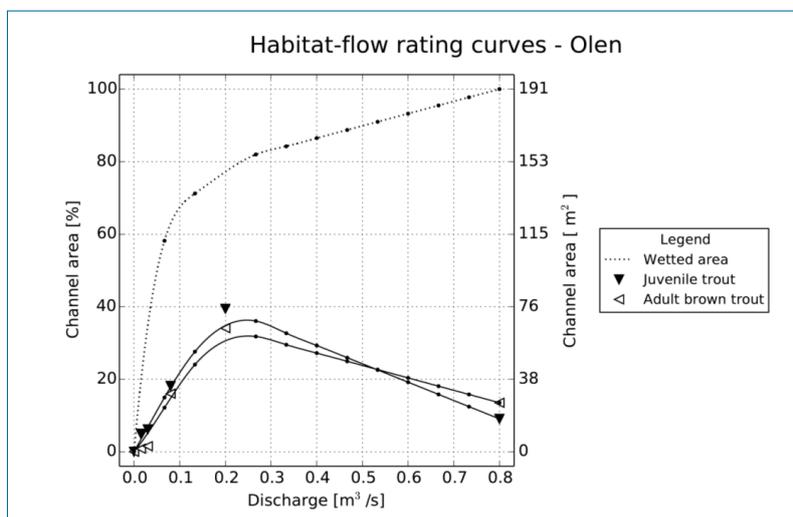


Figura 1.10: Esempio di curva habitat-portata, per la trota fario adulta e giovane, nel torrente Olen ([16]).

La curva habitat-portata descrive la variazione spaziale di habitat in funzione della portata; associando a tale curva, attraverso la 1.3, la serie storica di portate in alveo, sarà possibile rappresentare la variazione temporale di habitat disponibile rappresentata in figura 1.11.

$$H_d(t) = H(Q(t)) \quad (1.3)$$

H rappresenta la relazione H-Q per la determinata specie e stadio vitale, $Q(t)$ e $H_d(t)$, rispettivamente la portata registrata e l'habitat disponibile al tempo t . Sulle serie storiche di habitat vengono effettuate analisi statistiche che permettono di stabilire diverse metriche di interesse.

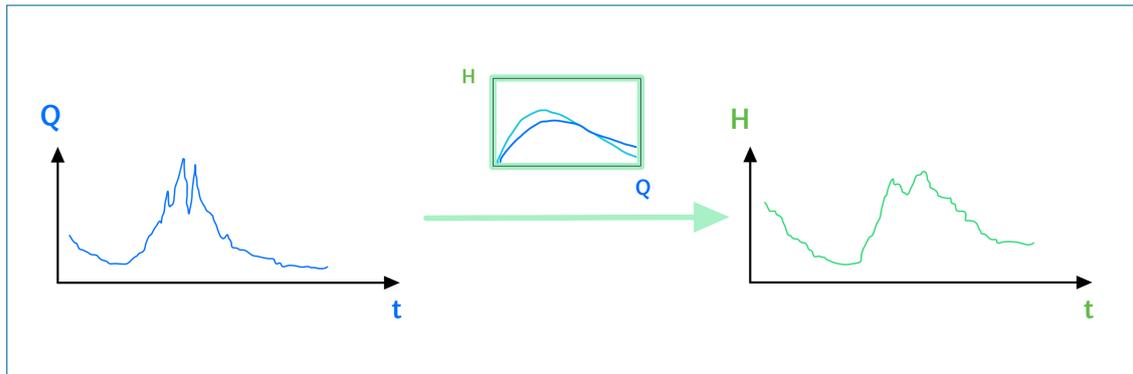


Figura 1.11: Generazione serie storiche di habitat.

Ai fini dell'applicazione della metodologia MesoHABSIM è fondamentale la valutazione dell'indice di integrità di habitat, dell'habitat fluviale IH; tale indice viene utilizzato per valutare l'integrità dell'habitat in presenza di alterazioni antropiche sia di carattere idrologico, come derivazioni e prelievi, sia morfologico, come di-salvei. L'IH è composto dai sub-indici ISH e ITH, ricavati dalle serie storiche di habitat, rappresentanti rispettivamente l'Indice di disponibilità Spaziale e Temporale di Habitat.

L'indice ISH valuta la quantità media di habitat persa per una determinata specie e stadio vitale, in un determinato periodo di tempo, a seguito di un'alterazione. Per la valutazione di impatto e la determinazione dei deflussi ecologici, l'ISH è valutato, con la 1.4, su base annuale.

$$ISH = \min \left(\left\{ \begin{array}{l} 1 - \frac{|A_{Hd,r} - A_{Hd}|}{A_{Hd,r}}, \frac{|A_{Hd,r} - A_{Hd}|}{A_{Hd,r}} \leq 1 \\ 0, \frac{|A_{Hd,r} - A_{Hd}|}{A_{Hd,r}} > 1 \end{array} \right. \right)_{specie} \quad (1.4)$$

Dove:

$A_{Hd,r}$ indica il valore medio della serie storica in condizioni di riferimento (assenza dell'alterazione).

A_{Hd} indica la quantità media di habitat disponibile in condizioni alterate.

$A_{Hd,r}$ e A_{Hd} sono espresse in m^2 o in % dell'alveo bagnato. Il valore minimo di ISH tra quelli valutati per ciascuna specie e stadio vitale viene scelto come valore rappresentativo per tutelare qualsiasi esemplare presente nel corso d'acqua.

Attraverso l'indice ITH viene invece valuta, per la comunità ittica, la variazione temporale nella durata degli eventi in cui l'habitat si mantiene sotto la soglia di habitat disponibile definita dall' AQ_{97} (m^2 di habitat corrispondente al 97° percentile di portata in condizioni non alterate). Dalla serie temporale di habitat, fissando la soglia AQ_{97} , viene costruita la curva di durata sotto-soglia (CDS o UCUT, Con-

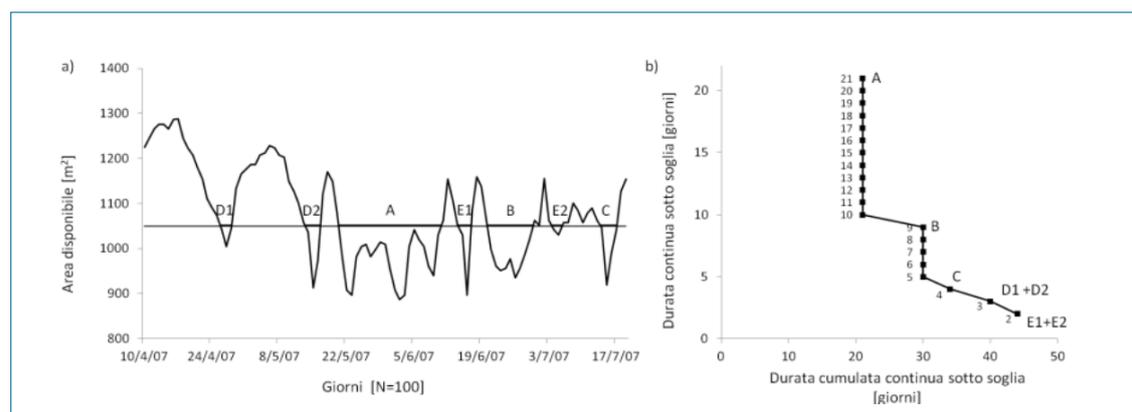


Figura 1.12: Costruzione della curva UCUT da serie temporale di habitat disponibile, [16].

tinuous Under-Threshold curves, [9]), che riporta sull'asse delle ordinate il valore della durata continua sotto soglia, mentre sull'asse delle ascisse la durata cumulata continua sotto soglia (un esempio di costruzione della CDS è osservabile in 1.12). Per la valutazione dell'impatto di una derivazione, per una determinata specie e stadio vitale, la curva sopra descritta viene costruita sia per condizioni di assenza della derivazione (dalla serie storica di habitat in condizioni di riferimento), sia per condizioni alterate. Attraverso la distanza media tra le due curve viene determinata l'alterazione degli eventi di stress.

Con la 1.5 viene calcolato il parametro AGS che rappresenta l'alterazione complessiva nel numero di giorni di stress, [10].

$$AGS = \frac{1}{d_{max,r}} \sum_{k=1}^{k=d_{max,r}} \left(\frac{|d_{c,AQ97} - d_{c,r,AQ97}|}{d_{c,r,AQ97}} \right) \quad (1.5)$$

Dove:

$d_{max,r}$ rappresenta la massima durata continua sotto-soglia in condizioni di riferimento, mentre $d_{c,AQ97}$ e $d_{c,r,AQ97}$ rappresentano rispettivamente il numero di giorni di durata cumulata continua sotto-soglia in condizioni alterate e di riferimento.

Con la 1.6, viene determinato l'indice ITH per il tratto analizzato, che è un numero compreso tra 0 e 1 corrispondente al minimo tra i valori calcolati per ciascuna specie e stadio vitale.

$$ITH = \min \left(e^{-0.38AGS} \right)_{specie} \quad (1.6)$$

Dalla 1.7 viene determinato il valore dell'Indice di integrità dell'Habitat fluviale (IH) il cui valore è inversamente proporzionale al grado di alterazione.

$$IH = \min(ISH, ITH) \quad (1.7)$$

L'integrità dell'habitat fluviale, come da Direttiva Quadro delle Acque, viene suddivisa nelle 5 classi in tabella 1.2.

Tabella 1.2: Classi di integrità dell'habitat secondo l'Indice IH, [16].

IH	CLASSE
$IH \geq 0.80$	ELEVATO
$0.60 \leq IH < 0.80$	BUONO
$0.40 \leq IH < 0.60$	SUFFICIENTE
$0.20 \leq IH < 0.40$	SCADENTE
$IH < 0.20$	PESSIMO

La metodologia MesoHABSIM rappresenta quindi uno strumento fondamentale, e supporto decisionale, nel processo di pianificazione della gestione delle risorse idriche superficiali; tramite il confronto della disponibilità di habitat nelle condizioni di riferimento e alterate, permette la valutazione di diversi scenari di future alterazioni.

1.2 SimStream Web

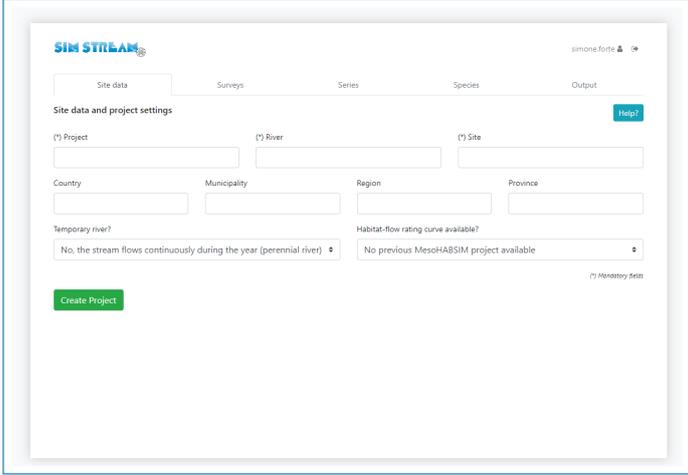
SimStream-Web rappresenta la versione operante come servizio web del software SimStream che rappresenta lo strumento informatico di riferimento, a livello internazionale, per l'applicazione della metodologia MesoHABSIM. Nel presente lavoro di tesi è stato utilizzato SimStream-Web per l'elaborazione dei dati e l'applicazione della metodologia (ci si riferirà nel seguito SimStream-Web come SimStream).

Le funzionalità di SimStream-Web sono molteplici; in primo luogo è possibile scaricare risultati dal database di progetti già creati, oppure creare una nuova sessione di lavoro, per la valutazione dell'habitat fluviale, e ancora calcolare l'indice IH dall'integrazione di risultati di progetti precedenti.

Un nuovo progetto sarà sviluppabile per fiumi a regime idrologico perenne o temporaneo.

1.2.1 Input SimStream-Web

Nel presente lavoro di tesi si è fatto riferimento a fiume perenni. Nel seguito ci si è soffermati sugli aspetti fondamentali del software analizzati, quali la formattazione dei dati in Input e la tipologia di output. All'avvio della sessione di calcolo, sono inseribili le caratteristiche generali del sito analizzato, come osservabile in fig. 1.13.



The screenshot displays the SimStream-Web web interface. At the top, there is a navigation menu with tabs for 'Site data', 'Surveys', 'Series', 'Species', and 'Output'. The 'Site data' tab is currently selected. Below the navigation, the page title is 'Site data and project settings'. The form contains several input fields: 'Project' (marked with a red asterisk), 'River' (marked with a red asterisk), and 'Site' (marked with a red asterisk). Below these are fields for 'Country', 'Municipality', 'Region', and 'Province'. There are two dropdown menus: 'Temporary river?' with the selected option 'No, the stream flows continuously during the year (perennial river)', and 'Habitat-flow rating curve available?' with the selected option 'No previous MesoHABSIM project available'. A green 'Create Project' button is located at the bottom left of the form area. The user's name 'simone.forte' is visible in the top right corner.

Figura 1.13: Schermata SimStream-Web per il caricamento delle generalità del progetto.

Dopo l'inserimento dei campi obbligatori, relativi al nome del progetto, fiume e sito del rilievo, saranno inseribili, dalla schermata in fig. 1.14, i dati idro-morfologici del rilievo, per i fiumi permanenti:

- Data del rilievo.
- Portata al momento del rilievo (m^3/s), valori da inserire in ordine crescente.
- Shapefile delle unità idro-morfologiche (HMUs), costituito da 4 file di diversa estensione (.shp, .shx, .dbf, .prj)
- File di testo .txt delle misure puntuali di velocità, profondità e substrato.

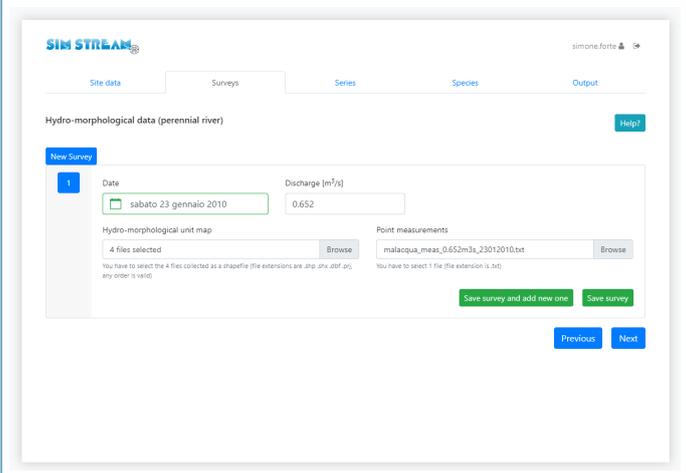
The screenshot shows the 'New Survey' form in the SimStream Web application. The form is titled 'Hydro-morphological data (perennial river)'. It includes a 'Date' field with the value 'sabato 23 gennaio 2010', a 'Discharge (m³/s)' field with the value '0.652', and two file selection sections: 'Hydro-morphological unit map' with '4 files selected' and 'Point measurements' with 'malacqua_meas_0.652m3s_23012010.txt'. There are 'Browse' buttons for both file sections, and 'Save survey and add new one' and 'Save survey' buttons at the bottom. Navigation buttons 'Previous' and 'Next' are also visible.

Figura 1.14: Schermata SimStream-Web per il caricamento dei dati dei rilievi.

Inseriti i precedenti dati, viene testata dal software la consistenza fisica tra tutti i rilievi inseriti, tramite le seguenti regole:

- L'area bagnata deve aumentare all'aumentare della portata defluente in alveo.
- La distribuzione delle velocità e delle profondità devono aumentare all'aumentare della portata defluente in alveo.
- La distribuzione dei substrati deve essere simile tra rilievi con simile estensione dell'area bagnata.

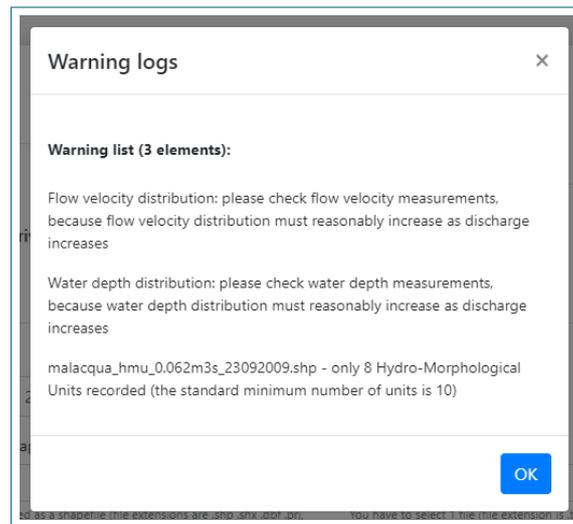


Figura 1.15: Messaggio di warning durante il caricamento dei rilievi in SimStream-Web.

Nel caso in cui queste regole non venissero rispettate, sarà attivato in automatico un messaggio di warning, come osservabile in figura 1.15.

Viene inoltre valutata la corretta formattazione dello shapefile e del file delle misurazioni puntuali.

Per lo shapefile delle HMUs è richiesto che queste siano della tipologia "poligono", non presentino buchi interni o auto intersezioni, come osservabile in figura 1.16. La tabella .dbf deve inoltre rispettare l'intestazione e ordine delle colonne seguente: HMU_NUM, HMU_TYPE, Z_MAX, Z_MIN, CONNECTIV, BOULDER, CANOP_SHAD, OVERHA_VEG, ROOTS, SUBMER_VEG, EMERG_VEG, UNDERC_BAN, WOODY_DEBR, RIPRAP, SHALL_MARG.

Le tipologie di unità selezionabili sono le seguenti POTHOLE, CASCADE, RAPID, RIFFLE, STEP, POOL, GLIDE, DUNE, AQUAT VEG, SEC CHAN, FLOOD LAKE, WETLAND, ARTIF ELEM, WATERFALL, PLUNGE POOL, BACKWATER, ROCK GLIDE, ISOLATED POND.

La presenza/assenza della cover deve essere indicata rispettivamente con True/False e le quote minima e massima dell'unità devono essere in metri, con il "." come separatore decimale.

Per il file delle misure puntuali deve essere rispettato il seguente ordine delle colonne: HMU_NUM, HMU_TYPE, PNTNUM, DEPTH, VELOCITY, SUBSTRATE. Il tipo di unità deve rispettare le stesse regole del file delle HMUs, PNTNUM deve essere un numero progressivo per ciascuna unità, DEPTH e VELOCITY devono utilizzare come separatore decimale il "." ed essere espresse rispettivamente in metri (m) e metri al secondo (m/s) e il substrato deve essere di una delle seguenti tipologie: GIGALITHAL, MEGALITHAL, MACROLITHAL, MESOLITHAL,

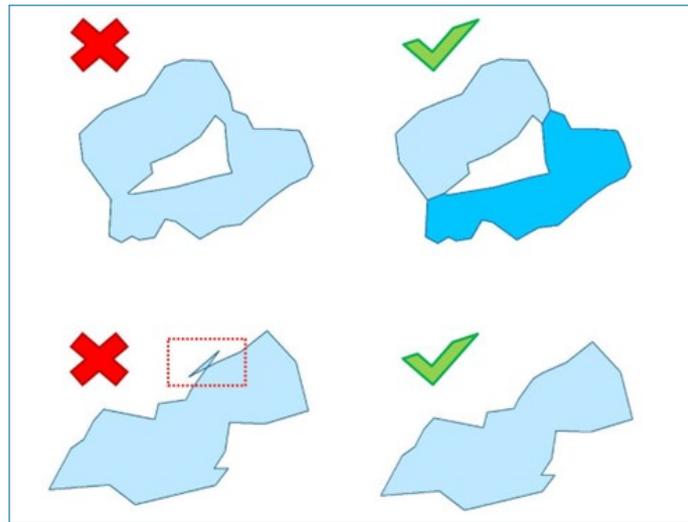


Figura 1.16: Geometria shapefile HMU non accettata da SimStream-Web, presenza di buchi (in alto) e auto intersezioni (in basso).

MICROLITHAL, AKAL, PSAMMAL, PELAL, DETRITUS, XYLAL, SAPROPEL, PHYTAL.

Caricati correttamente i dati, è possibile proseguire verso la pagina successiva; nel caso in cui la corretta formattazione di almeno un campo non venisse rispettata, non sarà invece possibile avanzare e occorrerà reinserire il dato previa modifica.

Nella schermata in figura 1.17 è prevista la possibilità di inserimento della serie temporale di portate, opportunamente formattata, prima del passaggio alla schermata di selezione della specie, dello stadio vitale e del modello biologico di interesse (fig. 1.18).

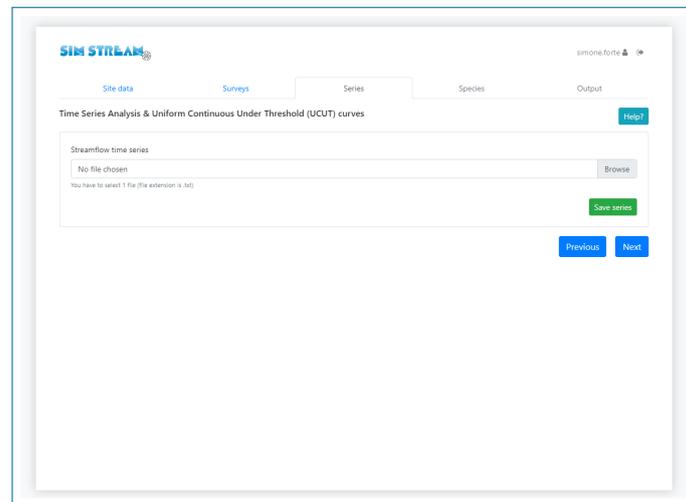


Figura 1.17: Schermata SimStream-Web di inserimento serie temporale di portata.

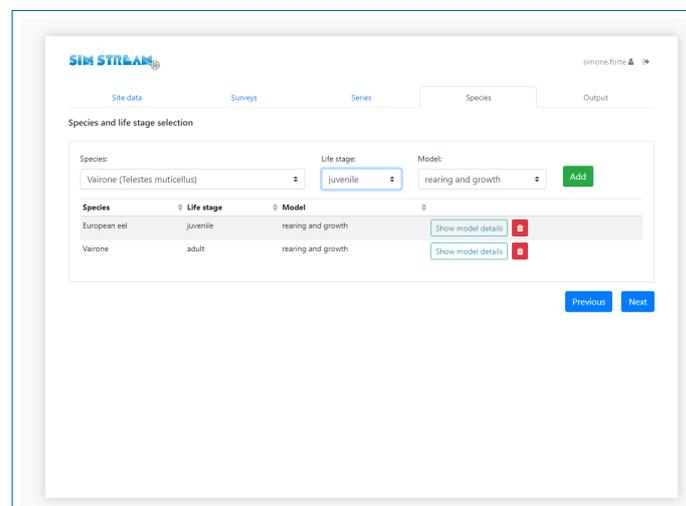


Figura 1.18: Schermata SimStream-Web di selezione specie, stadio vitale e modello biologico.

Per ciascuna specie selezionata è visualizzabile il modello statistico (es. per lo scazzone adulto, fig. 1.19), con cui viene valutata la disponibilità di habitat.

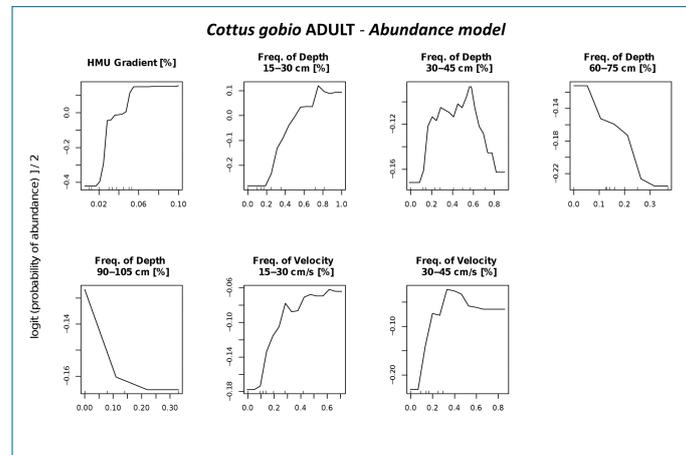


Figura 1.19: Dettagli modello biologico scazzone adulto (*Cottus gobio*).

1.2.2 Output SimStream-Web

Dopo la selezione è possibile passare alla finestra degli output, osservabile in figura 1.20; i file selezionati saranno raccolti nell'apposita cartella compressa, scaricabile dalla sezione di download.

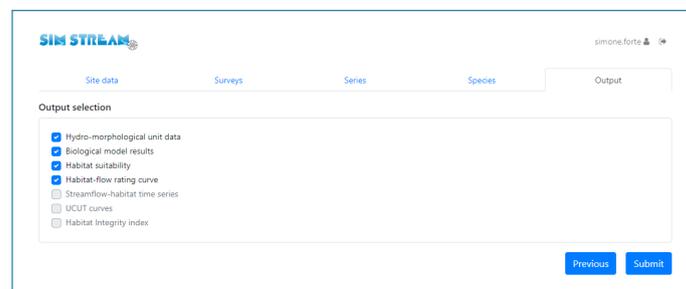


Figura 1.20: Schermata SimStream-Web di selezione output.

Gli output selezionabili sono i seguenti:

- Hydro-morphological unit data (shapefile): shapefile GIS delle UMI per ciascuna condizione di deflusso (fig. 1.21).
- Hydro-morphological unit data (.txt): file di testo contenente descrizione fisica delle HMUs.
- Hydro-morphological unit data (.pdf): file riassuntivo della distribuzione di frequenza di profondità, velocità e substrato, per ciascuna condizione di deflusso (fig. 1.22).
- Habitat suitability (shapefile): mappa di idoneità di habitat, per ciascuna specie e stadio vitale considerati, in ciascuna delle condizioni di deflusso analizzate (fig. 1.23).
- Habitat suitability (.txt): file di testo riportante l'idoneità di habitat, per ciascuna specie e stadio vitale di interesse, in ciascuna delle condizioni di deflusso analizzate.
- Biological model (shapefile): risultato del modello biologico, per ogni specie e stadio vitale di interesse, per ciascuna delle condizioni di deflusso analizzate. Sono riportati i risultati di due modelli statistici binari, rispettivamente di Assenza/Presenza per classificare ciascun meso-habitat come non_idoneo/idoneo e di Presenza/Abbondanza per la classificazione in Idoneo/Ottimale (in fig. 1.24 si è riportato come esempio il solo modello Assenza/Presenza).

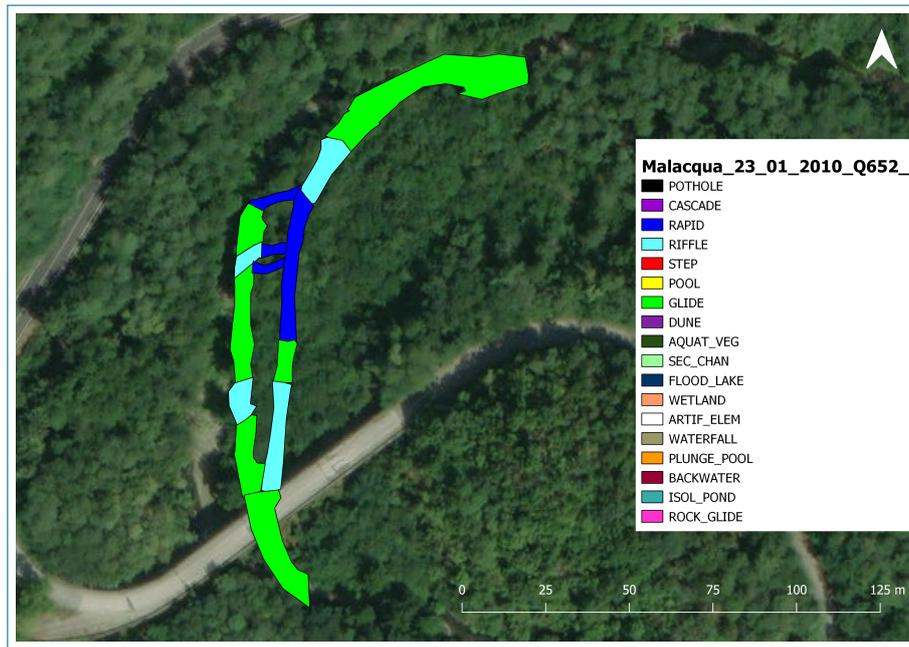


Figura 1.21: Output Simstream-Web. Shapefile unità idro-morfologiche, Malaqua, 0.652 m³/s.

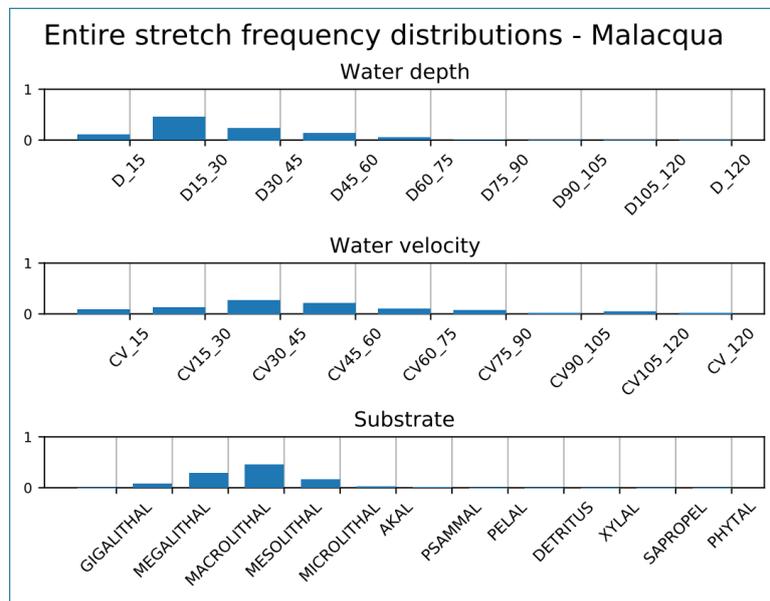


Figura 1.22: Output Simstream-Web. Distribuzioni di frequenza, Malaqua, 0.652 m³/s.



Figura 1.23: Output SimStream-Web. Shapefile mappa idoneità di habitat per vairone adulto (*Telestes muticellus*), Malacqua, $0.652 \text{ m}^3/\text{s}$.

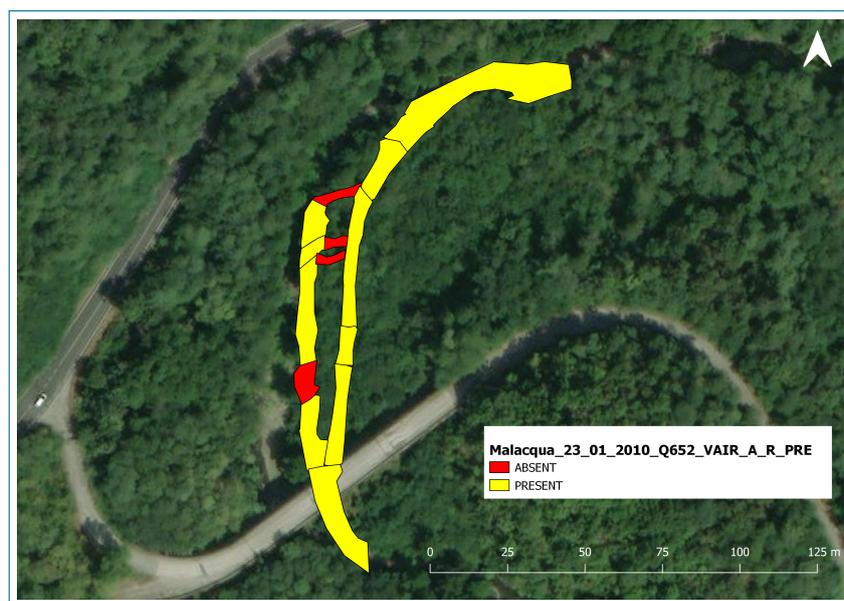
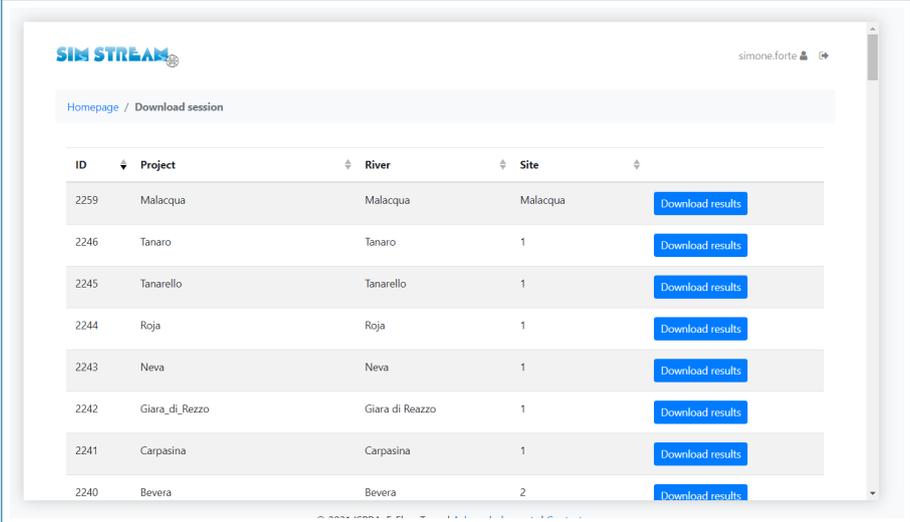


Figura 1.24: Output SimStream-Web. Shapefile mappa risultato modello biologico di Assenza/Presenza per vairone adulto (*Telestes muticellus*), Malacqua, $0.652 \text{ m}^3/\text{s}$, 23/01/2010.

Nel caso in cui siano stati inseriti almeno 3 rilievi, con portata maggiore di zero, sarà possibile ricavare anche la curva habitat-portata.

Qualora sia stata caricata anche la serie storica di portate, saranno selezionabili anche la serie temporale di habitat, la curva UCUT e l'indice IH.

Terminato l'inserimento dei dati e la selezione degli output è possibile il scaricarli dall'apposita schermata di download (figura 1.25).



ID	Project	River	Site	
2259	Malacqua	Malacqua	Malacqua	Download results
2246	Tanaro	Tanaro	1	Download results
2245	Tanarello	Tanarello	1	Download results
2244	Roja	Roja	1	Download results
2243	Neva	Neva	1	Download results
2242	Giara di Rezzo	Giara di Rezzo	1	Download results
2241	Carpasina	Carpasina	1	Download results
2240	Bevera	Bevera	2	Download results

Figura 1.25: Schermata SimStream-Web per download output.

Capitolo 2

Organizzazione Dati Idro-morfologici e Biologici

Ai fini dell'applicazione della metodologia MesoHABSIM, a partire dal 2009, sono stati raccolti una moltitudine di dati idro-morfologici e biologici che, con il presente lavoro di tesi, sono stati raccolti, formattati e organizzati in un unico database a scala nazionale che ha permesso di realizzare un quadro aggiornato dei rilevamenti eseguiti. Il dato idro-morfologico (previa elaborazione in SimStream) è stato associato a quello biologico, grazie al codice univoco HMU_CODE, permettendo la generazione di nuovi modelli di distribuzione di specie, implementabili in SimStream stesso, che ampliano il panorama di applicazione della metodologia MesoHABSIM.

Di seguito saranno illustrate le tipologie di struttura dei dati riscontrate durante l'analisi, strettamente legate alla presenza o meno di una procedura standardizzata di raccolta dato in campo, quindi alcune tipologie di elaborazioni eseguite per poter rendere il dato idro-morfologico correttamente caricabile in SimStream.

Si vuole puntualizzare, come nella sotto sezione 2.1 verrà discusso lo schema standard attuale del dato idro-morfologico e biologico raccolto in campo; nella sezione 2.2 verrà poi presentata la struttura del database in cui questi dati verranno raccolti e la loro formattazione standard.

2.1 Formato pregresso e formattazione dati per applicazione MesoHABSIM

2.1.1 Dati idro-morfologici

A partire dal 2012, anno in cui la metodologia è stata adattata al contesto italiano (Veza et al. ([15], [18], [19])), il formato di raccolta dei dati idro-morfologici ha visto un'evoluzione; solo con lo sviluppo di MapStream, è stato definito lo standard di acquisizione attuale, descritto al punto 1.1.1.1, che permette la raccolta di dati immediatamente utilizzabili in SimStream.

Per poter elaborare i dati idro-morfologici con tale software, è infatti necessario che questi rispettino il formato descritto nella sezione 1.2. Per una corretta e completa applicazione della metodologia MesoHABSIM è infatti necessario disporre di dati idro-morfologici correttamente strutturati e prontamente caricabili in SimStream; si basti pensare, ad esempio, come per realizzare l'analisi della disponibilità spazio-temporale di habitat siano necessari rilievi, eseguiti nello stesso sito a diverse condizioni di portata e di come un dato già rilevato, ma non correttamente formattato, sia di fatto inutilizzabile insieme a dati più recenti. Nel presente lavoro di tesi si sono quindi modificati i dati idro-morfologici relativi a una moltitudine di rilievi.

Occorre specificare come, tra i dati analizzati, erano spesso presenti nomenclature di unità idro-morfologiche differenti dallo standard attuale; è stata preliminarmente eseguita la loro modifica (ad esempio le UMI di tipo RUN sono state convertite in GLIDE, mentre le RUFFLE in RIFFLE).

Durante l'analisi sono stati riscontrati molteplici formati di dato grezzo; sono state quindi effettuate numerose formattazioni rilievo-specifiche sia sugli shapefile delle HMUs (.shp e .dbf) che sul file di testo delle misure puntuali di profondità, velocità e substrato.

Di seguito si riportano alcune delle problematiche più comuni per gli shapefile dei rilievi (.dbf e .shp) e nel file delle misure (.txt) e le modifiche apportate per rendere il dato idro-morfologico conforme allo standard di input di SimStream.

2.1.1.1 Modifica shapefile HMU - Tabella attributi e file delle forme

- Tabella attributi shapefile (.dbf)

In figura 2.1 si riporta l'esempio del torrente Malacqua (23 gennaio 2010); è osservabile l'aspetto della tabella attributi prima e dopo la sua formattazione allo standard richiesto; tramite il software OpenOfficeCalc, con cui è possibile manipolare file .dbf, sono stati convertiti i nomi delle unità e sono state aggiunte le colonne mancanti, avendo cura di rispettare il corretto formato dei campi, indicando la presenza/assenza delle diverse covers con i valori True/False.

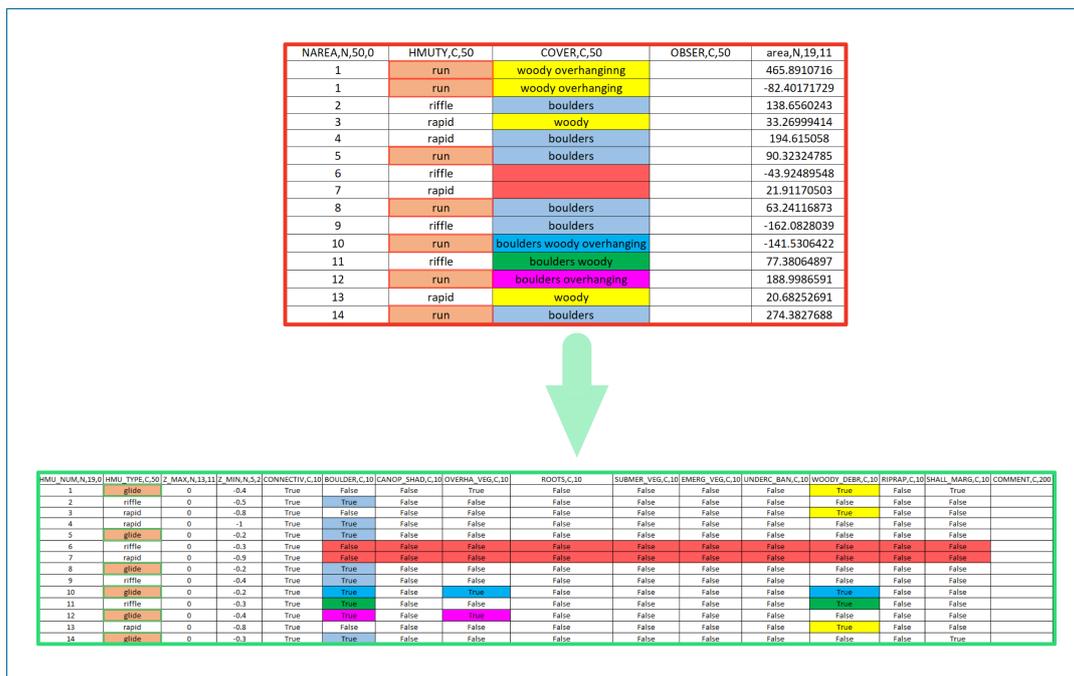


Figura 2.1: Elaborazioni su tabella attributi dello shapefile HMU.

I rilievi più recenti, eseguiti tramite software MapStream, non presentano problematiche di questo tipo.

- **File delle forme (.shp)** I poligoni degli shapefile devono essere conformi alle richieste illustrate a pagina 20; in figura 2.2 è riportato l'esempio di un poligono "self-intersecting" (a sinistra) e un poligono con la presenza di un buco interno (a destra); è stata quindi effettuata la formattazione, tramite editing dei vertici, in ArcMAP.

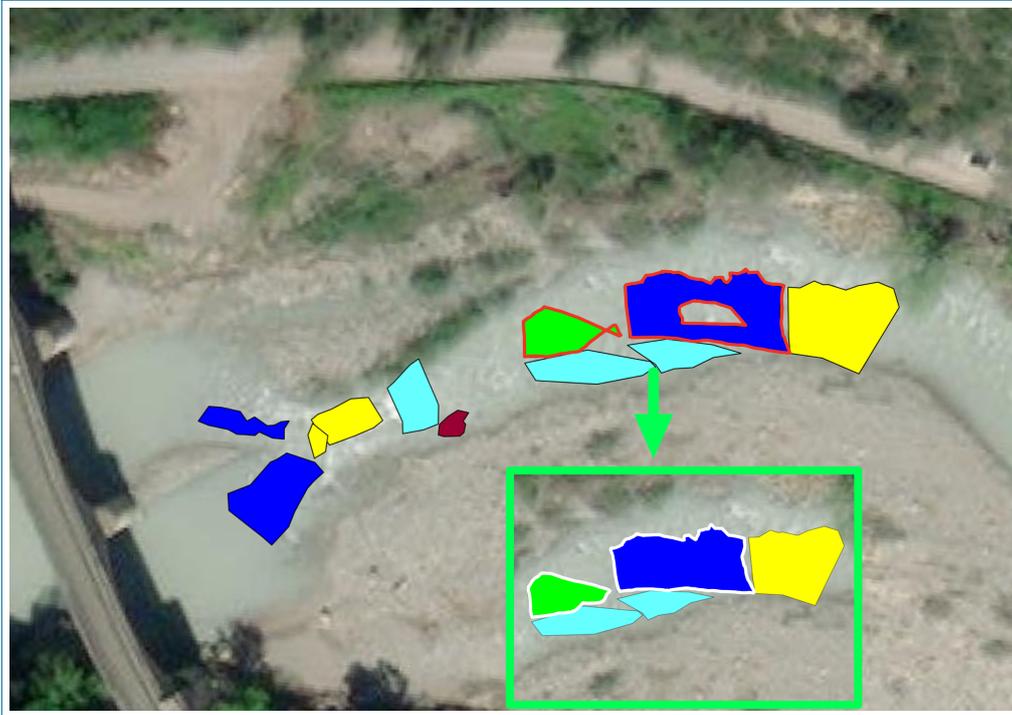


Figura 2.2: Elaborazioni su poligono dello shapefile HMU.

Occorre specificare come la geometria non conforme allo standard SimStream non sia prerogativa dei rilievi cronologicamente più distanti, bensì una problematica eventualmente sempre presente.

2.1.1.2 Modifica file di testo misure puntuali

Il file delle misure in figura 2.3 si riferisce nuovamente al rilievo effettuato sul Malacca a gennaio 2010; il file originale (a sinistra) differisce notevolmente da quello che è stato formattato (a destra); sono presenti disformità sia per l'ordine delle colonne (alcune risultano mancanti), che per le unità di misura di velocità e profondità e la nomenclatura dei substrati. Tramite il programma Excel, velocità, profondità e substrato sono state convertite conformemente a quanto riportato a pagina 20 ed è stata aggiunta la colonna "HMU_TYPE" ricavando l'attributo dalla tabella attributi del file HMU; tramite apposito script in Visual Basic for Applications (VBA) è stata associata la tipologia di unità in funzione del numero delle stesse.

In figura 2.3, e in tutto il seguito del documento, la formattazione di celle con riempimento a strisce verticali indicherà porzioni di tabelle omesse nella rappresentazione.

NPOINT	DEPTH	VEL	SUBSTRATE	NHABITAT
1	25.00	39.60	6-20	1.00
2	18.00	53.40	6-20	1.00
3	15.00	43.70	2-6	1.00
4	12.00	20.40	2-6	1.00
5	32.00	38.80	6-20	1.00
6	29.00	35.10	40-	1.00
7	48.00	27.50	40-	1.00
8	40.00	48.60	6-20	1.00
9	10.00	28.60	2-6	1.00
10	18.00	32.60	2-6	1.00
11	61.00	24.70	6-20	1.00
12	53.00	27.10	40-	1.00
13	48.00	38.30	40-	1.00
13	50.00	48.60	6-20	1.00
15	19.00	38.30	6-20	1.00
16	70.00	41.80	40-	1.00
17	48.00	42.20	40-	1.00
17	22.00	57.80	6-20	1.00
17	31.00	88.70	6-20	1.00
18	20.00	88.80	40-	1.00
19	10.00	43.30	2-6	1.00
19	45.00	87.20	40-	1.00
21	65.00	61.40	2-6	1.00
20	10.00	32.00	2-6	1.00
21	11.00	13.20	2-6	1.00
21	71.00	47.60	2-6	1.00

HMU_NUM	HMU_TYPE	PNTNUM	DEPTH	VELOCITY	SUBSTRATE
1	GLIDE	1	0.25	0.396	MESOLITHAL
1	GLIDE	2	0.18	0.534	MESOLITHAL
1	GLIDE	3	0.15	0.437	MICROLITHAL
1	GLIDE	4	0.12	0.204	MICROLITHAL
1	GLIDE	5	0.32	0.388	MESOLITHAL
1	GLIDE	6	0.29	0.351	MEGALITHAL
1	GLIDE	7	0.48	0.275	MEGALITHAL
1	GLIDE	8	0.4	0.486	MESOLITHAL
1	GLIDE	9	0.1	0.286	MICROLITHAL
1	GLIDE	10	0.18	0.326	MICROLITHAL
1	GLIDE	11	0.61	0.247	MESOLITHAL
1	GLIDE	12	0.53	0.271	MEGALITHAL
1	GLIDE	13	0.48	0.383	MEGALITHAL
1	GLIDE	14	0.5	0.486	MESOLITHAL
1	GLIDE	15	0.19	0.383	MESOLITHAL
1	GLIDE	16	0.7	0.418	MEGALITHAL
1	GLIDE	17	0.48	0.422	MEGALITHAL
1	GLIDE	18	0.22	0.578	MESOLITHAL
1	GLIDE	19	0.31	0.887	MESOLITHAL
1	GLIDE	20	0.2	0.888	MEGALITHAL
1	GLIDE	21	0.1	0.433	MICROLITHAL
1	GLIDE	22	0.45	0.872	MEGALITHAL
1	GLIDE	23	0.65	0.614	MICROLITHAL
1	GLIDE	24	0.1	0.32	MICROLITHAL
1	GLIDE	25	0.11	0.132	MICROLITHAL
1	GLIDE	26	0.71	0.476	MICROLITHAL

Figura 2.3: Elaborazioni su file di testo misure puntuali.

I rilievi eseguiti tramite software MapStream non presentano problematiche di questo tipo.

2.1.2 Dati biologici

Per i campionamenti biologici, a oggi, non esiste un software che fornisca un dato campionato con formattazione standardizzata. A titolo esemplificativo si riportano alcuni tipi di formati analizzati; le tipologie sono le più disparate. In figura 2.4 è riportato il caso di dato registrato su carta (Malacqua 2010); ad oggi il supporto cartaceo è ancora utilizzato per la raccolta in campo e successiva trascrizione a computer (esempio in figura 2.5, campionamento in Friuli Venezia Giulia, marzo 2022).

SCHEDA RILEVAMENTO FAUNA ITTICA N°1..... DEL 13/04/10

Bacino: VARA Corso d'acqua: MALACQUA
 Località: Coordinate: TRAIANO
 Lunghezza tratto: Larghezza media sezione:
 Tipologia fluviale: Temperatura acqua: 11,6°C
 pH: Conduttività elettrica:
 Ossigeno (mg/l): Ossigeno (% di saturazione):

Specie	Lunghezza (mm)	Peso (g)	Note
BARBO COMUNE	66	4	
VAIRONE	58	1	
CAVEDANO	88	5	
BARBO COMUNE	61	3	
VAIRONE	59	2	
GHIOTTO	71	5	
VAIRONE	49	0+	
GHIOTTO	65	4	
GHIOTTO	70	4	
VAIRONE	62	2	
VAIRONE	60	2 3	
GHIOTTO	65	2 3	
GHIOTTO	55	3	
VAIRONE	49	1	
VAIRONE	58	2	
VAIRONE	49	2	
VAIRONE	39	0+	
GHIOTTO	45	0+	
GHIOTTO	39	0+	
ATIPICO	150	43	00069085
<hr/>			
BARBO COMUNE	120	22	
BARBO COMUNE	81	6	
VAIRONE	64	3	
VAIRONE	63	3	
BARBO COMUNE	72	4	
BARBO COMUNE	59	2	
BARBO COMUNE	74	5	
BARBO COMUNE	61	2	
VAIRONE	37	0+	

20
PASSATA

Figura 2.4: Dato biologico registrato su supporto cartaceo.

Corso d'acqua:	Torrente Natisona	Specie	LT (cm)	1=giovani 2=adulti	Mesohabitat	Note
Data campionamento:	06/09/2021	<i>Alburnus arborella</i>	2.5	1	5_glide	
Bacino:	Isonzo	<i>Alburnus arborella</i>	3.5	1	5_glide	
Località:	Ponte San Quirino, confluenza col Torrente Alberone	<i>Alburnus arborella</i>	4.0	1	5_glide	
Comune:	Cividale del Friuli	<i>Barbus plebejus</i>	2.5	1	4_riffle	
Coordinate:	46° 6'34.25"N; 13°28'28.29"E	<i>Barbus plebejus</i>	2.5	1	5_glide	
Specie target:	Alborella, Barbo comune, Cobite comune, Ghiozzo padano	<i>Barbus plebejus</i>	2.5	1	7_pool	
Temperatura (°C):	18.2	<i>Barbus plebejus</i>	2.7	1	10_glide	
Ossigeno disciolto (mg/l):	8.8	<i>Barbus plebejus</i>	2.7	1	2_pool	
Saturazione ossigeno (%):	90	<i>Barbus plebejus</i>	2.7	1	4_riffle	
Conduttività (µS/cm):	305	<i>Barbus plebejus</i>	2.8	1	4_riffle	
Mesohabitat presenti:	Backwater, glide, pool, rapid, riffle	<i>Barbus plebejus</i>	3.0	1	1_pool	
		<i>Barbus plebejus</i>	3.0	1	10_glide	
		<i>Barbus plebejus</i>	3.0	1	4_riffle	
		<i>Barbus plebejus</i>	3.0	1	4_riffle	
		<i>Barbus plebejus</i>	3.0	1	5_glide	
		<i>Barbus plebejus</i>	3.0	1	5_glide	
		<i>Barbus plebejus</i>	3.0	1	5_glide	
		<i>Barbus plebejus</i>	3.0	1	5_glide	
		<i>Barbus plebejus</i>	3.0	1	5_glide	
		<i>Barbus plebejus</i>	3.0	1	7_pool	
		<i>Barbus plebejus</i>	3.0	1	7_pool	
		<i>Barbus plebejus</i>	3.0	1	7_pool	
		<i>Barbus plebejus</i>	3.0	1	7_pool	
		<i>Barbus plebejus</i>	3.0	1	5_glide	
		<i>Barbus plebejus</i>	3.1	1	5_glide	
		<i>Barbus plebejus</i>	3.2	1	4_riffle	
		<i>Barbus plebejus</i>	3.4	1	10_glide	
		<i>Barbus plebejus</i>	3.4	1	4_riffle	
		<i>Barbus plebejus</i>	3.5	1	10_glide	
		<i>Barbus plebejus</i>	3.5	1	10_glide	
		<i>Barbus plebejus</i>	3.5	1	10_glide	

Figura 2.5: Dato biologico registrato su supporto cartaceo e trascritto su file Excel.

In tabella 2.1 è invece riportato l'esempio di dato registrato su foglio Excel da dispositivo mobile.

Anche se le grandezze campionate sono pressoché le stesse, l'introduzione di una procedura standard per la raccolta velocizzerebbe le successive operazioni di editing e inserimento nel database descritto in seguito.

Per queste ragioni, è stata sviluppata l'app (con Microsoft Power Apps), descritta nel capitolo 3.1, per la raccolta di dati biologici in campo.

Tabella 2.1: Dato biologico registrato su supporto informatico (foglio Excel).

HMU_NUM	HMU_TYPE	FISH_NUM	FISH_TYPE	LENGTH	LIFE_STAGE	ADUNDANCE			
1	pool	1	caninus	70					
				65					
				110					
						barbo x canino	120		
							60		
						cagnetta	93	a	
						anguilla	295		
							312		
							330		
							298		
							157		
							690		
			2	rapid	1	barbo x canino	100		
							182		
							142		
122									
118									
						caninus	107		
							80		
							142		
							112		
							155		
							63		
							66		
							91		
							80		
							125		
						barbo x canino	102		
							103		
						caninus	78		
							80		
							69		
							70		
						barbo x canino	52		
						cagnetta	81	a	
				100	a				
				56	j				

2.2 Struttura Database

Per la generazione di nuovi modelli di distribuzione di specie è necessario disporre di una base di dati in cui il dato biologico sia chiaramente ed univocamente legato al dato idro-morfologico; la generazione di nuovi modelli implementabili in SimStream permetterà di ottenere mappe georeferenziate di presenza e di abbondanza per un numero sempre maggiore di specie. Di seguito verrà illustrata l'organizzazione stabilita per i dati e gli strumenti creati per la loro manipolazione e accorpamento. L'idea di base è quella di popolare, grazie ad un apposito tool di inserimento dati ("Hydro Bio Tool"), due tabelle in due fogli (sheets) della cartella Excel, "MesoHABSIM DataBase", contenenti rispettivamente i dati biologici e idro-morfologici, messe in relazione dal codice univoco HMU_CODE, per poterli estrarre al fine dell'implementazione dell'algoritmo Random Forest. Gli

strumenti sono attivabili da apposite cartelle Excel.xlsm (con attivazione di macro), e sono stati creati attraverso Userform, gestite dal linguaggio VBA, con cui è resa possibile l'interazione con l'utente e l'avvio di macro apposite per la gestione e il trattamento di dati.

2.2.1 Schema generale "MesoHABSIM DataBase"

Nell'allegato [A1-SchemaDB](#) sono osservabili le relazioni tra gli strumenti creati che hanno permesso di raggiungere la finalità desiderate. I file sopra citati sono organizzati in un'unica cartella "MESO DB", contenente i seguenti file:

- **"Hydro Bio Tool.xlsm"**: cartella automatizzata Excel che contiene al suo interno l'omonimo tool per la validazione dei dati di Input, la generazione di schede di output sito specifiche e il popolamento delle tabelle dati contenute nel file "MesoHABSIM DataBase.xlsm".
- **"MesoHABSIM DataBase.xlsm"**: cartella automatizzata Excel in cui sono raccolti, grazie allo strumento presente in "Hydro Bio Tool" tutti i dati, idro-morfologici e biologici, formattati allo standard prescelto. Sono inoltre presenti numerosi strumenti per l'estrazione e la manipolazione dei dati.
- **"RELIEFS DATA"**: Cartella contenente le schede di output generate con lo strumento "Hydro Bio Tool".
- **"SPECIES DATA"**: Cartella contenente i file per la generazione di modelli biologici ottenuti con lo strumento di estrazione implementato nel "MesoHABSIM DataBase".

Nell'allegato [A1-SchemaDB](#) è osservabile come, all'apertura della cartella Excel contenente **"Hydro Bio Tool"** (*a*), verrà avviata in automatico la prima schermata dello strumento di inserimento dati (*b*), che permetterà di inserire le generalità del rilievo (come regione, nome del fiume, coordinate, portata e data). Due schede standardizzate di input (*c*), il cui formato scelto verrà descritto nella sezione [2.3](#), verranno selezionate e manipolate dall'apposito strumento **"Hydro Bio Tool"** creato, previa validazione (*b*). Dalla selezione tramite la finestra (*d*), sarà possibile l'estrazione delle schede rilievo-specifiche (in formato .xlsm), **"Hydro Sheet River *m3s ddmmyyyy"**, **"Bio Sheet River *m3s ddmmyyyy"**, **"Resume AJ *m3s ddmmyyyy"** e **"Resume All *m3s ddmmyyyy"**, all'interno della **"RELIEFS DATA"** verrà creata automaticamente cartella **"RIVER *m3s ddmmyyyy"** contenente i file sopra citati (*e*).

Il tool di inserimento dati popolerà, con i dati idro-morfologici e biologici esportati

singolarmente nella cartella "**RIVER *m3s ddmmyyyy**", i rispettivi sheets Excel, "**Bio Data**" e "**Hydro Data**", del file "**MesoHABSIM DataBase**" (*f*), Da quest'ultimo sarà possibile estrarre, tramite apposito tool (*g*), schede specie-specifiche "**LifeStage Specie yyymmdd**" (*h*), per la generazione dei modelli di distribuzione di specie; queste saranno salvate nella cartella "**Specie**" creata in automatico (se non ancora esistente).

Di seguito ci si riferirà alle schede di input dello strumento di inserimento dati come **scheda dato grezzo biologico** e **scheda dato grezzo idro-morfologico**., allo strumento "Hydro Bio Tool" come **tool/strumento/maschera di inserimento dati** e ai dati idro-morfologici e biologici contenuti nei rispettivi sheets della cartella "MesoHABSIM DataBase" come "**Database idro-morfologico**" e "**Database biologico**".

Il funzionamento dettagliato dei diversi strumenti verrà illustrato a partire dalla sezione 3.2. A partire dalla prossima sezione saranno invece descritte le diverse tipologie di input e output, per i diversi strumenti creati, al fine di comprendere la logica di funzionamento del sistema di raccolta e organizzazione dati.

2.3 Tipologia dati di input

Per l'esecuzione delle precedenti operazioni è stato necessario stabilire un formato standardizzato per i dati grezzi di input alla maschera di inserimento. È importante sottolineare come questi debbano contenere tutte le informazioni necessarie alla generazione di modelli biologici.

2.3.1 Input dato grezzo idro-morfologico

Per il dato idro-morfologico si è scelto di utilizzare come input il file .txt generato da SimStream, di cui un esempio è riportato nell'allegato [A2-InputIdro](#) (caso rilievo Roja, 10/03/2022), in quanto contiene al suo interno tutte le informazioni relative ai descrittori fisici richiesti per ciascuna unità campionata. In tabella 2.2 sono riportate le grandezze presenti nel suddetto file.

Tabella 2.2: Formato scheda dati di input grezzi idro-morfologici

CARATT. GENERALI:	HMU_NUM	HMU_TYPE	AREA	SLOPE	SDVEL	FROUDE						
ATTRIBUTI HMU:	CONNECTIV	BOULDER	CANOP_SHAD	OVERHA_VEG	ROOTS	SUBMER_VEG	EMERG_VEG	UNDERC_BAN	WOODY_DEBR	RIPRAP	SHALL_MARG	
TIPOLOGIA HMU:	POTHOLE	CASCADE	RAPID	RIFFLE	STEP	POOL	GLIDE	DUNE	AQUAT_VEG			
	SEC_CHAN	FLOOD_LAKE	WETLAND	ARTIF_ELEM	WATERFALL	PLUNGE_POOL	BACKWATER	ROCK_GLIDE	ISOL_POND			
FREQ. PROFONDITA':	D_15	D15_30	D30_45	D45_60	D60_75	D75_90	D90_105	D105_120	D_120			
FREQ. VELOCITA':	CV_15	CV15_30	CV30_45	CV45_60	CV60_75	CV75_90	CV90_105	CV105_120	CV_120			
FREQ. SUBSTRATO:	GIGALITHAL	MEGALITHAL	MACROLITHAL	MESOLITHAL	MICROLITHAL	AKAL	PSAMMAL	PELAL	DETRITUS	XYLAL	SAPROPEL	PHYTAL

Nell'inserimento di questo dato nell'apposito tool di inserimento, non sono state poste particolari restrizioni, se non quella di formato .txt, essendo lui stesso l'output di un software (SimStream-Web) che effettua una validazione, come descritto dettagliatamente nella sezione 1.2.

2.3.2 Input dato grezzo biologico

Per il dato biologico si è scelto di utilizzare come standard il formato riportato nell'allegato [A3-InputBio](#) (caso campionamento Roja, 10/03/2022).

Il dato biologico grezzo sarà sempre inseribile purché vengano rispettati il nome e l'ordine delle colonne; gli attributi di ciascuna osservazione (es. stadio vitale o nome della specie) potranno essere modificati grazie al convertitore implementato all'interno di "Hydro Bio Tool", attivato in automatico nel caso in cui vengano riscontrate discrepanze rispetto ai valori accettati; la possibilità di modifica del dato grezzo direttamente dalla suddetta maschera di modifica consentirà una maggiore versatilità nella tipologia di formato utilizzato durante il campionamento.

2.4 Esportazione output con "Hydro Bio Tool"

Inseriti i dati biologici e idro-morfologici all'interno dello strumento di inserimento, e verificata la corrispondenza tra le unità, saranno attivate le check box in figura 2.6, da cui saranno selezionabili le tipologie di output desiderate.

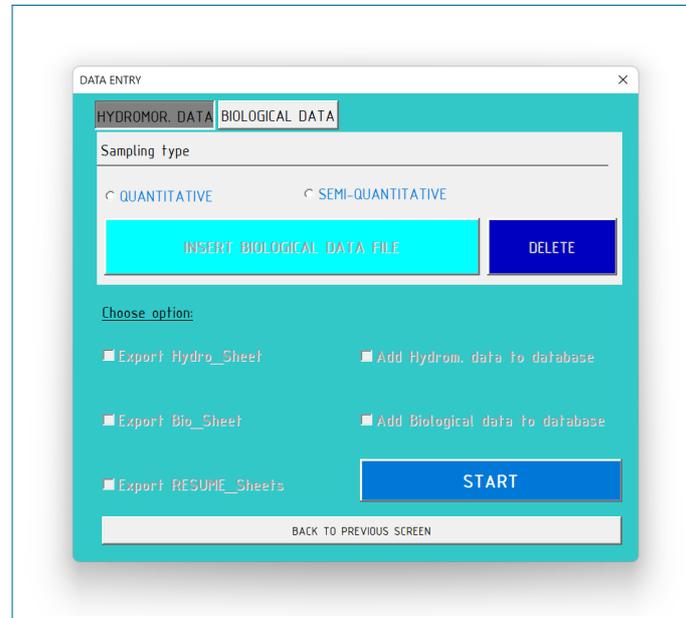


Figura 2.6: Check box per selezione output ("Hydro Bio Tool")

È da specificare come nel database saranno inserite soltanto le unità per cui è stato eseguito sia il campionamento biologico che il rilievo idro-morfologico. I modelli di distribuzione di specie saranno infatti costruiti sulla base della presenza/assenza e presenza/abbondanza delle sole unità campionate; l'inserimento di unità in cui non è stato effettuato il campionamento biologico genererebbe casi aggiuntivi di assenza di specie, distorcendo le previsioni dei modelli costruiti con un data set non veritiero; la descrizione del codice "HMU_CODE" utilizzato per la caratterizzazione e il confronto delle unità sarà descritto nella sotto sezione 2.5. Il tool di inserimento rimuoverà, dal dato idro-morfologico, tutte quelle unità in cui tale codice non è presente nel dato biologico. A titolo esemplificativo si riporta il caso del torrente Argentina dove il rilievo idro-morfologico è stato effettuato per 37 HMU, mentre il campionamento biologico per sole 20 unità; nel database (e nelle schede sito-specifiche per il dato idro-morfologico) verranno inserite le sole 20 unità che hanno HMU_CODE in comune con il dato biologico.

2.4.1 Formato schede di output

Di seguito sono riportati alcuni esempi delle tabelle selezionabili come output. Si rimarca il fatto di come queste contengano tutti i dati fisici e biologici necessari messi in relazione dall'HMU_CODE.

"Hydro_Sheet_River_*m3s_ddmmyyyy.xlsm"

In tabella 2.3 è riportata la scheda di output idro-morfologico per il fiume Sangone. Associando ai dati di output di SimStream (celle evidenziate in azzurro), le caratteristiche del sito in esame (celle evidenziate in verde), è generato il codice univoco dell'unità (celle evidenziate in arancione).

Tabella 2.3: Esempio di Hydro_Sheet_River_*m3s_ddmmyyyy

REGION	RIVER	UTM_E	UTM_N	DATE	DISCHARGE	HMU_CODE	HMU_NUM	HMU_TYPE	OUTPUT SIMSTREAM
PIEMONTE	SANGONE	378512	4987805	12/07/2021	0.13	1-0.13-378512-4987805-12/07/2021	1	GUIDE	...
PIEMONTE	SANGONE	378512	4987805	12/07/2021	0.13	2-0.13-378512-4987805-12/07/2021	2	RIFFLE	...
PIEMONTE	SANGONE	378512	4987805	12/07/2021	0.13	3-0.13-378512-4987805-12/07/2021	3	POOL	...
PIEMONTE	SANGONE	378512	4987805	12/07/2021	0.13	4-0.13-378512-4987805-12/07/2021	4	GUIDE	...

Bio_Sheet_River_*m3s_ddmmyyyy

In tabella 2.4 è riportata la scheda di output biologico per il fiume Sangone. La scheda risulterà ordinata per numero di HMU crescente, in modo da rispettare lo stesso ordinamento della scheda precedente (output ordinato automaticamente da SimStream nello stesso modo).

Tabella 2.4: Esempio di Bio_Sheet_River_*m3s_ddmmyyyy

RIVER	HMU_CODE	SAMPLING_TYPE	HMU_NUM	HMU_TYPE	FISH_NUM	SPECIES_LAT	LENGTH_cm	WEIGHT_g	LIFE_STAGE	ABUNDANCE
SANGONE	1-0.13-378512-4987805-12/07/2021	S	1	GUIDE	1	Barbus caninus			A	
SANGONE	1-0.13-378512-4987805-12/07/2021	S	1	GUIDE	2	Barbus plebejus			A	
SANGONE	1-0.13-378512-4987805-12/07/2021	S	1	GUIDE	3	Padogobius martensii			J	X
SANGONE	1-0.13-378512-4987805-12/07/2021	S	1	GUIDE	4	Padogobius martensii			A	X
SANGONE	1-0.13-378512-4987805-12/07/2021	S	1	GUIDE	5	Phoxinus phoxinus			A	
SANGONE	1-0.13-378512-4987805-12/07/2021	S	1	GUIDE	6	Squalius cephalus			A	
SANGONE	1-0.13-378512-4987805-12/07/2021	S	1	GUIDE	7	Squalius cephalus			J	
SANGONE	1-0.13-378512-4987805-12/07/2021	S	1	GUIDE	8	Telestes muticellus			J	X
SANGONE	1-0.13-378512-4987805-12/07/2021	S	1	GUIDE	9	Telestes muticellus			A	X
SANGONE	2-0.13-378512-4987805-12/07/2021	S	2	RIFFLE	1	Barbus caninus			A	X
SANGONE	2-0.13-378512-4987805-12/07/2021	S	2	RIFFLE	2	Barbus caninus			J	X
SANGONE	2-0.13-378512-4987805-12/07/2021	S	2	RIFFLE	3	Barbus plebejus			A	
SANGONE	2-0.13-378512-4987805-12/07/2021	S	2	RIFFLE	4	Phoxinus phoxinus			A	
SANGONE	2-0.13-378512-4987805-12/07/2021	S	2	RIFFLE	5	Telestes muticellus			J	
SANGONE	2-0.13-378512-4987805-12/07/2021	S	2	RIFFLE	6	Telestes muticellus			A	X
SANGONE	3-0.13-378512-4987805-12/07/2021	S	3	POOL	1	Barbus plebejus			A	X
SANGONE	3-0.13-378512-4987805-12/07/2021	S	3	POOL	2	Barbus plebejus			J	
SANGONE	3-0.13-378512-4987805-12/07/2021	S	3	POOL	3	Padogobius martensii			A	
SANGONE	3-0.13-378512-4987805-12/07/2021	S	3	POOL	4	Salmo trutta			J	
SANGONE	3-0.13-378512-4987805-12/07/2021	S	3	POOL	5	Squalius cephalus			A	X
SANGONE	3-0.13-378512-4987805-12/07/2021	S	3	POOL	6	Squalius cephalus			J	
SANGONE	3-0.13-378512-4987805-12/07/2021	S	3	POOL	7	Telestes muticellus			A	
SANGONE	3-0.13-378512-4987805-12/07/2021	S	3	POOL	8	Telestes muticellus			J	
SANGONE	4-0.13-378512-4987805-12/07/2021	S	4	GUIDE	1	Telestes muticellus			J	X
SANGONE	4-0.13-378512-4987805-12/07/2021	S	4	GUIDE	2	Phoxinus phoxinus			J	

Si può notare come, al dato biologico grezzo di input, siano state aggiunte le colonne evidenziate in azzurro: "SAMPLING_TYPE" è rappresentativa del tipo di campionamento svolto (Q – Quantitativo, S – Semi quantitativo), "RIVER" indica il nome del fiume mentre "FISH_NUM" è un numero crescente utile per la visualizzazione del numero di esemplari campionati per ciascuna unità. Come riportato in precedenza, la presenza dell'HMU_CODE permetterà di relazionare univocamente i descrittori fisici della tabella 2.3 e le informazioni biologiche della 2.4; in figura 2.7 è rappresentata, tramite associazione cromatica, la relazione tra i suddetti codici per le tabelle sopra citate.

HMU_CODE_IDRO	HMU_CODE_BIO
1-0.13-378512-4987805-12/07/2021	1-0.13-378512-4987805-12/07/2021
2-0.13-378512-4987805-12/07/2021	1-0.13-378512-4987805-12/07/2021
3-0.13-378512-4987805-12/07/2021	1-0.13-378512-4987805-12/07/2021
4-0.13-378512-4987805-12/07/2021	1-0.13-378512-4987805-12/07/2021
	1-0.13-378512-4987805-12/07/2021
	1-0.13-378512-4987805-12/07/2021
	1-0.13-378512-4987805-12/07/2021
	1-0.13-378512-4987805-12/07/2021
	1-0.13-378512-4987805-12/07/2021
	2-0.13-378512-4987805-12/07/2021
	2-0.13-378512-4987805-12/07/2021
	2-0.13-378512-4987805-12/07/2021
	2-0.13-378512-4987805-12/07/2021
	2-0.13-378512-4987805-12/07/2021
	3-0.13-378512-4987805-12/07/2021
	3-0.13-378512-4987805-12/07/2021
	3-0.13-378512-4987805-12/07/2021
	3-0.13-378512-4987805-12/07/2021
	3-0.13-378512-4987805-12/07/2021
	3-0.13-378512-4987805-12/07/2021
	3-0.13-378512-4987805-12/07/2021
	3-0.13-378512-4987805-12/07/2021
	4-0.13-378512-4987805-12/07/2021
	4-0.13-378512-4987805-12/07/2021

Figura 2.7: Associazione cromatica HMU CODE dato biologico e idro-morfologico.

Resume_AJ_River_*m3s_ddmmyyyy e Resume_All_River_*m3s_ddmmyyyy

L'output in tabella 2.5, per ogni unità campionata, specie trovata e stadio vitale, riporterà:

- 0: se specie ASSENTE nell'unità.
- 1: se specie PRESENTE nell'unità.
- 2: se specie ABBONDANTE nell'unità.

Tabella 2.5: Esempio di Resume_AJ_*m3s_ddmmyyyy

ADULT	HMU_NUM	HMU_TYPE	Barbus caninus	Barbus plebejus	Padogobius martensii	Phoxinus phoxinus	Salmo trutta	Squalius cephalus	Telestes muticellus
	1	GLIDE	1	1	2	1	0	1	2
	2	RIFFLE	2	1	0	1	0	0	2
	3	POOL	0	2	1	0	0	2	1
	4	GLIDE	0	0	0	0	0	0	0

JUVENILE	HMU_NUM	HMU_TYPE	Barbus caninus	Barbus plebejus	Padogobius martensii	Phoxinus phoxinus	Salmo trutta	Squalius cephalus	Telestes muticellus
	1	GLIDE	0	0	2	0	0	1	2
	2	RIFFLE	2	0	0	0	0	0	1
	3	POOL	0	1	0	0	1	1	1
	4	GLIDE	0	0	0	1	0	0	2

Se, per nuovi rilievi inseriti, non fosse stata ancora stabilita la soglia tra esemplari adulti e giovani, l'output precedente risulterebbe come in tabella 2.6; per avere comunque un'indicazione di presenza (1) e assenza (0) delle diverse specie è stata aggiunta la scheda "Resume_All_*m3s_ddmmyyyy" (tab. 2.7).

Tabella 2.6: Esempio di Resume_AJ_*m3s_ddmmyyyy per rilievo biologico senza soglia A/J.

ADULT	HMU_NUM	HMU_TYPE	Barbus caninus	Barbus plebejus	Padogobius martensii	Phoxinus phoxinus	Salmo trutta	Squalius cephalus	Telestes muticellus
	1	GLIDE	0	0	0	0	0	0	0
	2	RIFFLE	0	0	0	0	0	0	0
	3	POOL	0	0	0	0	0	0	0
	4	GLIDE	0	0	0	0	0	0	0

JUVENILE	HMU_NUM	HMU_TYPE	Barbus caninus	Barbus plebejus	Padogobius martensii	Phoxinus phoxinus	Salmo trutta	Squalius cephalus	Telestes muticellus
	1	GLIDE	0	0	0	0	0	0	0
	2	RIFFLE	0	0	0	0	0	0	0
	3	POOL	0	0	0	0	0	0	0
	4	GLIDE	0	0	0	0	0	0	0

Tabella 2.7: Esempio di Resume_All_*m3s_ddmmyyyy.

ADULT & JUVENILE	HMU_NUM	HMU_TYPE	Barbus caninus	Barbus plebejus	Padogobius martensii	Phoxinus phoxinus	Salmo trutta	Squalius cephalus	Telestes muticellus
	1	GLIDE	1	1	1	1	0	1	1
	2	RIFFLE	1	1	0	1	0	0	1
	3	POOL	0	1	1	0	1	1	1
	4	GLIDE	0	0	0	1	0	0	1

2.5 Codici univoci per il confronto di HMUs

Per identificare univocamente ciascuna unità presente nel database, è stato introdotto il codice **"HMU_CODE"** che verrà utilizzato sia per identificare univocamente ogni unità, che per le motivazioni descritte a pagina 38, è infatti necessario valutare la corrispondenza tra le unità campionate in entrambe le tipologie di rilievo. In figura 2.8 è osservabile la composizione del codice; ciascuno di questi identificherà "quell'unità (HMU_NUM), campionata nelle specifiche condizioni di deflusso (DISCHARGE), in quel sito (UTM_E e UTM_N), quello specifico giorno (DATE)". Il codice univoco è generato automaticamente nell'Hydro Bio Tool, una volta validati i dati in ingresso, tramite l'associazione dell'informazione contenuta nei file grezzi (HMU_NUM) e i valori delle generalità del sito inseriti manualmente dall'utente.

HMU_CODE				
1-0.39-395643-5052696-22/11/2010				
HMU_NUM	DISCHARGE_m3/s	UTM_E	UTM_N	DATE
1	0.39	395643	5052696	22/11/2010

Figura 2.8: Codice identificativo dell'unità "HMU_CODE".

Per permettere alcune operazioni specifiche sui dati è stato necessario introdurre anche il codice **"CODE"**. Come osservabile in 2.9, trattandosi dell'HMU_CODE deturpato dell'HMU_NUM, questo sarà univoco per il campionamento ed identificherà tutte le unità campionate in un determinato luogo, lo stesso giorno e nelle stesse condizioni di portata.

CODE			
0.39-395643-5052696-22/11/2010			
DISCHARGE_m3/s	UTM_E	UTM_N	DATE
0.39	395643	5052696	22/11/2010

Figura 2.9: Codice identificativo del sito di campionamento "CODE".

2.6 Componenti "MesoHABSIM DataBase"

2.6.1 Sheets DataBase

I dati opportunamente formattati dallo strumento descritto in precedenza saranno raccolti nella cartella Excel "MesoHABSIM DataBase.xlsm" ed organizzati negli sheets in figura 2.10.



Figura 2.10: Sheets cartella di lavoro "MesoHABSIM DataBase".

- "List CODE": sheet contenente l'elenco aggiornato (estratto riportato in tab. 2.8) dei rilievi caricati nel database.
- "Hydro DATA": sheet contenente la raccolta di tutti i dati idro-morfologici formattati (tab. 2.9).
- "Bio DATA": sheet contenente la raccolta di tutti i dati biologici formattati (tab. 2.10). È da specificare come l'inserimento della colonna RIVER, HMU_NUM e HMU_TYPE sia ridondante in quanto insita nell'HMU_CODE; a fronte di una maggiore chiarezza durante la consultazione si è comunque scelto di mantenerle.
- "Species Extraction": sheet ausiliario per l'estrazione del file in tabella 2.11 e 2.12, utilizzabile come input per la generazione dei modelli biologici; saranno generati due output separati per lo stadio vitale adulto e giovane.

Tabella 2.11: Tabella estrazione sanguinerola (*Phoxinus phoxinus*) adulta (estratto).

RIVER	HMU_CODE	HMU_NUM	HMU_TYPE	OUTPUT SIMSTREAM	SPECIES_LAT	LIFE_STAGE	Adu_ASS0_PRE1_ABB2	Adu_ASS0_PRE1	Adu_PRE0_ABB1
CHISOLA	1-1-382698-4978317-26/05/2021	1	GLIDE	...	Phoxinus phoxinus	A	1	1	0
CHISOLA	2-1-382698-4978317-26/05/2021	2	RIFFLE	...	No Phoxinus phoxinus found		0	0	Not found
TREBBIA	6-1.73-546860-4980443-02/07/2020	6	POOL	...	Phoxinus phoxinus	A	1	1	NOT_AV
TREBBIA	7-1.73-546860-4980443-02/07/2020	7	GLIDE	...	Phoxinus phoxinus	A	1	1	NOT_AV
TREBBIA	8-1.73-546860-4980443-02/07/2020	8	RIFFLE	...	No Phoxinus phoxinus found		0	0	Not found
TREBBIA	9-1.73-546860-4980443-02/07/2020	9	BACKWATER	...	Phoxinus phoxinus	A	1	1	NOT_AV
TREBBIA	10-1.73-546860-4980443-02/07/2020	10	GLIDE	...	Phoxinus phoxinus	A	1	1	NOT_AV

Tabella 2.12: Tabella estrazione sanguinerola (*Phoxinus phoxinus*) giovane (estratto).

RIVER	HMU_CODE	HMU_NUM	HMU_TYPE	OUTPUT SIMSTREAM	SPECIES_LAT	LIFE_STAGE	Juv_ASS0_PRE1_ABB2	Juv_ASS0_PRE1	Juv_PRE0_ABB1
NATISONE	1-0.422-382142-5107396-06/09/2021	1	POOL	...	Phoxinus phoxinus	J	1	1	0
NATISONE	2-0.422-382142-5107396-06/09/2021	2	POOL	...	Phoxinus phoxinus	J	1	1	0
TREBBIA	7-1.73-546860-4980443-02/07/2020	7	GLIDE	...	Phoxinus phoxinus	J	1	1	NOT_AV
TREBBIA	8-1.73-546860-4980443-02/07/2020	8	RIFFLE	...	Phoxinus phoxinus	J	1	1	NOT_AV
TREBBIA	9-1.73-546860-4980443-02/07/2020	9	BACKWATER	...	Phoxinus phoxinus	J	1	1	NOT_AV
TREBBIA	10-1.73-546860-4980443-02/07/2020	10	GLIDE	...	Phoxinus phoxinus	J	1	1	NOT_AV

2.6.2 Tools aggiuntivi

Sono state implementate delle funzionalità specifiche per il DataBase MesoHABSIM, accessibili tramite l'apposito pulsante presente nella barra di accesso rapido del file Excel (fig. 2.11).

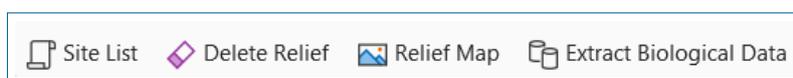


Figura 2.11: Strumenti per la gestione del Database MesoHABSIM.

- **"Site List"**: genera la lista dei rilievi presenti nel database segnalando l'eventuale assenza del dato biologico o idro-morfologico
- **"Delete Relief"**: permette la cancellazione di alcuni siti dal database sulla base del codice CODE.
- **"Relief Map"**: genera una mappa rappresentante la distribuzione spaziale dei rilievi.
- **"Extract Biological Data"**: tool necessario per l'estrazione dei file per la generazione di modelli biologici.

I tools sopra citati sono descritti accuratamente alla sezione 3.3.

2.7 Modello Random Forest per Lampreda padana (Lampetra zanandreae)

L'organizzazione dei dati nell'apposito database agevolerà la generazione di nuovi modelli di distribuzione di specie; nel presente lavoro è stato generato il primo modello quantitativo di assenza/presenza della lampreda padana (Lampetra zanandreae) allo stadio vitale giovanile (ammocete). Grazie allo strumento descritto al 3.3.4, è stato possibile estrarre i dati relativi a tutte le unità idro-morfologiche di quei rilievi in cui è stata individuata la presenza dell'ammocete in almeno una HMU.

2.7.1 Area di studio

Per la generazione del modello sono stati considerati 5 tratti fluviali di 4 fiumi piemontesi, all'interno dei quali sono state campionate 42 HMUs attraverso il software MapStream. La mappa dei siti di riferimento è riportata in figura 2.12.

2.7.2 Configurazione modello RF

Per la generazione del modello Random Forest di presenza per l'ammocete di lampreda padana è stato utilizzato il pacchetto "randomForest", implementato in R, che opera attraverso l'approccio top-down greedy della divisione ricorsiva binaria ([4]). Partendo dalla parte superiore dell'albero (dataset di partenza) vengono prodotti due rami che suddividono lo spazio dei predittori in due nuove regioni (approccio top-down), queste divengono a loro volta classi da cui verranno generati 2 rami figli per ognuna; la divisione si arresta quando al nodo terminale sono associate il numero minimo di osservazioni prestabilito (parametro nodesize). Ogni nuova divisione verrà effettuata valutando l'ottimo in quel preciso punto (approccio greedy), minimizzando l'indice di Gini, il cui valore è inversamente proporzionale alla purezza di un nodo. L'errore di test del modello è stato stimato attraverso l'OOB-Error. Dall'analisi della distribuzione della classe di presenza/assenza, come osservabile in figura 2.13, è evidenziato un non bilanciamento tra gli elementi della classe; la frequenza di presenze della lampreda risulta infatti pari a 0.26 (prevalenza).

Come specificato in [14], la presenza all'interno del dataset di un numero non proporzionale di istanze, appartenenti a classi diverse, genera una distorsione nei risultati con migliori performance nella predizione della classe maggioritaria. Inoltre l'accuratezza, parametro con cui viene valutata la predizione del modello, non sarebbe una scelta adatta (un modello, costruito su un data set contenente l'1% di presenze e che predicesse tutte le istanze come assenze, avrebbe un'accuratezza

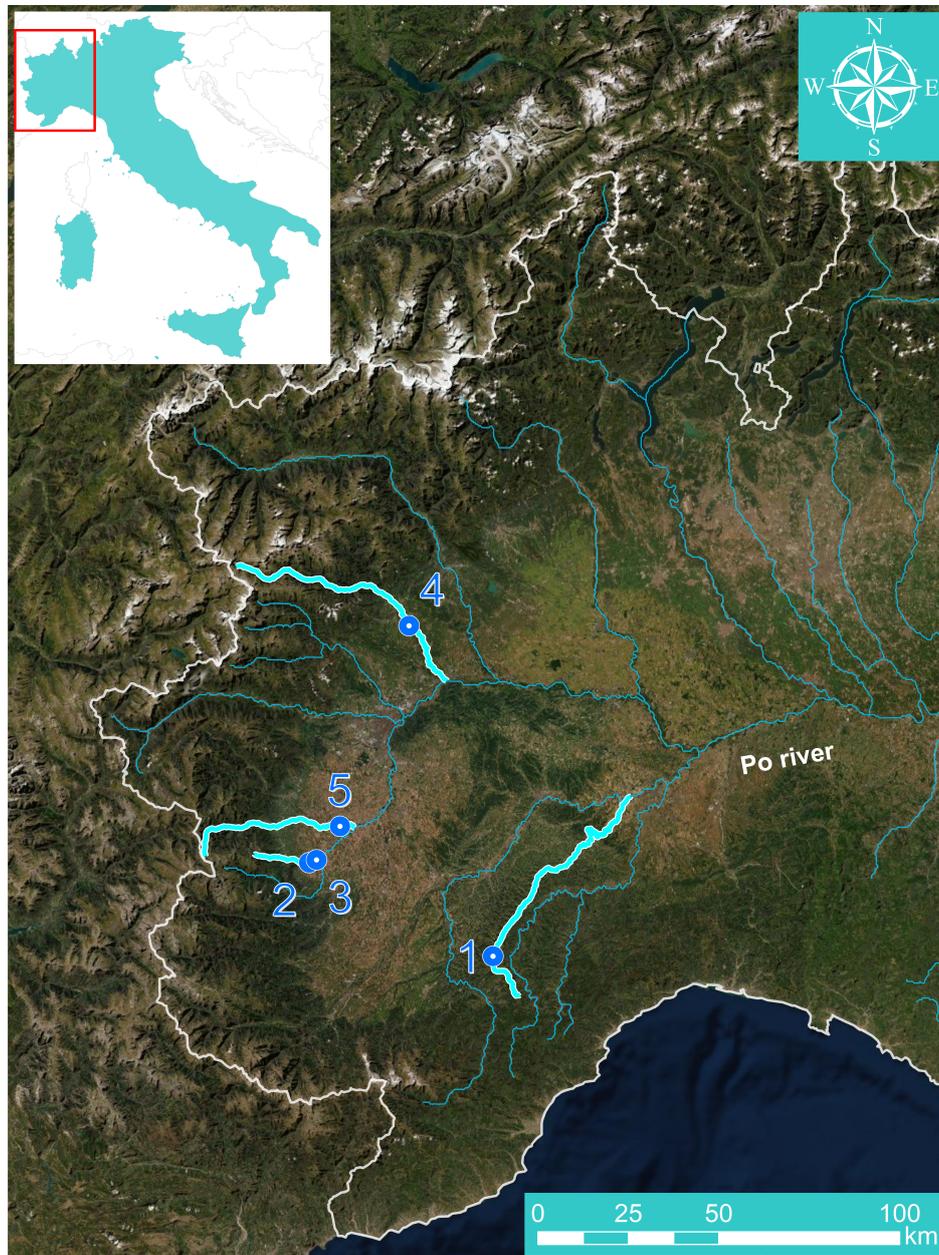


Figura 2.12: Siti analizzati per la generazione del modello statistico di distribuzione della Lampetra zanandreai allo stadio vitale giovanile. I fiumi analizzati sono i seguenti: 1 - Belbo, 2/3 - Ghiandone, 4 - Orco, 5 - Pellice.

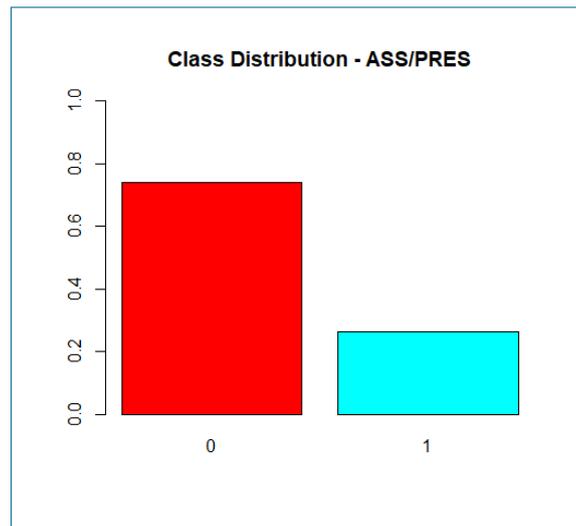


Figura 2.13: Valori classe assenza/presenza pre oversampling.

del 99%). Per la generazione del modello in analisi è stata adottata la tecnica dell'oversampling (sovra-campionamento random del data set di addestramento), al fine di ottenere una più alta prevalenza e ridurre le problematiche sopra citate. Come riportato in [12], il tuning dei parametri può incrementare le performance predittive dell'algoritmo; nella generazione del modello è stato quindi implementato per i seguenti iperparametri:

- **n_{tree}**: numero di alberi della foresta.
- **m_{try}**: numero di variabili estratte casualmente su cui viene valutato ogni split.
- **s_{ampsize}**: numero di istanze estratte per l'addestramento di ogni albero della foresta.

Come numero minimo di osservazioni al nodo terminale (`nodesize`) si è scelto di utilizzare il valore 1 (è infatti noto come questo parametro influenzi debolmente le performance del modello ([12]):

Al fine di minimizzare l'OOB-Error è stata eseguita, tramite algoritmo BORUTA ([5]), implementato in R, la selezione delle variabili, in ordine di importanza, che generassero il miglior modello parsimonioso possibile. Scelti i migliori predittori, è stata valutata la correlazione tra le variabili, attraverso la matrice di correlazione di Spearman per evitare così la presenza di variabili altamente correlate (coeff. di correlazione di Spearman > 0.7).

Il modello parsimonioso, costruito con i predittori estratti dall'algoritmo sopra citato, è stato rappresentato graficamente con le curve PDP (Partial Dependence

Plots,[2]) che rappresentano la relazione parziale tra il singolo descrittore ambientale e la probabilità di presenza, espressa in termini di funzione logit, che può assumere valori positivi o negativi, permettendo un'immediata analisi del trend della funzione, rispetto al suo valore numerico.

Capitolo 3

Raccolta, Elaborazione ed Immagazzinamento Dati

3.1 App "Bio Sampling" per raccolta dati biologici in campo.

Al fine di velocizzare e standardizzare la raccolta del dato biologico in campo, è stata sviluppata l'App "BIO SAMPLING", utilizzabile su dispositivo mobile e grazie alla quale il dato raccolto viene periodicamente salvato sul cloud (One Drive). Per la sua generazione è stata utilizzata la piattaforma di sviluppo low-code Power Apps, prodotta da Microsoft.

3.1.1 Inserimento caratteristiche generali sito di campionamento

In figura 3.1 è osservabile la prima schermata dell'app, da cui sarà possibile inserire le caratteristiche generali del rilievo quali: regione, nome del fiume, portata (m^3/s), data e tipologia di campionamento.

Come verrà specificato nella sotto sezione 3.1.2, il salvataggio del dato è influenzato dalla presenza di rete internet; dal momento che i campionamenti si svolgono spesso in condizioni di sua assenza, è stata inserita una visualizzazione immediata dello stato della rete tramite la variazione della proprietà fill (colore di riempimento) dell'oggetto circle; come osservabile in fig. 3.2, in condizioni di presenza di rete, l'icona risulterà blu, mentre apparirà rossa in caso di assenza.

Nella seguente trattazione, i colori assunti dal testo o dai riquadri hanno la funzione di richiamare la parte dell'immagine che essi descrivono. Come osservabile in figura 3.3, nella **Input text** per l'inserimento del **Nome Fiume (River Name)** è inseribile qualsiasi carattere alfanumerico, mentre, per la **Portata (Discharge)**, impostando

Figura 3.1: Schermata app per inserimento generalità sito di campionamento.

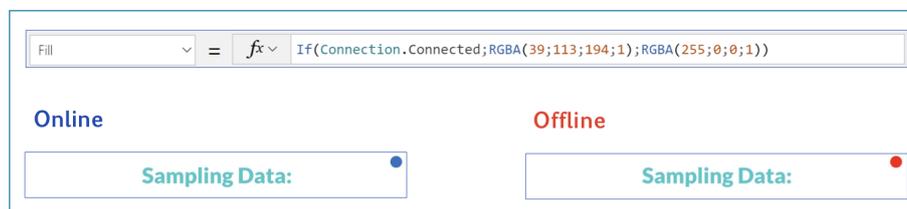


Figura 3.2: Visualizzazione stato rete e codice per la sua gestione.

la proprietà della **Input text** "VirtualKeyboardMode = Numeric", l'inserimento sarà vincolato ai soli caratteri numerici.

La **Regione** è selezionabile dall'elenco a discesa presente nella relativa **Combo box**, mentre è possibile inserire la **Data del Rilievo** tramite il **Date picker**. La **tipologia di campionamento biologico** è selezionabile spuntando l'apposita **Check box**. Terminato l'inserimento di tutte le richieste, si aprirà automaticamente la **casella di Riepilogo**; nel caso in cui si desiderasse modificare qualsiasi input sarà sufficiente disattivare il comando toggle (indicato dalla freccia di colore bianco), effettuare la modifica e riattivare lo stesso.

Selezionando l'icona di avanzamento, **freccia** in fig. 3.3, verrà generata la Raccolta "**_DATI_FIUME**", tabella contenente le informazioni inserite in precedenza, osservabile in figura 3.4.

The image displays the 'Bio Sampling' app interface. The main screen is titled 'Sampling Data:' and contains several input fields: 'Select Region' (a dropdown menu), 'River Name', 'Discharge' (with a unit of 'm3/s'), 'Date' (with a calendar icon), and 'Sampling type' (with radio buttons for 'Quantitative' and 'Semi-Quantitative'). A calendar widget on the right shows the date '16 giugno 2022'. Below the main form, a summary view shows the entered data: 'Region : Liguria', 'River Name : Tanaro', 'Discharge m3/s : 1.25', 'Date : 15/05/2022', and 'Sampling Type : Q'. A large blue arrow points downwards from the summary view, indicating the next step in the process.

Figura 3.3: Dettaglio comandi di inserimento caratteristiche generali del sito di campionamento.

Attraverso la funzione "SaveData", che consente il salvataggio e la memorizzazione del dato nella cache locale del dispositivo, come specificato nella [Documentazione di Power Apps di Microsoft](#), sarà possibile salvare le informazioni contenute nella raccolta. All'apertura dell'app (azione "OnStart"), attraverso la funzione "LoadData", avverrà il caricamento delle informazioni salvate; nel caso di blocco del dispositivo o chiusura dell'app sarà sufficiente riaprire quest'ultima senza che avvenga la perdita di informazioni. Prima del passaggio alla schermata di inserimento dati (ScreenHome), gestita dalle apposite istruzioni, da cui saranno effettivamente registrabili gli esemplari campionati, verrà generata la stringa contenente il nome con cui sarà salvato periodicamente il dato biologico.

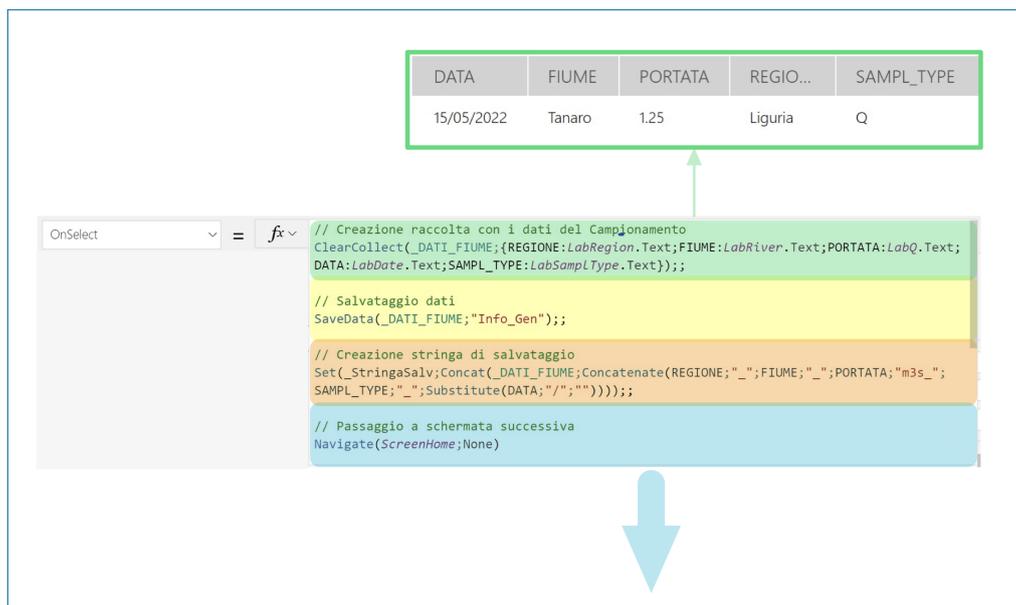


Figura 3.4: Dettaglio codice per operazioni avviate da barra di avanzamento.

3.1.2 Registrazione dato biologico

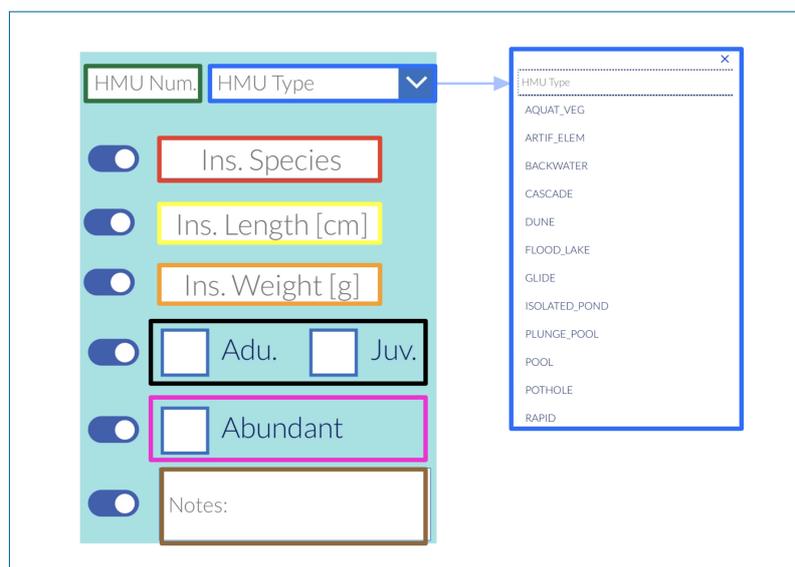
Dalla schermata in figura 3.5 saranno registrabili gli esemplari ritrovati durante il campionamento. Per ritornare alla schermata precedente sarà sufficiente premere l'icona evidenziata dall'apposita [freccia](#).

The screenshot shows a mobile application interface for biological data registration. At the top, there is a large white area with a blue arrow pointing upwards. Below this is a form with a light blue background. The form contains the following elements: a text input for 'HMU Num.', a dropdown menu for 'HMU Type', a toggle switch, and a vertical scroll bar with a '+' sign. Below these are several rows, each with a toggle switch and a text input: 'Ins. Species', 'Ins. Length [cm]', 'Ins. Weight [g]', 'Adu.' and 'Juv.' (with checkboxes), 'Abundant', and 'Notes'. A fish icon is located in the bottom right corner of the form area. At the bottom of the screen is a grey button labeled 'Export Definitive CSV'.

Figura 3.5: Schermata di registrazione dato biologico.

Come osservabile in figura 3.6, di default, i parametri inseribili per ciascun esemplare campionato sono i seguenti:

- **HMU Num**: numero dell'unità idro-morfologica campionata, formato input numerico.
- **HMU Type**: tipologia unità idro-morfologica campionata, selezionabile dall'elenco a discesa collegato (stesse scelte effettuabili in MapStream).
- **Species**: specie dell'esemplare considerato, formato input libero.
- **Length [cm]**: lunghezza dell'esemplare misurato (solo per campionamento quantitativo), formato input numerico.
- **Weight [g]**: peso dell'esemplare misurato (solo per campionamento quantitativo), formato input numerico.
- **Adu – Juv**: Stadio vitale dell'esemplare considerato (Adu – Adulto, Juv – Giovane), selezionabile da apposita Check box.
- **Abu**: Check box da selezionare per specie abbondante.
- **Notes**: eventuali note inseribili per lo specifico esemplare (es. sesso, caratteristiche distintive), formato input libero.



The image shows a screenshot of a mobile application interface for biological data collection. The main form has a light blue background and contains several input fields and checkboxes, each highlighted with a colored border: 'HMU Num.' (green), 'HMU Type' (blue dropdown), 'Ins. Species' (red), 'Ins. Length [cm]' (yellow), 'Ins. Weight [g]' (orange), 'Adu.' and 'Juv.' (black), 'Abundant' (pink), and 'Notes' (brown). To the right, a dropdown menu for 'HMU Type' is open, displaying a list of hydro-morphological types: AQUAT_VEG, ARTIF_ELEM, BACKWATER, CASCADE, DUNE, FLOOD_LAKE, GLIDE, ISOLATED_POND, PLUNGE_POOL, POOL, POTHOLE, and RAPID.

Figura 3.6: Dettaglio comandi di registrazione dato biologico.

Dopo l'inserimento, che avverrà selezionando il pulsante in figura 3.7, l'esemplare campionato verrà visualizzato nella galleria (in alto della 3.7) e tutti i campi verranno resettati (opzione di default).



Figura 3.7: Dettaglio opzioni e procedura di registrazione dato biologico.

Durante l'attività di campo, non sempre verranno registrati tutti i parametri (per esempio il peso), inoltre è pratica molto comune registrare sequenzialmente gli esemplari con le stesse caratteristiche, ragione per cui la cancellazione di tutte le informazioni (anche quelle in comune) dopo ogni inserimento rallenterebbe molto le procedure. Per ovviare a questa problematica, sono stati predisposti controlli **toggle** a fianco di ciascuna casella di input, che permetterà di attivare, soltanto le caselle relative ai parametri che si stanno effettivamente rilevando; questa tipologia di controllo permette un più rapido inserimento del dato e una maggiore fluidità durante la digitazione dei valori, qualità necessarie e da ricercare durante la raccolta in campo, dal momento che il focus (spostamento automatico alla casella di inserimento) non avverrà per le caselle indesiderate. Per le stesse finalità è stata prevista la possibilità di inserire il valore di un campo e successivamente bloccarlo tramite lo stesso toggle in modo da evitarne la cancellazione a seguito

dell'inserimento.

Si ipotizzi, per esempio, di dover registrare 15 sanguinerole adulte, ritrovate nella stessa unità, i cui unici parametri variabili sono lunghezza e stadio vitale. In figura 3.7 è quindi riportata la configurazione della schermata per questo esempio di campionamento quantitativo, con registrazione di soli lunghezza e stadio vitale (Adu – Juv), per tutti gli esemplari di sanguinerole dell'unità idro-morfologica POOL 1. A seguito dell'inserimento di ciascun esemplare, la proprietà peso e abbondanza non saranno registrabili, inserendo invece specie, numero e tipo di unità, e fissandoli tramite toggle, questi non verranno resettati. È evidente (ed è stato testato durante attività di campo descritte successivamente), come controlli di questo tipo velocizzino drasticamente le procedure di campionamento.

Al fine di migliorare la chiarezza della schermata è stato inoltre scelto di diminuire la dimensione delle caselle delle proprietà non inseribili (solamente per stadio vitale e abbondanza).

Si vuole evidenziare ancora la presenza del pulsante di ordinamento, comando che permetterà l'inversione dell'ordine di visualizzazione degli esemplari campionati, in base al numero di campionamento (visualizzando quindi in testa il primo o l'ultimo esemplare campionato), facilitandone la selezione durante le eventuali operazioni di modifica dettagliate al punto 3.1.3.

In figura 3.8 è rappresentato il dettaglio del pulsante di inserimento e la sua evoluzione dopo ogni selezione (incremento del **numero di pesci campionati**). Si è poi scelto di effettuare il backup dei dati su OneDrive (se rete internet disponibile), ogni 20 esemplari inseriti; viene quindi indicato il **numero di esemplari caricabili prima che avvenga il salvataggio**.

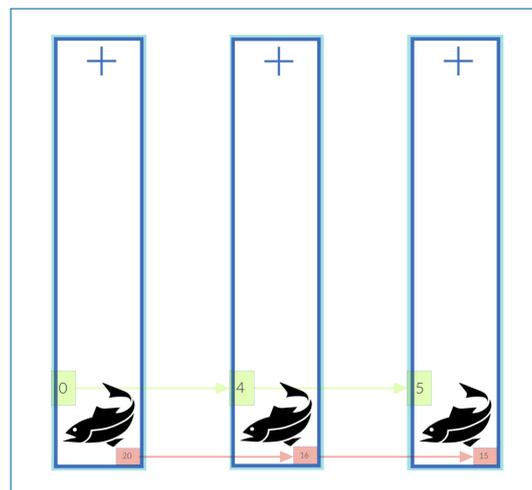


Figura 3.8: Dettaglio pulsante ed evoluzione post inserimento.

Nell'allegato [A4-CodInsEs](#) è riportato il codice richiamato ad ogni selezione del pulsante di inserimento.

Con la sezione A viene gestita la registrazione dell'esemplare misurato nella raccolta "_BIO_SAMPLING"; il formato della raccolta è riportato in figura 3.9.

ABUNDANCE	HMU_NUM	HMU_TYPE	INFO	LENGTH_cm	LIFE_STAGE	NOTE	SPECIE	WEIGH...
	1	POOL	1_Tanaro	7	A		Sanguinerola	
	1	POOL	2_Tanaro	6.5	A		Sanguinerola	
	1	POOL	3_Tanaro	6.2	A		Sanguinerola	

Figura 3.9: Esempio raccolta _BIO_SAMPLING.

Nella sezione B viene gestito l'eventuale reset della casella di inserimento.

Con la C, ogni 20 esemplari registrati e in condizione di disponibilità di rete, il backup automatico del dato, nella cartella "Rilievo_Bio" (OneDrive), su file .csv denominato con le caratteristiche generali del sito di campionamento, come osservabile in fig. 3.10. Si sottolinea come il salvataggio della portata in "m3s" sia dovuta all'impossibilità di salvare con stringhe contenenti il carattere "/" e di come ci si riferisca in realtà a m^3/s .



Figura 3.10: Cartella di salvataggio e file di backup .csv (Q in m^3/s).

Con la sezione D viene effettuato il salvataggio del dato nella memoria del dispositivo così da evitarne la perdita.

L'esportazione della raccolta _BIO_SAMPLING in OneDrive avviene attraverso il flusso cloud istantaneo "MESO_EXPORT" di Microsoft Power Automate (PA) attivato da PowerApps. Come osservabile in figura 3.11, il trigger è fornito in PA (premendo il pulsante di inserimento verrà avviato il salvataggio); la prima azione del flusso è quella di comporre l'input sulla base della struttura della raccolta _BIO_SAMPLING, per poi creare il file JSON (JavaScript Object Notation), quindi il .csv che verrà salvato nella cartella selezionata su OneDrive.

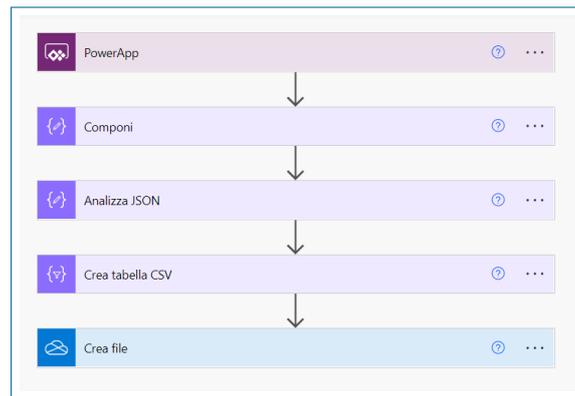


Figura 3.11: Flusso Power Automate per il salvataggio dei dati.

3.1.3 Modifica dati inseriti

All'interno dell'app è stata prevista una finestra di modifica dei dati inseriti; In figura 3.12 è riportata la sequenza di operazioni da seguire per modificare uno dei campi. Le fasi in cui si articola la modifica sono le seguenti:

- Selezione esemplare da modificare e apertura schermata di modifica tramite apposita icona.
- Avvio modifica tramite apposito pulsante.
- Modifica tramite caselle di input e salvataggio con apposito pulsante.
- Ritorno alla schermata principale a modifica avvenuta.

Tramite le frecce di avanzamento, sarà possibile ritornare, in qualsiasi momento, alla schermata di inserimento, annullando tutte le modifiche effettuate.

Come esempio viene riportata la modifica della Sanguinerola, 61-esimo esemplare campionato nel GLIDE 4, che è stata erroneamente inserita nella POOL 5 e a cui si vuole aggiungere la nota "Femmina con uova".

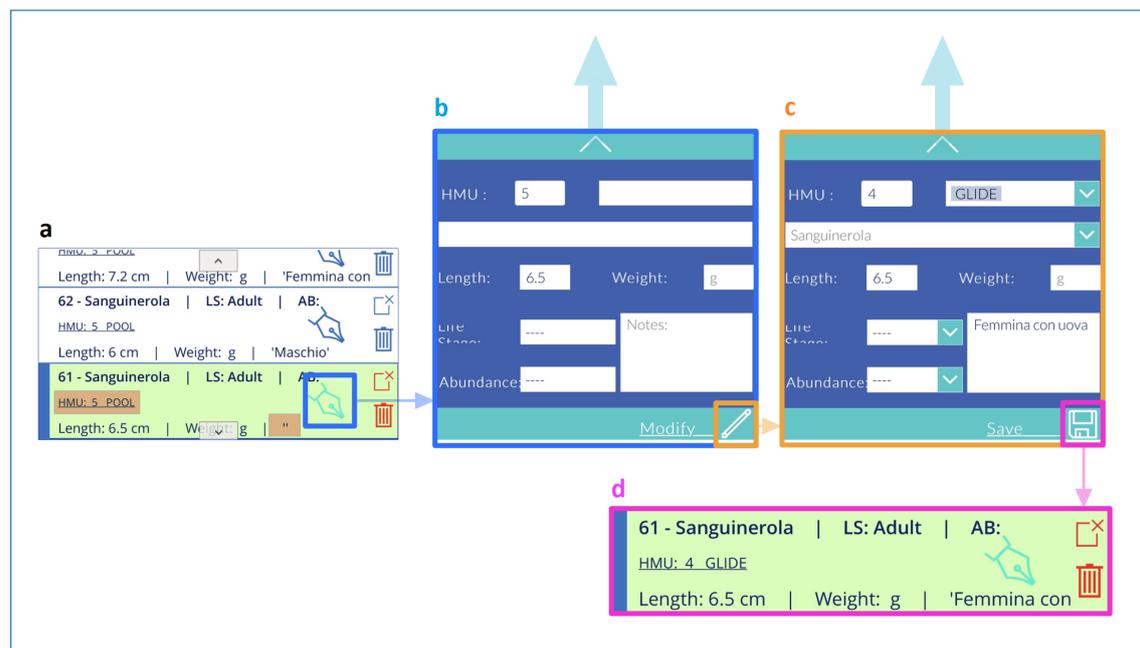


Figura 3.12: Sequenza comandi per modifica esemplare.

Attraverso i pulsanti in figura 3.13 sarà possibile rimuovere il solo esemplare selezionato oppure tutti gli inserimenti.

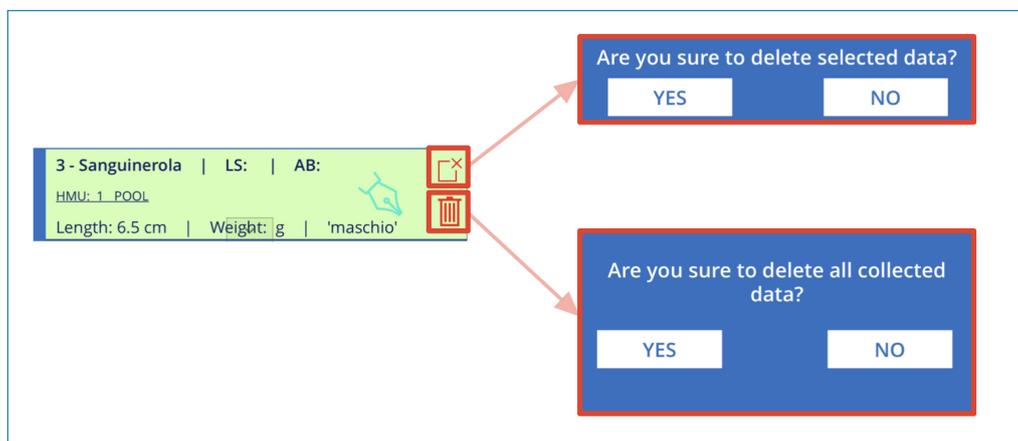


Figura 3.13: Pulsanti per eliminazione esemplari inseriti.

3.1.4 Salvataggio file campionamento finale

Terminato il campionamento sarà possibile esportare il file .csv nella rispettiva cartella collegata su OneDrive; è necessario specificare come esportando definitivamente il file, ogni elemento salvato sulla memoria locale del dispositivo verrà rimosso; è quindi consigliabile verificare preventivamente il corretto salvataggio su drive del file di Backup "Region_RiverName_Qm3s_SamplType_Date.csv".

In figura 3.14 è riportata la sequenza di operazioni che porterà, selezionando il pulsante "Export Definitive CSV", al salvataggio del file finale, nell'apposita cartella OneDrive, "Def_Regione_NomeFiume_Qm3s_SamplType_Date.csv". Affinchè il pulsante di esportazione sia selezionabile è necessario l'inserimento di almeno un dato (si noti la differenza di aspetto tra la 3.5 e la 3.14). Anche in questo caso, si sottolinea come la portata sia da intendersi in m^3/s .

L'app appena descritta è stata testata, quindi modificata, attraverso due campagne di campionamento in Liguria tra marzo e maggio 2022; le attività di campo e di ottimizzazione dell'applicazione sono descritte rispettivamente ai punti 4 e 5.1.



Figura 3.14: Flusso di operazioni per salvataggio file .csv.

3.2 "Hydro Bio Tool" per inserimento ed elaborazione dati.

La formattazione del dato grezzo e l'inserimento all'interno del database, come anticipato nella sezione 2.4, sono stati gestiti tramite la creazione di "Hydro Bio Tool", contenuto all'interno dell'omonima cartella di lavoro Excel. Di seguito sarà illustrato dettagliatamente il suo funzionamento.

3.2.1 Inserimento e Validazione Generalità del Rilievo

All'apertura della cartella sopra citata si aprirà in automatico la schermata iniziale del tool, da cui saranno inseribili le caratteristiche generali del sito del rilievo quali regione, nome del fiume, portata (m³/s), coordinate nel sistema di riferimento WGS84/UTM zona 32N e data. Per convenzione, come coordinate rappresentative, si è scelto di utilizzare quelle del baricentro della bounding box contenente tutte le unità idro-morfologiche campionate (come osservabile in figura 3.15).

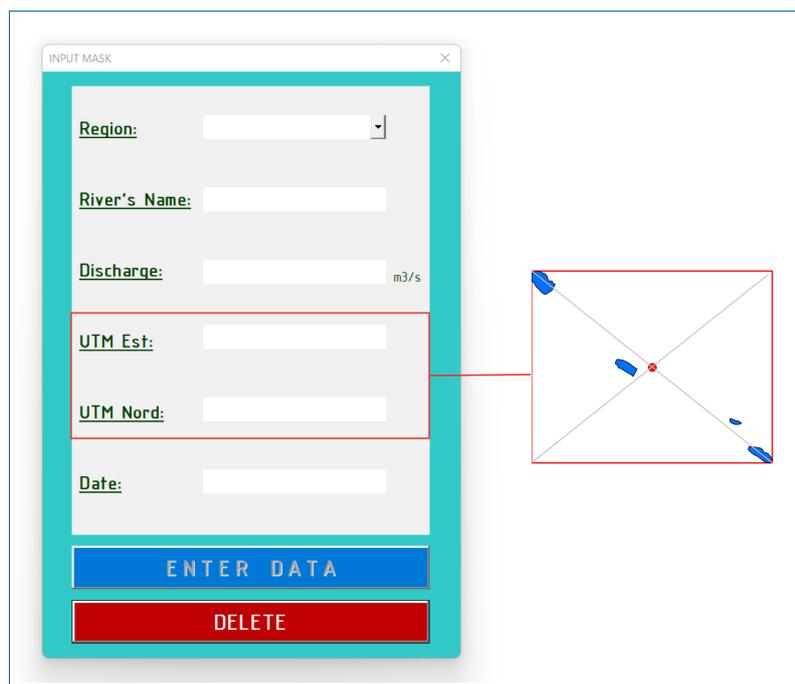
The image shows a software interface titled "INPUT MASK" with a close button (X) in the top right corner. The interface contains several input fields: "Region:" with a dropdown menu, "River's Name:" with a text input field, "Discharge:" with a text input field and the unit "m³/s" to its right, "UTM Est:" with a text input field, "UTM Nord:" with a text input field, and "Date:" with a text input field. At the bottom of the form are two buttons: a blue "ENTER DATA" button and a red "DELETE" button. To the right of the form, a map is displayed within a red rectangular bounding box. The map shows a red dot at the center, representing the centroid of the sampled units. A red line connects the "UTM Est:" field to the map.

Figura 3.15: Schermata di inserimento generalità del rilievo e bounding box HMU.

Prima di poter inserire i successivi input, è stata predisposta la validazione dei campi sopra citati; mentre la regione italiana è selezionabile dall'apposito menù a discesa e il nome del fiume non presenta vincoli di formato, la portata e le coordinate

devono essere campi numerici e la data deve presentare il formato "dd/mm/yyyy" (in figura 3.16 è riportato un esempio di validazione dei campi). Con il pulsante "DELETE", sarà possibile resettare tutti i campi inseriti, mentre, dopo la corretta validazione dei dati, sarà abilitato il comando "ENTER DATA" per il passaggio alla schermata successiva descritta nella sottosezione 3.2.2.

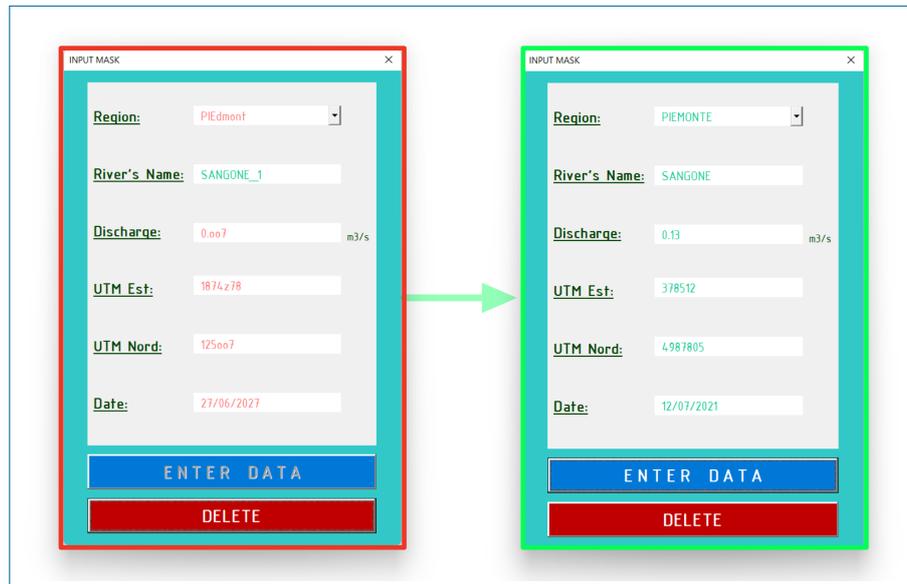


Figura 3.16: Validazione caratteristiche generali del rilievo inserite.

3.2.2 Inserimento e Validazione Dati Grezzi di Input

Nella schermata successiva sarà possibile selezionare i dati grezzi idro-morfologici e biologici, che verranno validati ed elaborati, e gli output desiderati. In figura 3.17 è osservabile la schermata nella configurazione di inserimento del dato biologico (a) ed idro-morfologico (b).

Tramite la selezione del pulsante "BACK TO PREVIOUS SCREEN" sarà possibile ritornare alla schermata precedente, mentre i pulsanti di inserimento permetteranno la selezione dei file di input (descritta al 3.2.3 e 3.2.4).

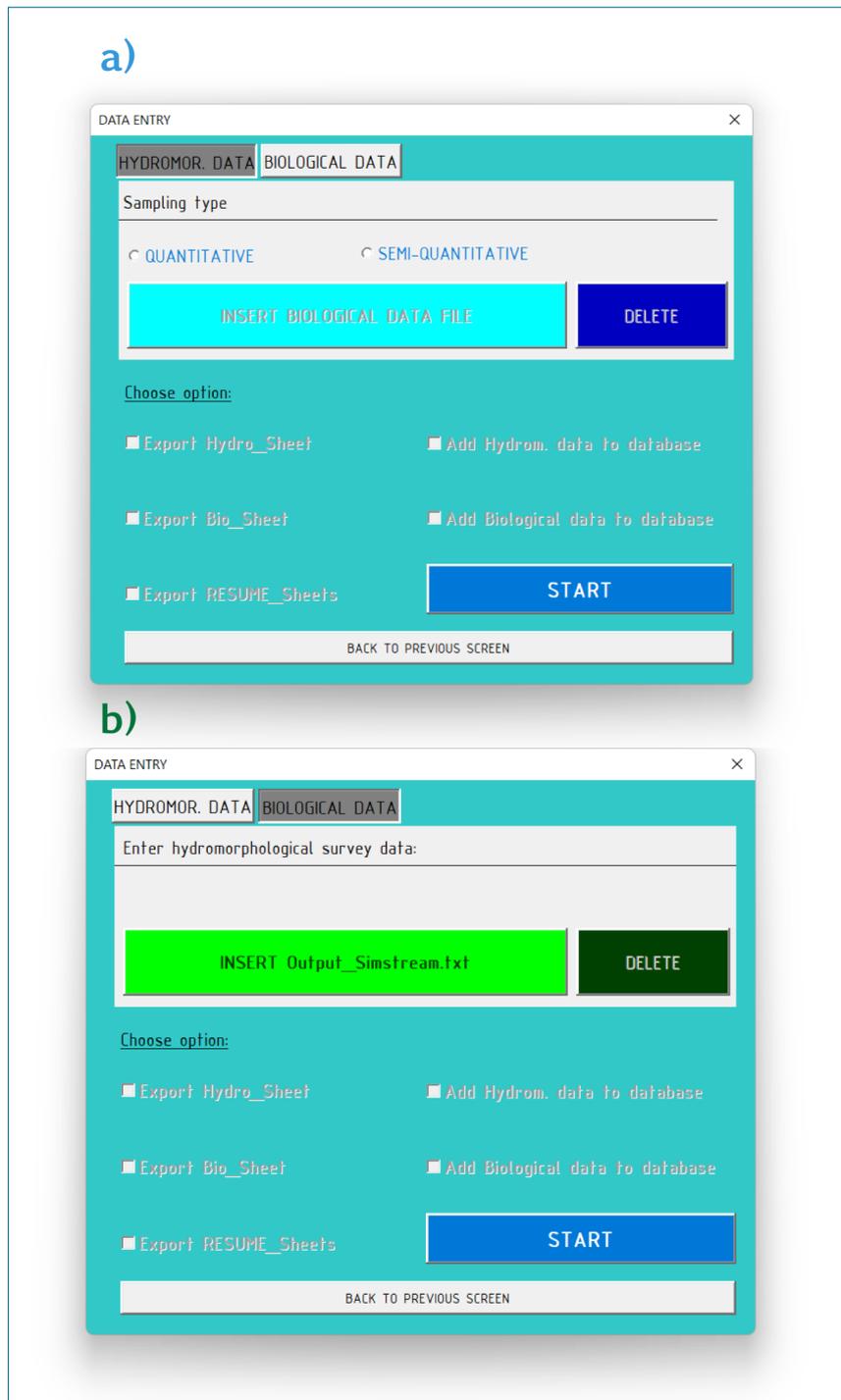


Figura 3.17: Schermata di inserimento del dato grezzo biologico (a) ed idro-morfologico, (b).

3.2.3 Selezione dato idro-morfologico

Dalla selezionando del pulsante "INSERT Output_SimStream.txt", osservabile in figura 3.18, verrà attivata la finestra di dialogo attraverso cui potrà essere selezionato l'apposito file di input (generato da SimStream-Web). Come anticipato, per questa tipologia di dato non è prevista alcuna validazione, essendo già eseguita in SimStream-Web. L'estensione del file richiesta è la **.txt** (che tramite apposito filtro in VBA è stata impostata come unica selezionabile).

Al file appena importato (contenente i descrittori fisici di ciascuna unità), saranno aggiunte quelle che sono le caratteristiche generali del sito, inserite dalla schermata precedente, e verrà generato in automatico il codice "HMU_CODE"; il file elaborato sarà salvato nello sheet ausiliario "Hydro_Sheet" e sarà esportabile (a seguito dell'inserimento del dato biologico).

Attraverso il pulsante DELETE le precedenti operazioni potranno essere annullate e sarà possibile inserire i dati di un nuovo rilievo.

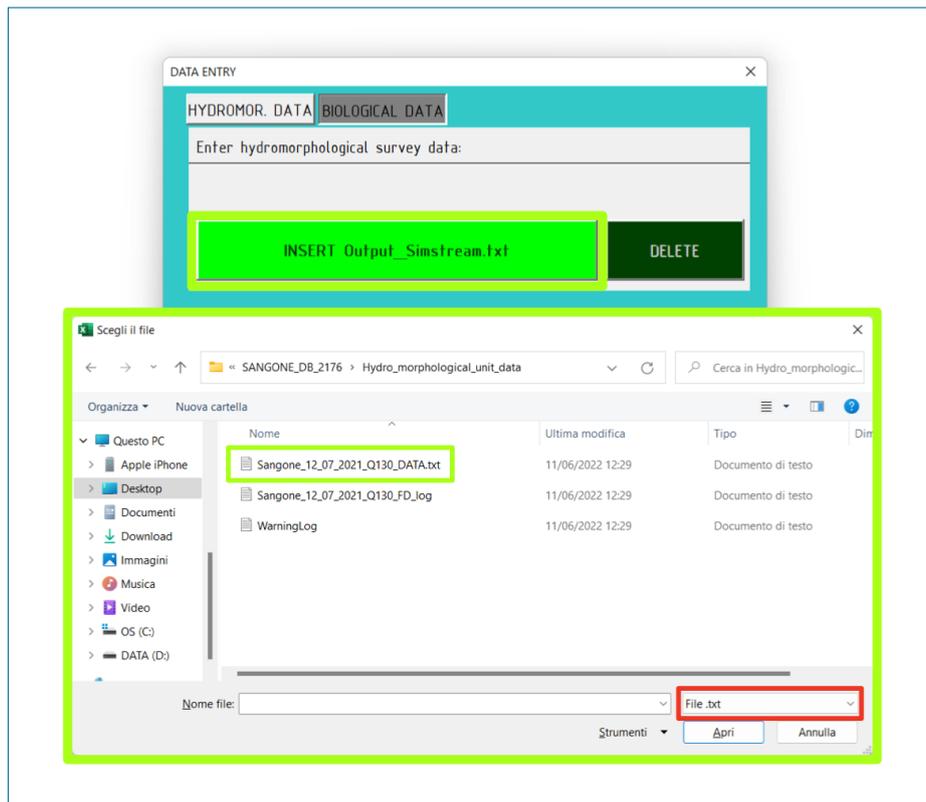


Figura 3.18: Schermata di selezione dato di input grezzo idro-morfologico.

3.2.4 Selezione dato biologico

In figura 3.19 è descritta l'interfaccia per l'inserimento del dato grezzo biologico, accessibile dal pulsante "INSERT BIOLOGICAL DATA FILE", solo dopo la selezione della tipologia di campionamento (quantitativa o semi quantitativa).

Il file da importare dovrà essere salvato in formato **.xls***.

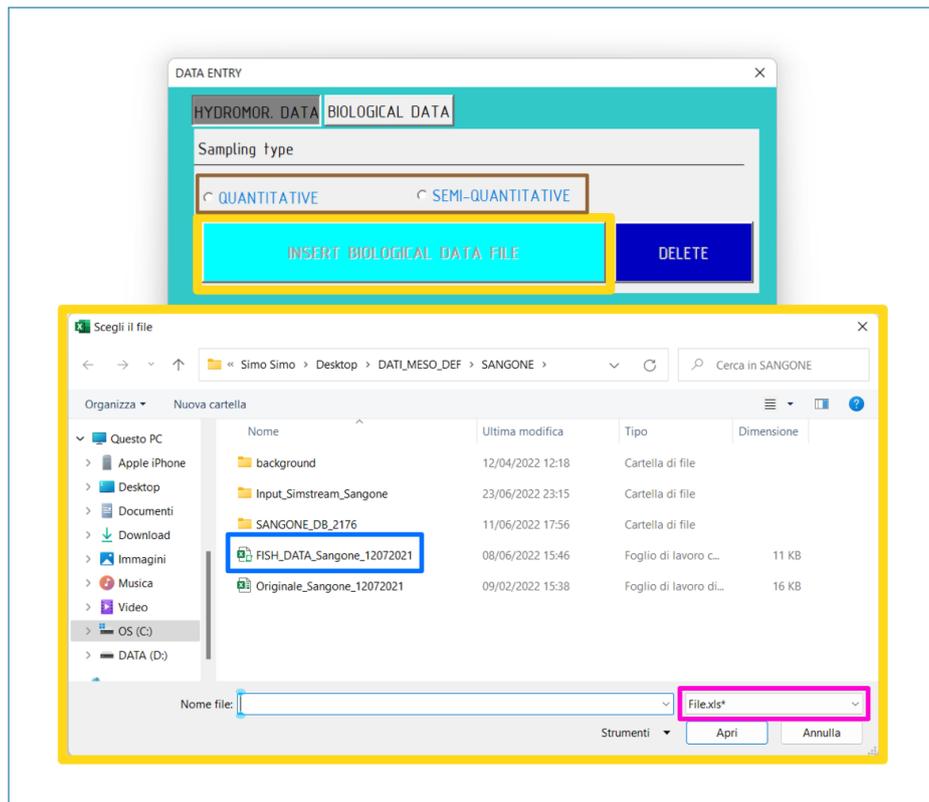


Figura 3.19: Schermata di selezione dato di input grezzo biologico.

3.2.5 Validazione dato biologico

Come discusso non è previsto, a livello nazionale, un formato informatizzato standard per la raccolta di dati biologici; è stata quindi proposta una formattazione base di riferimento per i dati grezzi che verranno elaborati e validati con il tool in descrizione. Proprio per questa eterogeneità tra i dati di partenza, è stato scelto di vincolare il solo numero delle grandezze inseribili, lasciando la maggior libertà possibile ai valori assumibili dai singoli campi, rimandando ad un apposito convertitore, interno al tool stesso, la conversione delle grandezze. Nel dato biologico di input dovranno essere presenti e ordinati i seguenti campi:

1. HMU_NUM: Numero dell'unità di ritrovamento dell'esemplare.
2. HMU_TYPE: Tipologia unità di ritrovamento dell'esemplare.
3. SPECIES_LAT: specie dell'esemplare.
4. LENGTH_cm: lunghezza in centimetri dell'esemplare.
5. WEIGHT_g: peso in grammi dell'esemplare
6. LIFE_STAGE: stadio vitale dell'esemplare.
7. ABUNDANCE: eventuale abbondanza, della specie e stadio vitale dell'esemplare, nell'unità di campionamento.

3.2.5.1 Validazione numero campi.

Non dovranno essere presenti colonne con intestazione mancanti (tra le prime 7 colonne) ma potranno essere inseriti valori nulli per i dati non campionati (ad es. nel caso in cui durante un campionamento non fosse stato misurato il peso degli esemplari, dovrà comparire comunque l'intestazione della colonna "WEIGHT_g" ma non sarà richiesto alcun valore per i campi).

File con un numero di colonne maggiore o minore rispetto allo standard non saranno inseribili; per evidenziare eventuali incongruenze, come osservabile in 3.20, sono stati predisposti messaggi di errore, attenzione e informazione per la guida dell'operatore nell'inserimento.

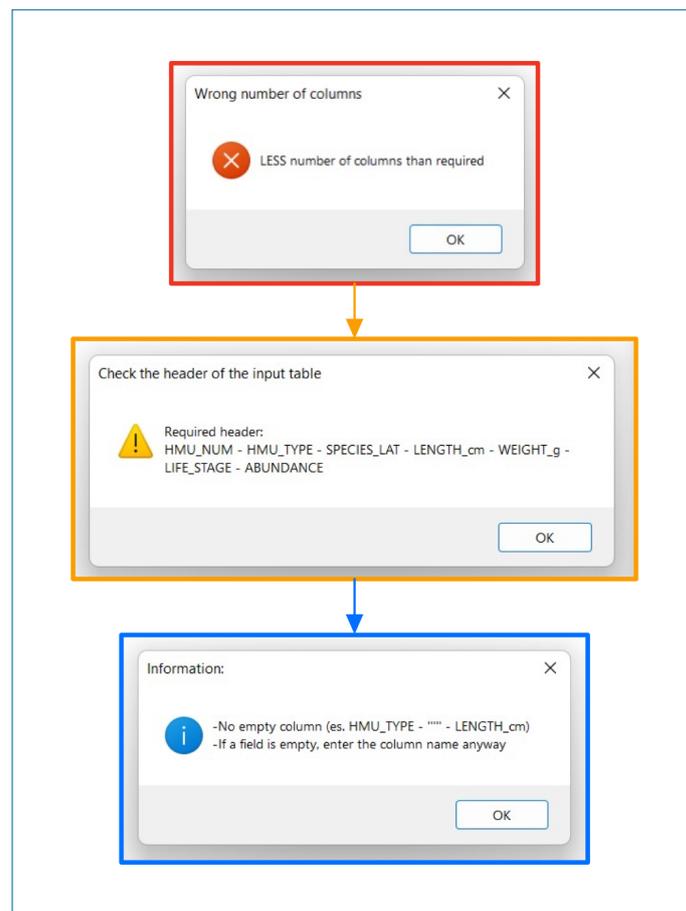


Figura 3.20: Messaggio di errore, attenzione e informazione per l'inserimento dei dati biologici.

3.2.5.2 Validazione intestazione campi.

Confermato il numero dei campi sarà valutata la correttezza del loro nome e indirettamente il loro ordine. In figura 3.21 è osservabile il caso di errata nomenclatura dell'intestazione (HMU_N) e l'inversione delle colonne evidenziate (con i relativi messaggi di errore che ne segnaleranno la presenza all'interno del tool). Sarà l'operatore a decidere se correggere automaticamente il nome oppure interrompere le operazioni.

The screenshot displays a data table with the following columns: HMU_N, HMU_TYPE, LENGTH_cm, LENGHT_cm, WEIGHT_g, LIFE_STAGE, and ABUNDACE. The table contains 32 rows of data. Two error dialog boxes are overlaid on the table:

- INCORRECT COLUMN HEADER: A**: This dialog box points to the HMU_N column. The message reads: "HMU_N different by HMU_NUM. You want to change the name of the column?" with "Si" and "No" buttons.
- INCORRECT COLUMN HEADER: C**: This dialog box points to the LENGHT_cm column. The message reads: "LENGHT_cm different by SPECIES_LAT. You want to change the name of the column?" with "Si" and "No" buttons.

HMU_N	HMU_TYPE	LENGTH_cm	LENGHT_cm	WEIGHT_g	LIFE_STAGE	ABUNDACE
1	GLIDE		Barbus caninus		ADULT	
1	GLIDE		Barbus plebejus		ADULT	
1	GLIDE		Padogobius martensii		JUVENILE	X
1	GLIDE		Padogobius martensii		ADULT	X
1	GLIDE		Phoxinus phoxinus		ADULT	
1	GLIDE		Squalius cephalus		ADULT	
1	GLIDE		Squalius cephalus		JUVENILE	
1	GLIDE		Telestes muticellus		JUVENILE	X
1	GLIDE		Telestes muticellus		ADULT	X
2	RIFFLE		Barbus caninus		ADULT	X
2	RIFFLE		Barbus caninus		JUVENILE	X
2	RIFFLE		Barbus plebejus		ADULT	
2	RIFFLE		Phoxinus phoxinus		ADULT	
2	RIFFLE		Telestes muticellus		JUVENILE	
2	RIFFLE		Telestes muticellus		ADULT	X
3	POOL		Barbus plebejus		ADULT	X
3	POOL		Barbus plebejus		JUVENILE	
3	POOL		Padogobius martensii		ADULT	
3	POOL		Salmo trutta		JUVENILE	
3	POOL		Squalius cephalus		ADULT	X
3	POOL		Squalius cephalus		JUVENILE	
3	POOL		Telestes muticellus		ADULT	
3	POOL		Telestes muticellus		JUVENILE	
4	GLIDE		Telestes muticellus		JUVENILE	X
4	GLIDE		Phoxinus phoxinus		JUVENILE	

Figura 3.21: Errata nomenclatura intestazioni e inversione ordine colonne.

3.2.5.3 Validazione formato campi.

Successivamente il tool di inserimento dati, valuterà il corretto formato di ciascun valore dei campi e consentirà l'eventuale conversione; in tabella 3.1 sono riportati i valori, o la tipologia di valori accettati. L'elenco delle specie è stato reperito da [6] e rappresentano circa il 75% delle specie pesci d'acqua dolce autoctone italiane.

Tabella 3.1: Valori campi accettati dato biologico.

HMU_NUM	HMU_TYPE	SPECIES_LAT	LENGTH_g	WEIGHT_g	LIFE_STAGE	ABUNDACE
NUMERICO	AQUAT_VEG	Alburnus albidus	NUMERICO	NUMERICO	A	X
	ARTIF_ELEM	Alburnus alburnus			J	-----
	BACKWATER	Alburnus arborella				NOT_AV
	CASCADE	Alosa fallax				
	DUNE	Anguilla anguilla				
	FLOOD_LAKE	Barbatula barbatula				
	GLIDE	Barbus caninus				
	ISOLATED_POND	Barbus plebejus				
	PLUNGE_POOL	Chondrostoma soetta				
	POOL	Cobitis taenia				
	POTHOLE	Cottus gobio				
	RAPID	Cyprinus carpio				
	RIFFLE	Esox cisalpinus				
	ROCK_GLIDE	Esox lucius				
	SEC_CHAN	Gobio benacensis				
	STEP	Gobio gobio				
	WATERFALL	Lampetra fluviatilis				
	WETLAND	Lampetra planeri				
		Lampetra zanandreai				
		Padogobius bonelli				
		Padogobius nigricans				
		Perca fluviatilis				
		Phoxinus phoxinus				
		Protochondrostoma genei				
		Rutilus pigus				
		Sabanejewia larvata				
		Salaria fluviatilis				
		Salmo cettii				
		Salmo marmoratus				
		Salmo trutta				
		Sarmarutilus rubilio				
		Scardinius erythrophthalmus				
		Squalius cephalus				
		Squalius lucumonis				
		Squalius squalus				
		Telestes muticellus				
		Thymallus thymallus				
		Tinta tinca				

La possibilità di inserire nuove specie, in aggiunta a quelle già selezionabili, verrà illustrata al punto 3.2.5.4.

3.2.5.4 Convertitore valori campi.

Si rimarca come non esista un formato informatizzato standard per la raccolta del dato biologico; i dati da aggiungere al database differiranno notevolmente nella forma. Per evitare eccessive operazioni di modifica sul dato biologico di campo, si è deciso di impedire l'inserimento solo per numero di colonne non conforme allo standard; eventuali discrepanze nel formato dei dati saranno gestite e modificate direttamente dal convertitore in figura 3.22.

Questo strumento riconoscerà tutti gli attributi con valore diverso da quelli in tabella 3.1 che compariranno nell'elenco a sinistra; selezionando il valore da modificare, l'opportuna scelta dal menù a tendina (a destra) e premendo "CONVERT", verrà eseguita in automatico la conversione per tutti i dati con valore non preconfigurato.

BIO VALIDATION

Change:

ADULT with J

JUVENILE

CONVERT

If you don't want convert the name, and add it as a possible future choice, press "STOP"

STOP

Figura 3.22: Convertitori campi dato di input grezzo biologico.

In figura 3.23 viene riportato un esempio di dato grezzo non conforme allo standard e la modifica dei valori non accettabili tramite convertitore.

Nel caso in cui si desiderasse inserire un nuovo valore tra le future scelte selezionabili (es. inserimento di una nuova specie), occorrerà premere il pulsante "STOP" e rispondere affermativamente al pop-up. Questa possibilità è implementata per tutti i valori dei campi convertibili da convertitore.

The diagram illustrates the data conversion process. A central table shows the data being processed:

HMU_NUM	HMU_TYPE	SPECIES_LAT	WEIGHT_g	LIFE_STAGE	ABUNDANCE
1	GLID	Ang. NON LAT		ADULT	no
2	POOL	Anguilla anguilla		JUVENILE	Si

Four 'Change' dialog boxes are shown, each with a 'CONVERT' button and a 'STOP' button. The dialog boxes are connected to the table by colored lines: blue for GLID, green for NO, pink for Ang. non lat, and yellow for ADULT. The 'STOP' button in the pink dialog box leads to a Microsoft Excel pop-up window with the message: 'Add Ang. non lat as a new type of SPECIES_LAT?' and buttons for 'Si' and 'No'.

Figura 3.23: Esempio di dato biologico con valori non conformi allo standard prescelto e loro modifica via convertitore.

3.2.5.5 Selezione output

Il dato in ingresso validato sarà ordinato per numero di unità crescente e, selezionando le operazioni desiderate (fig. 3.24), sarà possibile procedere con l'esportazione degli output avviate con il pulsante "START". Per ogni nuovo rilievo verrà generata una cartella (interna a "RELIEF DATA") contenente le schede desiderate; inoltre i dati appena formattati potranno essere aggiunti al Database MesoHABSIM (per la descrizione del formato delle schede si rimanda al punto 2.4).

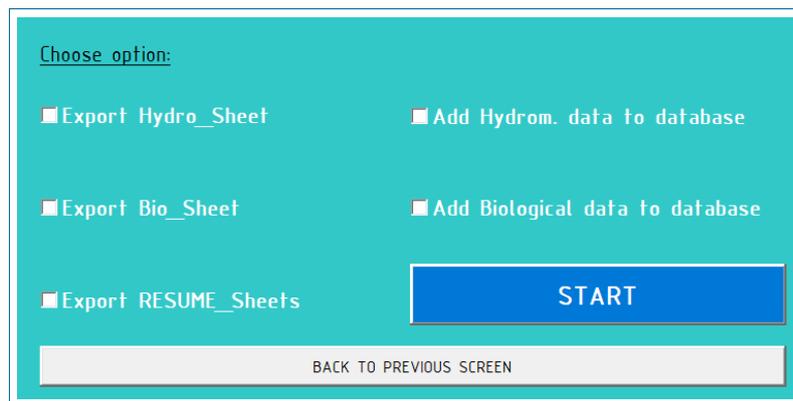


Figura 3.24: Schermata di selezione output.

Validazione corrispondenza HMU_CODE

Come anticipato nella sezione 2.4, nel database saranno inserite soltanto le unità per cui è stato eseguito sia il campionamento biologico che il rilievo idro-morfologico. In [A5-CodCorrHMU](#) si è riportato il codice VBA che valuta la corrispondenza tra unità, rimuovendo quelle senza corrispondenza.

3.3 "MesoHABSIM Database" per Immagazzinamento ed Estrazione dati.

I dati idro-morfologici e biologici, derivanti dai rispettivi rilievi, ed elaborati e formattati con "Hydro Bio Tool", saranno raccolti nel file .xslm "MesoHABSIM Database", dove saranno interrogabili e modificabili con tutti gli strumenti propri di Excel (es. strumenti di filtraggio e ordinamento). Data l'unicità della metodologia in esame, sono stati inoltre sviluppati degli strumenti specifici per le richieste e le necessità riscontrate. Nella presente sezione saranno descritti dettagliatamente tali strumenti (introdotti al punto 2.6.2), attivabili dalla barra di accesso rapido ed osservabili in figura 2.11.

3.3.1 "Site List" per elenco siti

È stato previsto uno strumento in grado di fornire una lista di tutti i siti inseriti nel database, evidenziando l'eventuale mancanza di una tipologia di una delle due categorie di dato. Attivando l'icona "Site List" (fig. 2.11) verrà avviato un algoritmo che, a partire dagli "HMU_CODE" dei dati idro-morfologici e biologici, genererà la lista dei "CODE" univoci presenti in almeno una delle due categorie di dato (schema esemplificativo osservabile in figura 3.25).

DATI IDROMORFOLOGICI		DATI BIOLOGICI	
HMU_CODE		HMU_CODE	
2-0.25-406216-4865888-29/07/2015		2-0.25-406216-4865888-29/07/2015	
3-0.25-406216-4865888-29/07/2015		2-0.25-406216-4865888-29/07/2015	
		2-0.25-406216-4865888-29/07/2015	
1-0.417-626263-5099824-11/10/2016		2-0.25-406216-4865888-29/07/2015	
2-0.417-626263-5099824-11/10/2016		2-0.25-406216-4865888-29/07/2015	
1-14.5-725210-5192677-19/09/2018		2-0.25-406216-4865888-29/07/2015	
2-14.5-725210-5192677-19/09/2018		2-0.25-406216-4865888-29/07/2015	
1-0.39-395643-5052696-22/11/2010		28-0.007-657504-5099156-21/10/2016	
2-0.39-395643-5052696-22/11/2010		28-0.007-657504-5099156-21/10/2016	
CODE		CODE	
0.25-406216-4865888-29/07/2015		0.25-406216-4865888-29/07/2015	
0.417-626263-5099824-11/10/2016			
14.5-725210-5192677-19/09/2018			
0.39-395643-5052696-22/11/2010		0.007-657504-5099156-21/10/2016	

Figura 3.25: Confronto CODE dati biologici e idro-morfologici.

3.3.2 "Delete Relief" per rimozione dati

Per la rimozione di un eventuale sito di campionamento indesiderato è stato sviluppato il tool "Delete Relief", attivabile dall'icona apposita in fig. 2.11, e rappresentato nella 3.26. Saranno eliminati tutti i dati idro-morfologici e/o biologici (in base alla selezione della checkbox) aventi lo stesso codice CODE inserito in input.

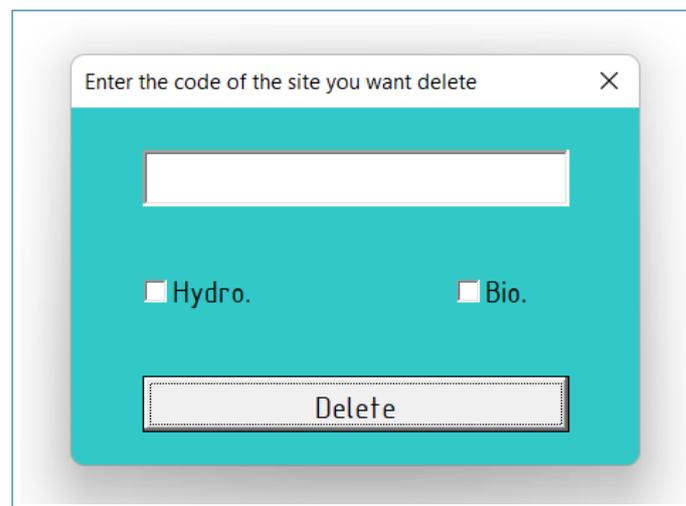


Figura 3.26: Interfaccia tool "Delete Relief".

L'algoritmo avviato dal pulsante "Delete" (riportato in allegato [A6-CodDelRel](#)), cui compete l'eliminazione dei dati, confronterà il CODE inserito in input con quelli della tipologia di dato prescelto, eliminando le eventuali corrispondenze.

3.3.3 "Relief Map" per distribuzione spaziale rilievi

È stato generato un comando, attivabile dall'apposita icona in fig. 2.11, che a partire dalle coordinate dell'Italia, generasse un grafico all'interno della quale siano rappresentati tutti i rilievi presenti nell'elenco List_CODE (coordinate UTM Est e Nord ricavate dal codice CODE). In figura 3.27 è riportato un esempio di grafico ottenuto.



Figura 3.27: Grafico siti rilievi ottenuto tramite tool "Relief Map".

3.3.4 "Extract Biological Data" per estrazioni dati biologici

Si vuole rimarcare come, per la generazione dei modelli di distribuzione di specie è necessario disporre di un dato che contenga al suo interno le informazioni relative sia ai descrittori fisici dell'habitat fluviale, sia l'assenza, presenza e abbondanza della specie in ciascuna unità; si vuole cioè ottenere un file che contenga, per lo stadio vitale della specie selezionata, tutte le unità registrate nei rilievi in cui sia stato trovato almeno un esemplare avente queste caratteristiche. È stato quindi sviluppato lo strumento "Extract Biological Data", osservabile in figura 3.28 e attivabile dall'apposita icona in fig. 2.11, che permetterà l'estrazione della scheda apposita per la generazione del modello biologico.

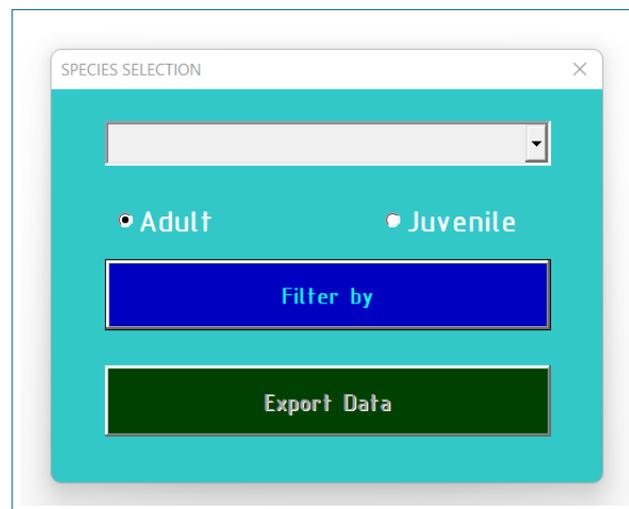


Figura 3.28: Strumento "Extract Biological Data" per l'estrazione di dati biologici.

La specie desiderata è selezionabile, dall'apposito elenco, tra tutte le specie caricate nel database biologico (figura 3.29 a sinistra), mentre lo stadio vitale è definibile dal pulsante di opzione selezionato (figura 3.29 a destra).

Effettuate le selezioni sarà possibile filtrare e comporre l'output desiderato. L'algoritmo di estrazione si articola nei seguenti punti:

- Duplicazione dati biologici in sheets ausiliario "Species_Extraction" in cui verranno eseguite le operazioni di eliminazione dei dati superflui.
- Dichiarazione variabile SPECIE_PESCE (contenente la specie selezionata es. *Telestes muticellus*).
- Dichiarazione variabile STADIO_VITALE (contenente lo stadio vitale selezionato es. Adult).

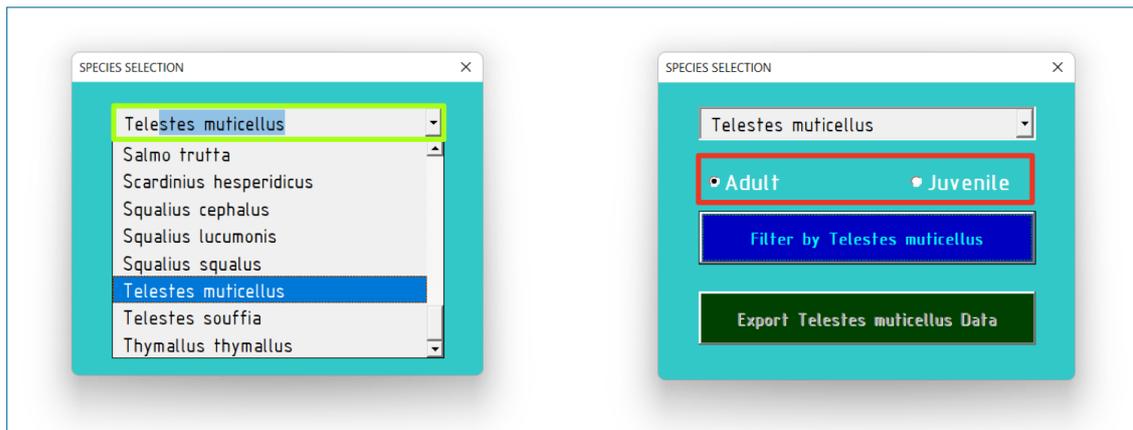


Figura 3.29: Selezione specie (sinistra) e stadio vitale (destra) da "Extract Biological Data".

- Rimozione dati duplicati pesci della stessa unità e stadio vitale.
- Rimozione pesci = SPECIE_PESCE con stadio vitale \neq STADIO_VITALE e modifica nomi specie \neq SPECIE_PESCE.
- Rimozione esemplari duplicati nella stessa unità e di specie sia diversa a SPECIE_PESCE.

Arrivati a questo punto saranno presenti tutti gli esemplari della SPECIE_PESCE e STADIO_VITALE selezionati e tutte le unità in cui il quella specie non è stata trovata, comprese quelle in cui non è stata trovata in nessuna unità; occorrerà quindi rimuovere quest'ultime.

- Generazione dei vettori HMU_CODE_SLSPECIE() e CODE_SLSPECIE() contenenti i codici di tutte le unità e i siti in cui la SPECIE_PESCE dello STADIO_VITALE è stata ritrovata.
- Rimozione siti in cui la SPECIE_PESCE dello STADIO_VITALE NON è stata ritrovata in nessuna unità, tramite confronto CODE unità nello sheets estrazioni con CODE_SLSPECIE.
- Rimozione unità che in cui è stato campionato sia la SPECIE_PESCE dello STADIO_VITALE che almeno un altro esemplare di specie diversa da SPECIE_PESCE; eliminazione unità con HMU_CODE uguale a HMU_SLSPECIE MA specie diversa da SPECIE_PESCE (unità ancora duplicata).
- Confronto HMU_CODE unità rimanenti con database idro-morfologico per associazione dati fisici e biologici.

In tabella 3.3 è osservabile un esempio di file di output generato.

Nella colonna "Adu_ASS0_PRE1_ABB2" sarà presente l'indicazione di ASSENZA (0), PRESENZA (1) o ABBONDANZA (2) della specie. Le colonne "Adu_ASS0_PRE1" e "Adu_PRE0_ABB1" indicano rispettivamente la ASSENZA (0) / PRESENZA (1) e PRESENZA (0) / ABBONDANZA (1). I modelli binari di distribuzione di specie saranno sviluppati proprio da queste etichette di classe. L'etichetta "Not found" della colonna "Adu_PRE0_ABB1" indica un'unità in cui il pesce è assente e quindi dovrà essere scartata nella generazione del modello PRESENZA/ABBONDANZA mentre l'etichetta "NOT_AV" indicano dati di rilievi in cui non si hanno, e non si potranno avere, informazioni sull'abbondanza della specie (campionamenti semi-quantitativi di cui non può essere ricavato l'istogramma di densità).

Tabella 3.3: Estrazione dato biologico con "Extract Biological Data" per il vairone (*Telestes muticellus*) adulto.

RIVER	HMU_CODE	HMU_NUM	HMU_TYPE	OUTPUT SIMSTREAM	SPECIES_LAT	LIFE_STAGE	Adu_ASS0_PRE1_ABB2	Adu_ASS0_PRE1	Adu_PRE0_ABB1
ARGENTINA	2-0.25-406216-4865888-29/07/2015	2	BACKWATER	...	Telestes muticellus	A	1	1	0
ARGENTINA	3-0.25-406216-4865888-29/07/2015	3	GUIDE	...	Telestes muticellus	A	2	1	1
ARGENTINA	4-0.25-406216-4865888-29/07/2015	4	GUIDE	...	Telestes muticellus	A	2	1	1
ARGENTINA	5-0.25-406216-4865888-29/07/2015	5	RIFFLE	...	Telestes muticellus	A	1	1	0
ARGENTINA	6-0.25-406216-4865888-29/07/2015	6	RIFFLE	...	Telestes muticellus	A	2	1	1
ARGENTINA	7-0.25-406216-4865888-29/07/2015	7	BACKWATER	...	Telestes muticellus	A	1	1	0

TREBBIA	5-1.73-546860-4980443-02/07/2020	5	POOL	...	No Telestes muticellus found		0	0	Not found
TREBBIA	6-1.73-546860-4980443-02/07/2020	6	POOL	...	Telestes muticellus	A	1	1	NOT_AV
TREBBIA	7-1.73-546860-4980443-02/07/2020	7	GUIDE	...	No Telestes muticellus found		0	0	Not found
TREBBIA	8-1.73-546860-4980443-02/07/2020	8	RIFFLE	...	Telestes muticellus	A	1	1	NOT_AV
TREBBIA	9-1.73-546860-4980443-02/07/2020	9	BACKWATER	...	Telestes muticellus	A	1	1	NOT_AV
TREBBIA	10-1.73-546860-4980443-02/07/2020	10	GUIDE	...	No Telestes muticellus found		0	0	Not found

Dopo aver ricavato la tabella di output, il pulsante di esportazione in figura 3.28 sarà attivato e sarà possibile esportare i dati; nel caso in cui non fosse mai stata eseguita l'estrazione per la specie selezionata, verrà creata, all'interno della cartella "RELIEF DATA", la cartella "Specie" in cui verrà esportato il file appena creato con il nome "LS_Specie_Date.xlsm".

Capitolo 4

Attività di Campo e Test Strumenti

Nel periodo 7/11 marzo e 16/20 maggio 2022 sono stati eseguiti i rilievi idro-morfologici e i campionamenti biologici in 9 tratti fluviali di 8 corsi d'acqua, nelle provincie di Savona e Imperia, dove è stata testata la funzionalità e l'efficacia dell'app descritta alla sezione 3.1; i dati raccolti sono stati successivamente trattati con gli strumenti descritti nelle sezioni 3.2 e 3.3.

4.1 Area di studio

I campionamenti di maggio e marzo sono stati effettuati negli stessi 9 siti, osservabili in figura 4.1.

Di seguito verranno descritte le attività svolte, partendo dalla raccolta dati in campo, fino al loro caricamento nel Database MesoHABSIM.

In ciascun sito, e per ogni unità (o gruppo di unità), è stato eseguito prima il campionamento biologico, per non influenzarne i risultati e, solo successivamente, quello idro-morfologico (è facile immaginare come gli esemplari presenti all'interno delle unità, vengano disturbati e possano disperdersi al passaggio dell'operatore generando una possibile sottostima del dato campionato).

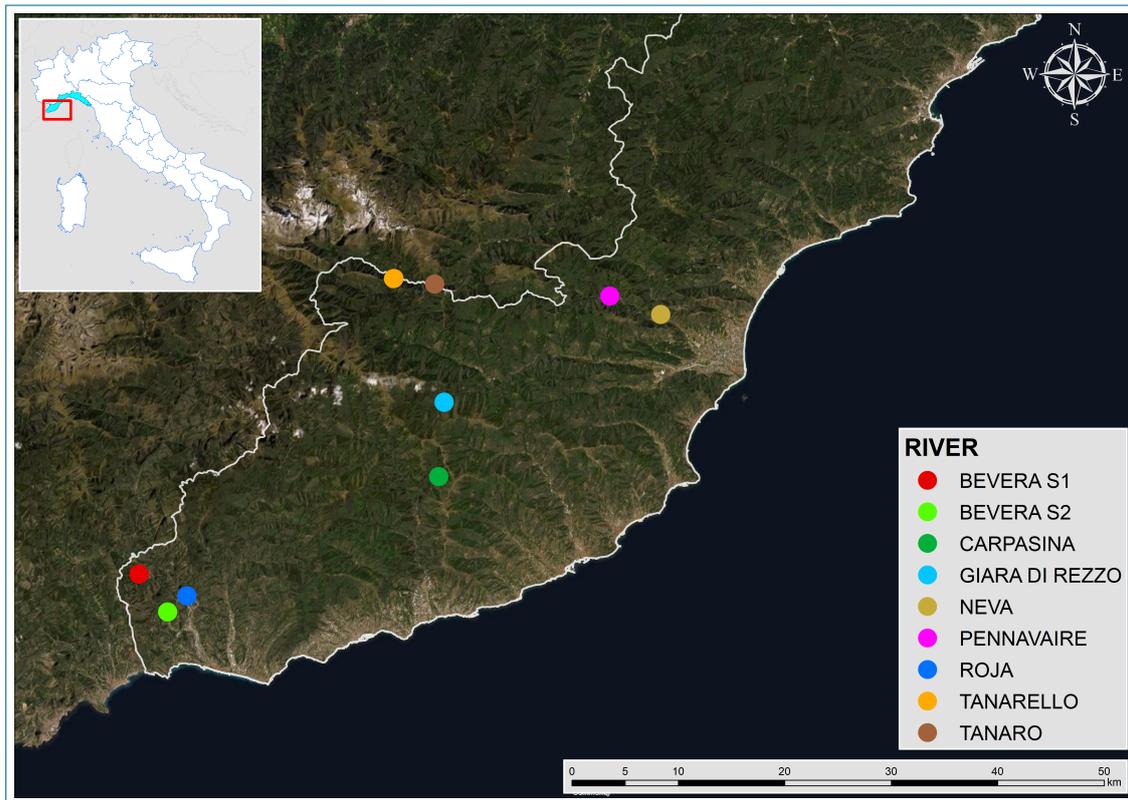


Figura 4.1: Mappa dei siti campionati in Liguria nei mesi di marzo e maggio 2022.

4.1.1 Campionamento biologico

4.1.1.1 Ricerca pesce per unità idro-morfologica.

Ciascuna unità idro-morfologica è stata campionata, da valle verso monte, mediante elettrostorditore spallabile, come osservabile in figura 4.2 e 4.3.

Durante le operazioni ci si è serviti di retino da pesca per facilitare la cattura dei pesci (figura 4.4).

Gli esemplari campionati in ciascuna unità sono stati riposti, in funzione della dimensione, in un secchio (dimensione minore, esempio in fig. 4.5) oppure in una nassa (dimensione maggiore, fig. 4.6).



Figura 4.2: Campionamento HMU con elettroscorditore spallabile.



Figura 4.3: Dettaglio retino elettroscorditore spallabile.

4.1.1.2 Misura esemplari

Ciascun esemplare, compatibilmente alle sue dimensioni, è stato misurato dalla punta del muso all'estremità della pinna caudale (lunghezza totale), tramite ittiometro (dimensioni minori), fig. 4.7, o con bindella metrica (dimensioni maggiori), fig. 4.8.



Figura 4.4: Utilizzo di retino da pesca come ausilio durante le operazioni di campionamento biologico.



Figura 4.5: Raccolta esemplari in apposito secchio.



Figura 4.6: Raccolta esemplari in apposita nassa.



Figura 4.7: Misurazione esemplare di anguilla con ittiometro.



Figura 4.8: Misurazione esemplare di anguilla con bindella metrica.

4.1.1.3 Registrazione dato e Test app

La registrazione del dato è stata eseguita, da telefono cellulare, sia su foglio Excel che tramite App, come osservabile in figura 4.9.

Per ciascun esemplare sono stati registrati i valori di numero e tipo di meso-habitat, specie, lunghezza e stadio vitale.

È così stato possibile testare il funzionamento dell'app, confrontando la funzionalità delle 2 procedure e i dati raccolti.



Figura 4.9: Test app sviluppata e registrazione pesce.

4.1.1.4 Liberazione esemplari

Dopo la fase di misura, ciascun esemplare è stato liberato nella stessa HMU in cui era stato precedentemente campionato (in figura 4.10 è osservabile un esemplare di anguilla appena liberato).



Figura 4.10: Esempio di anguilla appena liberato nella medesima unità di ritrovamento (foto subacquea).

4.1.2 Rilievo idro-morfologico

Il rilievo idro-morfologico, eseguito seguendo le procedure proprie della metodologia MesoHABSIM [16], si è articolato nelle fasi seguenti.

4.1.2.1 Calibrazione telemetro laser

Le unità idro-morfologiche sono state mappate con telemetro laser (Trupulse 360R, Laser Technology, Inc., Centennial, CO, USA); previa calibrazione dell'inclinometro e della bussola elettronica, sensori interni allo strumento. In figura 4.11 è osservabile la procedura di calibrazione.



Figura 4.11: Operazione di calibrazione telemetro laser.

4.1.2.2 Mappatura unità idro-morfologiche

In figura 4.12 è rappresentata l'individuazione dei meso-habitat di un tratto fluviale. La mappatura delle unità individuate è stata eseguita mediante il software MapStream, operante in ambiente QGIS, da computer da campo "rugged" (Panasonic TOUGHBOOK FZ-M1, Panasonic Corporation, Kadoma, Japan), come osservabile in figura 4.13. Il perimetro di ciascuna HMU è stato quindi tracciato con il telemetro, installato su treppiede fotografico (figura 4.14); per ciascuna unità sono stati inseriti numero, tipologia e attributi (es. covers presenti e eventuali note).

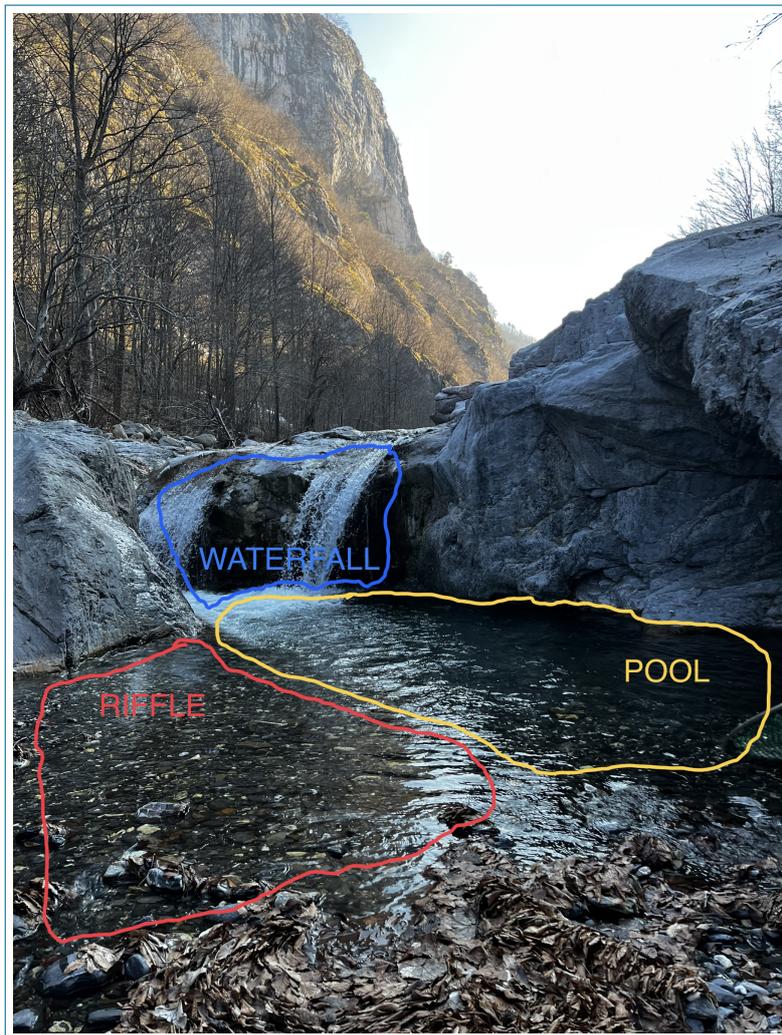


Figura 4.12: Distinzione meso-habitat all'interno di un tratto fluviale.

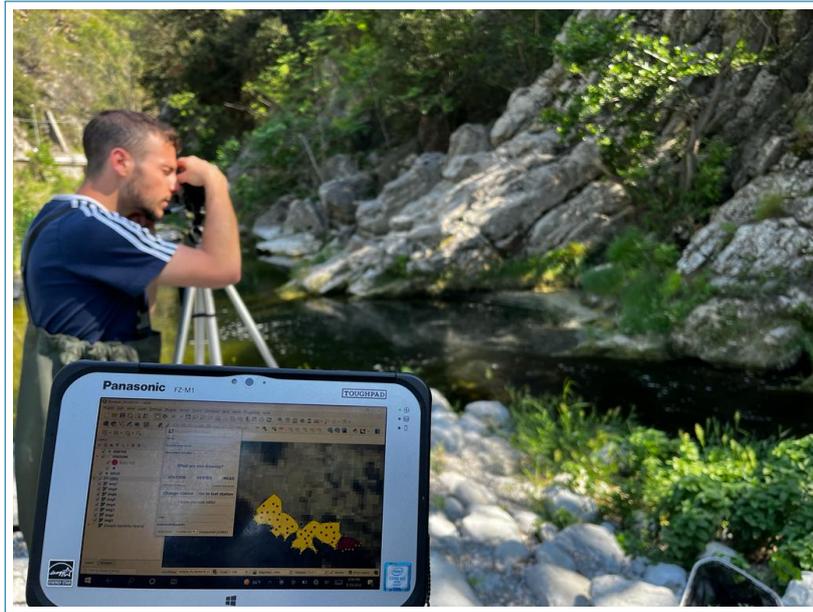


Figura 4.13: Utilizzo MapStream da computer da campo e mappatura meso-habitat tramite telemetro laser.



Figura 4.14: Telemetro laser montato su treppiede.

4.1.2.3 Rilevamento punto stazione con antenna

È noto come durante la mappatura dei meso-habitat, seguendo una poligonale con i punti stazione, vengano commessi degli errori di posizionamento su quest'ultimi. Per una corretta georeferenziazione dei diversi punti stazione, è stato utilizzato il ricevitore RTK GNSS (Emlid Reach RS2, Emlid, Budapest, Hungary), montato su apposita asta (fig. 4.15), che permette di ottenere la posizione del punto con accuratezza centimetrica.



Figura 4.15: Antenna e asta per la misura del punto stazione.

4.1.2.4 Misurazioni puntuali di profondità, velocità e substrato

Per ciascun punto stabilito all'interno dell'unità è rilevata la tipologia di substrato e, tramite il correntometro OTT MF pro (OTT HydroMet), sono misurate velocità e profondità della corrente; la posizione di ciascun punto campionato e le grandezze misurate sono registrate in MapStream. In figura 4.16 è osservabile la procedura di misurazione e la distribuzione dei punti campionati.

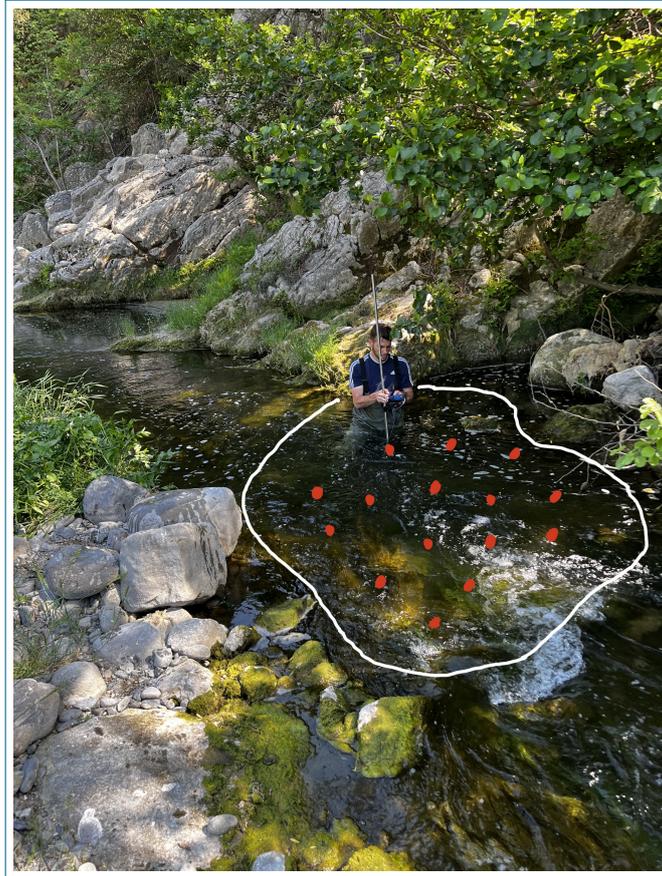


Figura 4.16: Distribuzione punti e misurazione di profondità, velocità e substrato.

4.1.2.5 Misurazione portata

Come riportato nell'apposito manuale [16], occorre misurare la “portata defluente in alveo a riferimento delle condizioni di deflusso che determinano la specifica distribuzione spaziale di unità morfologiche rilevate nel sotto tratto”: Si è quindi selezionata la sezione trasversale in cui eseguire la misura con correntometro OTT MF pro e bindella metrica.

4.2 Trattamento dati

4.2.1 Elaborazione dati biologici

In figura 4.1 è osservabile un esempio di dato biologico, registrato tramite app, a cui sono state apportate successive e contenute modifiche (come l'ordinamento delle colonne), il dato raccolto è stato poi elaborato con il tool di inserimento dati per poi essere caricato nell'apposita sezione del DataBase MesoHABSIM.

Tabella 4.1: Dato biologico raccolto con apposita app, Carpasina, 16/05/2022

ABUNDANCE	HMU_NUM	HMU_TYPE	INFO	LENGTH_cm	LIFE_STAGE	NOTE	SPECIE	WEIGHT_g
1	1	RAPID	1_Carpasina	17	A		Fario	
1	1	RAPID	2_Carpasina	14	J		Iridea	
1	1	RAPID	3_Carpasina	4	J		Fario	
1	1	RAPID	4_Carpasina	4.5	J		Fario	
1	1	RAPID	5_Carpasina	4	J		Fario	
2	2	RIFFLE	6_Carpasina	15			Fario	
2	2	RIFFLE	7_Carpasina	12			Fario	
2	2	RIFFLE	8_Carpasina	16			Fario	
2	2	RIFFLE	9_Carpasina	4	J		Fario	
2	2	RIFFLE	10_Carpasina	4	J		Fario	
2	2	RIFFLE	11_Carpasina	4	J		Fario	
3	3	GLIDE	12_Carpasina	3.5			Fario	
3	3	GLIDE	13_Carpasina	5			Fario	
3	3	GLIDE	14_Carpasina	4			Fario	
3	3	GLIDE	15_Carpasina	4			Fario	
3	3	GLIDE	16_Carpasina	4			Fario	
4	4	RIFFLE	17_Carpasina	4.5			Fario	
4	4	RIFFLE	18_Carpasina	3.5			Fario	
5	5	GLIDE	19_Carpasina	4.5			Fario	
5	5	GLIDE	20_Carpasina	4.5			Fario	
5	5	GLIDE	21_Carpasina	4			Fario	
5	5	GLIDE	22_Carpasina	3.5			Fario	
5	5	GLIDE	23_Carpasina	4.5			Fario	
5	5	GLIDE	24_Carpasina	4			Fario	
5	5	GLIDE	25_Carpasina	4			Fario	
6	6	POOL	26_Carpasina	25			Fario	
6	6	POOL	27_Carpasina	17			Fario	
6	6	POOL	28_Carpasina	15			Fario	
6	6	POOL	29_Carpasina	14.5		Da incubatoio	Fario	

4.2.2 Elaborazione dati idro-morfologici

Dopo il rilievo idro-morfologico, da QGIS sono stati salvati i seguenti file:

- Mappa di Background: file .tiff della mappa di background.
- Shapefile di:
 - STATION: shapefile puntuale della posizione delle stazioni di misura valutate con telemetro.
 - VERTEX: shapefile puntuale dei vertici delle unità mappate con il telemetro.
 - HMU: shapefile poligonale delle unità idro-morfologiche mappate.
 - MEAS: shapefile puntuale della posizione delle misure di profondità, velocità e substrato.
 - RTK GNSS: shapefile puntuale della posizione delle stazioni valutate con il ricevitore RTK GNSS (ricavato da file .csv).

È stata poi eseguita la traslazione di tutte le stazioni dello shapefile STATION, e di tutte le HMU mappate dalla medesima stazione, verso i corrispondenti punti dello shapefile RTK GNSS (in modo che il punto stazione e il punto antenna si sovrappongano). I file prima della traslazione sono rappresentati in fig. 4.17; lo shapefile delle HMU correttamente georeferenziate, elaborato tramite ArcMap, è invece rappresentato in figura 4.18.

Per poter eseguire correttamente le operazioni di traslazione, nella colonna commenti della tabella attributi delle HMU, deve essere inserito il numero di stazione da cui sono state mappate (operazione eseguibile direttamente con MapStream), con cui collegarsi al rispettivo numero del punto antenna.

La tabella attributi dello shapefile MEASURE è stata convertita in formato .txt, richiesto per il caricamento in SimStream-Web, mentre le misure di profondità sono state convertite da cm a m.

I dati idro-morfologici sono quindi stati caricati in SimStream-Web e il file .txt di output delle HMU è stato a sua volta utilizzato come input allo strumento "Hydro Bio Tool" per essere infine caricato nell'apposita sezione del database MesoHASIM.



Figura 4.17: Shapefile rilievo in campo, necessaria traslazione verso i punti Antenna.



Figura 4.18: Shapefile HMU georeferenziato rispetto ai punti Antenna.

Capitolo 5

Risultati

Di seguito verrà descritta l'ottimizzazione apportata all'app ed i risultati ottenuti, dall'elaborazione dei dati mediante i tools introdotti nei capitoli precedenti.

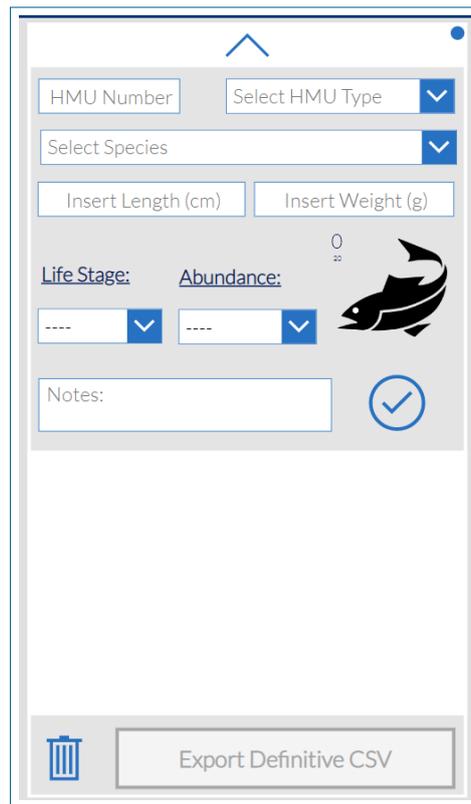
5.1 Ottimizzazione App di registrazione dato biologico

L'app per la raccolta dati biologici, che ha portato a definire un formato informatizzato standard per il file grezzo di campionamento biologico, ha visto una procedura di test e ottimizzazione in campo.

La prima versione è stata testata durante i campionamenti, svolti in Liguria, a marzo 2022, ed ha evidenziato alcune problematiche nella fluidità di utilizzo. Il funzionamento e il salvataggio del dato di campo sono stati verificati correttamente, mentre la registrazione è risultata ancora troppo lenta e macchinosa rispetto alla scrittura su foglio Excel.

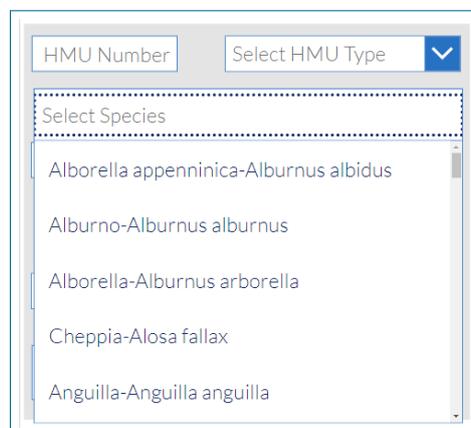
Il formato del dato di output, tra le due versioni dell'app, è rimasto inalterato, così come la schermata per l'inserimento dei dati generali del sito del rilievo; le modifiche maggiori sono state apportate sulla schermata di inserimento dati per gli esemplari campionati (interfaccia della prima versione in figura 5.1)

Nella prima versione, a seguito di ogni inserimento, venivano resettati automaticamente i campi di lunghezza, peso, stadio vitale e abbondanza dell'esemplare. La specie del pesce era selezionabile dall'elenco in figura 5.2 e il dover ricercare la selezione, tra le diverse voci disponibili, comportava l'aggiunta di operazioni superflue.



The screenshot shows a mobile application interface for recording biological data. At the top, there is a navigation arrow and a status indicator. Below this, the form includes several input fields: 'HMU Number' (text), 'Select HMU Type' (dropdown), 'Select Species' (dropdown), 'Insert Length (cm)' (text), and 'Insert Weight (g)' (text). There are also dropdown menus for 'Life Stage' and 'Abundance'. A fish icon is positioned to the right of the 'Abundance' dropdown. Below these fields is a 'Notes' text area and a blue checkmark icon. At the bottom of the screen, there is a trash can icon and a button labeled 'Export Definitive CSV'.

Figura 5.1: Schermata di registrazione dato biologico (prima versione app).



This screenshot shows a dropdown menu for selecting species. The menu is open, displaying a list of species names. The options are: 'Alborella appenninica-Alburnus albidus', 'Alburno-Alburnus alburnus', 'Alborella-Alburnus arborella', 'Cheppia-Alosa fallax', and 'Anguilla-Anguilla anguilla'. The dropdown is framed by a dashed border.

Figura 5.2: Elenco di selezione specie (prima versione app).

La suddetta configurazione sarebbe comparabile all'utilizzo del foglio Excel, in termini di velocità di inserimento, qualora tutti i pesci della stessa specie venissero registrati in successione, eventualità troppo spesso non verificata durante l'attività

di campo.

La problematica è stata risolta modificando l'interfaccia di inserimento e introducendo l'utilizzo dei toggle, descritti al punto 3.1.2, per selezionare i campi per cui evitare il reset. Questa modifica ha permesso di velocizzare notevolmente l'inserimento del dato a discapito di una leggera variabilità nel formato della specie inserita, dal momento che il nome della stessa specie (es. vairone) non sarà univoco per tutto il rilievo ma potrà variare leggermente (es. vai o vairo). L'aumento di variabilità è stata accettata in quanto la conversione dei nomi sarà semplicemente e rapidamente gestibile tramite il convertitore del tool di inserimento dati. La variabilità di inserimento del nome non deve essere inoltre vista come caratteristica unicamente negativa; la libertà di formato ha introdotto in certi casi versatilità nell'utilizzo dell'app.

In figura 5.3 è osservabile il confronto tra le due interfacce di inserimento.

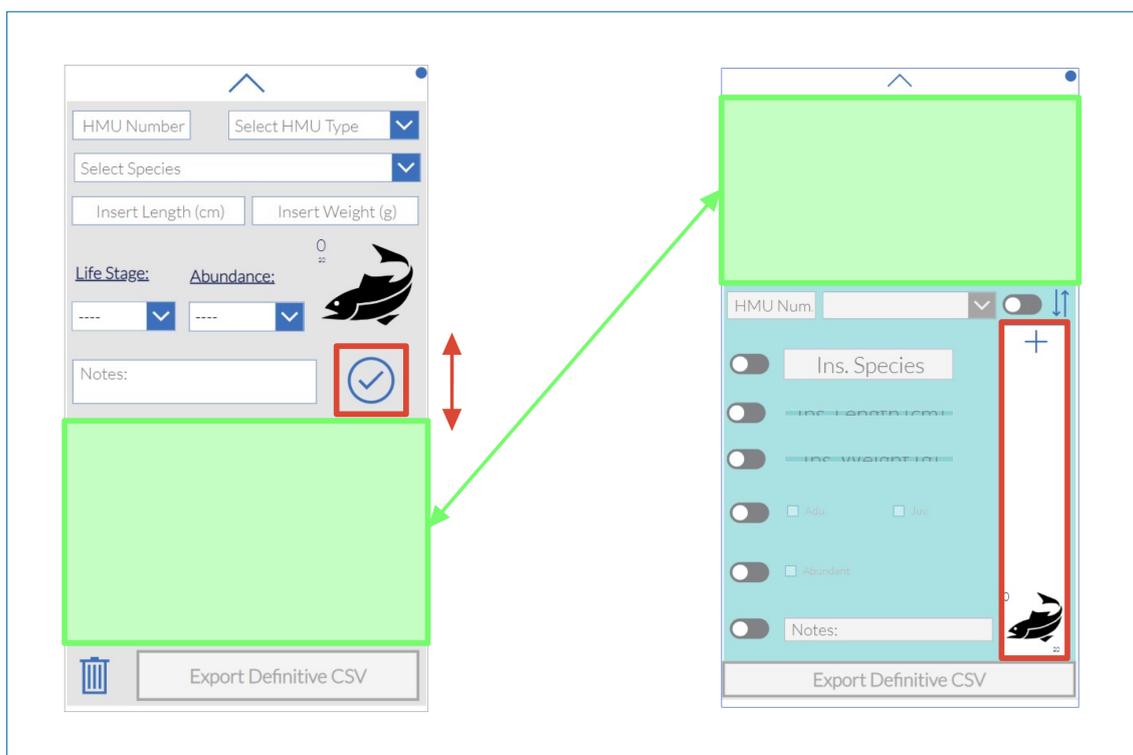


Figura 5.3: Confronto interfaccia di inserimento dati tra prima (sinistra) e seconda (destra) versione app.

Dopo la compilazione dei campi, nella versione a sinistra, l'inserimento era gestito dalla selezione del pulsante riquadrato; questo è risultato troppo piccolo e distante dalla tastiera aperta automaticamente dal dispositivo durante la compilazione, introducendo movimenti eccessivi che, oltre al rallentamento delle operazioni, ri-

sultavano scomodi e poco compatti.

L'inversione di posizione della galleria di visualizzazione ha permesso di compatte le operazioni di raccolta che ne hanno beneficiato in comodità e rapidità.

Tra la prima (marzo) e seconda (maggio) versione le altre funzionalità di modifica, salvataggio ed eliminazione del dato sono state mantenute tali.

5.2 Confronto dato grezzo campionato

Durante le attività di raccolta dati sono stati utilizzati come strumento di registrazione sia il foglio Excel compilato da mobile che l'App. Il confronto tra i file di campionati è riportato in figura 5.4.

DATO APP								DATO EXCEL							
ABUNDANCE	HMU_NUM	HMU_TYPE	INFO	LENGTH_cm	LIFE_STAGE	NOTE	SPECIE	WEIGHT_g	HMU_NUM	HMU_TYPE	FISH_NUM	FISH_TYPE	LENGTH	LIFE_STAGE	ABUNDANCE
	1	POOL	1_Tanaro	14.5	A		Scazzone		1	pool		scazz	145	a	
	1	POOL	2_Tanaro	13	A		Scazzone						130	a	
	1	POOL	3_Tanaro	13.8	A		Scazzone						138	a	
	1	POOL	4_Tanaro	14.2	A		Scazzone						142	a	
	1	POOL	5_Tanaro	27.3	A		Scazzone					farlio	273	a	
	2	BACKWATER	6_Tanaro				No fish		2	backwatwe		no			
	3	RIFFLE	7_Tanaro	11.7	J		Fario		3	riffle		farlio	117	j	
	3	RIFFLE	8_Tanaro	10.5	J		Fario					farlio	105	j	
	3	RIFFLE	9_Tanaro	11.6	A		Scazzone					scazz	116	a	
	3	RIFFLE	10_Tanaro	11	A		Scazzone						100	a	
	3	RIFFLE	11_Tanaro	13.2	A		Scazzone						132	a	
	4	POOL	12_Tanaro	11.4	A		Scazzone		4	pool		ibr marmo farlio	352	a	
	4	POOL	13_Tanaro	12.6	A		Scazzone						114	a	
	4	POOL	14_Tanaro	15.1	A		Scazzone						126	a	
	4	POOL	15_Tanaro	11.1	J		Scazzone						151	a	
	4	POOL	16_Tanaro	35.2	A		ibrido farlio X marmo						111	j	
	5	RAPID	17_Tanaro	24.5	A		ibrido farlio X marmo		5	rapid		ibrido marmorata farlio	245	a	
	6	POOL	18_Tanaro	13.5	J		Fario		6	pool		farlio	135	j	
	6	POOL	19_Tanaro	6	J		Scazzone					scazz	60	j	
	7	POOL	20_Tanaro	14.2	J		Fario		7	pool		farlio	142	j	
	7	POOL	21_Tanaro	13	J		Fario						130	j	
	7	POOL	22_Tanaro	34	A		Iridea					iridea	340	a	
	8	RIFFLE	23_Tanaro				No fish		8	riffle		no fish			
	9	RIFFLE	24_Tanaro	16.2	J		ibrido Fario Marmo		9	riffle		ibrido marmorata farlio	162	j	
	10	RIFFLE	25_Tanaro	15.2	J		Fario		10	riffle		farlio	152	j	
	10	RIFFLE	26_Tanaro	12.1	J		Fario					farlio	121	j	
	11	POOL	27_Tanaro	6.6	J		Scazzone		11	pool		scazz	66	j	
	11	POOL	28_Tanaro	6.3	J		Scazzone						63	j	
	12	RAPID	29_Tanaro				No fish		12	rapid		no fish			
	13	POOL	30_Tanaro	13.3	J		Fario		13	pool		farlio	133	j	
	13	POOL	31_Tanaro	11	J		Fario						110	j	
	13	POOL	32_Tanaro	6	J		Scazzone					scazz	60	j	

Figura 5.4: Confronto dato biologico campionato con apposita app (sinistra) e file Excel (destra).

Il file registrato con l'app, a fronte di velocità di inserimento comparabili, è risultato più strutturato del file registrato con Excel e le manipolazioni del dato, che dovranno essere eseguita prima del suo caricamento del database, sono decisamente più contenute. In questi termini, tanto più elevato sarà il numero di esemplari campionati, maggiori saranno i benefici apportati dall'utilizzo dell'app.

5.3 Elaborazioni dati di campo

I dati di campo rilevati sono stati elaborati con gli strumenti informatici a disposizione. Di seguito sono riportati gli output ottenuti.

Viene illustrato il caso del rilievo effettuato sul Roja (Liguria) il 10/03/2022, a portata pari a $1 \text{ m}^3/\text{s}$. Gli stessi risultati sono stati ricavati per ciascun campionamento effettuato.

5.3.1 Output SimStream-Web

I dati, raccolti seguendo le procedure della metodologia MesoHABSIM, sono stati elaborati in SimStream-Web. Dagli output forniti, per ciascun rilievo, specie selezionata e stadio vitale, sono state ricavate le schede riassuntive rappresentanti il mosaico di unità idro-morfologiche mappate, i risultati dei modelli biologici di assenza/presenza e presenza/abbondanza e le mappe di idoneità di habitat.

In allegato [A7-RojaSS10032022](#) è riportato l'esempio della schede ottenute per il fiume Roja (10/03/2022, $1 \text{ m}^3/\text{s}$), facendo riferimento alla sanguinerola (*Phoxinus phoxinus*). Data la sola presenza di 2 rilievi eseguiti in altrettante condizioni di deflusso, non è stato possibile elaborare la curva habitat-portata.

5.3.2 Output "Hydro Bio Tool"

I dati, previa elaborazioni in SimStream-Web, sono stati elaborati con "Hydro Bio Tool", ottenendo le schede riassuntive in allegato [A8-HydrRoja10032022](#), [A9-BioRoja10032022](#), [A-ResAJRoja10032022](#), [A-ResAllRoja10032022](#), descritte al capitolo [2.4.1](#).

La consultazione di queste schede permetterà di ricavare informazioni riassuntive, sito specifiche, relative ai descrittori fisici dei meso-habitat e alla quantità e tipologia (specie e stadio vitale) degli esemplari campionati.

5.4 Analisi "MesoHABSIM DataBase"

Dalle schede riportate al punto 5.3.2 sono ricavabili informazioni sito-specifiche, mentre, dopo l'inserimento di tutti dati idro-morfologici e biologici all'interno del database, sono state ricavate informazioni globali riferite ai siti campionati.

In Tabella 5.1 è riportato il numero di rilievi analizzati, unità idro-morfologiche mappate, specie presenti nel database ed esemplari campionati; a partire da febbraio 2009, sono stati effettuati 117 rilievi in 9 regioni italiane (principalmente concentrati al Nord); sono stati registrati e mappati 1303 meso-habitat diversi all'interno della quale sono stati ritrovati più di 19'000 pesci, di 38 specie diverse.

Tabella 5.1: Informazioni generali riassuntive per rilievi idro-morfologici e biologici.

REGIONI	9
N° RILIEVI	117
N° HMU	1303
N° SPECIE	38
N° INDIVIDUI	19130

Nelle tabelle 5.2 sono riportate le condizioni massime (Aurino, $14.5 m^3/s$) e minime (Albedosa, $0.001 m^3/s$) di portata in cui sono stati eseguiti i campionamenti, il numero minimo (Roccia, 1 HMU) e massimo (60 HMUs) di unità campionate per rilievo, il numero di campionamenti quantitativi (106) e semi-quantitativi (11) effettuati e il numero minimo (Vallone d'Elva, 0) e massimo (Natisone, 856) di esemplari individuati per rilievo.

In tabella 5.3 è osservabile come la regione dove sono stati effettuati più rilievi sia il Piemonte (61 rilievi per 297 HMUs campionate), mentre in Emilia Romagna è stato eseguito un solo rilievo (10 HMU campionate).

Tabella 5.2: a) Condizioni massime e minime di portata rilievi. b) Numero minimo e massimo di unità campionate. c) Numero di campionamenti quantitativi e semi-quantitativi effettuati. d) Numero minimo e massimo di esemplari individuati.

a)	PORTATA (m3s)	Q min	Sito	Data
		0.001	Albedosa	24/08/2009
		Q max	Sito	Data
		14.5	Aurino	19/09/2018

b)	N° UNITA' CAMPIONATE	N° min	Sito	Data	Q (m3s)
		1	Roccia	24/03/2009	0.005
		N° max	Sito	Data	Q (m3s)
		60	Ayasse	22/11/2010	0.3

c)	N° CAMPIONAMENTI	
	QUANTI	SEMI
	106	11

d)	N° ESEMPLARI CAMPIONATI	N° min	Sito	Data	Q (m3s)
		0	Vallone d'Elva	01/04/2009	0.037
		N° max	Sito	Data	Q (m3s)
		856	Natisone	06/09/2021	0.422

Tabella 5.3: Numero di rilievi e unità campionate per regione.

REGIONE	N° RILIEVI	N° HMU
ABRUZZO	2	61
EMILIA ROMAGNA	1	10
FRIULI VENEZIA GIULIA	16	225
LIGURIA	22	267
LOMBARDIA	5	156
PIEMONTE	61	297
TRENTINO ALTO ADIGE	4	88
VALLE D'AOSTA	4	181
VENETO	2	18
Totale	117	1303

Sono poi stati valutati il numero di esemplari campionati (totale, adulti e giovani) per ciascuna tipologia di unità idro-morfologica e il numero di esemplari campionati per tipologia e numero di unità campionate. Come osservabile in tabella 5.4, il maggior numero di esemplari totali, adulti e giovani, è stato ritrovato in unità POOL (rispettivamente 7165, 3195 e 3970) che è risultato allo stesso tempo la tipologia di unità maggiormente rappresentata, mentre il numero di esemplari ritrovati per tipo di meso-habitat, rapportato al numero di unità campionate ha evidenziato come nei GLIDE siano state ritrovate le densità più alte di esemplari (rispettivamente 11.75 esemplari/unità adulti, 12.55 esemplari/unità giovani e 24.30 esemplari/unità totali).

Tabella 5.4: Analisi esemplari campionati per tipologia di unità.

HMU TYPE	N°	N° ADULT	ADULT/TOT_hmu	N° JUVENILE	JUV/TOT_hmu	TOTAL	TOT/TOT_hmu
POOL	318	3195	10.05	3970	12.48	7165	22.53
GLIDE	228	2680	11.75	2861	12.55	5541	24.30
RIFFLE	316	2080	6.58	2264	7.16	4344	13.75
RAPID	151	565	3.74	431	2.85	996	6.60
BACKWATER	71	145	2.04	445	6.27	590	8.31
AQUAT_VEG	18	145	8.06	32	1.78	177	9.83
PLUNGE_POOL	57	90	1.58	44	0.77	134	2.35
STEP	104	45	0.43	79	0.76	124	1.19
CASCADE	32	20	0.63	9	0.28	29	0.91
WETLAND	2	5	2.50	15	7.50	20	10.00
ARTIF_ELEM	2	5	2.50	2	1.00	7	3.50
ISOLATED_POND	3	0	0.00	3	1.00	3	1.00
WATERFALL	1	0	0.00	0	0.00	0	0.00
	1303	8975		10155		19130	

Si sono poi esaminate le specie maggiormente rappresentate in totale e suddivise per stadio vitale. Si sono rappresentate solo quelle per cui sono stati individuati almeno 50 esemplari. Per una migliore rappresentazione grafica si è preferito separare le specie per cui sono stati campionati tra i 50 e 500 individui (fig. 5.5), e quelle che superano tale soglia (figura 5.6); da quest'ultimo grafico è evidenziato come il vairone (*Telestes muticellus*, 6189 esemplari), sia stato ritrovato più del triplo delle volte della trota fario (*Trota fario*, 1810 esemplari), seconda specie più campionata.

Si sono poi rappresentati il numero di esemplari individuati suddivisi per tipologia di unità e rapportati al numero di unità campionate per ciascuna tipologia. Si sono analizzati solo quei tipi di meso-habitat rappresentati almeno 50 volte e le specie rappresentate da almeno 50 esemplari. Per una migliore rappresentazione grafica si è preferito separare le specie per cui sono stati campionati tra i 50 e 500 individui (fig. 5.7), e quelle che superano tale soglia (figura 5.8). Grazie a questi grafici è possibile osservare la tipologia di unità in cui è stato campionato il maggior numero di esemplari per specie e numero di unità. Si riporta come esempio il vairone (*Telestes muticellus*), in cui, dei 6189 esemplari campionati, 2048 erano in GLIDE (in totale 228 unità di questo tipo), si può quindi affermare

come, in ogni GLIDE, sono stati campionati in media, circa 9 vaironi). Si vuole sottolineare come questo dato sia solamente indicativo ma può allo stesso tempo dare un'indicazione di massima dell'idoneità di una certa tipologia di habitat per un determinato stadio vitale).

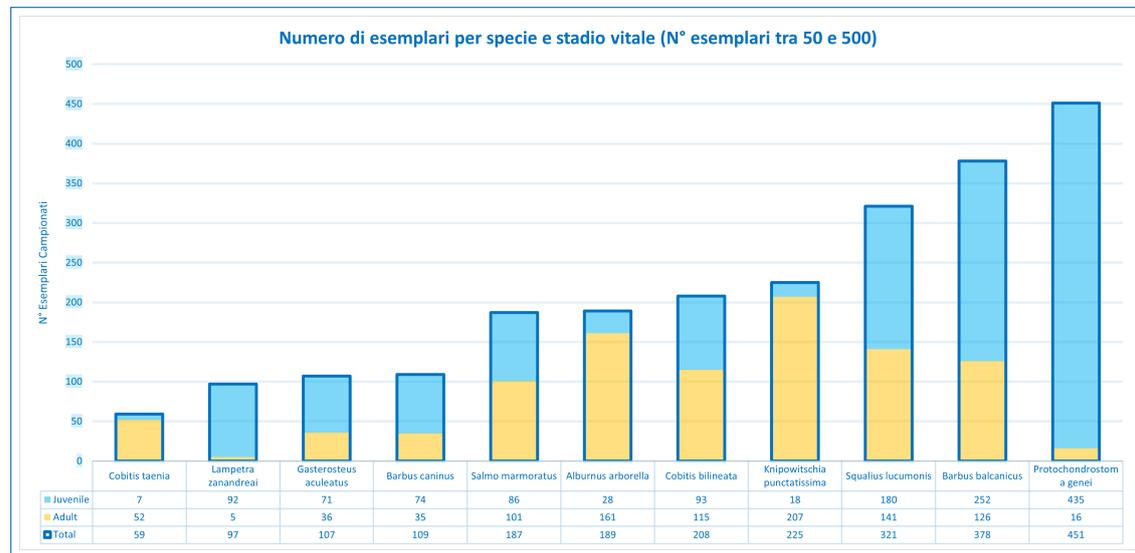


Figura 5.5: Numero di esemplari campionati per specie e stadio vitale (N° esemplari tra 50 e 500).

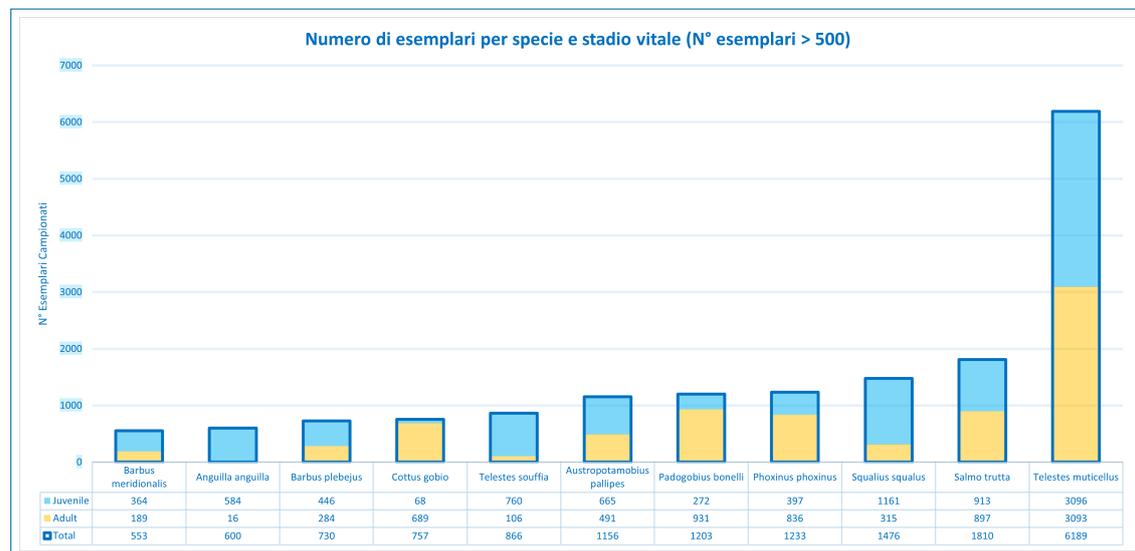


Figura 5.6: Numero di esemplari campionati per specie e stadio vitale (N° esemplari maggiore di 500).

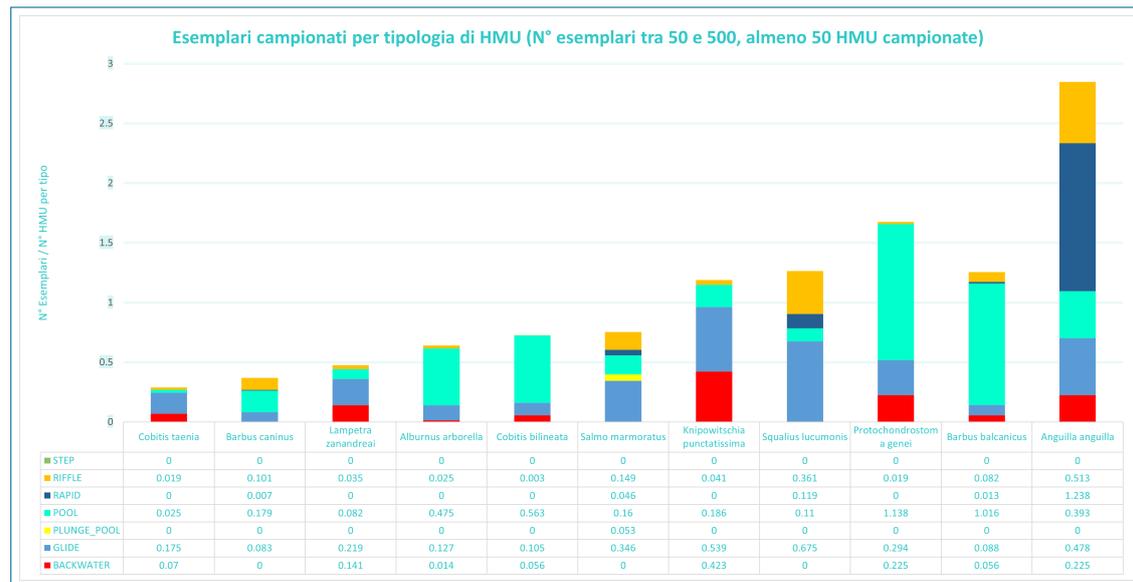


Figura 5.7: Numero di esemplari campionati per specie, per ciascuna tipologia di unità (N° HMU maggiori di 50, N° esemplari tra 50 e 500).

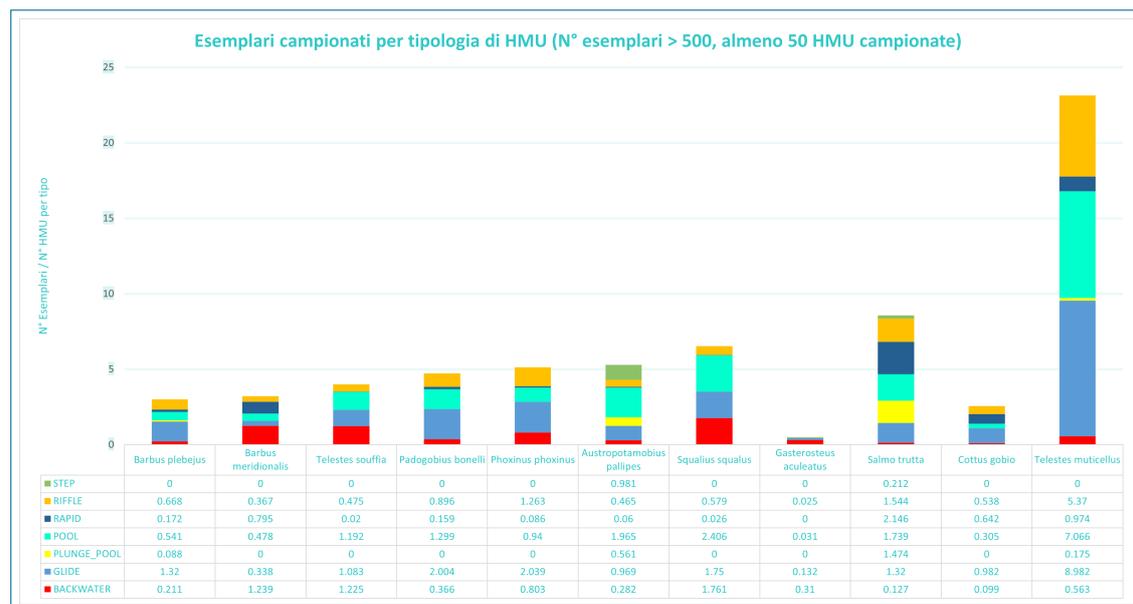


Figura 5.8: Numero di esemplari campionati per specie, per ciascuna tipologia di unità (N° HMU maggiori di 50, N° esemplari maggiori di 500).

Sono poi state realizzate mappe georeferenziate per la rappresentazione spaziale del dato campionato.

Nell'allegato **A12-MapQS** è riportata la distribuzione spaziale dei campionamenti effettuati, classificati in funzione della tipologia (quantitativa, Q, o semi-quantitativa, S).

Nell'allegato **A13-IndHMU** è osservabile la mappa dei rilievi classificati in funzione del numero di individui per numero di meso-habitat campionati.

Nell'allegato **A14-IndAJAll** è presente la mappa dei rilievi con annesso numero di esemplari classificati per stadio vitale.

Nell'allegato **A15-ElRel** sono infine riportati i siti campionati con il rispettivo CO-DE univoco e l'identificativo ID, utile per la loro visualizzazione negli allegati.

Sono poi state prodotte le mappe di ritrovamento di specie per ciascuna regione analizzata; in figura 5.9 sono evidenziate, tra le regioni in cui è stato eseguito almeno un campionamento, quelle in cui è stato ritrovato almeno un esemplare (esempio trota fario). È fondamentale sottolineare come l'assenza della specie in alcune regioni della mappa non debba essere intesa come la sua assenza in quel territorio; la fig. 5.9 indicherà solamente se la specie, nei campionamenti eseguiti, è stata ritrovata o meno tra gli esemplari campionati.

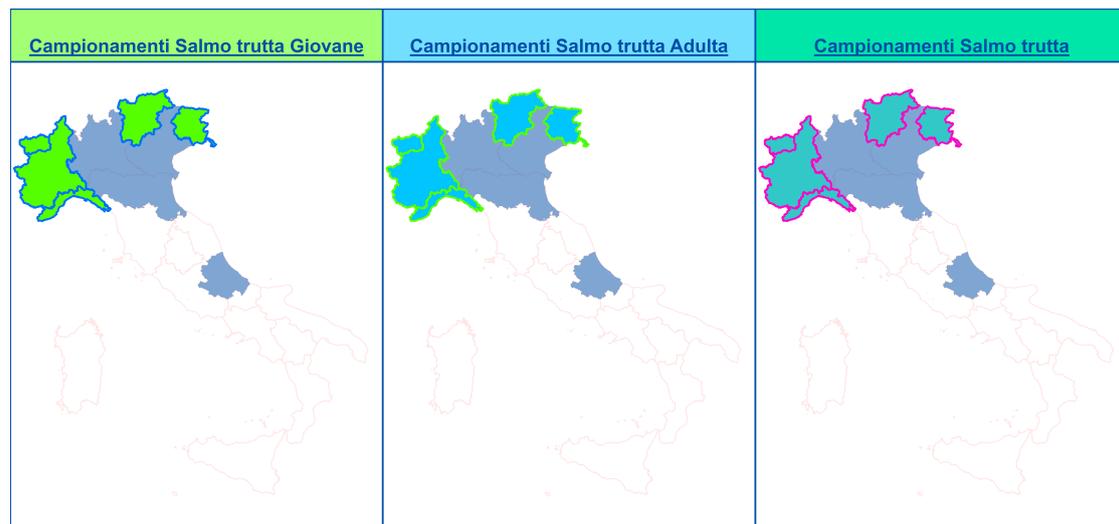


Figura 5.9: Regioni con campionamenti di trota fario (*Salmo trutta*) adulta, giovane e totale per regioni campionate.

5.5 Preferenza di habitat per ammocete di Lampetra zanandreai

Nell'allegato [A16-ExtLamZan](#) è riportata la scheda che, fornita come input all'apposito script R, permetterà la generazione del modello di distribuzione di specie. In tabella 5.5 sono osservabili i siti analizzati e il numero di ammoceti individuati in ciascuno di essi.

Tabella 5.5: Analisi esemplari campionati per tipologia di unità.

RIVER	DISCHARGE_m3s	DATE	N° HMU	N° Lz_amm
BELBO	0.15	25/06/2020	14	1
GHIANDONE	0.577	31/07/2020	6	34
GHIANDONE	1	04/03/2021	4	40
ORCO	3	10/10/2019	12	1
PELLICE	2	12/07/2021	6	1
			42	77

La valutazione dell'importanza delle variabili, tramite algoritmo BORUTA, ha permesso di selezionare la seguenti 10 variabili per il modello parsimonioso: MICROLITHAL, PSAMMAL, AKAL, SHALLMARG, CV60_75, SLOPE, MESOLITHAL, D15_30, CV_15, CV75_90. In figura 5.10 è osservabile il ranking di importanza valutata dall'algoritmo. Sono state selezionate come variabili indipendenti del modello, i descrittori relativi alla presenza o assenza di zone marginali nell'HMU, pendenza della superficie libera, classi di substrato, velocità (CV) e profondità della corrente (D). Per la descrizione delle seguenti variabili si rimanda alla tabella 1.1.

Il valore settato per gli iper-parametri è il seguente:

- **ntree**: 600, valore superiore al doppio rispetto al numero di alberi (240) per cui viene stabilizzato l'OOB-Error (come osservabile in figura 5.11).
- **mtry**: 3, radice quadrata del numero di variabili indipendenti.
- **samplesize**: 42, totale delle osservazioni di input (ricavato da processo iterativo per la determinazione del valore che minimizza l'OOB-Error).

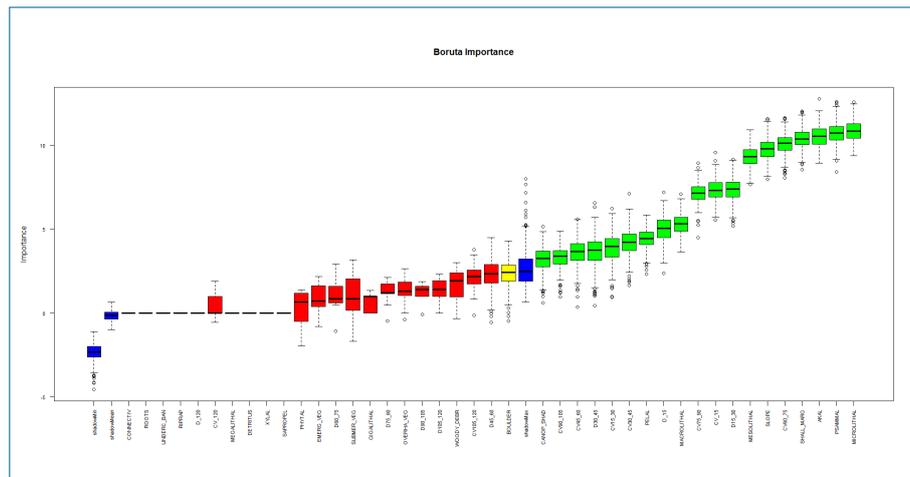


Figura 5.10: Importanza variabili da algoritmo BORUTA.

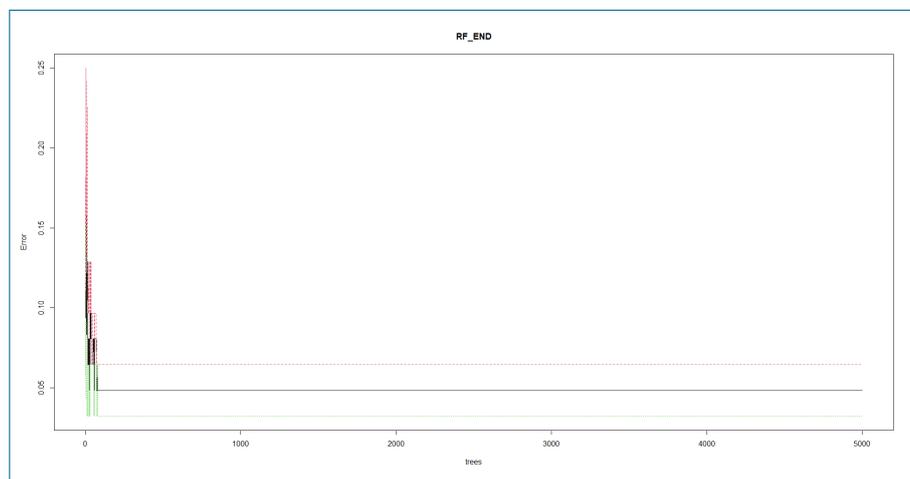


Figura 5.11: Stabilizzazione dell' E_{OOB} in funzione del numero di alberi, valore di stabilizzazione: 240 alberi.

Le performance dei parametri sono state valutate in termini di accuratezza (accuracy), sensibilità (sensitivity), specificità (specificity) e TSS (true skill statistics); i valori, riportati in tabella 5.6, sono indicatori di ottime performance predittive.

Tabella 5.6: Performance predittive modello statistico *Lampreda padana*.

Accuracy	Sensitivity	Specificity	TSS
95.2%	96.8%	93.6%	90.3%

Dall'analisi delle Partial Dependence Plots (PDP), osservabili in figura 5.12 e ordinate in ordine di importanza delle variabili predittive, è osservabile l'effetto che i singoli descrittori ambientali hanno sulla probabilità di presenza (p.p.) della *Lampetra zanandrea*. Risulta evidente come le variabili che maggiormente influenzano la presenza della specie siano tipologie di substrato; PSAMMAL e AKAL risultano positivamente correlati alla p.p., mentre MICROLITHAL, la variabile più importante, è correlata negativamente alla p.p.. La presenza di zone marginali all'interno dei mesohabitat caratterizzate da acque poco profonde (15-30 cm) e basse velocità (0-15 cm/s) aumenta la probabilità di presenza dell'ammocete. La maggiore distribuzione di zone a velocità di corrente più elevata (60-75 cm/s e 75-90 cm/s) diminuiscono la p.p., così come gradienti più elevati e substrati a granulometria più grossolana (MESOLITHAL). L'analisi statistica realizzata mediante l'algorit-

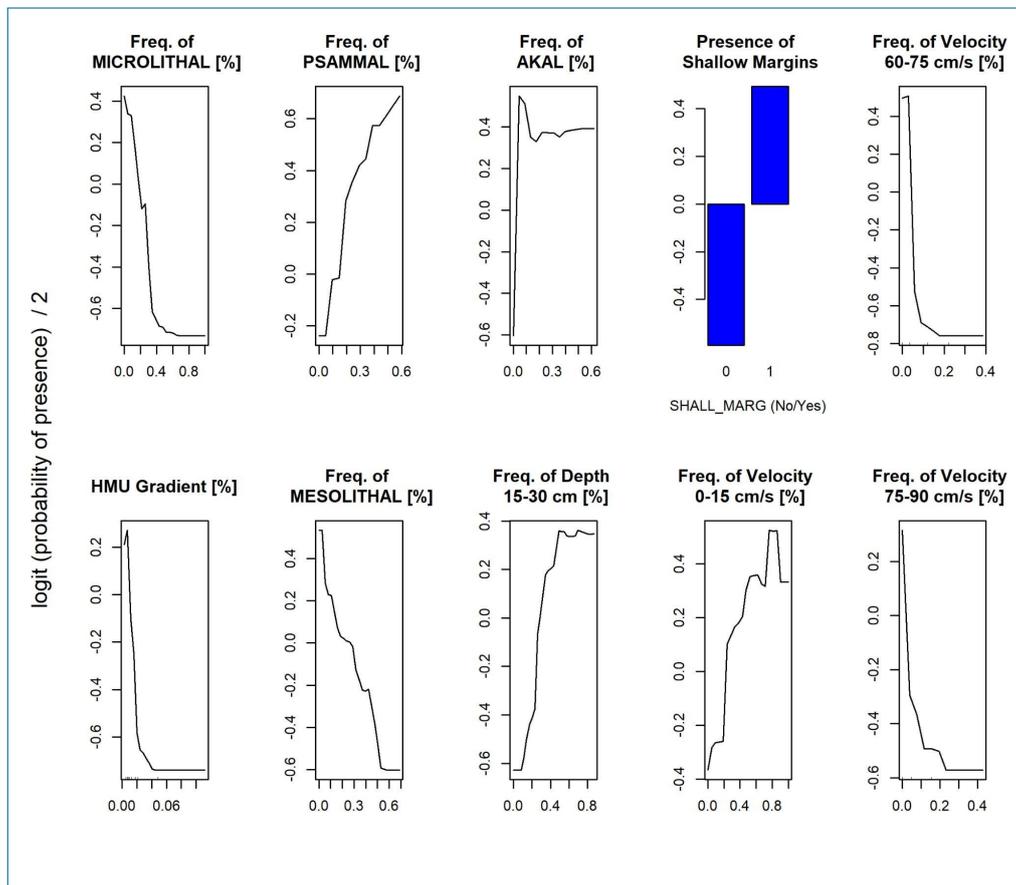


Figura 5.12: *Partial Dependence Plots per il modello parsimonioso oversampled per ammocete di lampreda padana.*

mo RF ha quindi permesso di ricavare modelli ecologicamente consistenti per la specie considerata, evidenziando come gli ammoceti prediligano generalmente zone a bassa velocità e profondità di corrente e substrati fini.

Si vuole sottolineare il fatto che con il campionamento di nuovi dati relativi alla specie in analisi sarà consigliabile un'ulteriore validazione del modello; allo stesso tempo però i risultati ottenuti sono consistenti con quelle che si ipotizza possano essere le preferenze di habitat per la specie in analisi.

Conclusioni

Per la corretta valutazione della disponibilità spazio-temporale di habitat è necessario disporre di dati opportunamente formattati e correttamente georeferenziati. Allo stesso modo per la generazione di nuovi modelli di distribuzione di specie sono necessari dati idro-morfologico e biologici, univocamente legati in base al meso-habitat di campionamento, che permettono l'associazione dell'habitat fisico alle risposte ecologiche delle comunità animali (nel presente elaborato ci si è concentrati sulla fauna ittica).

I dati, raccolti a partire dal 2009, sono stati organizzati in un database a scala nazionale, ed unico nel suo genere, che ha richiesto la definizione di una formattazione standard al fine di legare tra loro anche rilievi cronologicamente distanti. L'organizzazione di una tipologia di dato con finalità così specifiche ha richiesto la creazione di appositi strumenti per l'elaborazione del dato proveniente dai rilievi ("Hydro Bio Tool") e per la gestione del database stesso.

La mancanza di un software per la raccolta dati in campo ha portato alla realizzazione dell'apposita App descritta alla sezione 3.1. L'introduzione di uno strumento specifico per la raccolta di dati biologici ha permesso di velocizzarne la registrazione, mantenendo allo stesso tempo un'ottima capacità di adattamento nei confronti della variabilità dei dati raccolti, garantendo inoltre il salvataggio periodico sul cloud, riducendo quindi al minimo la possibilità di una sua perdita. Lo strumento sviluppato, in aggiunta ai sistemi informatizzati ben consolidati per i rilievi idro-morfologici, come MapStream, ha evidenziato come gli strumenti informatici siano sempre più imprescindibili per lo sviluppo di scienze come l'ecologia fluviale.

Grazie all'organizzazione del dato in un'unica banca dati viene quindi agevolata la validazione dei modelli distribuzione di specie esistenti e la creazione di nuovi modelli statistici, andando così ad estendere il campo di utilizzo della metodologia MesoHABSIM verso un numero sempre crescente di specie.

Si vuole ancora sottolineare come un continuo lavoro di aggiornamento e ampliamento del database consentirà la creazione di modelli sempre più robusti, sviluppando così un mezzo sempre più potente per la salvaguardia della vita all'interno dell'ambiente lotico.

La generazione di nuovi modelli biologici viene implementata grazie all'apprendi-

mento statistico, in particolare attraverso la tecnica del Random Forest. Nel presente lavoro è stato sviluppato un primo modello quantitativo di presenza per l'ammocete di lampreda padana (*Lampetra zanandreae*). Il modello risultante è il migliore, in termini di performance predittive, tra tutti i possibili modelli generabili con diverse combinazioni dei predittori; la generazione del modello parsimonioso ha permesso così di isolare le variabili che influenzano maggiormente la probabilità di presenza della specie considerata e di fornirne un'interpretazione ecologica. Grazie alle tecniche delle Partial Dependence Plots è stato possibile visualizzare la relazione parziale tra la probabilità di presenza e i descrittori ambientali del modello, evidenziando i trend non lineari che legano le due variabili. Dall'analisi della preferenza di habitat per la lampreda è emerso come questa prediliga fondali a substrato fine e a bassa profondità e velocità di corrente.

Bibliografia

- [1] Ken D Bovee. *A guide to stream habitat analysis using the instream flow incremental methodology*, volume 1. Western Energy and Land Use Team, Office of Biological Services, Fish and . . . , 1982.
- [2] Leo Breiman. Random forests. *Machine learning*, 45(1):5–32, 2001.
- [3] Robin Genuer and Jean-Michel Poggi. Random forests. In *Random Forests with R*, pages 33–55. Springer, 2020.
- [4] Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani. *An introduction to statistical learning*, volume 112. Springer, 2013.
- [5] Miron B Kursa and Witold R Rudnicki. Feature selection with the boruta package. *Journal of statistical software*, 36:1–13, 2010.
- [6] Giovanni Negro, Stefano Fenoglio, Emanuele Quaranta, Claudio Comoglio, Isabella Garzia, and Paolo Vezza. Habitat preferences of italian freshwater fish: A systematic review of data availability for applications of the mesohabsim model. *Frontiers in Environmental Science*, page 305, 2021.
- [7] Piotr Parasiewicz. Mesohabsim: A concept for application of instream flow models in river restoration planning. *Fisheries*, 26(9):6–13, 2001.
- [8] Piotr Parasiewicz. The mesohabsim model revisited. *River Research and Applications*, 23(8):893–903, 2007.
- [9] Piotr Parasiewicz, Joseph N Rogers, Paolo Vezza, Javier Gortázar, Thomas Seager, Mark Pegg, Wiesław Wiśniewolski, and Claudio Comoglio. Applications of the mesohabsim simulation model. *Ecohydraulics: an integrated approach*, pages 109–124, 2013.
- [10] Piotr Parasiewicz, Kathleen Ryan, Paolo Vezza, Claudio Comoglio, Thomas Ballesterio, and Joseph N Rogers. Use of quantitative habitat models for establishing performance metrics in river restoration planning. *Ecohydrology*, 6(4):668–678, 2013.

-
- [11] N Leroy Poff and Julie KH Zimmerman. Ecological responses to altered flow regimes: a literature review to inform the science and management of environmental flows. *Freshwater biology*, 55(1):194–205, 2010.
- [12] Philipp Probst, Marvin N Wright, and Anne-Laure Boulesteix. Hyperparameters and tuning strategies for random forest. *Wiley Interdisciplinary Reviews: data mining and knowledge discovery*, 9(3):e1301, 2019.
- [13] M Rinaldi, B Belletti, F Comiti, L Nardi, L Mao, and M Bussetтини. Sistema di rilevamento e classificazione delle unità morfologiche dei corsi d’acqua (sum)–versione aggiornata 2016. *Roma, Italy: Istituto Superiore per la Ricerca e la Protezione Ambientale (ISPRA). Manuali e Linee Guida*, 132(2016):178, 2016.
- [14] Pang-Ning Tan, Michael Steinbach, and Vipin Kumar. *Introduction to data mining*, chapter 6. Pearson Education India, 2016.
- [15] P Vezza, P Parasiewicz, M Rosso, and C Comoglio. Defining minimum environmental flows at regional scale: application of mesoscale habitat models and catchments classification. *River research and applications*, 28(6):675–792, 2012.
- [16] P Vezza, A Zanin, and P Parasiewicz. Manuale tecnico-operativo per la modellazione e la valutazione dell’integrità dell’habitat fluviale. *ISPRA–Manuali e Linee Guida*, 154:2017, 2017.
- [17] Paolo Vezza, Rafael Muñoz-Mas, Francisco Martinez-Capel, and A Mouton. Random forests to evaluate biotic interactions in fish distribution models. *Environmental Modelling & Software*, 67:173–183, 2015.
- [18] Paolo Vezza, P Parasiewicz, Olle Calles, M Spairani, and Claudio Comoglio. Modelling habitat requirements of bullhead (*cottus gobio*) in alpine streams. *Aquatic sciences*, 76(1):1–15, 2014.
- [19] Paolo Vezza, Piotr Parasiewicz, Michele Spairani, and Claudio Comoglio. Habitat modeling in high-gradient streams: the mesoscale approach and application. *Ecological Applications*, 24(4):844–861, 2014.