
SEZIONE

METODOLOGIE E DATABASE



Gruppo di lavoro

Questo progetto è stato possibile grazie all'impegno, professionalità e passione di molte persone che hanno collaborato con la LIPU, a titolo professionale o di volontariato, nella raccolta e nell'elaborazione dei dati.

Coordinamento generale:

Patrizia Rossi

LIPU

Via Trento, 49 - 43122 Parma - Telefono 0521 273043 - E-mail: patrizia.rossi@lipu.it

Gruppo di lavoro LIPU: Patrizia Rossi (coordinatore generale), Laura Silva (segreteria e coordinamento generale), Ambra De Lorentiis (inserimento dati).

Hanno collaborato anche Claudio Celada, Marco Gustin, Giovanni Albarella, Giorgia Gaibani.

Hanno collaborato:

FaunaViva

Viale Sarca, 78 - 20125 Milano - Telefono 02 36591561

Gruppo di lavoro FaunaViva: Elisabetta de Carli, Lia Buvoli, Gianpiero Calvi, Paolo Bonazzi.

Ha inoltre collaborato: Lorenzo Fornasari e Jacopo Tonetti.

D.R.E.Am. Italia Soc. Coop. Agr. For.

Via Garibaldi, 3 - Pratovecchio (AR) - Telefono 0575 529514

Gruppo di lavoro D.R.E.Am.: Guido Tellini Florenzano, Simonetta Cutini, Tommaso Campedelli, Guglielmo Londi.

INDICE

1.IL METODO DI CENSIMENTO.....	4
1.1.TECNICA DI RILEVAMENTO.....	4
1.2.COMPILAZIONE DELLA SCHEDA DI RILEVAMENTO.....	4
1.3.DISTRIBUZIONE DEI PUNTI D'ASCOLTO.....	8
2.STRUTTURA DELLA BANCA DATI DEL PROGETTO MITO2000 E PROCEDURE DI VALIDAZIONE.....	11
2.1.LA STRUTTURA DELLA BANCA DATI.....	11
2.2.ARCHIVIO DEGLI ERRORI.....	16
3.SELEZIONE DEI DATI PER L'ANALISI.....	21
4.METODI E SOFTWARE.....	23
5.CALCOLO DELLE TENDENZE DI POPOLAZIONE DELLE SPECIE.....	26
5.1.REGIONI.....	26
5.2.ITALIA.....	27
6.CALCOLO DEGLI INDICI AGGREGATI.....	29
6.1.REGIONI.....	29
6.2.ITALIA	30
7.BIBLIOGRAFIA.....	31

1.IL METODO DI CENSIMENTO

1.1. TECNICA DI RILEVAMENTO

La tecnica di rilevamento prescelta è quella dei punti di ascolto senza limiti di distanza (Blondel *et al.*, in Ralph & Scott 1981). La durata di ciascun punto è di 10 minuti (Fornasari *et al.*, 2002). I punti sono stati eseguiti compatibilmente con l'inizio della stagione riproduttiva in base alle condizioni locali (latitudine, quota delle stazioni), indicativamente in maggio e giugno. I rilevamenti sono stati iniziati poco dopo l'alba (evitando giornate con condizioni meteorologiche sfavorevoli, ossia con vento forte o precipitazioni intense), visitando ogni stazione una sola volta per anno. Si richiede ai rilevatori di distinguere tra gli uccelli visti e sentiti entro un raggio di 100 m e oltre tale raggio.

Allo scopo di poter trasformare il dato relativo al numero d'individui rilevati in stima del numero di coppie nidificanti, ai rilevatori si richiede di unire al numero di uccelli osservati alcuni semplici codici relativi alle caratteristiche dell'osservazione.

L'archiviazione dei dati e la trasformazione del numero di individui in coppie viene effettuata tramite un software appositamente realizzato denominato AEGITHALOS.

1.2. COMPILAZIONE DELLA SCHEDA DI RILEVAMENTO

Indicazioni generali

Per ciascun punto è richiesto di riportare sul fronte della scheda (Figura 1.2): 1. il nome del rilevatore; 2. il nome di eventuali osservatori; 3. la particella UTM (10x10 km) nella quale si sta eseguendo uno dei punti d'ascolto del programma randomizzato; 4. il numero della stazione UTM (1x1 km); 5. le coordinate del punto in cui si rileva, ricavate dalla lettura di un GPS, e la proiezione utilizzata; 6. un toponimo contenuto all'interno del quadrato di un km di lato in cui si sta rilevando; 7. la data; 8. l'ora (legale) d'inizio del censimento; 9. le condizioni meteorologiche, in particolare la copertura del cielo e il vento; 10. la risposta alla domanda "hai già visitato almeno una volta questa stazione"; 11. la risposta alla domanda "hai fatto il punto d'ascolto nella stessa posizione esatta?".

Per ciascun punto viene inoltre richiesto di indicare sul retro della scheda (Figura 1.1): 1. le caratteristiche ambientali nel raggio di 100m secondo le Categorie CORINE Land Cover; 2. eventuali note.

categorie CORINE Land Cover (INDICARE L'ESTENSIONE PERCENTUALE NEL RAGGIO DI 100 m: 1%, 5%, 10% ecc.)			
1 Superfici artificiali		3 Boschi ed aree semi-naturali	
1.1 Zone edificate*		3.1 Boschi	
1,1,1 Aree edificate continue		3.1,1 Boschi di latifoglie	
1,1,2 Aree edificate discontinue		3,1,1,1 Leccio o/e sughera	
1.2 Infrastrutture		3,1,1,2 Querce caducifoglie (cerro, roverella, farnetto, rovere, farnia)	
1,2,1 Aree commerciali e industriali		3,1,1,3 Latif. mesofile e mesotermofile (acero-frassino, carpino nero-orniello)	
1,2,2 Reti viarie e ferroviarie ed aree associate		3,1,1,4 Castagno	
1,2,3 Porti		3,1,1,5 Faggio	
1,2,4 Aeroporti		3,1,1,6 Specie igrofile (salici, pioppi, ontani)	
1.3 Terreni artefatti		3,1,1,7 Latifoglie non native	
1,3,1 Aree estrattive e minerarie		3.1,2 Boschi di conifere	
1,3,2 Discariche		3,1,2,1 Pini mediterranei e cipressete	
1,3,3 Cantieri		3,1,2,2 Pini montani e oromediterranei (p.nero, p.laricio, p.silvestre, p.loricato)	
1.4 Aree di verde attrezzato		3,1,2,3 Abete bianco e/o abete rosso	
1,4,1 Parchi urbani		3,1,2,4 Larice e/o pino cembro	
1,4,2 Impianti sportivi		3,1,2,5 Conifere non native	
2 Zone agricole		3,1,3 Boschi misti di conifere e latif.	
2.1 Seminativi*		3.2 Associazioni arbustive o erbacee	
2,1,x Seminativi		3.2,1 Aree a pascolo naturale e praterie*	
2,1,3 Risaie		3,2,1,1 Praterie continue	
2.2 Colture permanenti		3,2,1,2 Praterie discontinue con arbusti e/o alberi sparsi	
2,2,1 Vigneti		3.2,2 Arbusteti	
2,2,2 Frutteti (anche minori)		3.2,3 Arbusteti a sclerofille	
2,2,3 Oliveti		3,2,3,1 Macchia alta	
2,2,4 Altre colture permanenti (arboricoltura da legno)		3,2,3,2 Macchia bassa e garighe	
2.3 Prati permanenti*		3.3 Aree aperte con vegetazione scarsa o nulla	
2,3,1 Prati permanenti (prati a sfalcio)		3,3,1 Spiagge, dune e piane sabbiose	
2.4 Aree agricole eterogenee		3,3,2 Rocce nude, falesie, rupi, affioramenti (senza vegetazione)	
2,4,1 Colture stratificate (colture annuali associate a colture arboree)		3,3,3 Aree con vegetazione rada (calanco, pietraia)	
2,4,2 Mosaici agrari (coltivi annuali, pascoli e colture arboree)		3,3,4 Aree percorse da incendi	
2,4,3 Mosaici di colture agrarie e vegetazione naturale (anche incolti)		3,3,5 Ghiacciai e nevai perenni	
2,4,4 Aree agro-forestali (pascoli o colture sotto copertura arborea)		5 Corpi d'acqua	
4 Zone umide		5.1 Acque interne (alveo incluso)	
4.1 Zone umide interne		5,1,1 Corsi d'acqua e canali	
4,1,1 Paludi		5,1,2 Laghi, stagni ed altri bacini	
4,1,2 Torbiere		5.2 Acque marine	
4.2 Zone umide costiere		5,2,1 Lagune costiere	
4,2,1 Paludi salmastre		5,2,2 Estuari	
4,2,2 Saline		5,2,3 Acque costiere	
4,2,3 Fasce intertidali		Portata del traffico	
		strade assenti	a
		sterrate e poderali	b
		asfaltate a bassa circolazione	c
		asfaltate ad alta circolazione	d
		superstrade e autostrade	e
		Colture dominanti	
		mais	a
		altri cereali	b
		ortaggi	c
		girasole	d
		barbabetola	e
		foraggere	m
		serre	f
		terreno arato	g
		altro (specificare)	h
		Tipologie lotiche	
		Torrenti	a
		Riali	b
		Fiumi	c
		Fossi	d
		Canali e rogge	e
		Fiumare	f
		Tipologie lentiche	
		Lanche e stagni	a
		Laghi	b
		Valli e vasche di colmata	c
		Invasi di ritenuta	d
		Tipologia della palude	
		Tifeti	a
		Fragmiteti	b
		Magnocariceti	c
		Acquitrini e cariceti	d
NOTE			

Figura 1.2: La scheda utilizzata per i rilevamenti di campagna (retro).

Dati ornitici

Sulla scheda sono presenti tre colonne per la raccolta dei dati relativi all'avifauna, in cui viene chiesto di indicare: 1. il nome comune della specie rilevata; 2. gli individui osservati entro i 100 metri dal punto di rilevamento; 3. quelli rilevati oltre i 100 metri. E' richiesto altresì di utilizzare, in abbinamento al numero di uccelli osservati, i seguenti codici:

- C maschio in canto o mostrante qualche altra manifestazione territoriale (come nel caso di Columbiformi, Piciformi e Galliformi);
- M maschio non in canto;
- F femmina;
- j giovani non atti al volo o appena involati (con l'indicazione della quantità);
- r attività riproduttiva (trasporto imbeccata, asportazione di sacche fecali, trasporto di materiale per il nido, ecc.);
- V soggetti in volo di trasferimento, la cui presenza non è strettamente connessa alla stazione di rilevamento.

Per una corretta valutazione dell'avifauna presente è necessario che vengano indicati anche gli individui che si allontanano dal cerchio di 100 m di raggio perché disturbati dall'arrivo del rilevatore.

Le osservazioni devono essere registrate singolarmente e non "facendone una somma", poiché il loro significato in termini di numero di coppie potrebbe, nei due casi, essere diverso. Ad esempio, se il rilevatore vede 4 passeri in volo e successivamente ne vede altri 16 sempre in volo, dovrà indicare "4V, 16V" e non "20V".

L'abbinamento dei codici è permesso nei seguenti casi:

- Mr maschio impegnato in qualche attività riproduttiva;
- Fr femmina impegnata in qualche attività riproduttiva;
- MV maschio in volo di trasferimento;
- FV femmina in volo di trasferimento.

Il censimento comprende il conteggio sia dei piccioni torraioli sia dei piccioni selvatici, considerati separatamente. In generale si richiede di segnare qualunque specie identificata, includendo anche quelle non autoctone o naturalizzate.

Indicazioni ambientali

E' stato richiesto ai rilevatori di indicare nel raggio di 100 m dal punto di rilevamento:

a) la copertura percentuale delle variabili ambientali elencate (Categorie CORINE Land Cover, cfr. Fig. 2), utilizzando valori percentuali con incrementi successivi del 5%; la somma dei valori deve ovviamente ammontare a 100. In caso di categorie rappresentate con copertura ridotta segnalarne la presenza indicando come simbolo standard un "+" corrispondente circa all'1%;

b) la struttura della rete viaria, barrando una delle caselle corrispondenti alle tipologie di strade presenti;

c) le colture dominanti, nel caso siano presenti zone agricole, barrando una delle caselle corrispondenti o, nel caso di colture non elencate, barrando la casella "altro" e specificando la coltura;

d) la tipologia delle acque lotiche, barrando una delle caselle corrispondenti;

- e) la tipologia delle acque lentiche, barrando una delle caselle corrispondenti;
- f) la tipologia della palude, barrando una delle caselle corrispondenti.

1.3. DISTRIBUZIONE DEI PUNTI D'ASCOLTO

La distribuzione delle particelle da campionare, e delle relative stazioni d'ascolto, è fornita ai rilevatori dal coordinamento nazionale. Le particelle da campionare sono state selezionate principalmente in base a due criteri: 1) devono essere state visitate almeno una volta nel periodo 2000-2011; 2) devono preferibilmente contenere una percentuale significativa di ambienti agricoli e forestali (per i dettagli si veda il piano di campionamento).

L'esplorazione di ciascuna particella UTM 10x10 comporta l'esecuzione generalmente di 15 punti d'ascolto da eseguirsi in altrettanti quadrati di un km di lato, a loro volta individuati, nella fase iniziale del progetto, in base a una procedura di randomizzazione statistica. Solamente per le particelle al confine nazionale e sulla costa è previsto un numero minore di stazioni, in proporzione alla superficie della particella stessa. Per rispettare il più possibile le caratteristiche di "casualità" della selezione dei quadrati, il punto d'ascolto veniva eseguito il più vicino possibile al centro del quadrato (compatibilmente con le condizioni presenti). Attualmente la scelta delle stazioni da coprire viene fatta in maniera prioritaria su quelle stazioni che nel periodo 2000-2011 sono state visitate il maggior numero di volte ed avendo dal 2010 richiesto ai rilevatori di indicare le coordinate esatte del punto d'ascolto, è stata data l'indicazione di svolgere il punto d'ascolto esattamente nel solito posto.

L'obiettivo principale dei rilievi è quello di eseguire i rilevamenti nel maggior numero di particelle e - stazioni già rilevate nel passato. A tale fine l'organizzazione centrale ha fornito ai coordinamenti locali e/o ai singoli rilevatori, la cartografia necessaria aggiornata, di regola basata sulla cartografia IGMI in scala 1:25.000 messa a disposizione tramite servizio WMS dal Geoportale Nazionale del Ministero dell'Ambiente.

Nelle Figure 1.3, 1.4 e 1.5 viene riportato un esempio di carta, e i suoi dettagli, distribuita ai rilevatori, in cui sono indicate le stazioni prioritarie (le 15 stazioni con il numero maggiore di ripetizioni tra il 2000 e il 2011), le stazioni secondarie da scegliere in sostituzione di una o più stazioni prioritarie eventualmente impossibili da raggiungere (le stazioni che nel periodo 2000-2011 sono state visitate almeno una volta), e i punti esatti, con le relative coordinate, in cui svolgere i punti d'ascolto.

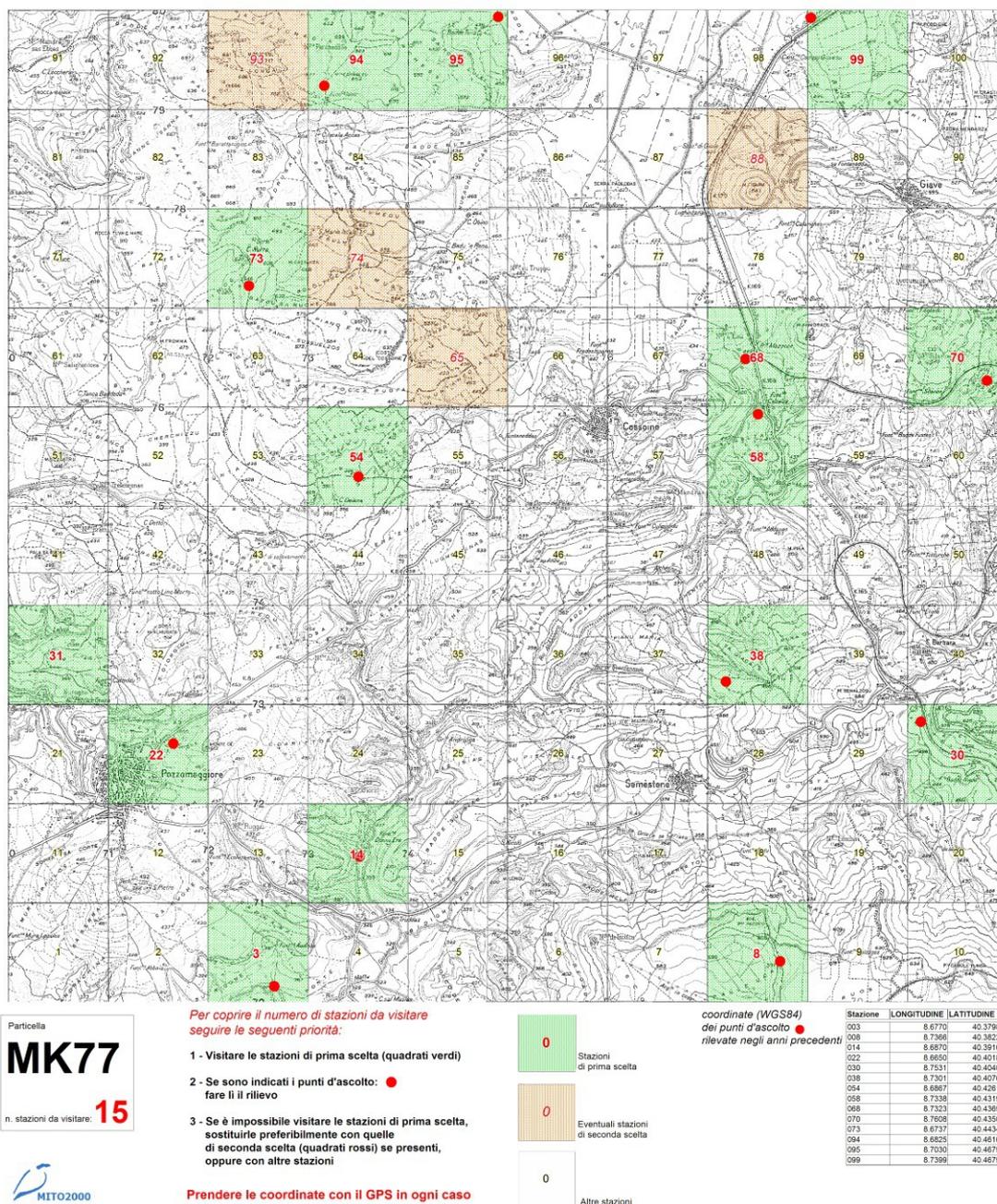


Figura 1.3: Esempio di cartografia utilizzata dai rilevatori. Le aree verdi indicano le stazioni di prima scelta, ovvero quelle che in passato avevano un numero maggiore di ripetizioni, quelle rosse le stazioni di seconda scelta, ovvero quelle che negli anni precedenti sono state censite almeno una volta. I punti rossi indicano la localizzazione esatta dei punti d'ascolto, derivante dalle coordinate registrate generalmente sul campo dagli stessi rilevatori, durante le sessioni precedenti.

2. STRUTTURA DELLA BANCA DATI DEL PROGETTO MITO2000 E PROCEDURE DI VALIDAZIONE

Il FBI (*Farmland Bird Index*) e il WBI (*Woodland Bird Index*) sono due strumenti utili ad individuare lo stato di conservazione rispettivamente degli ambienti agricoli e forestali, attraverso la descrizione degli andamenti delle popolazioni di uccelli che in questi vivono. Per poter essere calcolati necessitano di una serie di dati raccolti a lungo termine e su scala nazionale. A questo scopo ben si presta il progetto MITO2000 (Monitoraggio Italiano Ornitologico), che iniziato, dopo un anno di prova, nel 2000 e tuttora in corso, ha prodotto, in 13 anni di rilievi (2000-2012), una banca dati sull'avifauna nidificante che contiene informazioni raccolte in maniera standardizzata ed omogenea su tutto il territorio italiano (pur con alcune differenze tra le regioni nell'intensità di copertura nei vari anni).

Il progetto prevedeva originariamente un piano di campionamento randomizzato che utilizza come unità di campionamento le particelle 10x10 km ed un piano specifico per i rilievi nelle ZPS (Zone di Protezione Speciale) e le ZIO (Zone di Interesse Ornitologico), utilizzando come metodo di censimento quello delle stazioni di ascolto (si veda la sezione precedente). La banca dati del MITO2000 ha una struttura complessa che raccoglie, oltre alle informazioni sulla presenza degli uccelli, anche le caratteristiche ambientali dei siti in cui sono stati effettuati i rilievi, fornendo uno strumento utile per indagare in maniera analitica le relazioni tra gli uccelli e l'ambiente. Anche questa caratteristica rende particolarmente adatta la banca dati del progetto MITO2000 per applicazioni come il calcolo del FBI e del WBI.

Come ogni banca dati, anche quella del progetto MITO2000 può contenere inevitabilmente errori di varia natura, che possono generarsi in tutte le fasi del percorso di ciascun dato, dal momento della raccolta a tutte le fasi di gestione e aggiornamento della banca dati stessa. Tali errori che possono generarsi ad ogni passaggio (come del resto insegna la teoria dell'informazione) nonostante alcuni controlli vengano comunque fatti nelle varie fasi del processo di informatizzazione e gestione dei dati, possono, almeno in alcuni casi, essere comunque individuati in una fase di controllo specifica sulla banca dati completa. Con lo scopo quindi di fornire i dati corretti per il calcolo del FBI e del WBI, l'intero database del progetto MITO2000 viene sottoposto ad una procedura di "validazione", cercando di definire le principali tipologie di errore presenti e identificando procedure idonee all'individuazione di tali errori e alla loro correzione ove possibile.

In questo documento sono prima descritte le caratteristiche del database del progetto, poi vengono descritti i metodi e le procedure adottate nel processo di validazione.

2.1. LA STRUTTURA DELLA BANCA DATI

La banca dati del progetto MITO2000 è costituita da un database relazionale realizzato utilizzando la tecnologia Microsoft SQL Server. La Banca Dati Online (di seguito BDO) è localizzata su un server web professionale dedicato e si avvale di una procedura di backup automatico quotidiano.

Per l'accesso alla BDO, la visualizzazione ed il download dei dati delle osservazioni è stata realizzata un'interfaccia web di gestione dedicata al progetto. L'accesso all'interfaccia web è consentito solamente ad utenti dotati di privilegi di gestione, che possono visualizzare ed operare solo sulle aree della BDO per cui godono dei relativi privilegi.

Per poter operare con le ordinarie procedure di validazione e correzione dei dati, nonché

per ottenere *set* di dati "di lavoro", è possibile utilizzare una procedura di *download* dall'interfaccia web della BDO. I dati selezionati per il *download*, tramite un applicativo della BDO vengono trasferiti in un apposito spazio FTP, in cartelle dedicate ai singoli utenti. I dati vengono forniti in file in formato di testo (.CSV), corrispondenti alle tabelle della BDO.

La digitalizzazione dei dati dei rilevamenti è svolta mediante un applicativo *desktop* multi-piattaforma, realizzato in ambiente Adobe AIR. L'applicativo si appoggia su una banca dati locale basato sulla tecnologia SQL Lite.

L'applicativo locale consente, agli utenti dotati dei necessari diritti di accesso, di interfacciarsi con la BDO per effettuare la procedura di accodamento dei dati delle osservazioni.

L'accodamento dei dati nella BDO, qualora digitalizzati mediante software differente dall'applicativo dedicato, può avvenire anche tramite trasferimento FTP di file con formattazione standard. Procedendo in questo modo, l'accodamento dei dati avviene tramite un applicativo dedicato che processa i file caricati in apposite cartelle FTP, ne verifica la coerenza con le specifiche della banca dati e, qualora le specifiche siano adeguate, procede all'accodamento dei dati. Anche questa procedura è accessibile solo ad utenti che godano dei necessari diritti di gestione.

Utilizzando lo stesso applicativo è possibile, successivamente alle operazioni di validazione e correzione, caricare dati aggiornati nella BDO che sostituiscano quelli in essa già presenti. I dati sostituiti sono oggetto di backup, ai fini di poter sempre risalire alle versioni precedenti gli aggiornamenti effettuati.

Per poter utilizzare i dati del progetto in maniera semplice ai fini della validazione e delle analisi previste, il contenuto della BDO viene periodicamente trasferito in archivio locale "di lavoro", in formato Microsoft Access.

Dall'archivio locale vengono estratti i dati nel formato .dbf, comodamente utilizzabili all'interno di software GIS (MapInfo e QGIS) per poter procedere anche alle verifiche di tipo spaziale.

Gli archivi, allegati alla presente relazione, sono due, uno dei quali contenente i dati stazionali con le relative informazioni sull'ambiente "SCHEDE_2009_2012", mentre l'altro contenente i dati sugli uccelli "SPECIE_2009_2012".

L'archivio "SCHEDE_2009_2012" si articola in due blocchi principali: quello contenente i dati relativi alla localizzazione delle stazioni (Tabella 2.1) e quello contenente le informazioni relative ai rilevatori, alla cronologia del rilevamento e alle condizioni meteorologiche (Tabella 2.2).

I dati che riguardano gli uccelli, presenti in "SPECIE_2009_2012", sono il numero di individui censiti ed il numero di coppie nelle due fasce di distanza, entro ed oltre 100 metri dal punto d'ascolto (Tabella 2.3). Ogni record, identificato con un numero progressivo nel campo "ID" è legato ai dati stazionali da un codice identificativo "ID_GEN", univoco e indipendente da qualsiasi altra informazione presente nell'archivio.

Tabella 2.1: Struttura dei campi del blocco relativo ai dati stazionali dell'archivio "SCHEDE_2009_2012", contenente le informazioni relative alla localizzazione delle stazioni coperte dal progetto MITO2000. Nella colonna "Origine" abbiamo inserito "I" per i campi riempiti dall'informatizzatore durante la fase di inserimento dei dati e "D" per quelli derivati, ad esempio utilizzando programmi GIS.

Nome del campo	Struttura	Descrizione	Origine
ID_GEN	Carattere, 7	Identificativo univoco, con struttura CNNNNNN, che collega i dati stazionali ai record degli uccelli; i numeri sono progressivi in ordine crescente regolare e indipendenti da altri campi.	D
SITO_ANNO	Carattere, 15	Identificativo di controllo contenente le tre informazioni di base riguardanti la localizzazione. E' costruito con la seguente struttura Particella_Stazione_anno	D
PROGRAMMA	Carattere, 2	RA=programma randomizzato; ZP=ZPS; ZI=ZIO	I
AREA	Carattere, 9	Codice della particella UTM 10x10 km	I
COD_ZZ	Carattere, 9	Codice della ZPS, SIC e ZIO (zone di interesse ornitologico)	I
NOME_ZZ	Carattere, 82	Nome della ZPS, SIC e ZIO (zone di interesse ornitologico)	I
REGIONE_STAZ	Carattere, 3	Codice della regione costituito dalle prime tre lettere, attribuito a livello di stazione, dipendentemente dalla posizione del suo centroide	D
REGIONE_PART	Carattere, 3	Codice della regione costituito dalle prime tre lettere, attribuito a livello di particella in base alla regione di appartenenza del maggior numero di stazioni	D
STAZ	Carattere, 4	Numero della stazione, compreso tra 1 e 100, nella forma SNNNX	I
TOPONIMO	Carattere, 58	La località in cui viene realizzato il punto d'ascolto riportata sulla scheda dal rilevatore.	I
X_NOD	Intero	Longitudine del centroide della stazione 1x1km, secondo la proiezione UTM (ED50) fuso 32	D
Y_NOD	Intero	Latitudine del punto, vedi XNODUTM	D
X_WGS84	Decimale	Longitudine del punto d'ascolto, coordinate rilevate sul campo secondo la proiezione WGS84	I
Y_WGS84	Decimale	Latitudine del punto d'ascolto, vedi X_WGS84	I
VERIFICHE	Carattere, 17	Campo di servizio, utilizzato per i controlli	D
ERRORI	Carattere, 2	Sono riportati numeri progressivi per le segnalazioni di errori e sigle per le correzioni. Le spiegazioni di tali codifiche sono riportate nell'archivio degli errori.	I

Tabella 2.2: Struttura dei campi del blocco relativo ai dati stazionali dell'archivio "SCHEDE_2009_2012", contenente le informazioni relative alla localizzazione delle stazioni coperte dal progetto MITO2000. Nella colonna "Origine" abbiamo inserito "I" per i campi riempiti dall'informatizzatore durante la fase di inserimento dei dati e "D" per quelli derivati, ad esempio utilizzando programmi GIS.

Nome del campo	Struttura	Descrizione	Origine
DATA	Data	Il giorno del rilievo riportato in formato data	I
ANNO	Intero	Anno del rilievo, espresso come numero intero	D
ORA	Intero	E' registrata solamente la fascia oraria in cui è stato fatto il rilievo fino al 2010, dal 2011 l'orario completo di ora e minuti	I
RILEVATORE	Carattere, 3	Nome di colui che ha effettuato il rilievo	I
OSSERVATORE	Carattere, 9	Nomi di coloro che accompagnano il rilevatore	I
CIELO	Carattere, 2	1=sereno, 2=nuvole per ¼, 3=nuvole per ½, 4=nuvole per ¾, 5=coperto, 6=pioggia e 9=nebbia	I
VENTO	Carattere, 2	A=assente, B=debole, C=moderato, D=forte, E=molto forte	I
VIS_PRE	Carattere, 2	Risposta alla domanda "Hai già visitato almeno una volta questa stazione?"	I
RIP_EX	Carattere, 2	Risposta alla domanda "Hai fatto il punto d'ascolto nella stessa posizione esatta?"	I

Tabella 2.3: Struttura dell'archivio "SPECIE_2009_2012", contenente i dati sugli uccelli raccolti durante il programma di monitoraggio MITO dal 2000 al 2011. Nella colonna "Origine" abbiamo inserito "I" per i campi riempiti durante la fase di inserimento dati e "D" per quelli derivati.

Nome del campo	Struttura	Descrizione	Origine
verifiche	Carattere, 15	Campo di servizio per la gestione e validazione.	D
Id_gen	Carattere, 7	Identificativo univoco, con struttura DNNNNNN, che collega ogni record ai dati stazionali; i numeri sono progressivi in ordine crescente regolare	D
Id	Intero	Numero progressivo che identifica in maniera univoca ogni contatto relativo agli uccelli registrato	D
Cod_euring	Intero	Codice Euring del taxon, specie, genere, o ibrido interspecifico; il riferimento è la lista CISO-COI degli uccelli italiani pubblicata nel 2009 sul volume 33 (1) di Avocetta	I
Coppie_entro	Decimale, 5,1	Numero di coppie calcolate da Aegithalos secondo lo specifico algoritmo dentro la fascia di 100 metri dal punto d'ascolto. Il dato non è presente nel caso di rilevamenti archiviati senza l'utilizzo dell'apposito software (es. dati provenienti da programmi regionali propri)	D
Coppie_oltre	Decimale, 5,1	Numero di coppie calcolate da Aegithalos secondo lo specifico algoritmo dentro la fascia di 100 metri dal punto d'ascolto. Il dato non è presente nel caso di rilevamenti archiviati senza l'utilizzo dell'apposito software	D
Coppie_tot	Decimale, 5,1	Numero totale di coppie. Il dato non è presente nel caso di rilevamenti archiviati senza l'utilizzo dell'apposito software	D
Specie_ent	Carattere, 20	Codice relativo al contatto (C, R, V...) entro 100 m	I
Specie_olt	Carattere, 20	Codice relativo al contatto (C, R, V...) oltre 100 m	I

Nome del campo	Struttura	Descrizione	Origine
Ind_entro	Intero	Numero di individui contattati dentro la fascia di 100 metri dal punto d'ascolto	I
Ind_oltre	Intero	Numero di individui contattati fuori dalla fascia di 100 metri dal punto d'ascolto	I
Ind_tot	Intero	Numero totale di individui censiti	I

2.2. ARCHIVIO DEGLI ERRORI

Per la gestione degli errori abbiamo creato un archivio di servizio contenente l'elenco degli errori individuati nel database, la tipologia di errore, il nome di file parziali contenenti tali errori estratti tramite SQL dal database e la soluzione adottata. Tale archivio è un utile strumento di lavoro, in quanto permette di mantenere la memoria delle modifiche effettuate e di segnalare eventuali dati dubbi, senza dover ricorrere necessariamente alla separazione e all'isolamento di tali dati dagli archivi principali. Questa operazione, infatti, non permetterebbe di sottoporre questi dati a tutte le successive fasi di validazione.

In entrambi gli archivi i dati dubbi sono segnalati con un codice numerico preceduto dalla lettera "E" in appositi campi e questo permette di selezionarli facilmente ed escluderli ogni qualvolta si utilizzano i dati per le analisi statistiche.

Una volta che i dati sono stati sottoposti a validazione ed hanno subito la correzione, ai codici di errore viene sostituita la lettera "E" con la lettera "C".

Procedure di validazione

Di seguito riportiamo prima le operazioni di validazione preliminari applicate ad entrambi gli archivi, poi, separatamente, quelle operate sui differenti settori degli archivi. Per ogni verifica riportiamo se è stata eseguita per campioni con controlli sulle schede cartacee (*a campione*) oppure su tutti i dati, con l'ausilio di procedure informatiche (*completa*).

Verifiche generali

Per ogni campo, di entrambi gli archivi, sono state create tabelle in cui vengono riportati i valori inseriti, allo scopo di individuare la presenza di valori o codici inesistenti e al fine di uniformare i dati secondo le codifiche stabilite. Riportiamo a titolo di esempio il lavoro svolto per il campo "rilevatore" in "SCHEDE_2009_2012". Estraendo dell'elenco dei rilevatori può capitare che risulti che lo stesso rilevatore sia stato registrato in differenti modi: il signor Mario Rossi, può essere riportato con il nome e cognome scritti per esteso, nell'ordine citato o invertito, oppure abbreviando il nome, riportandolo come M. Rossi. Se alle tre voci si attribuisce un codice che identifica il signor Mario Rossi, si possono sostituire facilmente con un unico codice di riferimento utilizzando procedure automatiche.

Inoltre, per ogni campo numerico (che, in questi archivi dovrebbe contenere solo valori positivi) è stato inserito il valore -1, in tutti quei casi in cui la cella non è stata riempita, in modo da poterla discernere da quelle contenenti valori uguali a zero.

Infine abbiamo verificato la congruità del campo di azione di ciascun rilevatore. La distribuzione delle stazioni coperte ogni anno da un singolo rilevatore può essere un'indicazione utile per mettere in luce un eventuale errore di localizzazione. Per questo motivo è stato controllato per via grafica il campo di azione di ogni rilevatore, evidenziando e controllando eventuali incongruenze.

Verifiche spaziali sui dati stazionali

Le seguenti verifiche riguardano i campi dell'archivio "SCHEDE_2009_2012" contenenti le informazioni relative alla localizzazione dei dati stazionali del programma randomizzato, riportati nella Tabella 2.1. Queste verifiche hanno riguardato:

1. per i dati del 2012, la corrispondenza tra il piano di monitoraggio e le particelle presenti nell'archivio in modo da verificare la presenza di tutti i dati previsti (*completa*);
2. conteggio del numero di ripetizioni a livello di stazione (*completa*): prevedendo il

metodo di campionamento la ripetizione delle stesse stazioni ogni anno, un modo utile per individuare eventuali errori di numerazione delle stazioni è tramite il confronto diretto con i dati presenti in archivio degli anni passati, abbiamo quindi conteggiato quante volte ogni stazione riportata del 2012 risulta essere stata ripetuta;

3. confronto fra stazione e punto d'ascolto (*completa*): tramite le coordinate raccolte sul campo dal rilevatore con un GPS è stato possibile, con strumenti GIS, cartografare la posizione esatta del punto d'ascolto. Sono state confrontate le stazioni riportate con quelle in cui ricade il punto descritto dalle coordinate (Figura 2.1);

091	092	093	094	095	096	097	098	099	100
081	082	083	084	085	086	087	088	089	090
071	072	073	074	075	076	077	078	079	080
061	062	063	064	065	066	067	068	069	070
051	052	053	054	055	056	057	058	059	060
041	042	043	044	045	046	047	048	049	050
031	032	033	034	035	036	037	038	039	040
021	022	023	024	025	026	027	028	029	030
011	012	013	014	015	016	017	018	019	020
001	002	003	004	005	006	007	008	009	010

Figura 2.1: I quadratini evidenziati di giallo indicano le stazioni riportate sulla scheda dal rilevatore, i pallini rossi indicano il punto d'ascolto proiettato via GIS tramite le coordinate raccolte sul campo dal rilevatore. Sono evidenziati con i cerchi blu due possibili tipologie di errori: il punto cade fuori dalla particella oppure cade fuori dalla stazione.

4. presenza di stazioni con numerazioni progressive (*completa*): è possibile che alcuni rilevatori abbiano inserito al posto del numero che identifica la stazione un numero progressivo che indica l'ordine con cui le stazioni sono state coperte. In questo caso, una volta individuata la serie, la correzione prevede un controllo sulle coordinate, sui toponimi oppure una richiesta di correzione direttamente ai rilevatori;
5. presenza di dopponi nel numero della stazione (*completa*): la correttezza nel riportare il numero della stazione è di fondamentale importanza per avere la giusta localizzazione del rilievo. Il doppone oltre ad essere di per sé un errore può indicare l'assenza di una stazione. Anche in questo caso per risolvere l'errore è stato fatto un controllo sul materiale cartaceo;
6. corrispondenza tra stazione e toponimo (*a campione*): la presenza del toponimo è un'informazione preziosa per l'individuazione di eventuali errori commessi nel riportare il numero della particella. La verifica è stata fatta via GIS utilizzando come base cartografica i tipi dell'IGMI in scala 1:25000, scaricabile liberamente dal servizio WMS del Portale Cartografico Nazionale nel sito del Ministero

dell'Ambiente.

Verifiche su rilevatore e cronologia del rilevamento

Le seguenti verifiche riguardano i campi dell'archivio "SCHEDE_20096_2012" riportati nella Tabella 2.2:

1. congruenza dell'anno nei campi ANNO e DATA (*completa*);
2. verifica della congruenza degli orari (*completa*): l'orario suggerito per svolgere i rilievi va da poco dopo l'alba alle 11 del mattino. E' però possibile per motivi logistici che alcuni rilievi siano stati realizzati al di fuori dell'orario standardizzato. Abbiamo così deciso di segnalare sistematicamente come errore soltanto gli orari notturni.

Verifiche sui dati degli Uccelli

Per quanto riguarda i campi dell'archivio "SPECIE_2009_2012", riportati in Tabella 2.3, sono state eseguite le seguenti verifiche:

1. presenza unica di ogni specie in ogni stazione (*completa*): i record dell'archivio non rappresentano i singoli contatti con gli uccelli o con i gruppi di uccelli, bensì gli individui appartenenti ad una determinata specie, contattati durante un rilevamento, espressi in forma di numero di individui e numero di coppie. E' perciò necessario che ogni specie non compaia più di una volta in ogni punto d'ascolto (identificato da Sito_anno);
2. controllo e aggiornamento della check-list (*completa*): la presenza di errori nel campo Cod_euring può dipendere da identificazioni sbagliate delle specie di uccelli, da errori commessi durante la compilazione della scheda o durante la fase di immissione dei dati e dall'uso di codici euring inesistenti.

In preparazione delle procedure di validazione sono state estratte tutte le occorrenze di codice Euring presenti nell'archivio, rilevate durante tutti gli anni di monitoraggio e per ognuna il numero di volte in cui è stata riportata.

Per prima cosa è stata verificata la presenza di codici realmente esistenti tramite un confronto utilizzando come riferimento la lista CISO-COI degli uccelli italiani pubblicata recentemente (Fracasso et al., 2009). In alcuni casi tuttavia nell'archivio sono presenti codici Euring non conformi alla suddetta lista, per specie o sottospecie non incluse nella lista stessa (per le quali è stata consultata la lista ufficiale dell'Euring) oppure per risolvere situazioni particolari. Ad esempio è rimasto presente nell'archivio il codice Euring 15912 (passera d'Italia) che non trova corrispondenza nella lista, in quanto, al momento, non è chiarita la posizione sistematica delle popolazioni *italiae*. Per il momento, in questo caso, abbiamo mantenuto per il taxon "italiae", il codice 15912. Un discorso diverso va fatto per le sottospecie che sono state elevate a specie, come ad esempio la sterpazzolina. Nel recente passato vi era una unica specie identificata dal codice Euring 12650, invece con l'ultimo aggiornamento abbiamo almeno due specie: la sterpazzolina comune (12650) e la sterpazzolina di Moltoni (12652). Per poter discernere in futuro le due nuove specie da quello non identificate nel passato, sono state trasformate tutte le sterpazzoline contattate fino al 2009 in sterpazzolina *sp.* (12659), che giocoforza comprende le attuali due specie.

Allo scopo di individuare eventuali specie rare da sottoporre a verifica, a tutti i codici Euring sono state attribuite le categorie proposte per la stesura della sopraccitata lista CISO-COI, ovvero le categorie AERC, raccomandate dall'Association of European Records and Rarities Committees (cfr. www.aerc.eu), le categorie di status generale e le categorie di status riproduttivo. Il primo

controllo eseguito riguardava la presenza/assenza di specie per le quali è richiesta l'omologazione dalla COI, la Commissione Ornitologica Italiana, che nell'ambito del CISO si occupa delle rarità. Successivamente sono state controllate sul cartaceo, ed in certi casi con un confronto diretto con i rilevatori, quelle specie con un numero molto basso di contatti.

3. controllo dei dati fuori areale: per ogni specie è stata prodotta una cartina di distribuzione da confrontare con gli areali noti per le specie stesse. A titolo di esempio riportiamo nella Figura 2.2 la cartina dei dati di cincia alpestre *Poecile montanus* presenti nell'archivio dove è evidente la presenza di un dato fuori areale. Questi dati sono stati selezionati e verificati con ulteriori procedure. In alcuni casi si è proceduto ad una verifica ambientale utilizzando il Portale Cartografico Nazionale del Ministero dell'Ambiente, in cui sono messe a disposizione ortofoto a colori per tutto il territorio nazionale importabili su programmi GIS; su queste carte sono state proiettate le localizzazioni, in modo da verificarne anche la compatibilità ambientale, ad esempio controllando che una specie di ambienti agricoli non risulti rilevata nel mezzo di un bosco o viceversa. Altre procedure di verifica hanno riguardato il controllo delle schede cartacee e, ove necessario e possibile, sono stati contattati i rilevatori stessi;



Figura 2.2: Particelle in cui sono stati contattati individui di Cincia alpestre *Poecile montanus*. Oltre alle localizzazioni nell'area alpina, che rientrano nell'areale noto della specie, è indicata con una freccia una localizzazione al di fuori di questa area. A seguito di questo controllo si è potuto appurare che si trattava di un'errata attribuzione di codice.

4. controllo del numero delle coppie e degli individui (*completa*): il numero delle coppie e degli individui in una stazione deve essere sempre rispettivamente

maggiore di 0.5 e di 1;

5. presenza di numeri non verosimili per le specie non gregarie (*completa*): abbiamo stilato un elenco delle specie non gregarie per le quali sono state stabilite delle soglie specie specifiche di numero di individui possibilmente contattabili durante un rilevamento. Successivamente sono stati estratti e verificati tutti i record superiori a queste soglie.

3. SELEZIONE DEI DATI PER L'ANALISI

Sia a livello regionale che a scala nazionale, ai fini del calcolo degli andamenti di popolazione delle specie ornitiche indicatrici degli ambienti agricoli e forestali sono stati considerati solo i dati provenienti dal programma randomizzato: ciò garantisce la produzione di risultati rappresentativi dell'intero territorio di interesse, scopo precipuo degli indicatori compositi.

Dai dati provenienti dal programma randomizzato sono stati eliminati i record contrassegnati da codici di errore che ne compromettessero l'affidabilità ai fini del calcolo degli indici di popolazione (si veda a tale proposito il paragrafo 2.2).

Quale unità territoriale di analisi è stata utilizzata la particella 10x10 km. Il singolo punto di ascolto è caratterizzato da una elevata variabilità dovuta a diversi fattori stocastici; essa può essere ridotta aggregando i dati di diversi campionamenti puntiformi. Tale accorgimento permette inoltre di limitare, in fase di analisi, il fenomeno noto come sovradisersione, ovvero la presenza nei dati di una variabilità maggiore rispetto a quella attesa in base al modello utilizzato per l'analisi dei dati stessi (in questo caso la regressione di Poisson, vedi oltre).

Al fine di utilizzare i dati a livello locale le particelle sono state attribuite alle differenti regioni. Le particelle di confine sono state attribuite ad una determinata regione, quando almeno 6 punti fossero ricaduti entro i confini regionali. In pochi casi dunque le particelle sono state attribuite ad entrambe le regioni (se ad esempio in una particella 9 punti fossero ricaduti in una data regione e 6 in quella confinante). Sono state escluse da questa "doppia" attribuzione le particelle di confine rilevate da un coordinamento regionale nell'ambito di programmi finanziati da enti locali.

Poiché per confrontare correttamente gli indici di popolazione tra anni, è necessario poter disporre di serie temporali relative alle stesse unità di campionamento, sono poi stati selezionati per le analisi i dati raccolti nelle particelle visitate in almeno due stagioni riproduttive nel periodo compreso tra il 2000 ed il 2012.

Il programma randomizzato prevede di norma 15 campionamenti puntiformi all'interno di una particella di 10x10 km. Il numero effettivo di campionamenti condotti in una particella, soprattutto nel passato, non è tuttavia sempre stato lo stesso a causa di diverse motivazioni di ordine pratico ed organizzativo. Inoltre, stante l'esecuzione dei 15 rilievi previsti, la disponibilità effettiva di dati potrebbe essere inferiore a causa di problemi nei dati che li rendono non utilizzabili (ad esempio l'errata localizzazione della stazione di rilevamento – cfr. paragrafo 2.2). Si è a tal fine scelta una soglia minima di numero di stazioni per particella pari a 7. Dalla banca dati per le analisi sono dunque stati eliminati i dati relativi alle particelle nelle quali non erano state visitate almeno 7 stazioni per due anni.

Per effettuare correttamente il confronto tra anni diversi per la stessa particella è necessario disporre di numeri uguali di stazioni. Per ogni particella è dunque stato individuato il numero minimo annuale di stazioni visitate, selezionando successivamente, per ogni anno in cui la particella è stata visitata, lo stesso numero di stazioni. Se ad esempio in una particella sono state visitate un anno 8 stazioni, e un altro anno 12 stazioni, sono state eliminate 4 stazioni dal secondo anno. In pochi casi si è preferito eliminare l'annualità di una determinata particella che presentava un numero minimo di stazioni molto più basso rispetto alla media: la regola che ha guidato queste scelte è stata quella di minimizzare il numero di dati scartati garantendo la migliore copertura temporale possibile. Quindi, ad esempio, se per tutti gli anni di copertura di una determinata particella erano state coperte 15 stazioni, ma solamente per un anno 7 stazioni, si è preferito eliminare quest'ultima annualità, piuttosto che eliminare 8 stazioni

da tutti gli altri anni.

La selezione delle stazioni all'interno della particella è stata operata conservando le stazioni visitate nel maggiore numero di anni mentre, a parità di copertura, la selezione è stata casuale.

Da tutte queste operazioni è derivato un archivio utilizzabile a scala nazionale di 451687 dati.

A differenza degli anni precedenti quest'anno si è deciso di utilizzare per il calcolo dei trend il numero degli individui invece del numero delle coppie, sulla base di differenti motivazioni.

In primo luogo i dati hanno una diversa fonte di origine, essendo in alcune regioni raccolti ed archiviati in banche dati diverse da quelle del progetto MITO2000; ciò non permette di trattare gli stessi in maniera uniforme in fase di stima del numero di coppie.

Vi sono poi motivi legati alla non uniformità di interpretazione del significato etologico delle vocalizzazioni di alcune specie. Alcune vocalizzazioni, ad esempio quelle delle specie del genere *Passer*, vengono interpretate come canto di tipo territoriale da alcuni rilevatori, come vocalizzazioni di altro genere senza alcun significato riproduttivo da altri. L'algoritmo che converte il numero di individui in coppie sulla base dei codici di attività utilizzati in abbinamento al numero di individui osservati o sentiti durante i rilievi di campo (si veda paragrafo 2.2), può dare origine a stime che possono differire a seconda dell'interpretazione data alla vocalizzazione dal rilevatore. Tale condizione può divenire problematica quando ad effettuare i rilevamenti nella medesima particella siano rilevatori differenti di anno in anno. Pertanto per ovviare a problemi di sovrastima o sottostima del numero di coppie che possono poi riflettersi sul calcolo degli andamenti si è optato per l'utilizzo nelle analisi del numero di individui.

4. METODI E SOFTWARE

I dati relativi agli uccelli comuni nidificanti in Italia vengono analizzati con metodi statistici sviluppati appositamente per l'analisi di serie temporali di conteggi contenenti diverse osservazioni mancanti. Questi metodi vengono applicati tramite un programma *freeware* sviluppato da *Statistics Netherlands* e denominato TRIM (*TREnds and Indices for Monitoring data*). TRIM è stato sviluppato appositamente per l'analisi dei dati di conteggio derivanti dal monitoraggio delle popolazioni selvatiche finalizzata alla produzione di indici temporali e tendenze demografiche. L'utilizzo di TRIM viene raccomandato da EBCC ai fini della comparabilità degli indici provenienti dai diversi Paesi europei.

I modelli implementati in TRIM appartengono alla classe dei modelli log-lineari (Agresti, 1990, capitolo 5; McCullagh & Nelder, 1989, capitolo 6), cui ci si riferisce anche con il termine "regressione di Poisson".

La stima dei parametri per i modelli log-lineari si basa sul metodo della massima verosimiglianza (ML- *maximum likelihood*). Con esso vengono calcolati i parametri, i relativi errori ed effettuati i test statistici.

Nel caso della regressione di Poisson queste procedure si fondano su alcuni assunti, in particolare la distribuzione indipendente di Poisson dei conteggi. Tali assunti sono spesso violati nei monitoraggi faunistici poiché la varianza reale è solitamente maggiore di quanto atteso per una distribuzione di Poisson (fenomeno noto come sovradisersione), e perché i conteggi non sono indipendenti ma i valori registrati in un dato anno dipendono da quelli registrati l'anno precedente (fenomeno noto come correlazione seriale).

TRIM può utilizzare una differente procedura di stima dei parametri che tiene in considerazione questi due fenomeni: si tratta del metodo delle Equazioni di Stima Generalizzate (GEE - Generalised Estimating Equations; Liang & Zeger, 1986).

A prescindere dal metodo utilizzato, la stima dei parametri avviene per mezzo di un processo iterativo. TRIM permette di settare il numero massimo di iterazioni.

TRIM utilizza 3 modelli per l'analisi di tendenze e indici che però possono essere formalizzati matematicamente in un modello generale denominato, nel manuale d'uso del programma, come *linear (switching) trend*

Il modello lineare di base è un modello con effetti del sito (i) e con un effetto lineare (su scala logaritmica) del tempo (j).

Questo modello può esser formulato con la seguente equazione:

$$\text{Ln}(\mu_{ij}) = \alpha_i + \beta(j-1) \quad (1)$$

in cui si evince che per ogni sito i i valori attesi (μ_{ij}) sono funzione lineare del tempo.

Il modello (1) implica una crescita costante (del logaritmo) dei valori attesi di conteggio da un anno a quello successivo. Il secondo modello lineare denominato *switching trend model* permette invece che il parametro che definisce la pendenza della retta di regressione cambi in corrispondenza di determinati anni denominati *changepoint*. Infine quando viene consentito un cambio di pendenza per tutti gli anni (eccetto l'ultimo), si ha il modello più complesso, denominato *time-effect*. Quest'ultimo, in teoria, è quello che descrive con maggior precisione le variazioni interannuali delle dimensioni di popolazione.

TRIM fornisce due prodotti principali:

- indici annuali
- tendenze sull'intero periodo

Gli indici annuali costituiscono il conteggio totale per quel determinato anno diviso per il conteggio totale dell'anno di riferimento (nel nostro caso il 2000). Essi costituiscono dunque il fattore di incremento o decremento rispetto all'anno di riferimento. TRIM calcola due tipi di indici: quelli basati sul modello e quelli basati sui valori, detti *imputed*, della matrice Particella x Anno.

Per ogni combinazione di Particella ed Anno il valore *imputed* corrisponde al conteggio reale nel caso in cui il dato esista; corrisponde invece al valore stimato dal modello se un dato è mancante (ovvero se la tal particella non è stata visitata il tal anno).

Gli indici basati sui valori *imputed* sono quelli utilizzati per il calcolo degli indicatori compositi e per la produzione dei grafici riportati nella relazione. Per comodità di interpretazione i valori degli indici vengono rappresentati come valori percentuali considerando pari al 100% l'anno di riferimento (2000).

Oltre agli indici TRIM stima la tendenza delle popolazioni sull'intero periodo considerato: questa è la pendenza della retta di regressione del logaritmo dell'indice. Questa tendenza viene detta addittiva ed è utilizzata comunemente nei lavori scientifici. TRIM calcola anche la tendenza moltiplicativa, più facilmente interpretabile. Si tratta del cambiamento percentuale medio per anno. Se questo corrisponde ad esempio a 1.08, significa che c'è un incremento medio dell'8% ogni anno (l'indice avrà valore 1 nel primo anno, valore $1 \times 1.08 = 1.08$ il secondo anno, valore $1.08 \times 1.08 = 1.17$ il terzo anno, ecc.).

TRIM in base ai risultati dei modelli, per una determinata specie permette di definire alcune categorie di andamento delle popolazioni nidificanti. La definizione di tali categorie viene effettuata tenendo in considerazione il valore della variazione media annua (tendenza moltiplicativa), ma anche il suo grado di "incertezza", costituito dall'intervallo di confidenza al 95% (ovvero il valore della tendenza ± 1.96 volte l'errore standard della tendenza stessa). Gli andamenti vengono classificati nel seguente modo:

- Incremento marcato - incremento annuo statisticamente significativo maggiore del 5%;
- Incremento moderato - incremento statisticamente significativo, ma con valore di variazione non significativamente maggiore del 5%;
- Stabilità - assenza di incrementi o diminuzioni statisticamente significative e variazione media annua generalmente inferiore al 5%;
- Diminuzione moderata - diminuzione statisticamente significativa, ma con valore di variazione non significativamente maggiore del 5%;
- Diminuzione marcata - diminuzione annua statisticamente significativa maggiore del 5%;
- Andamento incerto - assenza di incrementi o diminuzioni statisticamente significative e variazione media annua generalmente superiore al 5%. Ricadono in questa categoria le specie per le quali, a partire dai dati analizzati, non è possibile definire statisticamente una tendenza in atto. L'incertezza statistica deriva da molteplici fattori tra i quali possiamo ad esempio includere la presenza di valori molto dissimili dell'indice di popolazione da un anno con l'altro o la diversa tendenza calcolata nelle unità di campionamento (in alcune particelle la specie può aumentare, mentre in altre diminuire). Per le specie più abbondanti e meglio distribuite l'inclusione nella categoria non significa necessariamente che l'andamento non sia realistico;
- Dati insufficienti - i dati di presenza della specie sono in numero troppo scarso per

poter calcolare indici di popolazione annuali descrittivi dell'andamento, anche di tipo incerto, in corso. Si è scelto di considerare in questa categoria le specie per le quali il numero di casi positivi (ovvero il numero di volte che, nel periodo considerato, è stato rilevato almeno un individuo della specie nelle particelle selezionate per le analisi) è risultato pari o inferiore a 26 (corrispondente ad una media di due casi positivi per anno). La scelta di applicare criteri più rigidi di esclusione delle specie dalle analisi che nel passato è legato alla necessità di ottenere indicatori più realistici e meno soggetti a oscillazioni ampie e repentine.

5. CALCOLO DELLE TENDENZE DI POPOLAZIONE DELLE SPECIE

5.1. REGIONI

Per mezzo del software TRIM, dunque, per ogni specie indicatrice è stata effettuata la stima degli indici di popolazione relativi agli anni dal 2000 al 2012 e la loro tendenza nel periodo considerato.

Quando possibile è stato selezionato il modello *time-effect* (= *linear switching trend* con tutti i *change point*). Qualora si sia verificata mancanza di dati o di osservazioni per una determinata specie in determinati anni, tali anni non sono stati considerati come *change point*: in questi casi è cioè stato utilizzato il modello *linear switching trend* con il maggior numero possibile di *change point*. Il metodo di stima dei parametri del modello è stato impostato in modo da tenere in considerazione sia sovradisersione dei dati che correlazione seriale. Il numero massimo di iterazioni nel processo di stima dei parametri è stato settato a 500 aumentandolo rispetto al valore di default di 200 ed aumentando la probabilità di convergenza dei modelli, ovvero la probabilità di giungere alla migliore soluzione possibile delle equazioni su cui i modelli si basano.

Prima di produrre le analisi è stata individuata una soglia minima di abbondanza delle specie per procedere all'analisi dei dati utilizzando criteri più rigidi rispetto al passato, al fine di cercare di ottenere indici compositi più realistici e con oscillazioni meno marcate e repentine.. La soglia è stata individuata in base al numero di dati presenti per una determinata specie utilizzando il valore dei conteggi positivi (per ogni specie un conteggio positivo corrisponde ad un dato di presenza relativo ad una determinata particella in un determinato anno). Mettendo in relazione il numero di dati utilizzati e i risultati (output) degli anni precedenti, è stata individuata la soglia di 26 conteggi positivi che corrispondono ad una media minima annuale di due particelle in cui la specie sia stata rilevata.

Le specie con numero di conteggi positivi maggiore della soglia stabilita sono state sottoposte ad analisi dei dati. Queste possono produrre risultati statisticamente coerenti ma ecologicamente privi di significato. È dunque prassi comune verificarne i risultati controllando in particolare valori o *pattern* aberranti.

A tal fine è stata implementata una procedura di controllo degli output di TRIM mutuata da documenti tecnici dell'EBCC. La procedura ha previsto il controllo dei casi in cui si fossero verificate le seguenti condizioni:

- Valore dell'indice (in forma percentuale) > 1000 o < 0.5
- Valore del trend (*multiplicative*) > 1.5 o < 0.6
- Valore dell'errore standard del trend > 0.5

Qualora si fosse verificata una delle situazioni sopra descritte il modello è stato controllato nel dettaglio cercando di capire se i valori aberranti fossero determinati da problematiche nella struttura dei dati.

In diversi casi è risultato che i problemi nella stima dei parametri erano dovuti alla difficoltà di calcolare alcuni cambiamenti di direzione dell'abbondanza in corrispondenza di *change point* caratterizzati da un basso numero di dati. È dunque stata effettuata una rimozione manuale dei *change point* problematici. Nei casi in cui tale accorgimento ha portato ad un miglioramento del modello, lo stesso è stato sostituito al modello iniziale. Quando invece i problemi nei risultati sono rimasti tali il modello è stato considerato non valido e pertanto non presentato nei risultati.

In alcuni casi TRIM non è stato in grado di effettuare la stima dei parametri del modello

prima del raggiungimento del limite di iterazioni. In questo caso i risultati non sono stati presentati e ciò è stato riportato nella relazione con l'espressione "il modello non converge".

5.2. ITALIA

Per ciascuna delle specie selezionate, abbiamo calcolato gli indici annuali di popolazione e la tendenza generale mediante il software TRIM (Pannekoek & van Strien, 2001; van Strien et al., 2001), come raccomandato dall'EBCC. Tra le numerose opzioni possibili, abbiamo selezionato le seguenti:

1. calcolo della tendenza di popolazione utilizzando il modello "time effect";
2. quest'anno, coerentemente a quanto fatto nel 2012, abbiamo utilizzato la pesatura differenziale dei dati, poiché il recupero di un elevato numero di particelle, in parte effettuate nel 2012, in parte negli anni precedenti, concentrate per lo più in poche regioni, rischiava di compromettere il corretto calcolo dell'indice. Il peso attribuito a ciascun dato è stato calcolato in funzione sia del livello di copertura della regione amministrativa, sia di quello della regione ornitologica di appartenenza della particella a cui il dato afferisce;
3. sempre a livello di approfondimenti abbiamo cercato di evidenziare l'esistenza di andamenti significativamente differenti tra sei zone ornitologiche (vedi di seguito) confrontando, dove possibile, ovvero dove l'andamento è risultato certo, la tendenza di popolazione di ciascuna specie in ogni singola zona (Figura 5.1).
4. Utilizzando i dati prodotti dal progetto MITO2000, è stato possibile identificare, sulla base della presenza e dell'abbondanza di ogni singola specie, distinte zone ornitologiche presenti in Italia, ciascuna caratterizzata da specifici parametri qualitativi e di abbondanza riguardanti gli uccelli (Londi *et al.*, 2010). Per raggiungere questo obiettivo sono stati usati i dati semi-quantitativi relativi alle 103 specie target del progetto MITO2000, raccolti nel periodo 2000-2003. L'intero territorio italiano è stato quindi suddiviso in unità spaziali 10x10 Km, successivamente raggruppate per mezzo di *cluster analysis* a seconda del grado di somiglianza dei parametri ornitologici che le caratterizzavano. In totale sono state identificate 6 zone ornitologiche (Figura 5.1), ciascuna con le sue specie indicatrici:
 1. la zona alpina, caratterizzata dalla presenza dello spioncello, della bigiarella, dell'organetto, del merlo dal collare, del sordone, della cincia bigia alpestre e della nocciolaia;
 2. i rilievi prealpini e appenninici, caratterizzati dalla presenza della cincia mora e del picchio muratore;
 3. i sistemi collinari, caratterizzati dalla presenza del canapino comune, della sterpazzolina comune o della sterpazzolina di Moltoni e della tottavilla;
 4. le pianure, caratterizzate dalla presenza della cutrettola e del cannareccione;
 5. i rilievi mediterranei, caratterizzati dalla presenza della magnanina e del passero solitario;
 6. le "steppe" mediterranee, caratterizzate dalla presenza della calandrella, della calandra e della cappellaccia.

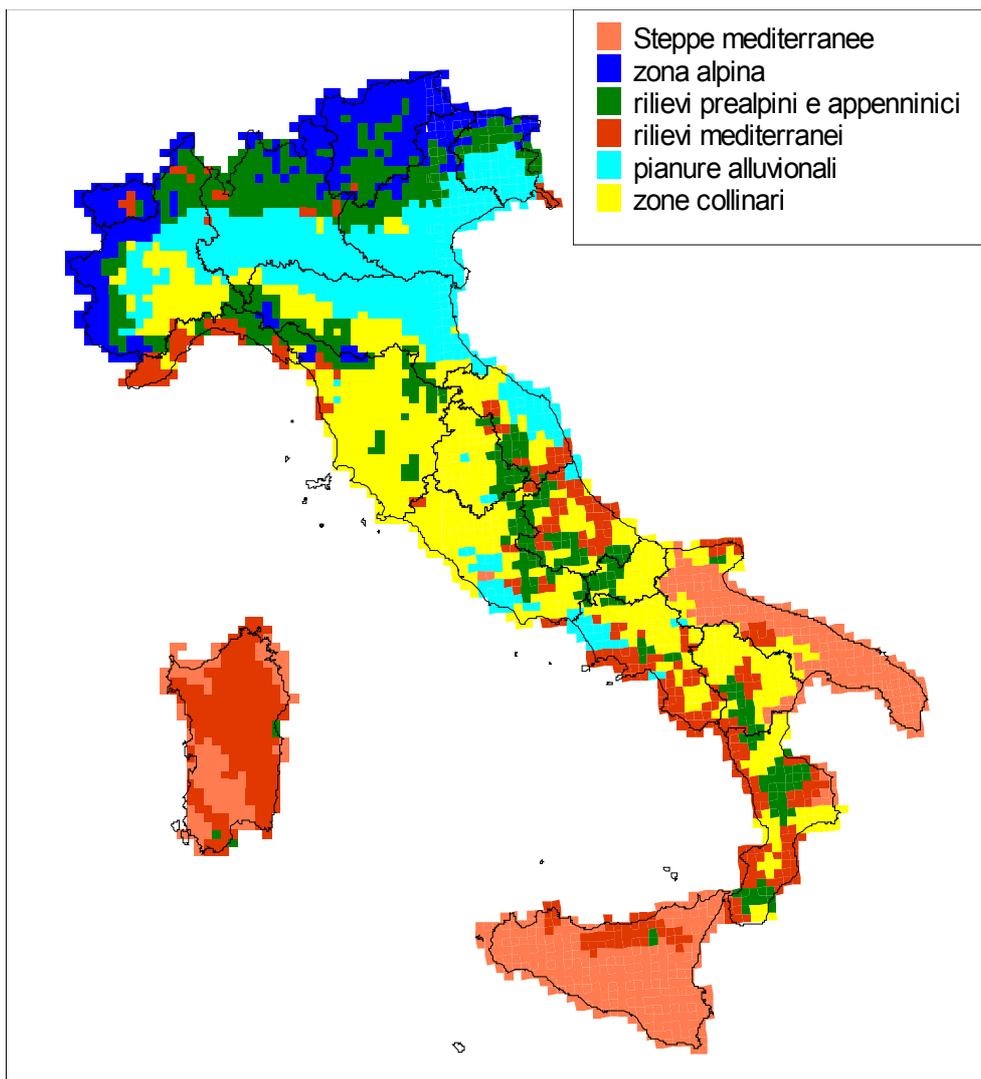


Figura 5.1: Distribuzione delle "zone ornitologiche" italiane.

6. CALCOLO DEGLI INDICI AGGREGATI

6.1. REGIONI

Gli indici compositi (FBI e WBI) vengono calcolati come media geometrica degli indici relativi alle singole specie (Gregory *et al.*, 2005; van Strien *et al.*, 2012).

Esistono numerosi indicatori che mirano a descrivere cambiamenti nella biodiversità ma non è ancora chiaro quale di essi sia più appropriato nell'intento. Una recente analisi delle proprietà matematiche di tali indicatori ha tuttavia dimostrato che la media geometrica è l'indicatore che si comporta meglio in relazione ad alcune desiderabili proprietà matematiche degli indicatori di biodiversità, con il solo punto debole di una elevata sensibilità all'aggiunta o all'eliminazione di alcune specie al sistema monitorato (van Strien *et al.*, 2012).

Una delle proprietà della media geometrica è quella di essere "robusta" in relazione all'influenza delle singole specie (Gregory & van Strien, 2010) e ciò grazie anche al fatto, nel caso specifico, che gli indicatori sono fatti sul logaritmo dei singoli indici. Questa proprietà può essere testata qualitativamente rimuovendo di volta in volta ognuna delle singole specie componenti l'indicatore e ricalcolando lo stesso (Figura 6.1).

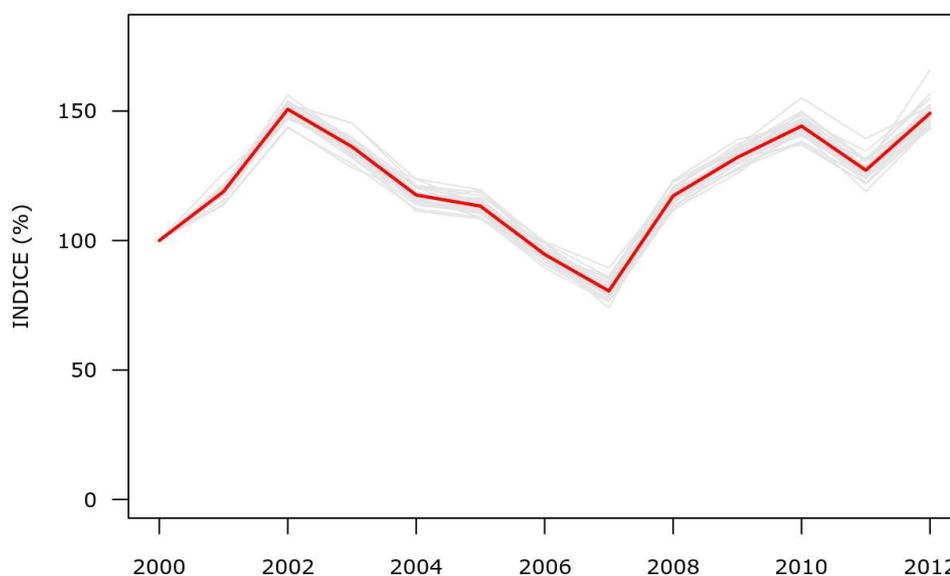


Figura 6.1: Una delle proprietà matematiche della media geometrica è quella di ben rappresentare l'andamento complessivo del set di specie utilizzate per il calcolo dell'indicatore composito. Nell'esempio in figura sono illustrati l'andamento dell'indicatore composito calcolato con tutte le specie (linea rossa) e gli andamenti dei diversi indicatori calcolati rimuovendo di volta in volta ognuna delle specie componente l'indicatore stesso. Lo scarto di questi ultimi dall'indicatore completo è molto ridotto.

Naturalmente maggiore è il numero di specie indicatrici utilizzate per il calcolo dell'indicatore composito e minore sarà l'influenza delle singole specie sull'indicatore.

Proprio per questa ragione nel calcolo degli indicatori regionali sono stati utilizzati anche gli indici relativi alle specie per le quali la tendenza demografica è stata classificata come incerta (vedi capitolo 4 – andamento incerto).

La media geometrica, come affermato in precedenza è sensibile alla scomparsa di specie

(valore dell'indice di una determinata specie in un determinato anno pari a zero). Le specie il cui indice risulti pari a zero in uno degli anni di indagine andrebbero dunque rimosse dal set delle specie indicatrici poichè la media geometrica di un insieme di numeri contenenti uno zero è pari a zero. Quando l'indice di una determinata specie scende sotto il 5%, in accordo con le indicazioni di EBCC, il suo valore nel calcolo dell'indice viene tenuto pari a 5%. Ciò al fine di non rimuovere specie dall'indicatore, garantendo che ognuna di esse possa mantenere la propria influenza sull'indicatore stesso.

6.2. ITALIA

Gli indici compositi (Farmland Bird Index, Indice delle specie delle praterie montane e Woodland Bird Index) vengono calcolati, come media geometrica degli indici relativi alle singole specie (Gregory *et al.*, 2005; van Strien *et al.*, 2012).

A scala nazionale, avendo a disposizione la maggior parte delle specie con andamenti definiti, si è scelto di non utilizzare per il calcolo degli indici aggregati le specie con andamento incerto (per l'elenco delle specie e i relativi andamenti si veda la Sezione "Italia – Farmland Bird Index, Woodland Bird Index e Andamenti di popolazione delle specie nel periodo 2000-2012").

7. BIBLIOGRAFIA

- Agresti A., 1990. *Categorical data analysis*. John Wiley, New York, USA.
- Blondel J., Ferry C., & Frochot B., 1981. Point counts with unlimited distance. *Stud. Avian Biol.*
- Fornasari L., de Carli E., Brambilla S., Buvoli L., Maritan E. e Mingozi T., 2002. Distribuzione dell'Avifauna nidificante in Italia: primo bollettino del progetto di Monitoraggio MITO2000. *Avocetta*, 26: 59-115.
- Gregory R.D. & van Strien A.J., 2010. Wild bird indicators: using composite population trends of birds as measures of environmental health. *Ornithol Sci* 9: 3-22.
- Gregory R.D., van Strien A., Vorisek P., Mayling A.W.G., Noble D.G., Foppen R.P.B. & Gibbons D.W., 2005. Developing indicators for European birds. *Philosophical Transactions of The Royal Society*, B 360: 269-288.
- Liang K.Y. & Zeger S., 1986. Longitudinal Data Analysis Using Generalized Linear Models. *Biometrika*, 73(1):13-22.
- Londi G., Tellini Florenzano G., Campedelli T. & Fornasari L., 2010. An ornithological zonation of Italy. In: Bermejo A. (ed.). *Bird Numbers 2010 "Monitoring, indicators and targets"*. Book of abstracts of the 18th Conference of the European Bird Census Council-SEO/BirdLife, Madrid, pp. 77.
- McCullagh P. and Nelder J.A., 1989. *Generalized Linear Models*. London: Chapman and Hall.
- Pannekoek J. & van Strien A.J., 2001. TRIM 3 Manual. TRends and Indices for Monitoring Data. Research paper No. 0102. Statistics Netherlands, Voorburg, The Netherlands.
- van Strien A.J., Pannekoek J. & Gibbons D.W., 2001. Indexing European bird population trends using results of national monitoring schemes: a trial of a new method. *Bird Study* 48: 200-213.
- van Strien A.J., Soldaat L.L. & Gregory R.D., 2012. Desirable mathematical properties of indicators for biodiversity change. *Ecological Indicators*, 14:202-208.