

DIPARTIMENTO DI CHIMICA, BIOLOGIA E BIOTECNOLOGIE DELL'UNIVERSITÀ  
DEGLI STUDI DI PERUGIA

e

ISTITUTO DI RICERCA PER LA PROTEZIONE IDROGEOLOGICA, DEL CONSIGLIO  
NAZIONALE DELLE RICERCHE

# ANALISI GIS RELATIVA DELL'IDONEITÀ AMBIENTALE DEL RETICOLO UMBRO PER LA FAUNA ITTICA

CONTRATTO TRA

IL DIPARTIMENTO DI CHIMICA, BIOLOGIA E BIOTECNOLOGIE DELL'UNIVERSITÀ DEGLI STUDI DI PERUGIA

E

L'ISTITUTO DI RICERCA PER LA PROTEZIONE IDROGEOLOGICA, DEL CONSIGLIO NAZIONALE DELLE RICERCHE

PER

L'"ANALISI GIS RELATIVA DELL'IDONEITÀ AMBIENTALE DEL RETICOLO UMBRO PER LA FAUNA ITTICA".

## INFORMAZIONI DOCUMENTO

IDENTIFICATIVO PRODOTTO	D01-01
TIPO PRODOTTO	RAPPORTO
RESPONSABILE WP	DOTT. IVAN MARCHESINI
AUTORI	DOTT. IVAN MARCHESINI

## STORICO REVISIONI DOCUMENTO

DATA	VERSIONE	MODIFICHE APPORTATE	AUTORI MODIFICHE
25/08/2025	2.1	TERZA EMISSIONE DOCUMENTO	DOTT. IVAN MARCHESINI

# SOMMARIO

Il documento illustra le attività svolte dall'Istituto di Ricerca per la Protezione Idrogeologica (IRPI) del Consiglio Nazionale delle Ricerche (CNR), nell'ambito del Contratto tra il Dipartimento di Chimica, Biologia e Biotecnologie dell'Università degli Studi di Perugia (DCBB-UNIPG) e l'Istituto di Ricerca per la Protezione Idrogeologica per "ANALISI GIS RELATIVA DELL'IDONEITÀ AMBIENTALE DEL RETICOLO UMBRO PER LA FAUNA ITTICA"

In particolare, il documento descrive le diverse attività svolte in relazione all'allegato tecnico del contratto.

# INDICE

Informazioni Documento.....	2
Storico Revisioni Documento.....	2
Sommario.....	3
Indice.....	4
Documenti Applicabili e di Riferimento.....	5
Acronimi e Abbreviazioni.....	6
1 Organizzazione ed elaborazione su gis di dati dati relativi alla distribuzione e abbondanza della fauna ittica.....	7
2 Dati acquisiti da fonti esterne.....	10
3 Organizzazione ed elaborazione su GIS di serie di dati meteo-climatici di origine diverse .....	11
4 Acquisizione, organizzazione ed elaborazione su GIS di serie di dati meteo-climatici relativi agli scenari di cambiamento climatico.....	12
5 Analisi GIS relativa ai cambiamenti climatici sul territorio regionale.....	12
6 Analisi esplorativa dei dati.....	21
7 Modellazione dell'abbondanza delle specie nel reticolo idrografico.....	29
7.1 Analisi della correlazioni tra le variabili.....	29
7.2 Training dei modelli.....	31
7.3 Applicazione dei modelli.....	34
8 References.....	55

# DOCUMENTI APPLICABILI E DI RIFERIMENTO

DA-001 Allegato tecnico al contratto oggetto del presente rapporto

## ACRONIMI E ABBREVIAZIONI

DCBB-UNIPG	Dipartimento di Chimica, Biologia e Biotecnologie dell'Università degli Studi di Perugia
CNR IRPI	Istituto di Ricerca per la Protezione Idrogeologica, del Consiglio Nazionale delle Ricerche
SIR	Sistema Informativo Regionale
ECA	European Climate Assessment
E-OBS	European daily high-resolution gridded dataset
ESGF	Earth System Grid Federation
DKRZ	German Climate Computing Center
CDO	Climate Data Operators (software)
GRASS GIS	Geographic Resources Analysis Support System (software)
EPSG	Geodetic Parameter Dataset
RCP	Representative Concentration Pathways

# 1 ORGANIZZAZIONE ED ELABORAZIONE SU GIS DI DATI DATI RELATIVI ALLA DISTRIBUZIONE E ABBONDANZA DELLA FAUNA ITTICA

Il DCBB-UNIPG ha fornito a CNR IRPI due set di dati i formato excell e un file in formato shapefile.

I due file in formato excell erano “ambientali+ittici\_3\_serie.xlsx” e “ambientali+ittici\_4\_serie.xlsx” e contenevano i dati ambientali e ittici rilevati da DCBB-UNIPG in stazioni di campionamento dislocate sull'intero reticolo idrografico umbro. Nel file con 3 serie di dati (“ambientali+ittici\_3\_serie.xlsx”), contenente 129 siti di campionamento, l'abbondanza delle specie ittiche è espressa in termini di densità, mentre nel file con 4 serie di dati (“ambientali+ittici\_4\_serie.xlsx”), contenente 117 siti di campionamento, l'arco temporale considerato è più esteso ma per le specie ittiche sono riportate le classi di abbondanza anziché le densità perché nel corso del primo monitoraggio non sono stati raccolti dati quantitativi.

In entrambi i file sono presenti le stesse variabili. Nelle tabelle seguenti vengono descritte distinguendo quelle ambientali da quelle relative all'abbondanza delle specie. Su entrambe le tabelle si descrive la presenza o assenza di lacune nei dati

Tabella 1: Variabili ambientali

Variabili ambientali e identificativi	Descrizione	Presenza lacune (3 serie)	Presenza lacune 4 serie
Sampling site code	Codice del sito (e.g. “01BAGN01”)		
Sampling period	Codice periodo (e.g. “1° livello”)		
Data	Data del campionamento		
Distanza dalla sorgente	Km		
Bacino sotteso	Kmq		
Quota	Metri sul livello del mare		
Tacqua	Gradi centigradi	si	si
pH	valore	si	si
Conducibilità	µS/cm		
O2	?	si	si

BOD	?	si	si
NNO3	?	si	si
NNO2	?	si	si
NNH3	?	si	si
SO4	?	si	si
PPO4	?	si	si
Cl	?	si	si
C.O.D.	?	si	si
PTot	?	si	si
EBI	?	si	si
Classe qualità EBI	?	si	si
Velocità	?	si	si
Portata	?	si	si

Tabella 2: Abbondanza specie ittiche

Specie	Lacune 3 serie	Lacune 4 serie	Specie	Lacune 3 serie	Lacune 4 serie
Alburnus alburnus alborella			Leuciscus lucumonis		
Anguilla anguilla			Leuciscus souffia		
Barbus barbus			Luciobarbu s graellsii		
Barbus plebejus			Micropteru s salmoides		
Barbus tyberinus			Oncorhyinc hus mykiss		
Blicca bjerckna			Padogobiu s martensi		
Carassius auratus			Padogobiu s nigricans		

Chondrostoma genei		Perca fluviatilis
Chondrostoma soetta		Pseudoras bora parva
Cyprinus carpio		Rhodeus sericeus
Cobitis taenia		Rutilus erythrophth almus
Cottus gobio		Rutilus rubilio
Ctenopharyngodon idellus		Rutilus rutilus
Esox lucius		Salmo trutta
Gambusia affinis		Salvelinus fontinalis
Gasterosteus aculeatus		Scardinius erythrophth almus
Gobio gobio		Silurus glanis
Gymnocephalus cernuus	si	Stizostedio n lucioperca
Ictalurus melas		Thymallus thymallus
Lepomis gibbosus		Tinca tinca
Leuciscus cephalus		
Squalius cephalus	si	

Il layer shapefile fornito è invece chiamato stazioni.shp  
il file contiene i seguenti campi:

Nome campo	Esempio contenuto
CODNEW	01amar01

COD	01AMAR01
FIUME	amara
X_COORD	311952,0000000000000000
y_COORD	4802189,0000000000000000
BACINO	chiascio

L'osservazione del contenuto delle due tabelle mette in evidenza che tra i layer geografico shapefile e i due file Excell esistono campi comuni che possono essere utilizzati per collegare le informazioni. In particolare i campi "Sampling site code" dei file Excell e il campo COD (o anche CODNEW) per il layer shapefile.

## 2 DATI ACQUISITI DA FONTI ESTERNE

Tra i DEM più accurati disponibili su scala nazionale il TINITALY (Tarquini et al., 2007; Tarquini & Nannipieri, 2017) rappresenta sicuramente uno dei prodotti più accurati ed è stato acquisito quindi ai fini del presente lavoro. Si tratta di un DEM con risoluzione 10 metri al suolo, costruito a partire da curve di livello e punti quotati delle carte tecniche regionali disponibili in Italia.

Il reticolo idrografico è un altro elemento importante per le finalità delle attività di progetto. Allo scopo di disporre di un reticolo accurato in termini di completezza e posizionamento si è deciso di acquisire il reticolo fluviale digitalizzato nell'ambito del database DBPrior10k realizzato dal Centro Interregionale di Coordinamento e Documentazione per le Informazioni Territoriali tra il 2003 ed il 2007 ( [https://www.cisis.it/?page\\_id=4078](https://www.cisis.it/?page_id=4078) ). Si tratta di un reticolo molto accurato, nel quale sono rappresentati anche i corsi d'acqua minori e in cui sono tracciati, con accuratezza, anche i canali con tracciato artificiale, modificati e rettificati dalle autorità, nelle aree di pianura alluvionale.

Al fine di derivare indici climatici storici per l'area di studio, sono stati acquisiti da Copernicus Climate Data Store, i dati E-OBS (Cornes et al., 2018). Si tratta di dati grigliati (formato raster) ottenuti, su scala europea, a partire dal monitoraggio di stazioni a terra. I dati riguardano temperatura media, minima e massima giornaliera e precipitazioni cumulate in 24 ore. Il grigliato ha risoluzione al suolo di circa 0.1 gradi, circa 9 km alle latitudini dell'area di studio). I dati sono disponibili come serie continua a partire dal 1 gennaio 1950 e fin al 31 dicembre 2021. I dati originali sono forniti in formato NetCDF con uno "scale factor" (fattore di riduzione finalizzato a trattare numeri interi e non decimali ai fini di minimizzare le dimensioni del dato) di 0.01 per le temperature e 0.1 per le precipitazioni.

Per quanto attiene le proiezioni climatiche l'attenzione è stata invece rivolta a dati, sempre disponibili su scala giornaliera, prodotti grazie ai risultati dell' "European Union FP6 integrated Project "WATCH", del EU FP6 Integrated Project "ENSEMBLES" oltre che dall'"European Commission Joint Research Centre PESETA II project" (Dosio et al., 2012; Dosio & Paruolo, 2011). I dati sono stati richiesti agli autori e forniti in formato NetCDF. Si tratta in realtà di un "ensemble" di 11 "Regional Climate Models" (RCM) che sono stati

derivati a partire da diversi “Global Climate Models” (GCM) e sottoposti a tecniche di “bias correction” al fine di rimuovere errori (sovrastime o sottostime evidenti) che derivano dal “downscaling” del GCM in RCM. I dati sono stati calcolati sia su base storica (al fine di verificare la capacità delle tecniche di modellare i dati esistenti, sia su base future, sotto l’assunzione di “Representative Concentration Pathway (RCP)” 8.5.

### 3 ORGANIZZAZIONE ED ELABORAZIONE SU GIS DI SERIE DI DATI METEO-CLIMATICI DI ORIGINE DIVERSE

I dati E-OBS sono stati registrati all’interno di una location di GRASS GIS<sup>1</sup> definita nel sistema di riferimento EPSG:4326. Sono stati utilizzati i dati di temperatura media (tg), massima (tx) e minima (tn) e i dati di precipitazione (rr) ottenibili dalla specifica piattaforma Copernicus<sup>2</sup>. I dati sono stati corretti per lo *scalefactor* sopra ricordato.

L’analisi dei dati relativi al dataset E-OBS è stata svolta al fine di disporre di informazioni utili per caratterizzare il clima passato e presente. I calcoli sono stati effettuati su base raster, calcolando quindi per ogni cella il valore dell’indice di interesse, inoltre gli indici sono stati calcolati per più periodi: 1961-1990, 1991-2020, 1976-1990, 2006-2020. Più in particolare, utilizzando le serie temporali di temperature e precipitazioni sono stati calcolati i seguenti indici climatici (mappe):

- 1) Rx1day Maximum 1-day precipitation Let RR<sub>ij</sub> be the daily precipitation amount on day i in period j. The maximum 1-day value for period j are Rx1day<sub>j</sub> = max (RR<sub>ij</sub>)
- 2) Rx5day Maximum consecutive 5-day precipitation Let RR<sub>kj</sub> be the precipitation amount for the 5-day interval ending k, period j. Then maximum 5-day values for period j are Rx5day<sub>j</sub> = max (RR<sub>kj</sub>)
- 3) SDII (intensità media nei giorni piovosi, in cui la pioggia è superiore a 1 mm)
- 4) #R10mm (numero di giorni con pioggia sopra a 10 mm)
- 5) #R20mm (numero di giorni con pioggia sopra a 20 mm)
- 6) #PRCPTOT (precipitazione cumulata media annua)
- 7) #FD Annual count of days when TN (daily minimum temperature) < 0°C
- 8) #SU Number of summer days Annual count of days when TX (daily maximum temperature) > 25°C
- 9) #ID Number of icing days Annual count of days when TX (daily maximum temperature) < 0 °C
- 10) #TR Number of tropical nights Annual count of days when TN (daily minimum temperature) > 20 °C
- 11) temperature medie stagionali: Dicembre, Gennaio, Febbraio (DJF) – Marzo, Aprile Maggio (MAM) – Giugno, Luglio, Agosto (JJA) – Settembre, Ottobre, Novembre (SON).
- 12) precipitazioni medie stagionali: Dicembre, Gennaio, Febbraio (DJF) – Marzo, Aprile Maggio (MAM) – Giugno, Luglio, Agosto (JJA) – Settembre, Ottobre, Novembre (SON).

<sup>1</sup><https://grass.osgeo.org/>

<sup>2</sup>[https://surfobs.climate.copernicus.eu/dataaccess/access\\_eobs.php#datafiles](https://surfobs.climate.copernicus.eu/dataaccess/access_eobs.php#datafiles)

## **4 ACQUISIZIONE, ORGANIZZAZIONE ED ELABORAZIONE SU GIS DI SERIE DI DATI METEO-CLIMATICI RELATIVI AGLI SCENARI DI CAMBIAMENTO CLIMATICO**

I dati climatici giornalieri relativi alle proiezioni climatiche hanno richiesto un procedimento più articolato.

Innanzitutto trattandosi di un “ensemble” di 11 modelli di proiezioni climatiche è stato necessario trattarli tutti e quindi ripetere le procedure per tutti i modelli. I dati sono stati forniti in formato NetCDF ma con un sistema di riferimento “rotated pole” (Griglia centrata fittiziamente sull’equatore) che è spesso utilizzato da chi si occupa di modellazione di cambiamenti climatici.

L’analisi del dato iniziale ha consentito di verificare che la definizione del sistema “rotated pole” soffriva inoltre di alcune inaccuratezze. La trasformazione in coordinate geografiche (EPSG:4326) è stata quindi eseguita assegnando prima ai dati NetCDF una più adeguata definizione del rotated pole (definizione tratta da altri modelli forniti dal nodo ESGF del DKRZ) e poi riproiettando i dati nel sistema EPSG:4326 mediante il software CDO (Climate Data Operators<sup>3</sup>).

Successivamente, analogamente a quanto fatto per i dati E-OBS, i dati sono stati registrati in GRASS GIS, sebbene sia stato necessario trattarli su 11 mapset differenti. L’analisi è stata eseguita su tutti i dati dello scenario RCP 8.5.

I dati di precipitazione nel dato originale sono espressi in termini di flusso di precipitazione ( $\text{kg m}^{-2} \text{s}^{-1}$ ) e sono stati quindi convertiti in mm/giorno moltiplicando il valore per 86400.

Anche i dati di temperatura sono stati corretti per portarli alla scala Celsius, sottraendo il valore 273.15 in quanto originariamente espressi in gradi Kelvin.

Anche in questo caso sono stati utilizzati i dati di temperatura media (tg), massima (tx) e minima (tn) e i dati di precipitazione (rr). Sono poi stati analizzati due intervalli temporali nel futuro: 2031-2065 e 2066-2100, calcolando tutte le 12 variabili descritte per i dati E-OBS.

## **5 ANALISI GIS RELATIVA AI CAMBIAMENTI CLIMATICI SUL TERRITORIO REGIONALE**

I dati relativi alla distribuzione e abbondanza della fauna ittica sul territorio regionale sono stati ulteriormente elaborati al fine di incrementare informazioni ad essi connesse.

Il lavoro si è principalmente basato sull’utilizzo del modello digitale di terreno TINITALY a 10 metri di risoluzione sul suolo e del reticolo fluviale digitalizzato nell’ambito del database DBPrior10k.

Utilizzando un apposito strumento di calcolo di GRASS GIS (r.carve) il reticolo idrografico è stato utilizzato per alterare il modello TINITALY e fare in modo che il reticolo fluviale fosse posizionato nelle corrette coordinate anche in aree pianeggianti in cui i reticoli derivanti dal DEM possono discostarsi da quelli reali (talvolta canalizzati/rettificati

<sup>3</sup><https://code.mpimet.mpg.de/projects/cdo>

artificialmente). Tecnicamente nelle celle del DEM interessate dalla presenza del reticolo si è determinata una incisione di circa 4 metri e pari a una larghezza di 20 metri. Questo ha poi consentito di generare, mediante l'utilizzo di un altro strumento di GRASS GIS (r.watershed) un reticolo idrografico raster denso del territorio umbro, andando a identificare tutte le aste idrografiche caratterizzate da bacini idrografici sottesi superiori a 100000 m<sup>2</sup> (0.1 km<sup>2</sup>).

Successivamente è stato costruito uno script in ambiente GRASS GIS che ha consentito di:

- associare ogni punto (stazione di campionamento) contenuto nel file stazioni.shp alla cella più prossima del reticolo idrografico raster
- associare ad ogni punto il valore del gradiente locale del corso d'acqua.

Per fare questa ultima operazione ci si è avvalsi di uno specifico tool di GRASS GIS (r.stream.slope) che consente di calcolare la pendenza del corso d'acqua su ogni cella raster che lo descrive. Il valore associato ad ogni stazione di campionamento è stato calcolato come il valore di pendenza medio derivante dalla misure della cella su cui la stazione insiste e da quelle immediatamente a monte e a valle della stessa.

Ad ogni stazione di campionamento sono stati poi associati i dati forniti da DCBB-UNIPG in merito al campionamento delle specie e al loro

Come risultato finale di queste elaborazioni sono stati quindi ottenuti due layer geografici vettoriali, corrispondenti alla "serie4" e alla "serie3" così come definiti nella sezione 1 del presente documento. Solo per poche stazioni non è stato possibile associare in maniera certa, la geolocalizzazione al relativo corso d'acqua.

La serie 4 risulta popolata da 114 stazioni di campionamento (contro le 117 iniziali). Per ogni stazione sono disponibili 4 campionamenti (effettuati in 4 diversi periodi) per un totale di 456 osservazioni.

La serie 3 risulta popolata da 126 stazioni di campionamento (contro le 129 iniziali). Per ogni stazione sono disponibili 3 campionamenti (effettuati in 3 diversi periodi) per un totale di 378 osservazioni.

Per ogni osservazione sono disponibili informazioni inerenti:

- la distanza delle coordinate fornite dal corso d'acqua (ottenuto dal DEM) più prossimo;
- il nome del corso d'acqua (da reticolo DBPrior10k) più prossimo
- una nota sull'incertezza dell'attribuzione del punto di campionamento all'asta fluviale
- il gradiente dell'asta fluviale in prossimità del punto di campionamento
- il codice del punto di campionamento (CODNEW)
- il nome del corso d'acqua originale (come da dati forniti da DCBB-UNIPG)
- area del bacino a monte del punto di campionamento (calcolata automaticamente dal DEM)
- la quota del DEM
- il periodi di campionamento (3 periodi per la Serie 3 e 4 periodi per la serie 4)
- la data del campionamento (espressa come timestamp YYYY-MM-DDT:00:00:00.000)

- le informazioni sulle specie campionate (41 colonne con dati di presenza assenza per la serie 4 e di densità per la serie 3)
- distanza dalla sorgente (come da dati forniti da DCBB-UNIPG ma non completo: 6 dati mancanti)
- bacino sotteso (come da dati forniti da DCBB-UNIPG ma non completo: 6 dati mancanti)
- quota (come da dati forniti da DCBB-UNIPG)
- temperatura dell'acqua (come da dati forniti da DCBB-UNIPG ma non completo: 15 dati mancanti per Serie4 e Serie 3)
- ph (come da dati forniti da DCBB-UNIPG ma non completo: 4 dati mancanti)
- conducibilità (come da dati forniti da DCBB-UNIPG)
- ossigeno (come da dati forniti da DCBB-UNIPG ma non completo: 15 dati mancanti per serie3 e 14 per serie 4)
- ossigeno (come da dati forniti da DCBB-UNIPG ma non completo: 15 dati mancanti per serie3 e 14 per serie 4)
- Domanda biochimica di ossigeno (come da dati forniti da DCBB-UNIPG ma non completo: 22 dati mancanti per serie 3 e 135 per serie 4)
- nno3 (come da dati forniti da DCBB-UNIPG ma non completo: 4 dati mancanti per Serie3 e 133 per Serie4)
- nno2 (come da dati forniti da DCBB-UNIPG ma non completo: 4 dati mancanti per Serie3 e 133 per Serie4)
- nnh3 (come da dati forniti da DCBB-UNIPG ma non completo: 4 dati mancanti per Serie3 e 20 per Serie4)
- solfati (come da dati forniti da DCBB-UNIPG ma non completo: 4 dati mancanti per Serie3 e 25 per Serie4)
- fosfati (come da dati forniti da DCBB-UNIPG ma non completo: 4 dati mancanti per Serie3 e 70 per Serie4)
- cloro (come da dati forniti da DCBB-UNIPG ma non completo: 4 dati mancanti per Serie3 e 18 per Serie4)
- domanda chimica di ossigeno (come da dati forniti da DCBB-UNIPG ma non completo: 4 dati mancanti per Serie3 e 166 per Serie4)
- fosforo totale (come da dati forniti da DCBB-UNIPG ma non completo: 4 dati mancanti per Serie3 e 157 per Serie4)
- indice biotico esteso (come da dati forniti da DCBB-UNIPG ma non completo: 37 dati mancanti per Serie3 e 28 per Serie4)
- classe qualità ebi (come da dati forniti da DCBB-UNIPG ma non completo: 37 dati mancanti per Serie3 e 28 per Serie4)
- velocità flusso (come da dati forniti da DCBB-UNIPG ma non completo: 78 dati mancanti per Serie3 e 73 per Serie4)
- portata (come da dati forniti da DCBB-UNIPG ma non completo: 70 dati mancanti per Serie3 e 66 per Serie4)

Ad ogni Serie sono associati i valori di presenza assenza delle specie ittiche riportate in tabella 2.

Inoltre entrambe le serie contengono valori di indici climatici calcolati sul dataset e-obs e sull'ensemble di 11 modelli climatici. Va fatto notare che gli indici climatici sono stati costruiti sia per caratterizzare le condizioni ambientali al momento del campionamento, sia per caratterizzare il clima più generale per diversi intervalli di tempo: 1961-1990, 1991-2020, 1976-1990, 2006-2020, 2031-2065, 2066-2100. I risultati dell'ensemble di modelli climatici sono stati riassunti con tre indicatori statistici: i quantili 0.1 (01), 0.5 (05) e 0.9 (09) che forniscono indicazioni sui valori di tendenza centrale e sulla dispersione dei valori.

#### Informazioni relative alla data di campionamento

- *precipitazioni cumulate su diverse durate prima della data di campionamento*
  - rr\_cum\_1days
  - rr\_cum\_3days
  - rr\_cum\_7days
  - rr\_cum\_14days
  - rr\_cum\_28days
- *temperatura media giornaliera mediata su diverse durate prima della data di campionamento*
  - tg\_avg\_1days
  - tg\_avg\_3days
  - tg\_avg\_7days
  - tg\_avg\_14days
  - tg\_avg\_28days
- *temperatura massima giornaliera mediata su diverse durate prima della data di campionamento*
  - tx\_avg\_1days
  - tx\_avg\_3days
  - tx\_avg\_7days
  - tx\_avg\_14days
  - tx\_avg\_28days
- *temperatura minima giornaliera mediata su diverse durate prima della data di campionamento*
  - tn\_avg\_1days
  - tn\_avg\_3days
  - tn\_avg\_7days
  - tn\_avg\_14days
  - tn\_avg\_28days

#### Informazioni climatiche dal 1961 al 1990, dal 1991 al 2020, dal 1976 al 1990 e dal 2006 al 2020

- *massima precipitazione cumulata giornaliera*

- rx1day\_1961\_1990
- rx1day\_1991\_2020
- rx1day\_1976\_1990
- rx1day\_2006\_2020
- *massima precipitazione cumulata su 5 giorni*
  - rx5day\_1961\_1990
  - rx5day\_1991\_2020
  - rx5day\_1976\_1990
  - rx5day\_2006\_2020
- *intensità media nei giorni piovosi, in cui la pioggia è superiore a 1 mm*
  - sdii\_1961\_1990
  - sdii\_1991\_2020
  - sdii\_1976\_1990
  - sdii\_2006\_2020
- *numero di giorni con pioggia sopra a 10 mm e sopra a 20 mm*
  - r10mm\_1961\_1990
  - r10mm\_1991\_2020
  - r20mm\_1961\_1990
  - r20mm\_1991\_2020
  - r10mm\_1976\_1990
  - r10mm\_2006\_2020
  - r20mm\_1976\_1990
  - r20mm\_2006\_2020
- *precipitazione media annua*
  - prcptot\_1961\_1990
  - prcptot\_1991\_2020
  - prcptot\_1976\_1990
  - prcptot\_2006\_2020
- *Numero di giorni in cui la temperatura minima giornaliera è inferiore a 0°C*
  - fd\_1961\_1990
  - fd\_1991\_2020
  - fd\_1976\_1990
  - fd\_2006\_2020
- *Numero di giorni in cui la temperatura massima giornaliera è superiore a 25°C*
  - su\_1961\_1990
  - su\_1991\_2020
  - su\_1976\_1990
  - su\_2006\_2020

- *Numero di giorni in cui la temperatura massima giornaliera è inferiore a 0°C*
  - id\_1961\_1990
  - id\_1991\_2020
  - id\_1976\_1990
  - id\_2006\_2020
- *Numero di giorni in cui la temperatura minima giornaliera è maggiore di 25°C*
  - tr\_1961\_1990
  - tr\_1991\_2020
  - tr\_1976\_1990
  - tr\_2006\_2020
- *Temperatura media nei trimestri Mar-May (03), Jun-Aug (06), Sep-Nov (09), Dec-Feb (12)*
  - tg\_1961\_1990\_average\_03
  - tg\_1961\_1990\_average\_06
  - tg\_1961\_1990\_average\_09
  - tg\_1961\_1990\_average\_12
  - tg\_1991\_2020\_average\_03
  - tg\_1991\_2020\_average\_06
  - tg\_1991\_2020\_average\_09
  - tg\_1991\_2020\_average\_12
  - tg\_1976\_1990\_average\_03
  - tg\_1976\_1990\_average\_06
  - tg\_1976\_1990\_average\_09
  - tg\_1976\_1990\_average\_12
  - tg\_2006\_2020\_average\_03
  - tg\_2006\_2020\_average\_06
  - tg\_2006\_2020\_average\_09
  - tg\_2006\_2020\_average\_12
- *Cumulata di precipitazione media nei trimestri Mar-May (03), Jun-Aug (06), Sep-Nov (09), Dec-Feb (12)*
  - rr\_1961\_1990\_sum\_03
  - rr\_1961\_1990\_sum\_06
  - rr\_1961\_1990\_sum\_09
  - rr\_1961\_1990\_sum\_12
  - rr\_1991\_2020\_sum\_03
  - rr\_1991\_2020\_sum\_06
  - rr\_1991\_2020\_sum\_09
  - rr\_1991\_2020\_sum\_12

- rr\_1976\_1990\_sum\_03
- rr\_1976\_1990\_sum\_06
- rr\_1976\_1990\_sum\_09
- rr\_1976\_1990\_sum\_12
- rr\_2006\_2020\_sum\_03
- rr\_2006\_2020\_sum\_06
- rr\_2006\_2020\_sum\_09
- rr\_2006\_2020\_sum\_12

Informazioni climatiche dal 2031 al 2065 e dal 2066 al 2100

- *massima precipitazione cumulata giornaliera*
  - rx1day\_2031\_2065\_q01
  - rx1day\_2031\_2065\_q05
  - rx1day\_2031\_2065\_q09
  - rx1day\_2066\_2100\_q01
  - rx1day\_2066\_2100\_q05
  - rx1day\_2066\_2100\_q09
- *massima precipitazione cumulata su 5 giorni*
  - rx5day\_2031\_2065\_q01
  - rx5day\_2031\_2065\_q05
  - rx5day\_2031\_2065\_q09
  - rx5day\_2066\_2100\_q01
  - rx5day\_2066\_2100\_q05
  - rx5day\_2066\_2100\_q09
- *intensità media nei giorni piovosi, in cui la pioggia è superiore a 1 mm*
  - sdii\_2031\_2065\_q01
  - sdii\_2031\_2065\_q05
  - sdii\_2031\_2065\_q09
  - sdii\_2066\_2100\_q01
  - sdii\_2066\_2100\_q05
  - sdii\_2066\_2100\_q09
- *numero di giorni con pioggia sopra a 10 mm e sopra a 20 mm*
  - r10mm\_2031\_2065\_q01
  - r10mm\_2031\_2065\_q05
  - r10mm\_2031\_2065\_q09

- r10mm\_2066\_2100\_q01
- r10mm\_2066\_2100\_q05
- r10mm\_2066\_2100\_q09
- r20mm\_2031\_2065\_q01
- r20mm\_2031\_2065\_q05
- r20mm\_2031\_2065\_q09
- r20mm\_2066\_2100\_q01
- r20mm\_2066\_2100\_q05
- r20mm\_2066\_2100\_q09
- *precipitazione media annua*
  - prcptot\_2031\_2065\_q01
  - prcptot\_2031\_2065\_q05
  - prcptot\_2031\_2065\_q09
  - prcptot\_2066\_2100\_q01
  - prcptot\_2066\_2100\_q05
  - prcptot\_2066\_2100\_q09
- *Numero di giorni in cui la temperatura minima giornaliera è inferiore a 0°C*
  - fd\_2031\_2065\_q01
  - fd\_2031\_2065\_q05
  - fd\_2031\_2065\_q09
  - fd\_2066\_2100\_q01
  - fd\_2066\_2100\_q05
  - fd\_2066\_2100\_q09
- *Numero di giorni in cui la temperatura massima giornaliera è superiore a 25°C*
  - su\_2031\_2065\_q01
  - su\_2031\_2065\_q05
  - su\_2031\_2065\_q09
  - su\_2066\_2100\_q01
  - su\_2066\_2100\_q05
  - su\_2066\_2100\_q09
- *Numero di giorni in cui la temperatura massima giornaliera è inferiore a 0°C*
  - id\_2031\_2065\_q01
  - id\_2031\_2065\_q05
  - id\_2031\_2065\_q09
  - id\_2066\_2100\_q01
  - id\_2066\_2100\_q05
  - id\_2066\_2100\_q09

- *Numero di giorni in cui la temperatura minima giornaliera è maggiore di 25°C*
  - tr\_2031\_2065\_q01
  - tr\_2031\_2065\_q05
  - tr\_2031\_2065\_q09
  - tr\_2066\_2100\_q01
  - tr\_2066\_2100\_q05
  - tr\_2066\_2100\_q09
- *Temperatura media nei trimestri Mar-May (03), Jun-Aug (06), Sep-Nov (09), Dec-Feb (12)*
  - tg\_2031\_2065\_average\_03\_q01
  - tg\_2031\_2065\_average\_03\_q05
  - tg\_2031\_2065\_average\_03\_q09
  - tg\_2031\_2065\_average\_06\_q01
  - tg\_2031\_2065\_average\_06\_q05
  - tg\_2031\_2065\_average\_06\_q09
  - tg\_2031\_2065\_average\_09\_q01
  - tg\_2031\_2065\_average\_09\_q05
  - tg\_2031\_2065\_average\_09\_q09
  - tg\_2031\_2065\_average\_12\_q01
  - tg\_2031\_2065\_average\_12\_q05
  - tg\_2031\_2065\_average\_12\_q09
  - tg\_2066\_2100\_average\_03\_q01
  - tg\_2066\_2100\_average\_03\_q05
  - tg\_2066\_2100\_average\_03\_q09
  - tg\_2066\_2100\_average\_06\_q01
  - tg\_2066\_2100\_average\_06\_q05
  - tg\_2066\_2100\_average\_06\_q09
  - tg\_2066\_2100\_average\_09\_q01
  - tg\_2066\_2100\_average\_09\_q05
  - tg\_2066\_2100\_average\_09\_q09
  - tg\_2066\_2100\_average\_12\_q01
  - tg\_2066\_2100\_average\_12\_q05
  - tg\_2066\_2100\_average\_12\_q09
- *Cumulata di precipitazione media nei trimestri Mar-May (03), Jun-Aug (06), Sep-Nov (09), Dec-Feb (12)*
  - rr\_sum\_2031\_2065\_sum\_03\_q01
  - rr\_sum\_2031\_2065\_sum\_03\_q05
  - rr\_sum\_2031\_2065\_sum\_03\_q09

- rr\_sum\_2031\_2065\_sum\_06\_q01
- rr\_sum\_2031\_2065\_sum\_06\_q05
- rr\_sum\_2031\_2065\_sum\_06\_q09
- rr\_sum\_2031\_2065\_sum\_09\_q01
- rr\_sum\_2031\_2065\_sum\_09\_q05
- rr\_sum\_2031\_2065\_sum\_09\_q09
- rr\_sum\_2031\_2065\_sum\_12\_q01
- rr\_sum\_2031\_2065\_sum\_12\_q05
- rr\_sum\_2031\_2065\_sum\_12\_q09
- rr\_sum\_2066\_2100\_sum\_03\_q01
- rr\_sum\_2066\_2100\_sum\_03\_q05
- rr\_sum\_2066\_2100\_sum\_03\_q09
- rr\_sum\_2066\_2100\_sum\_06\_q01
- rr\_sum\_2066\_2100\_sum\_06\_q05
- rr\_sum\_2066\_2100\_sum\_06\_q09
- rr\_sum\_2066\_2100\_sum\_09\_q01
- rr\_sum\_2066\_2100\_sum\_09\_q05
- rr\_sum\_2066\_2100\_sum\_09\_q09
- rr\_sum\_2066\_2100\_sum\_12\_q01
- rr\_sum\_2066\_2100\_sum\_12\_q05
- rr\_sum\_2066\_2100\_sum\_12\_q09

## 6 ANALISI ESPLORATIVA DEI DATI

La disponibilità dei dati ricondotti alle stazioni di campionamento consente di effettuare alcune analisi preliminari.

Nei grafici seguenti viene mostrata la temperatura media estiva e invernale misurata o prevista alle stazioni di campionamento per i periodi 1961-1990, 1991-2020, 2031-2065, 2066-2100

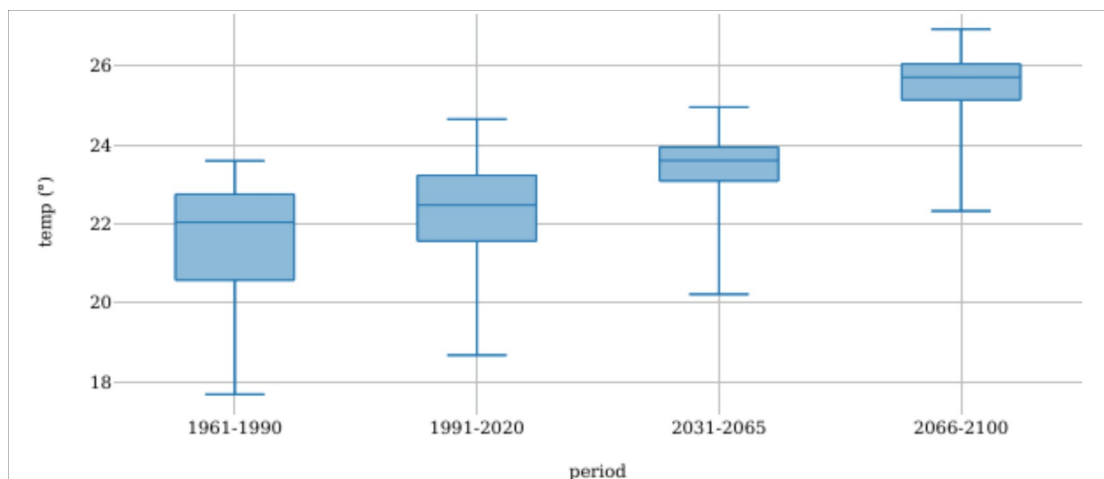


Fig. 1: Temperatura media nei mesi estivi in corrispondenza delle stazioni di campionamento

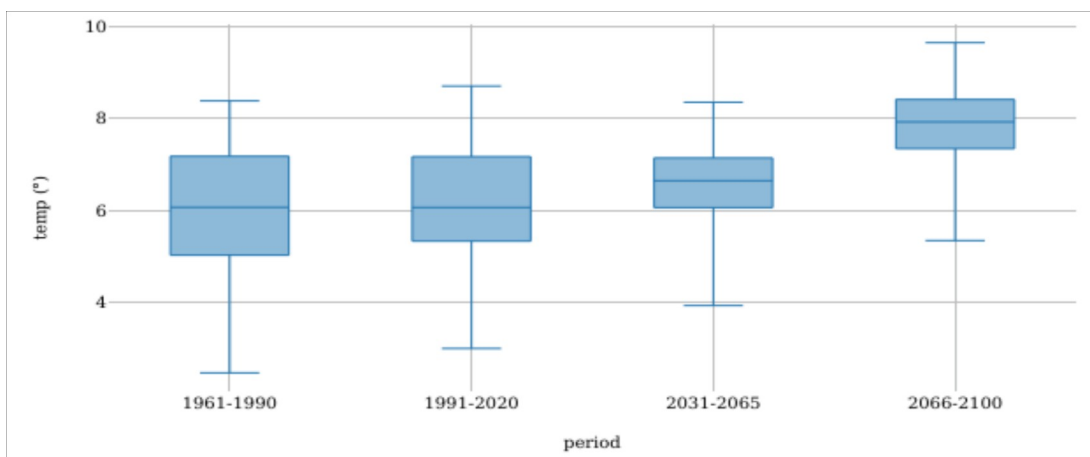


Fig. 2: Temperatura media nei mesi invernali in corrispondenza delle stazioni di campionamento

I dati mostrano un chiaro andamento crescente, più marcato nelle stagioni estive. Una analisi simile può essere effettuata sulle cumulate di precipitazione osservate nei mesi estivi ed invernali in corrispondenza delle stazioni di campionamento.

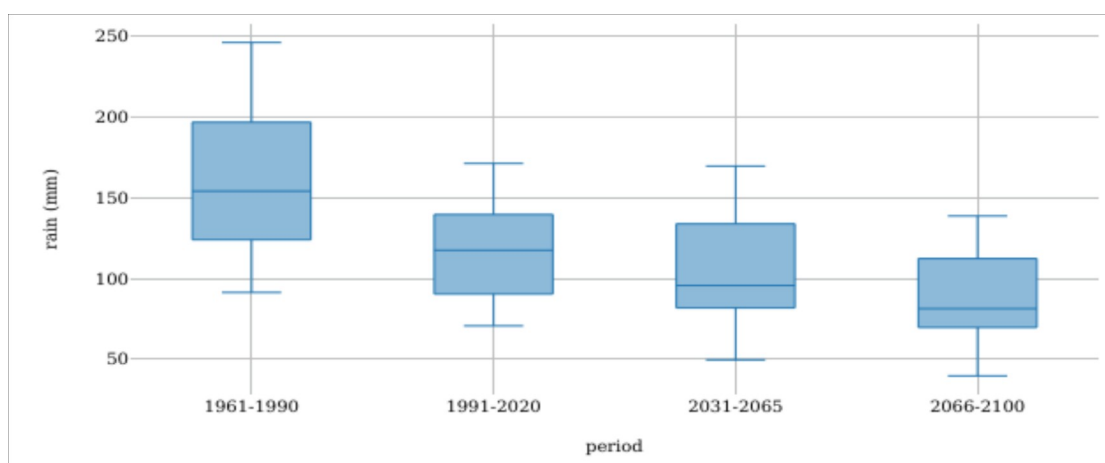


Fig. 3: Precipitazioni cumulate nei mesi estivi in corrispondenza delle stazioni di campionamento

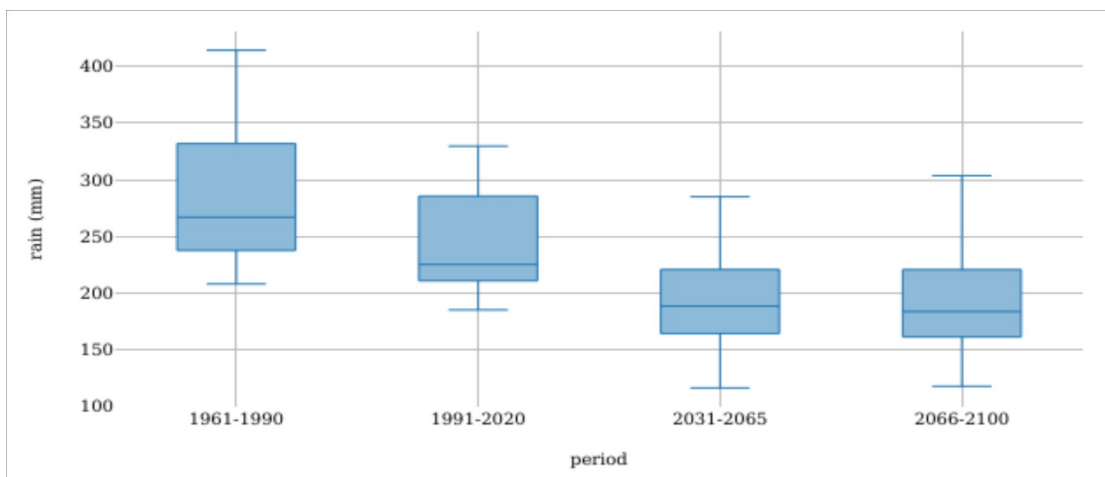


Fig. 4: Precipitazioni cumulate nei mesi invernali in corrispondenza delle stazioni di campionamento

I dati evidenziano una chiara tendenza alla diminuzione delle precipitazioni, con tendenze che sembrano significative.

Poichè nel database fornito da DCBB-UNIPG sono disponibili informazioni sulle date in cui sono stati effettuati i campionamenti e sono anche presenti i dati misurati di alcuni parametri ambientali relativi a quelle date (temperatura dell'acqua, portata idrica, velocità del flusso), sono state analizzate le potenziali relazioni tra questi dati misurati e i dati di temperatura e precipitazione derivati dal dataset e-obs. Tutti i grafici derivanti da questa analisi sono forniti in allegato al presente documento. Di seguito vengono presentati e discussi i risultati più significativi.

In Fig. 5 viene mostrata la relazione che esiste tra le temperature dell'acqua e dell'aria per i dati della Serie3 (disponibili nel periodo 1998-2017). Si osserva chiaramente che esiste una evidente correlazione tra i dati. Questo contribuisce a far ritenere che le variabili climatiche derivate dal dataset E-OBS siano capaci di rappresentare le condizioni ambientali del territorio di studio negli anni passati.

In Fig. 5 la stessa relazione viene mostrata per i dati della serie 4 (disponibili nel periodo 1989-2017). Anche in questo caso, sebbene con maggiore incertezza, si osserva un trend che ben correla temperatura dell'acqua e dell'aria.

Successivamente si è cercata anche una possibile correlazione tra portata misurata e precipitazioni. Un esempio è riportato in Fig. 7 in cui si mostra la relazione cercata tra la cumulata di precipitazione osservata al punto di campionamento e la portata misurata durante i campionamenti della Serie3. La scarsa correlazione osservata è probabilmente legata a una serie di fattori che includono: (i) la scarsa efficacia della pioggia puntuale nel determinare variazioni di portata sul corso d'acqua, (ii) la dimensione dei bacini drenanti al punto di campionamento che, essendo di dimensioni molto diverse, presentano tempi di corrivazione molto diversi tra loro, (iii) la presenza di afflussi sotterranei che, come nel caso del fiume Nera, contribuiscono a regolare i flussi superficiali e a renderli quasi indipendenti dalle trasformazioni afflussi-deflussi.

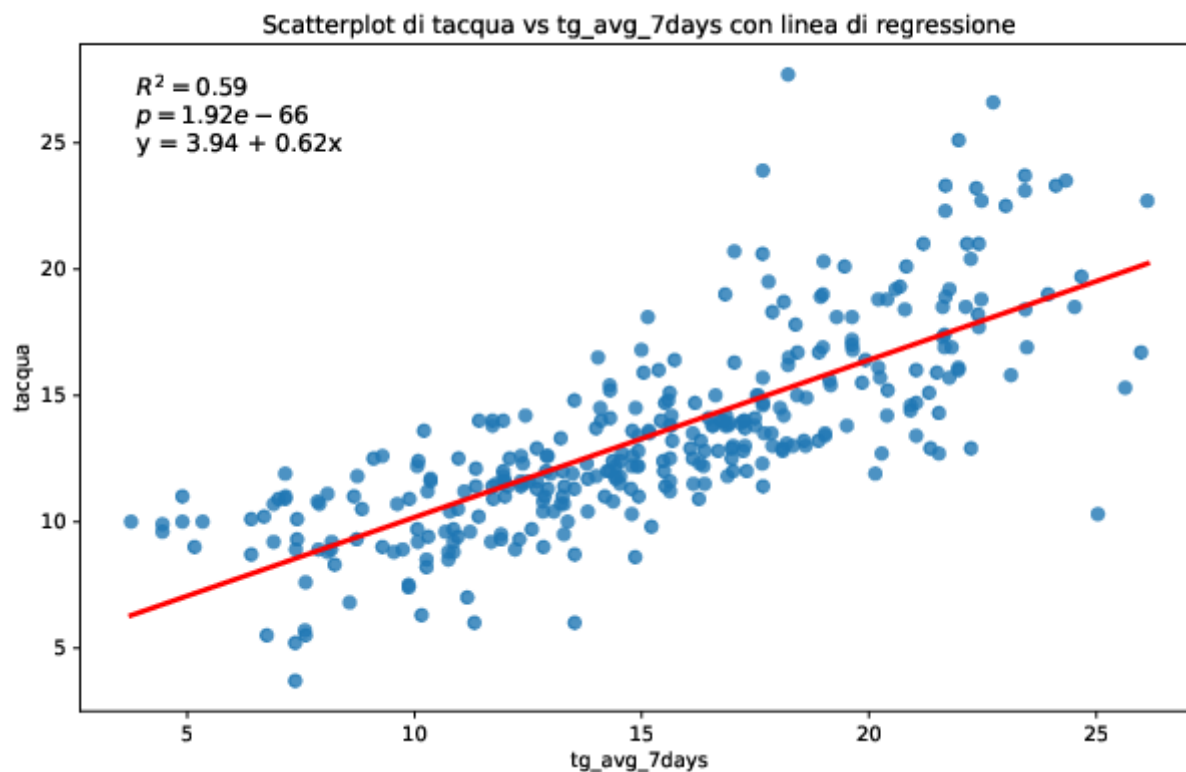


Fig. 5: relazione tra temperatura dell'acqua (tacqua, °) della Serie3 e temperatura dell'aria derivata da dati E-OBS. In particolare la temperatura media dell'aria nei 7 giorni precedenti il campionamento delle acque (tg\_avg\_7days, °)

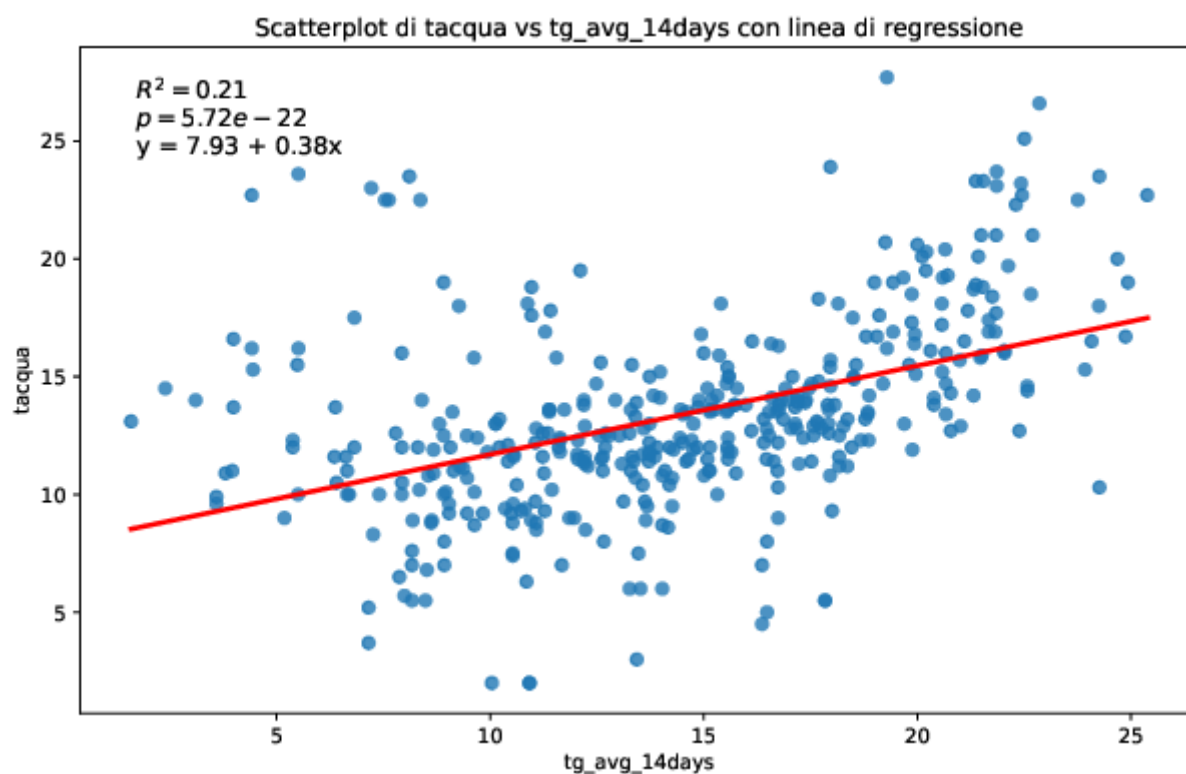


Fig. 6: relazione tra temperatura dell'acqua (tacqua, °) della Serie4 e temperatura dell'aria derivata da dati E-OBS. In particolare la temperatura media dell'aria nei 14 giorni precedenti il campionamento delle acque (tg\_avg\_14days, °)

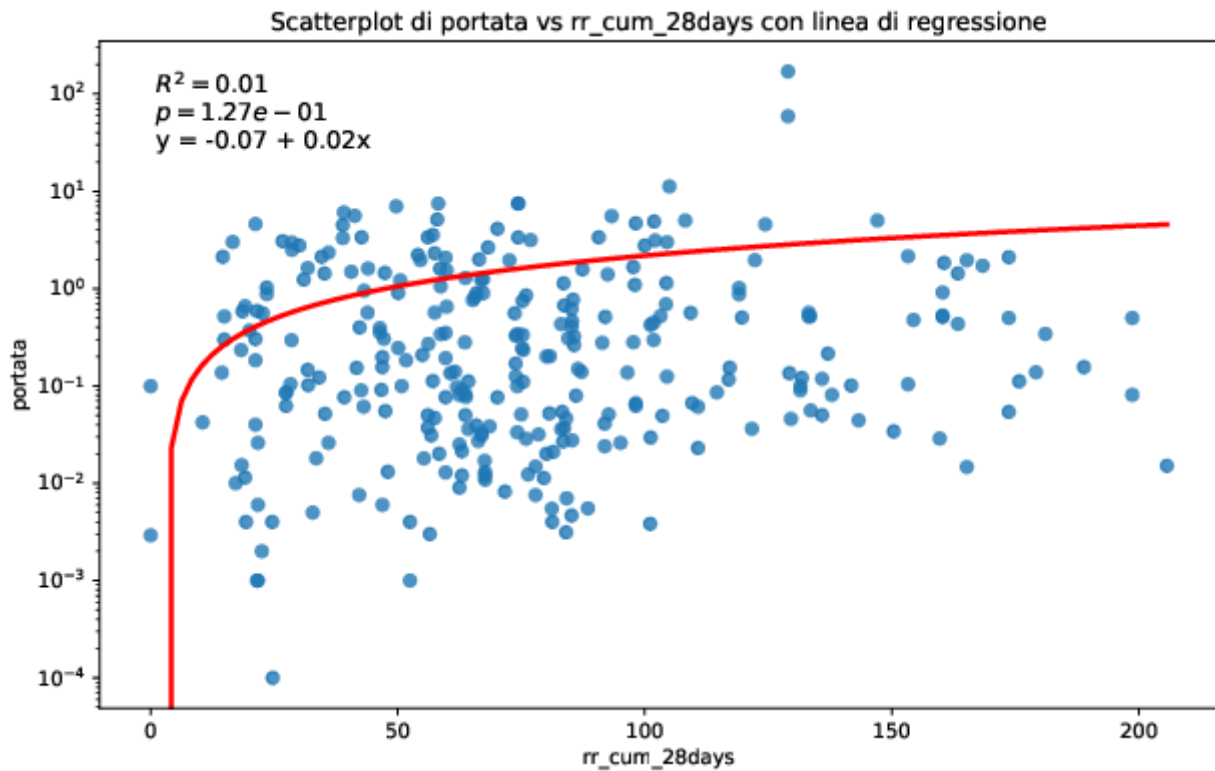


Fig. 7: relazione tra portata del corso d'acqua dell'acqua (portata, mc/s) della Serie3 e cumulata della precipitazione da dati E-OBS. In particolare viene mostrata la cumulata nei 28 giorni precedenti il campionamento delle acque (rr\_cum\_28days, mm)

La presenza di specie ittiche nei corsi d'acqua non deve però essere necessariamente ricondotta ai soli dati di temperatura e precipitazione osservati alle date di campionamento ma più genericamente alle condizioni climatiche e geomorfologiche più favorevoli al loro insediamento. Per verificare se possano essere osservate correlazioni tra l'abbondanza delle specie e le variabili climatiche/geomorfologiche, sono state eseguite una serie di analisi specifiche.

Le specie prese in considerazioni sono state le seguenti:

- leuciscus\_souffia
- leuciscus\_lucumonis
- rutilus\_rubilio
- cottus\_gobio
- padogobius\_nigricans
- leuciscus\_cephalus
- cobitis\_taenia
- salmo\_trutta
- barbus\_tyberinus

In Fig. 8 vengono mostrati i boxplot della area del bacino drenante (area accum, kmq) per le diverse classi di abbondanza delle specie (Serie4). Si osserva come alcune specie (cobitis\_taenia e leuciscus\_cephalus) si preentano soprattutto alla chiusura di bacini di dimensioni medie (e quindi probabilmente portate più elevate) e che altre specie sono

state campionate soprattutto in bacini di piccole dimensioni (e.g. *leuciscus\_souffia* e *rutilus\_rubilio*).

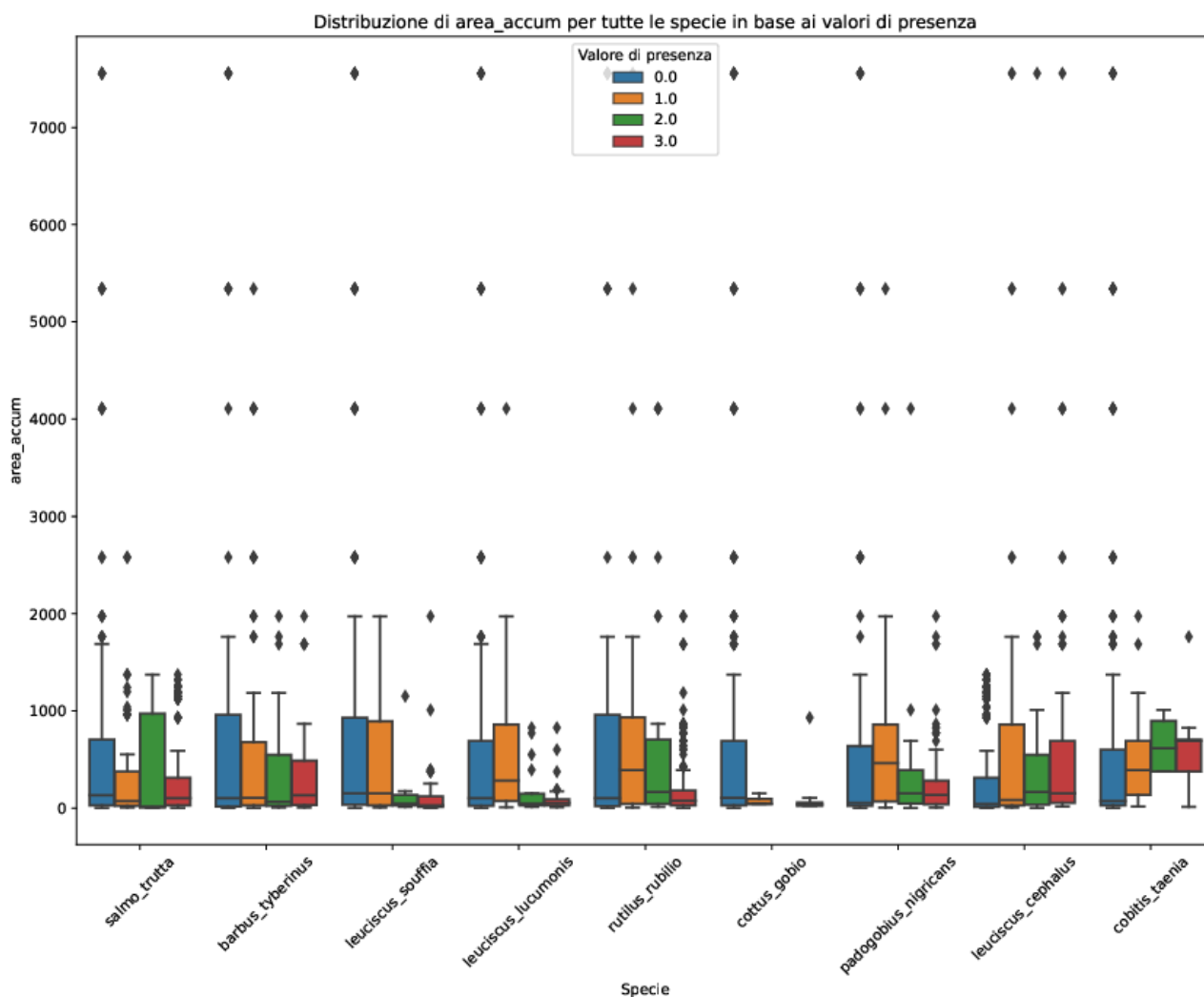
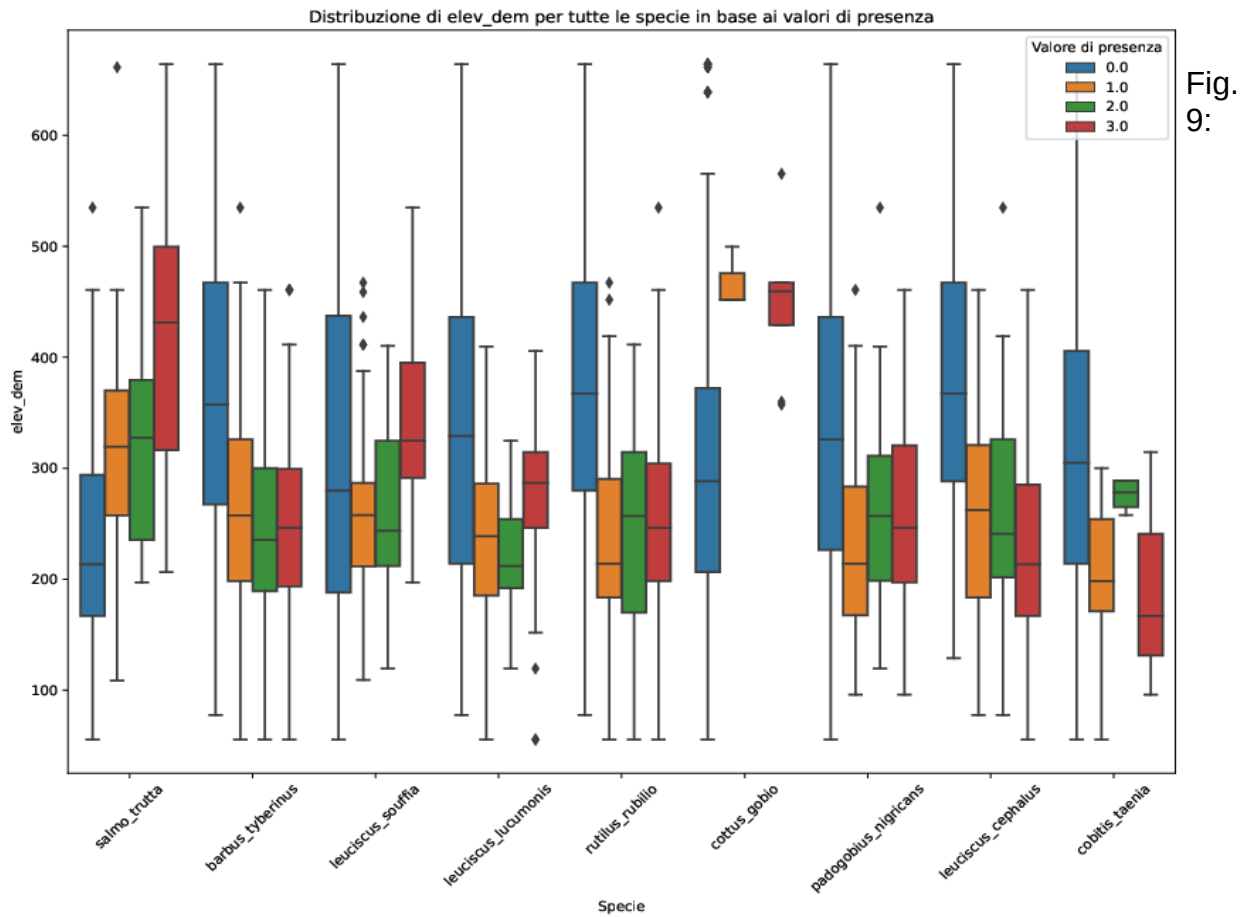


Fig. 8: Abbondanza delle specie (da 0 a 3 crescente) in funzione delle dimensioni del bacino a monte del punto di campionamento (area acc) espressa in kmq.

Anche la quota (Fig. 9) sembra avere un impatto sulla distribuzione della abbondanza delle specie. Si osserva infatti una maggiore abbondanza di *salmo\_trutta* e *cottus\_gobio* a quote più elevate, mentre un comportamento inverso si osserva per il *barbus\_tyberinus*.

Analogamente nelle Figure 10, 11 e 12 vengono mostrate le relazioni tra l'abbondanza delle specie e (i) la precipitazione totale media annua (prcptot\_1991-2020), (ii) la temperatura media nei tre mesi estivi (tg\_1991-202\_averae\_06), (iii) numero di giorni in cui la temperatura massima giornaliera è superiore a 25°C (su\_1991-2020). Per tutte le variabili sopra citate è possibile osservare una qualche correlazione con la presenza delle specie. La totalità dei grafici relativi a tutte le variabili climatiche e geomorfologiche sono riportati in allegato.



Abbondanza delle specie (da 0 a 3 crescente) in funzione della quota sul livello del mare (elev\_dem, m)

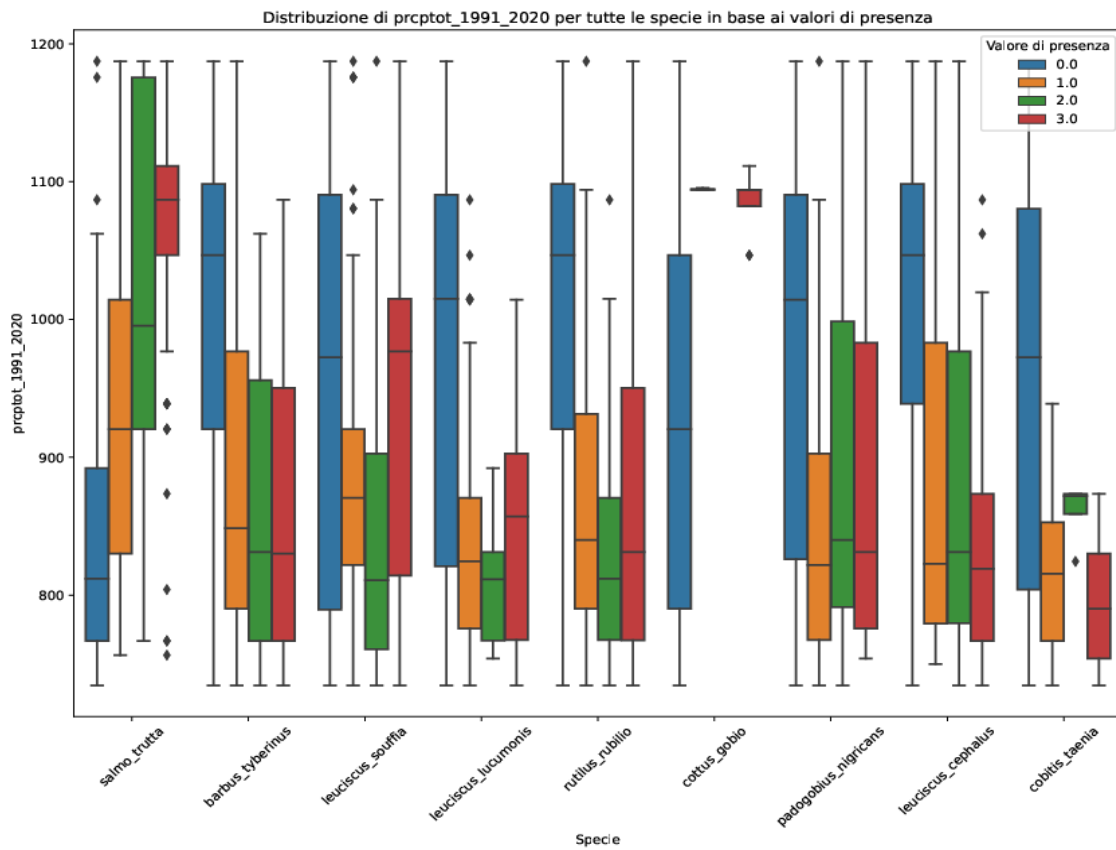


Fig. 10: Abbondanza delle specie (da 0 a 3 crescente) in funzione della precipitazione cumulata media annua (prcptot, mm)

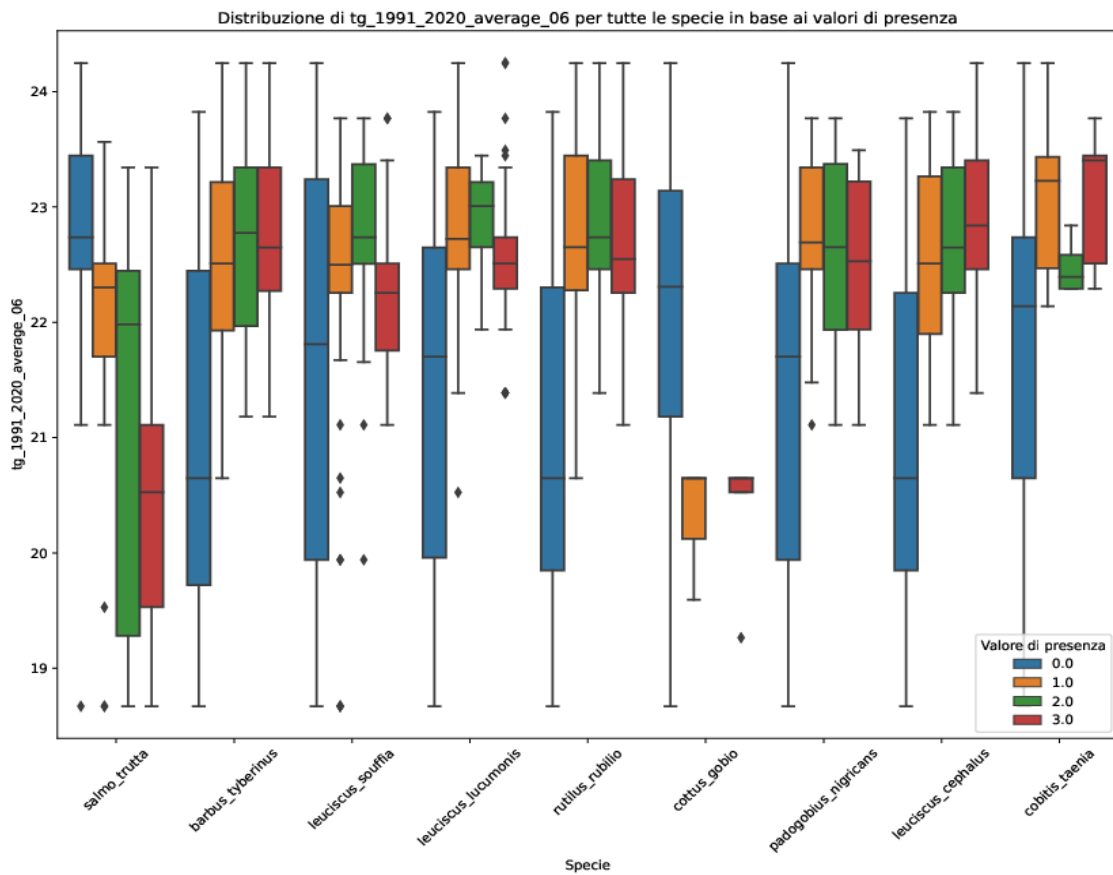


Fig. 11: Abbondanza delle specie (da 0 a 3 crescente) in funzione della temperatura media de mesi estivi (tg\_1991-2020\_average\_06, °)

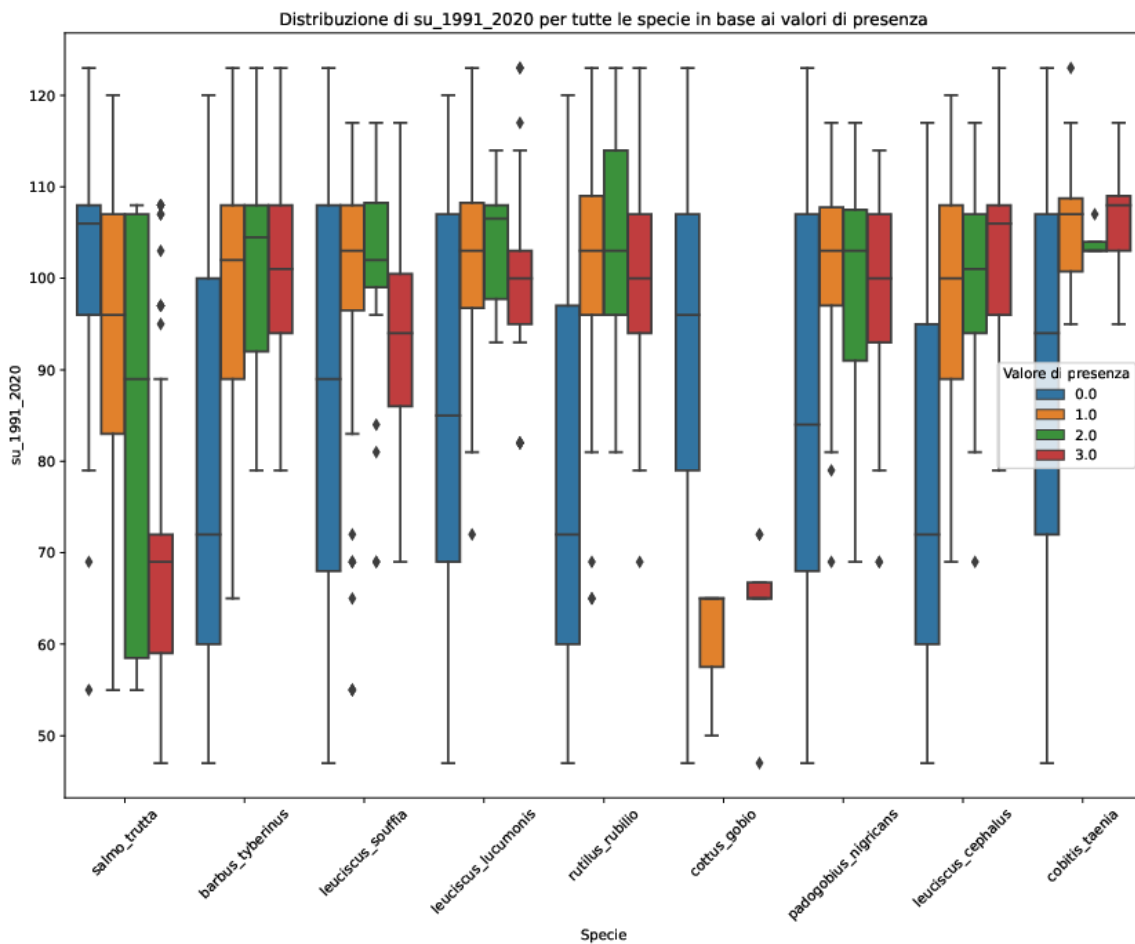


Fig. 12: Abbondanza delle specie (da 0 a 3 crescente) in funzione del numero di giorni in cui la temperatura massima giornaliera è superiore a 25°C (su\_1991-2020, #)

## 7 MODELLAZIONE DELL'ABBONDANZA DELLE SPECIE NEL RETICOLO IDROGRAFICO

La previsione della abbondanza delle specie sul reticolo idrografico è stata effettuata allenando modelli di machine learning sui dati forniti da DCBB-UNIPG e sulle variabili climatiche e geomorfologiche generate nell'ambito del presente lavoro.

### 7.1 ANALISI DELLA CORRELAZIONI TRA LE VARIABILI

Al fine di comprendere se qualche variabile potesse essere esclusa dall'analisi perché troppo correlata con altre variabili, è stata innanzi tutto costruita una matrice di correlazione (Fig. 13) tra le variabili che possono essere utilizzate per allenare modelli supervisionati con i dati della Serie4.

La matrice mostra che ci sono variabili tra loro molto correlate (coefficiente > 0.9). In particolare si osserva come l'intensità media dei giorni piovosi (sdii\_1991\_2020) sia molto correlata con le cumulate medie dei mesi autunnali e invernali (rr\_1991\_2020\_sum\_09, rr\_1991\_2020\_sum\_12). Il numero di giorni con pioggia sopra a 10 mm (r10mm\_1991\_2020) è correlato con tutte le cumulate medie trimestrali (rr\_1991\_2020\_sum\_03, rr\_1991\_2020\_sum\_06, rr\_1991\_2020\_sum\_09, rr\_1991\_2020\_sum\_12) e lo stesso accade per il numero di giorni con pioggia sopra a 20 mm (r20mm\_1991\_2020), a meno della variabile rr\_1991\_2020\_sum\_06 (cumulata nei

mesi estivi). La precipitazione media annua (prcptot\_1991\_2020) si correla con tutte le cumulate stagionali (anche se meno con quella estiva).

Le variabili r10mm\_1991\_2020 e r20mm\_1991\_2020 e prcptot\_1991\_2020 sono tra loro correlate e r20mm\_1991\_2020 è correlata con sdii\_1991\_2020. La massima pioggia cumulata in 5 giorni (rx5day\_1991-2020) è anch'essa correlata con sdii\_1991\_2020.

Le variabili che riguardano le temperature medie stagionali (tg\_1991\_2020\_average\_03, tg\_1991\_2020\_average\_06, tg\_1991\_2020\_average\_09, tg\_1991\_2020\_average\_12) sono tra loro fortemente correlate e sono molto correlate con il numero di giorni in cui la temperatura massima giornaliera è superiore a 25°C (su\_1991\_2020).

Piogge e temperature sono in genere correlate inversamente. Si osserva dal grafico mediante le relazioni (i) tra le temperature medie stagionali e le precipitazioni medie primaverili e autunnali e (ii) tra le precipitazioni medie annuali e il numero di giorni in cui la temperatura massima giornaliera è superiore a 25°C o le temperature medie stagionali.

Va sottolineato che si sta trattando di indici climatici trentennali e che quindi le correlazioni dirette o inverse che si osservano sono legate alla posizione del campionamento sul territorio in cui i dati sono stati acquisiti e non indicano, in questo caso, trend temporali.

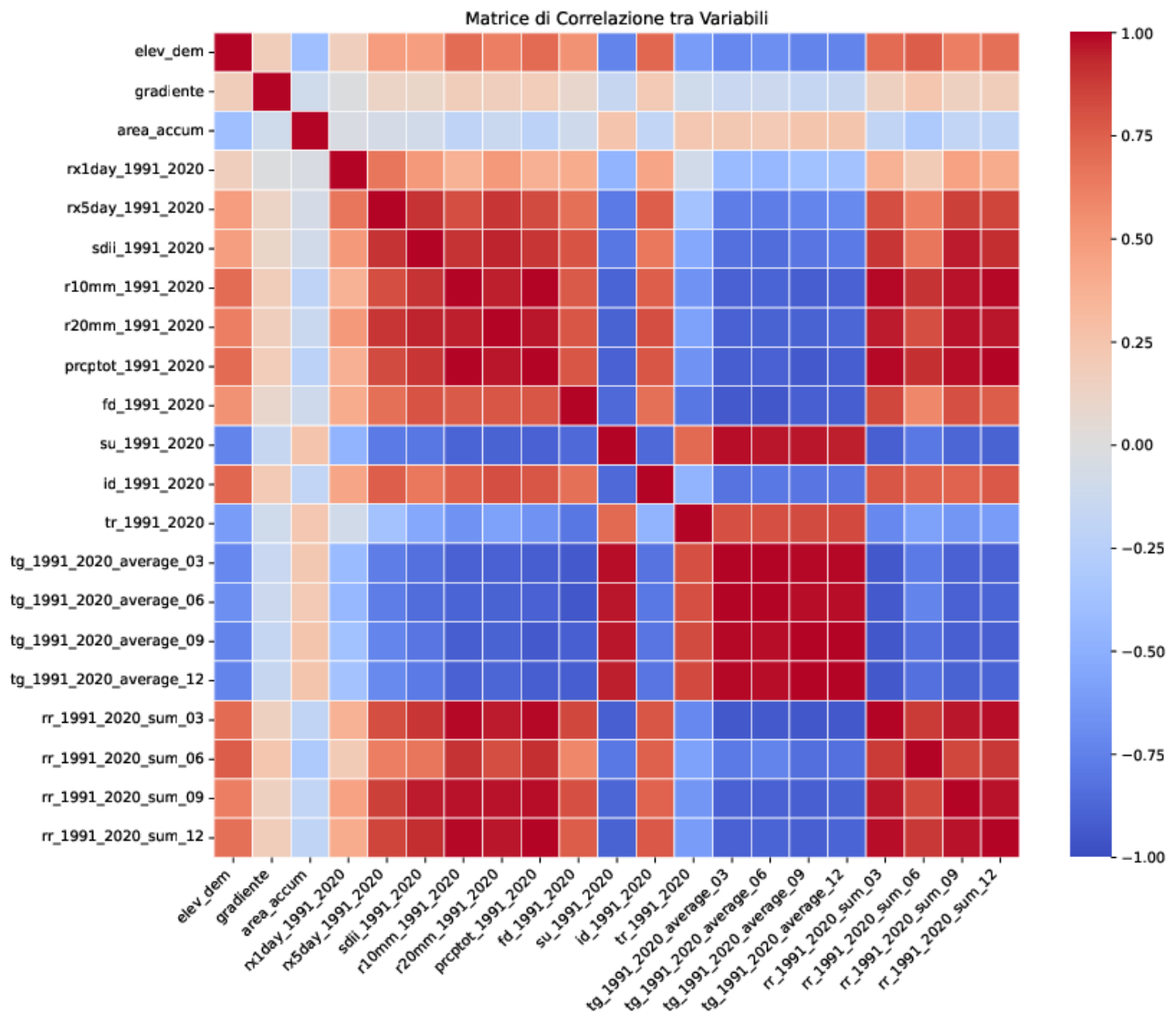


Fig. 13: Matrice di correlazione ottenuta per le variabili relative alla Serie4

L'analisi della matrice di correlazione consente di fare alcune considerazioni: Le variabili relative alle precipitazioni stagionali si correlano con molte altre variabili inerenti le precipitazioni. Tuttavia rispetto a queste ultime hanno il vantaggio di descrivere periodi dell'anno diversi e hanno quindi, complessivamente, un contenuto informativo più rilevante, ad esempio, rispetto alla precipitazione media annua. La variabile `sdii_1991_2020` è correlata solo con due medie stagionali e ha quindi un potenziale valore residuo nello spiegare l'abbondanza delle specie nelle altre due stagioni. Le variabili `rx5day_1991-2020` e `r20mm_1991_2020` sono correlata con `sdii_1991_2020` e possono quindi essere escluse.

Per le variabili relative alle temperature stagionali valgono le considerazioni fatte per quelle relative alle precipitazioni. Si può quindi escludere la variabile `su_1990_2020`.

Sebbene una correlazione negativa indichi che, mentre una variabile aumenta, l'altra tende a diminuire, entrambe possono ancora contenere informazioni utili per i modelli e quindi si decide di mantenerle.

Le variabili che vengono utilizzate per la modellazione sono quindi 17 e sono le seguenti:

`fd, id, rr_sum_03, rr_sum_06, rr_sum_09, rr_sum_12, rx1day, rx5day, sdii, tg_average_03, tg_average_06, tg_average_09, tg_average_12, tr, dem_carve, accum_kmq, stream_gradient`

in cui le ultime tre variabili rappresentano la quota, l'area del bacino drenante e la pendenza locale del fiume.

## 7.2 TRAINING DEI MODELLI

Per ciascuna delle 9 specie considerate è stato costruito un modello supervisionato basato sul un algoritmo di classificazione definito di Gradient Boosting.

Il **Gradient Boosting** è un potente metodo di apprendimento automatico che combina più modelli deboli (solitamente alberi di decisione) per creare un modello complessivo forte. Funziona in modo iterativo, addestrando ciascun nuovo albero per correggere gli errori commessi dai precedenti. Questo è ottenuto minimizzando una funzione di perdita (ad esempio, l'errore di previsione) tramite una tecnica chiamata "boosting" dei gradienti.

In sintesi, Gradient Boosting costruisce il modello passo dopo passo, cercando di migliorare le previsioni riducendo gli errori residui ad ogni iterazione.

Per l'applicazione al caso specifico di studio i modelli sono stati generati più volte per ciascuna specie considerata. Infatti, al fine di effettuare una cross validation, ciascun dataset è stato suddiviso in 10 "folds", in modo tale che sono stati costruiti 10 modelli validati su ognuna delle folds e allenati sulle rimanenti 9. Le statistiche di performance ottenute sono quindi da intendersi come medie tra quelle ottenute nei 10 run.

In tabella 3 sono mostrati gli indici che definiscono le performance dei modelli.

Particolarmente utile è l'indice "recall", che esprime, per ogni classe, il tasso di record correttamente classificati dal modello (true positive rate). In tabella viene riportata il recall per ogni classe e la media tra le classi. Si fa notare che per una classificazione a 4 classi un modello randomico (che classifica randomicamente quindi) restituisce un valore medio pari a 0.25. Un valore superiore indica quindi uno skill classificativo positivo del modello. Più il valore si avvicina a 1 maggiore è la performance.

Specie	Indice	Classi				avg
		0	1	2	3	
Barbus Tyberinus	precision	0.86	0.39	0.16	0.32	0.43
	recall	0.84	0.29	0.16	0.43	0.43
	f1-score	0.85	0.33	0.16	0.37	0.43
Cobitis Taenia	precision	0.50	0.45	0.73	0.59	0.57
	recall	0.85	0.56	0.50	0.22	0.53
	f1-score	0.63	0.50	0.59	0.32	0.51
Cottus Gobio	precision	0.58	0.72		0.94	0.75
	recall	0.97	0.67		0.38	0.67
	f1-score	0.72	0.69		0.54	0.65
Leuciscus Cephalus	precision	0.85	0.35	0.20	0.36	0.44
	recall	0.84	0.44	0.14	0.38	0.45
	f1-score	0.85	0.39	0.16	0.37	0.44
Leuciscus Lucumonis	precision	0.74	0.45	0.27	0.45	0.48
	recall	0.78	0.47	0.17	0.57	0.5
	f1-score	0.76	0.46	0.21	0.50	0.48
Leuciscus Souffia	precision	0.63	0.39	0.50	0.69	0.55
	recall	0.68	0.41	0.40	0.74	0.56
	f1-score	0.65	0.40	0.45	0.71	0.55
Padogobius Nigricans	precision	0.78	0.45	0.42	0.36	0.5
	recall	0.81	0.35	0.44	0.41	0.5
	f1-score	0.80	0.39	0.43	0.38	0.5
Rutilus Rubilio	precision	0.82	0.36	0.40	0.47	0.51
	recall	0.78	0.36	0.43	0.44	0.51
	f1-score	0.80	0.36	0.42	0.46	0.51
Salmo Trutta	precision	0.74	0.46	0.46	0.76	0.6
	recall	0.76	0.54	0.33	0.80	0.61
	f1-score	0.75	0.49	0.39	0.78	0.6

Tabella 3: indici di performance dei modelli

I risultati evidenziano che le migliori classificazioni complessive si ottengono per le specie Cottus Gobio e Salmo Trutta, anche se bisogna considerare che la prima specie risente del fatto che i dati di abbondanza sono molto sbilanciati, con moltissimi 0 e pochi dati nelle altre classi. I peggiori risultati si ottengono per Barbus Tyberinus e Leuciscus Cephalus.

Se si osservano i valori di recal per le singole classi di abbondanza di ogni specie si osserva una notevole variabilità con valori che oscillano tra 0.68 e 0.97 per la classe 0, 0.29 e 0.67 per la classe 1, 0.14 e 0.50 per la classe 2 e, infine, 0.22 e 0.80 per la classe 3. Si può in generale dire che i modelli sembrano essere più capaci di classificare le aree in cui le diverse specie non sono presenti rispetto a quelle in cui lo sono. Tra queste ultime la classe 3 sembra essere più facilmente riconosciuta, rispetto alla classe 1 e infine alla classe 2.

L'analisi della importanza delle variabili nel definire i diversi modelli supervisionati è stata effettuata facendo un rank dell'importanza nei diversi modelli (1 per la più importante e 17 per la meno importante) e sommando tutte le loro posizioni fino a ottenere uno score complessivo. Minore è lo score maggiore è la rilevanza della variabile. La tabella 4 mostra i risultati ottenuti.

<b>Variabile</b>	<b>Score</b>
dem_carve	42
rr_sum_03	49
accum_kmq	55
fd	66
sdi	68
tg_average_09	75
rr_sum_09	77
stream_gradient	84
rr_sum_06	84
rr_sum_12	86
tg_average_12	86
rx5day	92
tr	95
tg_average_03	97
tg_average_06	102
rx1day	106
id	113

La tabella 3 mostra che tra le variabili più rilevanti nei modelli ci sono le quote (dem\_carve), le precipitazioni medie nei mesi primaverili (rr\_sum\_03), la dimensione del bacino a monte del punto di interesse (accum\_kmq), il numero di giorni in cui la temperatura minima giornaliera è inferiore a 0°C (fd), l'intensità media dei giorni piovosi (sdii). Viceversa hanno meno importanza il numero di giorni in cui la temperatura massima giornaliera è inferiore a 0°C (id), la precipitazione massima giornaliera (rx1day), le temperature medie in primavera e estate (tg\_average\_03, tg\_average\_06) e le precipitazioni in inverno (rr\_sum\_12).

### 7.3 APPLICAZIONE DEI MODELLI

I modelli allenati sui dati campionati nella Serie4 sono stati successivamente applicati per predire gli scenari delle classi di abbondanza durante lo scenario climatico del campionamento (1991-2020), in un futuro a breve termine (2031-2065) e in uno scenario a lungo termine (2066-2100).

Nelle mappe che seguono (Fig. 14-22) sono mostrati i risultati ottenuti per le diverse specie.

Si osserva che i modelli sono adeguatamente in grado di predire l'abbondanza delle specie nel periodo del campionamento (confronto tra mappa "Campioni" e mappa "1991-2020") e che le predizioni delle classi di abbondanza a breve e lungo termine sono molto differenti in funzione delle specie, prefigurando, in alcuni casi, una quasi totale assenza o forte riduzione futura sul territorio regionale e in altri casi un'estensione delle aree di idoneità ambientale delle specie.

Analogamente a quanto fatto con la serie dei dati Serie4, sono stati utilizzati i dati presenti nella Serie3 per effettuare una analisi della possibile abbondanza di specie in termini di densità di esemplari.

In questo caso tuttavia i modelli hanno assegnato valori reali (con la virgola) a ciascun pixel del reticolo. Per consentire una visualizzazione del risultato adeguata è stato necessario procedere ad una riclassificazione e si è scelto di utilizzare una rappresentazione per quantili. Più in particolare sono stati considerati i quantili 0.25, 0.50, 0.75 e 0.95 (l'ultimo con lo scopo di visualizzare i valori all'estremo superiore). I quantili sono stati calcolati sulla distribuzione dei valori presenti nelle colonne della tabella della serie3.

Le figure dalla 23 alla 31 mostrano i risultati ottenuti. Nella consultazione delle mappe si deve porre attenzione alle scale che, poiché calcolate di volta in volta per ogni specie, possono risultare estremamente diverse.

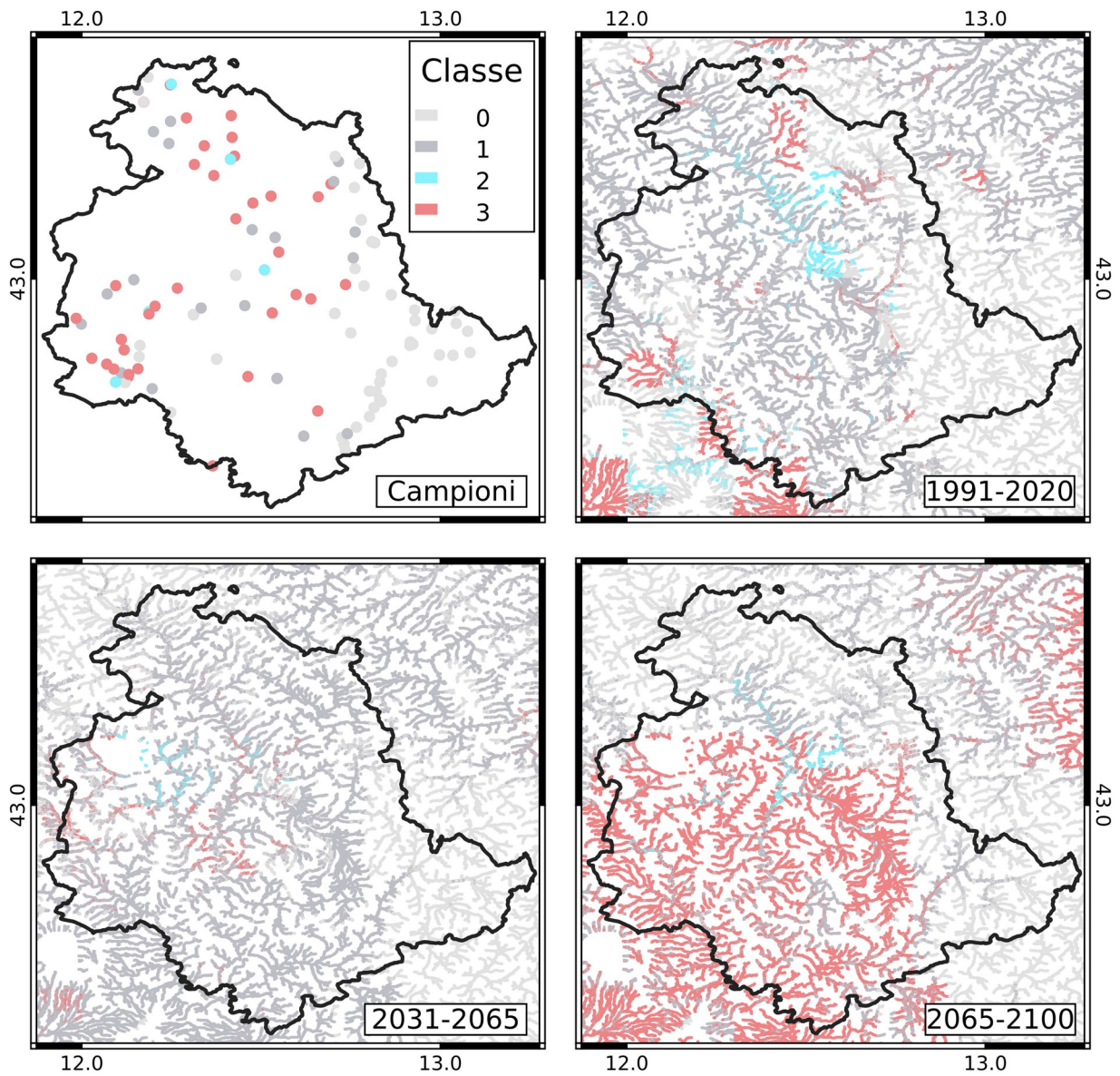


Figura 14: Dati di campionamento e scenari di abbondanza del *Barbus Tyberinus*

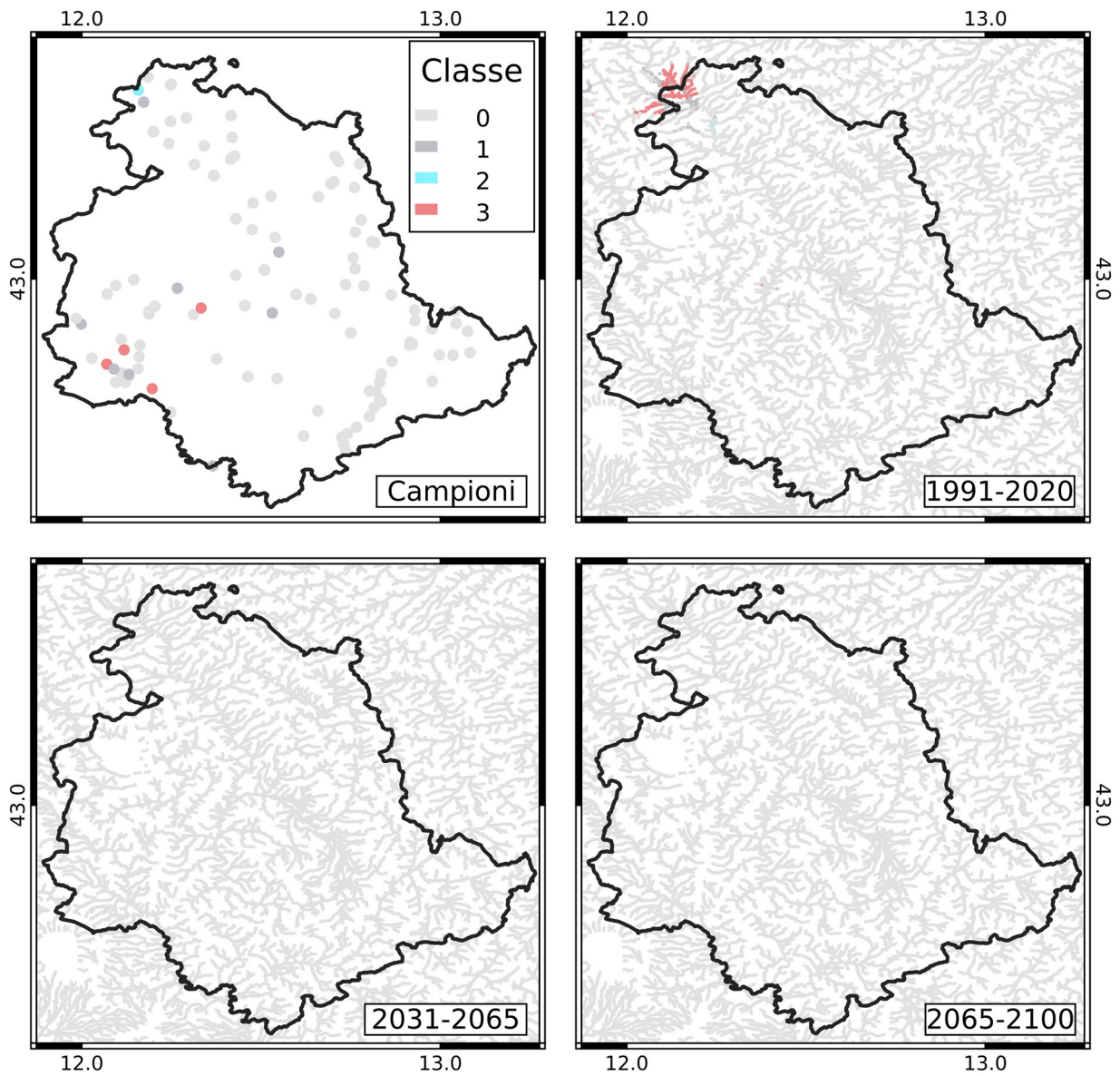


Figura 15: Dati di campionamento e scenari di abbondanza del *Cobitis Taenia*

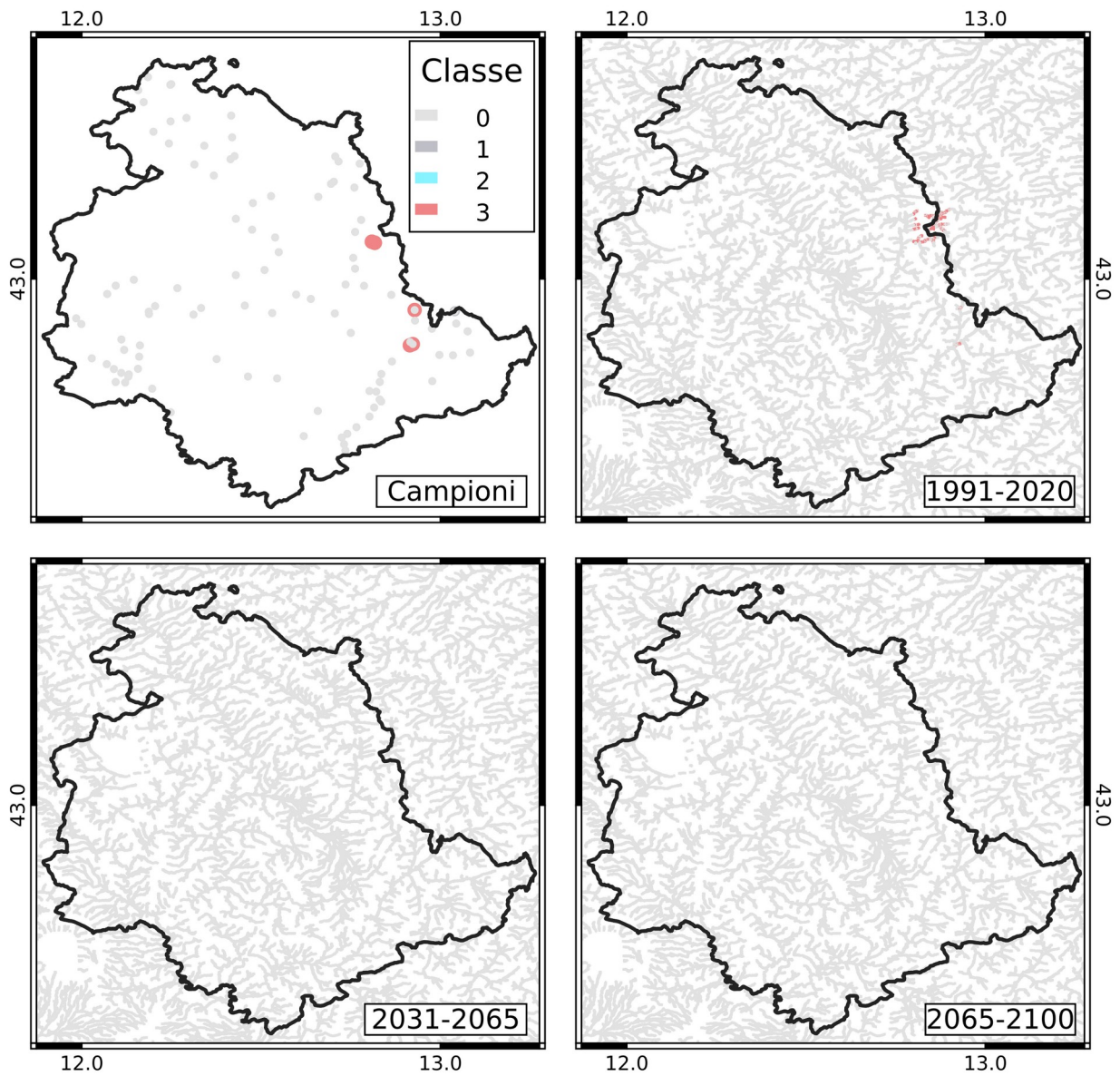


Figura 16: Dati di campionamento e scenari di abbondanza del *Cottus Gobio*

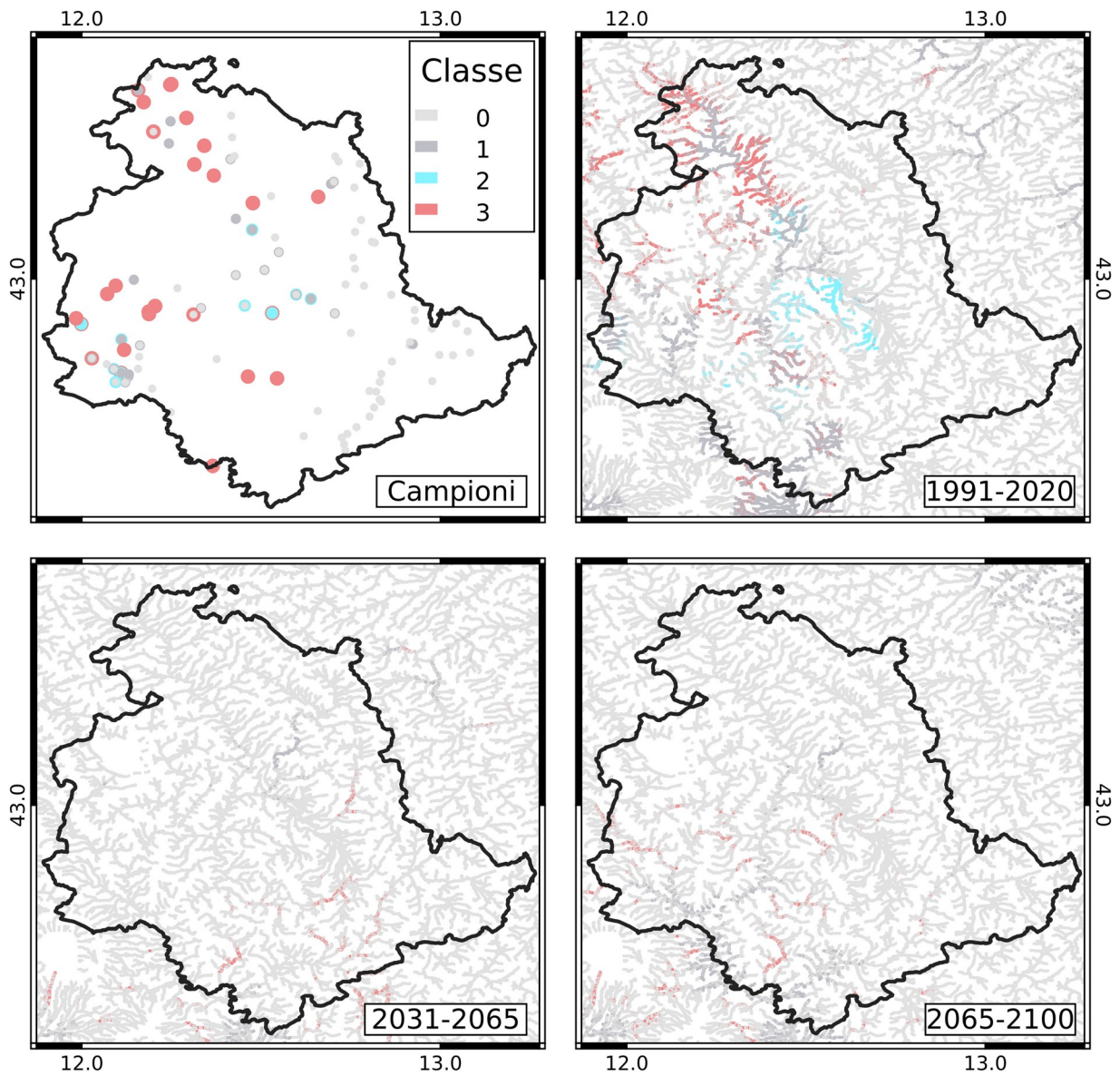


Figura 17: Dati di campionamento e scenari di abbondanza del leuciscus *Lucumonis*

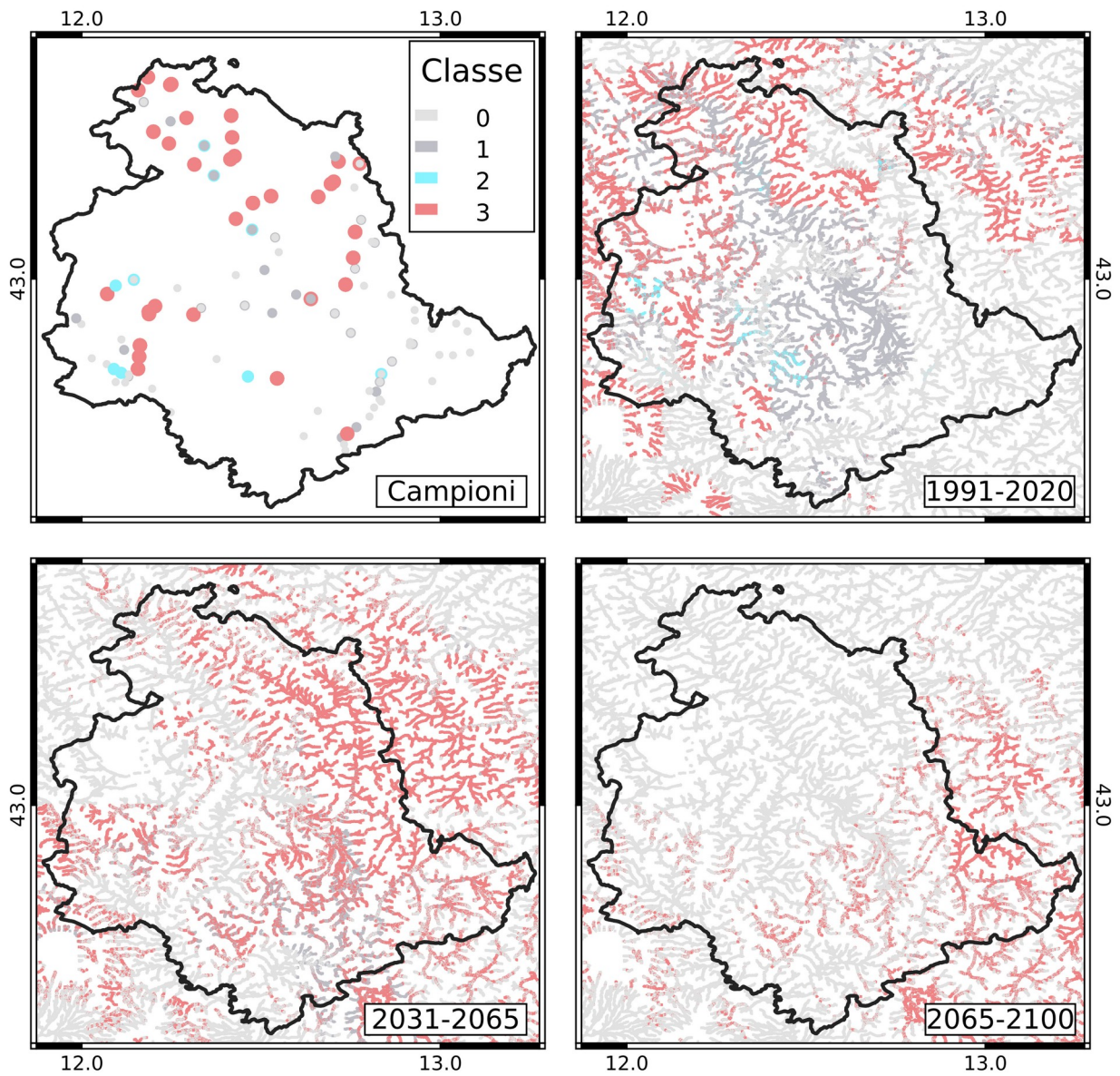


Figura 18: Dati di campionamento e scenari di abbondanza del leuciscus Souffia

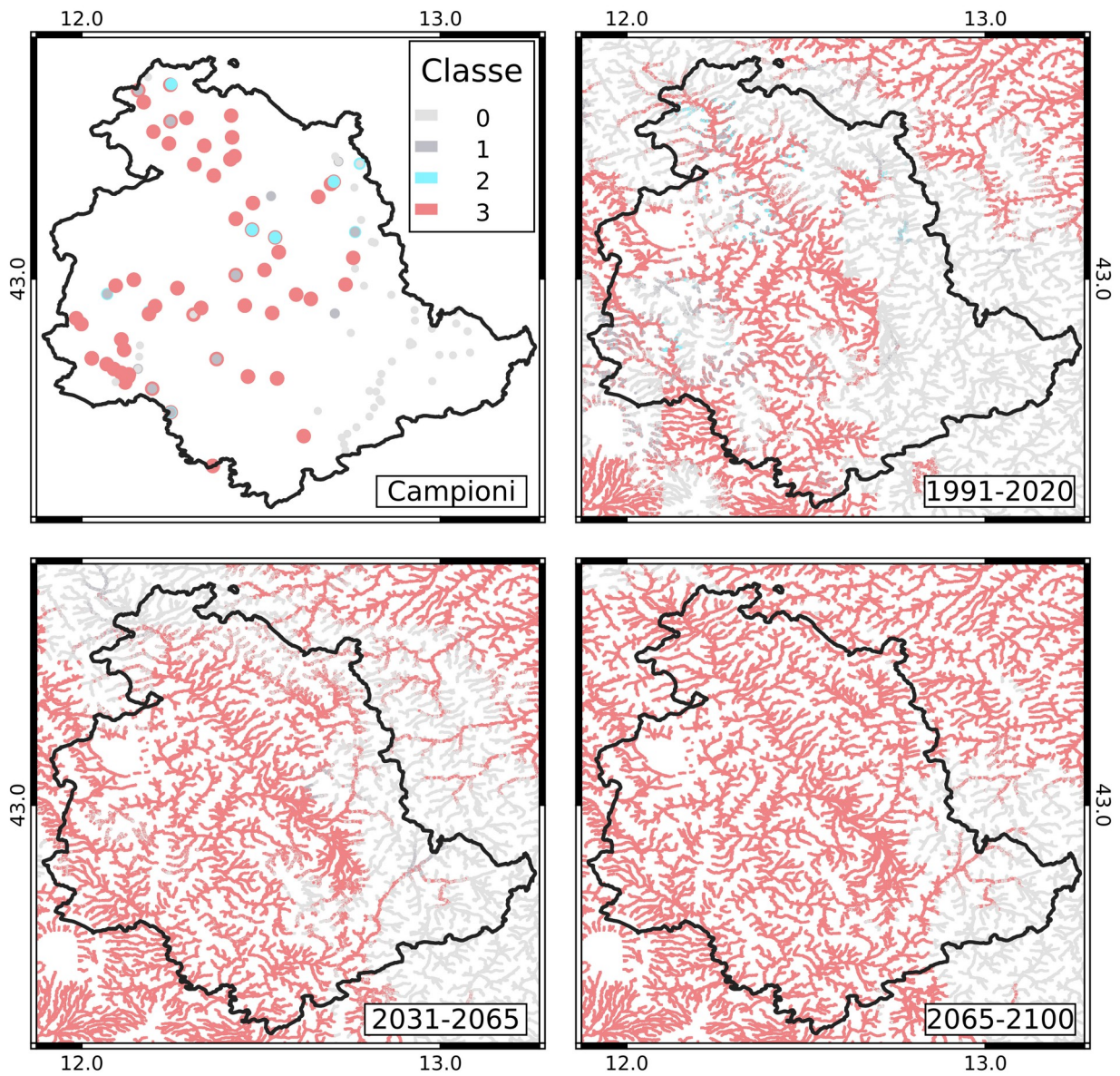


Figura 19: Dati di campionamento e scenari di abbondanza del leuciscus Cephalus

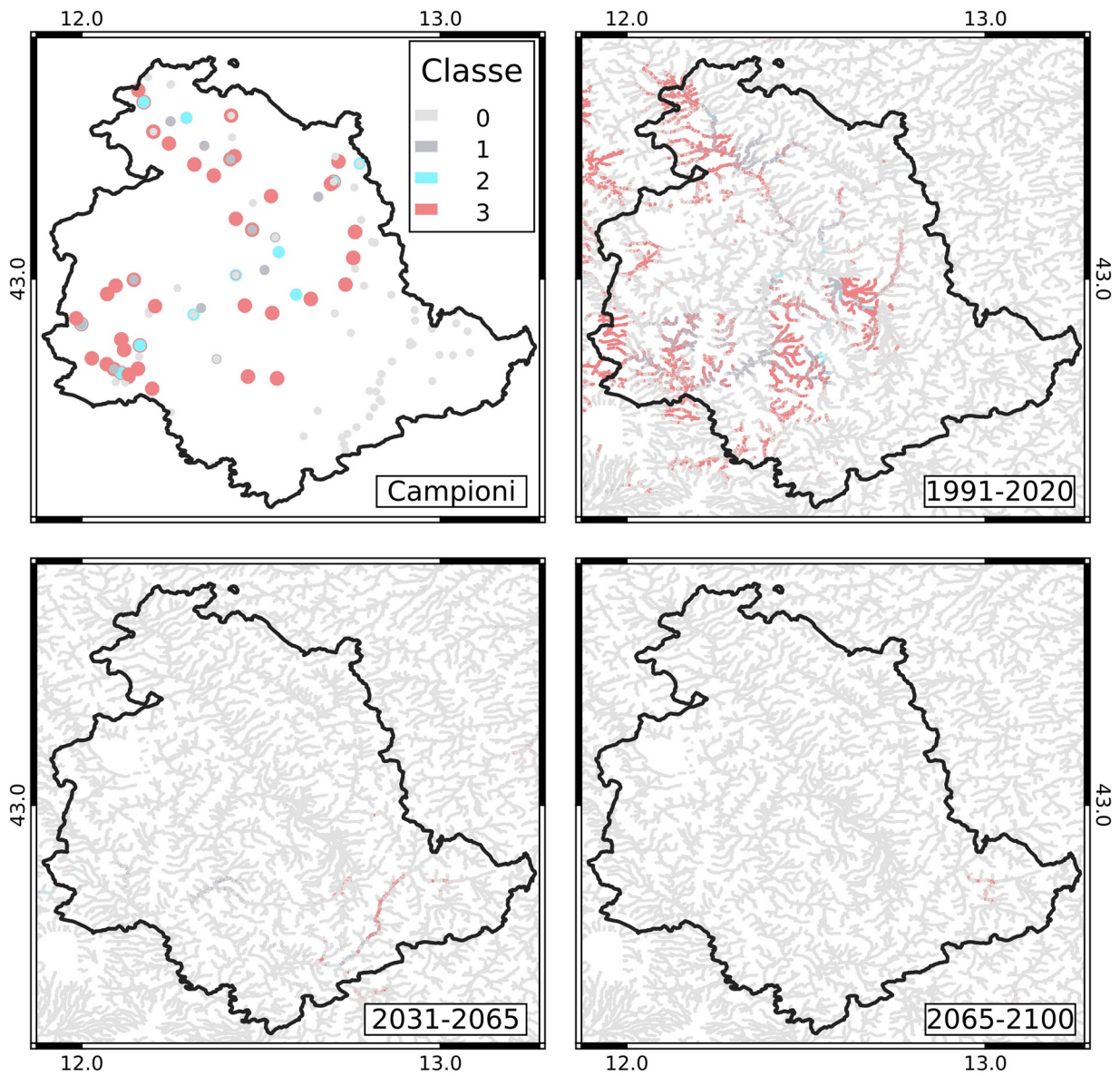


Figura 20: Dati di campionamento e scenari di abbondanza del *Padogobius Nigricans*

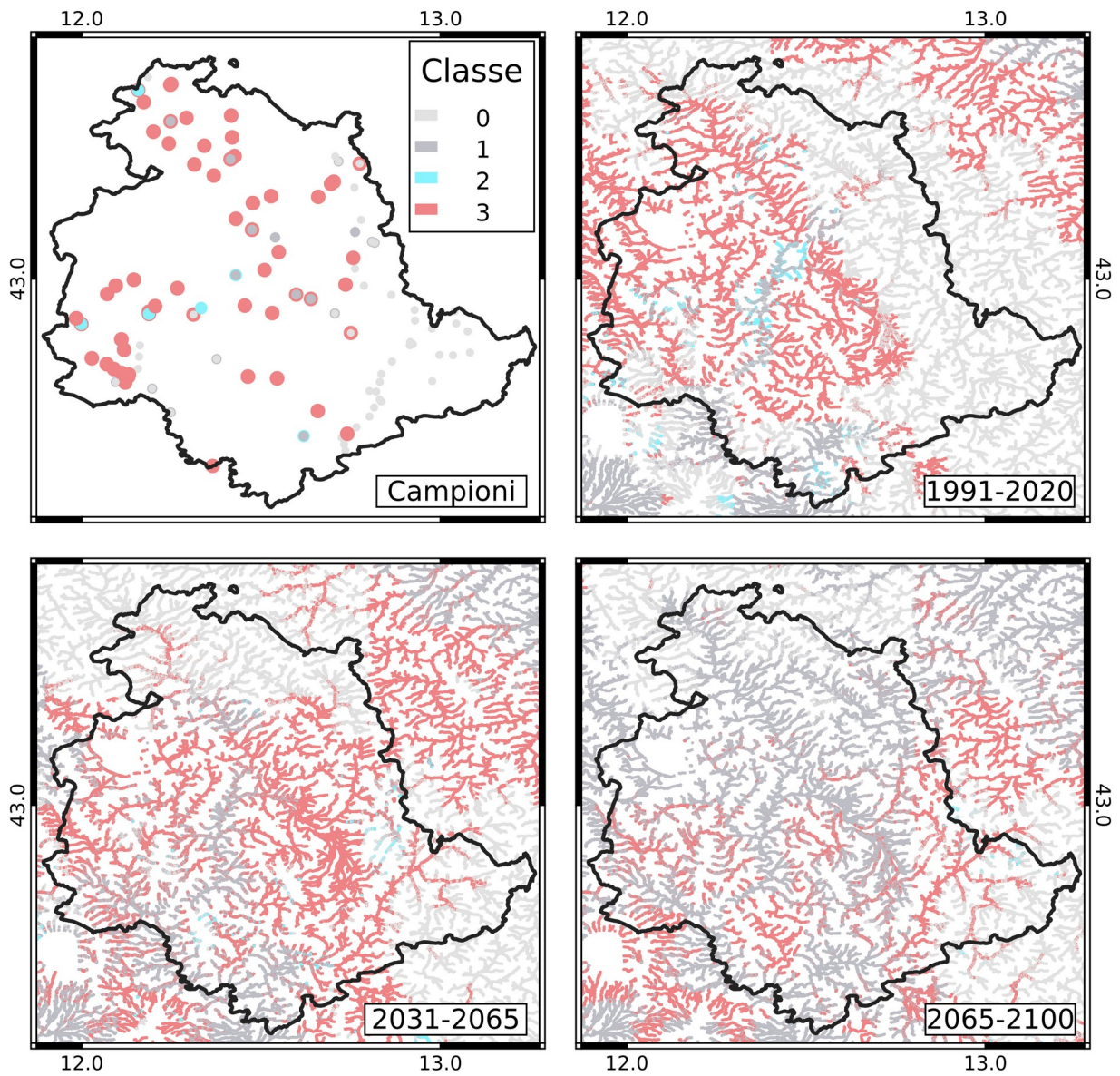


Figura 21: Dati di campionamento e scenari di abbondanza del *Rutilus Rubilio*

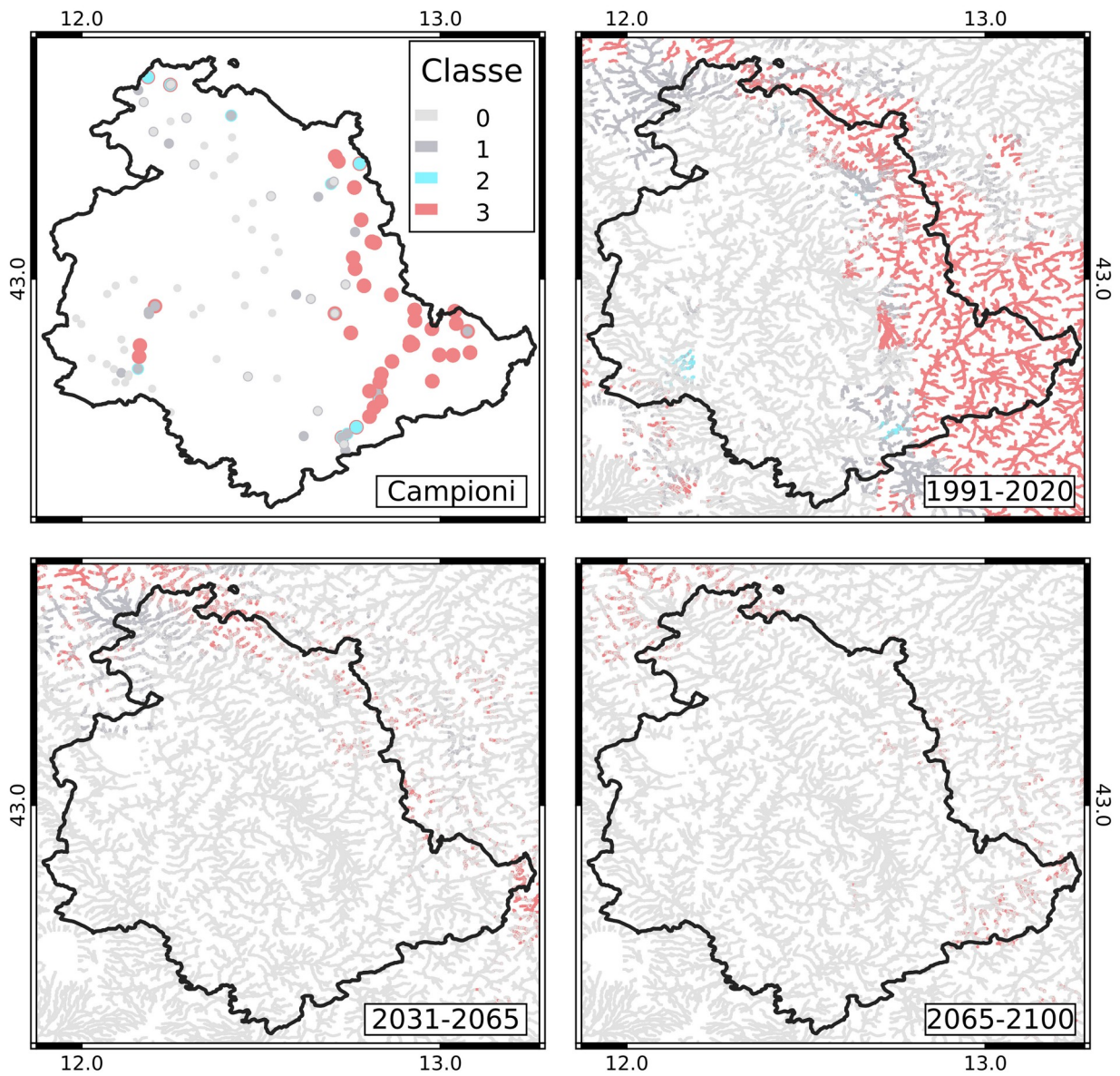


Figura 22: Dati di campionamento e scenari di abbondanza del *Salmo Trutta*

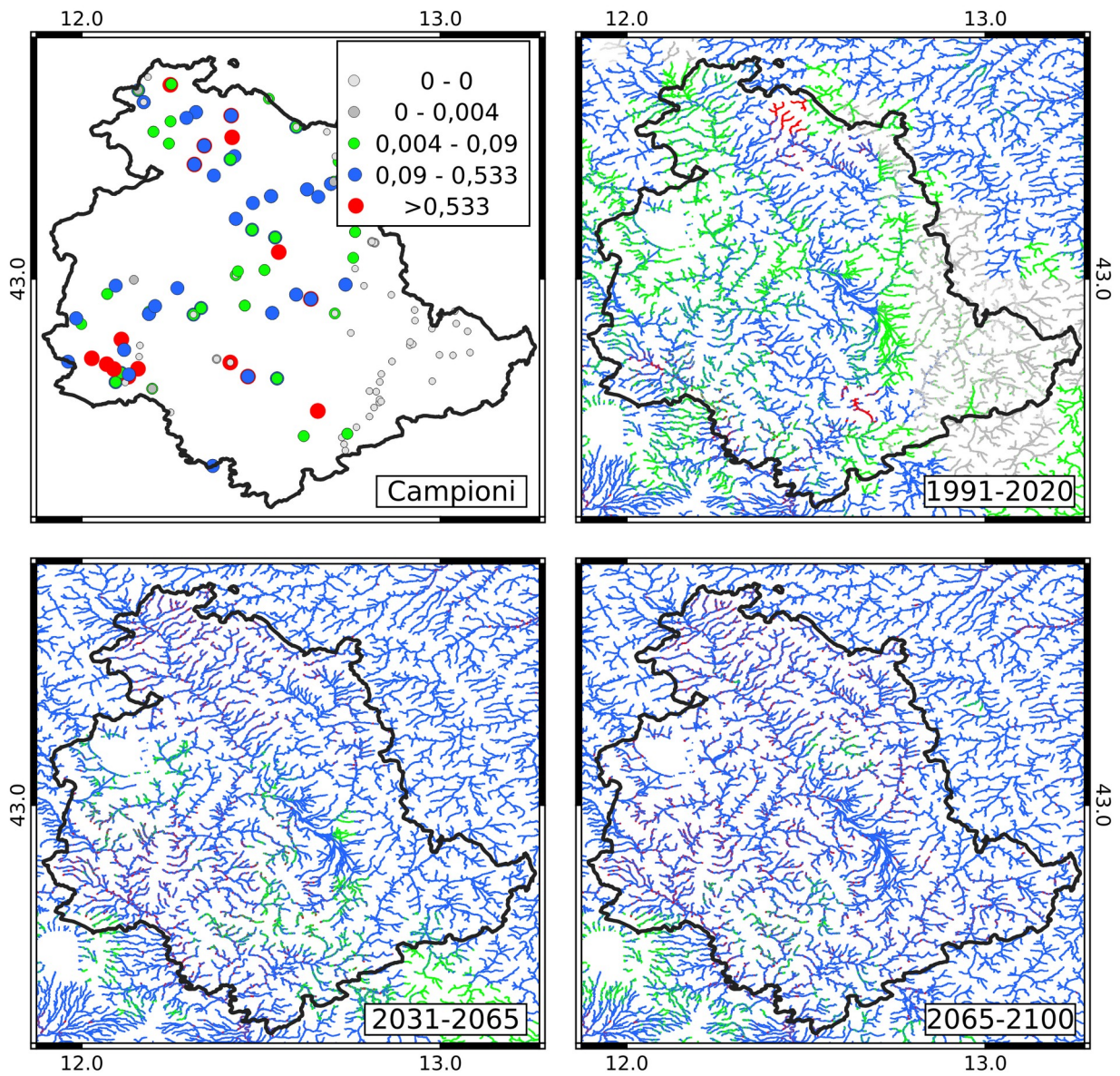


Figura 23: Dati di campionamento e scenari di abbondanza del *Barbus Tyberinus* modellati a partire dalla serie3

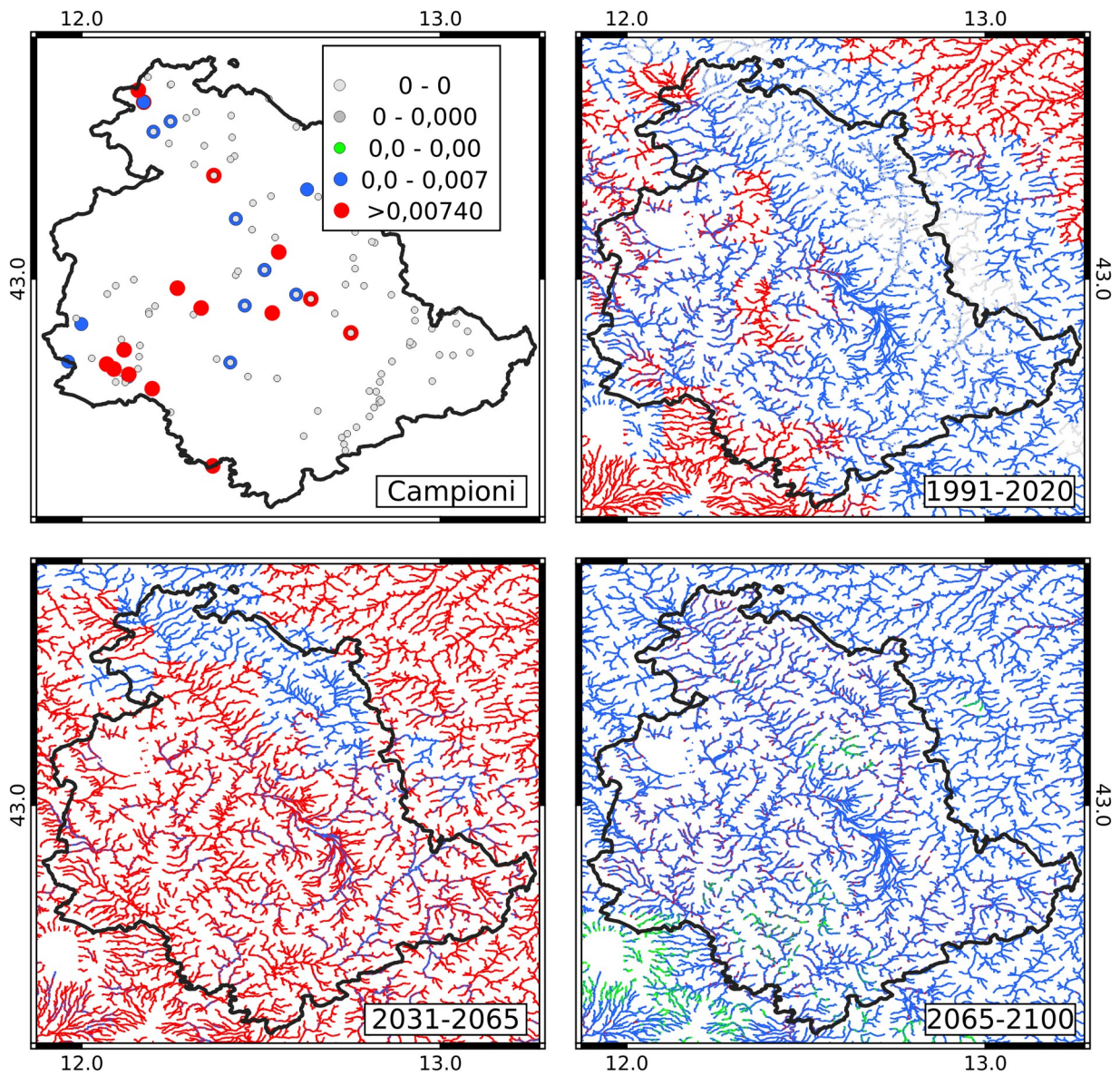


Figura 24: Dati di campionamento e scenari di abbondanza del *Cobitis Taenia* modellati a partire dalla serie3

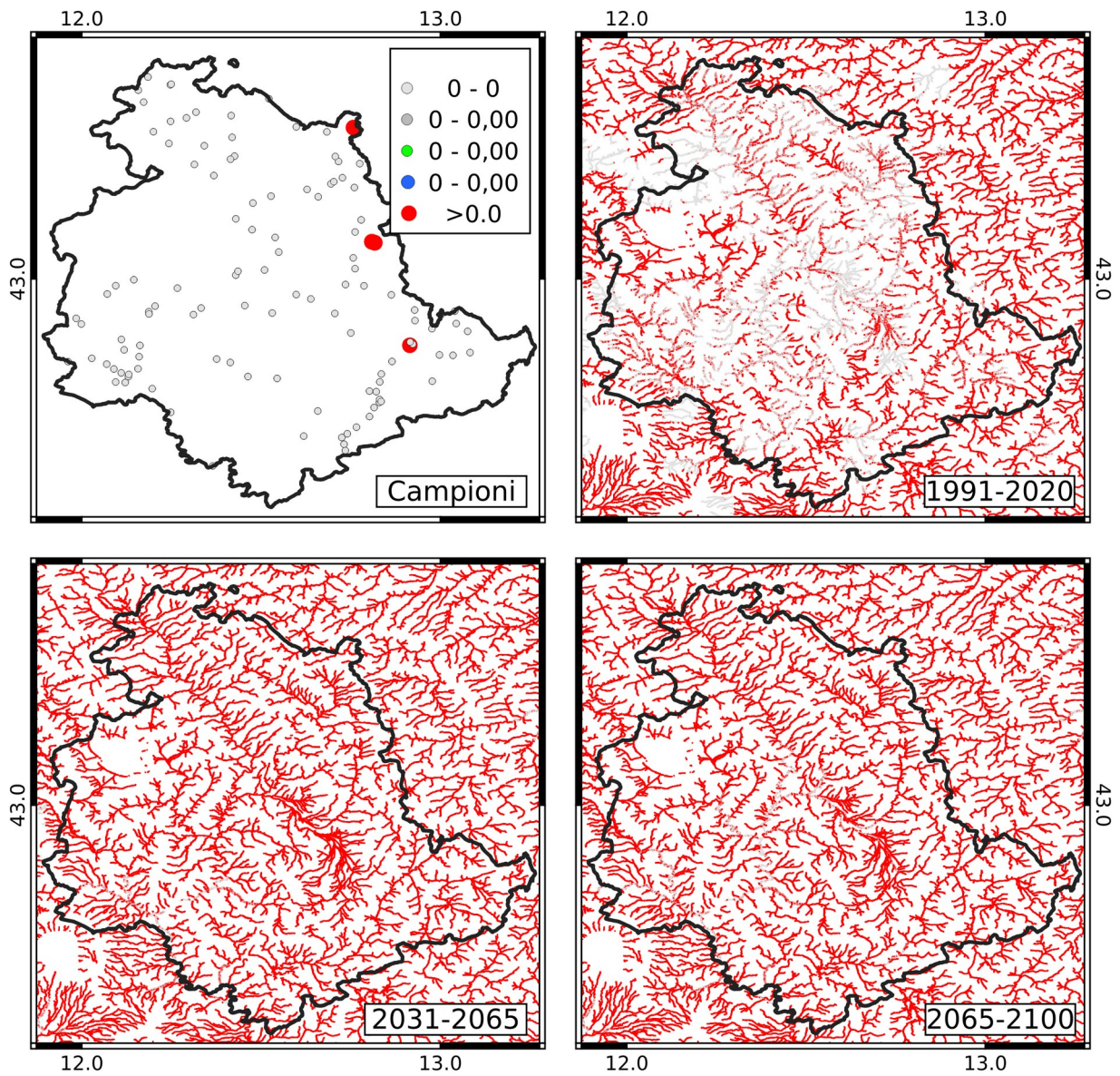


Figura 25: Dati di campionamento e scenari di abbondanza del Cottus Gobio modellati a partire dalla serie3

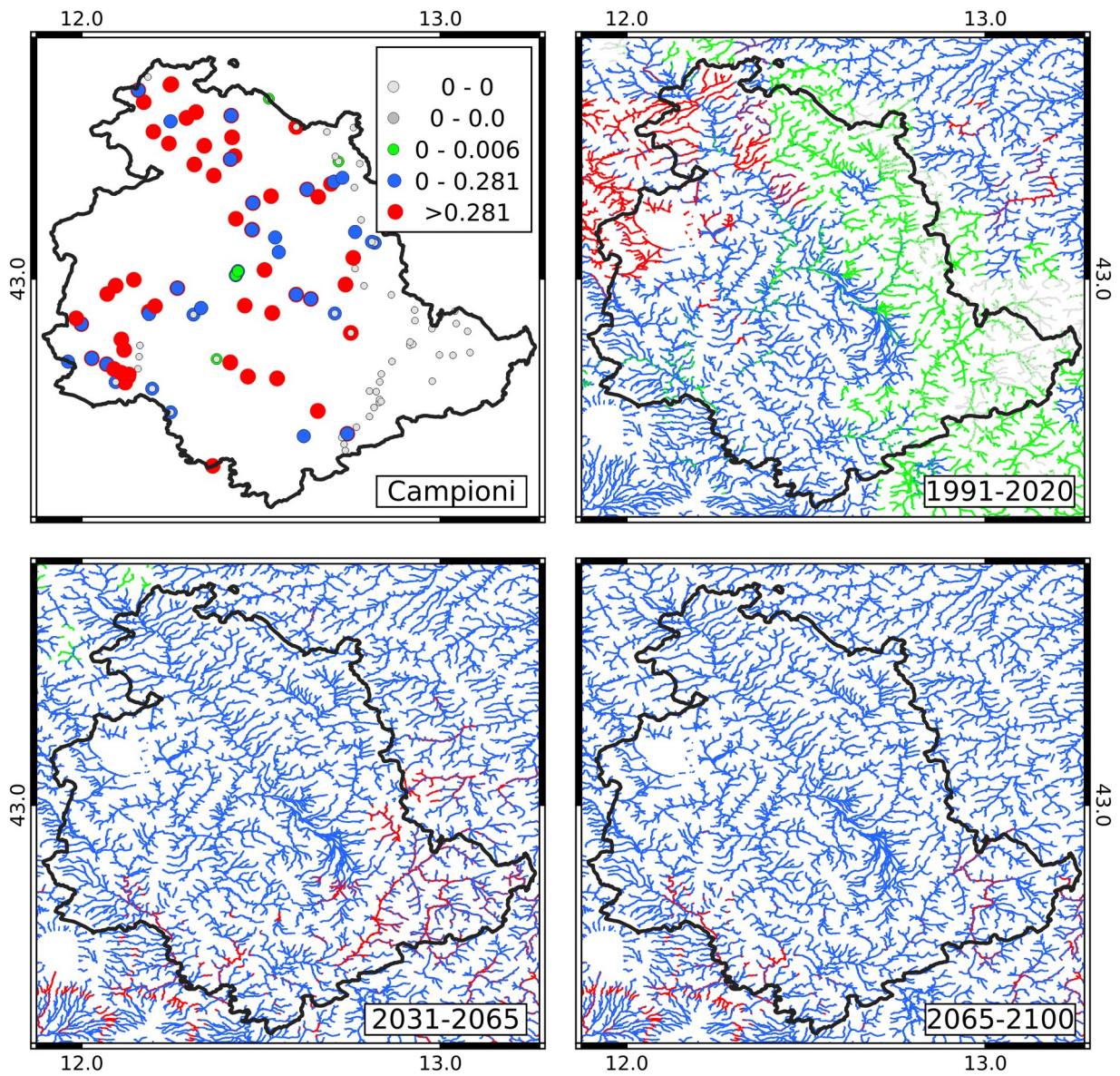


Figura 26: Dati di campionamento e scenari di abbondanza del *Leuciscus Lucumonis* modellati a partire dalla serie3

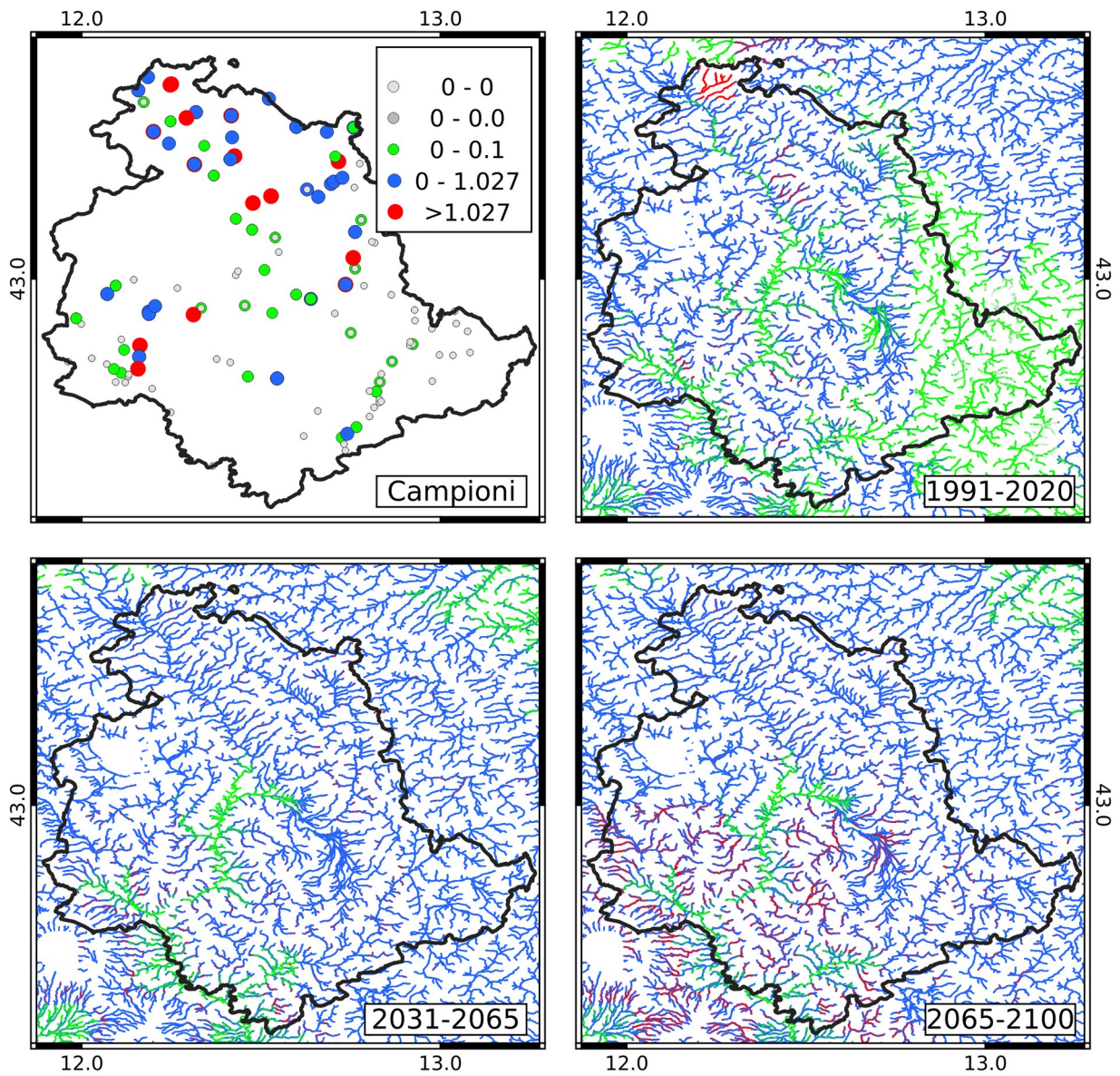


Figura 27: Dati di campionamento e scenari di abbondanza del *Leuciscus Souffia* modellati a partire dalla serie3

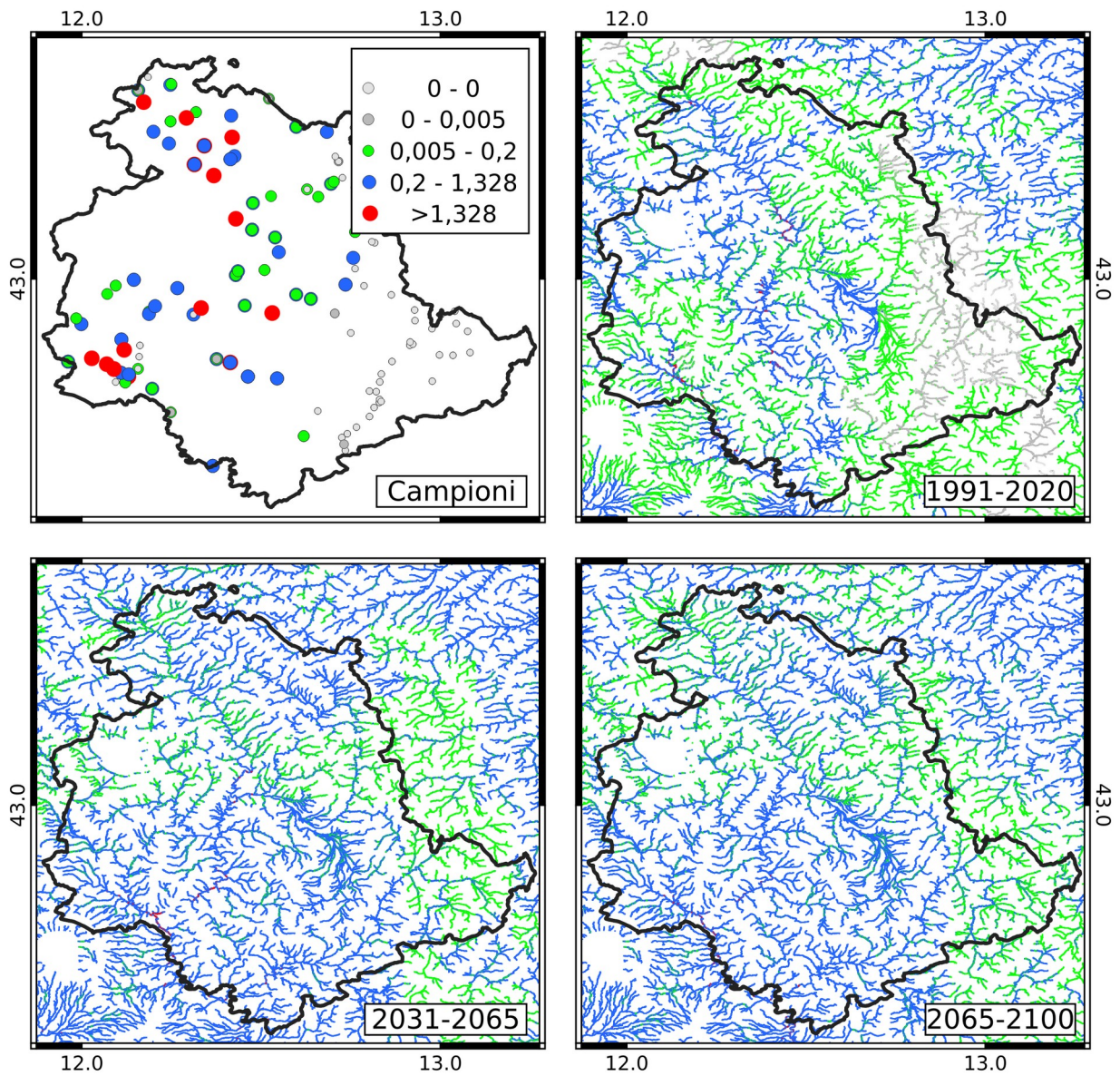


Figura 28: Dati di campionamento e scenari di abbondanza del *Leuciscus Cephalus* modellati a partire dalla serie3

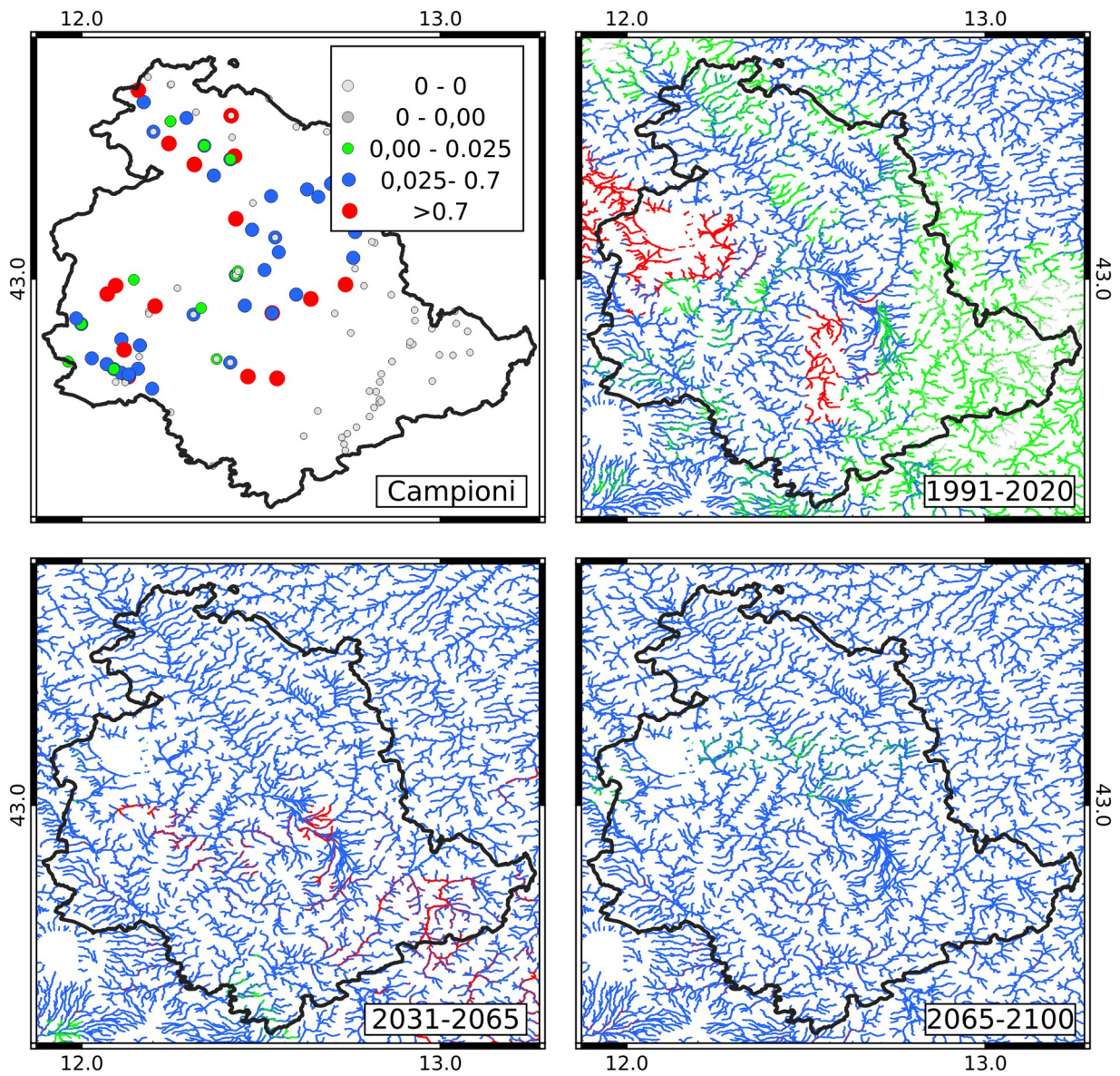


Figura 29: Dati di campionamento e scenari di abbondanza del *Padogobius Nigricans* modellati a partire dalla serie3

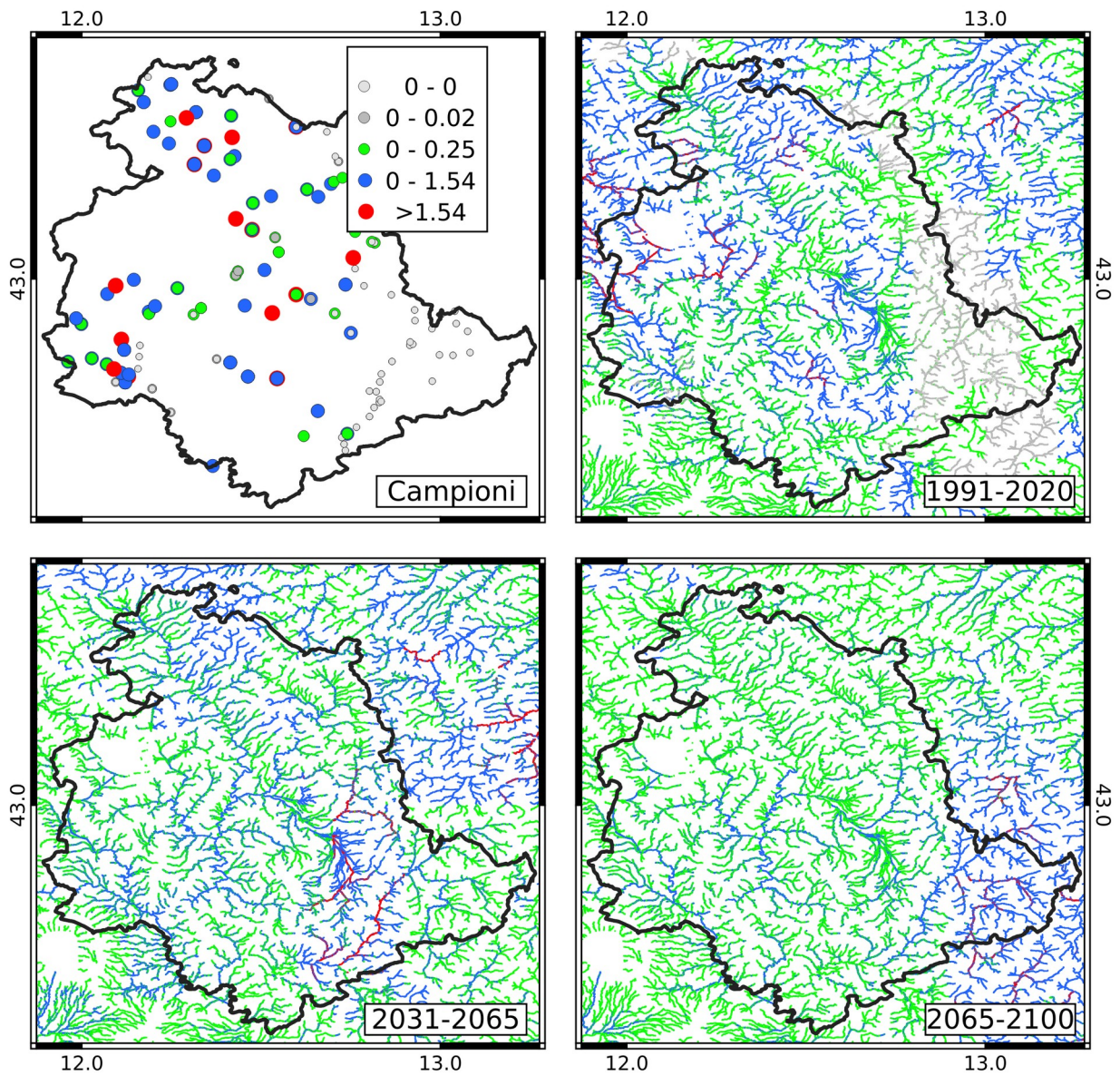


Figura 30: Dati di campionamento e scenari di abbondanza del *Rutilus Rubilio* modellati a partire dalla serie3

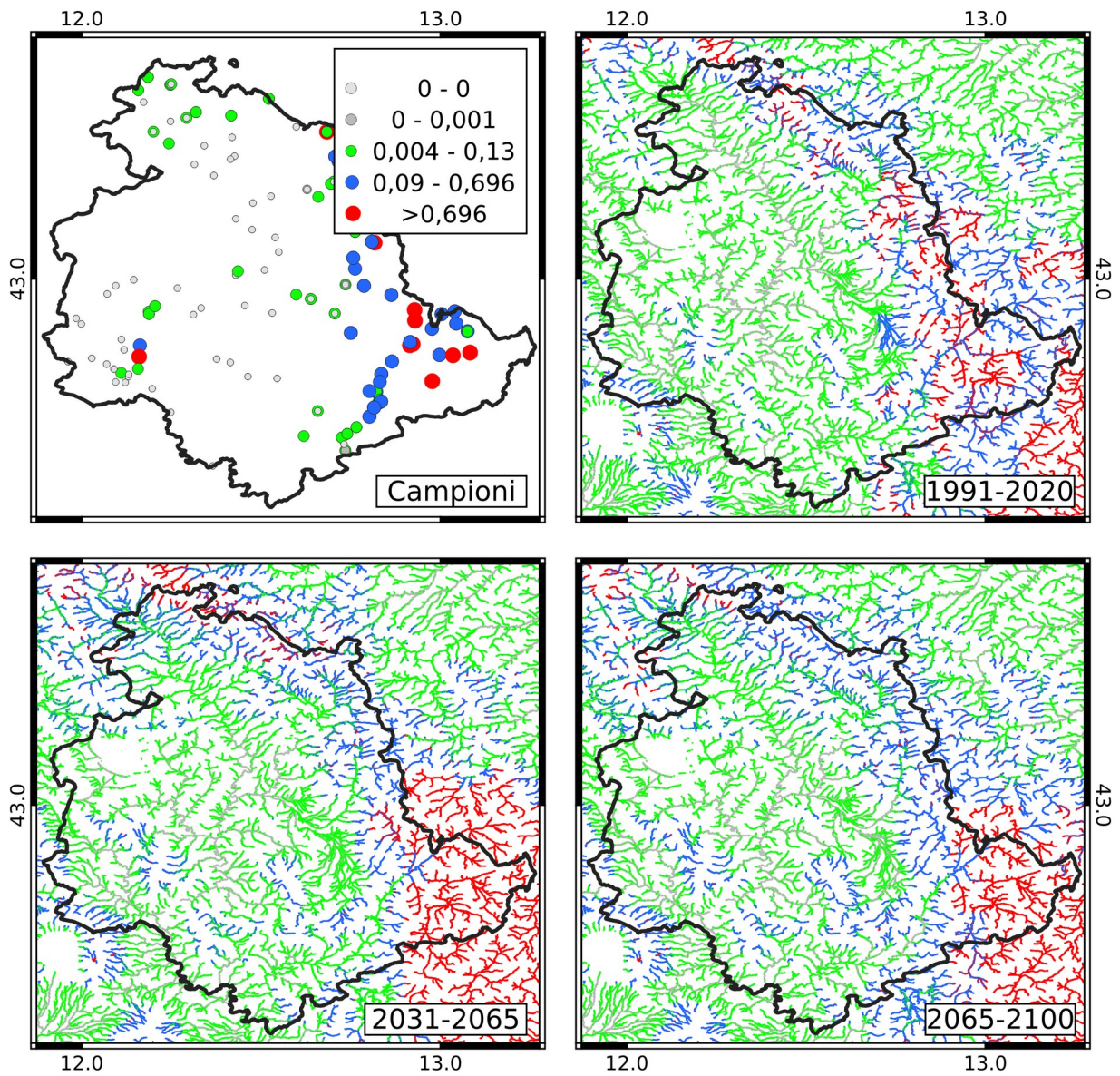


Figura 31: Dati di campionamento e scenari di abbondanza del Salmo Trutta modellati a partire dalla serie3

## 8 REFERENCES

- Cornes, R. C., van der Schrier, G., van den Besselaar, E. J. M., & Jones, P. D. (2018). An Ensemble Version of the E-OBS Temperature and Precipitation Data Sets. *Journal of Geophysical Research: Atmospheres*, *123*(17), 9391–9409. <https://doi.org/10.1029/2017JD028200>
- Dosio, A., & Paruolo, P. (2011). Bias correction of the ENSEMBLES high-resolution climate change projections for use by impact models: Evaluation on the present climate. *Journal of Geophysical Research: Atmospheres*, *116*(D16). <https://doi.org/10.1029/2011JD015934>
- Dosio, A., Paruolo, P., & Rojas, R. (2012). Bias correction of the ENSEMBLES high resolution climate change projections for use by impact models: Analysis of the climate change signal. *Journal of Geophysical Research: Atmospheres*, *117*(D17). <https://doi.org/10.1029/2012JD017968>
- Tarquini, S., Isola, I., Favalli, M., Mazzarini, F., Bisson, M., Pareschi, M. T., & Boschi, E. (2007). TINITALY/01: A new triangular irregular network of Italy. *Annals of Geophysics*.
- Tarquini, S., & Nannipieri, L. (2017). The 10m-resolution TINITALY DEM as a trans-disciplinary basis for the analysis of the Italian territory: Current trends and new perspectives. *Geomorphology*, *281*, 108–115. <https://doi.org/10.1016/j.geomorph.2016.12.022>