



GIORNATA ROMANA DI ORNITOLOGIA

dedicata a Gaspare Guerrieri

Metodi, strumenti e approcci nell'ornitologia di base e applicata

**Roma, Università degli studi Roma Tre
Viale Marconi, 446**

Sabato 24 Novembre 2012

***L'utilizzo di dati di presenza per la costruzione
di modelli di idoneità ambientale per le specie
di interesse***

a cura di

Dario Capizzi, Stefano Sarrocco, Marco Scalisi

ARP Biodiversità, Reti Ecologiche e Geodiversità

Obiettivi dello studio



- ✓ Utilizzare i dati di presenza registrati nell'*Atlante regionale degli uccelli nidificanti* per sviluppare modelli di idoneità ambientale per le specie di interesse (ma anche *comuni*)
- ✓ Selezionare un modello adeguato alla trattazione dei dati, sia faunistici sia ambientali
- ✓ Definire dei modelli predittivi sulla distribuzione delle specie
- ✓ Stimare l'estensione delle aree di distribuzione (*range*) e le consistenze delle popolazioni regionali



I dati di base:



Data records 56.876 (segnalazioni)
Data samples 3.470 (stazioni)

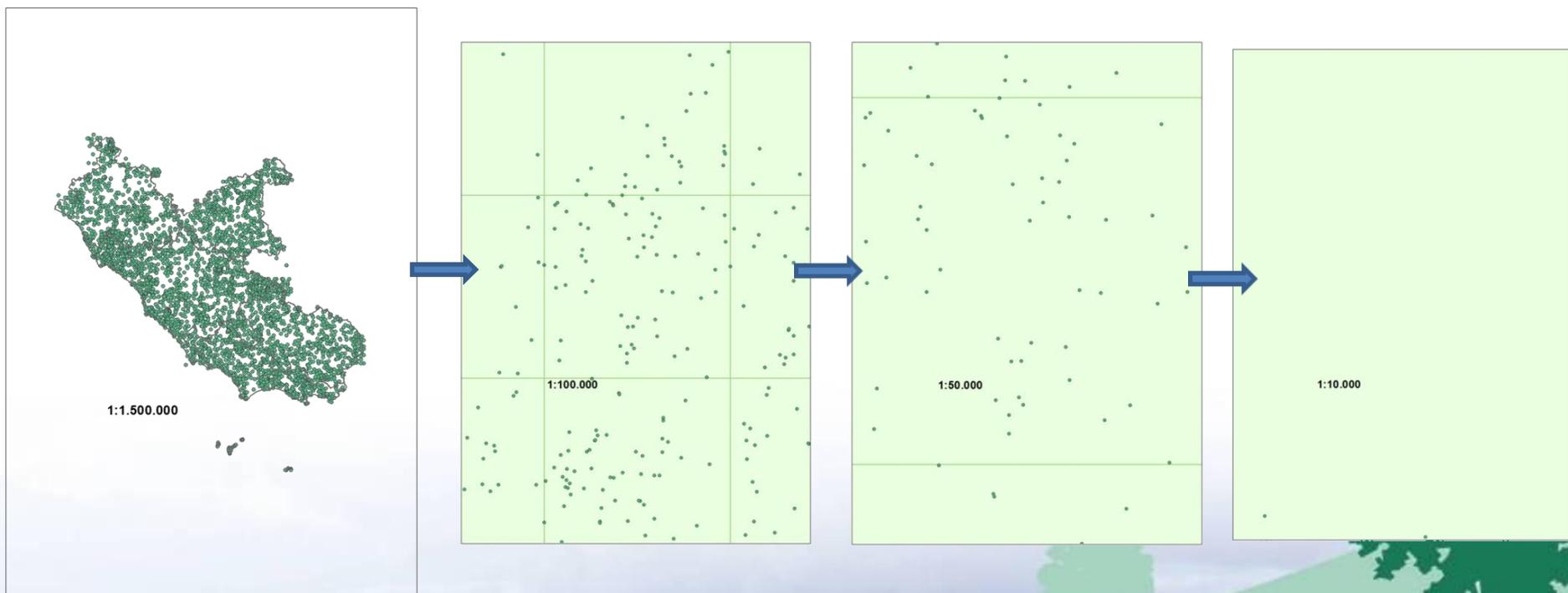
ARP
Agenzia Regionale Parchi

REGIONE LAZIO
ASSICURATO ALL'AMBIENTE E
PIU'NOI ECOLOGICAMENTE

NUOVO ATLANTE DEGLI UCCELLI NIDIFICANTI NEL LAZIO

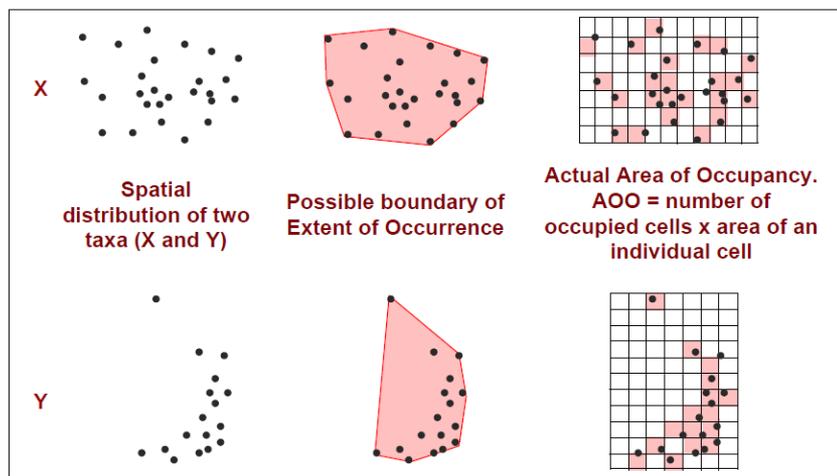
a cura di
Massimo Brunelli, Stefano Sarrocco, Ferdinando Corbi, Alberto Sorace, Aldo Boano,
Stefano De Felici, Gaspare Guerrieri, Angelo Meschini e Silvano Roma

I limiti della distribuzione delle specie per
stazioni di presenza: *pochi punti conosciuti
in una matrice estesa sconosciuta*



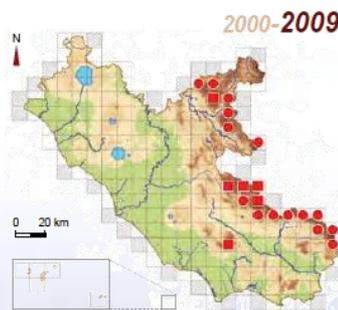
... seguendo gli standard per la rappresentazione della distribuzione delle specie

Figure 1.5. Two examples of the distinction between extent of occurrence and area of occupancy.



... e su superfici continue attraverso modelli predittivi della distribuzione

Aree ad alta e media
idoneità per la Balia dal
collare sull'Appennino
centrale (Monti Reatini)



1:10.000

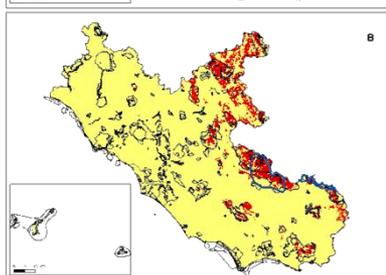
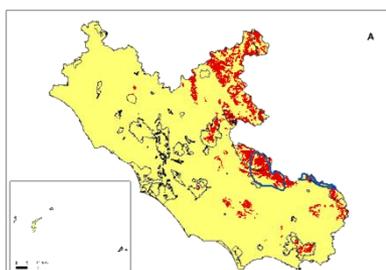
1:10.000



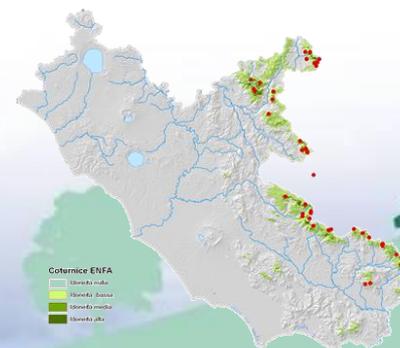
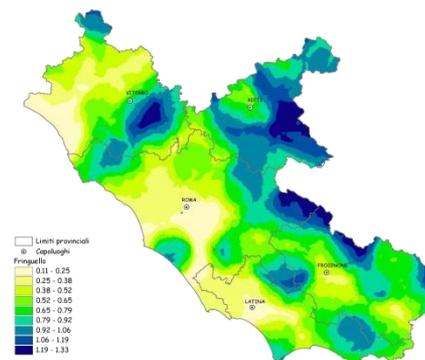
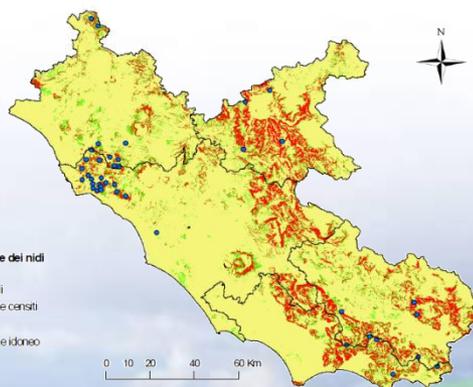
Numerosi modelli: deduttivi (deterministici) e induttivi

✓ con dati di presenza: ENFA, Maxent, DOMAIN, BIOCLIM;

✓ altri con presenza/assenza o abbondanze: GLMs (Generalized Linear Models) e suoi casi specifici (GAMs, Regressione Logistica, etc.), interpolazione con kriging (regression kriging)



Modello di distribuzione dei nidi di biancone nel Lazio



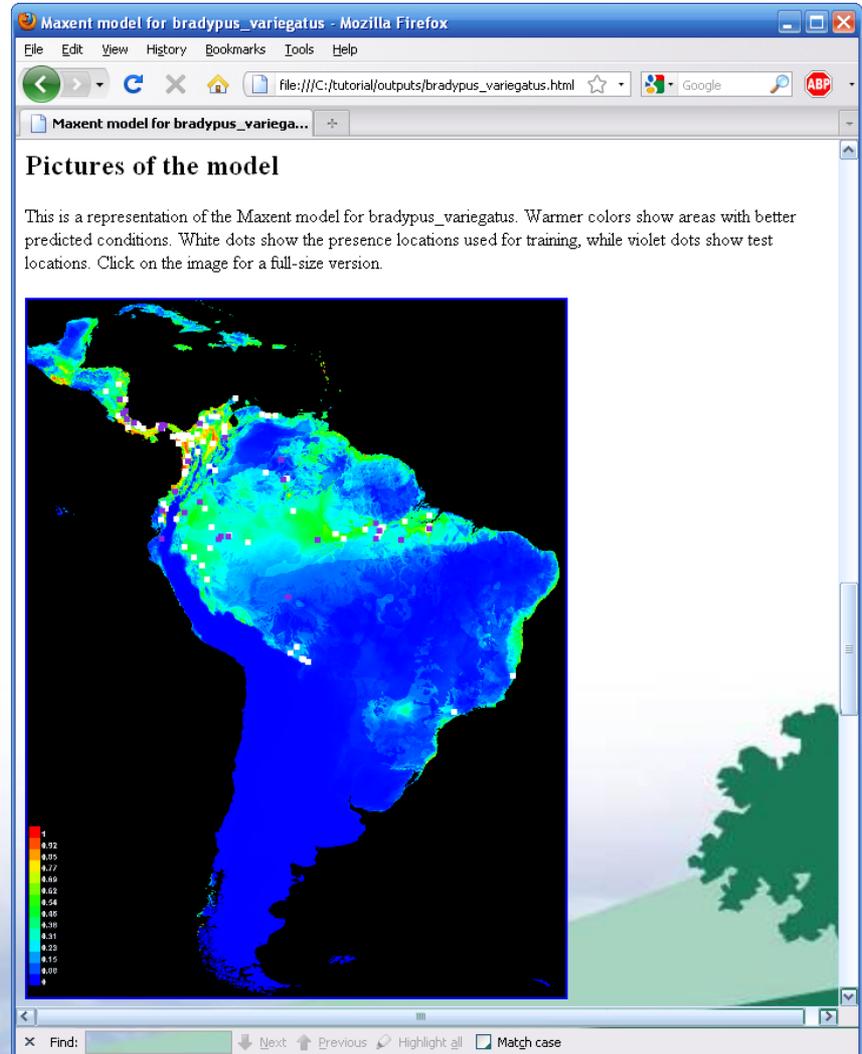
Maxent è un software *open source* utilizzato per svolgere predizioni e inferenze a partire da dati di presenza anche incompleti.

Stima la distribuzione delle specie restituendo una probabilità di presenza (idoneità)

[Steven J. Phillips](#), [Robert P. Anderson](#) and [Robert E. Schapire](#), Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, Vol 190/3-4 pp 231-259, 2006.

[Steven J. Phillips](#) and [Miroslav Dudik](#), Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography*, Vol 31, pp 161-175, 2008.

Dove fare il download del software maxent:
<http://www.cs.princeton.edu/~schapire/maxent/>
 This software may be freely downloaded and used for all educational and research activities



molti vantaggi...

- ✓ Utilizza solo dati di presenza, insieme ad informazioni ambientali nell'area di studio
- ✓ non è necessario raccogliere dati di assenza
- ✓ può utilizzare variabili ambientali continue o categoriche
- ✓ efficiente algoritmo deterministico per convergere ad un'ottimale distribuzione della probabilità
- ✓ restituisce una probabilità di presenza continua (tra 0 e 1)

...e alcuni svantaggi

- ✓ metodo statistico non testato ampiamente come altri metodi (GLM o GAM), poche linee guida disponibili
- ✓ possibile overfitting a causa della intercorrelazione delle variabili ambientali
- ✓ è necessario un software specifico, non compreso in altri pacchetti statistici
- ✓ necessità di predisporre di dati ambientali (layers) ad hoc

dati (samples) e variabili ambientali (layers) utilizzati



Features utilizzati per la realizzazione dei 9 modelli:

- ✓ 13 categorie di uso del suolo accorpando classi di Corine IV livello (trasformati in distanze di ogni punto della Regione dalla categoria trattata più vicina)
- ✓ Altitudine
- ✓ Distanza dai centri abitati

- ✓ 9 Specie (samples): Coturnice, Calandra, Calandrella, Calandro, Balia dal collare, Capinera, Cincia bigia, Fringuello, Strillozzo: 5576 records.

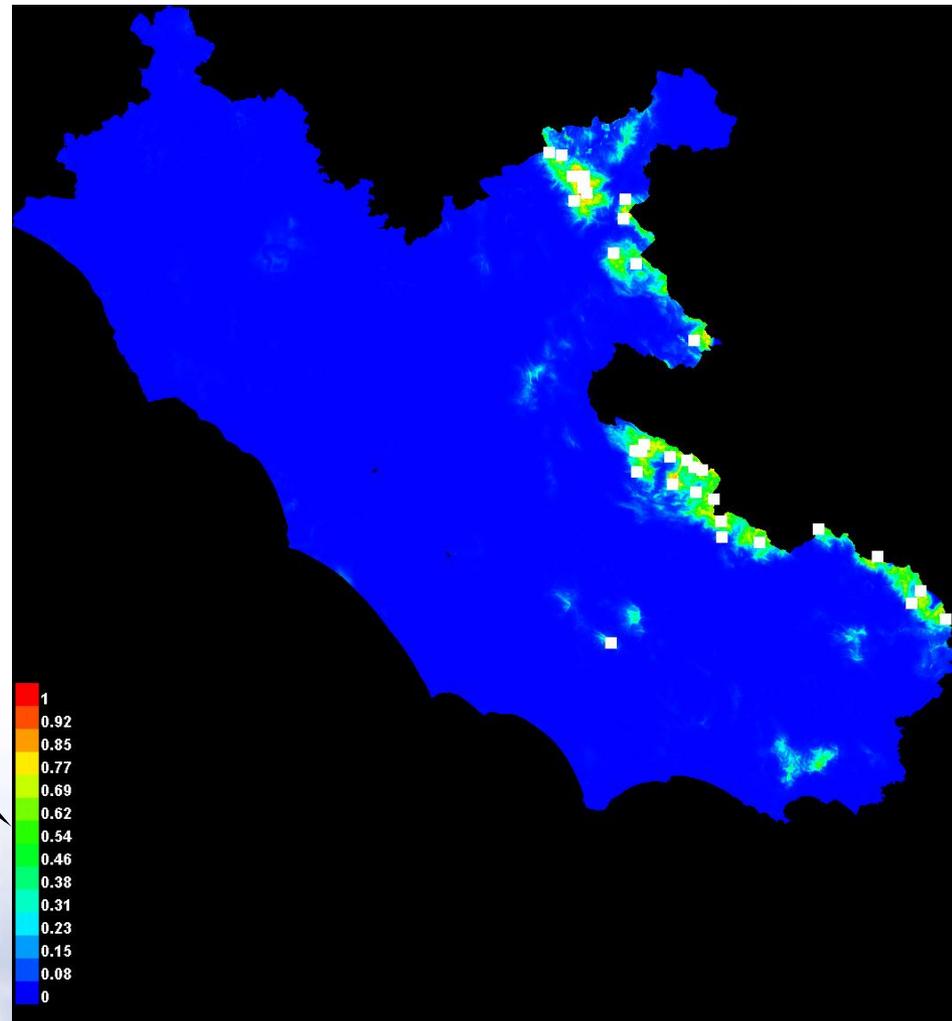
Pictures of the model

This is a representation of the Maxent model for **Balia dal collare**.

Warmer colors show areas with better predicted conditions.

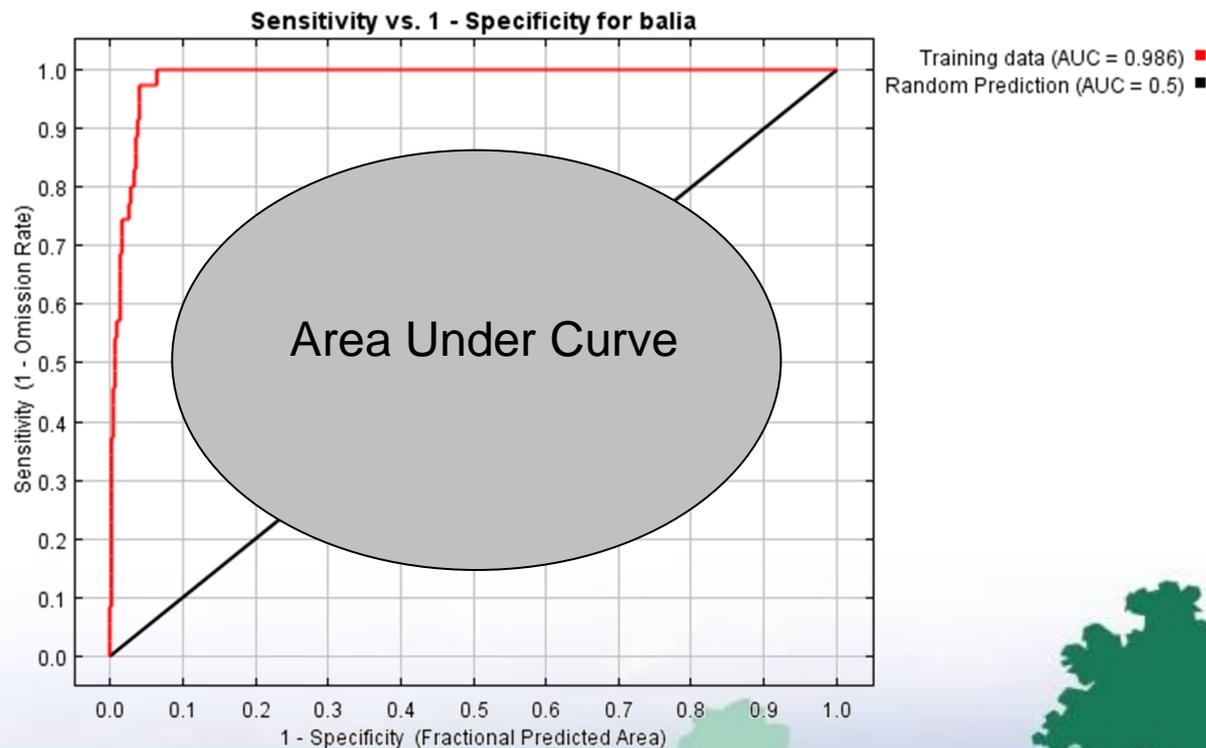
White dots show the presence locations used for training.

Probabilità
di presenza



Analysis of omission/commission

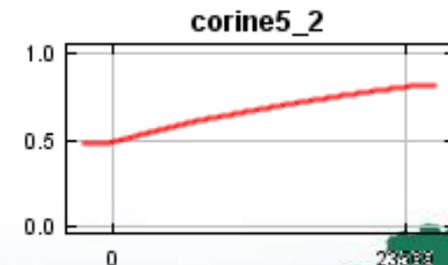
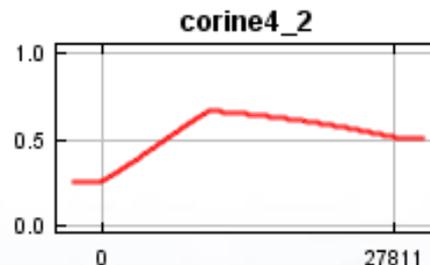
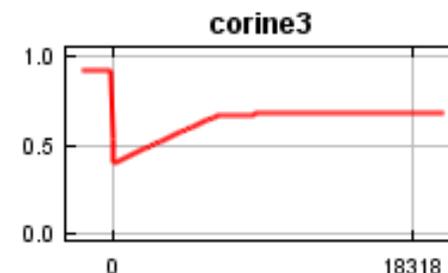
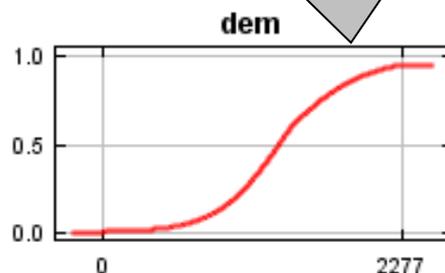
Receiver Operating Characteristic (ROC) curve e AUC (Area Under Curve): valuta la capacità discriminante tra omissioni e inclusioni (test perfetto=100%)



Response curves

These curves show how each environmental variable affects the Maxent prediction.

The curves show the marginal effect of changing exactly one variable, whereas the model may take advantage of sets of variables changing together.



Contributo al
modello (in%)

Decremento % di AUC
in seguito al permutare
dei valori delle variabili

Analysis of variable contributions

The following table gives estimates of relative contributions of the environmental variables to the Maxent model.

Variable	Percent contribution	Permutation importance
dem	46.7	72.4
corine3	22.7	2.7
corine4_2	13	0.9
corine5_2	5.7	0.3
corine10	3.4	2.8
aspectvar	2.2	4.8
popd	2.1	1.8
corine12	1.2	2.4
demvar	0.8	2.6
corine11	0.5	0.4
corine7_2	0.4	6
corine6_2	0.3	0.1
aspect	0.3	0
slopevar	0.2	0.3
abitatid	0.2	0.5
corine2	0.1	0.4
corine8_2	0.1	0.6
corine9_2	0	0.6
corine1	0	0.1
fiumi	0	0
slope	0	0.2
corine13	0	0

Quanto è buono il modello senza questa variabile?

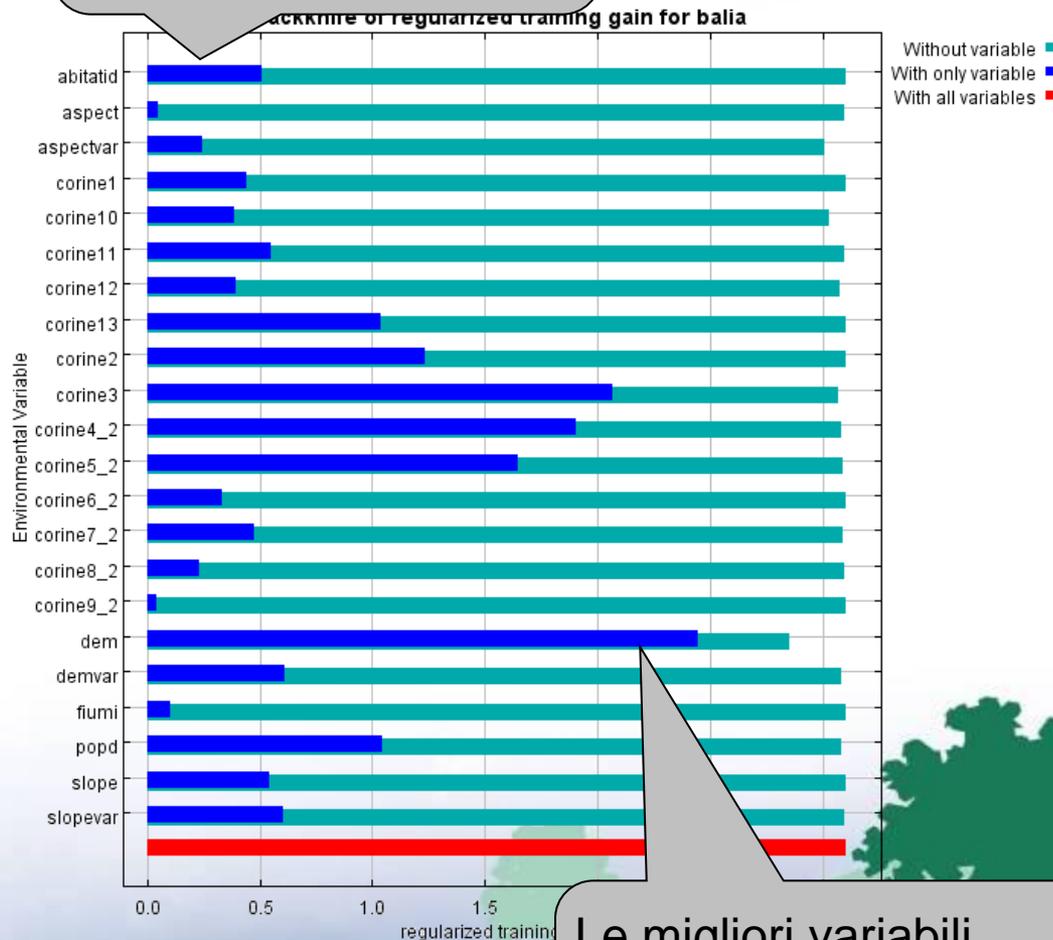


Analysis of variable contributions

Results of the jackknife test of variable importance.

The environmental variable with highest gain when used in isolation is **dem**.

The environmental variable that decreases the gain the most when it is omitted is **dem**, which therefore appears to have the most information that isn't present in the other variables.



Le migliori variabili hanno le barre blu lunghe

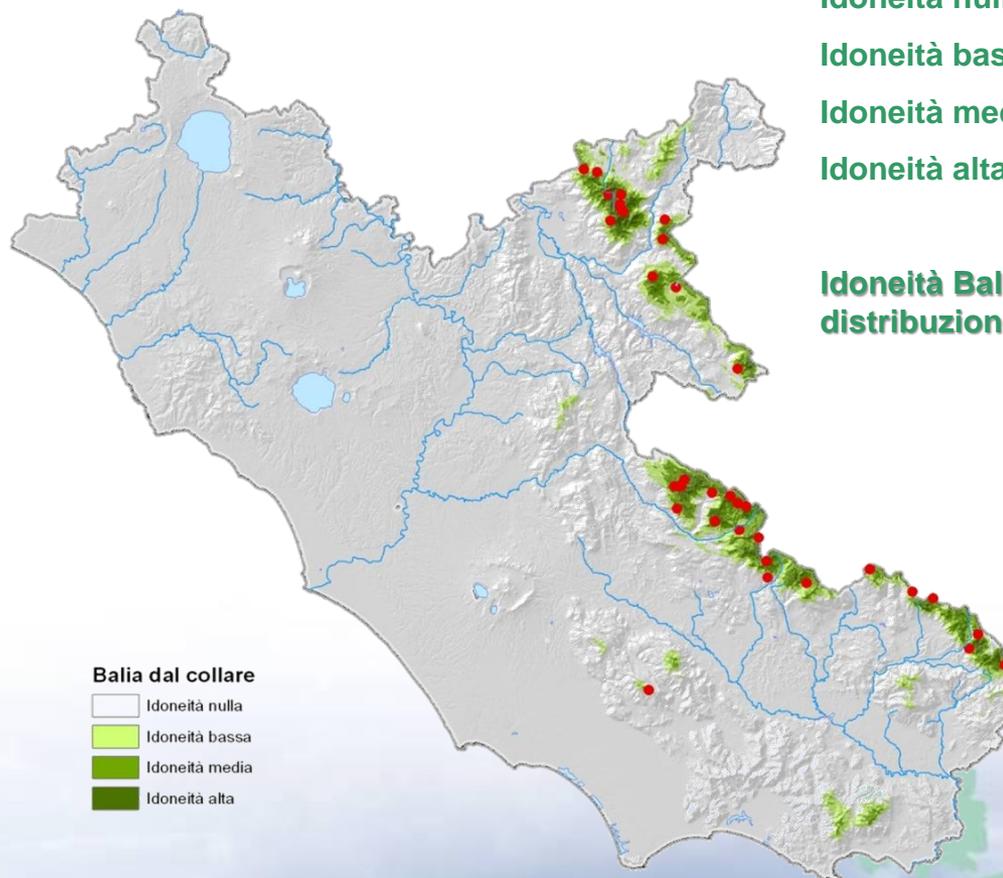
Test riassuntivi della predittività dei modelli delle 9 specie

Specie	<i>n</i> samples	AUC	Contributo variabili
Coturnice	42	0,992	altitudine (81,7);corine4_2(3,8);corine 10(3,8);corine13(2,4);aspectvar(1,2)
Calandra	60	0,946	corine1 (48);abitatid(7,6);altitudine_var(7,3);pendenza_var(7,1);corine11(5,6)
Calandrella	44	0,95	altitudinevar(32,4);corine1(18,8);corine7_2(14,5);abitatid(6,8);corine8_2(5);popd (4,6);corine11(3,1)
Calandro	178	0,922	corine6_2(20,9);abitatis(14,3);corine4_2(13,7);altitudine(11,4);corine8_2(8)
Capinera	2450	0,685	corine1(25,7);corine7_2(9,7);corine8_2(8,2);corine3(7,2);corine11(6,7)
Balia dal collare	37	0,986	altitudine (51,3); boschi (20,6); corine4_2(14,4);corine 5_2(4,6)
Cincia bigia	220	0,918	altitudine(53,5);corine1(8,7);corine7_2(5,2);corine5_2(4,1);corine2(3,9);pendenza_var(3,1)
Fringuello	1746	0,728	corine3(22,2); corine7_2(13,2); corine1(12,4);corine11(7,5); dem(6,5); corine4_2(5,7); corine9_2(4,7)
Strillozzo	799	0,781	abitatid(14,3);pendenza(12,6);corine7_2(10,8);corine5_2(9,7);corine1(8,7);corine4_2(8,2)

Il modello della Balia dal collare

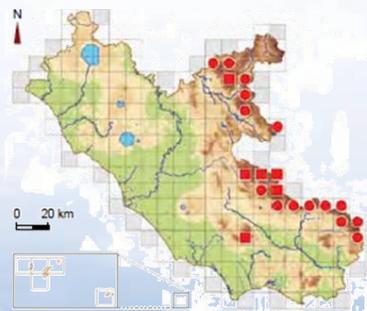
Idoneità nulla	0,0 - 0,1	1.618.589 ha
Idoneità bassa	0,1 - 0,3	49.331 ha
Idoneità media	0,3 - 0,6	37.519 ha
Idoneità alta	0,6 - 1,0	14.704 ha

Idoneità Balia dal collare/area di distribuzione potenziale: 101.554 ha

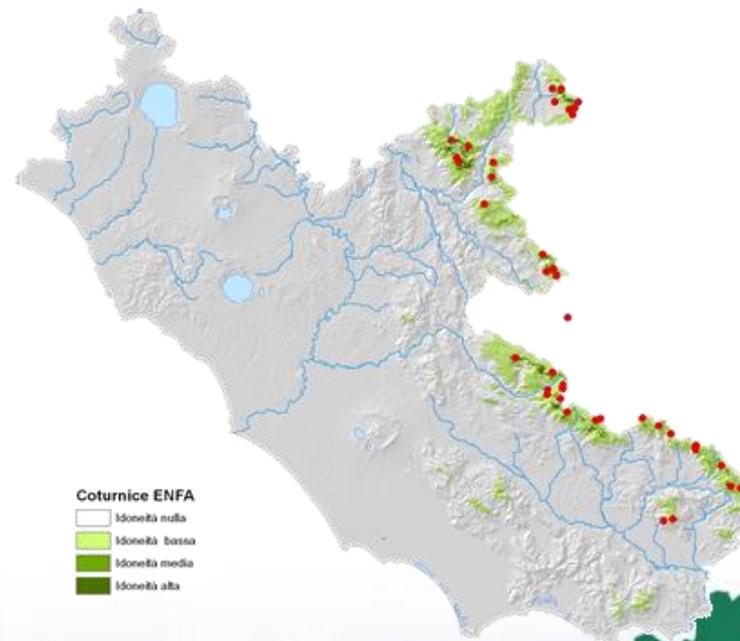
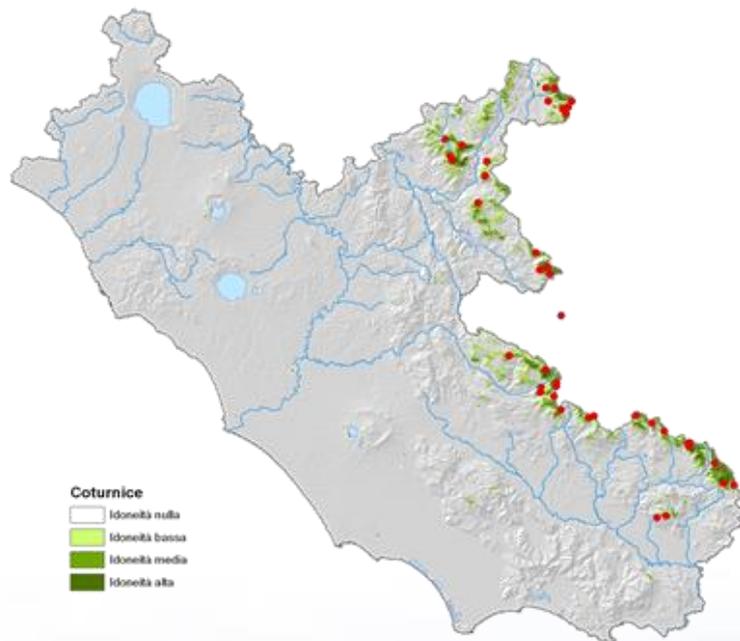


Balia dal collare

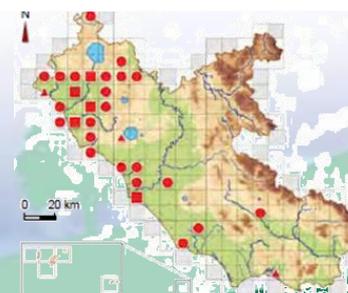
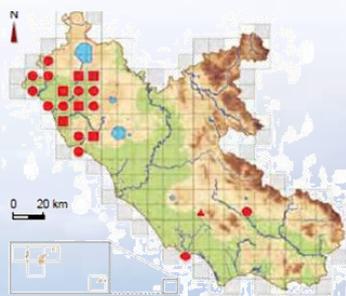
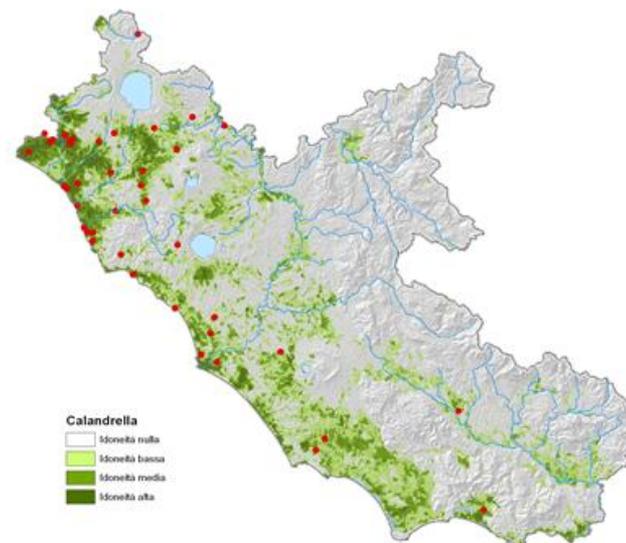
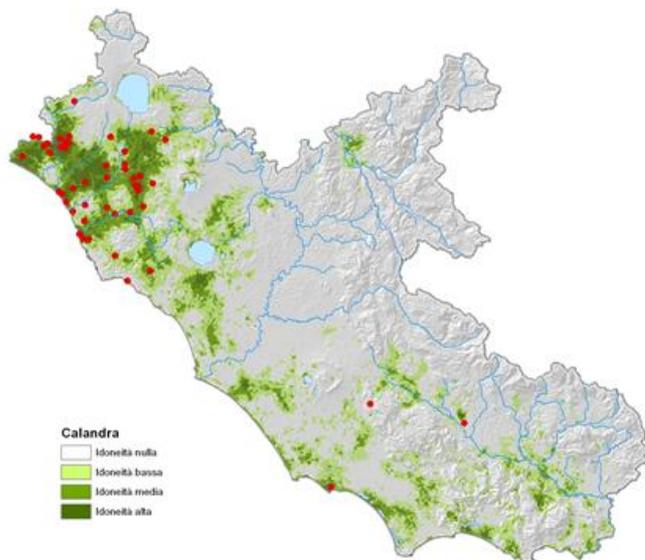
- Idoneità nulla
- Idoneità bassa
- Idoneità media
- Idoneità alta



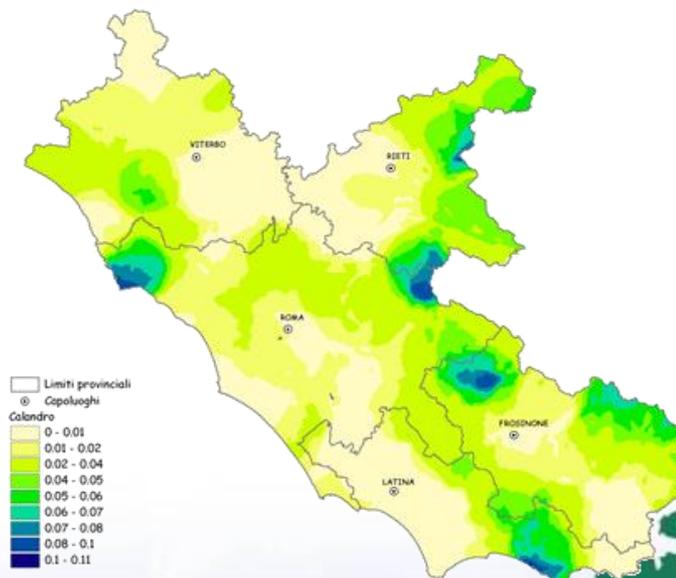
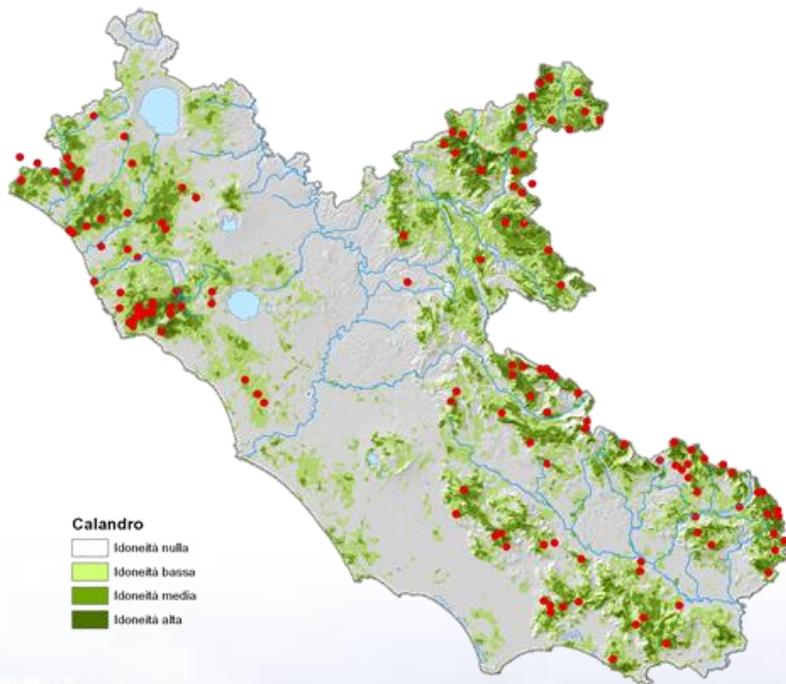
Il modello della Coturnice: maxent ed ENFA a confronto



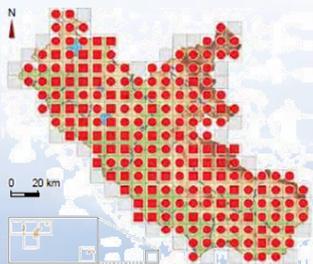
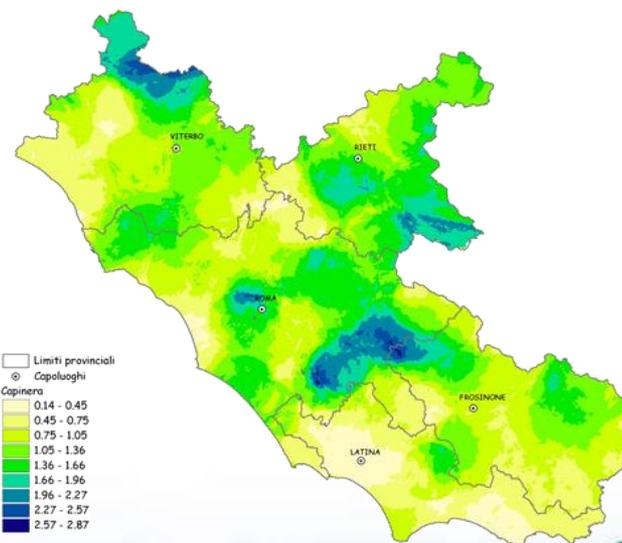
I modelli per Calandra e Calandrella



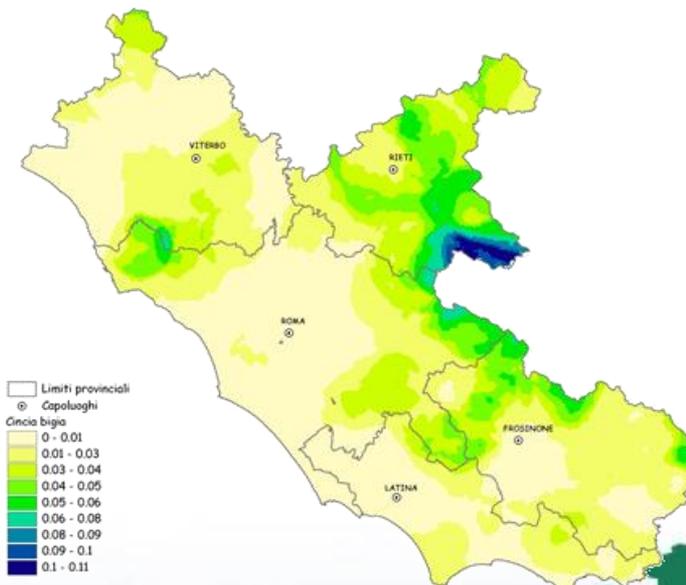
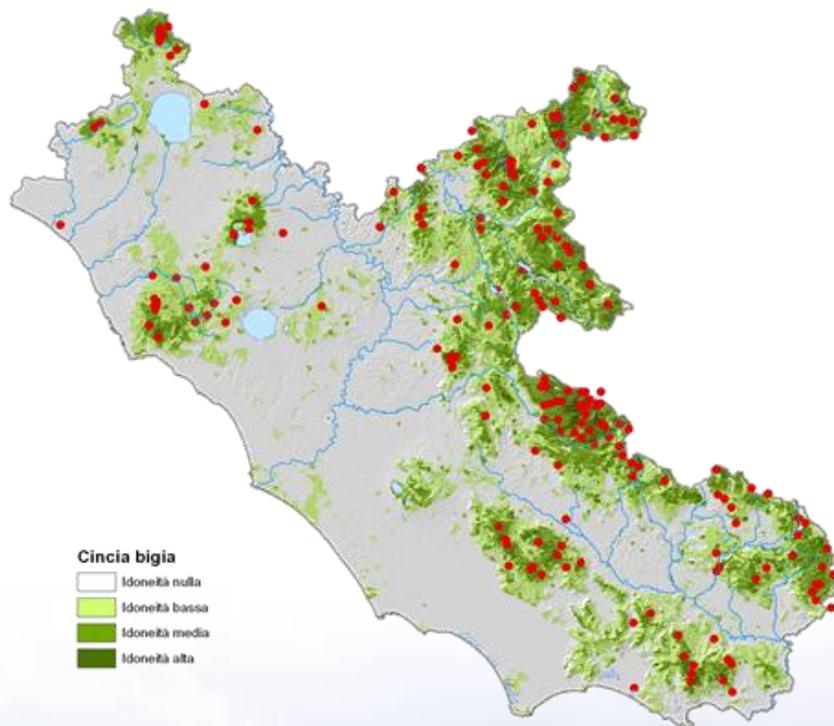
Il modello del Calandro: maxent e interpolazione con kriging



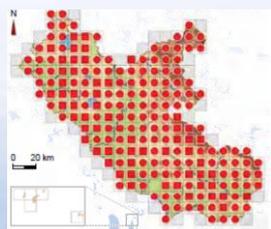
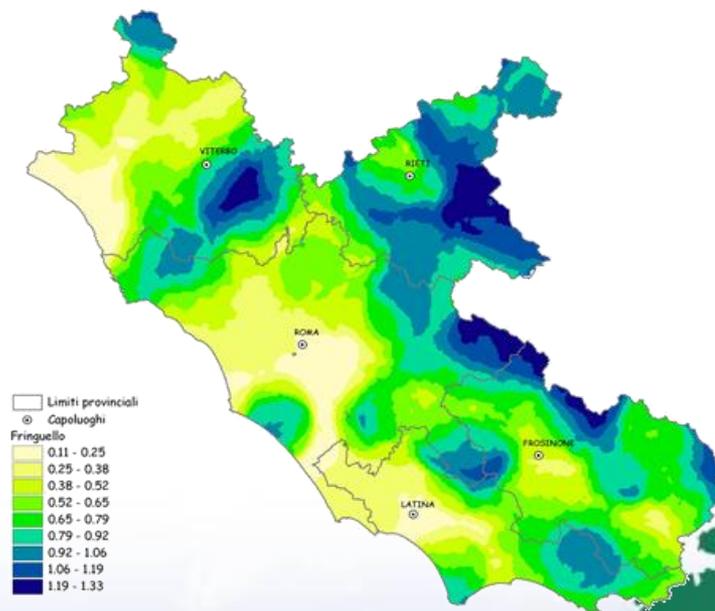
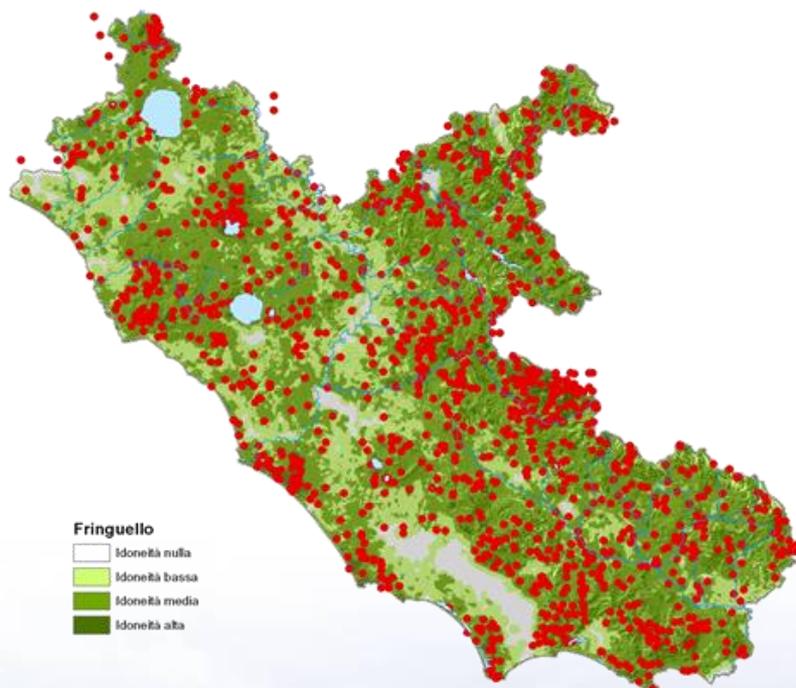
Il modello della Capinera: maxent e interpolazione con Kriging



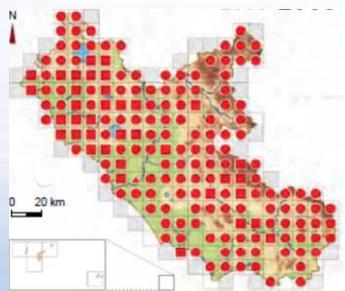
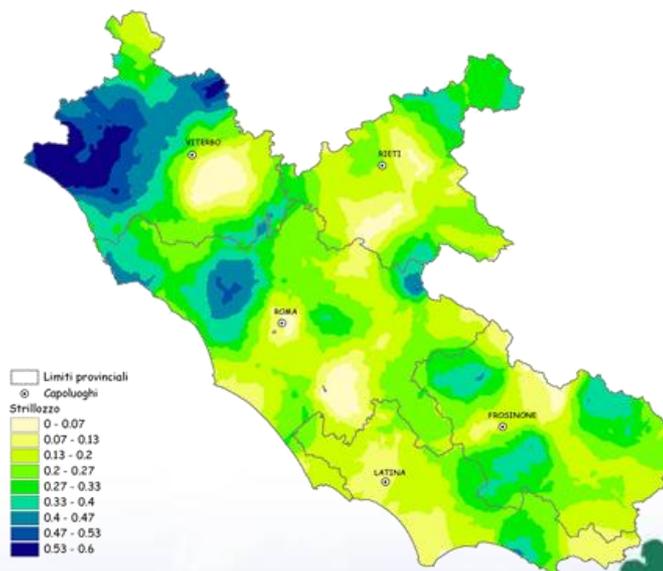
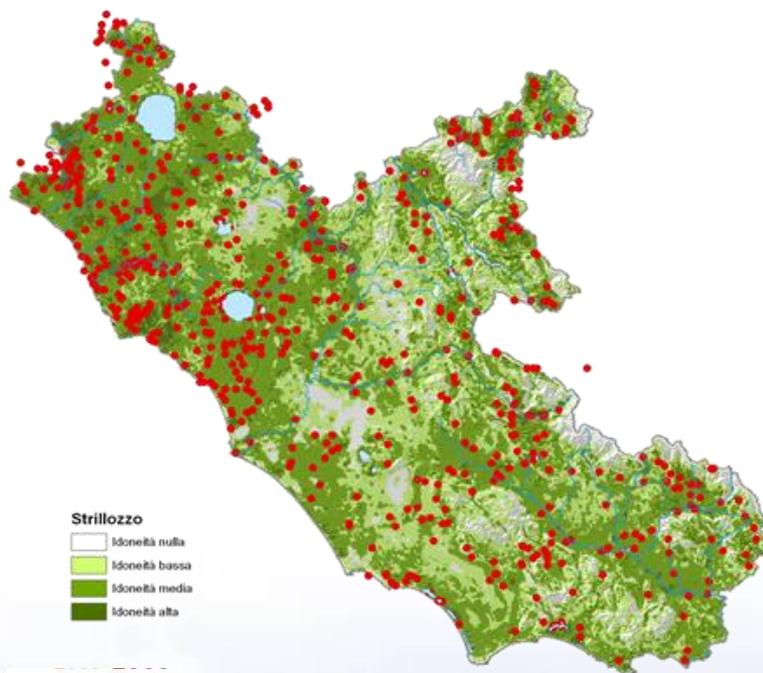
Il modello della Cincia bigia: maxent e interpolazione con kriging



Il modello del Fringuello: maxent e interpolazione con kriging



Il modello dello Strillozzo: maxent e interpolazione con kriging





Buoni propositi per sviluppi futuri

- ✓ aggiornare e migliorare i layers ambientali (features)
- ✓ ridurre i problemi di autocorrelazione (*overfitting*), minimizzare gli errori di omissione (false assenze) e commissione (false presenze)
- ✓ stimare la dimensione degli areali e delle consistenze regionali
- ✓ verificare la relazione tra gradi di idoneità e densità delle specie
- ✓ verifica dell'efficienza delle AAPP e rete Natura 2000 (*gap analysis*)

Grazie per l'attenzione

... grazie soprattutto a tutti coloro che hanno contribuito a vario titolo alla raccolta e alla sistematizzazione dei dati