

## MODELLI DI RISCHIO DI INNESCO DI INCENDIO IN VALLE D'AOSTA: ANALISI DELLE INTERAZIONI TRA COMPONENTI NATURALI E ANTROPICHE PER L'OTTIMIZZAZIONE DEI MODELLI

Cristiano Foderi<sup>1</sup>, Giorgio Vacchiano<sup>2</sup>

<sup>1</sup>GESAAF, Università di Firenze, Firenze (Firenze); cristiano.foderi@unifi.it

<sup>2</sup>DISAFA, Università di Torino, Grugliasco (Torino)

In questo lavoro sono stati impiegati modelli di probabilità del rischio di innesco degli incendi basati su un algoritmo di massima entropia implementato nel *software* Maxent. Il modello di rischio è stato implementato per il territorio della Regione Autonoma Valle d'Aosta partendo da un archivio regionale di inneschi registrati dal 1995 al 2009. L'archivio è stato suddiviso in base alla stagione e alla tipologia di vegetazione interessata dagli inneschi; sono stati così analizzati separatamente 3 *dataset*: i) incendi invernali in bosco; ii) incendi invernali su prati-pascoli ed incolti; iii) incendi estivi. I predittori considerati per la caratterizzazione degli inneschi sono stati individuati tra quelli che descrivono la morfologia, il clima e l'uso del suolo nel territorio investigato. La componente antropica è stata considerata includendo tra i predittori la distanza del punto di innesco dagli edifici e dalla rete viaria. Inoltre, per indagare le relazioni esistenti tra il rischio di innesco e la realtà socio-economica del territorio, sono stati inseriti i dati relativi al numero di capi al pascolo e alla consistenza numerica delle aziende zootecniche con animali al pascolo, ricavati dal più recente censimento ISTAT dell'agricoltura (2010). Per limitare gli errori causati da multi-collinearità dei predittori ambientali è stata condotta una analisi delle componenti principali (PCA). I modelli sono stati prodotti utilizzando il 70% dei punti dei *dataset* e validati utilizzando il restante 30% come test indipendente, selezionato tramite *bootstrap* su 5 reiterazioni di calcolo per ciascun *dataset*. I predittori ambientali sono stati analizzati con una procedura di *Leave One Out* con il metodo *JRR* per misurare l'apporto delle singole variabili all'accuratezza modello.

*Parole chiave*: incendi boschivi, MaxEnt, modello di rischio, PCA, innesco.

*Keywords*: forest fire, MaxEnt, risk model, PCA, ignition.

<http://dx.doi.org/10.4129/2cis-cf-mod>

### 1. Introduzione

Gli incendi boschivi hanno sempre rappresentato un pericolo per le foreste della Valle d'Aosta ed è ancora vivo in molte persone il ricordo dei disastrosi eventi del 1965, del 1973 e del 1981. Nel 1989 sono stati contati circa 370 incendi e nel 1990 ne sono stati registrati 362, con una superficie bruciata che superava i 1200 ettari di territorio e una distruzione di soprassuolo boschivo di circa 150 ettari. I fattori predisponenti e determinanti che concorrono all'innesco degli incendi boschivi sono molteplici. Essi dipendono dalla morfologia del territorio, dal clima, dalle caratteristiche della vegetazione e dalle attività umane.

Lo studio ha interessato il territorio della Regione Autonoma Valle d'Aosta analizzando 15 anni di incendi e prendendo in considerazione gli strati informativi che meglio caratterizzano i fenomeni di innesco di incendio nei diversi periodi dell'anno.

### 2. Materiali e metodi

#### 2.1 Dati di base e selezione dei dataset

L'archivio degli inneschi è stato estratto dalla banca dati geografica della Regione Autonoma Valle d'Aosta,

prendendo in esame gli eventi registrati dal gennaio 1995 al dicembre 2009, per un totale di 1077 incendi. Dalla analisi di frequenza degli inneschi, in relazione ai mesi dell'anno e alla tipologia di vegetazione prevalente interessata dagli incendi si è ritenuto opportuno, viste le finalità dello studio, suddividere l'archivio in 3 *dataset*. Sono stati così analizzati separatamente:

- i) 244 incendi estivi, registrati da maggio a settembre di ogni anno;
- ii) 305 incendi invernali in bosco, registrati da ottobre ad aprile di ogni anno sulle superfici a prevalente copertura forestale;
- iii) 507 incendi invernali su prati-pascoli ed incolti registrati da ottobre ad aprile di ogni anno sulle superfici in cui la copertura prevalente è costituita da prati e pascoli o aree a vegetazione incolta.

I predittori considerati per la caratterizzazione degli inneschi sono stati individuati tra quelli che descrivono la morfologia, il clima e l'uso del suolo nel territorio investigato e sono di seguito descritti.

Le variabili climatiche sono state ricavate dagli strati informativi disponibili dal sito worldclim :

(<http://www.worldclim.org>). La radiazione solare incidente è stata valutata attraverso il calcolo, in ambiente GIS, del *HeatLoad Index* (McCune e Keon, 2002). La

quota altimetrica è stata rappresentata attraverso il modello digitale di elevazione (DEM), dal quale è stato possibile calcolare le pendenze. Dal DEM è stato inoltre calcolato il *Southwestness index* (Franklin *et al.*, 2000) per valutare la propensione alla xericità del territorio. Lo strato informativo della vegetazione è stato ricavato dai dati riportati nel progetto *Corine Land Cover 2006* e attraverso il calcolo *Normalized Difference Vegetation Index* (NDVI)

La componente antropica è stata considerata includendo tra i predittori la distanza del punto di innesco dagli edifici e dalla rete viaria. Inoltre, per indagare le relazioni esistenti tra il rischio di innesco e la realtà socio-economica del territorio, sono stati inseriti i dati relativi al numero di capi al pascolo e alla consistenza numerica delle aziende zootecniche con animali al pascolo, ricavati dal più recente censimento ISTAT dell'agricoltura (2010). Il dato relativo alla consistenza numerica delle aziende, in rapporto al numero complessivo di capi di un territorio, rispecchia il livello di sfruttamento del territorio in termini di concentrazione del carico di pascolo: è noto infatti che situazioni di forte carico di pascolo comportano un contenimento della biomassa combustibile (Bovio e Marchi, 2010), limitando il rischio di incendi nel periodo invernale e svolgendo una funzione preventiva importante soprattutto nelle aree, sensibili al fuoco, boscate o limitrofe al bosco.

Tutti gli strati informativi descritti sono stati convertiti in formato raster con risoluzione a 500 m e sono riassunti nel seguente elenco:

- Temperatura media annuale (BIO1);
- Temperatura massima del mese più caldo (BIO5);
- Temperatura minima del mese più freddo (BIO6);
- Range annuale di temperature (BIO7);
- Temperatura media del trimestre più caldo (BIO10);
- Temperatura media del trimestre più freddo (BIO11);
- Precipitazioni annuali (BIO12);
- Precipitazioni del trimestre più caldo (BIO18);
- Modello digitale di elevazione del terreno (DEM);
- Distanza con strutture ed infrastrutture umane (DIST\_INFR);
- *Heatload Index* (McCune and Keon, 2002);
- Aziende con animali al pascolo dal censimento dell'agricoltura ISTAT 2010 (N\_FARMS);
- Numero di animali al pascolo dal censimento dell'agricoltura ISTAT 2010 (N\_GRAZ\_ANIMALS);
- *Normalized Difference Vegetation Index* (NDVI);
- Pendenza in gradi calcolata dal DEM (SLOPE);
- *Southwestness Index* calcolato come Coseno (esposizione - 225°), (Franklin *et al.*, 2000);
- Tipologia di vegetazione da *Corine Land Cover* (VEGETATION).

## 2.2 Analisi preliminare dei dataset e selezione delle variabili

Un ampio set di strati informativi legati ai predittori ambientali spesso sono intrinsecamente affetti da multi-collinearità, un problema statistico definito come un alto grado di correlazione tra covariate (Dormann, 2013). La multi-collinearità sorge, infatti, quando c'è un'elevata correlazione tra due o più variabili esplicative, che comporta come conseguenza il fatto che diano la

stessa informazione. In questa situazione il modello che verrà generato non dà la possibilità di interpretare ed attribuire un significato a ciascuna covariante. Per limitare gli errori causati da multi-collinearità dei predittori ambientali è stata condotta una analisi delle componenti principali (PCA).

Dai risultati delle PCA è stato possibile selezionare i predittori per ogni *dataset* sulla base dei punteggi di *loading*, registrati per le principali direttrici di covarianza, di ciascuna variabile.

## 2.3 Modellizzazione dei fenomeni di innesco

Per ciascun *dataset* è stato elaborato un modello utilizzando il programma Maxent (Phillips *et al.*, 2006), nella sua versione più recente (3.3.3.k).

Ciascun modello, prima confronta le informazioni ambientali disponibili per i punti di innesco con quelle disponibili per l'intero territorio (punti di *background*), poi, seguendo il principio di massima entropia, restituisce una distribuzione spaziale della probabilità di sviluppo di fenomeni simili sul territorio. L'algoritmo, individuando la distribuzione a entropia massima sotto le limitazioni imposte dai fattori ambientali, permette di ottimizzare le informazioni legate a ciò che si conosce, cioè i dati di presenza, evitando, al contempo, di produrre assunzioni circa ciò che non si conosce (Jaynes, 1990).

I modelli sono stati prodotti utilizzando il 70% dei punti dei *dataset* e validati utilizzando il restante 30% come test indipendente, selezionato tramite *bootstrap* su 5 iterazioni di calcolo per ciascun *dataset*.

## 2.4 Valutazione delle performance dei modelli

Per valutare le performance dei modelli è stato calcolato il valore dell'area sottesa alla curva ROC (*Receiver Operating Characteristic*), un metodo che è largamente impiegato per questi scopi (Fielding e Bell, 1997), e che ha il pregio di permettere confronti che non sono influenzati né dalle dimensioni del campione, né dal numero di fattori ambientali coinvolti.

La curva ROC (Fielding e Bell, 1997) è un indicatore della capacità di classificazione ottenibile da un modello predittivo della presenza di un fenomeno: in particolare, l'area (AUC) sottesa alla curva ROC (Fielding e Bell, 1997), che al massimo può assumere un valore pari a 1 (100%), indica la qualità del modello stesso.

Valori pari a 0,5 (50%) indicano un comportamento assolutamente casuale, valori di 0,6 (60%), ad esempio, indicano un modello che nel 40% dei casi effettua previsioni errate. L'analisi di questi valori, e la scelta di soglie efficienti, pur ponendo numerose difficoltà (Lobo *et al.*, 2008), permette di considerare efficienti i modelli il cui valore di AUC supera 0,8 (Manel *et al.*, 2001). I predittori ambientali sono stati analizzati con una procedura di *Leave One Out* con il metodo *Jackknife repeated replication* (JRR) (Quenouille, 1949) per misurare l'apporto delle singole variabili all'accuratezza modello.

Attraverso la procedura JRR è stato possibile ricalcolare le stime di AUC dei modelli escludendo di volta in volta un predittore ambientale. In questo modo si evidenzia quanto l'assenza di quel parametro incida sul comportamento del fenomeno indagato.

### 3. Risultati

Dalla analisi delle componenti principali (PCA) dei tre *dataset* emerge chiaramente la forte covariazione dei predittori considerati per la creazione dei modelli.

In particolare dai biplot in figura 1 si distinguono nettamente le direttrici di covariazione legate alle variabili climatiche e al DEM, così come è chiara la covariazione delle variabili utilizzate per indagare le relazioni esistenti tra il rischio di innesco e la realtà socio-economica del territorio legate alle attività di pascolo. Da queste considerazioni e in base ai punteggi di *Loadings* registrati dalla PCA sono state selezionate le variabili riassunte in tabella 1 e utilizzate per la produzione dei modelli di distribuzione di rischio per gli inneschi dei tre *dataset*. Di seguito sono riportati i risultati ottenuti per le tre condizioni analizzate in termini di attendibilità del modello (valore di AUC). Sono poi riportati i grafici dei risultati della procedura JRR per la valutazione del contributo delle singole variabili utilizzate per la creazione dei modelli.

Dal calcolo della AUC, a seguito della procedura JRR si evidenziano le diminuzioni registrate a seguito della esclusione di ogni variabile. Minore è il risultato registrato in corrispondenza della eliminazione di una variabile dal calcolo del modello, più questa variabile risulta informativa e quindi importante per la caratterizzazione e l'interpretazione di quel fenomeno.

#### 3.1 Incendi estivi

L'elaborazione del modello per gli inneschi in incendi estivi ha registrato un valore di AUC di 0,83 come media dei valori registrati nelle cinque reiterazioni di calcolo. I predittori che hanno fatto registrare la maggiore diminuzione in termini di AUC (Fig. 2) sono il *Southwestness Index*, il DEM e le precipitazioni nel trimestre più caldo (BIO18).

#### 3.2 Incendi invernali su aree boscate

L'elaborazione del modello per gli inneschi in incendi invernali in aree a prevalente copertura forestale ha registrato un valore di AUC di 0,92 come media dei valori registrati nelle cinque reiterazioni di calcolo.

I predittori che hanno fatto registrare la maggiore diminuzione in termini di AUC (Fig. 3) sono la tipologia di vegetazione e le temperature medie del trimestre più freddo (BIO11).

#### 3.3 Incendi invernali su aree incolte e pascoli

L'elaborazione del modello per gli inneschi in incendi invernali in aree a pascolo sugli incolti ha registrato un valore di AUC di 0,895 come media dei valori registrati nelle cinque reiterazioni di calcolo.

I predittori che hanno fatto registrare la maggiore diminuzione in termini di AUC (Fig. 4) sono la tipologia di vegetazione, la pendenza e le temperature minime del mese più freddo (BIO06).

Per quanto riguarda le variabili direttamente legate alle attività antropiche non si registrano decrementi particolarmente significativi in nessuna delle condizioni analizzate; di seguito sono riportati i grafici ROC, in figura 5, calcolati dalla procedura JRR per le sole variabili antropiche considerate escludendo tutte le altre.

### 4. Conclusioni

L'uso della PCA per selezionare i predittori ambientali ha permesso di generare modelli più facilmente interpretabili, con una accuratezza maggiore dell'80% in termini di valore di AUC. Nelle condizioni esaminate, la metodologia si è rivelata efficace nella valutazione dei fattori che influenzano maggiormente gli inneschi. Dall'analisi delle elaborazioni e delle rappresentazioni spaziali prodotte per i tre *dataset* sono state evidenziate le interazioni esistenti tra i fattori ambientali e i fattori legati alle attività umane che caratterizzano gli inneschi degli incendi.

I predittori che maggiormente influiscono sulla distribuzione della probabilità di rischio sono le precipitazioni, la quota altimetrica e il *Southwestness Index* in estate.

Nel periodo invernale invece risultano essere determinanti le condizioni legate alle temperature e la tipologia di vegetazione.

Si registra inoltre una generale tendenza alla diminuzione di probabilità all'aumento della distanza dalle infrastrutture e, seppur in modo marginale, si registra una maggiore probabilità nelle zone a più alta concentrazione di capi al pascolo per i *dataset* invernali.

A questo comportamento può aver concorso la scarsa risoluzione spaziale delle informazioni reperibili dal censimento dell'agricoltura 2010 di ISTAT, che per le sue finalità riporta dati aggregati e non di dettaglio per quanto riguarda le aziende e il numero di animali al pascolo.

Tabella 1. Variabili selezionate dopo l'analisi delle componenti principali per ridurre la multi-collinearità.  
 Table 1. Selected variables after principal components analysis to reduce multi-collinearity.

<i>Variable selection after PCA</i>	<i>Winter fires on grassland &amp; uncult</i>	<i>Winter forest fires</i>	<i>Summer fires</i>
bio01			
bio05			x
bio06	x		
bio07		x	x
bio10			
bio11		x	
bio12	x	x	x
bio18			x

(Segue Tabella 1)  
 (Table 1. Continued)

dist_infr	X	X	X
dtm	X	X	X
heatloadindex	X		
N_farm			
N_graz_animals	X	X	X
ndvi			X
slope	X	X	X
Southwestness	X	X	X
vegetation	X	X	X

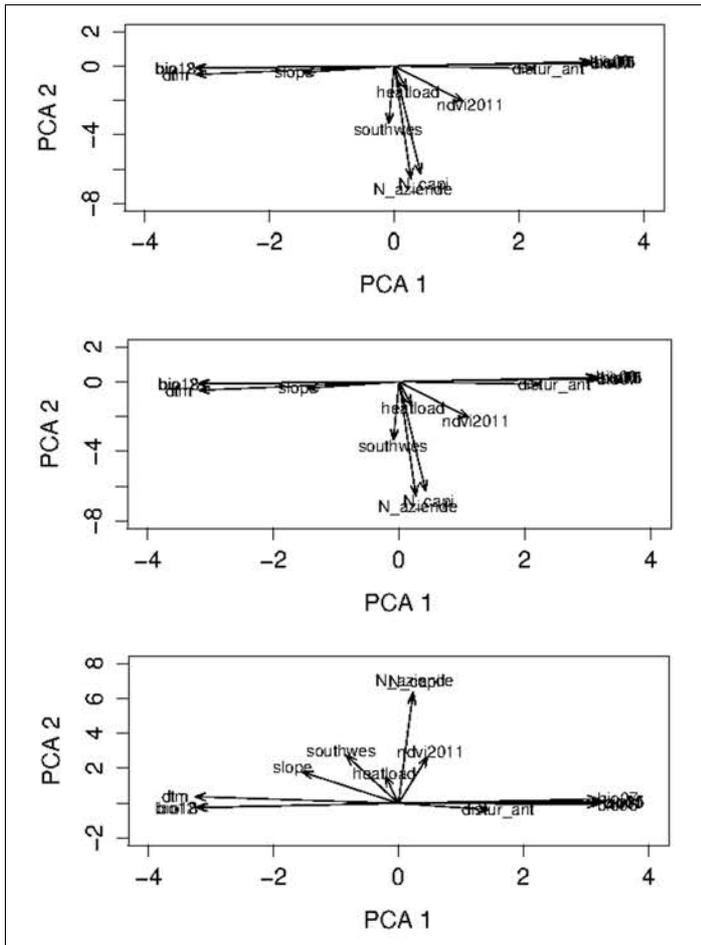


Figura 1. Biplot delle componenti principali, in ordine dal basso in alto: incendi estivi, incendi boschivi invernali e incendi invernali su pascoli ed aree incolte.  
 Figure 1. Principal components biplot, from bottom to top order: summer fires, winter forest fires and winter fires on grassland and uncultivated area.

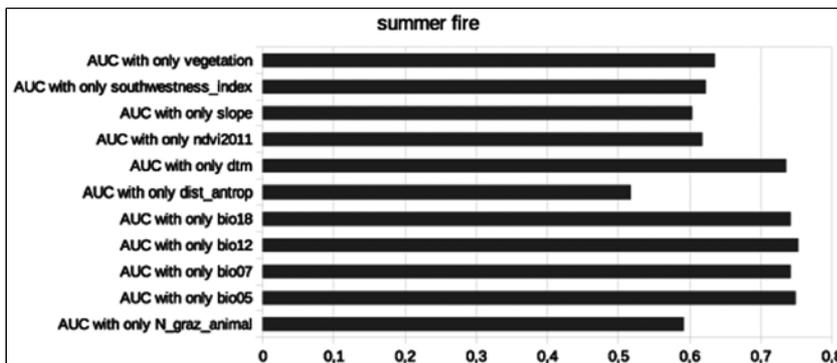


Figura 2. Diminuzione di AUC a seguito della esclusione di ogni variabile per gli incendi estivi.  
 Figure 2. AUC reduction after each variable exclusion in summer fires.

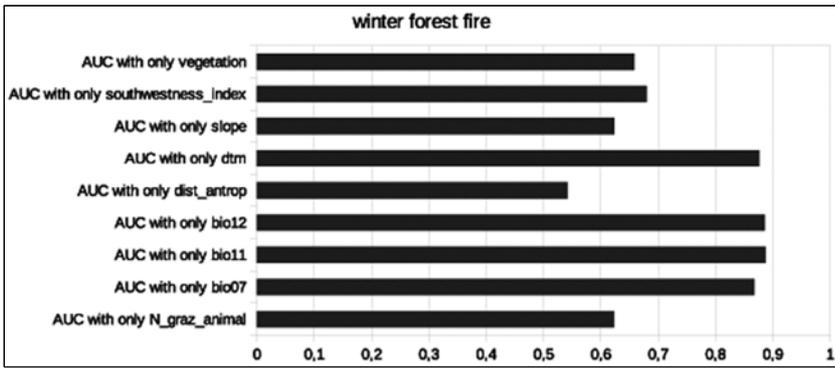


Figura 3. Diminuzione di AUC a seguito della esclusione di ogni variabile per gli incendi boschivi invernali.  
 Figure 3. AUC reduction after each variable exclusion in winter forest fires.

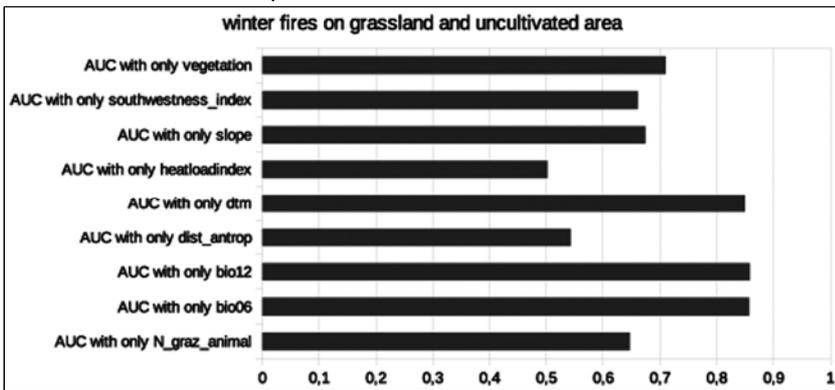


Figura 4. Diminuzione di AUC a seguito della esclusione di ogni variabile per gli incendi invernali su pascoli ed aree incolte.  
 Figure 4. AUC reduction after each variable exclusion in winter fires on grassland and uncultivated area.

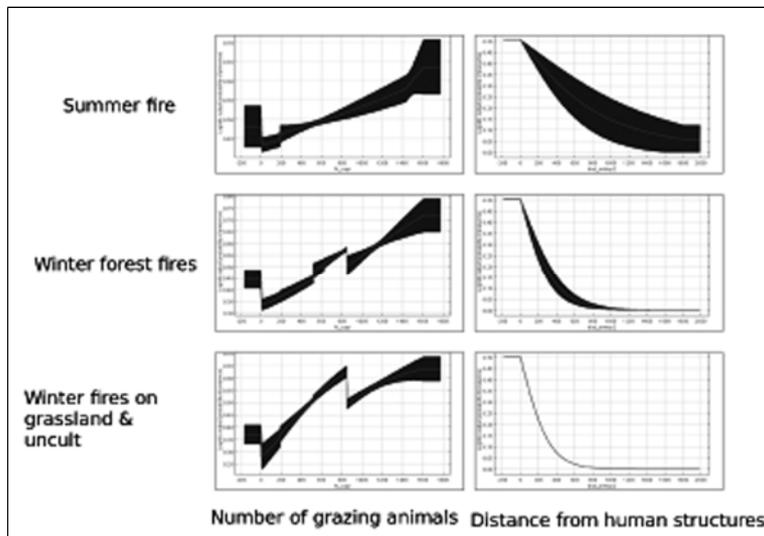


Figura 5. AUC delle variabili antropiche utilizzate.  
 Figure 5. AUC of used anthropogenic variables.

## SUMMARY

### Wildfire ignition risk modeling in Aosta Valley: natural and anthropogenic components interactions analysis for models optimization

In this work were used probability risk models of wildfire ignition based on the maximum entropy algorithm, implemented in Maxent software. The risk model has been

implemented for the Autonomous Region of Valle d'Aosta from a regional archive of ignitions recorded from 1995 to 2009. The archive has been split according to the season and the vegetation type affected by ignition; 3dataset were analyzed separately: i) winter forest fires; ii) winter fires on grassland and uncultivated area; iii) summer fires. The predictors considered for the characterization of ignitions were selected from among those that represent the expression of morphology, climate and

land use of the area investigated; distances from buildings and roads was considered to evaluate the anthropogenic component. In addition, to investigate relationship between the risk of ignition and the socio-economic reality of the territory, were included data for the number of animals grazing and the number of farms with grazing animals, derived from the most recent agriculture census ISTAT (2010).

To reduce errors caused by multi-collinearity of the environmental predictors was conducted a principal components analysis (PCA). The models were produced using 70% of the points of the datasets and validated using the remaining 30% as test independently selected by bootstrap on 5 calculation replicates for each dataset. The predictors were analyzed with a JRR Leave One Out procedure for measuring the contribution of the individual variables to models accuracy.

#### BIBLIOGRAFIA CITATA E DI RIFERIMENTO

- Bovio G., Marchi E., 2010 – *Selvicoltura: incendi, pascolo*. L'Italia Forestale e Montana, 65 (2):113-119.
- Dormann C.F., Elith J., Bacher S., Buchmann C., Carl G., Carré G., Marquéz J.R.G., Gruber B., Lafourcade B., Leitão P.J., Münkemüller T., McClean C., Osborne P.E., Reineking B., Schröder B., Skidmore A.K., Zurell D., Lautenbach S. 2013 – *Collinearity: a review of methods to deal with it and a simulation study evaluating their performance*. Ecography, 36 (1): 27-46.  
<http://dx.doi.org/10.1111/j.1600-0587.2012.07348.x>
- Fielding A., Bell J., 1997 – *A review of methods for the assessment of prediction errors in conservation presence/absence models*. Environmental Conservation, 24: 38- 49.  
<http://dx.doi.org/10.1017/S0376892997000088>
- Franklin J., McCullough P., Gray C., 2000 – *Terrain variables for predictive mapping of vegetation communities in Southern California*. In: Terrain Analysis: Principals and Applications. Wilson J. and Gallant J. (Eds.). John Wiley and Sons, New York, pp. 381.
- Jaynes E.T., 1990 – *Probability Theory as Logic*. In: Maximum-Entropy and Bayesian Methods. Fougère P.F. (ed.), Kluwer, Dordrecht, pp. 1-16.  
[http://dx.doi.org/10.1007/978-94-009-0683-9\\_1](http://dx.doi.org/10.1007/978-94-009-0683-9_1)
- Hijmans R.J., Cameron S.E., Parra J.L., Jones P.G., Jarvis A., 2005 – *Very high resolution interpolated climate surfaces for global land areas*. International Journal of Climatology, 25: 1965-1978.  
<http://dx.doi.org/10.1002/joc.1276>
- Lobo J.M., Jiménez-Valverde A., Real R., 2008 – *AUC: a misleading measure of the performance of predictive distribution models*. Global Ecology and Biogeography, 17 (2): 141-151.  
<http://dx.doi.org/10.1111/j.14668238.2007.00358.x>
- Manel S., Williams H.C., Ormerod S.J., 2001 – *Evaluating presence-absence models in ecology: the need to account for prevalence*. Journal of Applied Ecology, 38: 921-931.  
<http://dx.doi.org/10.1046/j.1365-2664.2001.00647.x>
- McCune B., Keon D., 2002 – *Equations for potential annual direct incident radiation and heat load*. Journal of Vegetation Science, 13: 603-606.  
<http://dx.doi.org/10.1111/j.1654-1103.2002.tb02087.x>
- Parisien M., Moritz M., 2009 – *Environmental controls on the distribution of wildfire at multiple spatial scales*. Ecological Monographs, 79: 127-154.  
<http://dx.doi.org/10.1890/07-1289.1>
- Parisien M., Snetsinger S., Greenberg J.A., Nelson C.R., Schoennagel T., Dobrowski S.Z., Moritz M.A., 2012 – *Spatial variability in wildfire probability across the western United States*. International Journal of Wildland Fire, 21(4): 313-327.  
<http://dx.doi.org/10.1071/WF11044>
- Phillips S., Anderson R., Schapire R., 2006 – *Maximum entropy modeling of species geographic distributions*. Ecological Modelling, 190: 231-259.  
<http://dx.doi.org/10.1016/j.ecolmodel.2005.03.026>
- Quenouille M.H. 1949 – *Approximate tests of correlation in time-series*. J. R. statist. Soc. B., 11: 68-84.