Multi-Scale Robust Modelling of Landslide Susceptibility: Regional Rapid Assessment and Catchment Robust Fuzzy Ensemble

Claudio Bosco, Daniele de Rigo, Tom Dijkstra,
Graham Sander, Janusz Wasowski
[c.bosco@lboro.ac.uk]

Loughborough University
Loughborough – UK

www.lboro.ac.uk

Why landslides?

- landslides can cause human injury, loss of life and economic damages, destroy construction works and cultural and natural heritage.

- landslides occur in many different geological and environmental conditions across Europe.

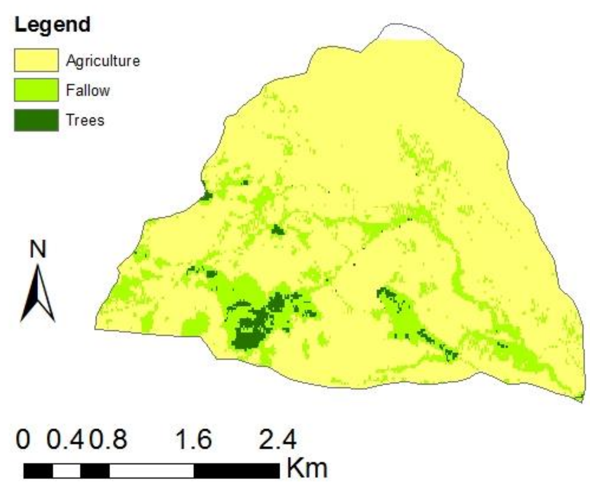
- Mapping or delineating areas prone to landslides is essential for land-use activities and management decision making. Unfortunately one of the main challenges in modelling landslides is related to the assessment of their spatial probability of occurrence.

Complexity:

Landslides are a complex phenomenon affected by many different Factors:

- **climate**
- **topography**
- **lithology** and
- **land cover** (in particular **forest resources, natural vegetation** and **agriculture**)

The Study Area



The study area (Rocchetta Sant'Antonio, Italy).
Google Earth, 2013 Google.


Landslide events are associated with a trigger such as an earthquake, a large storm, a rapid snowmelt, or a volcanic eruption. A landslide event may include a single landslide or many thousands and can be quantified by the frequency–area distribution of the triggered landslides.

Within the study area, precipitation is the main triggering factor for landslide occurrence.

The applied methodology

Estimating landslide spatial probability may be supported by many different analytical approaches: heuristic, deterministic and statistical. Despite the many different approaches, landslide susceptibility assessment still remains a challenge.

In order to improve the spatial prediction of landslides, a combined total of five different deterministic and statistical models have been applied and a new method based on an ensemble approach has been used for aggregating the modelling results.

The models

Each susceptibility zonation has been obtained by applying heterogeneous techniques as:

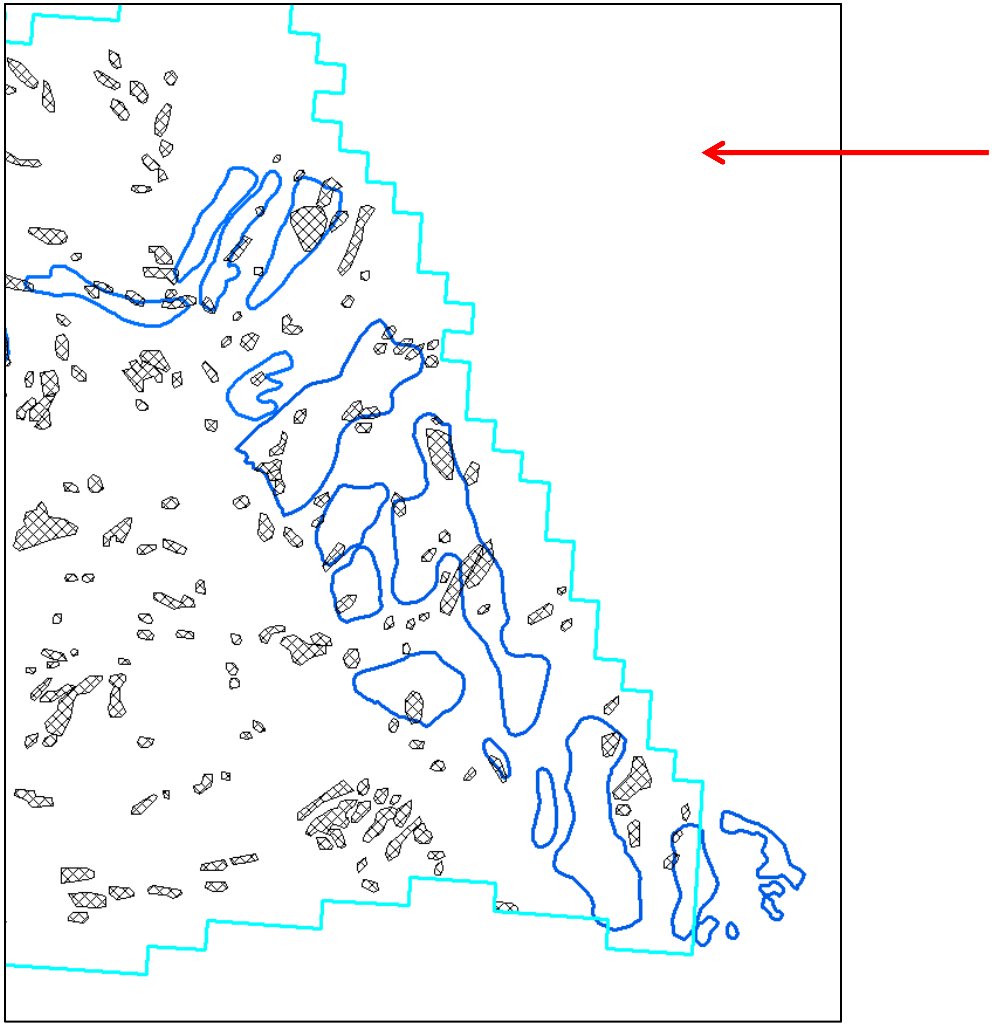
- logistic regression (LR),

- relative distance similarity (RDS),

- artificial neural network (ANN)

- two different landslide susceptibility techniques based on the infinite slope stability model (SINMAP and TransSlide).

The datasets



A database containing more than 200 landslides

A DEM with a resolution of 5 meters

A land cover map

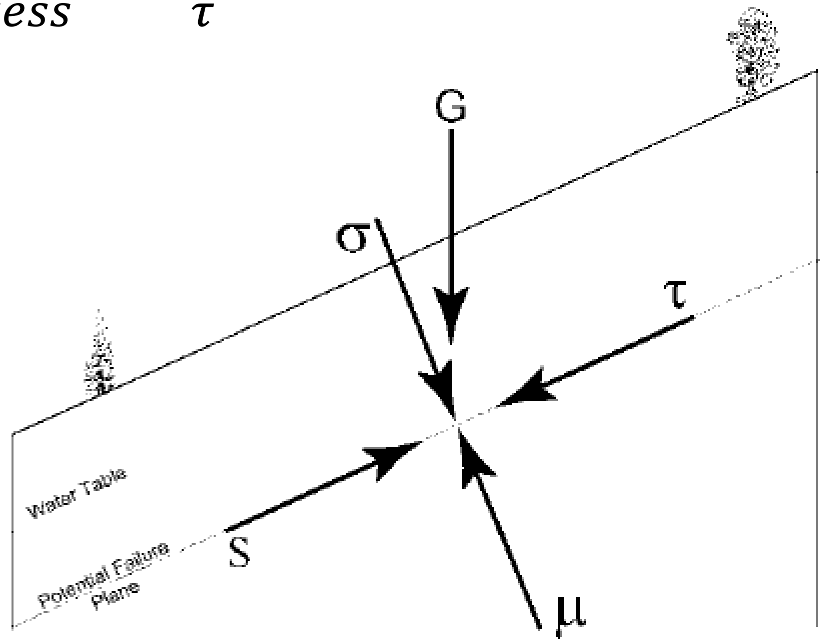
A lithological map

Deterministic models

Landslides are generally induced when the shear stress on the slope material exceeds the material's shear strength.

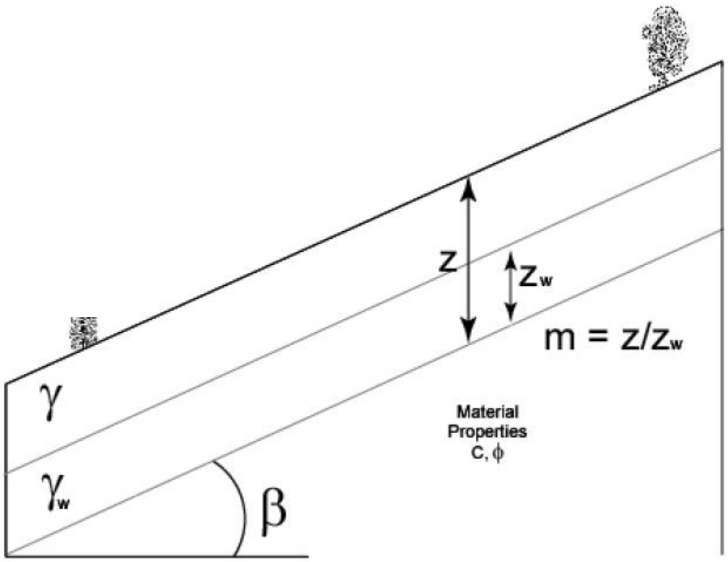
$$FS = \frac{resisting\ forces}{driving\ forces} = \frac{shear\ strenght}{shear\ stress} = \frac{S}{\tau}$$

Where FS is the factor of safety. If FS reaches values lower than unity, the slope will fail.



Definition diagram in the infinite slope stability model

$$FS = \frac{C + (\gamma - m\gamma_w) z \cos\beta \cos\beta \tan\phi}{\gamma z \sin\beta \cos\beta}$$

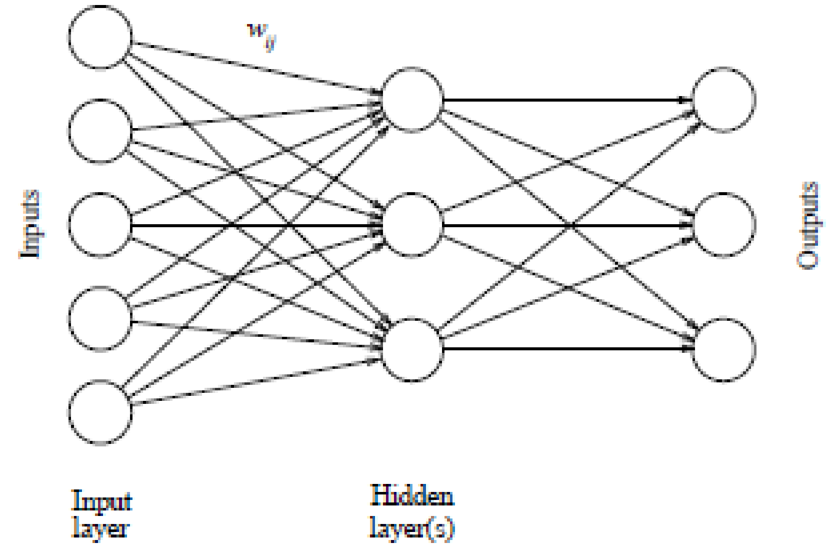


where C is the soil cohesion (kPa), γ is the soil unit weight (kN/m^3), γ_w is the water unit weight (kN/m^3), z is the vertical soil depth (m), z_w is the vertical water depth (m), β is slope angle (°) and ϕ is the internal friction angle (°).

Statistical techniques

Artificial Neural Network

A generic single layer feed-forward neural network



Logistic Regression

where β_0 is the intercept and β_j are the coefficients relating covariates x_j (j=1,2, …,k) and π is the probability of landslide occurrence

$$P(Y=1) = \pi = \frac{1}{\left[1 + e^{\left(\beta_0 + \sum_{j-1}^{k} \beta_j\, x_j\right)}\right]}$$

Relative Distance Similarity (RDS)

The RDS defines the relative distance between two values C_1 and C_2 of a given nonnegative covariate. The relative distance is a dimensionless number between 0 (maximum dissimilarity) and 1 (maximum similarity) and is simply the ratio between the minimum and the maximum value of the pair $\{C_1, C_2\}$:

$$\frac{min(C_1^j, C_2^j)}{max(C_1^j, C_2^j)}$$

The RDS index of a given multi-dimensional point *c* with respect to a set *A* of reference points involves the relative distance among the pairs $\{C^j_c, C^j_\alpha\}$ for each $\alpha \in A$ and each dimension *j* of the N^C covariates

The Ensemble model

The ensemble approach is a reproducible D-TM applied to the results of the array of models and is based on relative-distance similarity (RDS).

In the landslide application, the indices RDS_c^L and RDS_c^S express the possibility [0,1] for c to respectively belong to L (instable areas) or S (stable areas)

$$\begin{cases} RDS_c^L = \max\limits_{\alpha \in S^L} \left(\mathop{\Omega}_{j=1}^{N^c} \left(\dfrac{\max\left(\min\left(C_c^j, C_\alpha^j\right), \delta C^j\right)}{\max\left(\max\left(C_c^j, C_\alpha^j\right), \delta C^j\right)} \right) \right) \\ RDS_c^S = \max\limits_{\alpha \in S^S} \left(\mathop{\Omega}_{j=1}^{N^c} \left(\dfrac{\max\left(\min\left(C_c^j, C_\alpha^j\right), \delta C^j\right)}{\max\left(\max\left(C_c^j, C_\alpha^j\right), \delta C^j\right)} \right) \right) \end{cases}$$

Semantic array programming paradigm:

- accurate vector-based mathematical description of the model can simplify complex algorithm prototyping while moving mathematical reasoning directly into the source code (reproducible way)

- modularizing sub-models and autonomous tasks with a strong effort toward their *most concise* generalization and reusability in other contexts;

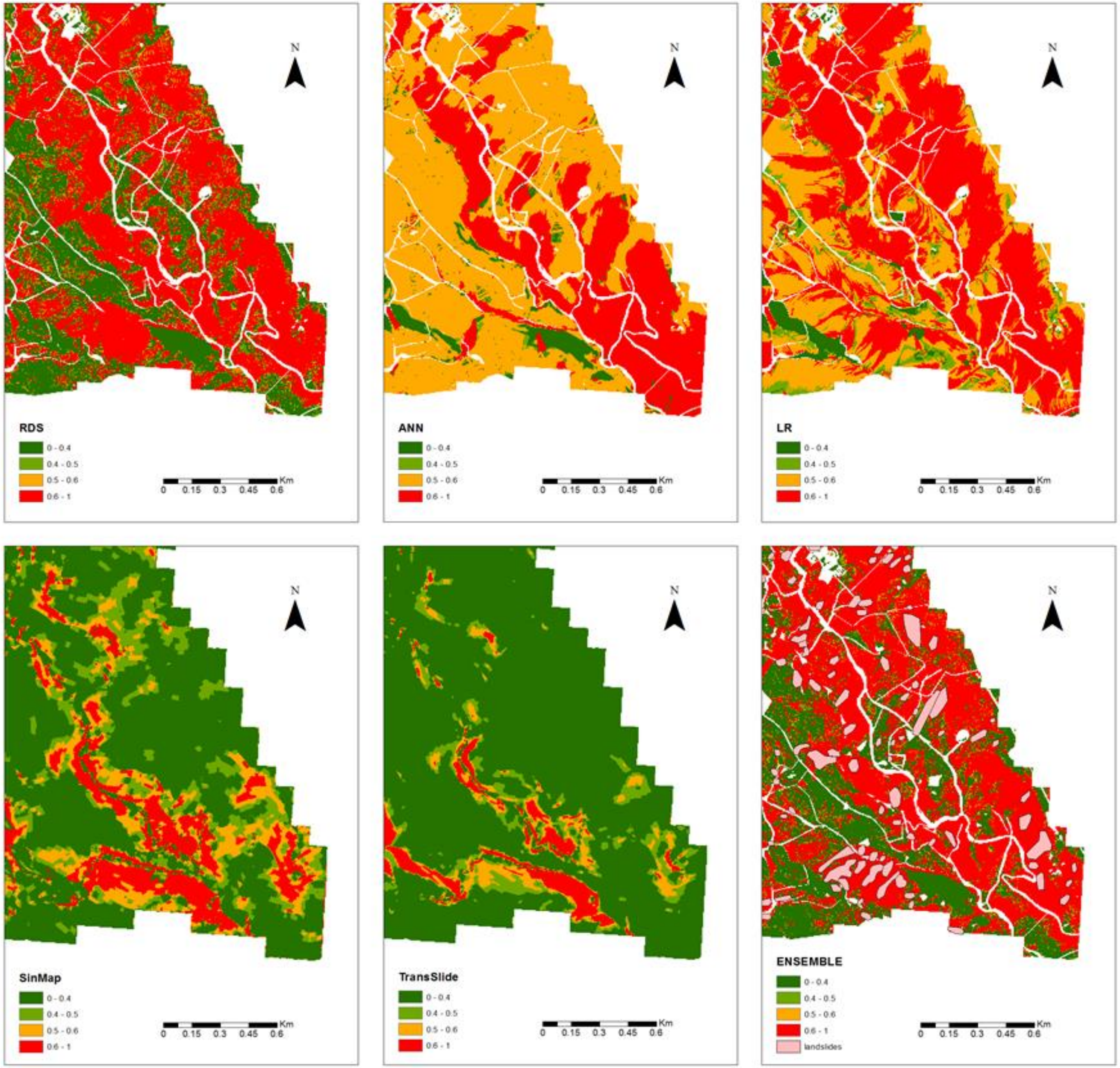
- *semantically* constraining – with terse array-based constraints – the information entered in and returned by each module instead of relying on external assumptions.

The modelling results

	RDS	ANN	LR	SINMAP	TransSl.	median	ENSEMB.
MAE	0.003	0.44	0.37	0.45	0.51	0.35	0.001
MAE U.	0.002	0.42	0.36	0.61	0.68	0.45	0
MAE S.	0.003	0.45	0.38	0.3	0.34	0.25	0.001
RMSE	0.02	0.47	0.43	0.54	0.58	0.4	0.019
RMSE U.	0.01	0.45	0.4	0.65	0.7	0.47	0
RMSE S.	0.03	0.48	0.46	0.4	0.42	0.32	0.026

Mean Absolute Error (MAE) and Root Mean Square Error (RMSE)

- the ensemble and RDS give the lowest errors.
- better performance of the statistical methods when compared with deterministic approaches.
- a straightforward unsupervised ensemble might prove useful even where no additional information is available (black box output data).


N
RDS
0 - 0.4
0.4 - 0.5
0.5 - 0.6
0.6 - 1
0 0.15 0.3 0.45 0.6 Km
N
ANN
0 - 0.4
0.4 - 0.5
0.5 - 0.6
0.6 - 1
0 0.15 0.3 0.45 0.6 Km
N
LR
0 - 0.4
0.4 - 0.5
0.5 - 0.6
0.6 - 1
0 0.15 0.3 0.45 0.6 Km
N
SinMap
0 - 0.4
0.4 - 0.5
0.5 - 0.6
0.6 - 1
0 0.15 0.3 0.45 0.6 Km
N
TransSlide
0 - 0.4
0.4 - 0.5
0.5 - 0.6
0.6 - 1
0 0.15 0.3 0.45 0.6 Km
N
ENSEMBLE
0 - 0.4
0.4 - 0.5
0.5 - 0.6
0.6 - 1
landslides
0 0.15 0.3 0.45 0.6 Km

The application of an ensemble approach, especially in data-poor regions, could potentially reduce the uncertainty and mitigate local poor performance associated with individual models, by excluding outlier estimations.

Limits of the presented approach

Because the quality of spatial landslides forecast is largely dependent on the quality of the available datasets, the bad performance of the physically based models could be linked with the lack of data.

The high performance showed by the RDS approach could be linked with the criterion used for the selection of the training and testing set of data. The possible presence of bias in using a similar technique for selecting the data and calculating the landslide susceptibility need to be further investigated.

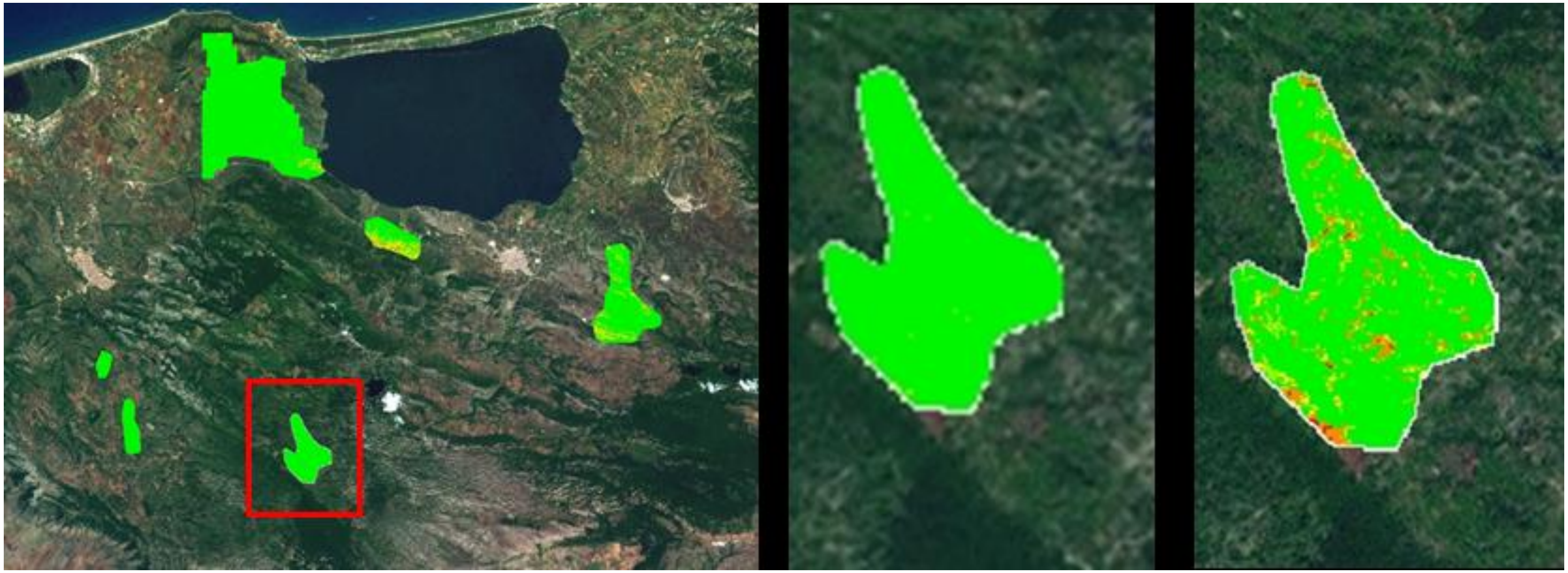
Although these preliminary results are promising, further research is required before this method can be used to communicate the findings with relevant authorities.

Post Fire Regional Scale Rapid Assessment

- wildfires have a major impact especially in southern Europe (Mediterranean region).

- besides their direct environmental impact, wildfires can result in future secondary effects such as shallow landslides and debris flows.

- an area of approximately 600 km2 was analysed calculating the changes in pre and post-fire landslide susceptibility. Six large fires were selected from the European Forest Fires Information System (EFFIS).



Pre and Post-re shallow landslide susceptibility in Northern Puglia (Italy).
Ground layer: Google Earth, c 2013 Google.

- almost 10% of the burnt area changes from stable to unstable conditions

- The application of the SINMAP model in post-fire landslide susceptibility analysis has to be considered as a first attempt for applying these techniques at a regional scale.

- Future research will be carried out to extend the multi-model ensemble architecture from catchment to regional scale, overcoming the possible unavailability of observations on stable and unstable areas.

Thank you