

Supplementary material

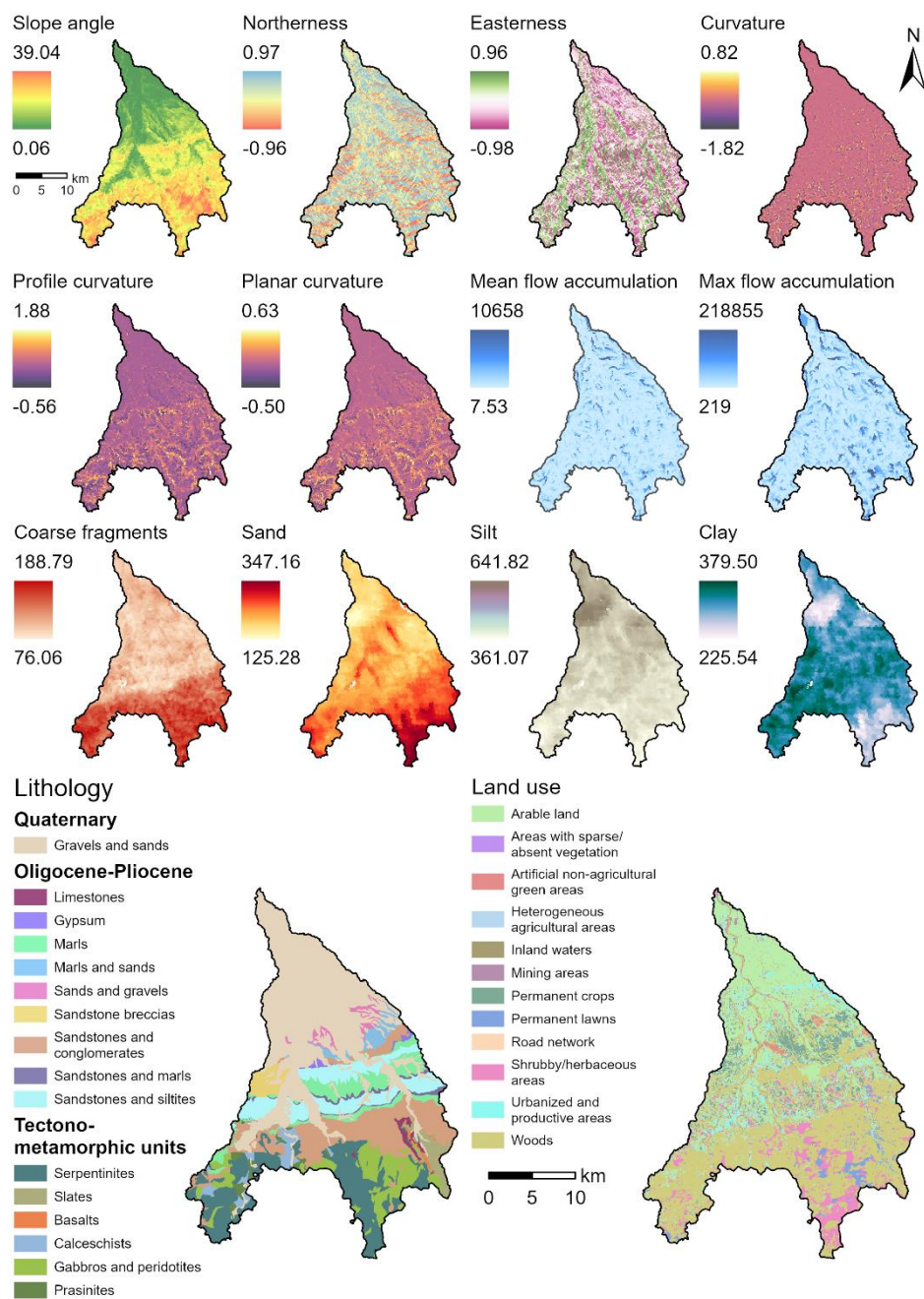


Figure S1. Maps of the areal distribution of the variables considered in the analyses.

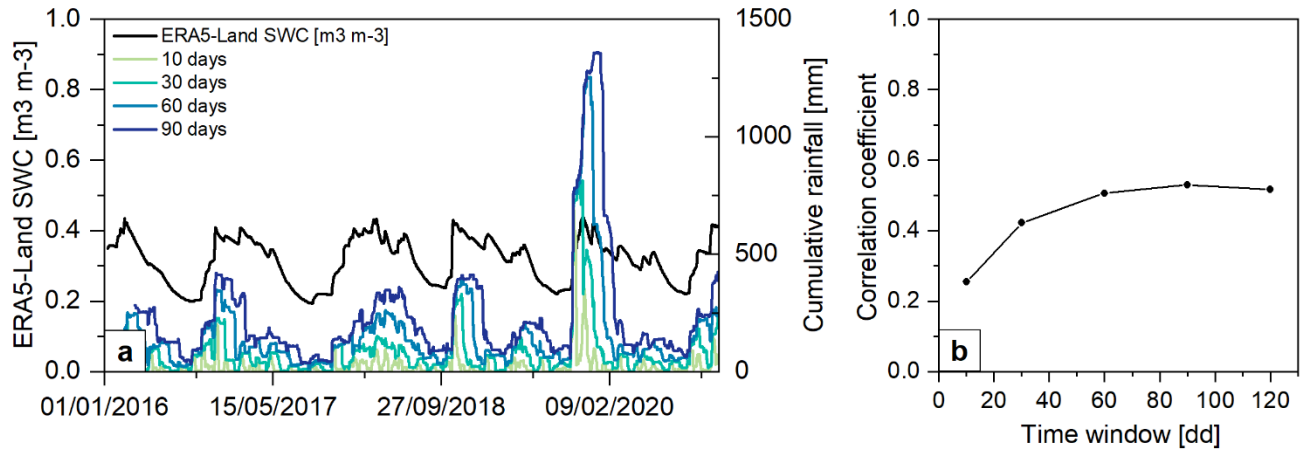
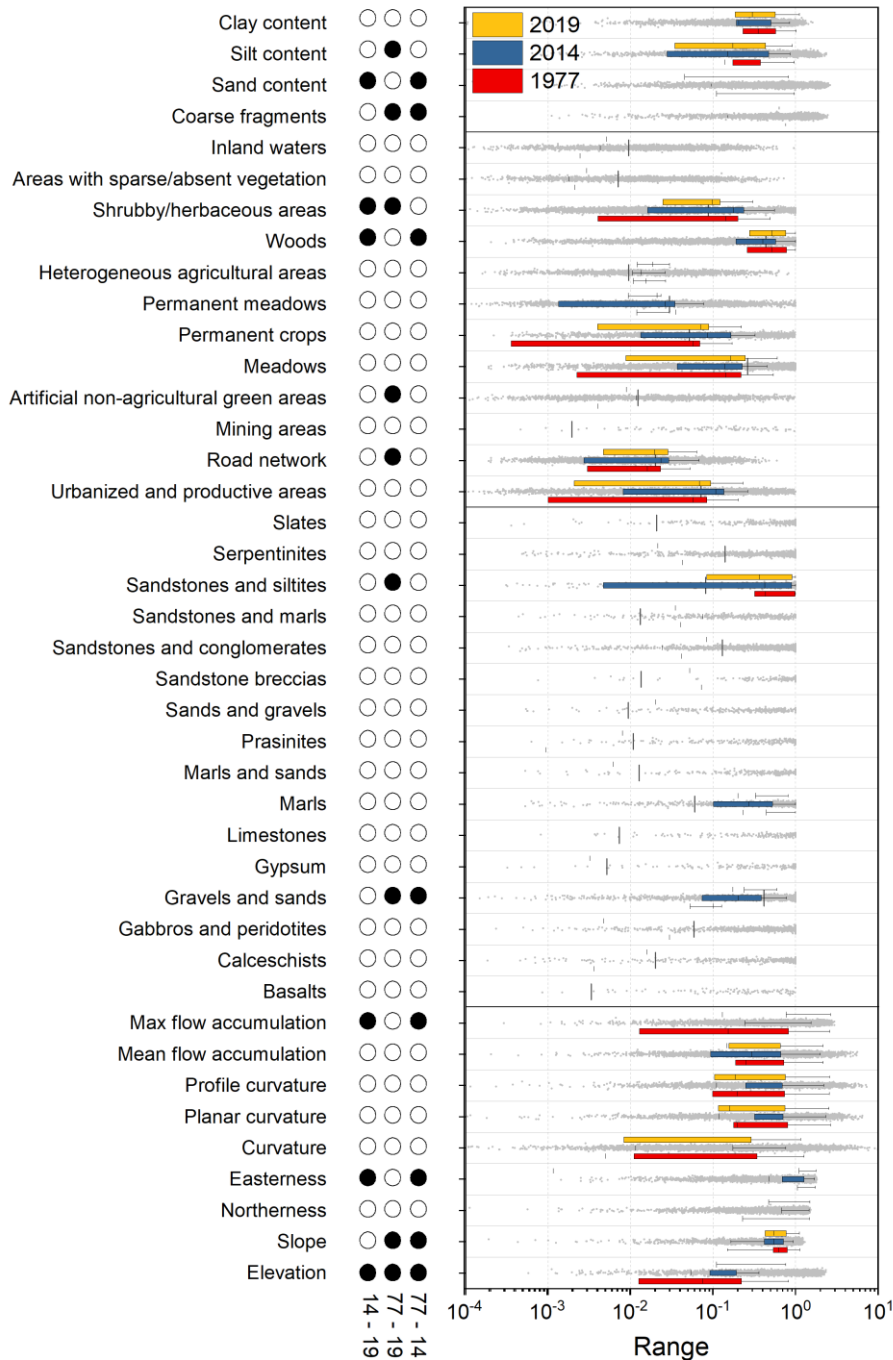
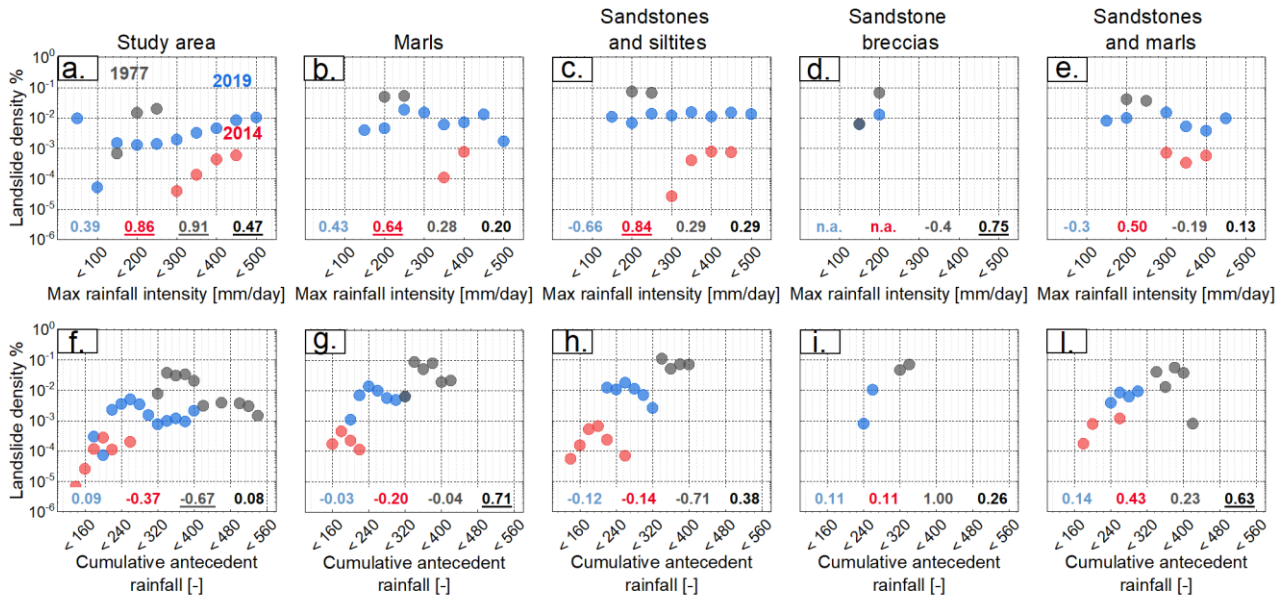


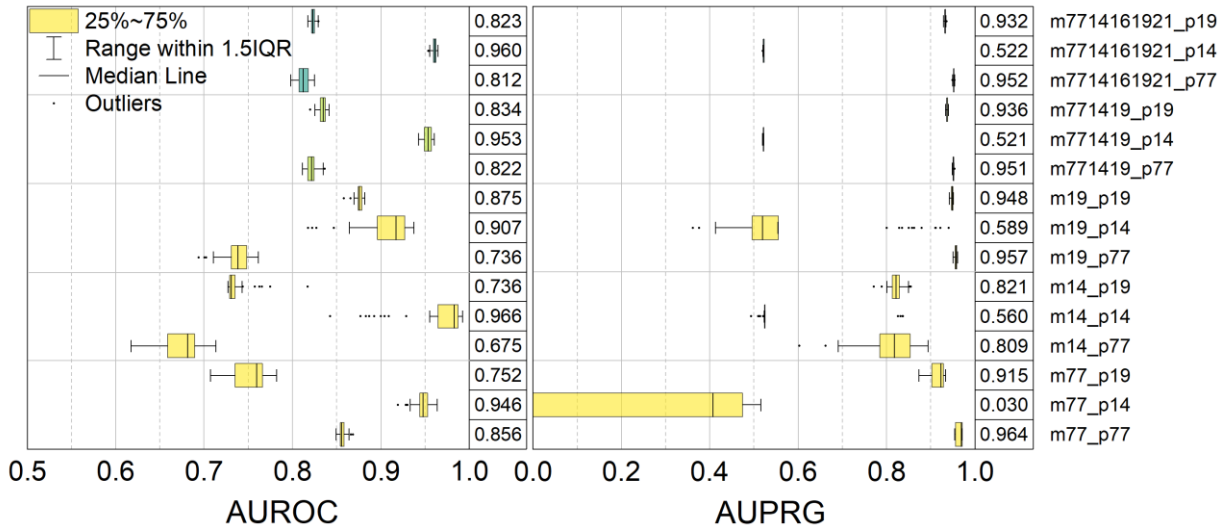
Figure S2. ERA5-Land estimated volumetric soil water content and cumulative rainfall time series (a) and correlation coefficient between the estimated volumetric soil water content and antecedent cumulative rainfall over different time windows (b).



10 **Figure S3. Distribution of the controlling variables in the slope units. Boxplots represent the normalized values distributions of slope units destabilized in each landslide event, while grey dots represent the entire values population. Dots above the variable names represent the significance of the population differences, according to the Dunn's test. A black dot means that the distributions are not statistically different at a significance level of 0.05.**



15 **Figure S4.** Scatterplots representing landslide density in each rainfall class, for the entire study area and for different lithologies. Spearman's rank order correlation coefficient between landslide density and rainfall classes is reported in each plot. Underlined values are statistically correlated at the 0.05 level.



20 **Figure S5.** Boxplots of the AUROC and AUPRG performance of all the tested models for the 50 repetitions, tested both on themselves and to forecast or hindcast other events. AUROC values can range between 0.5 and 1; the closest to 1, the best the performance of the model. More in details, if AUROC is equal to 0.5 there is no discrimination; if the AUROC value is between 0.7 and 0.8, the model performance is classified as acceptable; if the AUROC value is between 0.8 and 0.9 the model performance is excellent, while if the AUROC value is greater than 0.9 the model performance is outstanding (Hosmer and Lemeshow, 2000). Labels refer to the events used to train the model (m*), and to the events used to test the model (p*). Numbers next to the labels indicate the mean value of the distribution. All the distributions are statistically different at the 0.05 significance level, according to the Wilcoxon–Mann–Whitney test.

25