Grasslands cover almost one third of the world’s terrestrial surface. In Alpine environments grassland vegetation fulfills various key environmental purposes such as water reservoir, slope stabilizer and carbon sink or fodder for livestock. At the same time Alpine regions are more affected by climatic changes than other geographic zones, potentially resulting in earlier green-up phases or an elevated exposure to drought events, hampering the growth and vitality of grassland vegetation. The scope of this study is to build an algorithm capable of biomass estimation using Support Vector Machine approach on hyperspectral and Synthetic Aperture Radar (SAR) data. To that purpose, field campaigns were carried out during 2017 and 2019 in Val Mazia (South Tyrol, Italy), where hyperspectral spectroradiometer samples were collected, as well as leaf area index (LAI), soil moisture, and above ground biomass measurements. Copernicus Sentinel-1 IW SAR backscattering data were used to complete the dataset.

The spectroradiometer was used to simulate the hyperspectral data of the Italian Space Agency (ASI)’s PRISMA mission, launched on 22 March 2019. Since the number of bands is larger than the number of samples, a prediction approach based on machine learning risks to model noise. The following two solutions were tested and compared: (i) the number of bands was reduced by resampling the data to match specifications of Copernicus Sentinel-2 Multispectral Instrument (MSI), and (ii) the data was simulated using the PROSPECT model, increasing the sample size.

In the first case correlation $R^2$ of 0.37 was found. Discrepancies were observed for high biomass values, which could be explained by the small number of samples available shortly before harvest. To mitigate this effect, data were simulated for high biomass based on field average values and standard deviation within each date. $R^2$ increased to 0.71 in this case, confirming the above mentioned hypothesis regarding the dataset representativeness.

In the case of PROSPECT model, the parameters were found by iterating each one within ranges defined in the bibliography, until the spectral signatures matched the field observations. The resulting parameters were the input for data simulation. A genetic algorithm feature selection was run to reduce the number of features, discarding those with little or redundant information.
followed by an SVR model applied to the most sensitive bands resulting in an $R^2$ of 0.53. These initial results will be used as a basis for future investigations to improve the prediction model, for example by extending the dataset with new field campaigns, including more simulated data at biomass peak, as made with Sentinel-2 resampled dataset, or by adding further input variables, such as leaf area index. Furthermore, the procedure will be performed for fresh biomass and water content estimations.

The results obtained pave the way for future implementation of the tested algorithms on PRISMA hyperspectral and COSMO-SkyMed X-band SAR data in the future.

This research is part of the ongoing project ‘Development of algorithms for estimation and monitoring of hydrological parameters from satellite and drone’, funded by ASI under grant agreement n.2018-37-HH.0.