

Read me file:

This dataset contains six (6) files containing the environmental data and associated class assignment, training site locations for the datasets used in the paper: Impact of ecological redundancy on the performance of machine learning classifiers in vegetation mapping, Macintyre, Paul, Van Niekerk, Adriaan, Dobrowolski, Mark, Tsakalos, James, Mucina, Ladislav

For each table the following predictors are used:

Type of Predictor	Source	Type of Predictor	Source
Climate		Soil (Woodland)	
Bioclim 1–19		Available Water Content	
Average Precipitation (Prec 1–12)	BioClim dataset (Hijmans <i>et al.</i> 2005)	Bulk Density (Whole Earth)	
Maximum Temperature (Tmax 1–12)		Percentage Clay	
Minimum Temperature (Tmin 1–12)		Depth to regolith	Australian Soil Grid (Viscarra Rossel <i>et al.</i> 2015)
Mean Temperature (Tmean 1-12)		Depth of Soil	
Soil (Kwongan vegetation)		ECEC	
Conductivity		pHc	
Bulk Density		Percentage Sand	
Percentage Clay		Percentage Silt	
ECEC		Soil Organic Carbon	
Effective CaCO ₃		Total Nitrogen	
ESP		Total Phosphorus	
Exchangeable Acidity		Topographic	
Exchangeable Aluminium		Elevation	
Exchangeable Calcium		Slope (Degrees)	Derived from SRTM DEM
Exchangeable Potassium		Aspect	
Exchangeable Magnesium		Plan Curvature	
Exchangeable Sodium		Profile Curvature	
Exchangeable Aluminium (KCl)	Ground sampled	Convergence Index	
Exchangeable Hydrogen (KCl)		Closed Depressions	
Organic Carbon		Catchment Area	
pH (CaCl ₂)		Wetness Index	
pH (H ₂ O)		LS-Factor	
Percentage Sand		Channel Network Base Level	
Percentage Silt		Vertical Distance to Channel Network	
Total Carbon		Valley Depth	
Total Nitrogen		Relative Slope Position	
Total Phosphorus		Channel Network	
Water repellency			