Generating high-resolution land use and land cover maps for the greater Mariño watershed in 2019 with machine learning (Supplementary materials)

Table SM 1: Satellite images references.

Satellite product	**Tiles**	**Date**	**Specific image references**
TanDEM-X	S14W073 S14W074 S15W073 S15W074	-	TDM1_DEM__04_S14W073_DEM.tif TDM1_DEM__04_S14W074_DEM.tif TDM1_DEM__04_S15W073_DEM.tif TDM1_DEM__04_S15W074_DEM.tif
Pléiades	W074S14 W073S14	07/10/2019	2020-112_Vallet_SO20225702-1-01_DS_PHR1B_201910071516024_FR1_PX_W074S14_1205_06640.zip 2020-112_Vallet_SO20225702-2-01_DS_PHR1B_201910071515444_FR1_PX_W073S14_0207_04978.zip 2020-112_Vallet_SO20225702-3-01_DS_PHR1B_201910071515139_FR1_PX_W073S14_0407_04222.zip
Sentinel-2	T18LYK (South) T18LYL (North)	01/01/2018 – 30/10/2019	

Table SM 2: Summary of the field database.

Land-use land-cover description				Reference database			Validation database		
Level 1	**Level 2**	**Level 3**	**Code**	**Number of polygons (percentage of total number of polygons in the dataset)**	**Total surface in km^2 (percentage of total surface in dataset)**	**Average size of polygons (m^2)**	**Number of polygons (percentage of total number of polygons in the dataset)**	**Total surface in km^2 (percentage of total surface in dataset)**	**Average size of polygons (m^2)**
Agricultural areas	Agricultural areas	Sugar cane	1	23 (1.4%)	0.03 (0.2%)	1 582	23 (1.5%)	0.03 (0.2%)	1 582
		Pasture, fallow and feed	2	245 (14.4%)	0.69 (4.1%)	2 839	237 (15.5%)	0.69 (4.2%)	2 912
		Crop and alfalfa	3	353 (20.8%)	0.67 (4.0%)	1 906	322 (21%)	0.64 (4.0%)	2 023
		Fruit crop	4	52 (3.1%)	0.09 (0.5%)	1 737	52 (3.4%)	0.09 (0.5%)	1 737
Natural spaces and forest plantations	Woodlands	Polylepis mountain forest	5	23 (1.4%)	0.16 (0.9%)	7 010	23 (1.5%)	0.16 (0.9%)	7 010
		Podocarpus glomeratus mountain forest	6	26 (1.5%)	0.86 (5.1%)	33 186	26 (1.7%)	0.86 (5.3%)	33 186
		Dry forest	7	56 (3.3%)	1.26 (7.5%)	22 651	49 (3.2%)	0.99 (6.1%)	20 333
		Other tree vegetation	8	320 (18.8%)	1.93 (11.5%)	6 035	233 (15.2%)	1.83 (11.4%)	7 894
		Pine plantation	9	40 (2.4%)	0.17 (1.0%)	4 311	32 (2.1%)	0.12 (0.7%)	3 955
		Eucalyptus plantation	10	82 (4.8%)	0.99 (5.9%)	12 174	73 (4.8%)	0.98 (6.0%)	13 509
	Shrubs and natural grasslands	Mixed shrubland	11	83 (4.9%)	0.76 (4.5%)	9 202	80 (5.2%)	0.76 (4.7%)	9 510
		Dry shrubland and semi-arid steppe	12	46 (2.7%)	1.13 (6.7%)	24 574	45 (2.9%)	1.12 (6.9%)	25 105
		High-elevation grassland	13	53 (3.1%)	4.51 (26.8%)	85 018	51 (3.3%)	4.41 (27.7%)	88 085
	Rocks and natural bare soils	Rock and natural bare soil	14	44 (2.6%)	1.09 (6.5%)	24 890	44 (2.9%)	1.09 (6.7%)	24 890
		Beach and riverine rock	15	14 (0.8%)	0.03 (0.1%)	2 183	14 (0.9%)	0.03 (0.1%)	2 183
Water and glacier	Glacier	Glacier	16	6 (0.4%)	0.26 (1.5%)	44 271	6 (0.4%)	0.26 (1.6%)	44 271
	Water	Wetland	17	58 (3.4%)	0.68 (4.0%)	11 730	58 (3.8%)	0.67 (4.1%)	11 670
		Lake	18	35 (2.1%)	0.57 (3.4%)	16 559	26 (1.7%)	0.49 (3.0%)	18 851
		River network	19	19 (1.1%)	0.02 (0.1%)	1 176	19 (1.2%)	0.02 (0.1%)	1 176
Impervious areas	Impervious areas	Built-up area	20	120 (7.1%)	0.78 (4.7%)	6 571	118 (7.7%)	0.78 (4.8%)	6 645
			Total	1,698 (100%)	**16.75 (100%)**	**9 866**	1,531 (100%)	**16.18 (100%)**	**10 579**

Table SM 3: Data source and criteria used for post-processing MORINGA output map.

LULC class	Code	Data source for GIS cross-checking	Criteria	Reference
Sugar cane	1	Agricultural areas in 2018 (national map)	Intersect the agricultural surface layer and exclusively located near the Pachachaca River. Easily identifiable on the Pléiades image in false (bright red) and true (specific texture) colors.	(Livia Alejandro et al., 2021)
Pasture, fallow and feed	2	Agricultural areas in 2018 (national map)	Intersect the agricultural surface layer.	(Livia Alejandro et al., 2021)
Crop and alfalfa	3	Agricultural areas in 2018 (national map)	Intersect the agricultural surface layer.	(Livia Alejandro et al., 2021)
Fruit crop	4	Agricultural areas in 2018 (national map)	Intersect the agricultural surface layer. Easily identifiable on the Pléiades image (in true color): rows of trees with distinct texture	(Livia Alejandro et al., 2021)
Polylepis mountain forest	5		Located above 3500m. Distinctive texture and color (dark red, dense tree canopy) on the Pléiades image (in false color).	
Podocarpus glomeratus mountain forest	6	Vegetation cover (national map)	Intersect the mesoandean coniferous relic forest on the vegetation cover map.	(MINAM, 2015)
Dry forest	7		The only tree vegetation located in low altitude, easily identifiable on the Pléiades image (in false color): medium intensity red.	
Other tree vegetation	8	No specific criteria used.		
Pine plantation	9		Located above 3000m. Distinctive texture and color (bright red, dense tree canopy) on the Pléiades image (in false color).	
Eucalyptus plantation	10	Agricultural areas in 2018 (national map)	Intersect the agricultural surface layer . Very distinct from other agricultural LULC categories mapped in the national map of agricultural areas (i.e. it was the only tree vegetation)	(Livia Alejandro et al., 2021)
Mixed shrubland	11	Digital Elevation Model	Located above 2500m.	(Wessel, 2018)
Dry shrubland and semi-arid steppe	12	Digital Elevation Model	Located below 2500m, dry appearance in Pléiades image (in false color): they were less red and less dense than high altitude shrublands.	(Wessel, 2018)
High-elevation grassland	13	Digital Elevation Model	Located above 2500m.	(Wessel, 2018)
Rock and natural bare soil	14	No specific criteria used.		
Beach and riverine rock	15		Located along the river network.	
Glacier	16		Only located at the top of the Ampay mountain.	
Wetland	17		Located above 3500m. Wetlands appear in a bright red color on Pléiades image (in false color).	
Lake	18	Surface water maximum extent (GSWE)	Intersect the maximum water extent layer.	(Pekel et al., 2016)
River network	19	Surface water maximum extent (GSWE)	Intersect the maximum water extent layer.	(Pekel et al., 2016)
Built-up area	20	GHS built-up surface in 2018 (10m resolution)	Intersect the built-up areas cross-checking layer. Buildings easily identifiable on the very high-resolution Pleaides image (in true color).	(Pesaresi and Politis, 2022)
Road	21	OpenStreetMap	Intersect the OpenStreetMap major road layer.	(OpenStreetMap contributors, 2022)

Figure SM 1: Location of changes made during post-treatment.

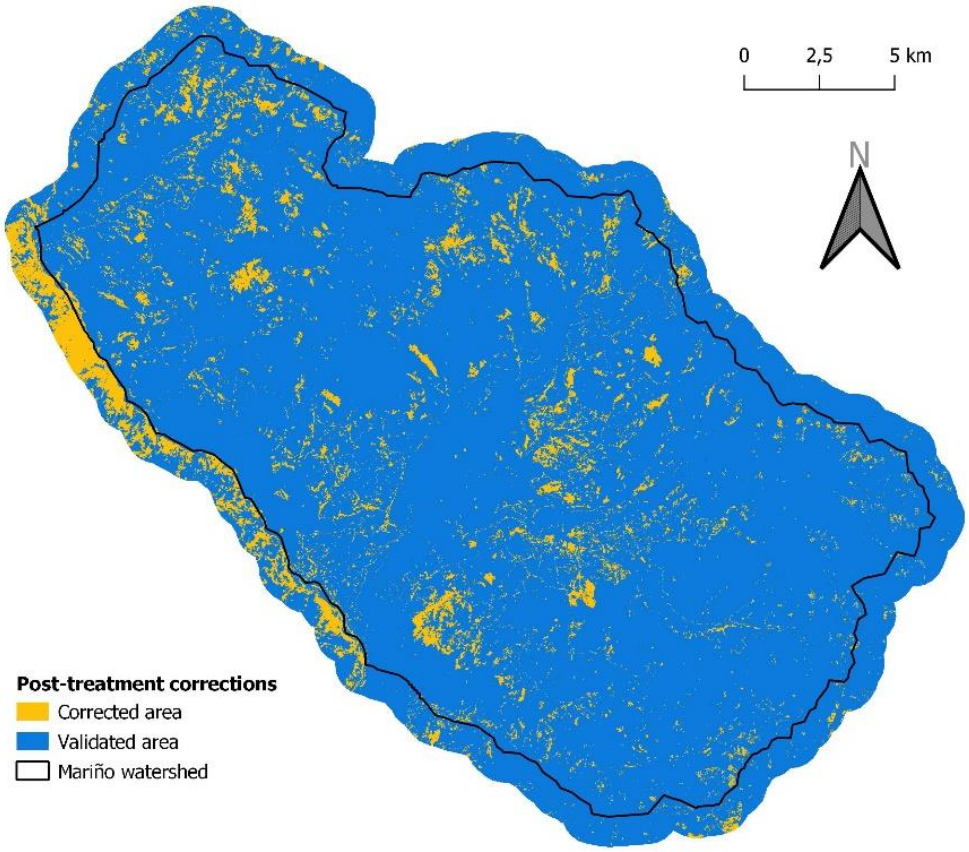

Table SM 4: Corrections applied to the MORINGA classification (rows) during post-treatment (columns). Grey cells on the diagonal correspond to the 91.5% of the study area to which no changes were applied during post-treatment. The percentages represent the proportion of changes made between different LULC classes during post-treatment, relative to the area that each LULC class occupy in the MORINGA output (i.e. they were calculated considering all pixels falling in the different LULC classes).

Moringa output \ Post-treatment	Sugar cane	Pasture, fallow and feed	Crop and alfalfa	Fruit crop	Polylepis mountain forest	Podocarpus glomeratus mountain forest	Dry forest	Other tree vegetation	Pine plantation	Eucalyptus plantation	Mixed shrubland	Dry shrubland and semi-arid steppe	High-elevation grassland	Rock and natural bare soil	Beach and riverine rock	Glacier	Wetland	Lake	River network	Built-up area	**Road**	Total
Sugar cane	100%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	100%
Pasture, fallow and feed	0%	95%	0%	0%	0%	1%	0%	1%	0%	0%	0%	0%	1%	0%	0%	0%	1%	0%	0%	0%	0%	100%
Crop and alfalfa	0%	0%	67%	0%	0%	0%	0%	0%	0%	0%	0%	27%	0%	4%	0%	0%	0%	0%	0%	0%	1%	100%
Fruit crop	0%	0%	1%	52%	0%	0%	0%	47%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	100%
Polylepis mountain forest	0%	0%	0%	0%	52%	0%	0%	29%	0%	1%	2%	0%	1%	13%	0%	0%	1%	0%	0%	0%	0%	100%
Podocarpus glomeratus mountain forest	0%	0%	0%	0%	0%	86%	0%	13%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	100%
Dry forest	0%	0%	0%	0%	0%	0%	96%	1%	0%	0%	2%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	100%
Other tree vegetation	0%	0%	0%	0%	0%	0%	0%	94%	0%	5%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	100%
Pine plantation	0%	0%	0%	0%	0%	1%	0%	30%	58%	0%	9%	0%	0%	2%	0%	0%	0%	0%	0%	0%	0%	100%
Eucalyptus plantation	0%	0%	0%	0%	0%	0%	0%	10%	1%	82%	4%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	100%
Mixed shrubland	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	98%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	100%
Dry shrubland and semi-arid steppe	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	17%	82%	0%	0%	0%	0%	0%	0%	0%	0%	0%	100%
High-elevation grassland	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	99%	1%	0%	0%	0%	0%	0%	0%	0%	100%
Rock and natural bare soil	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	99%	0%	0%	0%	0%	0%	0%	0%	100%
Beach and riverine rock	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	6%	0%	76%	12%	0%	0%	0%	0%	1%	4%	100%
Glacier	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	100%	0%	0%	0%	0%	0%	100%
Wetland	0%	1%	0%	0%	0%	0%	0%	1%	1%	0%	15%	0%	15%	3%	0%	0%	63%	0%	0%	0%	0%	100%
Lake	0%	0%	0%	0%	0%	0%	1%	2%	0%	0%	3%	1%	0%	57%	0%	0%	0%	36%	0%	0%	0%	100%
River network	0%	0%	0%	0%	0%	0%	4%	0%	0%	0%	0%	1%	0%	33%	5%	0%	0%	0%	57%	0%	0%	100%
Built-up area	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	3%	1%	11%	0%	0%	0%	0%	0%	77%	6%	100%
No data	0%	6%	0%	0%	0%	0%	3%	27%	0%	6%	24%	2%	19%	7%	0%	0%	4%	0%	0%	0%	0%	100%

Table SM 5: Description of the information provided for each polygon of the final LULC classification (the data is provided as a geopackage file, with one layer for each level of the nomenclature).

Attribute name	Data type	Description	Example
fid	Numeric	Unique ID	23151
Code1	Numeric	Level 1 land-use land-cover code associated	2
Level1	Text	Level 1 land-use land-cover name associated	Natural spaces and forest plantations
Code2	Numeric	Level 2 land-use land-cover code associated	3
Level2	Text	Level 2 land-use land-cover name associated	Shrubs and natural grasslands
Code3	Numeric	Level 3 land-use land-cover code associated	11
Level3	Text	Level 3 land-use land-cover name associated	Mixed shrubland
area_m2	Numeric	Area of the polygon (m^2)	17.683

Table SM 6: Description of the information recorded for each polygon of the field database (the data is provided as a geopackage file, for level 3 of the nomenclature).

Attribute name	Data type	Description	Example
fid	Numeric	Unique ID	71
Code1	Numeric	Level 1 land-use land-cover code associated	2
Level1	Text	Level 1 land-use land-cover name associated	Natural spaces and forest plantations
Code2	Numeric	Level 2 land-use land-cover code associated	2
Level2	Text	Level 2 land-use land-cover name associated	Woodlands
Code3	Numeric	Level 3 land-use land-cover code associated	10
Level3	Text	Level 3 land-use land-cover name associated	Eucalyptus plantation

References

Livia Alejandro, L., Sánchez Manayay, R., Galiano Uscapi, A., Cajas Ardiles, J., Arévalo Chong, E., Rosas Quispe, E., 2021. Atlas de la superficie agrícola del Perú. MIDAGRI - Dirección General de Estadística, Seguimiento y Evaluación de Políticas, Lima, Peru.

MINAM, 2019. Mapa nacional de ecosistemas del Perú - Memoria descriptiva. Dirección General de Ordenamiento Territorial Ambienta, Lima, Peru.

OpenStreetMap contributors, 2022. Planet dump retrieved from https://planet.osm.org.

Pekel, J.-F., Cottam, A., Gorelick, N., Belward, A.S., 2016. High-resolution mapping of global surface water and its long-term changes. Nature 540, 418–422. https://doi.org/10.1038/nature20584

Pesaresi, M., Politis, P., 2022. GHS-BUILT-S R2022A - GHS built-up surface grid, derived from Sentinel2 composite and Landsat, multitemporal (1975-2030). https://doi.org/10.2905/D07D81B4-7680-4D28-B896-583745C27085

Wessel, B., 2018. TanDEM-X Ground Segment – DEM Products Specification Document (TD-GS-PS-0021 No. Issue 3.2). EOC, DLR, Oberpfaffenhofen, Germany.