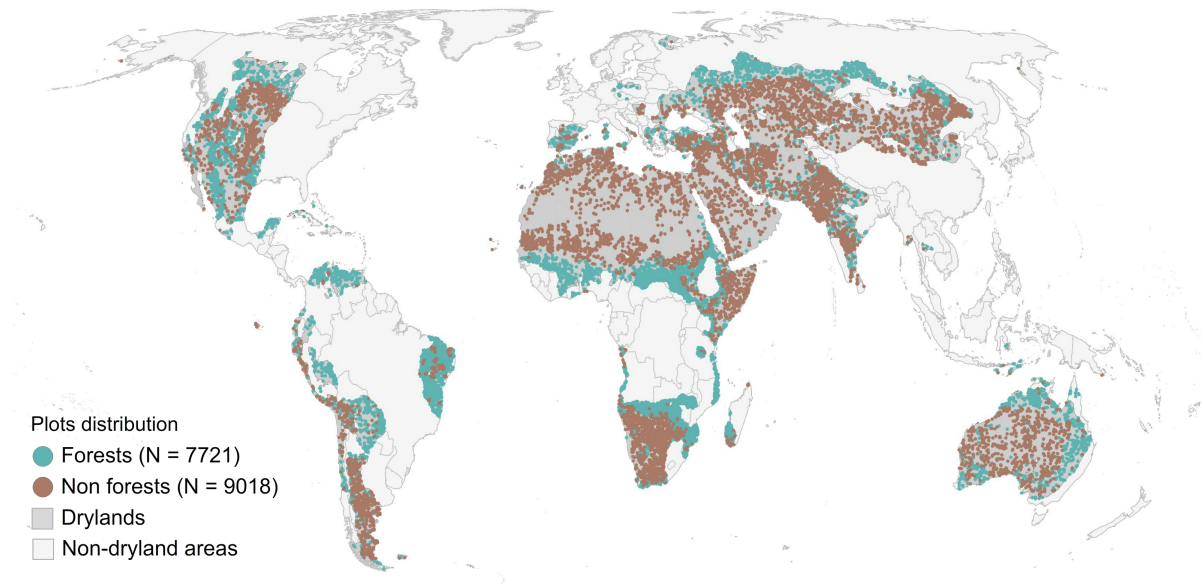**Supplementary information**

Climate legacies drive the distribution and future restoration potential of dryland forests

Emilio Guirado*, Manuel Delgado-Baquerizo, Jaime Martínez-Valderrama, Siham Tabik, Domingo Alcaraz-Segura, Fernando T. Maestre
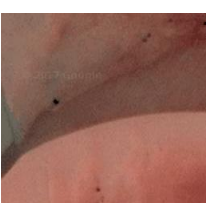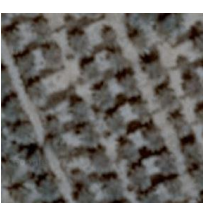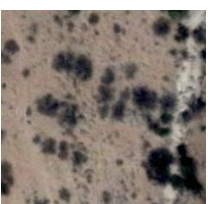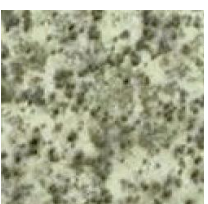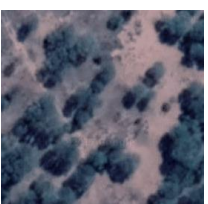
*Corresponding author: emilio.guirado@ua.es

Supplementary Fig. 1: Distribution of the Convolutional Neural Network-classified forest and non-forest plots used in the variation partitioning and random forest analyses.
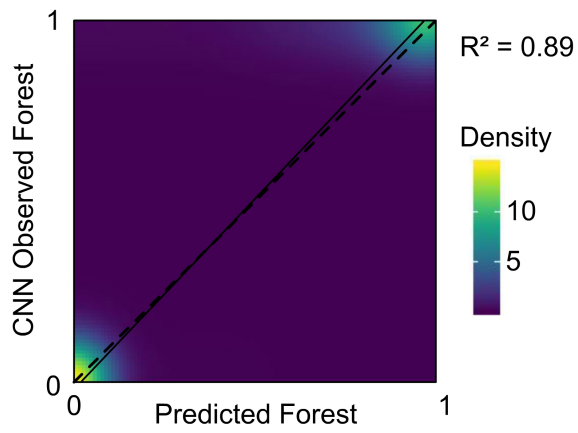
Supplementary Fig. 2: Density plot of observed (using Convolutional Neural Network, CNN) and predicted (using random forest) forest areas. The 1:1 and regression lines are shown in dashed and solid black, respectively (slope = 1.04; R^2 = 0.89), showing that the random forest model is not biased.

Forest class

	Hyper-arid	**Arid**	**Semi-arid**	**Dry sub-humid**
True Positive				
False Positive				
True Negative				
False Negative				

Supplementary Fig. 3: Examples of plots classified with the Convolutional Neural Networks across different aridity levels. A true positive is an outcome in which the model correctly predicts the forest class. Similarly, a true negative is an outcome where the model correctly predicts the non-forest class. A false positive is an outcome where the model incorrectly predicts the non-forest class (it predicts a forest when it should be a non-forested area). A false negative is an outcome where the model incorrectly predicts

the forest class (it predicts a non-forest when it should be a forest). Map data: Google, Maxar Technologies.

Supplementary Table 1. Climatic, soil and other environmental variables used in this study, with their abbreviations and sources (references are found in the main text).

Variable	Data Source	Abbreviation
Annual Mean Temperature	Worldclim[23]	AMT
Mean Diurnal Range (Mean of monthly (max temp - min temp))	Worldclim[23]	MDRM
Isothermality (MDRM/TARBIO) ($\times$100)	Worldclim[23]	IBIO
Temperature Seasonality (SD $\times$100)	Worldclim[23]	TS
Max Temperature of Warmest Month	Worldclim[23]	MTWM
Min Temperature of Coldest Month	Worldclim[23]	MTCM
Temperature Annual Range (MTWM-MTCM)	Worldclim[23]	TARBIO
Mean Temperature of Wettest Quarter	Worldclim[23]	MTWeQ
Mean Temperature of Driest Quarter	Worldclim[23]	MTDQ
Mean Temperature of Warmest Quarter	Worldclim[23]	MTWaQ
Mean Temperature of Coldest Quarter	Worldclim[23]	MTCQ
Annual Precipitation	Worldclim[23]	AP
Precipitation of Wettest Month	Worldclim[23]	PWM
Precipitation of Driest Month	Worldclim[23]	PDM
Precipitation Seasonality (CV)	Worldclim[23]	PSCV
Precipitation of Wettest Quarter	Worldclim[23]	PWeQ
Precipitation of Driest Quarter	Worldclim[23]	PDQ
Precipitation of Warmest Quarter	Worldclim[23]	PWaQ
Precipitation of Coldest Quarter	Worldclim[23]	PCQ
Albedo	MODIS/Terra[46]	Albedo
Evapotranspiration	MODIS/Terra[48]	ET
Soil Slope	ALOS[47]	Slope
Soil Moisture	TerraClimate[44]	Soil moisture
Elevation in metres above sea level	ALOS[47]	Elevation
PH in water soil	Soilgrids[25]	PH
Soil Organic Carbon	Soilgrids[25]	SOC
Soil Texture	Soilgrids[25]	Sands
Soil Nitrogen	Soilgrids[25]	Nit
Piezometric level balance of the aquifer (cumulative groundwater level height over time period)	GRACE[49]	Water Thickness

Supplementary Table 2. Indicators used to assess the accuracy of the forest/non-forest classification. True positives (TP) correspond to images that were correctly classified by the models, false positives (FP) correspond to images that were classified by the models in one class but were from another class, and false negatives (FN) correspond to images of the targeted class that were classified by the models into another class. The Precision metric measures the quality of the machine learning model in classification tasks. The Recall metric indicates how much the machine learning model can identify. The F1-measure value is an equal balance between Precision and Recall, measuring the quality and exhaustiveness of the model in classifying forest/non-forest areas (Precision, Recall, and F1-measure close to 1 is better).

					Forest			Non-Forest			Total
Regions	TP	FP	TN	FN	Precision	Recall	F1	Precision	Recall	F1	F1 mean
Global	234	32	344	28	0.88	0.89	0.89	0.92	0.91	0.92	0.90
1-NothAmerica	26	5	25	0	0.84	1.00	0.91	1.00	0.83	0.91	0.91
2-EastSouthAmerica	16	2	25	4	0.89	0.80	0.84	0.86	0.93	0.89	0.87
3-Europe	15	2	26	1	0.88	0.94	0.91	0.96	0.93	0.95	0.93
4-NorthernAfrica	14	0	37	5	1.00	0.74	0.85	0.88	1.00	0.94	0.89
5-Sahel	19	2	22	4	0.90	0.83	0.86	0.85	0.92	0.88	0.87
6-HornAfrica	27	6	22	2	0.82	0.93	0.87	0.92	0.79	0.85	0.86
7-MiddleEast	12	1	35	1	0.92	0.92	0.92	0.97	0.97	0.97	0.95
8-CentralAsia	20	0	37	1	1.00	0.95	0.98	0.97	1.00	0.99	0.98
9-Australia	22	2	21	1	0.92	0.96	0.94	0.95	0.91	0.93	0.93
10-SouthernAfrica	27	6	22	1	0.82	0.96	0.89	0.96	0.79	0.86	0.87
11-WestSouthAmerica	18	2	34	4	0.90	0.82	0.86	0.89	0.94	0.92	0.89
12-SouthwestAsia	18	4	38	4	0.82	0.82	0.82	0.90	0.90	0.90	0.86
Drylands											
Hyper-arid	3	2	41	1	NA	NA	NA	0.98	0.95	0.96	0.96
Arid	63	6	83	8	0.91	0.89	0.90	0.91	0.93	0.92	0.91
Semi-arid	79	11	108	10	0.88	0.89	0.88	0.92	0.91	0.91	0.90
Dry sub-humid	89	13	111	9	0.87	0.91	0.89	0.93	0.90	0.91	0.90

Tree Cover

<0.10	1	5	120	1	NA	NA	NA	0.99	0.96	0.98	0.98
10-40%	74	12	98	10	0.86	0.88	0.87	0.91	0.89	0.90	0.88
41-65%	79	6	66	11	0.93	0.88	0.90	0.86	0.92	0.89	0.89
66-100%	80	9	60	6	0.90	0.93	0.91	0.91	0.87	0.89	0.90

Supplementary Table 3. Summary of the indicators used to assess classification accuracy. True positives correspond to images that were correctly classified by the model, false positives correspond to images that were misclassified by the model, and false negatives correspond to images of the targeted class that were classified by the models into another class. See examples of true/false positives and negatives in Supplementary Fig. 3.

Accuracy indicator	Equation	Interpretation
Precision (or positive predictive value)	$\dfrac{true\ positives}{true\ positives + false\ positives}$	How many detected forests (or non-forests) are true.
Recall (or sensitivity)	$\dfrac{true\ positives}{true\ positives + false\ negatives}$	How many actual Forests (or Non-forests) are detected.
F1-measure (or F1-score)	$\dfrac{2 * Precision * Recall}{Precision + Recall}$	Index that evaluates the balance between Precision and Recall.

Supplementary Table 4. Importance of variables predicting the current forest distribution in global drylands from random forest permutation analysis. Current Climate (CC), mid-Holocene (MH). %IncMSE indicates the increase of the Mean Squared Error when a given variable is randomly permuted and p-value (for one-sided testing).

Variable	%IncMSE	p-value
Albedo	41.42877	0.01
PH	36.91288	0.01
Latitude	25.61112	0.01
Water Thickness	21.82914	0.01
Elevation	21.71287	0.01
PWaQ (CC)	19.8868	0.01
Slope	19.31883	0.01
Soil Moisture	18.85622	0.01
PDQ (MH)	18.48542	0.01
PCQ (MH)	18.33577	0.01
Longitude	17.9886	0.01
MDRM (CC)	17.78976	0.01
MDRM (MH)	17.782	0.01
PWaQ (MH)	16.29423	0.01
Sand	15.81528	0.01
MTDQ (CC)	15.73885	0.01
PCQ (CC)	15.66399	0.01
PDQ (CC)	15.14815	0.01
Evapotranspiration	14.875	0.01
Soil Organic Carbon	14.76194	0.01
MTWeQ (MH)	14.30748	0.01
MTCM (MH)	14.1081	0.01
MTWM (CC)	13.7493	0.01
PSCV (CC)	13.29451	0.01
PDM (MH)	13.25064	0.01

TARBIO (CC)	13.19653	0.01
MTWaQ (CC)	13.1097	0.01
MTCQ (MH)	12.98462	0.01
MTCM (CC)	12.9202	0.01
AP (CC)	12.8788	0.01
MTWeQ (CC)	12.48877	0.01
MTDQ (MH)	12.15014	0.01
MTWM (MH)	12.09465	0.01
IBIO (MH)	12.09444	0.01
IBIO (CC)	11.85019	0.01
Nitrogen in soil	11.72474	0.01
AP (MH)	11.56379	0.01
TS (MH)	10.99323	0.01
PSCV (MH)	10.89148	0.01
MTWaQ (MH)	10.79875	0.01
MTCQ (CC)	10.73665	0.01
PDM (CC)	10.71321	0.01
AMT (CC)	10.65666	0.01
QWeQ (MH)	10.63615	0.01
PWM (MH)	10.53345	0.01
TARBIO (MH)	10.52528	0.01
PWeQ (CC)	10.23841	0.01
TS (CC)	9.3317	0.01
PWM (CC)	8.66222	0.01
AMT (MH)	7.801763	0.01

Supplementary Data 1. Dataset for train Convolutional Neural Networks.

Supplementary Data 1 is available online as a comma separated values file under the

Figshare repository. https://doi.org/10.6084/m9.figshare.13635212

Supplementary Data 2. Dataset of 16,739 plots (0.5 ha) location and classification.

Supplementary Data 2 is available online as a comma separated values file under the

Figshare repository. https://doi.org/10.6084/m9.figshare.13635212