PNAS RESEARCH ARTICLE | ECOLOGY

The global biogeography and environmental drivers of fairy circles

Emilio Guirado[a,1], Manuel Delgado-Baquerizo[b], Blas M. Benito[a], José Luis Molina-Pardo[c], Miguel Berdugo[d,e], Jaime Martínez-Valderrama[f], and Fernando T. Maestre[a,g]

Edited by Susan Harrison, University of California, Davis, CA; received March 10, 2023; accepted August 2, 2023

Fairy circles (FCs) are regular vegetation patterns found in drylands of Namibia and Western Australia. It is virtually unknown whether they are also present in other regions of the world and which environmental factors determine their distribution. We conducted a global systematic survey and found FC-like vegetation patterns in 263 sites from 15 countries and three continents, including the Sahel, Madagascar, and Middle-West Asia. FC-like vegetation patterns are found in environments characterized by a unique combination of soil (including low nutrient levels and high sand content) and climatic (arid regions with high temperatures and high precipitation seasonality) conditions. In addition to these factors, the presence of specific biological elements (termite nests) in certain regions also plays a role in the presence of these patterns. Furthermore, areas with FC-like vegetation patterns also showed more stable temporal productivity patterns than those of surrounding areas. Our study presents a global atlas of FCs and provides unique insights into the ecology and biogeography of these fascinating vegetation patterns.

spatial patterns | grassland | drylands

Significance

Fairy circles (FCs) are intriguing regular vegetation patterns that have only been described in Namibia and Australia so far. We conducted a global and systematic assessment of FC-like vegetation patterns and discovered hundreds of FC-like locations on three continents. We also characterized the range of environmental conditions that determine their presence, which is restricted to narrow and specific soil and climatic conditions. Areas showing FC-like vegetation patterns also had more stable productivity over time than surrounding areas having non-FC patterns. Our study provides insights into the ecology and biogeography of these fascinating vegetation patterns and the first atlas of their global distribution.

Fairy circles (FCs) are regularly spaced patches of bare ground surrounded by tall grasses (1–3). So far, FCs have only been found in drylands of Namibia and Australia (4), where harsh conditions hamper plant regeneration (5). However, there are pronounced differences between Australian and Namibian FCs. The latter occurs on sandy soils with gaps surrounding vegetation patches showing higher soil moisture (6), whereas in Australian FCs, the finer soil texture induces the opposite response, i.e., higher runoff in gaps and higher infiltration under plant canopies (7). In addition, termites and ants have been found to be the main drivers of FCs in Namibia (8), but not in Australia, where the action of vegetation-water feedbacks has been proposed as the main mechanism driving the formation of FCs [(4); but recent evidence suggests that termites may also play a role there, (9)]. These differences, and the potential role of other factors in FC formation, such as allelopathic effects and soil nutrient deficits, have stimulated a lively scientific debate about the origin of FCs, which is conditioned mainly by the fact that so far there are only two places on Earth where these structures have been found (10–12).

Multiple hypotheses have been proposed to explain the formation of these vegetation patterns, which have fascinated scientists for decades. These include the action of climate [e.g., high interannual variability and narrow range of mean annual precipitation; (6)], nests of social ants or termites (13–16), complex and scale-dependent vegetation-ecohydrological feedbacks (7, 17–19), presence of allelopathic *Euphorbia* species (11, 20), soil characteristics such as sand and nutrient contents (21), or the integration of several of these factors (22). However, predictors of FCs are often evaluated in isolation of each other, and in most cases, each hypothesis proposed is valid only when explaining FC formation locally (4, 8, 11, 23). Together with the limited number of known locations of FCs to date, this has hampered our ability to understand where FCs occur at a global scale and what drives their distribution (24, 25). The consequences of FCs for ecosystem functioning and stability also remain largely unexplored despite the known positive influence of regular vegetation patterns on key ecosystem properties such as vegetation productivity (26).

Theoretical models explain the absence of FCs in mesic regions (22) and suggest that the extent of FCs in drylands, areas where the ratio of precipitation/potential evapotranspiration is less than 0.65 and that cover ~41% of the Earth's surface (27), may be larger than expected (22). Indeed, regular vegetation patterns are relatively common in drylands (28, 29). Identifying the existence of new locations that include FC-like vegetation patterns, i.e., dryland areas with a vegetation spatial structure characterized by circular gaps regularly spaced having a similar spatial structure to that of FCs found in Namibia and Australia (see *Materials and Methods* for more information), analyzing their effects on ecosystem functioning, and discovering their environmental drivers are key to better

Author contributions: E.G., M.D.-B., and F.T.M. designed research; E.G., B.M.B., J.L.M.-P., and M.B. performed research; E.G. and J.M.-V. contributed new reagents/analytic tools; E.G., B.M.B., M.B., and F.T.M. analyzed data; and E.G., M.D.-B., B.M.B., J.L.M.-P., M.B., J.M.-V., and F.T.M. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: emilio.guirado@ua.es.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2304032120/-/DCSupplemental.

Published September 25, 2023.

understanding the causes of the formation of these vegetation patterns, their biogeography, and ecological roles across global drylands (26, 30, 31).

Here, we used remote sensing and deep learning to conduct a systematic survey on 574,799 plots (~1-ha each) to quantify the extent of FC-like vegetation patterns in global drylands. Then, we combined this information with environmental data using machine learning to identify the main environmental factors predicting the distribution of FC-like vegetation patterns worldwide. We also performed a time-series analysis of productivity in each FC-like vegetation pattern location observed and in a systematic sample of other dryland ecosystems to test whether the stability of productivity of FC-like vegetation patterns is higher than that of other vegetation structures (32). Finally, we measured basic attributes of FC-like vegetation patterns (area, shape, and density) and related them to aridity, grazing pressure, and grass species richness to shed light on potential mechanisms underlying their formation and variability across global drylands. Our analyses simultaneously considered key environmental predictors, from climatic and edaphic characteristics to the presence of social insects, including all biotic and abiotic factors identified as important for FCs formation in the literature, plus additional ones that have never been evaluated before (e.g., albedo, topographic variables, soil electrical conductivity, aquifer trends, or wind speed; *Materials and Methods*).

Results and Discussion

Our analyses revealed 263 locations with FC-like vegetation patterns distributed across drylands worldwide. These include those already identified in Namibia and Western Australia, as well as areas never described before, including the Sahel, Western Sahara, Horn of Africa, Madagascar, Southwest Asia, or Central and Southwest Australia. By doing so, our study provides a global atlas of areas showing FC-like vegetation patterns and expands the known existence of this vegetation type to new countries and continents (Fig. 1). To further provide evidence of the similarities between the FC-like vegetation patterns described here and those previously described in the literature, we compared their spatial distribution using the Clark–Evans aggregation R index and the coefficient of variation of distance to the nearest neighbor (*SI Appendix*, Fig. S1). These are typical descriptors of vegetation pattern morphology

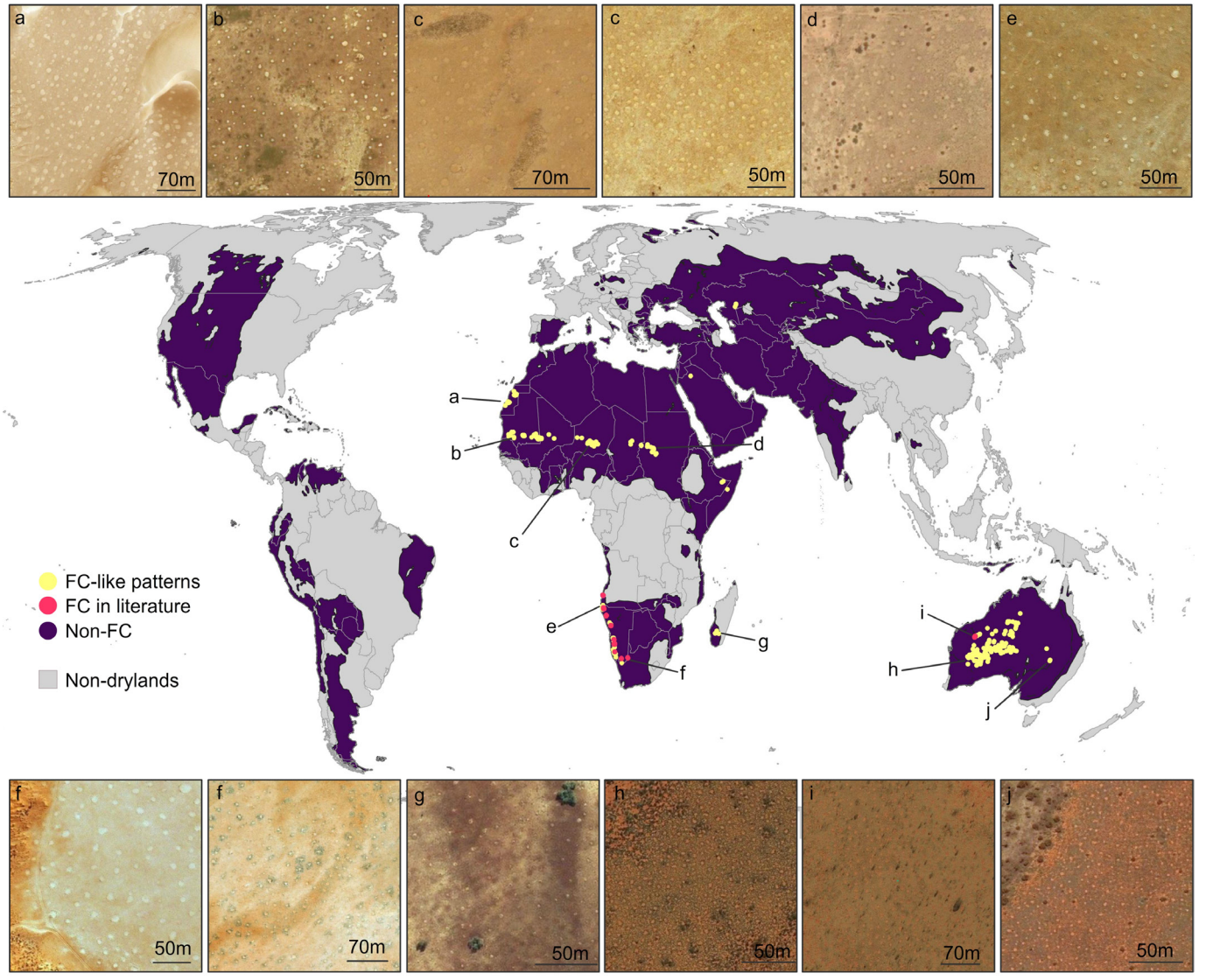

Fig. 1. Location of FC-like vegetation patterns detected in this study (yellow points) and those of FCs previously identified in the literature (red points) across global drylands. The panels show pictures of FC-like vegetation patterns detected by our analyses in different dryland regions. The panels *e*, *f*, and *i* are FC samples from locations identified in the literature (see references in *SI Appendix*, Table S1). Image source: Bing Maps (https://www.bing.com/maps/).

previously used to test similarities between FCs in Australia and Namibia (33). These analyses revealed that the morphological values of the FC-like vegetation patterns reported here and those of the FCs described in the literature were indistinguishable (*SI Appendix*, Figs. S1, S2, S3, and S15). Similar results were observed when evaluating their patterns using additional metrics defining the spatial structure of vegetation, such as the L-function (34) (*SI Appendix*, Fig. S2) and the pair correlation function (3, 23) (*SI Appendix*, Fig. S3). However, some locations in Chad or Kazakhstan were uncertain after conducting these analyses. The differences observed could be because these FC-like vegetation patterns could be transient, gradually disappearing over time and eventually transforming into uniform vegetation (35). In addition, we used Voronoi diagrams (36), an interpolation method based on Euclidean distance, which has been used to determine the hexagonal characteristic distribution of FC-like vegetation patterns (*SI Appendix*, Fig. S4). The average number of neighbors of the discovered FC-like vegetation patterns was 6.71, quite close to the 6.72 measured in the locations supporting FCs provided by the literature (*SI Appendix*, Tables S1 and S2). All these analyses support the identification of FC-like vegetation patterns described in this study resemble the FC spatial patterns previously reported in Australia and Namibia.

The FC-like vegetation patterns described here were associated with narrow and specific values of soil and climatic conditions. These patterns can be found in areas characterized by an arid climate and having either sparse grass cover or open shrublands with interspersed herbaceous vegetation (*SI Appendix*, Fig. S5). The large number of FC-like vegetation patterns locations described here, and the wide range of environmental conditions they encompass, allowed us to conduct the first attempt to understand the factors affecting the distribution of FC-like vegetation patterns at a global scale simultaneously considering all environmental factors expected to be important for the formation of FCs according to the literature (*SI Appendix*, Table S4 and *Materials and Methods*). We fitted two types of models: i) an explanatory model based on a Generalized Linear Regression Model (GLM) and ii) a predictive model based on Random Forest (RF) (37). Both models included environmental factors such as climate, edaphic characteristics, and the presence of social insects. Our explanatory model provided the importance of explanatory variables to determine the presence/absence of FC-like vegetation patterns. We found that soil moisture, forest land cover, soil nitrogen content, terrain slope, and mean annual precipitation were important factors negatively associated with the presence of FC-like vegetation patterns. In contrast, the percentage of sand in the soil, seasonality of precipitation, and mean wind speed were positively associated with the presence of FC-like vegetation patterns. Both termites and ants were of low importance in explaining the presence of FC-like vegetation patterns on a global scale (*SI Appendix*, Fig. S6). Our predictive RF-based model had a high predictive power, with 79.98 ± 1.96 percent of variance predicted. Such a high value indicates that the model can predict a large proportion of the variability observed in our response variable. We found that soil (nitrogen content, soil moisture, percentage of sand in the soil, and soil electrical conductivity) and climatic (mean annual precipitation and mean wind speed) properties were the most important factors predicting the presence of FC-like vegetation patterns (Fig. 2). Both termites and ants (*Anoplolepis* sp.) had a low importance as predictors of FC-like vegetation patterns at global scale (Fig. 2*A*). At regional scale, however, termites were an important predictor in Namibia, while they were less important in regions like the Sahel or Australia (*SI Appendix*, Fig. S7). This is consistent with previous studies showing the importance of social insects as drivers of FC patterns in Namibia (8, 38) but not in Australia [(23); but see ref. 9]. Our analyses further reveal that FC-like vegetation patterns are more likely to be present in hot deserts according to the Köppen climate classification [(39), *SI Appendix*, Fig. S8]. If we consider Whittaker's biomes (40), FC-like vegetation patterns are found within subtropical deserts and seasonal tropical savannas (*SI Appendix*, Fig. S9).

Regardless of the importance of the predictors evaluated, additional analyses conducted to extract partial dependence curves from RF analysis (*Materials and Methods*) showed that the existence of FC-like

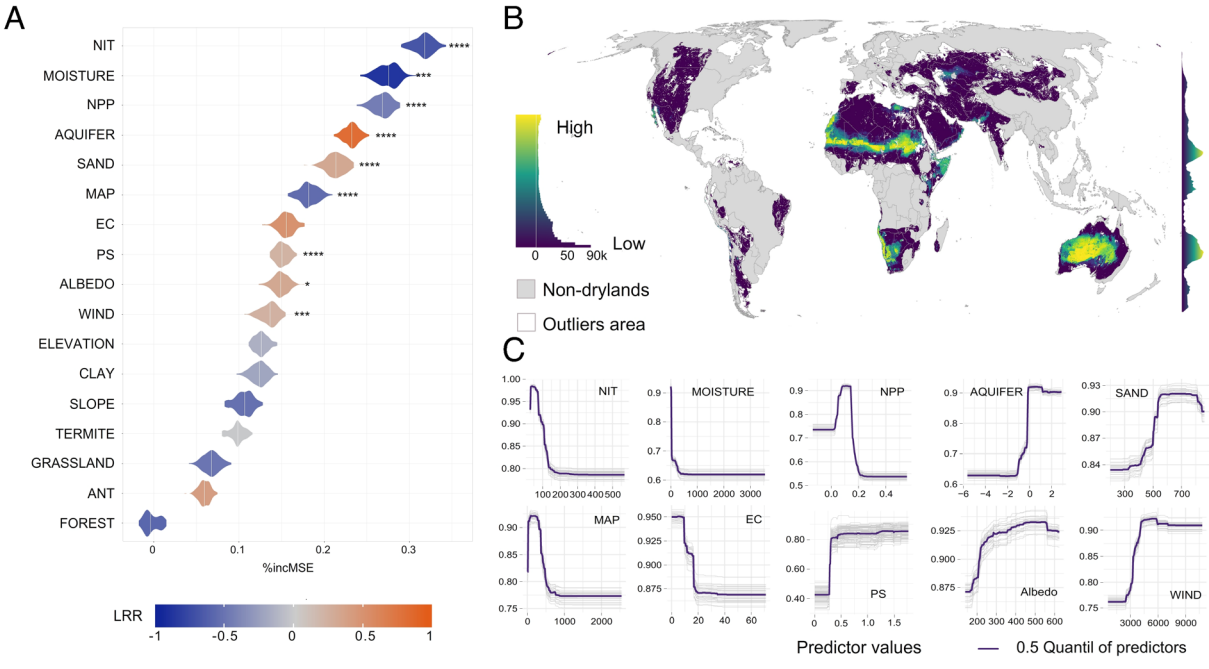

Fig. 2. Relative importance of the drivers of FC-like vegetation patterns (*A*) and probability of finding FC-like vegetation patterns across global drylands (*B*). Panel (*A*) shows the relative contribution of environmental drivers that predict the distribution of FC-like vegetation patterns and their weights. Partial dependence curves (quantile 0.5) for environmental drivers with significance above the median are shown in panel (*C*). Values close to 1 on the *y*-axis indicate values of the predictor where FC-like vegetation patterns would be expected. *$P \le 0.05$; **$P \le 0.01$; ***$P \le 0.001$; ****$P \le 0.0001$. Abbreviations can be found in *SI Appendix*, Table S4.

vegetation patterns is further restricted to narrow and specific values of some environmental variables (Fig. 2*C*). According to the predictive RF model used, we found FC-like vegetation patterns to be most likely present in soils with very low soil moisture, close to 2%, with limited nitrogen contents, between 0.025 and 0.1 g/kg, and with sand content between 52% and 80%. Also, FC-like vegetation patterns were most common in alkaline soils with a pH above 8.5, which typically have a low infiltration capacity, poor structure, and slow permeability (41). Concerning climatic variables, conditions favoring FC-like vegetation patterns include mean annual precipitation values ranging between 100 and 300 mm, and precipitation seasonality (coefficient of variation) higher than 0.26. FC-like vegetation patterns were found in areas with Normalized Difference Vegetation Index [NDVI as a proxy of NPP (42) values between 0.08 and 0.16, and with albedo reflectance values higher than 0.2.

With the information obtained from the RF analyses described above, we produced a global distribution map of FC-like vegetation patterns (Fig. 2*B*). Interestingly, there are areas suitable for FC-like vegetation patterns obtained from the model where these vegetation patterns have not been observed by our analyses using remotely sensed data. This is the case, for example, of the Baja California peninsula, the northwest coast of Libya, or areas near the border between India and Pakistan. In other words, a very restrictive and special set of climatic, edaphic, and nutrient availability conditions determine our predictions of the presence/absence of FC-like vegetation patterns. However, these factors are likely not the only ones predicting the distribution of FC-like vegetation patterns worldwide. In this regard, the impact of factors other than climatic and edaphic factors, including but not limited to, past and present human activities (31, 43, 44) may explain the absence of FC-like vegetation patterns in these areas.

We also measured the spatial distribution and the Clark–Evans R index of 10,620 centroids of FC-like vegetation in 54 sites of ~250 × 250 m, and the size, shape, and density of 4,234 FC-like vegetation in 22 of these sites, which were selected to be representative of the 17 ecoregions (45) where FC-like vegetation patterns were found (*Materials and Methods*). We related these features of FC-like vegetation patterns to i) the richness of grass species at the ecoregion level (28), ii) aridity index at a spatial scale of 1 km/pixel, and iii) grazing pressure (number of head of cattle, goats, and sheep in a census area of 10 × 10 km^2; *Materials and Methods*; Fig. 3*A* and *SI Appendix*, Fig. S10). The size of FC-like vegetation patterns decreased on less arid sites (Fig. 3*D*) but was not correlated by grass species richness or grazing pressure. The size, shape, and density of the FC-like vegetation patterns were correlated, so we grouped them as a

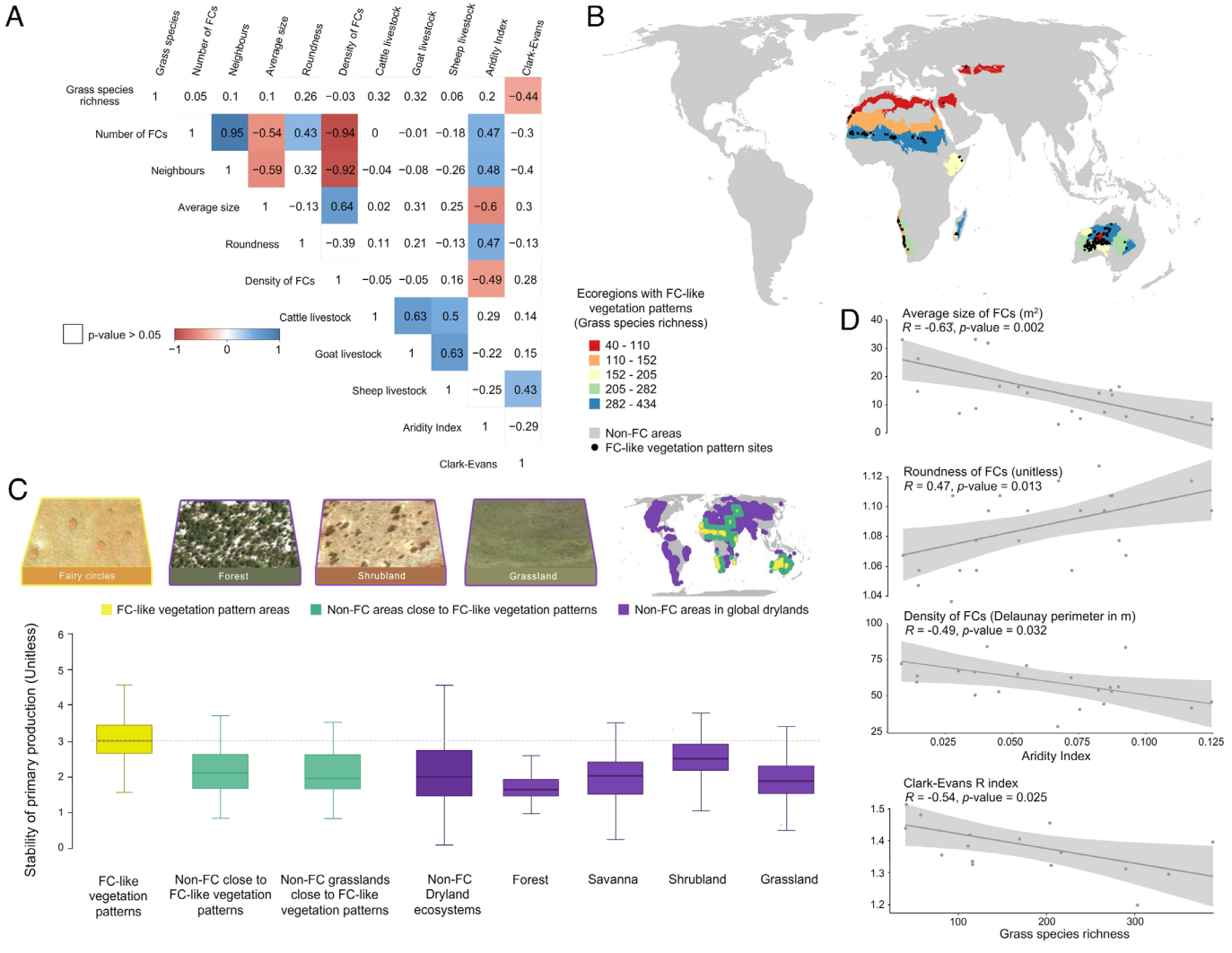

Fig. 3. Spearman correlations between FC-like vegetation patterns and grass species richness, grazing pressure, and the aridity index (*A*). Distribution of FC-like vegetation patterns in dryland ecoregions according to their grass species richness (*B*). Comparison of the stability of primary productivity between areas with FC-like vegetation patterns and both nearby (green) and global (purple) non-FC dryland ecosystems (*C*). Relationships between key structural features of FC-like vegetation patterns and the aridity index and between the Clark–Evans R index and grass species richness (*D*). Image data: Google, Maxar Technologies (49).

multivariate response variable to run a multivariate model to consider the covariance between them [(46); *Materials and Methods*]. Results of these analyses showed that neither grazing pressure nor grass species richness correlated with the size, shape, and density of FC-like vegetation patterns (*SI Appendix*, Table S5). Increases in aridity made FC-like vegetation patterns to become more circle-like and reduced the density of FCs (Fig. 3*D*). The Clark–Evans R index of FC-like vegetation patterns was associated with grass species richness (Fig. 3*D*) but not by grazing pressure nor by aridity. The observed changes along the global aridity gradient evaluated suggest that FC-like vegetation patterns could become more common in a warmer and drier world (47, 48).

The stability of primary productivity, measured using satellite data covering the 2000–2017 period (*Materials and Methods*), was significantly higher in areas with FC-like vegetation patterns than in nearby grassland areas with non-FC (median 0.67 vs. 0.21; unitless). We also compared the stability of FC-like vegetation patterns with that of other non-FC dryland ecosystems and found the same results (higher stability in areas with FC-like vegetation patterns, Fig. 3*C*). These findings provide the first empirical evidence of increased productivity stability, a key ecosystem property that is related to the stable provision of ecosystem services such as forage quantity (50), under FC-like vegetation patterns. Our results thus complement and substantially expand those from local-scale empirical studies showing higher vegetation productivity in dryland ecosystems having a regular spatial pattern of biological structures such as termite mounds (26, 51, 52). Our findings are also in line with predictions from mathematical models suggesting that vegetation spatial structures resulting from self-organizing processes confer higher resilience to ecosystems (22, 32, 53, 54). Whether FC-like vegetation patterns are more resistant and resilient to disturbances can be addressed by future studies using our database, which also will be useful to those interested in comparing the dynamic behavior of these patterns with that of other vegetation patterns found across drylands worldwide.

In summary, our work uncovered the existence of FC-like vegetation patterns in 263 sites from 15 countries and three continents, showing that they are much more common than previously found. The combination of arid climates with high temperature and high precipitation seasonality with soils characterized by low nutrients and coupled to specific biological drivers acting in specific regions (such as the presence of termites in Namibia) support these unique plant structures worldwide and help to better understand the spatial organization of vegetation across global drylands. The global atlas introduced here advances our understanding of the biogeography of FC-like vegetation patterns and will facilitate conducting future research about the characteristics and mechanisms underlying these enigmatic vegetation patterns in locations never studied so far. Our findings also help to reconcile ongoing debates about the formation of FCs. Multiple mechanisms can produce FC-like vegetation patterns and our results reveal both global and regionally specific correlates of these patterns. They indicate that the presence of FC-like vegetation in particular areas is influenced by drivers that may not be as important in others. Our findings thus highlight the need for site-specific experimentation to further identify the mechanisms underlying the formation of FCs. By additionally linking the emergence of FC-like vegetation patterns to a higher stability of vegetation productivity, our work also paves the way for further research on the functional implications of these vegetation structures, which make ecosystems more stable and may help them to avoid tipping points associated with climate change (31).

Materials and Methods

To delimit the extent of global drylands, we followed the United Nations Convention to Combat Desertification and the Convention on Biological Diversity (55), which defined drylands as those areas with an Aridity Index (precipitation/potential evapotranspiration, AI) < 0.65 (56). Here, we used the aridity map created by the United Nations Environment Programme (57). According to this definition, drylands can be divided into four areas: hyperarid (AI < 0.05), arid (0.05 ≤ AI < 0.2), semiarid (0.2 ≤ AI < 0.5), and dry subhumid (0.5 ≤ AI < 0.65).

We define FC-like vegetation patterns following the guidelines given by several authors to define FCs (3, 8, 12). FC-like vegetation patterns are vegetation patterns characterized by gaps regularly spaced and having a similar spatial structure to that of FCs found in Namibia and Australia. We considered that the size and shape (roundness) of FCs should also be similar (i.e., without significant differences) between FC-like vegetation patterns and FCs reported from the literature (*SI Appendix*, Table S1). Therefore, we do not consider in our definition the origin of the formation of FC-like vegetation patterns.

Automatic Identification of FC-like Vegetation Patterns. We gathered two new datasets of very high-resolution images (Datasets S1 and S2) and used a Convolutional Neural Network (CNN)–based model to distinguish between FC-like and non-FC vegetation patterns across global drylands. Our method works with globally available RGB imagery. We preferred to use these data rather than other high-resolution remote sensing data (e.g., LiDAR) because they are not globally available and thus can be only used to detect FC-like vegetation patterns locally (18, 58). The use of CNN models has already been proven to be effective for FC pattern detection (59). While this approach has been employed with aerial or unmanned aerial vehicle (UAV) imagery at local scales (60), it has never been used globally before. In this study, very-high-resolution satellite images (<1 m/pixel) provide sufficient detail to differentiate FC-like vegetation patterns (*SI Appendix*, Fig. S11).

We followed two steps to automatically detect FCs at the global scale. First, we used images from areas where the presence of FCs is well known (*SI Appendix*, Table S1) to train the CNN-based model. We used 15,032 1-ha images in Namibia and Australia in a binary system (50% of the images contained FCs and 50% Non-FC; Dataset S2). These plots were analyzed using the Inception v3 architecture (61), one of the most accurate CNN-based models available, and two optimization techniques: i) data augmentation (62) and ii) transfer learning (63). Data augmentation consists in artificially increasing the number of independent samples in the training dataset by applying specific transformations to the images (i.e., flipping 180°, margin cropping 10%, scale up the size of images in 10%, brightening pixel-level up to 50%, and darkening pixel-level up to 50%). Transfer learning consists of using the knowledge acquired from a previous problem to solve a new problem. Instead of starting the learning from scratch, a pretrained model is selected and retrained on the new problem with transfer learning. We used a pretrained CNN-based model using the ImageNet database [which contains 1,000 image categories including fauna, flora, artificial elements, and others; (64) with a learning rate of 0.001 and a decay factor of 16 every 30 epochs. As an optimization algorithm, we used RMSProp (65) with momentum and decay of 0.9, and epsilon of 0.1. Second, to detect the presence of FC-like vegetation patterns across global drylands, a systematic dataset of 574,799 1-ha sample plots with very high-resolution images from Bing maps (https://www.bing.com/maps) between September 1, 2020 and September 30, 2020 at zoom level 19 (~1 m/pixel) and a regular distribution (10 km apart) was used (Dataset S1). Our approach was inspired by Collect Earth (49), an Earth monitoring software developed by the Food and Agriculture Organization of the United Nations (FAO). It uses freely available very high–resolution (VHR) satellite images (e.g., from Google Earth, Bing Maps) as a base map. Here, we simply replaced the manual approach of Collect Earth, which is elaborated by photointerpretation experts by a deep learning–based model (i.e., CNN) that works an on-the-fly classification of the VHR satellite images from Bing Maps.

Assessing the accuracy of the CNN-based model. Our CNN-based model learned to locate FC-like vegetation patterns in satellite images by using convolution layers. These layers are composed of a set of filters that slide over the image performing mathematical operations to extract features from the image relevant to detect FC-like vegetation patterns. Each filter is activated based on specific patterns in the image, such as lines, edges, or in this case rounded shapes. As the convolutional neural network is trained, the filters are adjusted to detect

more complex patterns, allowing the neural network to learn to recognize specific FC-like vegetation patterns in satellite images. However, this approach also suffers from limitations, as CNN-based models are highly dependent on training data. If the training data are incomplete or biased, the model may be inaccurate. In addition, CNN-based models can be overfitted if the number of parameters is too high relative to the size of the training data. This would make it difficult to make decisions by the model and to detect patterns accurately. To ensure that the results of our CNN-based models are trustable, we used a robust, unbiased database to train the CNN-based model (see previous section) and assessed the accuracy of the classification it provided. For doing so, we selected all plots of FC-like vegetation patterns classified and a subset of non-FC plots (526 in total, 263 FC-like vegetation patterns and 263 non-FC plots systematically selected across drylands, see Dataset S3 and Movie S1). Two of us (E.G and J.L.M-P) tested, with photointerpretation of very high-resolution satellite images (~1 m/pixel), whether the CNN-based model correctly classified FC-like vegetation patterns based on two criteria: a) visual recognition of circular vegetation gap patterns and b) absence of shading to discard structures such as termite mounds (*SI Appendix*, Fig. S12). For higher certainty, three different image sources with sufficient spatial resolution to identify the smallest FCs (5 m in diameter) were used for these tests [Google Earth, Bing Maps, and Mapbox; (66) see *SI Appendix*, Fig. S11]. The accuracy metrics used were Precision (1), Recall (2), and the F1-measure (3), which were calculated according to (67)

$$Precision = \frac{True\ \ Positive}{True\ \ Positive\ \ +\ \ False\ \ Positive}, \quad [1]$$

$$Recall = \frac{True\ \ Positive}{True\ \ Positive\ \ +\ \ False\ \ Negative}, \quad [2]$$

$$F1 - measure = \frac{2\ \times\ Precision\ \times\ Recall}{Precision\ \ +\ \ Recall}, \quad [3]$$

where precision is the ratio of correctly predicted positive observations to the total number of correctly predicted positive observations, recall is the ratio of correctly predicted positive observations to all observations in the FC-like vegetation patterns class, and the F1-measure is the weighted average of precision and recall. The results of the global scale FC-like classification with the CNN-based model were satisfactory, with 263 true positives and seven false positives detected. This resulted in a precision of 0.97, and recall and F1-measure values of 1 and 0.98, respectively (F1-measure values close to 1 indicate that the CNN-based model is more accurate).

We also evaluated the performance of the binary classification provided by our model (i.e., FC-like and non-FC vegetation patterns). For doing so, we calculated the Receiver Operating Characteristic (ROC) curve (68). It is a scatter plot used showing the relationship between the true positive rate (sensitivity) and the false positive rate (1-specificity) as the classification threshold is varied. In addition, the area under the curve (AUC), a numerical measure representing model performance, was also calculated. The AUC obtained was 0.98, indicating that the model has a good fit detecting FC-like vegetation patterns.

Selection of Nonredundant Areas to Compare FCs and FC-like Vegetation Patterns. To avoid biasing the results by having some areas more overrepresented than others, we evaluated a representative sample of all sites with FC-like vegetation patterns detected. We considered the ecoregion, which is defined as a geographic territory determined by climate, geology, hydrography, fauna, and flora, as the unit of separation (45). We did this to ensure the independence of the sites used, i.e., that they do not constitute an overrepresentation of a given ecoregion in which there are more FC-like vegetation pattern locations than in others (*SI Appendix*, Figs. S13 and S14). First, we performed a photointerpretation analysis of all detected sites with FC-like vegetation patterns. Second, we compared the characteristics of the FC-like vegetation patterns identified (e.g., area, shape, density, and spatial patterns) to that of FCs reported in the literature to ensure that we obtained a representative sample of all FC-like vegetation patterns locations detected by the CNN-based model.

Comparison of the Spatial Patterns of Detected FC-like Vegetation Patterns and FCs Reported in the Literature. To assess our confidence in the new FC-like vegetation patterns reported, we i) compared these patterns with those from areas where FCs have been described in the literature, e.g., Namibia and Australia (*SI Appendix*, Table S1 and Dataset S1) and ii) established a level of certainty that reflects our confidence in the detection of the new FC-like vegetation patterns.

To compare the spatial patterns of FCs reported in the literature (Namibia and Australia; 4,10,23,112) and the FC-like vegetation patterns found in this study, we used the Clark–Evans aggregation R index with edge correction of Donnelly (33, 69), the nearest neighbor distance (70), the L-transform of the Ripley's K-function (L-function) with 99 simulations border correction (34), and the pair correlation function (3, 23) (PCF) with 30 simulations and edge correction in ~250 × 250 m plots (see code in *SI Appendix*, Software S1). To better visualize the results of these analyses, we used the stability of the nearest-neighbor distance (1–the coefficient of variation) on the *x*-axis and the Clark–Evans R index on the *y*-axis (*SI Appendix*, Fig. S1). A value Clark-Evans R index > 1 suggests ordering, while Clark–Evans index < 1 suggests clustering. Values of L(r) > 0 indicate an aggregation pattern of FCs, while values of L-function L(r) ≤ 0 suggest a segregated regular pattern of FCs (*SI Appendix*, Fig. S2). The pair correlation function [PCF; also known as the radial distribution function, g(r)] is used with the envelope method as implemented in the *spatstat* R package (71). The PCF was calculated with 30 simulations and the "translation" edge correction was used (*SI Appendix*, Software S1). This correction uses a displacement technique to adjust for edge effects. It consists of randomly shifting the points of the dataset within the observation window and calculating the point correlation function at each shift. The results are then averaged to obtain the final envelope. This correction is useful when precise information on edge effects is not available, and a robust estimate is desired. Values of g(r) = 0 indicate that the PCF did not detect any FC-like vegetation within the specified neighborhood radius. A value of g(r) < 1 suggests evenness or overdispersion, while g(r) > 1 suggests aggregation.

To determine the hexagonal distribution of the FC-like vegetation patterns identified here and those of FCs from Namibia and Australia, we used Voronoi diagrams (36), an interpolation method based on the Euclidean distance between the centroids of the FCs. The diagrams are created by joining the points to each other, tracing the bisectors of the joining segments. The intersections of these bisectors determine a series of polygons in two-dimensional space around a set of control points, such that the perimeter of the polygons generated is equidistant from neighboring points and designate their area of influence.

We also compared the size, shape, and density estimation from Delaunay triangulation (72) between the FCs reported in the literature and the FC-like vegetation patterns observed here. Finally, we used a Kruskal–Wallis test (73) to look for significant differences in the measurements described above (*SI Appendix*, Fig. S15). We conducted an analysis based on photointerpretation of satellite images to evaluate the size and shape of FC-like vegetation patterns. The perimeter of each FC was drawn, and we calculated the area (in m^2) of each FC with a Geographic Information System (QGIS, v.3.14 QGIS Development Team, 2021) (74). To determine the shape of the FCs, we used the Roundness shape index (RSI; (4), which calculates the equivalence of a polygon to a circle [*SI Appendix*, Table S2 and geospatial data (75)]:

$$RSI = (\$perimeter * \$perimeter / (4 * pi())) / \$area. \quad [4]$$

A RSI value of 1 would indicate a perfect circle.

Available images may have different spatial resolutions and may have been acquired in weather conditions that may make it difficult to visualize vegetation patterns (e.g., having clouds). Thus, we established a level of certainty in the FC-like vegetation patterns observed. Because some images of areas with similar vegetation patterns look better than others (e.g., due to resolution and other atmospheric factors), we created three confidence levels (i.e., definite, probable, and possible). The confidence level was assigned based on a combination of criteria: 1) The clarity and resolution of the satellite images. Images with a pixel size greater than the minimum diameter of FCs from the literature (76) (~2 m) were discarded and 2) analysis of significant differences in the FC-like vegetation patterns and FCs from the literature of the size, shape, density, and spatial patterns. The classification of new FC-like vegetation patterns according to confidence level is available in Dataset S3.

Assessing Potential Predictors of the Distribution of FC-like Vegetation Patterns. To identify the importance of the drivers in the distribution of FC-like vegetation patterns, we used key environmental predictors, including climate, edaphic variables, termite and ant distribution data, vegetation productivity,

reflectance satellite data, hydrogeological, land cover, and topographical information (*SI Appendix*, Table S4). The predictors used are described in more detail below.

Topography, climate, and soil variables. Elevation and slope were obtained from the Advanced Land Observation Satellite (77). A total of 19 standardized climatic variables were obtained for all the sites surveyed from WorldClim Global Climate Data (78). We used data at a 30 arc-second resolution (~1 km at the equator), the highest resolution available from this database. We gathered soil moisture data from TerraClimate (79), which is composed of data from Worldclim and the Japanese 55-y Reanalysis (JRA55) (80). Soil organic carbon stocks, texture (sand and clay contents), nitrogen content, and pH were obtained from Soilgrids (81).

Productivity, albedo, and aquifer level trends. The NDVI a proxy of aboveground net primary production (42), was obtained from the MODIS/Terra MOD13Q1 Version 6 product, which has a 250-m pixel spatial resolution (82). Albedo was obtained from the MODIS/Terra MCD43A3 Version 6 product at a 500-m pixel resolution (83). The equivalent liquid water thickness of aquifers was obtained by measuring monthly changes in gravity from the Gravity Recovery and Climate Experiment (84) as described in ref. 85.

Land cover, wind speed, and biological attributes. Global grassland and forest land cover types were obtained for 2016 from the MODIS/Terra MCD12Q1 Version 6 product at a 500-m pixel spatial resolution derived from the International Geosphere-Biosphere Programme (86). Mean wind speed was obtained from the Global wind atlas version 3. The Global Wind Atlas 3.0 is released in partnership with the World Bank Group, utilizing data provided by Vortex, using funding provided by the Energy Sector Management Assistance Program (ESMAP) (87). We used the global map of the termite richness (88) from several databases belonging to the Worldmap program (89) and the collections of the Natural History Museum, London (90). The spatial information of ant richness was obtained from Antmaps (https://www.antmaps.org), developed from the Global Ant Biodiversity Informatics database (91). Ant data include data on the species *Psammotermes allocerus* and on the richness of the genus *Anoplolepis*, which have been related to the formation of FCs in Namibia (2).

We recognize that the distribution of FC-like vegetation patterns may be affected by other potential predictors not considered in this study. For example, the hypothesis of FCs formation by the latex of *Euphorbia gummifera* and *Euphorbia damarama* has been suggested as a driver of FCs in Namibia (12). However, the distribution of these species is not sufficiently widespread for a global analysis (Gbif database, https://www.gbif.org, Accessed October 2021) and thus was not considered in our study.

Explanatory and Predictive Models of FC-like Vegetation Patterns. We fitted two types of models to explain and predict the presence of FC-like vegetation patterns: i) an explanatory model based on a GLM and ii) a predictive model based on RF. An explanatory model aims to explain the causal relationships between variables, i.e., how one variable influences another and how. This type of model focuses on understanding the underlying process that generates the data and seeks to identify the most important variables to explain the phenomenon studied. An explanatory model is useful when seeking to understand the causal relationships behind the data and can help to generate hypotheses and make informed decisions. On the other hand, a predictive model focuses on predicting the value of a variable of interest from other variables. This type of model is not necessarily concerned with understanding the underlying causal relationships between variables but rather focuses on finding patterns in the data that allow predicting the value of the variable of interest. A predictive model is useful when forecasting or making decisions based on the predictions it generates.

To select a reduced set of predictor variables for both explanatory and predictive models, we employed two criteria: correlation and the variance inflation factor (VIF) with R package spatialRF (92, 93). The full set of candidate predictor variables is shown in *SI Appendix*, Table S4, which includes all biotic and abiotic factors identified as important for FC formation in the literature. In addition, we have added additional predictor variables that have never been evaluated before (*SI Appendix*, Table S4). First, we calculated the spatial correlation coefficient between each pair of variables and specified a correlation threshold of 0.7 (94) to determine whether two variables are highly correlated with each other. Second, we performed an additional variable selection based on the VIF of each variable. We discarded all variables that have a VIF greater than 5 (95, 96). Finally, the variables selected were MAP, PS, mean wind speed, albedo, NPP, elevation, terrain slope, soil moisture, soil sand content, soil nitrogen content, soil electrical conductivity, aquifer trend, termites, and ants (see abbreviations in *SI Appendix*, Table S4).

Explanatory model: Explaining the importance of variables in the presence/absence of FC-like vegetation patterns. To explain the importance of predictor variables in the presence/absence of FCs, we used a GLM (97). The GLM assumes that the response variable has a probability distribution from the exponential family, which includes the normal, binomial, Poisson, and gamma distributions, among others. The model specifies a linear relationship between the predictor variables and the transformed mean of the response variable, where the transformation function depends on the choice of the response variable distribution. An exploratory analysis of the data was performed to identify possible problems, such as outliers or missing data. The variables were standardized. This avoids biases in the model and allows direct comparison of the estimated coefficients. Standardization was performed using the formula:

$$x\_standardized = (x - mean(x)) / sd(x),$$

where x is the original value of the variable, mean(x) is the mean of the variable, and sd(x) is the SD of the variable.

The data were divided into two sets: a training set and a test set. We used a random sampling technique to select 70% of the data for the training set and the remaining 30% for the test set (Dataset S3). This is done to ensure that the model fits the data well and can be generalized to new data. We used the "glm" function of R to fit a GLM. The model formula is specified using the dependent variable and the predictor variables. We set the "family" argument to "binomial" to indicate that this is a binary classification problem, and the "logit" link is specified, which is the canonical link for Logistic Regression. After fitting the model, the "vip" function from the "vip" library was used to calculate the relative importance of the predictor variables (98).

Finally, the fitted model was used to make predictions on the test set. We calculated a confusion matrix, which shows the number of true positives, true negatives, false positives, and false negatives. Some classification evaluation metrics (67), including precision, sensitivity or recall, and F1-measure (see refs. 1 and 3) above), were calculated to evaluate the model performance on the test set. To do the validation between observed and predicted data, a threshold of ≥0.5 probability for the presence of FC-like vegetation patterns was applied, i.e., all probabilities less than 0.5 were absences and all probabilities greater than or equal to 0.5 were the presence of FC-like vegetation patterns. The GLM delivered precision, recall, accuracy, and F1-measure values of 0.90, 0.85, 0.95, and 0.90, respectively. These metrics indicate that the model fitted the observed data well.

Predictive model: Mapping the global distribution of FC-like vegetation patterns and their predictor importance. To assess the importance of selected environmental predictors (MAP, PS, mean wind speed, albedo, NPP, elevation, terrain slope, soil moisture, soil sand content, soil nitrogen content, soil electrical conductivity, aquifer trend, termites, and ants, *SI Appendix*, Table S4) to predict the distribution of FC-like vegetation patterns, a RF model (99) was used on a sample of 526 sites that were previously classified using the CNN-based model (Dataset S3). This sample was balanced into 263 FC-like vegetation patterns sites and a systematic sample, i.e., separated in space with the same distance, of 263 non-FC patterns sites across drylands (Dataset S3). The objective with this analysis is to help fit predictive spatial regression, where the goal is to understand how a set of predictors and the spatial structure of the data influence the response variable (presence/absence of FC-like vegetation patterns). The significance of each predictor was determined by assessing the decrease in prediction accuracy when the predictor values are randomly permuted (unconditional permutation). We used the Mean Error Increase When Permuted (MEIP) algorithm, which is used to assess the importance of predictor variables in RF analyses. Instead of removing each variable from the dataset and measuring its effect on model accuracy, MEIP evaluates the importance of each variable by randomly permuting its values within the dataset. That is, the values of a variable are randomly exchanged across all observations in the dataset and the effect of this permutation on the accuracy of the model is measured. The MEIP measures the difference between the accuracy of the model in the original dataset and the accuracy of the model in the dataset with the permuted values. This difference is calculated for each variable in the model and averaged to obtain a measure of the importance of the variable in the model (100).

To know exactly the optimal range of values of the environmental predictors of FCs, we used the partial dependence curves of these predictors (101). These analyses helped us to understand how the response curve of a driver changes

when all other variables have low values. For instance, the quantile 0.5 for the construction of the partial dependence curve of the predictors (92) was used to plot driver values near to 1 representing optimal intervals for FC-like vegetation patterns. Low values of nitrogen (~50 cg/kg) and high for sand (>600 g/kg) described the best range of values for these predictors (Fig. 2*C*).

To determine the weights (positive, negative, or neutral directions) of the drivers of FC-like vegetation patterns (Fig. 2*A*), the Log-Response Ratio [LRR = log (FC-like vegetation patterns driver values/non-FC driver values)] was calculated (102). This ratio was calculated as the logarithm in base 10 of all driver values of FC-like vegetation pattern locations by the driver values of the non-FC locations. The results of the LRR indicate values greater than 0 when the value of the driver is higher in locations having FC-like vegetation patterns, and values less than 0 for the opposite case. For example, the LRR of the NPP driver is negative (−0.23) because the locations where FC-like vegetation patterns are found have lower NPP than those not having FC-like vegetation patterns.

Once these analyses were completed, and to map the distribution of FC-like vegetation patterns across global drylands, we used a RF regression analysis (103) using as inputs the most important variables revealed by the RF analysis with the 526 FC/non-FC plots previously classified using the CNN-based model. This model was built by finding the set of covariate combinations that most robustly predicted the training samples with a configuration of 999 decision trees in the model. The quality of the classification was tested and validated using a k-fold cross-validation method (104), where k (k = 5) models were trained from k subsets of the original data (total number of plots minus the total number of plots divided by k) and tested on k subsets of the remaining independent data (total number of plots divided by k). By combining the k iterations, we compared the original full dataset with the remaining independent data. To assess the accuracy of the predictions calculated from the RF-based model, we calculated how much the parameter space of the predictors differed from the original dataset. The modeling approach was then validated by returning the predicted values (*x*-axis) vs. the observed values (*y*-axis), following ref. 105.

To further provide more reliable predictions of the distribution of FC-like vegetation patterns across global drylands, we used the Mahalanobis distance. It is a measure of the distance between a point and a multivariate normal distribution. It is often used to detect outliers, since a point having a large distance from the normal distribution is considered an outlier. In our study, we used the Mahalanobis distance to calculate an outlier mask based on the location of points with FC-like and non-FC vegetation patterns. To do this, we first computed the covariance matrix of the sites. Second, we calculate the Mahalanobis distance between each point and the normal distribution, using the covariance matrix calculated earlier. Once we calculated the Mahalanobis distance for each point, we defined the threshold to determine which areas are outliers. Here, we have considered 0.05 of the probability of a larger value of chi-square with fourteen degrees of freedom (the number of predictor variables used in the RF analysis) (106). We could then create an outlier's mask that marks these areas on the map in white areas (Fig. 2*B*, see also *SI Appendix*, Figs. S16 and S17).

Multivariate Linear Mixed-Effects Model. We fitted a multiple regression model with four continuous response variables [Clark–Evans R index (33), average size, average roundness shape index, and density of FC-like vegetation] that depend on five continuous predictor variables (grass species richness, Aridity Index, and density of cattle, goat, and sheep; the later three variables act as proxies of grazing pressure) (*SI Appendix*, Software S1). Grass species richness values were obtained from the GBIF Plantae dataset (GBIF.org 2020) and were grouped into the global ecoregions defined in Dinerstein et al. (45) to avoid database sampling bias (28). Aridity index data were obtained from Global Aridity Index and Potential Evapotranspiration Database (56). Livestock density data were obtained from the Harvard Dataverse, V1 (107) as the number of heads of cattle, goats, and sheep per 100 km^2 (*SI Appendix*, Table S3).

The Stability of FC-like Vegetation Patterns. To quantify the stability of primary productivity, we used an 18-y period (2000–2017) of NDVI (5 data from Landsat 7 Enhanced Thematic Mapper Plus (ETM+) sensor (108) as a proxy for net primary production (42). Landsat ETM+ images were atmospherically corrected using the Landsat Ecosystem Disturbance Adaptive Processing System and included a cloud, shadow, water, and snow mask produced using the C Function of Mask, and a per-pixel saturation mask (109). NDVI was calculated as

$$\mathrm{NDVI} = (R_{NIR} - R_{red})/(R_{NIR} + R_{red}), \quad [5]$$

where R_{NIR} and R_{red} are the spectral reflectance near-infrared (0.77–0.90 μm) and in the red (0.63–0.69 μm) bands of Landsat ETM+. The NDVI calculation produces values between −1 and 1, where positive values indicate areas with vegetation, neutral are usually areas without vegetation cover, such as bare soil, and negative are clouds or water. The stability of primary productivity was obtained as the ratio between the mean annual NDVI calculated from 2000 to 2017 and the SD of the annual NDVI during this period (110).

We compared the stability of primary productivity of FC-like vegetation patterns with that of other dryland ecosystems. For doing so, we first obtained a representative sample of dryland ecosystems by using a regular grid of locations separated by 10 km and labeled according to their vegetation type (i.e., forests, savannas, shrublands, and grasslands; Fig. 3*C*). We then compared the stability of areas with FC-like patterns to that of nearby non-FC locations (located at a distance of 5 degrees, which is the minimum distance in which it was observed that there were no presence of FC-like vegetation patterns) using a Kruskal–Wallis test (73).

Assessing the Relationships between Environmental Factors and Key Features of FC-like Vegetation Patterns. We used Spearman's partial correlations (ppcor package in R software) (111) to explore the relationships between the size, shape, or density of the FC-like vegetation patterns and aridity index (precipitation/potential evapotranspiration), grazing pressure, and the richness of grass species (Fig. 3*A*). We focused on these predictors because i) aridity is a key driver of regular vegetation patterns in drylands (28), ii) we saw in photointerpretation analysis numerous herbivore tracks in areas showing FC-like vegetation patterns (*SI Appendix*, Fig. S10), and iii) the sites with FC-like vegetation patterns are mostly grasslands. Spearman rank correlations are a nonparametric approach that does not require normality of the data or homogeneity of variances, measures the strength and direction of the association between two ranked variables, and can be used to associate two variables regardless of whether they are ordinal, intervals, or ratios. Grass species richness values were obtained from the GBIF Plantae dataset (GBIF.org 2020) and were grouped into the global ecoregions defined in Dinerstein et al. (45) to avoid database sampling bias (28). Aridity index data were obtained from Global Aridity Index and Potential Evapotranspiration Database (56). Livestock pressure for cattle, goats, and sheep was obtained from the Harvard Dataverse V1 database (107).

Data, Materials, and Software Availability. Datasets S1 to S3, Geospatial data, and Source code in R language for the analyses conducted in this study have been deposited in Figshare (https://figshare.com/s/ed9e5e07f30c6fe8bfd5 (112); https://figshare.com/s/5c18dab9d697ee553ffe (113); https://figshare.com/s/d6ae64213be36e30e160 (114); https://figshare.com/s/7e0d9860cd170165adf5 (115); https://doi.org/10.6084/m9.figshare.17069954 (75); and https://figshare.com/s/69f1f0d9ee363977e83f (116)). All study data are included in the article and/or supporting information.

ACKNOWLEDGMENTS. We thank Dr. Emilio Rodríguez-Caballero and Dr. Cecilio Oyonarte for their insightful comments in our discussion about fairy circles and Dr. Javier Blanco for comments on an earlier draft. We also thank Dr. Siham Tabik and Prof. Francisco Herrera for providing the computer servers for the deep learning analysis and Prof. G. Darrel Jenerette for his comments and suggestions on the second version of this manuscript. This research was funded by the European Research Council (BIODESERT project, Grant agreement 647038), Generalitat Valenciana (CIDEGENT/2018/041), and the Spanish Ministry of Science and Innovation (EUR2022-134048). M.D-B. acknowledges support from TED2021-130908B-C41/AEI/10.13039/501100011033/Unión Europea NextGenerationEU/PRTR and from the Spanish Ministry of Science and Innovation for the I+D+i project PID2020-115813RA-I00 funded by MCIN/AEI/10.13039/501100011033. E.G. acknowledges the support by the Generalitat Valenciana and the European Social Fund (APOSTD/2021/188). M.B. is also supported by a Ramón y Cajal grant (RYC2021-031797-I).

Author affiliations: [a]Instituto Multidisciplinar para el Estudio del Medio 'Ramón Margalef', Universidad de Alicante, Alicante 03690, Spain; [b]Laboratorio de Biodiversidad y Funcionamiento Ecosistémico. Instituto de Recursos Naturales y Agrobiología de Sevilla (IRNAS), Consejo Superior de Investigaciones Científicas (CSIC), Sevilla 41012, Spain; [c]Centro de Colecciones Científicas de la Universidad de Almería, Almería 04120, Spain; [d]Crowther Lab, Department of Environmental Systems Science, Institute of Integrative Biology, ETH-Zürich, Zürich 8092, Switzerland; [e]Departamento de Biodiversidad, Ecología y Evolución, Facultad de Ciencias Biológicas, Universidad Complutense de Madrid, Madrid 28040, Spain; [f]Estación Experimental de Zonas Áridas, Consejo Superior de Investigaciones Científicas (CSIC), Almería 04120, Spain; and [g]Departamento de Ecología, Universidad de Alicante, Alicante 03690, Spain

1. T. Becker, S. Getzin, The fairy circles of Kaokoland (North–West Namibia) origin, distribution, and characteristics. *Basic Appl. Ecol.* **1**, 149–159 (2000).
2. N. Juergens, The biological underpinnings of Namib Desert fairy circles. *Science* **339**, 1618–1621 (2013).
3. S. Getzin, H. Yizhaq, W. R. Tschinkel, Definition of 'fairy circles' and how they differ from other common vegetation gaps and plant rings. *J. Veg. Sci.* **32**, e13092 (2021), 10.1111/jvs.13092.
4. S. Getzin *et al.*, Discovery of fairy circles in Australia supports self-organization theory. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 3551–3556 (2016).
5. S. Guan, F. Chen, K. Lv, J. Zhou, Y. Huang, Effects of soil water conditions on seedling regeneration in *Myricaria laxiflora* remnant populations. *Ecol. Res.* **35**, 524–532 (2020).
6. M. D. Cramer, N. N. Barger, Are Namibian 'fairy circles' the consequence of self-organizing spatial vegetation patterning? *PLoS ONE* **8**, e70876 (2013).
7. S. Getzin *et al.*, Adopting a spatially explicit perspective to study the mysterious fairy circles of Namibia. *Ecography* **38**, 1–11 (2015).
8. N. Jürgens *et al.*, Largest on earth: Discovery of a new type of fairy circle in Angola supports a termite origin. *Ecol. Entomol.* **46**, 777–789 (2021).
9. F. Walsh *et al.*, First Peoples' knowledge leads scientists to reveal 'fairy circles' and termite linyji are linked in Australia. *Nat. Ecol. Evol.* **7**, 610–622 (2023).
10. W. R. Tschinkel, The life cycle and life span of Namibian fairy circles. *PLoS ONE* **7**, e38056 (2012).
11. J. J. M. Meyer *et al.*, The allelopathic, adhesive, hydrophobic and toxic latex of *Euphorbia* species is the cause of fairy circles investigated at several locations in Namibia. *BMC Ecol.* **20**, 45 (2020).
12. J. M. Meyer, C. S. Schutte, N. Galt, J. W. Hurter, N. L. Meyer, The fairy circles (circular barren patches) of the Namib Desert–What do we know about their cause 50 years after their first description? *S. Afr. J. Bot.* **140**, 226–239 (2021).
13. S. Grube, The fairy circles of Kaokoland (Northwest Namibia)–Is the harvester termite *Hodotermes mossambicus* the prime causal factor in circle formation? *Basic Appl. Ecol.* **3**, 367–370 (2002).
14. N. Juergens *et al.*, Weaknesses in the plant competition hypothesis for fairy circle formation and evidence supporting the sand termite hypothesis. *Ecol. Entomol.* **40**, 661–668 (2015).
15. K. Vlieghe, M. Picker, V. Ross-Gillespie, B. Erni, Herbivory by subterranean termite colonies and the development of fairy circles in SW Namibia. *Ecol. Entomol.* **40**, 42–49 (2015).
16. F. J. Walsh, A. D. Sparrow, P. Kendrick, J. Schofield, Fairy circles or ghosts of termitaria? Pavement termites as alternative causes of circular patterns in vegetation of desert Australia. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E5365–E5367 (2016).
17. E. Meron, Pattern-formation approach to modelling spatially extended ecosystems. *Ecol. Modell.* **234**, 70–82 (2012).
18. S. Getzin *et al.*, High-resolution images and drone-based LiDAR reveal striking patterns of vegetation gaps in a wooded spinifex grassland of Western Australia. *Landsc. Ecol.* **37**, 829–845 (2021), 10.1007/s10980-021-01358-9.
19. S. Getzin, S. Holch, H. Yizhaq, K. Wiegand, Plant water stress, not termite herbivory, causes Namibia's fairy circles. *Perspect. Plant Ecol. Evol. Syst.* **57**, 125698 (2022).
20. J. J. M. Meyer, F. Senejoux, H. M. Heyman, N. L. Meyer, M. A. Meyer, The occurrence of triterpenoids from *Euphorbia gummifera* inside the fairy circles of Garub in the southern Namibian pro-desert. *S. Afr. J. Bot.* **98**, 10–15 (2015).
21. M. D. Cramer, N. N. Barger, W. R. Tschinkel, Edaphic properties enable facilitative and competitive interactions resulting in fairy circle formation. *Ecography* **40**, 1210–1220 (2017).
22. C. E. Tarnita *et al.*, A theoretical foundation for multi-scale regular vegetation patterns. *Nature* **541**, 398–401 (2017).
23. S. Getzin, H. Yizhaq, M. D. Cramer, W. R. Tschinkel, Contrasting global patterns of spatially periodic fairy circles and regular insect nests in drylands. *J. Geophys. Res. Biogeosci.* **124**, 3327–3342 (2019).
24. C. Fernandez-Oto, M. Tlidi, D. Escaff, M. G. Clerc, Strong interaction between plants induces circular barren patches: Fairy circles. *Philos. Trans. A Math. Phys. Eng. Sci.* **372**, 20140009 (2014), 10.1098/rsta.2014.0009.
25. Y. R. Zelnik, E. Meron, G. Bel, Gradual regime shifts in fairy circles. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 12327–12331 (2015).
26. R. M. Pringle, D. F. Doak, A. K. Brody, R. Jocqué, T. M. Palmer, Spatial pattern enhances ecosystem functioning in an African savanna. *PLoS Biol.* **8**, e1000377 (2010).
27. J. Yao *et al.*, Accelerated dryland expansion regulates future variability in dryland gross primary production. *Nat. Commun.* **11**, 1665 (2020).
28. F. T. Maestre *et al.*, Biogeography of global drylands. *New Phytol.* **231**, 540–558 (2021).
29. X. Dong, A trait-based approach to self-organized pattern formation in ecology. *Front Ecol. Evol.* **8**, 377 (2020).
30. M. Berdugo, S. Kéfi, S. Soliveres, F. T. Maestre, Plant spatial patterns identify alternative ecosystem multifunctionality states in global drylands. *Nat. Ecol. Evol.* **1**, 3 (2017).
31. M. Rietkerk *et al.*, Evasion of tipping in complex systems through spatial pattern formation. *Science* **374**, eabj0359 (2021).
32. L.-X. Zhao *et al.*, Fairy circles reveal the resilience of self-organized salt marshes. *Sci. Adv.* **7**, eabe1100 (2021), 10.1126/sciadv.abe1100.
33. P. J. Clark, F. C. Evans, Distance to nearest neighbor as a measure of spatial relationships in populations. *Ecology* **35**, 445–453 (1954).
34. BESAG, J, Contribution to the discussion on Dr Ripley's paper. *J. R. Stat. Soc.* **39**, 193–195 (1977).
35. L. F. Hou, S. P. Gao, G. Q. Sun, Two types of fairy circles coexist in a vegetation–water model. *Nonlinear Dyn.* **111**, 7883–7898 (2023).
36. F. Aurenhammer, Voronoi diagrams–A survey of a fundamental geometric data structure. *ACM Comput. Surv.* **23**, 345–405 (1991).
37. G. Shmueli, To explain or to predict? *Stat. Sci.* **25**, 289–310 (2010).
38. M. D. Picker, V. Ross-Gillespie, K. Vlieghe, E. Moll, Ants and the enigmatic Namibian fairy circles–Cause and effect? *Ecol. Entomol.* **37**, 33–42 (2012).
39. M. Belda, E. Holtanová, T. Halenka, J. Kalvová, Climate classification revisited: From Köppen to Trewartha. *Clim. Res.* **59**, 1–13 (2014).
40. R. H. Whittaker, *Communities and Ecosystems* (Macmillan Publishing, 1975).
41. P. Rengasamy, Transient salinity and subsoil constraints to dryland farming in Australian sodic soils: An overview. *Aust. J. Exp. Agric.* **42**, 351 (2002).
42. J. M. Paruelo, H. E. Epstein, W. K. Lauenroth, I. C. Burke, ANPP Estimates from NDVI for the central grassland region of The United States. *Ecology* **78**, 953–958 (1997).
43. S. Liu *et al.*, Modeling aboveground biomass of an alpine desert grassland with SPOT-VGT NDVI. *GISci. Remote Sens.* **52**, 680–699 (2015).
44. Z. Zang, X. Wu, Y. Niu, G. Mao, Analysis of the contributions of human factors and natural factors affecting the vegetation pattern in coastal wetlands. *Ecosyst. Health Sust.* **6**, 1827982 (2020).
45. E. Dinerstein, *et al.*, An ecoregion-based approach to protecting half the terrestrial realm. *BioScience* **67**, 534–545 (2017).
46. R. Drikvandi, G. Verbeke, A. Khodadadi, V. Partovi Nia, Testing multiple variance components in linear mixed-effects models. *Biostatistics* **14**, 144–159 (2013).
47. V. Deblauwe, P. Couteron, J. Bogaert, N. Barbier, Determinants and dynamics of banded vegetation pattern migration in arid climates. *Ecol. Monogr.* **82**, 3–21 (2012).
48. V. Deblauwe, N. Barbier, P. Couteron, O. Lejeune, J. Bogaert, The global biogeography of semi-arid periodic vegetation patterns. *Glob. Ecol. Biogeogr.* **17**, 715–723 (2008).
49. A. Bey *et al.*, Collect Earth: Land use and land cover assessment through augmented visual interpretation. *Remote Sens.* **8**, 807 (2016).
50. J. Ferner, S. Schmidtlein, R. T. Guuroh, J. Lopatin, A. Linstädter, Disentangling effects of climate and land-use change on West African drylands' forage supply. *Glob. Environ. Change* **53**, 24–38 (2018).
51. P. D'Odorico, F. Laio, L. Ridolfi, Patterns as indicators of productivity enhancement by facilitation and competition in dryland vegetation. *J. Geophys. Res.* **111**, G03010 (2006), 10.1029/2006jg000176.
52. J. Bonachela *et al.*, Termite mounds can increase the robustness of dryland ecosystems to climatic change. *Science* **347**, 651–655 (2015).
53. M. Berdugo *et al.*, Global ecosystem thresholds driven by aridity. *Science* **367**, 787–790 (2020).
54. J. Puigdefábregas, The role of vegetation patterns in structuring runoff and sediment fluxes in drylands. *Earth Surf. Process. Landf.* **30**, 133–147 (2005).
55. H. A. Diallo, The United Nations Convention to combat desertification. Villages in the future (2001), pp. 173–174.
56. A. Trabucco, R. J. Zomer, Data from "Global aridity index and potential evapotranspiration (ET0) climate database v2." Figshare. https://doi.org/10.6084/m9.figshare.7504448.v3. Accessed 15 January 2022.
57. UNEP-WCMC., A spatial analysis approach to the global delineation of dryland areas of relevance to the CBD Programme of Work on Dry and Subhumid Lands. Dataset based on spatial analysis between WWF terrestrial ecoregions (WWF-US, 2004) and aridity zones (CRU/UEA; UNEPGRID, 1991). Dataset checked and refined to remove many gaps, overlaps and slivers (July 2014) (2007).
58. P. Tao *et al.*, Recognition of ecological vegetation fairy circles in intertidal salt marshes from UAV LiDAR point clouds. *Int. J. Appl. Earth Obs. Geoinf.* **114**, 103029 (2022).
59. M. Al-Sarayreh, Z. Moayed, B. Bollard-Breen, J.-B. Ramond, R. Klette, *2016 International Conference on Image and Vision Computing New Zealand (IVCNZ)* (IEEE, 2016). https://ieeexplore.ieee.org/document/7804457/.
60. Y. Zhu *et al.*, *2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)* (IEEE, 2018). https://ieeexplore.ieee.org/document/8639450/.
61. C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, Z. Wojna, *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2016), pp. 2818–2826. https://dx.doi.org/10.1109/CVPR.2016.308.
62. S. Tabik, D. Peralta, A. Herrera-Poyatos, F. Herrera, A snapshot of image pre-processing for convolutional neural networks: Case study of MNIST. *Proc. Int. Joint Conf. Bioinforma. Syst. Biol. Intell. Comput.* **10**, 555 (2017).
63. D. Marmanis, M. Datcu, T. Esch, U. Stilla, Deep learning earth observation classification using ImageNet pretrained networks. *IEEE Geosci. Remote Sens. Mag.* **13**, 105–109 (2016).
64. L. Fei-Fei, J. Deng, K. Li, ImageNet: Constructing a large-scale image database. *J. Vis.* **9**, 1037–1037 (2010).
65. I. Karabayir, O. Akbilgic, N. Tas, A novel learning algorithm to optimize deep neural networks: Evolved gradient direction optimizer (EVGO). *IEEE Trans. Neural Netw. Learn Syst.* **32**, 685–694 (2021).
66. L. Yu, P. Gong, Google Earth as a virtual globe tool for Earth science applications at the global scale: Progress and perspectives. *Int. J. Remote Sens.* **33**, 3966–3986 (2012).
67. D. M. W. Powers, Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation. arXiv [Preprint] (2020). https://arxiv.org/abs/2010.16061 (Accessed 15 January 2023).
68. A. P. Bradley, The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognit.* **30**, 1145–1159 (1997).
69. K. Donnelly, "Simulations to determine the variance and edge-effect of total nearest neighbour distance" in *Simulation Studies in Archaeology*, I. Hodder, Ed. (Cambridge University Press, Cambridge/New York, 1978), pp. 91–95.
70. K. H. Hanisch, Some remarks on estimators of the distribution function of nearest neighbour distance in stationary spatial point processes. *Ser. Stat.* **15**, 409–412 (1984).
71. M. Myllymäki, T. Mrkvička, P. Grabarnik, H. Seijo, U. Hahn, Global envelope tests for spatial processes. *J. R. Stat. Soc. Series B Stat. Methodol.* **79**, 381–404 (2017).
72. M. Häfner, M. Liedlgruber, A. Uhl, A. Vécsei, F. Wrba, Delaunay triangulation-based pit density estimation for the classification of polyps in high-magnification chromo-colonoscopy. *Comput. Methods Programs Biomed.* **107**, 565–581 (2012).
73. A. Vargha, H. D. Delaney, The Kruskal–Wallis test and stochastic homogeneity. *J. Educ. Behav. Stat.* **23**, 170–192 (1998).
74. QGIS Development Team, *QGIS Geographic Information System* (Open Source Geospatial Foundation, 2021). https://qgis.org.
75. E. Guirado *et al.*, Location of the sample plots with FC-like vegetation patterns for the distance analysis. Figshare (2023), 10.6084/m9.figshare.17069954. Deposited 15 August 2023.
76. M. W. Van Rooyen, G. K. Theron, N. Van Rooyen, W. J. Jankowitz, W. S. Matthews, Mysterious circles in the Namib Desert: Review of hypotheses on their origin. *J. Arid Environ.* **57**, 467–485 (2004).
77. T. Hamazaki, Advanced land observation satellite (ALOS). 5 Outline of ALOS satellite system. *J. Japan Soc. Photogram. Remote Sens.* **38**, 25–26 (1999).
78. S. E. Fick, R. J. Hijmans, WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* **37**, 4302–4315 (2017).
79. J. T. Abatzoglou, S. Z. Dobrowski, S. A. Parks, K. C. Hegewisch, TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958–2015. *Sci Data.* **5**, 170191 (2018).
80. T. Tachikawa, M. Hato, M. Kaku, A. Iwasaki, *2011 IEEE International Geoscience and Remote Sensing Symposium* (IEEE, 2011). https://ieeexplore.ieee.org/document/6050017/.
81. T. Hengl *et al.*, SoilGrids250m: Global gridded soil information based on machine learning. *PLoS ONE* **12**, e0169748 (2017).
82. LP DAAC - MOD13Q1, https://lpdaac.usgs.gov/products/mod13q1v006/.
83. S. Alibakhshi, T. W. Crowther, B. Naimi, Land surface black-sky albedo at a fixed solar Zenith angle and its relation to forest structure during peak growing season based on remote sensing data. *Data Brief.* **31**, 105720 (2020).

84. V. Zlotnicki, S. Bettadpur, F. W. Landerer, M. M. Watkins, "Gravity recovery and climate experiment (GRACE): Detection of ice mass loss, terrestrial mass changes, and ocean mass gains" in *Encyclopedia of Sustainability Science and Technology*, R. A. Meyers, Eds. (Springer, New York, NY, 2012). 10.1007/978-1-4419-0851-3_745.
85. B. D. Tapley *et al.*, Contributions of GRACE to understanding climate change. *Nat. Clim. Change* **9**, 358–369 (2019).
86. T. R. Loveland *et al.*, An analysis of the IGBP global land-cover characterization process. *Photogramm. Eng. Remote Sens.* **65**, 1021–1032 (1999).
87. J. Badger *et al.*, Global Wind Atlas 3.0, a free, web-based application developed, the Technical University of Denmark (2019). https://globalwindatlas.info.
88. P. Eggleton, P. H. Williams, K. J. Gaston, Explaining global termite diversity: Productivity or history? *Biodivers. Conserv.* **3**, 318–330 (1994).
89. P. H. Williams, WORLDMAP: Priority areas for biodiversity. User manual for program version 3 (1992).
90. P. Gilbert, J. M. V. Harvey, K. S. Martin, "A catalogue of manuscripts" in *The Entomology Library of the Natural History Museum*,(Mansell Publishing Limited, London, UK, 1996), 10.5962/bhl.title.87389.
91. B. Guenard, M. D. Weiser, K. Gomez, N. Narula, E. P. Economo, The Global Ant Biodiversity Informatics (GABI) database: Synthesizing data on the geographic distribution of ant species (Hymenoptera: Formicidae). *Myrmecol. News/Osterr. Ges. Entomofaunistik* **24**, 83–89 (2017).
92. B. Benito, BlasBenito/spatialRF: SpatialRF: Easy Spatial regression with random forest (2021), 10.5281/zenodo.4745208.
93. W. Yoo *et al.*, A study of effects of multicollinearity in the multivariable analysis. *Int. J. Appl. Sci. Technol.* **4**, 9–19 (2014).
94. C. F. Dormann *et al.*, Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography* **36**, 27–46 (2013).
95. K. M. Marcoulides, T. Raykov, Evaluation of variance inflation factors in regression models using latent variable modeling methods. *Educ. Psychol. Meas.* **79**, 874–882 (2019).
96. N. Yuliantoro, V. Goeltom, I. B. Juliana, R. Pramono, A. Purwanto, Repurchase intention and word of mouth factors in the millennial generation against various brands of Boba drinks during the Covid 19 pandemic. *Afr. J. Hosp. Tour.* **8**, 1–11 (2019).
97. J. A. Nelder, R. W. Wedderburn, Generalized linear models. *J. R. Stat. Soc. Series A Gen.* **135**, 370–384 (1972).
98. B. Greenwell, B. Boehmke, B. Gray, Package 'vip'. Variable Importance Plots (2020), vol. **12**, pp. 343–366.
99. L. Breiman, Random forests. *Mach. Learn.* **45**, 5–32 (2001).
100. K. K. Nicodemus, J. D. Malley, C. Strobl, A. Ziegler, The behaviour of random forest permutation-based variable importance measures under predictor correlation. *BMC Bioinf.* **11**, 110 (2010).
101. M. N. Wright, A. Ziegler, ranger: A fast implementation of random forests for high dimensional data in C++ and R. *J. Stat. Softw.* **77**, 1–17 (2017).
102. L. V. Hedges, J. Gurevitch, P. S. Curtis, The meta-analysis of response ratios in experimental ecology. *Ecology* **80**, 1150–1156 (1999).
103. A. Lahouar, J. Ben Hadj Slama, Day-ahead load forecast using random forest and expert input selection. *Energy Convers. Manage.* **103**, 1040–1051 (2015).
104. R.R. Kohavi, *Proceedings of the 14th International Joint Conference on Artificial Intelligence (IJCAI'95)* (Morgan Kaufmann Publishers Inc., San Francisco, CA, 1995), vol. **2**, pp. 1137–1143.
105. G. Piñeiro, S. Perelman, J. P. Guerschman, J. M. Paruelo, How to evaluate models: Observed vs. predicted or predicted vs. observed? *Ecol. Model.* **216**, 316–322 (2008).
106. B. P. Mallavan, B. Minasny, A. B. McBratney, "Homosoil, a methodology for quantitative extrapolation of soil information across the globe" in *Digital soil mapping progress in soil science*, J. L. Boettinger, D. W. Howell, A. C. Moore, A. E. Hartemink, S. Kienast-Brown, Eds. (Springer, Academia, 2010), pp. 137–150.
107. M. Gilbert *et al.*, Global cattle distribution in 2015 (5 minutes of arc) Harvard Dataverse, V1 (2022).
108. B. L. Markham, W. C. Boncyk, D. L. Helder, J. L. Barker, Landsat-7 enhanced thematic mapper plus radiometric calibration. *Can. J. Remote Sens.* **23**, 318–332 (1997).
109. Z. Zhu, S. Wang, C. E. Woodcock, Improvement and expansion of the Fmask algorithm: Cloud, cloud shadow, and snow detection for Landsats 4–7, 8, and Sentinel 2 images. *Remote Sens. Environ.* **159**, 269–277 (2015).
110. S. Liu *et al.*, Phylotype diversity within soil fungal functional groups drives ecosystem stability. *Nat. Ecol. Evol.* **6**, 900–909 (2022).
111. S. Kim, ppcor: An R package for a fast calculation to semi-partial correlation coefficients. *Commun. Stat. Appl. Methods* **22**, 665 (2015).
112. E. Guirado *et al.*, Locations of fairy circle (FC)-like vegetation patterns and non-FC in global drylands. This dataset includes with 574,799 locations separated by 10 km distance between them. Dataset S1. https://figshare.com/s/ed9e5e07f30c6fe8bfd5. Deposited 9 March 2023.
113. E. Guirado *et al.*, Localization of 15,032 images with fairy circles (FC) located in Namibia and Australia obtained from areas reported in the literature. A binary system (7516 images for FC and 7516 for non-FC) was used to train the CNN-based model. Dataset S2. https://figshare.com/s/5c18dab9d697ee553ffe. Deposited 9 March 2023.
114. E. Guirado *et al.*, The dataset used for Random Forest analysis with 526 locations, 263 for fairy circles (FC)-like vegetation patterns and 263 for non-FC. Dataset S3. https://figshare.com/s/d6ae64213be36e30e160. Deposited 9 March 2023.
115. E. Guirado *et al.*, Location of the sample plots with FC-like vegetation patterns for the distance analysis. The data are in shapefile format. Dataset S4. https://figshare.com/s/7e0d9860cd170165adf5. Deposited 15 August 2023.
116. E. Guirado *et al.*, Source code in R language for the analyses conducted in this study (Random Forest, GLM, Multicollinearity, L-function, PCF and LMe analyses). Software S1. https://figshare.com/s/69f1f0d9ee363977e83f. Deposited 29 May 2023.