Received: 26 June 2019 | Revised: 18 October 2019 | Accepted: 30 October 2019
DOI: 10.1111/geb.13039

META-ANALYSIS

Global Ecology and Biogeography A Journal of Macroecology
WILEY

Deciphering the stability of grassland productivity in response to rainfall manipulation experiments

Ilaíne Silveira Matos[1] | **Imma Oliveras Menor**[2] | **Sami Walid Rifai**[2] | **Bruno Henrique Pimentel Rosado**[3]

[1]Ecology and Evolution Graduate Program, IBRAG, Universidade do Estado do Rio de Janeiro, PHLC, Rio de Janeiro, Rio de Janeiro, Brazil

[2]Environmental Change Institute, School of Geography and the Environment, University of Oxford, Oxford, United Kingdom

[3]Department of Ecology, IBRAG, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Rio de Janeiro, Brazil

Correspondence
Ilaíne Silveira Matos, Ecology and Evolution Graduate Program, Universidade do Estado do Rio de Janeiro, Rua São Francisco Xavier, IBRAG, PHLC, sala 202, 20550900, Rio de Janeiro, Rio de Janeiro, Brazil.
Email: ilaine.matos@gmail.com

Funding information
Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro, Grant/Award Number: Bolsa Jovem Cientista do Nosso Estado-JCNE, E-; Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Grant/Award Number: 88882.182435/2018-01; The Rufford Foundation, Grant/Award Number: 18749-1; Rufford Foundation, Grant/Award Number: 26, 203, 199, 2016, 88882, 182435 and 2018-01; Coordenação de Aperfeiçoamento de Pessoal de Nível Superior

Editor: John-Arvid Grytnes

Abstract

Aim: Rainfall manipulation experiments are essential tools for deciphering the mechanisms leading to variation in ecosystem stability across sites. Here, we gathered articles reporting results of experimental droughts on the above-ground biomass of grasslands to identify which indices have been used to assess stability, to evaluate the overall grassland responses to drought and to quantify the relative importance of drought characteristics and climatic conditions for explaining variation in stability.

Location: Global.

Time period: 1989–2018.

Major taxa studied: Grasslands.

Methods: We used meta-analytical approaches to evaluate overall grassland stability in terms of resistance, recovery and resilience, and multi-model inference to assess the relative importance of different moderators on explaining the variability of those three stability properties.

Results: Numerous indices of stability have been used, but they are inadequate for comparisons across sites. After applying standardized indices, we found that grasslands were resilient (biomass remained unchanged 1 year after drought) and exhibited a trade-off between low resistance (biomass was lost during drought) and high recovery (new biomass was produced after drought). Overall, climatic conditions and drought characteristics (intensity, duration and frequency) were not important to explain the differences in stability observed across grasslands.

Main conclusions: Grasslands are resilient, but if drought events last > 1 year, there might be long-term declines of biomass production owing to incomplete recovery. Despite the hundreds of experiments conducted in grasslands across the globe, the results are still inconclusive because of four important shortcomings: 50% of the studies have failed to create drought; 81% have not included recovery and resilience, assessing only resistance; 87% have not applied quantitative indices to assess stability; and < 1% of the studies were conducted on tropical grasslands. We discuss how to overcome those limitations to improve our ability to ensure stable grassland productivity under climate change.

KEYWORDS
extreme climatic events, meta-analysis, rain-out shelters, recovery, resilience, resistance

1 | INTRODUCTION

Grasslands cover more than one-fifth of the global land surface (Dixon, Faber-Langendoen, Josse, Morrison, & Loucks, 2014), and their productivity is strongly controlled by precipitation (Knapp & Smith, 2001). Given that the frequency, duration and intensity of droughts are predicted to increase (Dai, 2013), hundreds of experiments (Figure 1) have been conducted to investigate how changes in the amount and/or seasonality of rainfall could affect the biomass production of grasslands (Hoover, Wilcox, & Young, 2018) and, thereby, carbon sequestration and forage provision (Forrestel et al., 2017).

Those experiments, however, do not report consistent results (Stuart-Haëntjens et al., 2018; Wilcox et al., 2017). Some grasslands exhibited a rapid decline in productivity (Tilman & Haddi, 1992), whereas others maintained stable biomass production even after a decade of rainfall reduction (Grime et al., 2000). The underlying mechanism explaining such differences in ecosystem stability has not yet been identified (Hoover et al., 2018; Wilcox et al., 2017), and it remains one of the most fundamental questions in ecology (Sutherland et al., 2013).

Differences across sites could arise simply because of the lack of a common approach to define and to measure ecosystem stability. Stability has > 163 definitions catalogued (Grimm & Wissel, 1997) and can be measured through several different metrics, depending on the theoretical framework and the questions asked (Quinlan, Berbés-Blázquez, Haider, Peterson, & Allen, 2016). In the context of rainfall manipulation experiments, stability metrics have been applied to describe at least three distinct properties: resistance (RT), recovery (RC) and resilience (RS) (Figure 1). Resistance can be defined as the ability of the system to persist in its initial state during a perturbation, and thus it describes the immediate impacts of experimental droughts (Ingrisch & Bahn, 2018; Stuart-Haentjens et al., 2018). A grassland would be considered resistant, in terms of biomass production, if it did not lose biomass during drought (i.e., if the biomass in control and drought plots measured immediately after drought cessation were not significantly different). Recovery reflects the ability of the grassland to produce new biomass after drought cessation, and it can be obtained by comparing biomass in the drought plots at two different times: immediately after drought and, usually, 1 year after drought cessation (post-drought) (Carter & Blair, 2012). A grassland would be considered able to recover if the post-drought biomass was higher than the biomass obtained immediately after drought. Finally, resilience describes the ability of the system to persist in its initial state after a perturbation has ceased (Ingrisch & Bahn, 2018) and can be obtained by comparing biomass in control and drought plots during the post-drought period. A grassland would be considered resilient if the biomass did not differ significantly between those plots.

Resilience might be achieved either by a grassland that does not lose biomass during drought (high resistance) or by a grassland that does but is able to produce new biomass during the post-drought period, thus compensating for what has been lost (high recovery) (Ingrisch & Bahn, 2018). Given that plant traits that allow a higher drought resistance (i.e., sustained biomass production during water stress) may hinder a rapid biomass recovery after drought alleviation (Craven et al., 2018; Májeková, de Bello, Doležal, & Lepš, 2014), a trade-off might be expected between resistance and recovery (Hodgson, McDonald, & Hosken, 2015); grasslands that resist droughts should be less able to recover, and vice versa. However, a lack of correlation could also occur if the mechanisms that enable a community to be resistant were uncoupled to those that confer recovery (Nimmo, Mac Nally, Cunningham, Haslem & Bennett, 2015) or, in more diverse communities, where species with high recovery and others with high resistance can co-occur (Craven et al., 2018; Vogel, Scherer-Lorenzen, & Weigelt, 2012). Particularly in grasslands, uncertainties remain about the potential trade-offs between resistance and recovery and their relative importance to underpin the overall ecosystem resilience.

The absence of standardized methodologies and protocols (Knapp et al., 2017; Smith, Wilcox, Power, Tissue, & Knapp, 2017) could also lead to a high variation in grassland responses across studies. For instance, at least three distinct types of experimental droughts have been imposed (Figure 1): (a) the seasonality of rainfall is altered without significant changes in total amount (seasonality; Figure 1a); (b) the amount of rainfall is reduced by a fixed percentage for a short and interrupted period of time (amount “press”; Figure 1b); or (c) rainfall is reduced for long and continuous periods of time (amount “pulse”; Figure 1c). There are studies showing that the productivity of grasslands can be affected differently by changes in rainfall seasonality versus in the amount of rainfall (Peng et al., 2013) and in response to “press”- versus “pulse”-like droughts (Hoover & Rogers, 2016).

In experiments that reduced the amount of rainfall, the percentage reduction of mean annual precipitation (%PPT) has commonly been used as a measure of drought treatment intensity; however, it might be inadequate for comparing experimental droughts conducted across different years (Figure 1c; Hoover et al., 2018; Kreyling, Arfin Khan, et al., 2017; Vicca et al., 2012). For example, in Figure 1c, although the %PPT remained the same throughout the whole experimental period (−50%), larger absolute differences in precipitation between control and drought plots were observed in wetter (Drought 1) than in drier years (Drought 2). Moreover, when comparing stability across distinct grasslands, if the interannual variation in precipitation at one site is greater than that at another site, a drought of −50% intensity is likely to have less effect on the former site (Hoover et al., 2018; Kreyling, Arfin Khan, et al., 2017; Vicca et al., 2012). Therefore, for a proper comparison of drought treatment across studies it is necessary to use better indices of drought intensity, such as the standardized precipitation evapotranspiration index (SPEI). The SPEI represents the number of standard deviations by which the climatic water balance (precipitation minus potential evapotranspiration) differed from the monthly mean historical record for a given site (Vicente-Serrano, Beguería, & López-Moreno, 2010) and, thus, it can be used to characterize drought intensity irrespective of differences in annual precipitation.

Besides reflecting differences in the indices and/or in the characteristics of the experimental droughts applied, the observed variation in stability across grasslands might also be attributed to the distinct

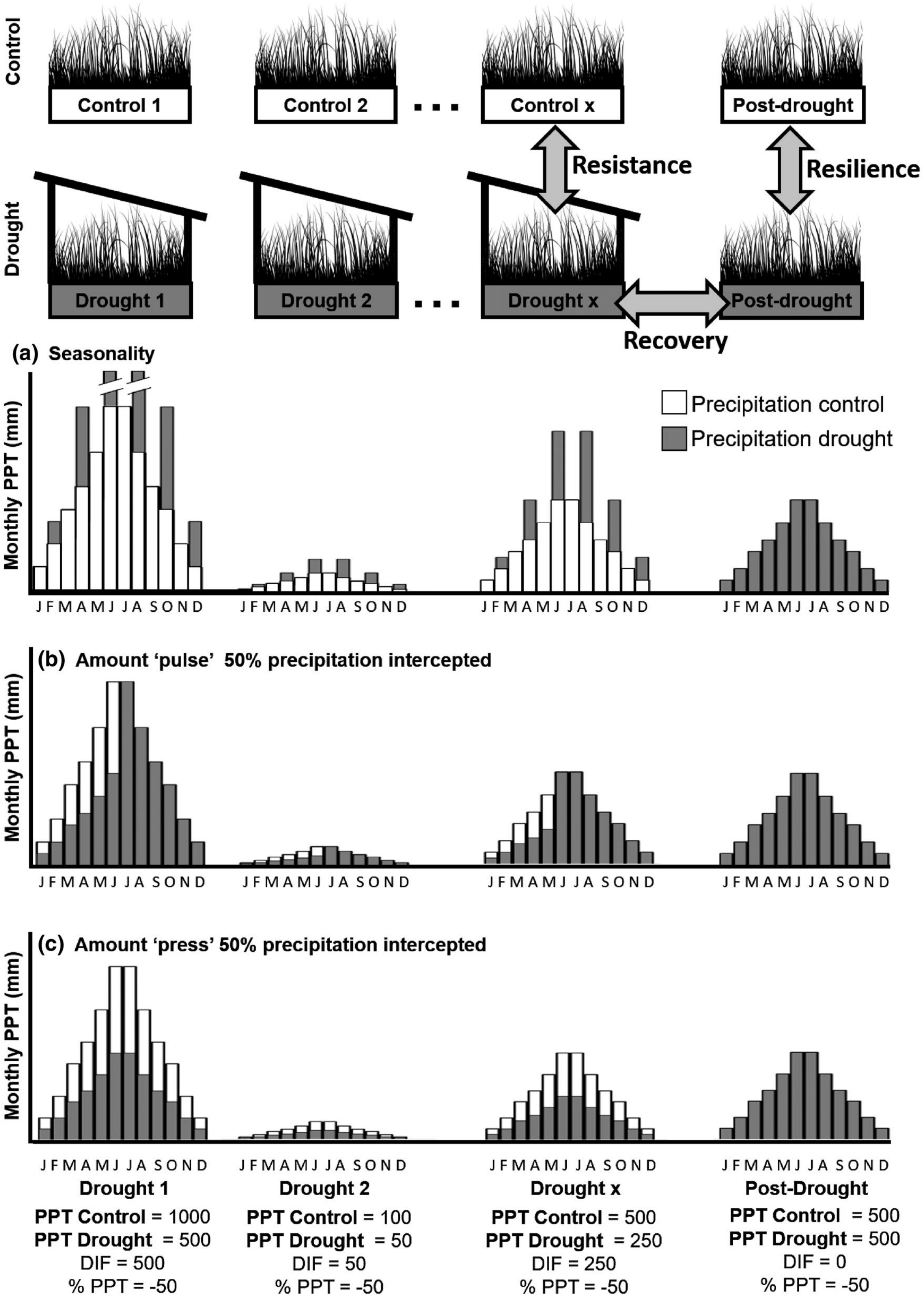

FIGURE 1 Overview of rainfall manipulation experiments. Rain-out shelters have been used to intercept distinct percentages of rainfall (%PPT), thus creating experimental drought events (in drought plots) during one or more sequential years ($n = 1, 2, \ldots, x$). Response variables, most frequently above-ground biomass, are measured on control plots and drought plots during the drought and 1 year after drought cessation (post-drought). Biomass data are used to estimate three stability properties: resistance (comparing biomass in drought versus control plots immediately after drought cessation); recovery (comparing biomass in drought versus post-drought plots); and resilience (comparing biomass in drought and control plots during the post-drought period). Monthly precipitation (PPT; in millimetres) recorded for control and drought plots (during drought and post-drought periods) is shown for three types of rainfall manipulation experiments: (a) seasonality: rainfall is repacked into larger events, thus resulting in longer dry intervals without changes in total mean annual precipitation; (b) amount "pulse": rainfall amount is reduced in drought plots by a fixed percentage during shorter and interrupted periods of time (e.g., growing season); and (c) amount "press": rainfall is reduced during longer and uninterrupted periods of time (e.g., all the year). Abbreviations: DIF: absolute differences in precipitation between control and drought plots; %PPT: percentage reduction of mean annual precipitation

climatic conditions experienced by plants (Kreyling, Arfin Khan, et al., 2017; Smith et al., 2017). For example, we could expect colder sites to be more stable than warmer sites, because higher temperatures can exacerbate the negative effects of drought by increasing vapour pressure deficit, inducing stomatal closure, hence reducing photosynthesis to support biomass production (De Boeck, Dreesen, Janssens, & Nijs, 2011). Likewise, mesic and moist sites could be more stable, because biomass production in xeric systems is already limited by water availability (Huxman et al., 2004). Conversely, xeric species are able to avoid and/or tolerate drier conditions, hence xeric grasslands could be more resistant to droughts (Knapp & Smith, 2001). Finally, sites with higher historical rainfall variability could be more stable in response to experimental droughts (Ciemer et al., 2019).

Here, we searched for published studies reporting effects of rainfall manipulation experiments on grasslands to identify which indices have been used to assess stability and whether they provide comparable results (i.e., whether rankings of grassland stability are consistent among different indices). We also evaluated the overall resistance, recovery and resilience of grasslands in response to experimental droughts and identified potential trade-offs between those properties. Finally, we quantified the relative importance of drought experimental characteristics (intensity, duration and frequency) and prevailing climatic conditions (temperature, precipitation and rainfall seasonality) in explaining differential stability across grasslands.

2 | MATERIALS AND METHODS

2.1 | Data search, inclusion criteria and metadata extraction

We searched for peer-reviewed scientific articles published before June 2018 that conducted rainfall manipulation experiments in grasslands. The search was conducted in two Web databases: Google Scholar [search string: (in subject: *grassland* in title: (*resilience* OR *resistance* OR *recovery* OR *stability* OR *vulnerability*) AND (*rain shelter* OR *rain exclusion* OR *rain manipulation* OR *drought* OR *precipitation*) AND (*plant* OR *tree* OR *grass* OR *vegetation*)] and ISI Web of Science [search string: (*grassland* AND (*drought* OR *rain** OR *precipitation*) AND (*resistance* OR *resilience* OR *recovery* OR *stability* OR *vulnerability*) AND (*plant** OR *tree* OR *grass* OR *shrub* OR *veget**)] and resulted in 278 relevant references (see Supporting Information Appendix S1). We only retained studies conducted in grasslands (i.e., herbs comprising > 50% of plant cover) composed by two or more species. We excluded experiments conducted in artificial conditions (e.g., greenhouses, growth chambers or individual pots) and in which precipitation was altered by means other than rain-out shelters. After cross-referencing, removal of duplicates, and application of the above set of inclusion criteria, we retained a total of 127 articles (204 study cases) published during 1989–2018 (a list of the data sources can be found in Supporting Information Appendix S2). From each study, we extracted metadata and recorded whether those studies computed indices of stability (i.e., resistance, recovery and/or resilience). Metadata were extracted directly from text or tables, or indirectly from figures using the Figure Calibration plugin in the ImageJ software, v.1.48. When elevation, mean annual precipitation (MAP) and/or mean annual temperature (MAT) were not provided in the original studies, they were derived using Geoplaner (http://www.geoplaner.com) and WorldClim (http://www.worldclim.org), respectively.

Subsequently, we applied a second set of inclusion criteria, retaining only those studies that provided sample sizes and above-ground biomass mean ± *SD* [or *SE* or confidence interval (CI)] for both control and drought plots. We decided to assess stability only in terms of above-ground biomass, because it was the response variable most frequently reported (83%; $n = 170$). In addition, we eliminated experiments altering only rainfall seasonality (as in Figure 1a) because of their limited sample size ($n = 23$) and particular experimental design, which precluded comparisons with studies that reduced the amount of rainfall. We retained experiments that reduced the amount of rainfall by applying both "press"- (Figure 1b) and "pulse"-like droughts (Figure 1c). After application of the second set of inclusion criteria, we retained a total 49 articles and 101 study cases (Supporting Information Appendix S2, Matos, Oliveras, Rifai, & Rosado, 2019).

2.2 | Drought intensity and rainfall seasonality calculation

For each of the 101 study cases, we calculate the SPEI as a measure of drought intensity. To calculate SPEI, we first obtained, for each study location, the monthly precipitation and potential evapotranspiration data from the TerraClimate product, which provides data of climatic water balance from 1958 to 2017 at a spatial resolution 2.5 arc min (*c*. 4 km; Abatzoglou, Dobrowski, Parks, & Hegewisch, 2018). Secondly, the approximate monthly precipitation inputs for the control and drought plots during the whole drought period were determined based on the information provided by each study (i.e., %PPT and drought starting and ending dates) and then the precipitation values in the original climatic dataset were altered accordingly. For instance, for a study that would have applied an experimental drought by intercepting 50% of rainfall from March to September 2005, we would simply halve the monthly precipitation values of TerraClimate for that specific period. Subsequently, we used the "SPEI" R package (Vicente-Serrano et al., 2010) to calculate the 1 month derivation of SPEI for the whole drought period. Negative/positive values of SPEI indicate drier/wetter conditions, as shown in Table 1.

The use of SPEI is justified by the fact that very few experiments provided the necessary data to calculate more meaningful metrics of drought intensity (Vicca et al., 2012), but there are two potential limitations to this approach. First, climatic water deficits do not necessarily lead to plant water deficits. Even if rainfall manipulation experiments succeed in translating precipitation deficits to soil water deficits, they might fail to alter the potential evapotranspiration in a manner similar to what occurs during real drought events (Beier et al., 2012; Kreyling, Arfin Khan, et al., 2017). To assess how the use of different indices (i.e., %PPT versus SPEI) could affect the assessment of grassland stability, we also classified drought intensity based on the %PPT values (Table 1) reported in the papers and compared mean effect sizes, for each category of drought intensity (i.e., non-drought, moderate, severe or extreme drought), using both SPEI and %PPT classifications (Supporting Information Appendix S6). Second, there could be error between the precipitation estimates extracted from TerraClimate and the real precipitation received at the experimental sites. To test this second potential limitation, we compared the MAP reported in the papers with MAP calculated from TerraClimate data.

Data obtained from TerraClimate were also used to compute rainfall seasonality (S; Feng, Porporato, & Rodriguez-Iturbe, 2013), which ranges from zero (no seasonality, with rainfall uniformly distributed throughout the year) to 3.58 (maximum seasonality, with rainfall concentrated in a single month); S was calculated as $D \times R/R_{max}$, where D is the relative entropy (i.e., a measure of the concentration of rainfall around the wet season), R is the long-term mean annual rainfall for each site, and R_{max} is the maximum mean annual rainfall in the dataset (R_{max} = 3,074 mm/year in ALPFOR research station, Switzerland; Feng et al., 2013).

TABLE 1 Classification of drought intensity based on standardized precipitation evapotranspiration indices (SPEI) (Tao, Borth, Fraedrich, Su, & Zhu, 2014) and percentage reduction of mean annual precipitation (%PPT) (Gao et al., 2019)

Category	SPEI	%PPT
Extreme drought	SPEI ≤ −2.00	%PPT ≥ 80
Severe drought	−1.50 ≤ SPEI ≤ −1.99	51 ≤ %PPT ≤ 79
Moderate drought	−1.00 ≤ SPEI ≤ −1.49	26 ≤ %PPT ≤ 50
Non-drought/near-normal conditions	SPEI ≥ −0.99	%PPT ≤ 25
Moderately wet	1.00 ≤ SPEI ≤1.49	–
Severely wet	1.50 ≤ SPEI ≤1.99	–
Extremely wet	SPEI ≥ 2.00	–

2.3 | Stability indices and effect sizes

To identify which indices of stability have been used and whether they provide comparable results, we first gathered all indices reported in the 127 articles. We then calculated the indices by using the biomass data provided in the subset of 101 study cases. Finally, we ranked grasslands according to their stability values and evaluated whether those rankings were consistent or varied depending on the indices used.

To assess overall grassland ecosystem stability in response to experimental droughts, we computed mean Hedges' g effect sizes for resistance, recovery and resilience (Borenstein, Hedges, Higgins, & Rothstein, 2009). We used mean effect sizes, instead of the stability indices reported in the papers, because in the former approach studies are weighted by their precision (more precise studies receive larger weights), thus providing a more unbiased overall effect (Borenstein et al., 2009). Hedges' g effect sizes for resistance were obtained as the difference of above-ground biomass between drought and control plots during the drought period, for recovery as the difference between post-drought and drought plots, and for resilience as the difference between drought and control plots during the post-drought period. Those differences were then divided by the pooled standard deviation and weighted by sample size. Consequently, effect sizes represented the number of standard deviations by which the treatment differed from the control group (Borenstein et al., 2009), and negative/positive values indicated decreases/increases in above-ground biomass in response to drought. Mean effect sizes for each stability property were estimated under random effect models using restricted maximum likelihood estimation and Knapp & Hartung adjustment within the "metaphor" R package (Viechtbauer, 2010). Linear regressions were used to test for potential trade-offs among stability properties. We checked graphically to ensure that the assumptions of normality of residuals and homogeneity of variance of the linear regressions were not violated.

2.4 | Effects of moderators

To verify whether grasslands varied in their stability in response to drought, we used Cochran's Q test, and we estimated the percentage of total variance that could be attributed to between-studies rather than to sampling error (I^2) (Borenstein et al., 2009). Significant differences from Cochran's Q tests ($p < .05$) and high percentages of between-studies heterogeneity (usually $I^2 > 50\%$) would indicate that grasslands varied significantly in their stability (Borenstein et al., 2009).

Subsequently, to quantify the relative importance of drought experimental features (duration, frequency and SPEI), prevailing climatic conditions (MAT and MAP, S) and elevation to explain the differential stability across grasslands, we used multi-model inference (Burnham & Anderson, 2002). In this analysis, all possible unique models involving the above set of moderators were fitted using random-effect meta-regression models and, owing to small sample size, they were ranked according to the Akaike's second-order corrected Information Criterion (AICc). For each moderator, we computed the model-averaged estimate and the relative importance. Relative importance was obtained by summing the Akaike weights of all models that included the moderator of interest. Values of relative importance > .8 indicated that the moderator was important to explain the variation in the stability property (Burnham & Anderson, 2002). Separate models were constructed for resistance, recovery and resilience using functions from the "glmulti" (Calcagno & Mazancourt, 2010) and "metaphor" R packages (Viechtbauer, 2010). Additionally, we fitted simple meta-regression models to evaluate the effect of moderators related to other experimental characteristics (vegetation type, shelter area, shelter type and control plot type), which are shown in the Supporting Information (Appendix S5).

2.5 | Sensitivity analysis

We performed a sensitivity analysis to test the influence of outliers, the non-independence effects, publication bias and effects related to the length of the post-drought period and to the type of drought intensity index (SPEI or %PPT). Sensitivity analyses are described in detail in the Supporting Information (Appendix S6).

3 | RESULTS

3.1 | Studies have imposed experimental droughts with very different characteristics, but some of them have failed to create a drought treatment

Metadata summary statistics are presented in Table 2 (see also Supporting Information Appendices S3 and S4). Rainfall manipulation experiments were conducted in 22 countries and 116 localities (Figure 2), particularly at mid- to high latitudes (30–60°) in the Northern Hemisphere (95%; $n = 110$). Droughts were imposed using rain-out shelters of different sizes (either automatic or permanent), mainly to reduce the amount of rainfall rather than to alter rainfall seasonality. Most studies conducted short-term experiments (< 1 year), and only 9% of them ($n = 4$) lasted > 10 years. Likewise, most studies imposed a single drought event (46%; $n = 95$), with a maximum of 15 sequential droughts (Fridley, Grime, Askew, Moser, & Stevens, 2011).

Usually, drought intensity was reported as the percentage reduction of MAP (%PPT), which ranged from 8 to 80% (Supporting Information Appendix S4). We found a reasonable agreement between the MAP values reported in the papers and those calculated from TerraClimate data (simple linear regression: $R^2 = 0.75$; d.f. = 99; $p < .0001$), with only three "outlier" studies (Cantarel, Bloor, & Soussana, 2013; De Boeck, Bassin, Verlinden, Zeiter, & Hiltbrunner, 2016; Hoeppner & Dukes, 2012), thus validating the SPEI as a reliable index of drought intensity.

Surprisingly, when we measured drought intensity using SPEI (Supporting Information Appendix S4), we found that in 50% of the studies ($n = 55$) the "drought" treatment did not experience a deficit in water balance, because the SPEI values were still in the range of near-normal conditions (Supporting Information Appendix S4). Moderate droughts were imposed in 28% ($n = 31$) of the studies, severe droughts in 14% ($n = 16$), and extreme drought events were conducted in only 8% ($n = 9$) of the studies. Different results were obtained when drought intensity was classified using %PPT (Supporting Information Table S6.2). Given that the %PPT classification disregarded the intra and interannual variability in precipitation, it overestimated drought intensity, especially for moderate droughts (Supporting Information Table S6.2 in Appendix S6). Regarding rainfall seasonality, in most of the sites the rainfall was well distributed throughout the year, with S values close to zero (Table 2; Supporting Information Appendix S4).

3.2 | Numerous indices have been used to measure ecosystem stability, limiting comparability across studies

Only 13% ($n = 17$) of the studies applied stability indices to assess grassland resistance, recovery and/or resilience in response to drought. We grouped those indices into six groups (Table 3), according to the type of stability property assessed [resistance (RT), recovery (RC), resilience (RS) or relative resilience (RRS)] and the way in which the stability was quantified (group 1, raw difference; group 2, simple ratio; groups 3 and 4, compounded ratio; group 5, natural logarithm ratio; group 6, standardized ratio). Although none of the selected studies had measured recovery as the raw difference (RC1) or compounded ratio (RC3 and RC4), nor resilience as the compounded ratio (RS3), we added those indices in our study to enable a full comparison among the six groups of indices. Thus, we obtained 19 stability indices in total (Table 3).

Those indices did not provide comparable results (Figure 3). When grasslands were ranked according to their stability values, the rankings varied depending on the group of indices considered. To illustrate better how and why results from different stability indices might not be comparable, we highlight in Figure 3 four study cases where grasslands showed distinct stability responses: Carter & Blair, 2012 (yellow); Yahdjian & Sala, 2006 (grey); Ingrisch et al., 2018 (red); and Shinoda, Nachinshonhor, & Nemoto, 2010 (blue). To facilitate understanding, we refer to these studies by their corresponding colours.

In group 1 indices, stability was obtained merely as raw differences in above-ground biomass between treatment and control plots. Consequently, results for those indices might be influenced strongly by the initial state of the system. That occurs because communities with higher initial biomass would always have more biomass to lose during drought and more biomass to recover

TABLE 2 Metadata extracted from rainfall manipulation experiments conducted on grasslands across the globe: 127 articles (204 study cases) that met the first set of inclusion criteria, 49 articles (101 study cases) that met the second set of inclusion criteria, and 13 articles (19 study cases) that reported above-ground biomass values for both drought and post-drought periods

Variables	127 articles (204 study cases)	49 articles 101 study cases)	13 articles (19 study cases)
Elevation (m a.s.l.)	612 (5–3,570)	530 (5–3,570)	500 (13–2,223)
MAT (°C)	8.5 (0.1–22.7)	8.7 (0.1–22.0)	9.0 (0.1–18.0)
MAP (mm/year)	771 (163–2,050)	680 (163–1,780)	587 (163–1,097)
Drought duration (days)	244 (17–2,310)	242 (17–2,310)	282 (38–2,310)
Drought frequency (*n*)	2 (1–15)	1 (1–12)	1 (1–12)
SPEI	–	−0.99 (−2.28 to 0.43)	−0.89 (−1.88 to −0.05)
Rainfall seasonality	–	0.02 (0.003–0.197)	0.03 (0.005–0.090)
Shelter area (m^2)	9 (0.5–900)	6; (0.5–900)	5.85; (0.5–900)
Vegetation type (%)	N: 40; S: 40; M: 20	N: 33; S: 45; M: 22	N: 53; S: 31; M: 16
Shelter type (%)	P: 91; A: 9	P: 94; A: 6	P: 89; A: 11
Control plot (%)	H:16; A: 84	H:19; A: 81	H: 26; A: 74
Drought type (%)	A: 90; S: 6; B: 4	A: 100; S: 0; B: 0	A: 100; S: 0; B: 0

Note: Values indicate the median and range for numerical variables and the percentage of study cases for categorical variables. Variables: site elevation (in metres a.s.l.); mean annual temperature (MAT; in degrees Celsius); mean annual precipitation (MAP; in millimetres per year); drought duration (in days): total number of days with rain-out shelters intercepting rainfall; drought frequency (*n*): number of sequential drought events; standardized precipitation evaporation index (SPEI) represents the number of standard deviations by which the climatic water balance (precipitation minus potential evapotranspiration) differs from the monthly mean historical record of a certain locality, with negative/positive values indicating drier/wetter conditions; rainfall seasonality indicates how monthly rainfall is distributed throughout the year; it ranges from zero (no seasonality, with rainfall uniformly distributed throughout the year) to 3.58 (maximum seasonality, with rainfall concentrated in a single month); rain-out shelter area (in square metres); vegetation type: natural (N; vegetation was not subjected to management), semi-natural (S; vegetation subjected to mowing, ploughing, fire, grazing and/or other disturbances) or manipulated (M; seeding in the field, transplantation of intact monoliths or mesocosms); rain-out shelter type: permanent (P; shelters covered vegetation throughout the experimental drought period) or automatic (A; shelters were able to detect rainfall and covered vegetation only during rainfall events); control plot type: ambient (A; unsheltered control plots receiving ambient precipitation) or historical (H; sheltered control plots, which are irrigated to receive amounts of rainfall corresponding to historical precipitation recorded for the study area); drought type: amount (A; precipitation amount reduced by a fixed percentage), seasonality (S; rainfall seasonality altered without significant changes in precipitation amount) or both (B; both types of drought were applied).

afterwards, thus leading to lower resistance and potentially higher recovery and resilience for those communities in comparison to communities with lower initial biomass (Wang, Yu, & Wang, 2007). Indices based on simple (group 2), compounded (groups 3) or natural logarithm ratios (group 5) could be used to avoid this problem, because they report changes between treatment and control plots in proportional terms. In group 4 indices, stability was also calculated as a ratio, but their results were more complicated to interpret. For the RT4 index, positive values indicate that biomass produced in drought plots was higher than in control plots (BD > BC; for example, as in the "yellow" study), whereas negative values indicate that biomass was lost during drought. A value equal to minus one indicates that the system collapsed during drought (BD = 0). In scenarios of full resistance (BD = BC), however, those indices were undefined because they resulted in divisions by zero.

Only group 6 indices expressed differences in above-ground biomass per unit change in precipitation. In the RT6 index, for instance, the magnitude of change in biomass between control and drought plots (BD/BC) was standardized by the magnitude of the drought treatment imposed (PPT_D/PPT_C). Thus, values equal to zero indicate that the system collapsed during drought (BD = 0), and progressively higher values indicate higher resistance. This standardization has been used to compare stability across studies that differed in the intensity of the drought treatment imposed. Therefore, a lower stability for a certain system over others would not be attributable merely to a greater proportional reduction in the amount of rainfall (Shinoda et al., 2010). However, as previously discussed, the percentage of rainfall intercepted might not be a reliable index of drought for comparison across studies.

Regarding the resilience indices, studies have measured this property using two distinct approaches: resilience per se (RS1–RS5)

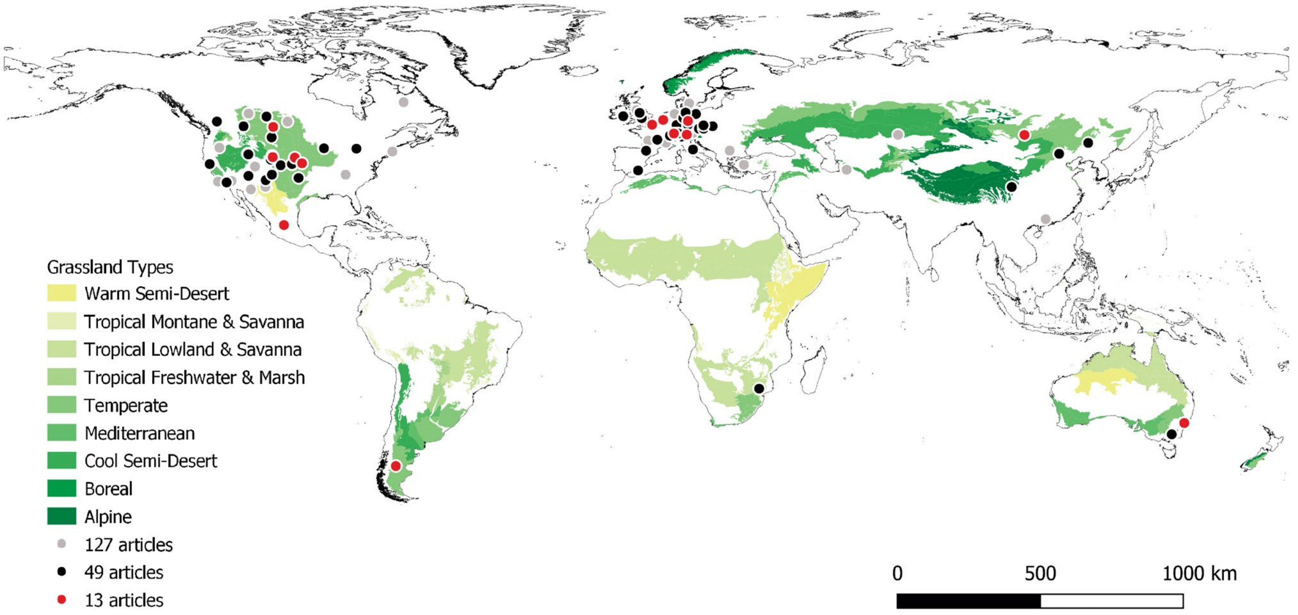

FIGURE 2 Global map displaying site location of rainfall manipulation experiments conducted in grasslands. Red dots indicate the location for 19 study cases (extracted from 13 articles that reported above-ground biomass data for the post-drought period), red + black dots indicate the location for 101 study cases (extracted from 49 articles that met the second set of inclusion criteria) and red + black + grey dots indicate the location of 204 study cases (extracted from 127 articles that met the first set of inclusion criteria). Grasslands types and distributions were obtained from Dixon et al. (2014)

TABLE 3 Stability indices used to assess resistance, recovery, resilience and relative resilience of grasslands in response to rainfall manipulation experiments

Group	Resistance (RT)	Recovery (RC)	Resilience (RS)	Relative resilience (RRS)
1	$RT1 = BD - BC$[2, 7, 9, 14, 15]	$RC1 = BD_{pd} - BD$*	$RS1 = BD_{pd} - BC_{pd}$[12]	$RRS1 = (BD_{pd}/BC_{pd}) - (BD/BC)$[5]
2	$RT2 = BD/BC$[4–6, 8, 12, 15]	$RC2 = BD_{pd}/BD$[2]	$RS2 = BD_{pd}/BC_{pd}$[2, 4–9, 12]	$RRS2 = [BC_{pd}/(BD_{pd} - BC_{pd})]/[BC/(BD - BC)]$[13]
3	$RT3 = (BD - BC)/BC$[3, 15–16]	$RC3 = (BD_{pd} - BD)/BD$*	$RS3 = (BD_{pd} - BC_{pd})/BC_{pd}$*	
4	$RT4 = BC/(BD - BC)$[13]	$RC4 = BD/(BD_{pd} - BD)$*	$RS4 = BC_{pd}/(BD_{pd} - BC_{pd})$[13]	
5	$RT5 = \ln(BD/BC)$[1, 2, 4, 11, 16, 17]	$RC5 = \ln(BD_{pd}/BD)$[2]	$RS5 = \ln(BD_{pd}/BC_{pd})$[9]	
6	$RT6 = \frac{(BD/BC)/(PPT_D/PPT_C)}{BC/PPT_C}$[10]	$RC6 = \frac{(BDpd/BD)/(PPT_{PD}/PPT_D)}{BC/PPT_C}$[10]		

Note: Stability indices classes: group 1, raw difference; group 2, simple ratio; groups 3 and 4, compounded ratio; group 5, natural logarithm ratio; group 6, standardized ratio; biomass in control (BC) and drought (BD) plots during the drought period; biomass in control (BC_{pd}) and drought (BD_{pd}) plots during the post-drought period; precipitation in control (PPT_C) and drought plots (PPT_D) during the drought period and in both plots during the post-drought period (PPT_{PD}). Numbers in square brackets indicate references reporting each index: [1] Byrne et al. (2017); [2] Carter & Blair (2012); [3] Hoekstra et al. (2015); [4] Hofer et al. (2016); [5] Ingrisch et al. (2018; [6] Kreyling, Dengler et al. (2017); [7] Lanta, Doležal, Zemková, and Lepš (2012); [8] Mariotte, Vandenberghe, Kardol, Hagedorn, & Buttler (2013); [9] Pfisterer & Schmid (2002); [10] Shinoda et al. (2010); [11] Stampfli et al. (2018); [12] Vogel et al. (2012); [13] Wagg et al. (2017); [14] Wang et al. (2007); [15] Wang et al. (2014); [16] Yahdjian & Sala (2006); [17] Dormann et al. (2017). *Indices not reported in any of the 127 references used in the meta-analysis, but derived here to fill all indices groups.

and relative resilience (RRS1 and RRS2). Indices of resilience per se quantified the ability of grasslands to reach non-drought biomass levels after drought alleviation. A grassland would be considered resilient whenever non-significant differences were detected between control and drought plots during the post-drought period ($BC_{pd} = BD_{pd}$). Alternatively, relative resilience indices quantified whether grasslands would be able to recover all the biomass previously lost during the drought. Thus, in RRS indices the resilience is weighted by the degree of damage incurred during the drought (i.e., RS is weighted by RT).

3.3 | On average, grasslands were resilient to drought and exhibited a trade-off between high recovery and low resistance

Resistance assumed significantly negative values (Figure 4), indicating that above-ground biomass was strongly reduced during drought ($n = 101$; mean = −1.12; 95% CI = −1.42, −0.82; $p < .01$; AICc = 379.28). In contrast, recovery was significantly positive ($n = 19$; mean = 1.12; 95% CI = 0.25, 1.99; $p = .01$; AICc = 78.98), because grasslands were able to produce new biomass after drought

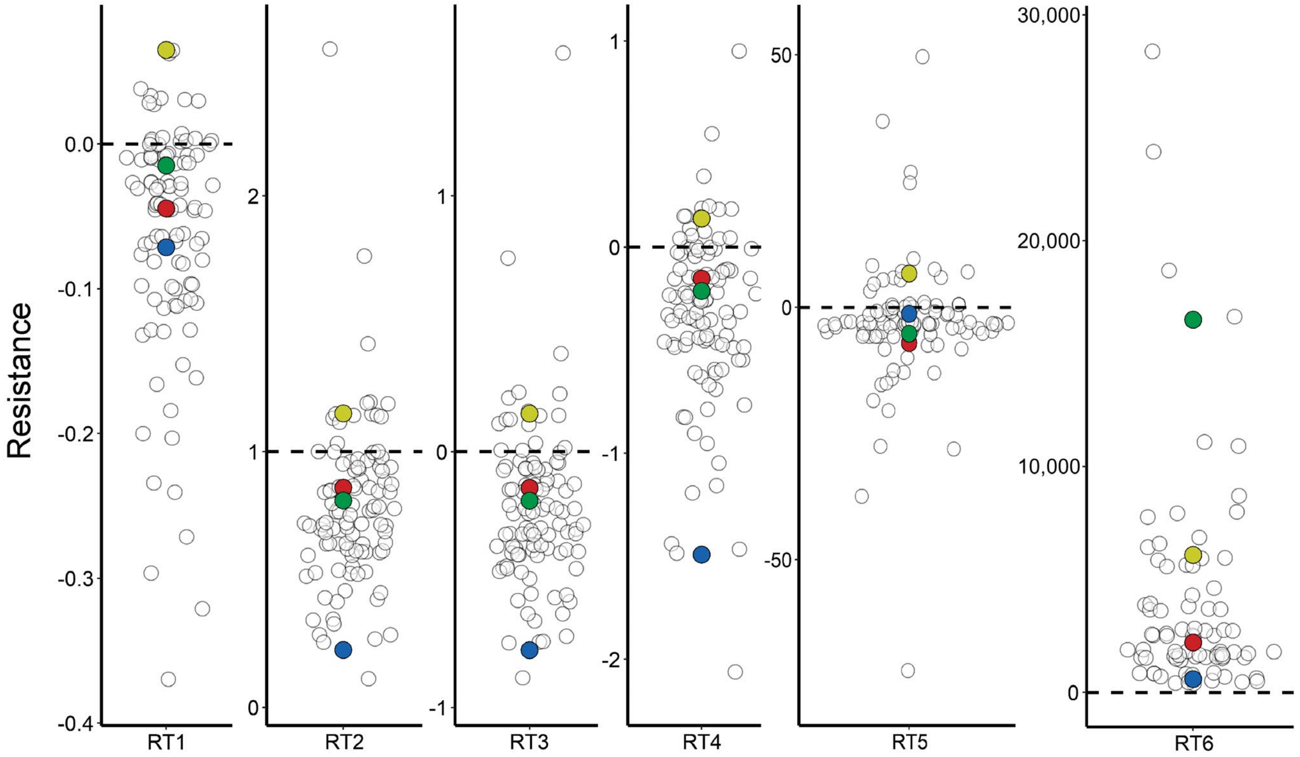

FIGURE 3 Comparisons across stability indices used to assess resistance (RT1–RT6, *n* = 101) of the above-ground biomass of grasslands in response to experimental droughts. In studies conducting sequential drought events, indices were computed using only the biomass data for the last drought event. Four study cases are highlighted (yellow: Carter & Blair, 2012; grey: Yahdjian & Sala, 2006; red: Ingrisch et al., 2018; and blue: Shinoda et al., 2010) to illustrate the incomparability among different stability indices. Dashed black lines indicate values for full resistance. Points are jittered to reduce overlapping. See Table 3 for resistance indices (RT1–RT6)

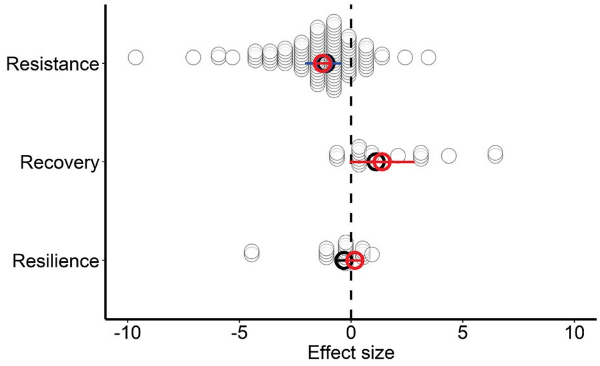

FIGURE 4 Hedges *g* effect sizes for resistance, recovery and resilience of the above-ground biomass of grasslands in response to experimental droughts. In studies reporting sequential drought events, effect sizes were computed using biomass data for only the last experimental drought. Negative/positive values indicate decreases/increases in biomass in treatment plots compared with control plots. Dots/lines indicate mean effect sizes and the 95% confidence interval (CI) for the whole dataset (blue; *n* = 101), for only studies that monitored post-drought responses (black; *n* = 19) and for only studies that succeed on creating drought conditions (red; *n* = 51 for resistance and *n* = 10 for recovery and resilience)

alleviation. The biomass produced after drought was, in general, sufficient to compensate for the biomass lost during drought, thus grasslands exhibited almost complete resilience (i.e., although resilience assumed a negative value, it was not significantly different from zero; *n* = 19; mean = −0.32; 95% CI = −0.66, 0.02; *p* = .06; AICc = 53.97).

Similar results were found when mean effect sizes were calculated for only the subset of studies that succeeded in creating experimental drought conditions (i.e., SPEI ≤ −1.0; RT: *n* = 51; mean = −1.28; 95% CI = −1.66, −0.89; *p* < .01; AICc = 181.69; RC: *n* = 10; mean = 1.39; 95% CI = −0.0, 2.85; *p* = .05; AICc = 42.65; RS: *n* = 10; mean = 0.16; 95% CI = −0.22, 0.56; *p* > .05; AICc = 19.96) and when resistance was estimated for only the 19 studies that also monitored the post-drought response (RT: *n* = 19; mean = −1.24; 95% CI = −2.06, −0.42; *p* < .01; AICc = 78.10) (Figure 4).

When resistance was calculated for each category of drought intensity (based on SPEI classification; Supporting Information Figure S6.1a in Appendix S6), we observed a decrease in resistance as drought intensity increased. Thus, we highlight that our results may be underestimated, because few studies have subjected grasslands to severe (*n* = 16) or extreme (*n* = 9) drought events. Finally, we found a trade-off between recovery and resistance (Figure 5a), but not between the other stability properties (Figure 5b,c).

3.4 | Ecosystem stability differed greatly among grasslands and drought characteristics, and climatic conditions did not generally explain these differences

We found significant variation in resistance (101 study cases: Q = 323.65; $p < .001$; $I^2 = 74.51\%$; 19 study cases: Q = 59.39; $p < .001$; $I^2 = 75.43\%$), recovery (Q = 59.24; $p < .001$; $I^2 = 79.18\%$) and resilience (Q = 29.06; $p = .0476$; $I^2 = 24.84\%$) among grasslands. However, drought characteristics (duration, frequency and SPEI), elevation and climatic conditions (MAT, MAP and *S*) were not considered important to explain the variability in resistance (Figure 6a) and recovery (Figure 6b) (i.e., importance values < .8). In contrast, resilience was explained by drought duration (grey bars in Figure 6c), with longer droughts resulting in lower resilience. When experiments that failed to create drought conditions were excluded, the results of Q-tests did not change for recovery (RC: $n = 10$; Q = 34.13; $p < .001$; $I^2 = 82.05\%$) and resistance (RT: $n = 51$; Q = 133.81; $p < .001$; $I^2 = 64.53\%$), but they became non-significant for resilience (RS: $n = 10$; Q = 5.97; $p = .74$; $I^2 = 0\%$). Consequently, drought duration was no longer important to explain the variability in grassland resilience (red bars in Figure 6c). This latter result, however, must be interpreted with caution, because for recovery and resilience the number of study cases was too small ($n = 10$) compared with the number of moderators tested ($n = 7$) (Calcagno & Mazancourt, 2010).

3.5 | Sensitivity analysis

Our results were robust to the influence of outliers (Supporting Information Table S6.1) and the non-independence effects, caused either by multiple locations (Supporting Information Table S6.4) or by multiple consecutive drought events (Supporting Information Figure S6. 2–4). Nonetheless, evidence of publication bias and effects related to the length of the post-drought period were detected for resilience (Supporting Information Table S6.5).

4 | DISCUSSION

We provide key insights into how grasslands respond to experimental droughts and we highlight four striking limitations to the methodological and analytical approaches in rainfall manipulation experiments conducted so far. First, although the stability of grasslands seems to be achieved mainly through recovery, 81% of the studies have only assessed grassland resistance, not monitoring post-drought productivity. Second, although the aim of most rainfall manipulation experiments is to expose grasslands to droughts, half of the experiments did not succeed doing so. Third, although the experiments were conducted to assess grassland stability, 87% of them did not calculate any index of stability, and those that did used distinct indices, which provide incomparable results. Fourth, although tropical grasslands contribute *c*. 38% of the world grassland area (Dixon et al., 2014), only one study (Arredondo et al., 2016) was conducted in the tropics (Figure 2).

We found that grassland above-ground biomass was resilient to experimental droughts, and this resilience was achieved more through recovery than through resistance (Figure 4), resulting in a trade-off between those two last properties (Figure 5a). This trade-off is consistent with previous studies (Ruppert et al., 2015) and supports theoretical assumptions that resistance and recovery might be considered as two alternative strategies to achieve ecosystem resilience (e.g., Hodgson et al., 2015). Such strategies could be predicted by the life history and evolutionary histories of resident plant species (Craven et al., 2018; Grime & Pierce, 2012; Majekova et al., 2014). Therefore, plant communities dominated by fast-growing, short-lived species (acquisitive resource-use strategy) would be expected to show low resistance, because they lack traits to sustain biomass production during water stress. However, they would show high recovery. After drought alleviation, their high investment

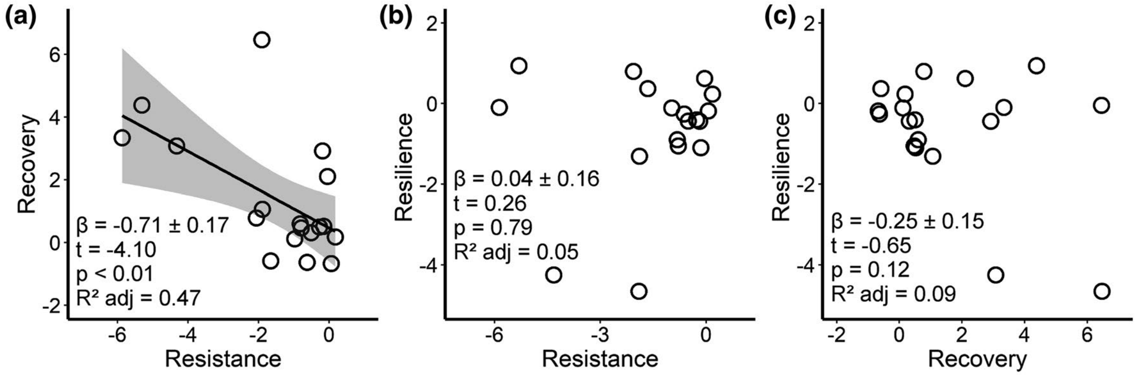

FIGURE 5 Linear regressions testing for potential trade-offs among stability properties. Resistance, recovery and resilience were computed as Hedges' *g* effect sizes measuring the effects of experimental droughts on above-ground biomass of 19 study cases. In studies reporting sequential drought events, effect sizes were computed using biomass data for only the last experimental drought. Linear regression results: coefficient estimate ± *SE* (β); *t*-statistics (*t*); *p*-value (*p*) and adjusted R^2 (R^2 adj)

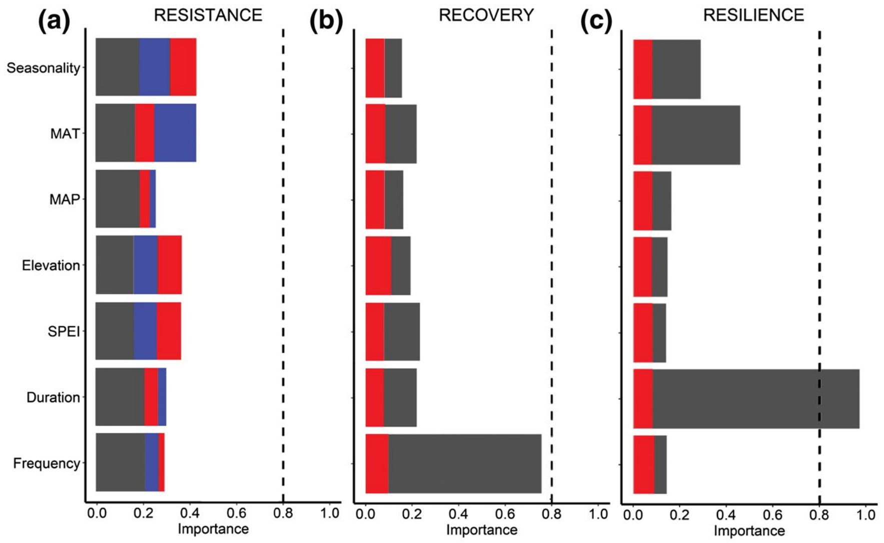

FIGURE 6 Model-averaged importance of seven moderators explaining: (a) resistance; (b) recovery; and (c) resilience of grasslands in terms of above-ground biomass in response to experimental droughts. Moderators: rainfall seasonality (seasonality; dimensionless); mean annual temperature (MAT; in degrees Celsius), mean annual precipitation (MAP; in millimetres per year), elevation above sea level (in metres a.s.l.), standardized precipitation evapotranspiration index (SPEI) as a measure of drought intensity; drought duration (in days); and drought frequency (number of sequential drought events). Blue bars indicate importance values calculated for the whole dataset (n = 101), grey bars for only studies that monitored post-drought responses (n = 19) and red bars for only studies that succeeded in creating drought conditions (red: n = 51 for resistance and n = 10 for recovery and resilience). The vertical dashed line drawn at .80 can be used as a cut-off to identify important moderators. In studies reporting sequential drought events, effect sizes were computed using biomass data for only the last experimental drought

in seed production and fast-growth rates, in comparison to other life-forms, would support a rapid return to the non-drought levels of biomass. An opposite response would be predicted for communities dominated by slow-growing, long-lived species (i.e., conservative resource-use strategy) (Craven et al., 2018; Majekova et al., 2014). Once grasslands are dominated by herbaceous species with an acquisitive strategy, we would indeed expect an overall resilience mainly driven by recovery (Stuart-Haëntjens et al., 2018).

The duration of drought was important to explain the variability in resilience (Figure 6c), with longer droughts resulting in decreased grassland resilience. Although this result is still preliminary, owing to the small number of studies monitoring post-drought biomass production (n = 19), it suggests that under future climatic regimes, where droughts are expected to last longer (Dai, 2013), there might be a gradual decline in overall grassland biomass production. Longer droughts could reduce resilience either by decreasing the ability of grasslands to recover by causing depletion of the seed bank and stored resources needed for re-establishment and resprouting of the drought-sensitive species or to resist by surpassing thresholds of tolerance and causing widespread mortality of the drought-tolerant species (Estiarte et al., 2016; Grime & Pierce, 2012; Ruppert et al., 2015).

For resistance and recovery, none of the moderators we tested was considered important to explain the variability across grasslands (Figure 6a,b). Previous meta-analysis also supported that climatic variables (MAP and MAT) are not major drivers to explain the variability in grassland above-ground resistance (n = 83; Wilcox et al., 2017), although other studies found a significant decline in resistance (n = 43) and resilience (n = 22) in xeric (higher MAP) and warmer (higher MAT) grasslands (Stuart-Haëntjens et al., 2018). Contrasting results have also been reported for the effect of drought characteristics. When drought intensity was measured as %PPT, a significant and negative relationship was found, with grassland resistance decreasing as drought intensity increased (n = 65; Gao, Zhang, Tang, & Wu, 2019). However, no effect was observed when drought intensity was measured as SPEI (n = 43; Stuart-Haëntjens et al., 2018). Such discrepancies across studies might be related to the different methods used to estimate the moderators (e.g., drought intensity as SPEI or %PPT intercepted), to assess stability (e.g., Hedges' g or log response ratio) and/or to the variation in sample size; hence, in the range of moderator values. For drought characteristics, the range of values evaluated was relatively narrow, because most of the studies evaluated the stability of grasslands in response to a single experimental drought event, with short duration and low intensity. Considering that stability and such drought characteristics may exhibit nonlinear relationships, the importance of those moderators must be re-evaluated in the future, as results of new experiments applying longer, more frequent and intense droughts become available.

Besides the drought characteristics and prevailing climatic conditions, the differential stability across grasslands might also be related to edaphic and/or intrinsic biological differences (Ploughe et al., 2019). Considerable evidence from both meta-analyses (Craven et al., 2018; Ruppert et al., 2015) and several individual studies suggests that different aspects of biodiversity [species diversity (Kreyling, Dengler, et al., 2017; Pfisterer & Schmid, 2002; Vogel et al., 2012; Wagg et al., 2017), functional diversity (grass, legumes and forbs) (Byrne, Adler, & Lauenroth, 2017; Hofer et al., 2016; Stampfli, Bloor, Fischer, & Zeiter, 2018), and plant life-history traits (annual versus perennial) (Hoeppner & Dukes, 2012; Wang et al., 2014)] can significantly influence the ecosystem stability of grasslands in response to drought. Moreover, soil characteristics, such as texture, depth and soil water content at field capacity, could also be important to explain the variation in resistance and recovery (Vicca et al., 2012; Wilcox et al., 2017) and could even influence the biodiversity–ecosystem stability relationships (Guerrero-Ramírez et al., 2017). To date, most of the studies have not reported soil characteristics, species and functional diversity at the plot level, nor provided the life-history traits for both dominant and subordinate plant species. We strongly recommend that future studies should provide this information, in order that further quantitative analysis could be able to assess the relative importance of both abiotic and biotic moderators simultaneously.

We also found a mismatch between how grasslands respond to drought and how experimenters have been assessing this response. Although it has been advocated that stability in the productivity of grasslands is achieved mainly through recovery (Stuart-Haëntjens et al., 2018), most of the experiments have focused only on the immediate impacts of drought (resistance). This incomplete view of the varied responses of grasslands risks overestimating the impacts of droughts (Donohue et al., 2016), because experiments show how the exposure to drought can result in significant biomass loss, but neglect the ability of grasslands partly or completely to recover the biomass that had been lost.

Conversely, impacts of droughts might be underestimated, because we found that 50% of the studies have failed to impose drought upon grasslands experimentally, when drought intensity was estimated as SPEI. Similar results were found by Slette et al. (2019), who investigated how ecologists define drought and found that for half of the evaluated drought studies the SPEI values were inside the range of nearly normal conditions. In fact, these authors recommended the use of standardized methods, such as SPEI, to characterize natural droughts owing to their ability to allow comparisons among studies and improving ecological predictions (Slette et al., 2019). In this sense, the answer to the question, “how can a drought experiment fail to create a drought treatment?”, is that this situation can happen particularly in experiments where rain-out shelters are installed on a site to intercept ≤ 50% of the rainfall (which would be considered as a moderate drought based on %PPT) during short and interrupted periods of time (Figure 1b). If the interannual variability of rainfall for that site is relatively large, and the experiment is conducted during a wetter year, the amount of rainfall intercepted may not be enough to simulate a drought event in terms of SPEI. This explains why classifications based on %PPT may overestimate the intensity of the experimental droughts (Supporting Information Table S6.2; Figure S6.4 in Appendix S6) and should be replaced by drought indices (such as SPEI) that consider the historical precipitation record for each locality (Knapp et al., 2017; Slette et al., 2019), especially when studies aim to compare grassland stability across years and/or across sites. Fortunately, Web tools are available to help quantify how much precipitation must be intercepted to impose a specific level of drought extremity for a given location (Lemoine, Sheffield, Dukes, Knapp, & Smith, 2016). Thus, whenever possible, grasslands should be subjected to multiple drought treatment levels, including more extreme droughts. This approach would allow identification of response thresholds (i.e., critical levels of drought extremity that result in an irreversible loss of biomass; Knapp et al., 2018) and provide the ability to predict which grasslands would be more affected in a future with increased frequency of extreme droughts (Kreyling, Jentsch, & Beier, 2014).

Identification of the most sensitive sites and the mechanisms responsible for such sensitivity is, indeed, one of the primary goals of those experiments. However, we demonstrate that the ranking of grassland stability is conditional upon the specific index of stability used (Figure 3). Previous studies have already highlighted this need to standardize indices of stability for a proper comparison across sites (Smith et al., 2017), and some metrics have been proposed for this purpose (Ingrisch & Bahn, 2018). We suggest that group 1 indices (raw difference) must be used only if the study aims to calculate the stability of a single system or to compare systems that have similar values of pre-drought above-ground biomass. To compare stability among grasslands that differ in their initial state, researchers could use indices of group 2 (simple ratio), group 3 (compounded ratio) or group 5 (natural logarithm ratio). Both group 4 (compounded ratio) and group 6 (standardized ratio) indices should be avoided, because the former provides results that are hard to interpret, and the latter use the percentage of rainfall intercepted for standardization, which might not be a reliable index of drought for comparison across studies. Given that each index has its advantages and limitations, it is unlikely that a single standard metric would be useful in all situations (Donohue et al., 2016; Quinlan et al., 2016) and would be able to account for all the different factors that might influence grassland stability (e.g., characteristics of drought treatment, prevailing climatic conditions during the drought period, intrinsic biological differences among grasslands). Moreover, when indices are used to assess stability, researchers tend to report only the results for the indices, omitting the biomass data used to calculate them. For example, 61% of the articles initially gathered in the present study could not be used to compute stability indices and Hedges’ *g* effect sizes, precisely because biomass data were not reported. Thus, regardless of the index selected, we strongly urge future studies to report biomass and precipitation data fully. These data could then be used to recalculate any of the stability indices and to conduct meta-analysis. Additionally, studies should report sufficient plot-level information (such as plant functional diversity and soil characteristics) for further assessment of moderator effects. Previous publications have provided excellent guidance about which information should be reported in studies to enable subsequent meta-analysis (Beier et al., 2012; Gerstner et al., 2017; Smith et al., 2014).

Finally, we highlight that caution is required in extrapolating our results to the tropics, and that future studies should expand the geographical extent of drought experiments to those understudied areas (Beier et al., 2012; Hoover et al., 2018; Stuart-Haëntjens et al., 2018). Collaborative initiatives, such as the Drought-Net (Knapp et al., 2017; Smith et al., 2017), could help and could also lead to the standardization of methodological approaches across studies, thus facilitating further meta-analytical assessments.

5 | CONCLUSION

To improve our mechanistic understanding of differential stability across grasslands and our ability to manage them to ensure stable productivity in a drier world, we provide the following recommendations for future rainfall manipulation experiments: (a) evaluate the three stability properties simultaneously; (b) impose multiple and more extreme drought treatments; (c) Use appropriate indices to estimate drought intensity; (d) report basic information necessary to calculate stability indices and to assess the effect of moderators; and (e) evaluate the stability of tropical grasslands. Despite the focus of our study upon grasslands, we argue that the main shortcomings listed here and the improvements suggested are generalizable and should be used to guide future experimental rainfall manipulation studies evaluating the effects of drought in different ecosystems.

ACKNOWLEDGMENTS

We are grateful to T. Moulton, J. Boelsums, F. Melo, C. Eller and J. Prevedello for their constructive comments on the first draft of this manuscript. We thank Rufford Foundation 18749-1, Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro - FAPERJ (Bolsa Jovem Cientista do Nosso Estado-JCNE, E-26/203.199/2016) and Prociência for financial support. This study was financed, in part, by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Brasil (CAPES), Finance Code 88882.182435/2018-01.

ORCID

Ilaíne Silveira Matos https://orcid.org/0000-0001-5557-5133
Imma Oliveras Menor https://orcid.org/0000-0001-5345-2236
Sami Walid Rifai https://orcid.org/0000-0003-3400-8601
Bruno Henrique Pimentel Rosado https://orcid.org/0000-0002-8924-8672

DATA ACCESSIBILITY

The data that support the findings of this study are openly available in Dryad: https://doi.org/10.5061/dryad.dz08kprt2.

REFERENCES

Abatzoglou, J. T., Dobrowski, S. Z., Parks, S. A., & Hegewisch, K. C. (2018). Terraclimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958–2015. *Scientific Data*, *5*, 170191.

Arredondo, T., Garcia-Moya, E., Huber-Sannwald, E., Loescher, H. W., Delgado-Balbuena, J., & Luna-Luna, M. (2016). Drought manipulation and its direct and legacy effects on productivity of a monodominant and mixed-species semi-arid grassland. *Agricultural and Forest Meteorology*, *223*, 132–140. https://doi.org/10.1016/j.agrformet.2016.03.011

Beier, C., Beierkuhnlein, C., Wohlgemuth, T., Penuelas, J., Emmett, B., Körner, C., … Hansen, K. (2012). Precipitation manipulation experiments – challenges and recommendations for the future. *Ecology Letters*, *15*, 899–911. https://doi.org/10.1111/j.1461-0248.2012.01793.x

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Chichester, West Sussex, UK: John Wiley and Sons.

Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multimodel inference: A practical information-theoretic approach*, 2nd ed. Fort Collins, CO: Springer.

Byrne, K. M., Adler, P. B., & Lauenroth, W. K. (2017). Contrasting effects of precipitation manipulations in two Great Plains plant communities. *Journal of Vegetation Science*, *28*, 238–249. https://doi.org/10.1111/jvs.12486

Calcagno, V., & de Mazancourt, C. (2010). Glmulti: An R package for easy automated model election with (generalized) linear models. *Journal of Statistical Software*, *34*, 1–29.

Cantarel, A. A. M., Bloor, J. M. G., & Soussana, J. F. (2013). Four years of simulated climate change reduces above-ground productivity and alters functional diversity in a grassland ecosystem. *Journal of Vegetation Science*, *24*, 113–126. https://doi.org/10.1111/j.1654-1103.2012.01452.x

Carter, D. L., & Blair, J. M. (2012). High richness and dense seeding enhance grassland restoration establishment but have little effect on drought response. *Ecological Applications*, *22*, 1308–1319. https://doi.org/10.1890/11-1970.1

Ciemer, C., Boers, N., Hirota, M., Kurths, J., Müller-Hansen, F., Oliveira, R. S., & Winkelmann, R. (2019). Higher resilience to climatic disturbances in tropical vegetation exposed to more variable rainfall. *Nature Geoscience*, *12*, 174–179. https://doi.org/10.1038/s41561-019-0312-z

Craven, D., Eisenhauer, N., Pearse, W. D., Hautier, Y., Isbell, F., Roscher, C., … Manning, P. (2018). Multiple facets of biodiversity drive the diversity–stability relationship. *Nature Ecology and Evolution*, *2*, 1579–1587. https://doi.org/10.1038/s41559-018-0647-7

Dai, A. (2013). Increasing drought under global warming in observations and models. *Nature Climate Change*, *3*, 52–58. https://doi.org/10.1038/nclimate1633

De Boeck, H. J., Bassin, S., Verlinden, M., Zeiter, M., & Hiltbrunner, E. (2016). Simulated heat waves affected alpine grassland only in combination with drought. *New Phytologist*, *209*, 531–541. https://doi.org/10.1111/nph.13601

De Boeck, H. J., Dreesen, F. E., Janssens, I. A., & Nijs, I. (2011). Whole-system responses of experimental plant communities to climate extremes imposed in different seasons. *New Phytologist*, *189*, 806–817. https://doi.org/10.1111/j.1469-8137.2010.03515.x

Dixon, A. P., Faber-Langendoen, D., Josse, C., Morrison, J., & Loucks, C. J. (2014). Distribution mapping of world grassland types. *Journal of Biogeography*, *41*, 2003–2019. https://doi.org/10.1111/jbi.12381

Donohue, I., Hillebrand, H., Montoya, J. M., Petchey, O. L., Pimm, S. L., Fowler, M. S., … Yang, Q. (2016). Navigating the complexity of ecological stability. *Ecology Letters*, *19*, 1172–1185. https://doi.org/10.1111/ele.12648

Dormann, C. F., Riedmatten, L. V., & Lorenzen, M. S. (2017). No consistent effect of plant species richness on resistance to simulated climate change for above- or below-ground processes in managed grasslands. *BMC Ecology*, *17*, 23. https://doi.org/10.1186/s12898-017-0133-0

Estiarte, M., Vicca, S., Peñuelas, J., Bahn, M., Beier, C., Emmett, B. A., … Janssens, I. A. (2016). Few multi-year precipitation-reduction experiments find a shift in the productivity–precipitation relationship. *Global Change Biology*, *22*, 2570–2581.

Feng, X., Porporato, A., & Rodriguez-Iturbe, I. (2013). Changes in rainfall seasonality in the tropics. *Nature Climate Change*, *3*, 811–815. https://doi.org/10.1038/nclimate1907

Forrestel, E. J., Donoghue, M. J., Edwards, E. J., Jetz, W., du Toit, J. C. O., & Smith, M. D. (2017). Different clades and traits yield similar grassland functional responses. *Proceedings of the National Academy of Sciences USA*, *114*, 705–710. https://doi.org/10.1073/pnas.1612909114

Fridley, J. D., Grime, J. P., Askew, A. P., Moser, B., & Stevens, C. J. (2011). Soil heterogeneity buffers community response to climate change in species-rich grassland. *Global Change Biology*, *17*, 2002–2011. https://doi.org/10.1111/j.1365-2486.2010.02347.x

Gao, J., Zhang, L., Tang, Z., & Wu, S. (2019). A synthesis of ecosystem aboveground productivity and its process variables under simulated drought stress. *Journal of Ecology*, *107*, 2519–2531. https://doi.org/10.1111/1365-2745.13218

Gerstner, K., Moreno-Mateos, D., Gurevitch, J., Beckmann, M., Kambach, S., Jones, H. P., & Seppelt, R. (2017). Will your paper be used in a meta-analysis? Make the reach of your research broader and longer lasting. *Methods in Ecology and Evolution*, *8*, 777–784. https://doi.org/10.1111/2041-210X.12758

Grime, J. P., Brown, V. K., Thompson, K., Masters, G. J., Hillier, S. H., Clarke, I. P., … Kielty, J. P. (2000). The response of two contrasting limestone grasslands to simulated climate change. *Science*, *289*, 762–765. https://doi.org/10.1126/science.289.5480.762

Grime, J. P., & Pierce, S. (2012). *The evolutionary strategies that shape ecosystems*. Chichester, West Sussex, UK: Wiley-Blackwell.

Grimm, V., & Wissel, C. (1997). Babel, or the ecological stability discussions: An inventory and analysis of terminology and a guide for avoiding confusion. *Oecologia*, *109*, 323–334. https://doi.org/10.1007/s004420050090

Guerrero-Ramírez, N. R., Craven, D., Reich, P. B., Ewel, J. J., Isbell, F., Koricheva, J., … Eisenhauer, N. (2017). Diversity-dependent temporal divergence of ecosystem functioning in experimental ecosystems. *Nature Ecology & Evolution*, *1*, 1639–1642. https://doi.org/10.1038/s41559-017-0325-1

Hodgson, D., McDonald, J. L., & Hosken, D. J. (2015). What do you mean, 'resilient'? *Trends in Ecology and Evolution*, *30*, 503–506. https://doi.org/10.1016/j.tree.2015.06.010

Hoekstra, N. J., Suter, M., Finn, J. A., Husse, S., & Lüscher, A. (2015). Do belowground vertical niche differences between deep- and shallow-rooted species enhance resource uptake and drought resistance in grassland mixtures? *Plant and Soil*, *394*, 21–34. https://doi.org/10.1007/s11104-014-2352-x

Hoeppner, S. S., & Dukes, J. S. (2012). Interactive responses of old-field plant growth and composition to warming and precipitation. *Global Change Biology*, *18*, 1754–1768. https://doi.org/10.1111/j.1365-2486.2011.02626.x

Hofer, D., Suter, M., Haughey, E., Finn, J. A., Hoekstra, N. J., Buchmann, N., & Lüscher, A. (2016). Yield of temperate forage grassland species is either largely resistant or resilient to experimental summer drought. *Journal of Applied Ecology*, *53*, 1023–1034. https://doi.org/10.1111/1365-2664.12694

Hoover, D. L., & Rogers, B. M. (2016). Not all droughts are created equal: The impacts of interannual drought pattern and magnitude on grassland carbon cycling. *Global Change Biology*, *22*, 1809–1820. https://doi.org/10.1111/gcb.13161

Hoover, D. L., Wilcox, K. R., & Young, K. E. (2018). Experimental droughts with rainout shelters: A methodological review. *Ecosphere*, *9*, e02088. https://doi.org/10.1002/ecs2.2088

Huxman, T. E., Snyder, K. A., Tissue, D., Leffler, A. J., Ogle, K., Pockman, W. T., … Schwinning, S. (2004). Precipitation pulses and carbon fluxes in semiarid and arid ecosystems. *Oecologia*, *141*, 254–268. https://doi.org/10.1007/s00442-004-1682-4

Ingrisch, J., & Bahn, M. (2018). Towards a comparable quantification of resilience. *Trends in Ecology and Evolution*, *33*, P251–259. https://doi.org/10.1016/j.tree.2018.01.013

Ingrisch, J., Karlowsky, S., Anadon-Rosell, A., Hasibeder, R., König, A., Augusti, A., … Bahn, M. (2018). Land use alters the drought responses of productivity and CO_2 fluxes in mountain grassland. *Ecosystems*, *21*, 689–703. https://doi.org/10.1007/s10021-017-0178-0

Knapp, A. K., Avolio, M. L., Beier, C., Carroll, C. J. W., Collins, S. L., Dukes, J. S., … Smith, M. D. (2017). Pushing precipitation to the extremes in distributed experiments: Recommendations for simulating wet and dry years. *Global Change Biology*, *23*, 1774–1782. https://doi.org/10.1111/gcb.13504

Knapp, A. K., Carroll, C. J. W., Griffin-Nolan, R. J., Slette, I. J., Chaves, F. A., Baur, L. E., … Smith, M. D. (2018). A reality check for climate change experiments: Do they reflect the real world? *Ecology*, *99*, 2145–2151. https://doi.org/10.1002/ecy.2474

Knapp, A. K., & Smith, M. D. (2001). Variation among biomes in temporal dynamics of aboveground primary production. *Science*, *291*, 481–484. https://doi.org/10.1126/science.291.5503.481

Kreyling, J., Arfin Khan, M. A. S., Sultana, F., Babel, W., Beierkuhnlein, C., Foken, T., … Jentsch, A. (2017). Drought effects in climate change manipulation experiments: Quantifying the influence of ambient weather conditions and rain-out shelter artifacts. *Ecosystems*, *20*, 301–3015. https://doi.org/10.1007/s10021-016-0025-8

Kreyling, J., Dengler, J., Walter, J., Velev, N., Ugurlu, E., Sopotlieva, D., … Jentsch, A. (2017). Species richness effects on grassland recovery from drought depend on community productivity in a multisite experiment. *Ecology Letters*, *20*, 1405–1413. https://doi.org/10.1111/ele.12848

Kreyling, J., Jentsch, A., & Beier, C. (2014). Beyond realism in climate change experiments: Gradient approaches identify thresholds and tipping points. *Ecology Letters*, *17*, 125–e1. https://doi.org/10.1111/ele.12193

Lanta, V., Doležal, J., Zemková, L., & Lepš, J. (2012). Communities of different plant diversity respond similarly to drought stress: Experimental evidence from field non-weeded and greenhouse conditions. *Naturwissenschaften*, *99*, 473–482. https://doi.org/10.1007/s00114-012-0922-4

Lemoine, N. P., Sheffield, J., Dukes, J. S., Knapp, A., & Smith, M. D. (2016). Terrestrial Precipitation Analysis (TPA): A resource for characterizing long-term precipitation regimes and extremes. *Methods in Ecology and Evolution*, *7*, 1396–1401. https://doi.org/10.1111/2041-210X.12582

Májeková, M., de Bello, F., Doležal, J., & Lepš, J. (2014). Plant functional traits as determinants of population stability. *Ecology*, *95*, 2369–2374. https://doi.org/10.1890/13-1880.1

Mariotte, P., Vandenberghe, C., Kardol, P., Hagedorn, F., & Buttler, A. (2013). Subordinate plant species enhance community resistance against drought in semi-natural grasslands. *Journal of Ecology*, *101*, 763–773. https://doi.org/10.1111/1365-2745.12064

Matos, I., Rifai, S., Oliveras, I., & Rosado, B. (2020). Data from: Deciphering the stability of grassland productivity in response to rainfall manipulation experiments, v3, Dryad, Dataset. https://doi.org/10.5061/dryad.dz08kprt2

Nimmo, D. G., Mac Nally, R., Cunningham, S. C., Haslem, A., & Bennett, A. F. (2015). Vive la résistance: Reviving resistance for 21st century conservation. *Trends in Ecology and Evolution*, *30*, 516–523. https://doi.org/10.1016/j.tree.2015.07.008

Peng, S., Piao, S., Shen, Z., Ciais, P., Sun, Z., Chen, S., … Chen, A. (2013). Precipitation amount, seasonality and frequency regulate carbon cycling of a semi-arid grassland ecosystem in Inner Mongolia, China: A modeling analysis. *Agricultural and Forest Meteorology*, *178–179*, 46–55. https://doi.org/10.1016/j.agrformet.2013.02.002

Pfisterer, A. B., & Schmid, B. (2002). Diversity-dependent production can decrease the stability of ecosystem functioning. *Nature*, *416*, 84–86. https://doi.org/10.1038/416084a

Ploughe, L. W., Jacobs, E. M., Frank, G. S., Smith, M. D., Dukes, J. S., & Skye, M. (2019). Community Response to Extreme Drought (CRED): A framework for drought-induced shifts in plant–plant interactions. *New Phytologist*, *222*, 52–69.

Quinlan, A. E., Berbés-Blázquez, M., Haider, L. J., Peterson, G. D., & Allen, C. (2016). Measuring and assessing resilience: Broadening understanding through multiple disciplinary perspectives. *Journal of Applied Ecology*, *53*, 677–687. https://doi.org/10.1111/1365-2664.12550

Ruppert, J. C., Harmoney, K., Henkin, Z., Snyman, H. A., Sternberg, M., Willms, W., & Linstädter, A. (2015). Quantifying drylands' drought resistance and recovery: The importance of drought intensity, dominant life history and grazing regime. *Global Change Biology*, *21*, 1258–1270. https://doi.org/10.1111/gcb.12777

Shinoda, M., Nachinshonhor, G. U., & Nemoto, M. (2010). Impact of drought on vegetation dynamics of the Mongolian steppe: A field experiment. *Journal of Arid Environment*, *74*, 63–69. https://doi.org/10.1016/j.jaridenv.2009.07.004

Slette, I. J., Post, A. K., Awad, M., Even, T., Punzalan, A., Williams, S., ... Knapp, A. K. (2019). How ecologists define drought, and why we should do better. *Global Change Biology*, *25*, 3193–3200. https://doi.org/10.1111/gcb.14747

Smith, M. D., Wilcox, K. R., Power, S. A., Tissue, D. T., & Knapp, A. K. (2017). Assessing community and ecosystem sensitivity to climate change – toward a more comparative approach. *Journal of Vegetation Science*, *28*, 235–237. https://doi.org/10.1111/jvs.12524

Smith, N. G., Rodgers, V. L., Brzostek, E. R., Kulmatiski, A., Avolio, M. L., Hoover, D. L., ... Niyogi, D. (2014). Toward a better integration of biological data from precipitation manipulation experiments into Earth system models. *Reviews of Geophysics*, *52*, 412–434. https://doi.org/10.1002/2014RG000458

Stampfli, A., Bloor, J. M. G., Fischer, M., & Zeiter, M. (2018). High land-use intensity exacerbates shifts in grassland vegetation composition after severe experimental drought. *Global Change Biology*, *24*, 2021–2034. https://doi.org/10.1111/gcb.14046

Stuart-Haëntjens, E., De Boeck, H. J., Lemoine, N. P., Mänd, P., Kröel-Dulay, G., Schmidt, I. K., ... Smith, M. D. (2018). Mean annual precipitation predicts primary production resistance and resilience to extreme drought. *Science of the Total Environment*, *636*, 360–366. https://doi.org/10.1016/j.scitotenv.2018.04.290

Sutherland, W. J., Freckleton, R. P., Godfray, H. C. J., Beissinger, S. R., Benton, T., Cameron, D. D., ... Wiegand, T. (2013). Identification of 100 fundamental ecological questions. *Journal of Ecology*, *101*, 58–67. https://doi.org/10.1111/1365-2745.12025

Tao, H., Borth, H., Fraedrich, K., Su, B., & Zhu, X. (2014). Drought and wetness variability in the Tarim River Basin and connection to large-scale atmospheric circulation. *International Journal of Climatology*, *34*, 2678–2684. https://doi.org/10.1002/joc.3867

Tilman, D., & Haddi, A. (1992). Drought and biodiversity in grasslands. *Oecologia*, *89*, 257–264. https://doi.org/10.1007/BF00317226

Vicca, S., Gilgen, A. K., Camino Serrano, M., Dreesen, F. E., Dukes, J. S., Estiarte, M., ... Granier, A. (2012). Urgent need for a common metric to make precipitation manipulation experiments comparable. *New Phytologist*, *195*, 518–522. https://doi.org/10.1111/j.1469-8137.2012.04224.x

Vicente-Serrano, S. M., Beguería, S., & López-Moreno, J. I. (2010). A Multi-scalar drought index sensitive to global warming: The Standardized Precipitation Evapotranspiration Index – SPEI. *Journal of Climate*, *23*, 1696.

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*, 1–48.

Vogel, A., Scherer-Lorenzen, M., & Weigelt, A. (2012). Grassland resistance and resilience after drought depends on management intensity and species richness. *PLoS ONE*, *7*, e36992.

Wagg, C., O'Brien, M. J., Vogel, A., Scherer-Lorenzen, M., Eisenhauer, N., Schmid, B., & Weigelt, A. (2017). Plant diversity maintains long-term ecosystem productivity under frequent drought by increasing short-term variation. *Ecology*, *98*, 2952–2961. https://doi.org/10.1002/ecy.2003

Wang, Y., Yu, S., & Wang, J. (2007). Biomass-dependent susceptibility to drought in experimental grassland communities. *Ecology Letters*, *10*, 401–410. https://doi.org/10.1111/j.1461-0248.2007.01031.x

Wang, Z., Silva, L. C. R., Sun, G., Luo, P., Chengxiang, M., & Horwath, W. R. (2014). Quantifying the impact of drought on soil-plant interactions: A seasonal analysis of biotic and abiotic controls of carbon and nutrient dynamics in high-altitudinal grasslands. *Plant and Soil*, *389*, 59–71. https://doi.org/10.1007/s11104-014-2337-9

Wilcox, K. R., Shi, Z., Gherardi, L. A., Lemoine, N. P., Koerner, S. E., Hoover, D. L., ... Luo, Y. (2017). Asymmetric responses of primary productivity to precipitation extremes: A synthesis of grassland precipitation manipulation experiments. *Global Change Biology*, *23*, 4376–4385. https://doi.org/10.1111/gcb.13706

Yahdjian, L., & Sala, O. E. (2006). Vegetation structure constrains primary production response to water availability in the Patagonian steppe. *Ecology*, *87*, 952–962. https://doi.org/10.1890/0012-9658(2006)87[952:VSCPPR]2.0.CO;2

BIOSKETCHES

Ilaíne S. Matos is a plant ecophysiologist who is particularly interested in deciphering the mechanisms underlying the differential stability of plant communities to climate change (www.sites.google.com/view/ilainematos/home). This study is part of her PhD thesis at the State University of Rio de Janeiro.

Imma Oliveras is an ecosystem ecologist who studies the vulnerability and resilience of ecosystems to disturbances and abiotic stressors.

Sami W. Rifai is a forest ecologist researching the interaction between tropical ecosystems and climate.

Bruno H. P. Rosado is a plant ecophysiologist researching ecohydrology and the responses of plant communities to climate.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Matos IS, Menor IO, Rifai SW, Rosado BHP. Deciphering the stability of grassland productivity in response to rainfall manipulation experiments. *Global Ecol Biogeogr*. 2019;00:1–15. https://doi.org/10.1111/geb.13039