Title 1: Assessing stability of grasslands' biomass production in response to experimental droughts

Or

Title 2: Deciphering ecological stability of grasslands in response to experimental droughts

Or

Title 3: Conceptual and methodological limitations of experimental droughts in grasslands

a Ilaíne S. Matos*, ilaine.matos@gmail.com

b Imma Oliveras

c Bruno H. P. Rosado*, brunorosado@gmail.com

a Ecology and Evolution Graduate Program, IBRAG, Rio de Janeiro State University, Rua São Francisco Xavier, PHLC, 20550900, Rio de Janeiro, Rio de Janeiro, Brazil, Phone: 55 2123340260

b Environmental Change Institute, School of Geography and the Environment, University of Oxford, South Parks Road, OX13QY, Oxford, UK

c Department of Ecology, IBRAG, Rio de Janeiro State University (UERJ), Rua São Francisco Xavier, PHLC, sala 220, 20550900, Rio de Janeiro, Rio de Janeiro, Brazil, Phone: 55 2123340260

Abstract

Rainfall manipulative experiments are essential for deciphering mechanisms leading to differential ecological stability across plant communities. However, the absence of standardized stability indices and common methodologies across studies has impaired such goal. Here, we conducted a meta-analysis of studies reporting effects of experimental droughts on grassland above-ground biomass production and we found that most grasslands were resilient (keeping biomass unchanged after drought) and exhibited a trade-off between low resistance (losing biomass during drought) and high recovery (producing new biomass after drought). Despite this general pattern, ecological stability varied across studies. This variability was not merely related to differences in drought treatments or in prevailing climatic conditions, but it might reflect intrinsic biotic differences among grasslands. We call for an urgent need to standardize indices and experimental methodologies to provide comparable and consistent results across studies. We also call for tropical studies on ecological stability in response to drought.

Key-words: extreme climatic events; meta-analysis; rain-out shelters; resistance; recovery; resilience;

Main text

Grasslands cover a large proportion of the land surface[1], and their productivity is strongly controlled by precipitation[2]. Considering that the frequency, duration and intensity of droughts are predicted to increase[3,4], hundreds of rainfall manipulative experiments have been conducted (Fig. 1) to investigate the effects of those shifted precipitation regimes on grassland above-ground biomass production. These experiments do not report consistent results[5-7], and the underlying mechanisms explaining differences in grasslands ecological stability in response to drought have not yet been identified[6,8].We argue that these inconsistencies arise from two main issues: a) the lack of standardized experimental methodologies and protocols on manipulative experiments; and b) the absence of both consensual definitions and standardized indices to assess ecological stability[9,10].

Indeed, methodological differences difficult comparability among studies[8,11]. To start with, studies differ on the type of experimental drought imposed; while some studies only altered rainfall seasonality, others reduced precipitation amounts by a fixed percentage (Figs. 1a-c). As those drought experiments are in reality mostly precipitation deficit experiments (meteorological drought), scientists have often relied on that percentage of rainfall intercepted to scale drought treatment intensity[8]. However, it has been shown that percentages are inappropriate for comparing sites with different climate (Fig. 1) and more meaningful metrics are needed[8,12,13]. As a result of these methodological constraints, we still do not know whether the observed variation in stability responses might be attributed to intrinsic biological differences across systems or might merely reflect differences in the drought treatment or in the prevailing climatic conditions experienced by plants[10,13].

Second, studies have also varied in their definition of ecological stability and in the indices used to assess it[9,10]. Although there is still no consensus in the literature, it has been suggested that ecological stability should be measured through three distinct properties: resistance (RT: the system ability to keep its biomass unchanged during drought); recovery (RC: the system ability to produce biomass after drought); and resilience (RS: the system ability to keep its biomass unchanged after drought)[14,15]. Previous attempts to quantitatively review the results of rainfall manipulative experiments, however, have mostly focusing on system resistance[5-7]. Therefore, we are still uncertain about the potential trade-offs between those properties and their relative importance to underpin the overall grasslands ecological stability in response to drought[16]. Here, we searched for published studies reporting effects of experimental droughts on grasslands (see Methods). We compile 127 publications, that met our first set of inclusion criteria (Supplementary Note 1), and we identify which indices have been used to assess ecological stability properties and whether those indices provide comparable results. We further conducted a meta-analysis, using a subset of 49 studies (101 grasslands) that met a second set of inclusion criteria (Supplementary Note 1), in order to: (1) evaluate the overall grasslands resistance, recovery and resilience in response to drought; (2) identify potential trade-offs between those properties; and (3) quantify the relative importance of drought experimental features and prevailing climatic conditions to explain the differential stability across grasslands.

Results

Studies have imposed experimental droughts with very different characteristics, but some of them have failed in creating a real drought event. Rainfall manipulative experiments were conducted in 22 countries and 116 localities (Fig. 2), either in natural (41 %; N = 83), seminatural (40 %; N = 82) or manipulated grasslands (19 %; N = 39). Most of the studies (95 %; N = 110) were performed at mid-to-high latitudes (30–60°) in the northern hemisphere. Climatically, experiments were carried out in locations with mean annual temperature (MAT) ranging from 0.1° to 22.7°C (median = 8.3°C) and mean annual precipitation (MAP) ranging from 163 to 2,050 mm (median = 771 mm); while elevation (asl) ranged from 5 to 3,570 m (median= 530 m). Droughts were imposed using either permanent (91 %; N = 186) or automatic shelters (9 %; N = 18), with widely different sizes (range: 0.5 to 900 m^2, median = 9 m^2). Most studies reduced total rainfall amount (89 %; N = 181) rather than altered rainfall seasonality (7 %; N = 15) or both (4 %; N = 8). Drought duration ranged from 17 to 2,310 days, with most studies being short-term experiments (< one year) and only 9% of them (N = 4) lasting more than ten years. Similarly, most studies addressed a single drought event (46 %; N = 95), with a maximum of 15 sequential droughts (median = 2). In most of the studies (74 %; N = 150) drought intensity was measured as the percentage of rainfall intercepted by shelters, which ranged from 8 to 80 % reduction in annual precipitation. However, some authors defined experimental droughts in terms of their statistical extremeness (23 %; N = 41), that is, as the probability of occurrence based on precipitation historical records[11]. Extremeness ranged from 1 to 1000-year, where a 100-year drought event is the one that has 1% chance of occurrence in any given year. Moreover, historic control plots (sheltered plots receiving rainfall amounts correspondent to historical site precipitation) were used only in a few experiments (16 %; N = 32), in contrast to ambient controls that were used in most studies (unsheltered plots receiving ambient precipitation; 84 % N = 172). To allow for comparison on drought intensity across sites with distinct climate, we computed the standardized precipitation evapotranspiration index (SPEI) for a subset of 101 grasslands. We found that 27 % of the studies (N = 27) imposed moderate droughts (10-year extremeness), while 15 % (N = 15) created severe droughts (20-year) and only 7% (N = 7) conducted extreme drought events (50-year). Surprisingly, in 50% of the studies (N = 50) the ‘drought’ treatment actually did not experience a deficit in water balance, as the SPEI values felt in the range of near normal conditions. We also computed the SPEI for control plots either before (pre-drought), during (drought) and after (post-drought) the experimental drought period to evaluate if the control plots were exposed to abnormal dryer/wetter years. We found that control plots were usually exposed to near normal conditions, excepted in two study cases, where control plots were subjected to moderate wet conditions (Supplementary Table 2).

Numerous indices have been used to measure ecological stability, and their results were not comparable across studies. Only 13 % (N = 16) of the studies applied ecological stability indices to assess grassland resistance, recovery and/or resilience in response to drought. Firstly, we classified these indices in six groups, according to the index calculation method (Table 1). Subsequently, we calculated all those indices using above-ground biomass data extracted from studies fulfilling selection criteria 2 (101 grasslands, Supplementary Note 1). Finally, when we ordered grasslands according to their values on recovery, resistance, and resilience, different rankings were obtained depending on the indices considered (Supplementary Note 2). Therefore, results from studies using different groups of stability indices were not comparable.

On average, grasslands were resilient to drought and exhibited a trade-off between high recovery and low resistance. We used mean effect sizes to assess grasslands ecological stability in response to experimental droughts (Fig. 3). While resistance assumed significantly negative values (Fig. 3a), thus indicating that above-ground biomass production was strongly reduced during drought; recovery was significantly positive (Fig. 3b), indicating that grasslands were able to produce new biomass after drought alleviation. The biomass produced after drought was in general sufficient to compensate for the biomass lost during drought, thus grasslands exhibited nearly full resilience (Fig. 3c). There was also a trade-off between recovery and resistance (Fig. 4a). However, no predictable relationships were found between the others stability properties (Fig. 4 b-c).

Stability largely differed among grasslands, but drought features and climatic conditions were, in general, not important to explain those differences. We found large variation in resistance, recovery and resilience among grasslands, as indicated by the significant Q-tests and by the high percentages of between-studies heterogeneity (usually $I^2 > 50$ %) (Fig. 3). However, the variability in resistance and recovery was not explained by drought features (duration, frequency and intensity -SPEI) neither by climatic (MAT and MAP) or topographic (elevation) conditions. As shown in the Fig. 5 importance values for all those moderators remained below 0.8 (Fig.5 a-b). In contrast, variability in resilience was largely explained by drought duration, with longer droughts resulting in lower resilience (Fig. 5c). Results for simple meta-regressions testing the effect of moderators related to other experimental features are shown in Supplementary Table 3.

Sensitivity analysis. Our results were robust for outliers influence and non-independence effects (Supplementary Note 3). Nonetheless, evidence of publication bias was detected for resilience.

Discussion

This study provided key insights into how grasslands respond to experimental droughts and also highlighted the striking lack of consistency on methodological and analytical approaches to study the ecological stability of grasslands to drought.

We found that grasslands above-ground biomass production was, on average, resilient to experimental droughts (Fig. 3). Although we argue that this result might be underestimated, since few studies have subjected grasslands to severe or extreme drought events. Grasslands ecological stability was achieved more via recovery than via resistance (Fig. 3), thus leading to a trade-off between these properties (Fig.4a). This trade-off is consistent with previous studies[14,33] on supporting theoretical assumptions that resistance and recovery might be considered as two alternative strategies to achieve ecological stability[15,16]. It has been suggested that those strategies could be predicted by the life and evolutionary histories of resident plant species[34-36]. Therefore, plant communities dominated by fast-growing short-lived species (acquisitive resource-use strategy) would be expected to show low resistance, as they lack traits to sustain biomass production under water stress; but high recovery, as their high investment in seed production and fast-growth rates would support a rapid re-establishment and a shortly return to the non-drought biomass levels after drought alleviation. An opposite response would be predicted for communities dominated by slow-growing and long-lived species (i.e. conservative resource-use strategy)[34-36]. Since grasslands are mostly dominated by herbaceous species with an acquisitive strategy, we would indeed expect an overall resilience mainly driven by recovery than by resistance[5].

None of the six moderators we tested (drought duration, frequency and intensity, mean annual temperature, mean annual precipitation, elevation) were important to explain the variability in resistance and recovery across grasslands (Fig. 5 a-b). Therefore, the differential stability did not merely reflected differences in the drought treatments imposed or in the prevailing topo-climatic conditions, but they might be related to intrinsic biological differences across grasslands. In fact, considerable evidence from individual studies suggests that species diversity[13,25,28,29]; functional diversity (grass, legumes and forbs)[17,20,27] ; and plant life history traits (annual versus perennial)[31,37] can significantly influence grasslands' stability in response to drought. Moreover, in a meta-analysis assessing the effects of natural droughts on productivity, it was found that variation in resistance and recovery across dryland communities was largely explained by plant life history, with perennial plants showing high resistance but low recovery, and annual plants exhibiting the opposite pattern[33]. To date, most of the studies have not reported species and functional diversity at plot-level, nor provided the life history traits (or any other traits related to the acquisitive-conservative strategies) for both dominant and subordinate plant species. Therefore, we strongly recommend future studies to provide such information, so that further quantitative analysis could be able to simultaneously assess the relative importance of both abiotic and biotic moderators to explain the differential stability across grasslands in response to experimental droughts.

Contrastingly, we found that variability in resilience was largely explained by drought duration (Fig. 5c), with longer droughts resulting in decreased grassland's resilience. Longer droughts could reduce grasslands' resilience either by decreasing their ability to recover from droughts (e.g. by causing depletion of seed bank and storage resources needed for re-establishment and resprouting of drought-sensitive species) or their ability to resist to droughts (e.g. by surpassing thresholds of tolerance and causing widespread mortality of drought-tolerant species)[33,35,38]. Under future climatic regimes where droughts are expected to be longer, and also more frequent and intense, we will might see a gradual decline in overall grasslands' biomass production. However, the evidence of publication bias detected for resilience indicates that these results must be interpreted with caution.

To further improve our mechanistic understanding of differential stability across grasslands and our ability to manage them to ensure a stable productivity in a drier world, we firmly encourage future experimental drought studies to: (1) Standardize stability indices: we found that different stability indices might provide largely divergent results (Supplementary Note 2); thus, reinforcing the need for standardization across studies[9,10]; (2) Report basic information necessary to compute stability indices and to assess moderators' effects: we show that each stability index has its pros and cons (Supplementary Note 2), thus regardless of the metric eventually selected as standard, we recommend studies to fully report biomass and precipitation data, which could be further used to recalculate stability using any of the indices. Moreover, studies should report sufficient site-level characteristics for further assessment of moderator effects. Previous publications have provided an excellent guidance about which information studies should report[39-41]; (3) Evaluate the three stability properties simultaneously: we found that the majority of studies have only assessed grasslands resistance; thus, possibly overestimating the negative effects of droughts over grasslands production and also impairing a better assessment of the trade-offs between stability properties[16]. We then recommend the evaluation of both drought (resistance) and post-drought system responses (recovery and resilience); (4) Impose multiple and more extreme drought treatments: we showed that many studies have actually failed to impose experimental droughts and have been assessed drought intensity by using inappropriate metrics (percentage of rainfall intercept). Future experiments must consider local historical precipitation records[12,42]; and must use more meaningful metrics (such as SPEI) to determine the magnitude of the drought treatment[8,13,33]. Whenever possible, they should also impose multiple drought levels, including more extreme droughts, in order to better identify thresholds in grasslands response[43]; and (5) Evaluate stability of tropical grasslands: most of the experiments were conducted in the northern hemisphere (Fig. 2), thus caution is required in extrapolating our results to tropical grasslands. Future studies should then expand the geographic extent of drought experiments[8,39]. Collaborative initiatives, such as the Drought-Net[10], could help on this and also could lead to the standardization of methodological approaches across studies, thus facilitating further assessments of how and why plant communities differ in their stability in response to drought.

Despite our study was focused on grasslands, we argue that the main shortcomings listed here and improvements suggested should guide future experimental studies evaluating effects of drought in different ecosystems

Methods

Data search, inclusion criteria and meta-data extraction

We searched for peer-reviewed scientific articles published before June 2018 that conducted rainfall manipulative experiments on grasslands across the globe. 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 Supplementary Note 1 for the list of included and excluded references). We further applied a first set of inclusion criteria, retaining only studies where: (1) experiments were conducted in the field (not greenhouses, growth chambers or individual pots) in natural, seminatural and/or manipulated plant communities (e.g. seeding in the field, transplantation of intact monoliths, mesocosms) composed by at least two species (not monocultures); (2) experiments were conducted in grasslands (i.e. herbs comprising more than 50 % of plant cover); (3) precipitation amount was decreased and/or rainfall seasonality was experimentally altered via rain-out shelters; and (4) response variables were measured in adult vascular plants (not seedlings) on both drought (treatment) and non-drought (control) plots. After cross-referencing, removal of duplicates, and application of the above inclusion criteria, we retained a total of 127 articles (204 study cases) published during 1989 and 2018, from which we extracted the following metadata: study location (coordinates, country and locality); elevation (asl m); mean annual temperature (MAT, °C); mean annual precipitation (MAP, mm); vegetation type (natural, seminatural or manipulated); shelter area (m^2); shelter type (permanent - shelters covered vegetation all over the experimental drought period; or automatic - shelters were able to detect rainfall and only covered vegetation during rainfall events, thus minimizing the potential undesired shelters effects on microclimatic conditions, such as increases in air and soil temperatures and vapour pressure deficit); control plot type (ambient - unsheltered control plots receiving ambient precipitation; or historic - sheltered control plots which are irrigated to receive rainfall amounts correspondent to historical precipitation); drought type (amount - precipitation amount reduced by a fixed percentage; or seasonality - rainfall seasonality altered without significant changes in precipitation amount); drought duration (total number of days with shelters intercepting rainfall); drought frequency (number of sequential drought events) and drought magnitude (either reported as the percentage of rainfall intercepted by shelters; or as extremeness, that is, the probability of occurrence based on precipitation historical records); response variables measured on plants and whether those studies reported indices of stability (Supplementary Table 1). Subsequently, we computed stability indices and Hedges' g effect sizes for a subset of 49 articles (101 study cases) which met a second set of inclusion criteria (Supplementary Table 2): (1) precipitation amount was reduced in drought plots, i.e. experiments altering only rainfall seasonality were eliminated from the quantitative analysis due to their small sample size (N = 23) and their peculiar experimental design which precluded comparisons to amount drought types; (2) studies provided sample sizes (n) and mean above-ground biomass ± standard deviation (or standard error or confidence interval) measured for the whole community (not individual species) in both control and drought plots. We decided to assess stability only in terms of above-ground biomass production, because it was the response variable most frequently reported (83 %; N = 170); and (3) studies informed starting and ending dates of the experimental drought, as well as, the percentage of rainfall intercepted by rain-out shelters. That information was further needed to estimate drought intensity as SPEI (standardized precipitation evapotranspiration index). Metadata was extracted directly from text or tables, or indirectly from figures using the Figure Calibration plugin in the Image J software, version 1.48. When elevation, MAT, and/or MAP were not provided in the original studies, they were obtained from Geoplaner (http://www.geoplaner.com) and WorldClim (http://www.worldclim.org).

Ecological stability indices and effect sizes

We gathered all ecological stability indices reported in the 127 articles and grouped them according to the type of stability property assessed (resistance – RT; recovery – RC; resilience – RS; or relative resilience – RRS) and the way the stability was quantified (group 1 – raw difference; group 2 – simple ratio; group 3 and 4 – compounded ratio; group 5 – natural logarithm ratio; group 6 – standardized ratio). Although any of the selected studies have measured recovery as raw difference (RC1) or compounded ratio (RC3 and RC4), or resilience as compounded ratio (RS3) we decided to add those indices in the Table 1 to enable a full comparison across the six indices groups, thus resulting in 19 stability indices. We then computed all those indices, using the biomass data provided in the subset of 101 study cases, and evaluated whether those indices provided comparable results (Supplementary Note 2).

To assess overall grasslands stability in response to experimental droughts we computed mean Hedges' g effect sizes for RT, RC and RS. We used mean effect sizes, instead of selecting one of the six groups of stability indices, because in the former approach studies were weighted by their precision (more precise studies received larger weights), thus providing a more unbiased overall effect[44]. Therefore, Hedges' g effect sizes for RT (N=101) were obtained as the difference of above-ground biomass between drought and control plots during drought period (BD - BC); for RC (N = 19), as the difference between post-drought and drought plots (BD_{pd} - BD); and for RS (N = 19), as the difference between drought and control plots during post-drought period (BD_{pd} - BC_{pd}). 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[44]. Negative/positive values indicated decreases/increases in above-ground biomass. Mean effect sizes for each stability property were estimated under random effect-models using restricted maximum likelihood estimation and Knapp & Hartung adjustment within the 'metafor' R package[45]. Lastly, linear regressions were used to test for potential trade-offs among stability properties. Assumptions of normality of residuals and homogeneity of variance of the linear regressions were checked graphically.

SPEI computation

We calculate the SPEI, as a measure of experimental drought intensity, for each of the 101 study cases. We decided to used SPEI, instead of the percentage of rainfall intercepted, since the former index accounts for the interannual precipitation variability and enables reliable comparisons across studies regardless the season, location or climate[46]. To compute SPEI we firstly obtained, for each study location, the monthly precipitation and potential evapotranspiration data (Thornthwaite-Mather model) from the TerraClimate database, which provides data of climatic water balance from 1958 to 2017 at a spatial resolution 2.5 arc minutes (~ 4 km)[47]. 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. percentage of rainfall intercepted 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 have applied an experimental drought by intercepting 50% of rainfall from March to September 2005, we would simply divided by 2 the monthly precipitation values downloaded from TerraClimate for that specific period. Subsequently, we used the ‘SPEI’ R package[46] to calculate the one-month derivation of SPEI for the whole drought period. Lastly, we estimated the drought intensity for each drought event by averaging SPEI values for the drought period. We also evaluated potential confounding effects of interannual precipitation variability over experiments by computing the SPEI for the control plots. In this case, we used the original values of monthly precipitation extracted from TerraClimate and checked if control plots were exposed to naturally dryer/wetter conditions, either before (pre-drought; mean SPEI for 12 months before the starting drought date), during (drought; mean SPEI for the whole drought period) or after the experimental drought period (post-drought; mean SPEI for the period between the end of the drought and the post-drought biomass harvesting). SPEI for control plots were not computed for studies using historical control plots. For a given site, 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, with negative/positive values indicating dryer/wetter conditions as follow: extreme drought $\mathrm{SPEI} < -2$; severe drought: $-1.50 < \mathrm{SPEI} < -1.99$; moderate drought: $-1.00 < \mathrm{SPEI} < -1.49$; near normal conditions: $-0.99 < \mathrm{SPEI} < 0.99$; moderate wet:$1.00 < \mathrm{SPEI} < 1.49$; severe wet: $1.50 < \mathrm{SPEI} < 1.99$; and extreme wet: $\mathrm{SPEI} > 2.0$[46,48]. Regarding the extremeness, droughts classified as moderate based on their SPEI values have a 10-year probability of occurrence (or 10 events at each 100 years); severe droughts a 20-year and extreme droughts a 50-year probability[48]. We recognize two main limitations of using SPEI as a measure of experimental drought intensity. Firstly, climatic water deficits do not necessarily lead to plant water deficits; and secondly, there could be mismatches between the precipitation inputs extracted from TerraClimate and the actual precipitation received in the experimental sites.

However, as very few experiments provided the necessary data to calculate more meaningful metrics of drought intensity[12], we believe our approach is justified.

Effects of moderators

To verify whether grasslands varied in their stability in response to drought we tested for significant between-studies heterogeneity using the Cochran's Q test, and estimated the percentage of total variance that could be attributed to between-studies rather than to sampling error (I^2). Subsequently, to quantify the relative importance of drought experimental features (duration, frequency and intensity- SPEI) and prevailing climatic (MAT and MAP) and topographic conditions (elevation) to explain the differential stability across grasslands we used multi-model inference based on information theory[49]. In this analysis all possible unique models involving the above set of moderators were fitted using random-effect meta-regression models, and ranked according to the second-order Akaike information criterion (AICc). For each moderator, we computed the model-averaged estimate and relative importance. Relative importance was obtained by summing the Akaike weights of all models that included the moderator of interest. Values of relative importance greater than 0.8 indicated that the moderator was important to explain the variation in the stability properties. Separated models were constructed for resistance, recovery and resilience using functions from the 'glmulti'[50] and 'metafor' R packages. Additionally, we fitted simple meta-regression models to evaluate the effect of moderators related to experimental features (vegetation type; shelter area; shelter type; control plot type; and drought discreteness).

Sensitivity analysis

We performed sensitivity analysis, using the 'metafor' R package, to test for the effect of outliers, publication bias and non-independence problems. To deal with non-independence in studies performed with more than two levels of species biodiversity, we restricted our analysis to the most diverse treatment, that is, the treatment with the highest number of species. When a single study reported results from experiments conducted in different locations, each locality was treated as an independent study case. Subsequently, we used hierarchical models to test for non-independence of studies conducted at the same locality. In studies reporting sequential drought events, stability indices and effect sizes were computed using only the biomass data for the last drought event (final). To evaluate the extent to which our results were robust to this decision, we also computed the effect sizes using biomass data averaged across all drought events (averaged) and presented their results in the Supplementary Note 3.

Data availability

Falta

References

1. Dixon, A. P., Faber-Langendoen, D., Josse, C., Morrison, J., & Loucks, C. J. Distribution mapping of world grassland types. J. Biogeogr. 41, 2003–2019 (2012).

2. Knapp, A.K., & Smith, M.D. Variation among biomes in temporal dynamics of aboveground primary production. Science 291, 481-484 (2001).

3. Dai, A. Increasing drought under global warming in observations and models. Nature Clim. Change 3, 52-58 (2013).

4. Min, S.-K., Zhang, X., Zwiers, F.W. & Hegerl, G.C. Human contribution to more-intense precipitation extremes Nature 470, 378–381 (2011).

5. Stuart-Haëntjens, E., et al. Mean annual precipitation predicts primary production resistance and resilience to extreme drought. Sci. Total Environ. 636, 360–366 (2018).

6. Wilcox, K. R., et al. Asymmetric responses of primary productivity to precipitation extremes: a synthesis of grassland precipitation manipulation experiments. Global Chang. Biol. 23, 4376–4385 (2017).

7. Wu, Z., Dijkstra, P., Koch, G. W., Peñuelas, J., & Hungate, B. A. Responses of terrestrial ecosystems to temperature and precipitation change: a meta-analysis of experimental manipulation. Global Chang. Biol. 17, 927–942 (2011).

8. Hoover, D. L., Wilcox, K. R., & Young, K. E. Experimental droughts with rainout shelters: A methodological review. Ecosphere 9, e02088 (2018).

9. Ingrisch, J., & Bahn, M. Towards a comparable quantification of resilience. Trend. Ecol. Evol. 33, P251-259 (2018).

10. Smith, M. D., Wilcox, K. R., Power, S. A., Tissue, D. T. & Knapp, A. K. Assessing community and ecosystem sensitivity to climate change – toward a more comparative approach. J. Veg. Sci. 28, 235-237 (2017).

11. Knapp, J., et al. Pushing precipitation to the extremes in distributed experiments: recommendations for simulating wet and dry years. Ecosystems 20, 301 (2017).

12. Vicca, S., et al. Urgent need for a common metric to make precipitation manipulation experiments comparable. New Phytol. 195, 518–522 (2012).

13. Kreyling, J., et al. Drought effects in climate change manipulation experiments: quantifying the influence of ambient weather conditions and rain-out shelter artifacts. Ecosystems 20, 301-3015 (2017).

14. Gazol, A., Camarero, J. J., Anderegg, W. R., Vicente-Serrano, S. M. & Fortin, M. Impacts of droughts on the growth resilience of Northern Hemisphere forests. Global Ecol. Biogeogr. 26, 166-176 (2017).

15. Hodgson, D., McDonald, J. L., & Hosken, D. J. What do you mean, ‘resilient’? Trend. Ecol. Evol. 30, 503–506 (2015).

16. Donohue, I., et al. On the dimensionality of ecological stability. Ecol. Lett. 16, 421–429 (2013).

17. Byrne, K. M., Adler, P. B., & Lauenroth, W. K. Contrasting effects of precipitation manipulations in two Great Plains plant communities. J. Veg. Sci. 28, 238–249 (2017).

18. Carter, D. L., & Blair, J. M. High richness and dense seeding enhance grassland restoration establishment but have little effect on drought response. Ecol. Appl. 22, 1308–1319 (2012).

19. Hoekstra, N. J., Suter, M., Finn, J. A., Husse, S., & Lüscher, A. Do belowground vertical niche differences between deep- and shallow-rooted species enhance resource uptake and drought resistance in grassland mixtures? Plant Soil. 394, 21–34 (2015).

20. Hofer, D., et al. Yield of temperate forage grassland species is either largely resistant or resilient to experimental summer drought. J. Appl. Ecol. 53, 1023–1034 (2016).

21. Ingrisch, J., et al. Land use alters the drought responses of productivity and CO2 fluxes in mountain grassland. Ecosystems 21, 689–703 (2018).

22. Kreyling, J., et al. Species richness effects on grassland recovery from drought depend on community productivity in a multisite experiment Ecol. Lett. 20, 1405-1413 (2017).

23. Lanta, V., Doležal, J., Zemková, L., & Lepš, J. Communities of different plant diversity respond similarly to drought stress: experimental evidence from field non-weeded and greenhouse conditions. Naturwissenschaften 99, 473–482 (2012).

24. Mariotte, P., Vandenberghe, C., Kardol, P., Hagedorn, F., & Buttler, A. Subordinate plant species enhance community resistance against drought in semi-natural grasslands. J. Ecol. 101, 763–773 (2013).

25. Pfisterer, A. B., & Schmid, B. Diversity-dependent production can decrease the stability of ecosystem functioning. Nature 416, 84–86 (2002).

26. Shinoda, M., Nachinshonhor, G. U., & Nemoto, M. Impact of drought on vegetation dynamics of the Mongolian steppe: A field experiment. J. Arid Environ. 74, 63–69 (2010).

27. Stampfli, A., Bloor, J. M. G., Fischer, M., & Zeiter, M. High land-use intensity exacerbates shifts in grassland vegetation composition after severe experimental drought. Global Chang. Biol. 24, 2021–2034 (2018).

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

29. Wagg, C., et al. Plant diversity maintains long-term ecosystem productivity under frequent drought by increasing short-term variation. Ecology 98, 2952–2961 (2017).

30. Wang, Y., Yu, S., & Wang, J. Biomass-dependent susceptibility to drought in experimental grassland communities. Ecol. Lett. 10, 401–10 (2007).

31. Wang, Z., et al. 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 Soil. 389, 59–71 (2014).

32. Yahdjian, L., & Sala, O. E. Vegetation structure constrains primary production response to water availability in the Patagonian steppe. Ecology 87, 952–962 (2006).

33. Ruppert, J. C., et al. Quantifying drylands' drought resistance and recovery: The importance of drought intensity, dominant life history and grazing regime. Global Chang. Biol. 21, 1258–1270 (2015).

34. Grime, J.P. Plant strategies, vegetation processes and ecosystem properties. 2nd (John Wiley & Sons, Chichester, UK, 2001).

35. Grime, J. P. and Pierce, S. The evolutionary strategies that shape ecosystems. (Wiley-Blackwell, Chichester, UK, 2012).

36. Reich, P. B. The world-wide 'fast–slow' plant economics spectrum: a traits manifesto. J. Ecol. 102, 275–301 (2014).

37. Hoeppner, S. S., & Dukes, J. S. Interactive responses of old-field plant growth and composition to warming and precipitation. Global Chang. Biol. 18, 1754–1768 (2012).

38. Estiarte, M., et al. Few multi-year precipitation-reduction experiments find a shift in the productivity-precipitation relationship. Global Chang. Biol. 22, 2570-2581 (2016).

39. Beier, C., et al. Precipitation manipulation experiments- challenges and recommendations for the future. Ecol. Lett. 15, 899–911 (2012).

40. Smith, N. G., et al. Toward a better integration of biological data from precipitation manipulation experiments into Earth system models Rev. Geophys. 52, 412–434 (2014).

41. Gerstner, K., et al. Will your paper be used in a meta-analysis? Make the reach of your research broader and longer lasting. Methods Ecol. Evol. 8, 777-784 (2017).

42. Lemoine, N. P., et al. Terrestrial Precipitation Analysis (TPA): A resource for characterizing long-term precipitation regimes and extremes. Methods Ecol. Evol. 7, 1396-1401 (2016).

43. Knapp, A. K., et al. A reality check for climate change experiments: Do they reflect the real world? Ecology 99, 2145-2151 (2018).

44. Borenstein, M., Hedges, L.V., Higgins, J.P.T. & Rothstein, H.R. Introduction to meta-analysis (Wiley, Chichester, 2009).

45. Viechtbauer, W. Conducting Meta-Analyses in R with the metafor Package. J. Stat. Softw. 36, 1–48 (2010).

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

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

48. World Meteorological Organization, Standardized Precipitation Index: user guide (World Meteorological Organization, Geneva, Switzerland, 2012).

49. Burnham, K.P., & Anderson, D.R., Model selection and multimodel inference: a practical information-theoretic approach, 2nd ed. (Springer-Verlag, New York, 2002).

50. Calcagno, V., de Mazancourt, C. glmulti: An R package for easy automated model election with (generalized) linear models. J. Stat. Softw. 34, 1 – 29 (2010).

Acknowledgements

We are grateful to T. Moulton, J. Boelsums and F. Melo for their constructive comments on the first draft of this manuscript. We also thank to S. Rifai for helping in obtaining and manipulating TerraClimate data. We thank RUFFORD FOUNDATION 18749-1, FAPERJ and Prociencia Program for financial support. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001

Figures

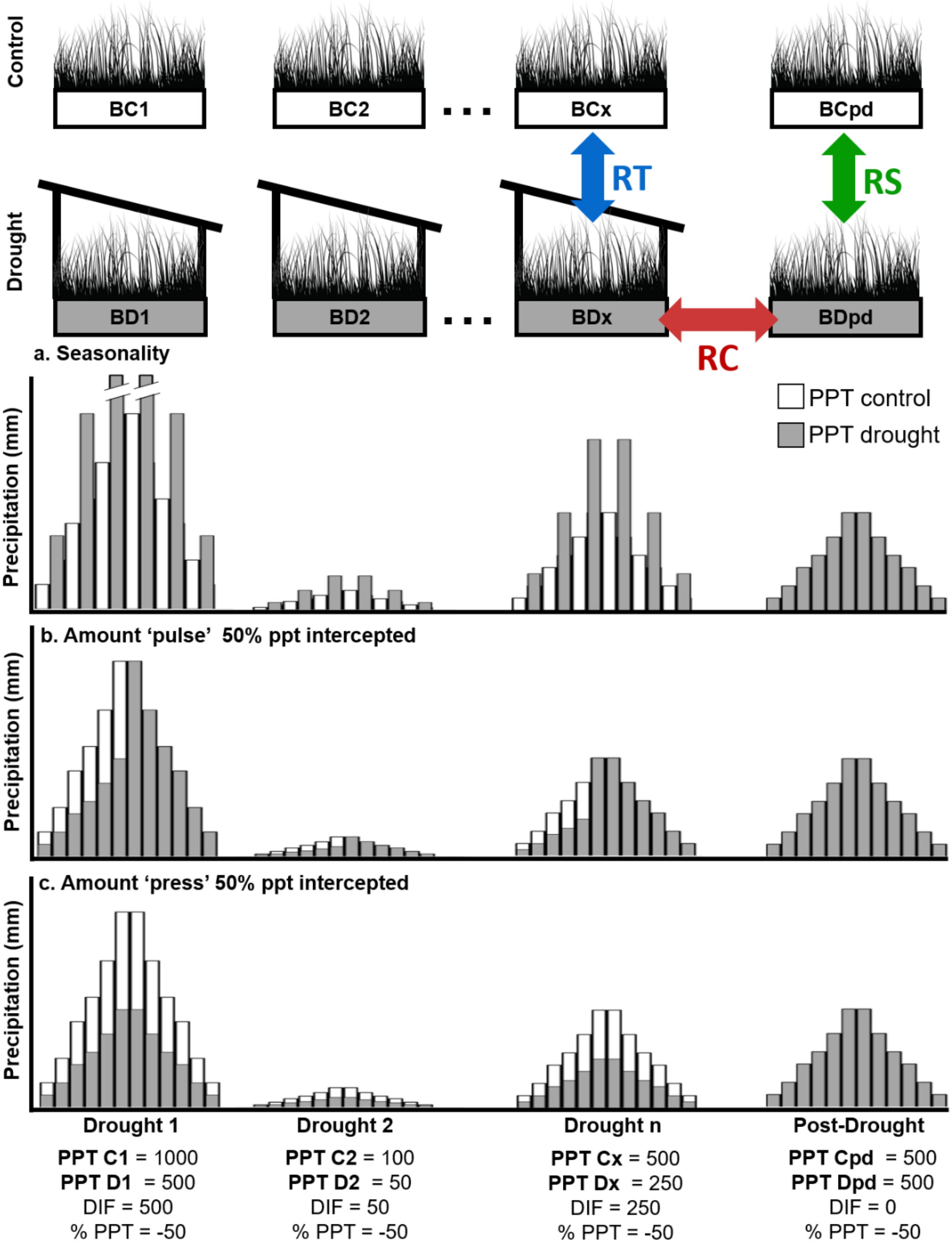

Figure 1. Overview of rainfall manipulative experiments. Experiments have used shelters to intercept rainfall in drought plots thus creating one or multiple sequential drought events (n=1, 2… x). Response variables, most frequently above-ground biomass, are measured on control (BC_n) and drought plots (BD_n) during the drought period, and after shelter removal (post-drought: BC_{pd} and BD_{pd}). Biomass data can be further used to estimate three stability properties: resistance (RT), the ability to keep biomass unchanged during drought ($BC_n = BD_n$); recovery (RC), the ability to produce

biomass after drought ($BD_{pd} > BD_n$); and resilience (RS), the ability to keep biomass unchanged after drought ($BC_{pd} = BD_{pd}$). Precipitation incomes (mm) for control ($PPTC_n$) and drought plots ($PPTD_n$) are shown for three droughts types: (a) seasonality, rainfall is repacked into larger events and longer dry intervals without changes in total amount; (b) amount 'pulse', rainfall amount is reduced by a fixed percentage (%PPT = -50 %) during shorter and interrupted periods; and (c) amount 'press', rainfall is reduced during longer and uninterrupted periods. Although the %PPT remains the same across drought period (-50 %), larger absolute differences between control and drought plots (DIF) are observed in wetter (Drought 1) than in drier years (Drought 2). Moreover, if the interannual variation in precipitation in one site is greater than that in another site, a drought of -50 % intensity is likely to have less effect on the former site. Therefore, %PPT is not a reliable index of drought intensity and it is not comparable across sites with distinct climate.

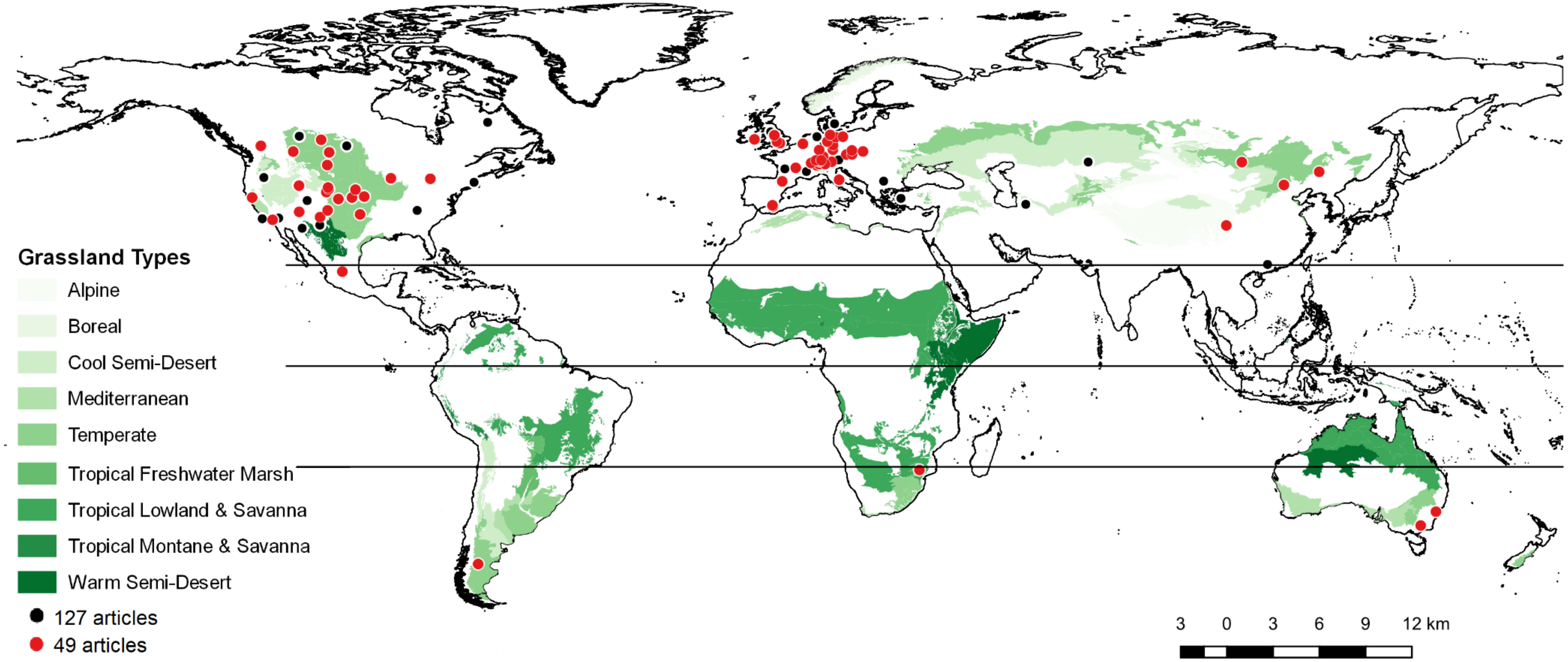

Figure 2. Global map displaying site location of rainfall exclusion experiments conducted in grasslands. Black dots indicate the location of 204 study cases (extracted from 127 articles that met the first set of inclusion criteria), while red dots indicate the location for 101 study cases (extracted from 49 articles that met the second set of inclusion criteria) that were used for quantitative analysis (inclusion criteria are listed in Supplementary Note 1). Grasslands types and distributions were obtained from Dixon et al. 2012.

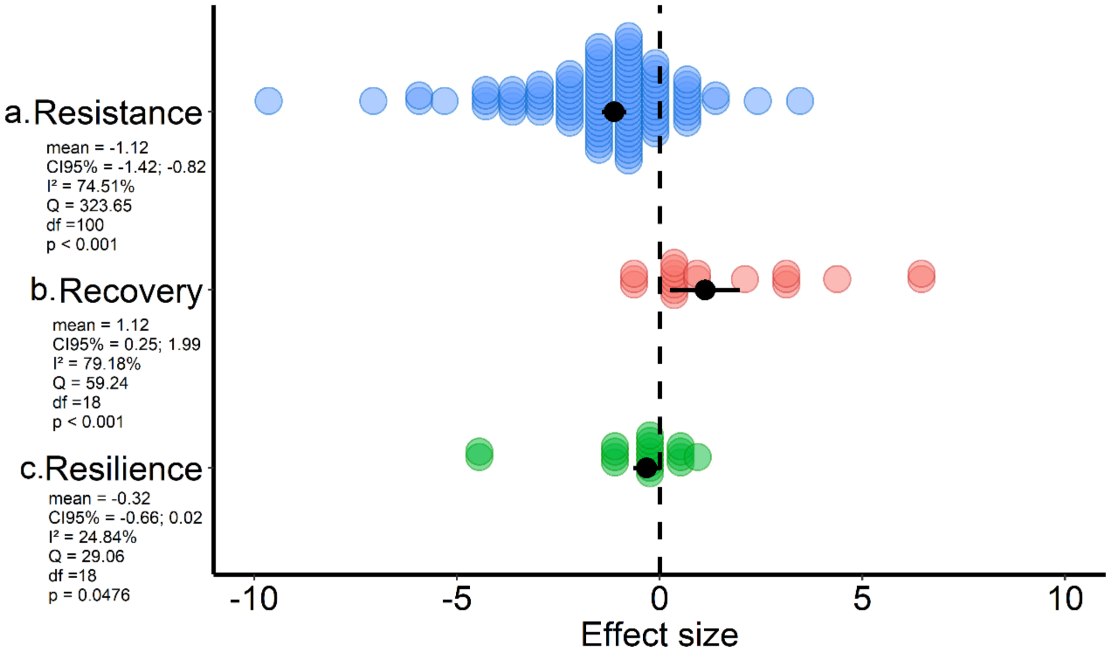

Figure 3. Hedges'g effect sizes for resistance (a), recovery (b) and resilience (c) of grasslands' above-ground biomass production (N = 101) in response to experimental droughts. In studies reporting sequential drought events effect sizes were computed using biomass data only for the last experimental drought. Negative/positive values indicate decreases/increases in biomass in treatment plots compared to control plots. Black dots/lines indicate mean effect sizes and its 95 % confidence interval. Q-statistics, degree of freedom (df); p-value and the proportion of the observed variance that reflects true differences among studies (I^2) testing for between-studies heterogeneity.

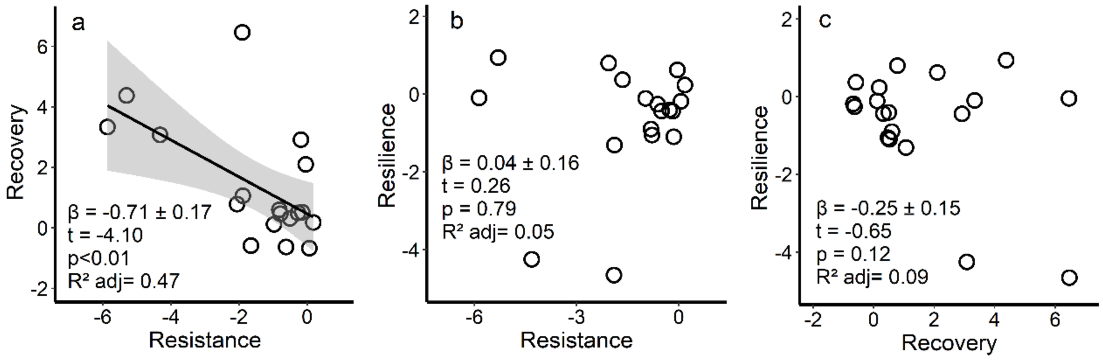

Figure 4. Linear regressions testing for potential trade-offs among stability properties. Resistance, recovery and resilience were computed as Hedges'g effect sizes measuring the effects of experimental droughts on above-ground biomass production of 101 grasslands. In studies reporting sequential drought events effect sizes were computed using biomass data only for the last experimental drought.

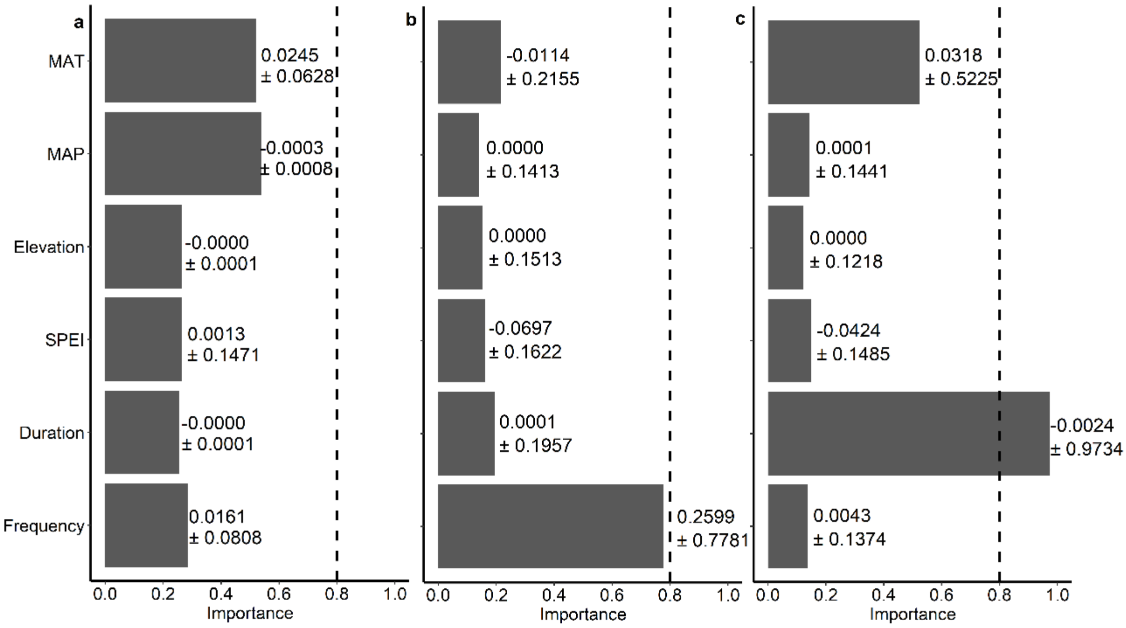

Figure 5. Model-averaged importance of six moderators explaining grasslands resistance (a), recovery (b) and resilience (c) in terms of above-ground biomass production in response to experimental droughts. Moderators: mean annual temperature (MAT, ºC); mean annual precipitation (MAP, mm); elevation asl (m); standardized precipitation evapotranspiration index (SPEI) as a measure of drought intensity; drought duration (days); and drought frequency (number of sequential drought events). Numbers next to the bars indicate model-averaged parameter estimates ± 95 % confidence interval. The vertical dashed line drawn at 0.80 can be used as a cut-off to identify important moderators. In studies reporting sequential drought events effect sizes were computed using biomass data only for the last experimental drought.

Tables

Table 1. Stability indices used to assess resistance (RT), recovery (RC), resilience (RS) and relative resilience (RRS) of grasslands in response to experimental droughts.

Group	Resistance (RT)	Recovery (RC)	Resilience (RS)	Relative resilience (RRS)
1	RT1 = BD – BC [18, 23, 25, 38, 31]	RC1 = BD_{pd} – BD*	RS1 = BD_{pd} – BC_{pd} [28]	RRS1 = (BD_{pd} / BC_{pd}) - (BD / BC) [21]
2	RT2 = BD/BC [20-22, 24, 28, 31]	RC2 = BD_{pd}/BD [18]	RS2 = BD_{pd} / BC_{pd} [18, 20-25, 28]	RRS2 = (BC_{pd} /(BD_{pd} - BC_{pd}))/(BC/ (BD - BC))[29]
3	RT3 = (BD - BC)/ BC [19, 31-32]	RC3 = (BD_{pd} - BD)/ BD*	RS3 = (BD_{pd} - BC_{pd})/ BC_{pd} *	
4	RT4 = BC/ (BD – BC) [29]	RC4 = BD/ (BD_{pd} – BD)*	RS4 = BC_{pd} /(BD_{pd} - BC_{pd}) [29]	
5	RT5 = ln (BD/BC) [17, 18, 20, 27, 32]	RC5 = ln (BD_{pd} / BD) [18]	RS5 = ln (BD_{pd} / BC_{pd}) [25]	
6	RT6 = $\frac{(BD/BC)/\ (PPT_D/PPT_C)}{BC/PPT_C}$[26]	RC6 = $\frac{(BDpd/BD)/\ (PPT_{PD}/PPT_D)}{BC/PPT_C}$[28]		

Stability indices classes: group 1 – raw difference; group 2 – simple ratio; group 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 between brackets indicate references reporting each index: 17. Byrne et al. 2017; 18. Carter & Blair 2012; 19. Hoekstra et al. 2015; 20. Hofer et al. 2016; 21. Ingrisch et al. 2018; 22. Kreyling et al. 2017; 23. Lanta et al. 2012; 24. Mariotte et al. 2013; 25. Pfisterer & Schmid 2002; 26. Shinoda et al. 2010; 27. Stampfli et al. 2018; 28. Vogel et al. 2012; 29. Wagg et al. 2017; 30. Wang et al. 2007; 31. Wang et al. 2014; 32. Yahdjian & Sala 2006; * indices not reported in any of the 127 references used in the meta-analysis, but derived here to fill all indices groups.

Additional Information

Supplementary information

Supplementary Note 1. List of inclusion criteria and references excluded from and included in the meta-analysis.

Supplementary Note 2. Comparisons among stability indices used to assess grasslands above-ground biomass resistance, recovery and resilience in response to experimental droughts.

Supplementary Note 3. Results of sensitivity analysis testing for outliers influence, publication bias and non-independence.

Supplementary Table 1. Metadata extracted from 127 articles (204 study cases) reporting on the effects of rainfall manipulative experiments on grasslands above-ground biomass production.

Supplementary Table 2. Metadata, stability indices and Hedges' g effect sizes computed for 56 articles (112 study cases) reporting on the effects of rainfall exclusion experiments on grasslands above-ground biomass production. In studies reporting sequential droughts, effect sizes were computed either using biomass data only for the last drought (final; 49 articles and 101 study cases) or using biomass averaged across sequential drought events (averaged; 54 articles and 106 study cases). SPEI: standardized precipitation evapotranspiration index; Resistance (RT1 - RT6), recovery (RC1 - RC6), resilience (RS1 - RS5), and relative resilience (RRS1 - RRS2) indices. Hedges' g effect sizes for resistance (RT), recovery (RC) and resilience (RS), associated effect sizes variance (RT var, RC var, RS var) and 95% confidence interval (95 % CI).

Supplementary Table 3. Results of simple meta-regressions testing the effects of moderators related to experimental features (vegetation type; shelter area; shelter type; control plot type; and drought discreteness) on the Hedges' g effect sizes for grasslands resistance (RT), recovery (RC) and resilience (RS) in response to experimental droughts. In studies reporting sequential drought events effect sizes were computed either using biomass data only for the last experimental drought (final; 49 articles and 101 study cases) or using biomass averaged across sequential drought events (averaged; 54 articles and 106 study cases). AICc: second-order Akaike information criterion (AICc); T^2: between studies variance; I^2: percentage of the observed variance that reflects true differences among studies; R^2: percentage of the true variance explained by the moderator; F or Q statistics and p-value for quantitative and qualitative moderators, respectively; df: degree of freedom; Estimate parameter [95 % confidence interval].

Competing interests

No competing interests to declare

Author contributions

B.H.P. Rosado outlined the scope which was developed and improved by I.S. Matos. I. Oliveras included analytical approaches and I.S. Matos led the writing of the manuscript and conducted the analysis. All authors contributed to editing, revising and writing.

Materials & Correspondence

Ilaíne Silveira Matos