Research

New Phytologist

Methods

Towards a statistically robust determination of minimum water potential and hydraulic risk in plants

Jordi Martínez-Vilalta[1,2], Louis S. Santiago[3], Rafael Poyatos[1,2], Llorenç Badiella[2], Miquel de Cáceres[1,4], Ismael Aranda[5], Sylvain Delzon[6], Alberto Vilagrosa[7] and Maurizio Mencuccini[1,8]

[1]CREAF, Bellaterra (Cerdanyola del Vallès), Catalonia E08193, Spain; [2]Universitat Autònoma de Barcelona, Bellaterra (Cerdanyola del Vallès), Catalonia E08193, Spain; [3]Department of Botany & Plant Sciences, University of California, 2150 Batchelor Hall, Riverside, CA 92521, USA; [4]Joint Research Unit CTFC – AGROTECNIO, Solsona 25280, Spain; [5]Centro de Investigación Forestal, Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria, Carretera Coruña Km 7.5, Madrid E-28040, Spain; [6]INRAE, BIOGECO, Univ. Bordeaux, Pessac 33615, France; [7]CEAM Foundation, Joint Research Unit University of Alicante-CEAM, Dept Ecology, University of Alicante, Carr. de San Vicente del Raspeig, PO Box 99, Alicante 03080, Spain; [8]ICREA, Pg. Lluís Companys 23, Barcelona 08010, Spain

Author for correspondence:
Jordi Martínez-Vilalta
Email: Jordi.Martinez.Vilalta@uab.cat

Received: *8 January 2021*
Accepted: *14 June 2021*

New Phytologist (2021) **232:** 404–417
doi: 10.1111/nph.17571

Key words: drought resistance, drought tolerance, extreme values, hydraulic risk, hydraulic safety margin, minimum water potential, physiological limits.

Summary

- Minimum water potential (Ψ_{min}) is a key variable for characterizing dehydration tolerance and hydraulic safety margins (HSMs) in plants. Ψ_{min} is usually estimated as the absolute minimum tissue Ψ experienced by a species, but this is problematic because sample extremes are affected by sample size and the underlying probability distribution.
- We compare alternative approaches to estimate Ψ_{min} and assess the corresponding uncertainties and biases; propose statistically robust estimation methods based on extreme value theory (EVT); and assess the implications of our results for the characterization of hydraulic risk.
- Our results show that current estimates of Ψ_{min} and HSMs are biased, as they are strongly affected by sample size. Because sampling effort is generally higher for species living in dry environments, the differences in current Ψ_{min} estimates between these species and those living under milder conditions are partly artefactual. When this bias is corrected using EVT methods, resulting HSMs tend to increase substantially with resistance to embolism across species.
- Although data availability and representativeness remain the main challenges for proper determination of Ψ_{min}, a closer look at Ψ distributions and the use of statistically robust methods to estimate Ψ_{min} opens new ground for characterizing plant hydraulic risks.

Introduction

Extreme values are important in shaping the physiology, ecology and evolution of plants (Gutschick & BassiriRad, 2003). Physiological extremes are frequently used to characterize the functional limits of species and populations, describing the operational limits of organisms and their tolerance to stressful environmental factors (e.g. Bozinovic *et al.*, 2011; Kingsolver & Buckley, 2017). For this reason, extreme values are frequently used as parameters constraining key variables in ecological models, particularly in the context of predicting responses to ongoing climate change. In plant ecophysiology, examples of such extreme values include maximum (Hoshika *et al.*, 2018) and minimum stomatal conductance (Duursma *et al.*, 2019), as well as the minimum tissue water potential characterizing extreme exposure to drought (Bhaskar & Ackerly, 2006).

The minimum water potential (Ψ_{min}) describes the largest xylem tension that a species (or population or individual plant) experiences in the field. Because in actively transpiring plants xylem water potentials decline from roots to leaves along the soil–plant–atmosphere continuum, minimum water potentials are usually determined in leaves or twigs. Low xylem water potentials imperil turgor maintenance and metabolic function (Tyree & Jarvis, 1982; Bartlett *et al.*, 2012) and increase the risk of xylem embolism and the corresponding hydraulic dysfunction (Tyree & Sperry, 1988; Tyree & Zimmermann, 2002). Our current understanding is that hydraulic failure as a result of xylem embolism triggers drought-induced mortality in plants (Adams *et al.*, 2017; Choat *et al.*, 2018; Brodribb *et al.*, 2020). Consequently, the hydraulic safety margin (HSM), usually quantified as the difference between Ψ_{min} and embolism resistance (e.g. Ψ_{PLC50}, or the water potential causing 50% loss

of hydraulic conductivity), has become a standard metric for characterizing plant vulnerability to drought (Meinzer *et al.*, 2009; Choat *et al.*, 2012; Delzon & Cochard, 2014). This variable has the merit of integrating a measure of absolute stress tolerance determined in the laboratory (Ψ_{PLC50}) with a measure of (extreme) exposure at the tissue level (Ψ_{min}), thus yielding a promising measure of mortality risk (Anderegg *et al.*, 2016; Benito Garzón *et al.*, 2018). Note, however, that hydraulic failure may be associated with other metrics such as the water potential causing 88% loss of conductivity, which is frequently used to calculate HSM in angiosperms (Urli *et al.*, 2013), or the remaining water transport capacity in the xylem (Venturas *et al.*, 2016). In addition, leaf shedding may complicate the link between measures of hydraulic risk and mortality, particularly in drought-deciduous species (Kolb & Davis, 1994).

The determination of Ψ_{min} is challenging. Some studies have modelled Ψ_{min} from estimated minimum soil water content using pedotransfer functions and information on soil texture (Benito Garzón *et al.*, 2018). This approximation has the advantage of providing spatially explicit estimates of Ψ_{min} (and HSM), but it fails to account for the fact that Ψ_{min} is also driven by several, covarying plant traits (Bhaskar & Ackerly, 2006; Martínez-Vilalta & Garcia-Forner, 2017). A more common approach is to estimate Ψ_{min} as the absolute minimum water potential measured in a given species or population. However, this is also problematic, as sample extremes are inherently biased and the magnitude of the bias depends on the probability distribution of the underlying variable and on sample size (Head *et al.*, 2012). In particular, if Ψ_{min} is estimated as the absolute minimum Ψ recorded for a given species, its magnitude will necessarily get more negative (more extreme) as more Ψ measurements are taken, which seems a highly undesirable property when different species with unequal sample size are compared. Importantly, sampling effort is likely to vary directionally with drought exposure because species in drier and more variable environments tend to be sampled more intensively than species from wetter environments, where drought is not perceived as an important stress factor.

An alternative approach to estimating Ψ_{min} involves determining the Ψ associated with mortality. When this is done through opportunistic sampling under extreme drought conditions in the field, the problem becomes the identification of the physiological point of no return (cf. Hammond *et al.*, 2019; Sapes *et al.*, 2019), thus avoiding measurements taken in drying tissue once plants are effectively dead. This can be achieved by monitoring measured branches to ensure they survive the drought (e.g. Venturas *et al.*, 2016), but this is not always feasible. An alternative approach is to conduct drought simulation experiments in controlled environments and apply rewatering trials to identify the point of no return (e.g. Brodribb & Cochard, 2009; Urli *et al.*, 2013). Even these types of study, however, may bring plants beyond any conditions they could actually experience in the field, and thus it seems better to refer to the corresponding magnitude as Ψ_{lethal} (cf. Liang *et al.*, 2021).

Clearly, a more robust way of determining Ψ_{min} (and other ecophysiological extremes) is needed. Percentiles may be used to characterize extreme values from a Ψ distribution, but they make use of all observations and are hence highly dependent on the representativeness of the sample. Extreme value theory (EVT) is the branch of statistics devoted to the study of extreme deviations from the median, focusing on the tails of the underlying probability distributions as a means to estimate extreme values (Coles, 2001; Reiss & Thomas, 2007; Embrechts *et al.*, 2013). The two most common approaches involve grouping data into blocks of equivalent sizes and then fitting the distribution of block maxima (or minima) using the generalized extreme value (GEV) distribution, or using the generalized Pareto distribution (GPD) to fit the tail of the distribution of raw values ('peak over threshold' or POT method; cf. Fig. 1) (see Supporting Information Notes S1 for an extended introduction to EVT methods).

Extreme value theory methods are widely applied in finance, hydrology and climatology, among other fields, but their use in ecology is rare and basically limited to the characterization of environmental extremes (Katz *et al.*, 2005; Denny *et al.*, 2009) and the likelihood of species extinction (Burgman *et al.*, 2012). We are only aware of two studies that apply EVT to address relevant questions in plant functional ecology, one to characterize seed mass distributions (Edwards *et al.*, 2015) and the other to determine maximum stomatal conductance (Murray *et al.*, 2019).

Here, we take advantage of EVT to provide a general framework aimed at improving the estimation of Ψ_{min}. We first characterize the statistical properties of water potential distributions and their variability within species, focusing on distribution tails and taking advantage of a new database of raw water potential values assembled for this study. We then compare different methods to estimate Ψ_{min} and the associated uncertainties and biases using two empirical databases and synthetic data characterizing typical water potential distributions. Finally, we assess the implications of our results for the characterization of hydraulic risk and its variability among species, and provide guidelines to estimate Ψ_{min} and other ecophysiological extremes in a statistically robust manner.

Materials and Methods

Empirical water potential datasets

The first dataset was collected by Martínez-Vilalta *et al.* (2014) and hence we identify it as the 'MV14' dataset. It consists of a global collection of paired predawn and midday leaf water potential seasonal time courses for 159 species (3562 individual observations in total). Note that we use the term leaf Ψ for simplicity, but all measurements we refer to in the manuscript were conducted in leaves or (leafed) twigs. These data were extracted from published papers and hence correspond to averages across individuals for a given site and date, as given typically in figures. This dataset was used to study the impact of sample size on the estimation of Ψ_{min} for the subset of 17 species with a reasonably large

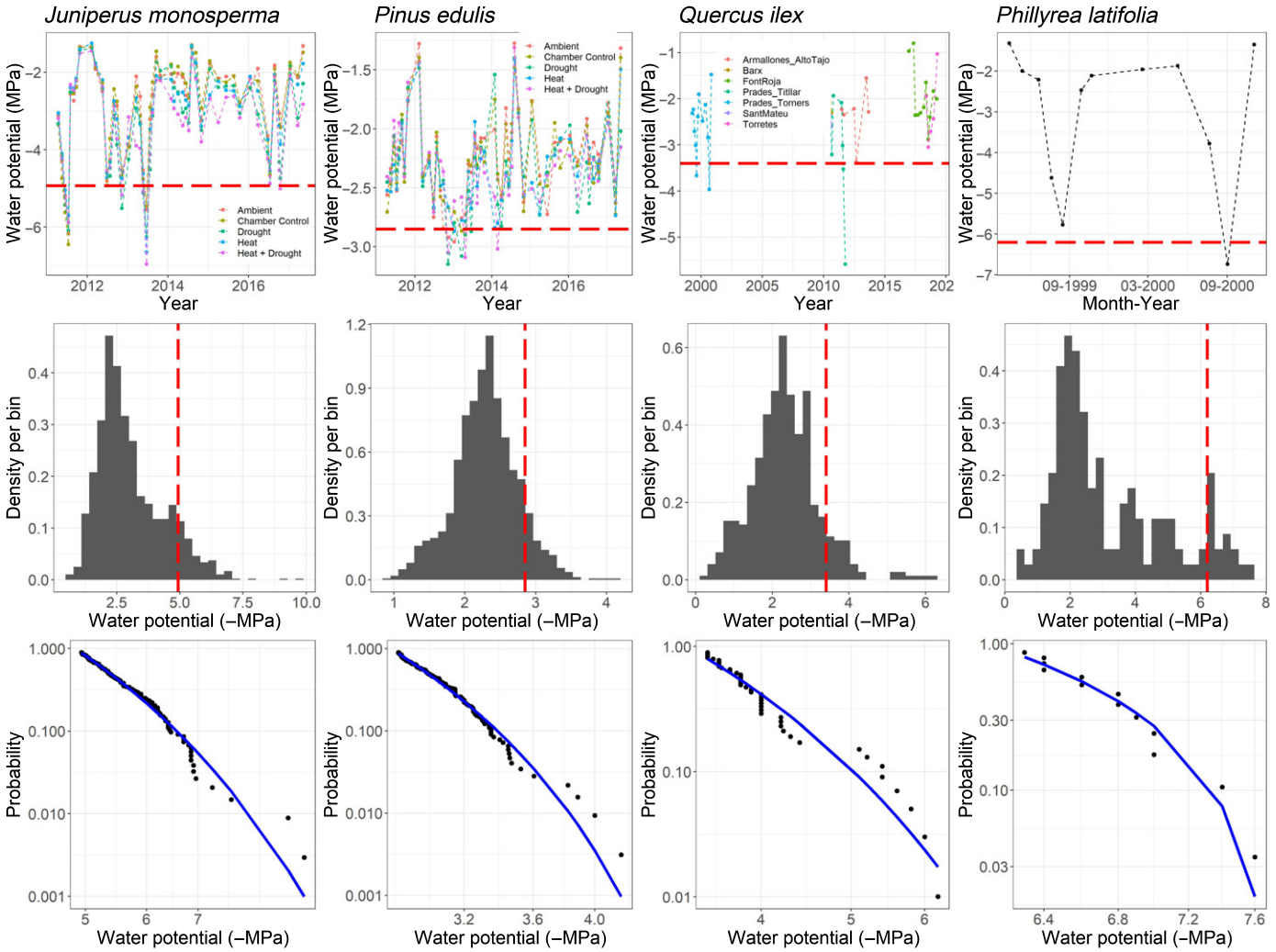

Fig. 1 Time series of measured midday leaf water potentials (top row), histograms showing the probability distributions of measured water potentials (middle row), and tails of the distributions (bottom row) for four representative species included in the 'raw' dataset. Each column corresponds to the species indicated on top. The dashed red lines in top and middle rows indicate the 90th percentile of the distribution, which is the threshold value used to define the tail in the bottom panel. The blue line in the bottom panel indicates the generalized Pareto distribution (GPD) fit to the tail of the corresponding distribution. Note the use of log scale in the *y*-axis of the bottom row panels. Data for *Juniperus monosperma* and *Pinus edulis* are pooled across experimental treatments.

sample size ($N \geq 50$ midday measurements, spanning one to 11 sites depending on the species) (Table S1).

A second empirical dataset including raw, individual-level measurements of midday water potentials was collected specifically for this study, using our own measurements and contributed or publicly available databases. This dataset contains a total of 7749 individual measurements of midday leaf water potential on 76 species, and we identify it as the 'raw' dataset. We selected data for the 26 species for which the number of measurements was ≥ 30 and there were at least 10 combinations of site and sampling date (e.g. two sites with five or more time points per site or one site with 10 or more time points) (Table S2). Note that an explicit separation of temporal and spatial effects would have been desirable in all analyses, but was not possible owing to data limitations.

Absolute Ψ_{min} values from the two previous databases ('MV14' and 'raw') were combined with an updated version of the global xylem functional traits (XFT) dataset by Choat *et al.* (2012) incorporating newer references covering primarily China, the Amazon and Australia (Sanchez-Martinez *et al.*, 2020). After aggregating to the species level this database contained Ψ_{min} data for 654 species ('species-level' dataset). Unlike the measurements in the 'MV14' and 'raw' datasets, which were taken on uncovered leaves (or twigs), the Ψ_{min} data in the XFT dataset combine measurements taken on covered and uncovered leaves. However, the effect of the pressure drop in leaves associated with transpiration is minimal in our case because Ψ_{min} is usually recorded under extreme conditions when stomata are closed (cf. Choat *et al.*, 2012; Martin-StPaul *et al.*, 2017).

Synthetic datasets

We generated two synthetic water potential datasets with realistic statistical properties to conduct a detailed analysis of the impact

of sample size and different analytical approaches on species-level estimates of Ψ_{min}. One of these datasets was generated from a Gamma distribution with scale and shape parameters equal to those obtained for the fitted *Juniperus monosperma* distribution, and simulated a right-skewed distribution typical of relatively anisohydric species with looser water potential regulation (Fig. S1). The second dataset was obtained using a Weibull distribution with parameters equal to those corresponding to *Pinus pinaster*, and represents a relatively isohydric species with tighter water potential regulation (Fig. S1).

Each of the synthetic datasets corresponded to a combination of data for two to 100 sites, each of which had been measured 10 times and with 10 plants sampled per time point, and hence provided an 'ideal' species-level coverage relative to what is possible in practice in field studies. Individual data points within a sampling event and seasonal measurements within a year were sampled from the same distribution (constant scale and shape parameters). We also conducted additional simulations in which variability was allowed among sites to account for climatic as well as ecotypic variability. This variability was introduced by sampling the corresponding distribution parameters from multivariate normal distributions with the same means as discussed earlier and a realistic covariance matrix, based on the distribution fits for species with data from multiple populations/years.

Data analysis

Water potential distributions and their tails (species level) We used the 'raw' dataset (including completely disaggregated, individual-level measurements) to study the tails of Ψ distributions (data above the 90th percentile). We compared the fit provided by three different probability functions with different tail shapes, corresponding to the log-normal (heavy-tailed), Gamma and Weibull (heavy-tailed depending on the shape parameter) distributions, with the fit provided by the GPD (using the probability weighted moment method). We also compared GPD fits using alternative thresholds to characterize the distribution tail, corresponding to 75th and 95th percentiles. In all analyses, water potential values were converted to positive (absolute) values before fitting the different probability distributions. The goodness-of-fit of different distributions was compared qualitatively using the Akaike information criterion (AIC).

Variance components and variability in Ψ and Ψ_{min} within species We used variance components analysis to estimate the contribution of species, site, year and date (day of the year) to total Ψ variability in the 'raw' dataset. We conducted additional analyses focusing on the species with largest sample size (*J. monosperma* and *Pinus edulis*) to further explore the components of temporal variability. In those models date was nested within season and season within year. All variance components analyses were conducted using the R package VCA (v.1.4.3).

To study the variability within species, we subset the data from the 'raw' dataset by combinations of site and year and selected only the combinations with sufficient data ($N \geq 30$ individual measurements). For each of the six species that had several such combinations we compared the full Ψ distributions among combinations of site and year using a *k* sample Anderson–Darling test (R package KSAMPLES v.1.2.9), which gives more weight to the tails of the distributions than do alternative methods (Scholz & Stephens, 1987). In addition, we also compared the Ψ_{min} values estimated for different combinations of site and year (within species) using the following methods: the lowest midday leaf water potential (highest absolute value) recorded for a given species (cf. Choat *et al.*, 2012); the 99th percentile of the sampled values for the species; and the 99th percentiles of the Gamma, Weibull, and GPD distributions (parameterized as described earlier) fitted to the water potential data.

Effect of sample size and estimation method on Ψ_{min} determination (species level) To assess the impact of sample size and estimation method on the determination of species-level Ψ_{min} we used both the 'MV14' and 'raw' datasets. As earlier, Ψ_{min} was calculated as: the lowest midday leaf water potential, the 99th percentile of the sample, or the 99th percentile of the Gamma, Weibull and GPD distributions fitted to the species data. In all cases we obtained Ψ_{min} estimates from the whole dataset for each species and from random subsamples, including 80%, 60%, 40%, 20% and 10% of the data, as well as for $N = 10$ (in that latter case to compare a common, low sample size across species). Computations were repeated 100 times per species using different random subsamples.

To account for all sources of variability in water potential observations in an idealized situation, we conducted a similar analysis using the synthetic datasets. In this case the underlying probability distribution was known and hence the methods to estimate Ψ_{min} that we compared were: the lowest midday leaf water potential observed for a given 'species'; the average Ψ_{min} across populations (cf. Bartlett *et al.*, 2016); the 99.9th percentile of the sampled values; the 99.9th percentile of the underlying distribution used to generate the data; the 90th percentile of the GEV distribution fitted to the Ψ_{min} data, blocking by population; and the 99.9th percentile of the GPD fitted to the water potential data. Here we used more extreme percentiles than with the empirical datasets (99.9th instead of 99th as a reference) because sample size was much larger and included all sources of variability (across individuals, time and populations). Similarly, the 90th percentile was used when fitting the GEV distribution (instead of the 99.9th percentile) because in that case we are fitting (blocked) Ψ_{min} data as opposed to raw Ψ data. Because we are blocking by population, each of which corresponds to 100 Ψ values, the 90th percentile of the resulting Ψ_{min} should be comparable to the 99.9th percentile of the underlying raw data.

Estimating hydraulic risk We used the disaggregated data of the 'raw' dataset (full sample size) to calculate the species-level HSM using different estimates of Ψ_{min}, including the lowest Ψ_{min} value, the average Ψ_{min} across populations and years, the 99th percentiles of the Weibull, Gamma and GPD distributions, and the 90th percentile of the GEV distribution fitted to the Ψ_{min} data, blocking by population and year (the latter only for the only two species for which the number of blocks was > 10: *Pinus*

halepensis and *Quercus ilex*). These estimates were computed for the 17 species for which stem Ψ_{PLC50} data were present in the 'species-level' database. We also used the fitted GPD and GEV distributions to calculate the probability of Ψ_{min} reaching a value equal to Ψ_{PLC50}. To get a Ψ_{PLC50} value representative of the maximum resistance to embolism at the species level, and also to avoid methodological issues that may overestimate vulnerability to embolism (Cochard *et al.*, 2013; see also Venturas *et al.*, 2019), in these analyses we estimated Ψ_{PLC50} as the minimum value recorded for a given species. This minimum Ψ_{PLC50} was tightly related to the species mean Ψ_{PLC50} ($r = 0.93$, slope = 0.85 ± 0.09) and to more extreme measures of hydraulic loss such as Ψ_{PLC88} ($r = 0.91$, slope = 1.10 ± 0.15), and our results remained qualitatively identical if these other metrics were used (data not shown).

Assessing implications for global patterns of hydraulic safety margins Finally, we assessed whether methodological issues in Ψ_{min} estimation could be affecting the global relationship between Ψ_{min} and Ψ_{PLC50} as reported by Choat *et al.* (2012) and similar studies. To that end, we first assessed the relationship between Ψ_{PLC50} and sample size (for Ψ_{min}) for both the species in our 'raw' dataset and globally using the updated XFT database (with individual site-specific values retained instead of being aggregated by species as in our 'species-level' dataset). A positive relationship between $|\Psi_{PLC50}|$ and sample size would indicate that more resistant species are more intensively sampled, and hence a potential, directional bias in the estimation of Ψ_{min}. We then compared the relationship between Ψ_{min} and Ψ_{PLC50} for the species in our 'raw' dataset using two different estimates of Ψ_{min}: the absolute minimum (as in Choat *et al.* (2012)) and the GPD-based estimate described earlier, which is expected to be more robust to sample size effects. In both models $|\Psi_{min}|$ and $|\Psi_{PLC50}|$ were log-transformed to normalize the distribution of the residuals. In this case Ψ_{PLC50} was estimated as the mean Ψ_{PLC50} for a given species, for consistency with the original analysis by Choat *et al.* (2012), although other measures of hydraulic vulnerability were also tested (minimum Ψ_{PLC50}, Ψ_{PLC88}). Note that we could not perform a similar analysis using the whole 'species-level' dataset because the real sample size (number of individuals, sampling dates and sites used to estimate Ψ_{min}) is unknown in this case.

All analyses were conducted in the R environment (R v.3.6.2; The R Foundation for Statistical Computing). Distributions were fitted using the packages FITDISTRPLUS (v.1.1.1), EVD (GEV, v.2.3.3) and FEXTREMES (GPD, v.3042.82).

Results

The statistical distribution of water potential data

The distribution of leaf water potentials for the species included in the 'MV14' and 'raw' datasets are shown in Figs S2 and S3. As expected, the shapes of the distributions differed across species and were not always well captured by standard distributions (log-normal, Gamma, Weibull). In particular, some species showed bimodal distributions. This pattern was observed in some species for which several populations were included (e.g. *Quercus coccifera*), but also in species that had been measured over one single site and year (e.g. *Encelia actoni*, *Ericameria linearifolia* and *Purshia tridentata*; Fig. S3). As expected, when we focused on the tails of the distributions ('raw' dataset) the best fit was always provided by the GPD, which had AIC values that were on average 13, 14 and 24 units lower than those corresponding to the log-normal, Gamma and Weibull distributions, respectively (Table S3). Fig. 1 shows GPD fits to distribution tails for four representative species. The average value of the shape parameter of the GPD distribution was −0.38 and it was < 0 in 19 out of 26 species (Table S3), indicating that the distribution of values above the Ψ threshold was bounded in most cases (see Embrechts *et al.*, 2013). However, this shape parameter (unlike the corresponding, estimated Ψ_{min}) was relatively unstable and depended on the choice of threshold (Table S3), probably reflecting sample size limitations.

Variance components and within-species variability

The variance components analysis indicated that most of the variability in Ψ corresponded to differences among dates (57%) and sites (19%), whereas species accounted for 10% of the total variability (Table S4). A similar pattern emerged within species, where > 50% of the variability was observed among dates (within seasons and years) for the two species tested (Table S4). For five out of six species analysed, the distribution of Ψ differed among years/sites (*k* sample Anderson–Darling test, $P < 0.001$), the only exception being *Phillyrea latifolia*, for which we could only compare two consecutive years at the same site (Fig. 2). For three species (*P. halepensis*, *P. sylvestris* and *Q. ilex*) we could also group the observations by years within a site and compare the distributions by sites. The resulting Ψ distributions also differed significantly among sites ($P < 0.001$). The fact that Ψ distributions differed within species also resulted in substantially different Ψ_{min} estimates among sites and years, regardless of the estimation method that was used (Table S5). For estimates based on the GPD distribution, which again provided the best fit to the Ψ tail data (Table S5), Ψ_{min} values varied by as much as 5.7 MPa between different combinations of site by year in *P. halepensis*, and up to 4.6 MPa when comparing the same population (Orihuela) between two consecutive years.

Comparing Ψ_{min} estimation methods

Species-level Ψ_{min} differed substantially across estimation methods, both for the 'raw' (cf. values corresponding to maximum sample size in Fig. 3) and for the 'MV14' datasets (Fig. S4). These values were typically more extreme for either the 99th percentile of the Gamma distribution fitted to the data or for the lowest measured Ψ_{min}, with the latter being generally more extreme for species with a large sample size ($N > 100$). The Ψ_{min} estimates based on the 99th percentile of the GPD were usually similar to those obtained from the sample 99th percentile, but the

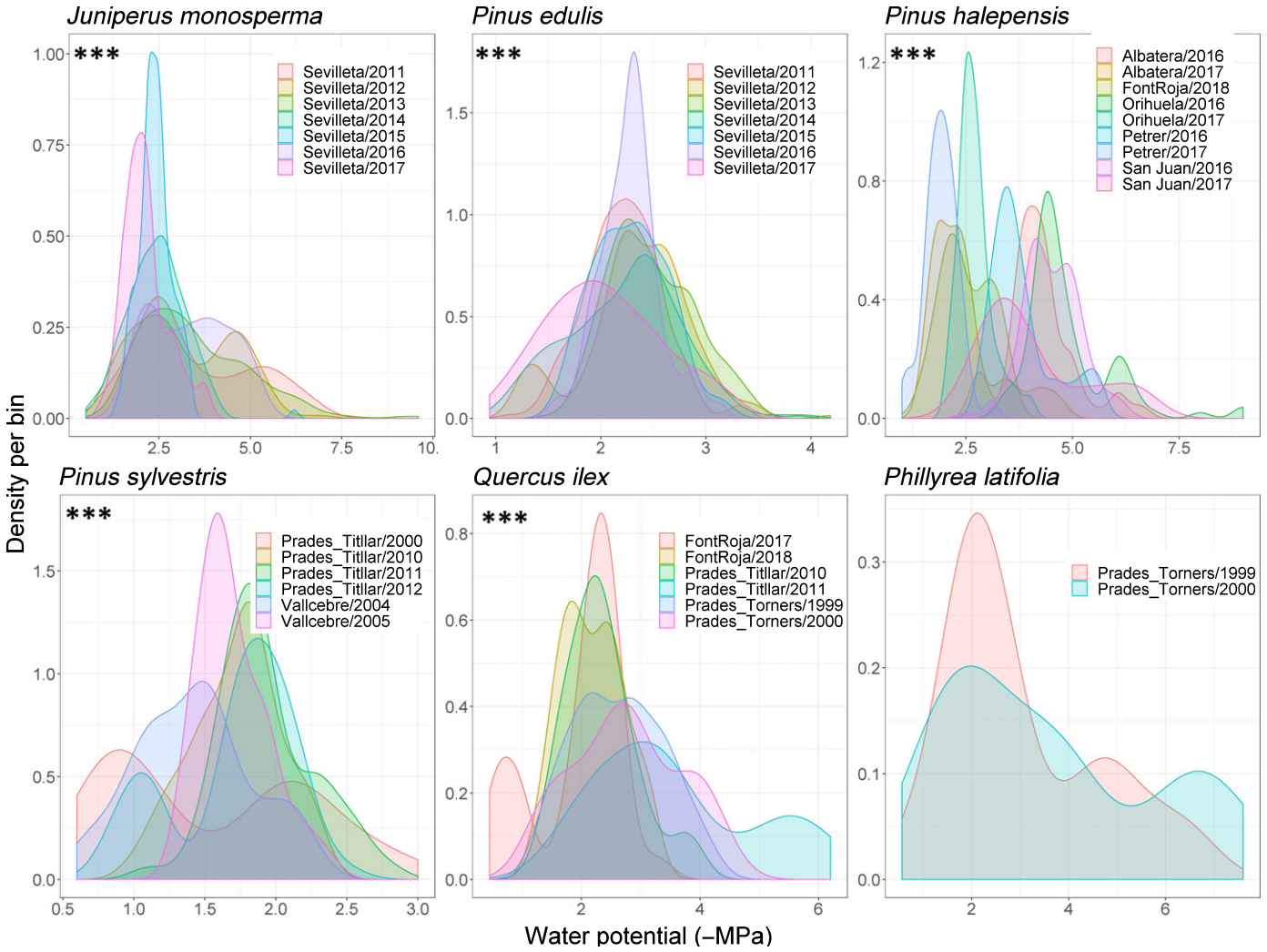

Fig. 2 Density distributions for different combinations of site and year for six representative species from the 'raw' dataset. Asterisks at the top left corner of a given panel indicate significant differences among sites/years for the corresponding species (***, $P < 0.001$). Data for *Juniperus monosperma* and *Pinus edulis* are pooled across experimental treatments.

former were less sensitive to sample size. The GPD-based estimates were very robust to the criteria used to set the threshold, with resulting Ψ_{min} values differing by 0.05 and −0.03 MPa on average between our base calculations (90th percentile) and those using the 75th or 95th percentiles, respectively (Table S3). In addition, GPD-based estimates were also reasonably robust to removing different proportions of the data with less extreme Ψ values (i.e. using only the data measured under relatively extreme conditions) (Table S6).

Effect of sample size on Ψ_{min} estimates

Sample size had a strong effect on Ψ_{min} when it was estimated as the lowest midday leaf water potential recorded for a given species, regardless of whether data came from the 'raw' or the 'MV14' dataset (Figs 3 and S4, respectively). In most species the effect tended to saturate at high sample sizes, but in some cases a plateau was not observed even for sample sizes > 100 (e.g. *Quercus coccifera* in Fig. 3; *Q. ilex* in Fig. S4). The bias associated with sample size was particularly conspicuous for heavily sampled species, for which estimates could differ by several MPa (e.g. between *c.* −6 and −10 MPa for *J. monosperma*; Fig. 3). As expected, Ψ_{min} estimates based on the 99th percentile of the sampled distribution were basically identical to the absolute minimums for sample sizes below 100, and shared the strong dependency on sample size. All the other Ψ_{min} estimates, based on percentiles of the fitted Gamma, Weibull or GPD distributions, were much less sensitive to sample size. However, they were all affected by sample size at low values ($N < 20$), at least for some species (Figs 3, S4).

Sample size also had a strong effect on Ψ_{min} estimated as the lowest midday leaf water potential recorded in the synthetic datasets, regardless of the distribution used to generate the data (Fig. 4). At high sample sizes, these estimates were always substantially more extreme than those obtained by the other methods, including the 99.9th percentile from the known

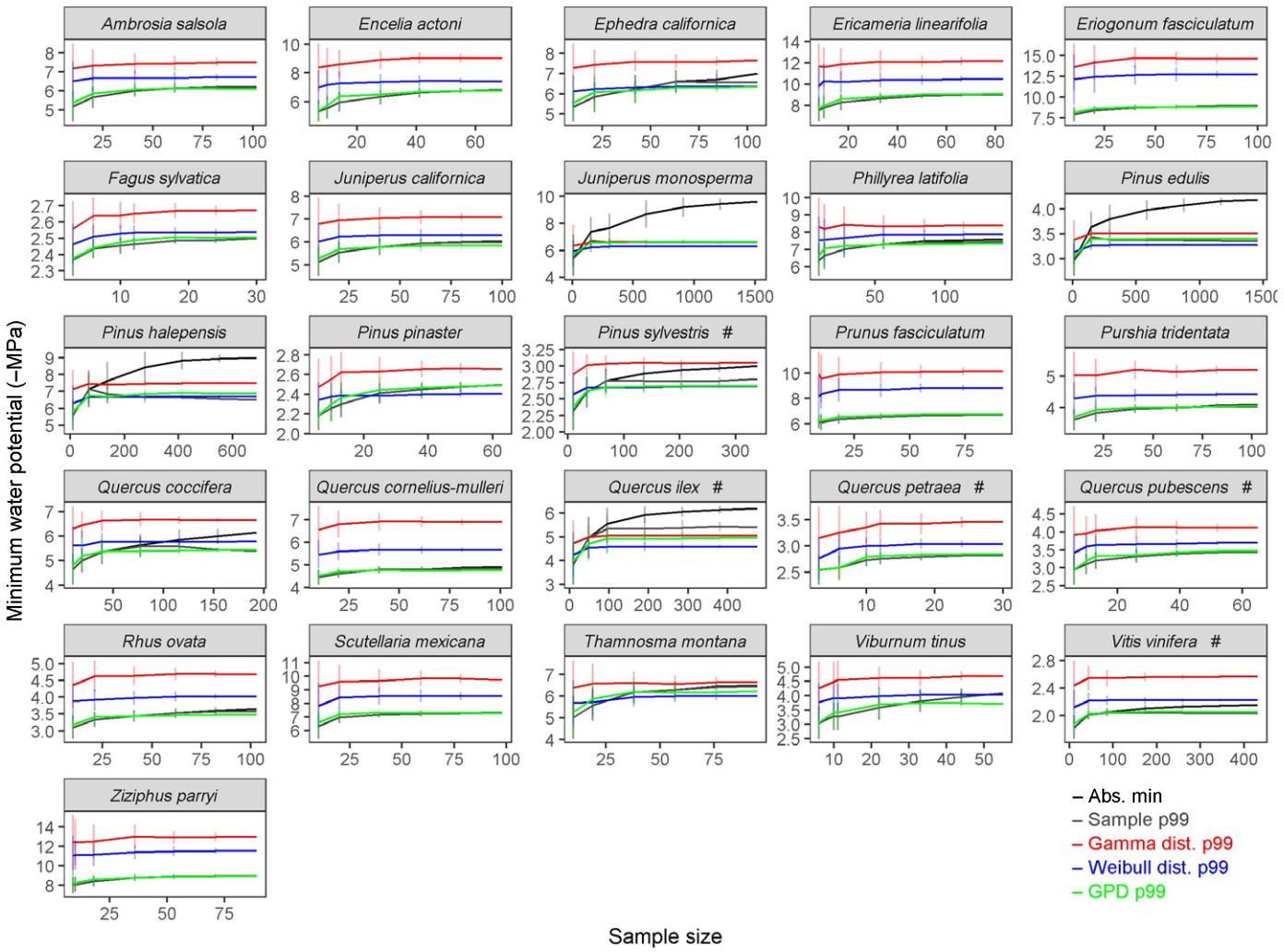

Fig. 3 Effect of sample size on minimum water potential estimates (absolute values in MPa) from the 'raw' dataset. For each species, different estimation methods are compared and shown in different colours (black lines: the absolute lowest value from the sample; grey lines: the 99th percentile from the sample, which is indistinguishable from absolute lowest values for sample sizes < 100; red lines: the 99th percentiles of the Gamma distribution fitted to the data; blue lines: the 99th percentiles of the Weibull distribution fitted to the data; green lines: the 99th percentile of the generalized Pareto distribution (GPD) fitted to the data). Error bars show SDs of 100 random repetitions for each species and sample size.

distribution. At the other extreme, when 'species' Ψ_{min} was estimated as the average Ψ_{min} across populations, the resulting values were very stable but always substantially less extreme than the 99.9th percentile from the known distribution (Fig. 4). Estimates obtained from the sample 99.9th percentile and those using the GEV and GPD distributions all converged to values indistinguishable from the 99.9th percentile of the original distribution (Fig. 4). These three methods were quite robust to low sample sizes down to $N = 10$ populations, although the estimate from the sample 99.9th percentile typically showed the highest standard deviations. For even lower sample sizes, GPD-based estimates appeared the most consistent, although uncertainty increased substantially in all cases, particularly for GEV-based estimates. These results remained qualitatively similar when variability in the parameters of the underlying distributions was introduced among populations (Fig. S5).

Estimating hydraulic risk

The calculated HSMs for the 17 species in the 'raw' dataset for which stem Ψ_{PLC50} was available were highly variable depending on the method used to estimate Ψ_{min}, with discrepancies of more than 1 MPa in many species and up to > 4 MPa in *P. halepensis* (Fig. 5). For all the species in which there was more than one combination of year and site, the HSM estimates using the absolute lowest Ψ_{min} were substantially lower than those using the mean Ψ_{min} across sites/years. The HSM estimates obtained using the Weibull and particularly the Gamma distributions tended to be substantially lower than those using the GPD distribution, whereas those based on the GPD and GEV distributions were similar for the two species for which the latter could be computed (Fig. 5).

The probability of individual water potential measurements reaching Ψ_{PLC50} calculated from the GPD (90th percentile

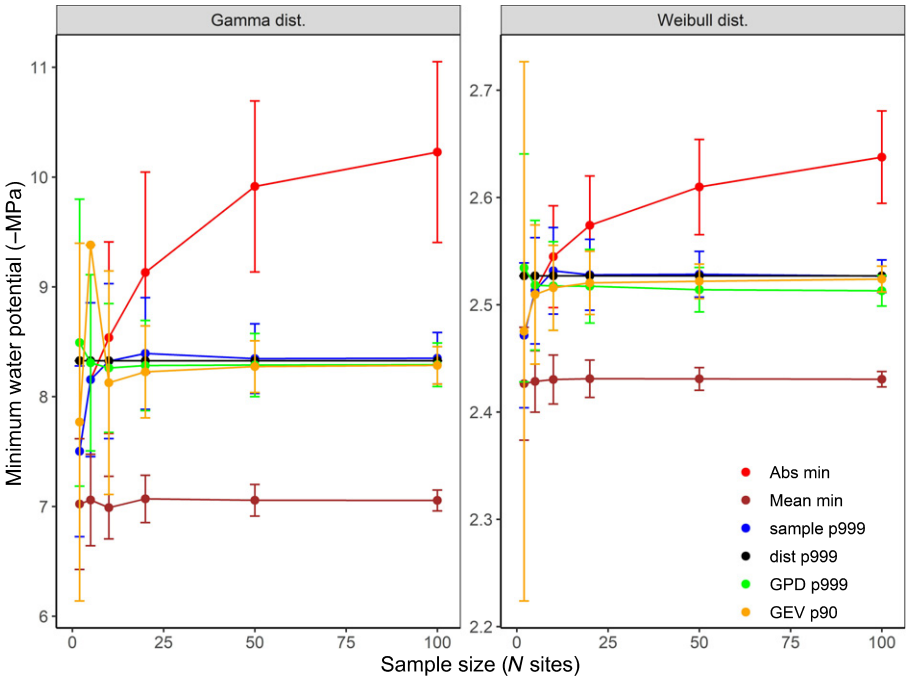

Fig. 4 Effect of sample size (number of sites, from two to 100) on minimum water potential estimates (absolute values in MPa) from the synthetic datasets. Each site comprised data for 10 measurement dates with 10 plants sampled per time point. Results are shown for two different distributions: a Gamma (left panel) and a Weibull distribution (right panel) (see text for details on parameterization and Supporting Information Fig. S1 for the corresponding histograms). Different estimation methods are compared and shown in different colours (red: the absolute lowest value; brown: the mean of the minimum values observed at different sites; blue: the 99.9th percentile from the whole sample; black: the 99.9th percentile of the known distribution; green: the 99.9th percentile of the generalized Pareto distribution (GPD) fitted to the data; orange: the 90th percentile of the generalized extreme value (GEV) distribution fitted to the data blocked by site). Error bars show SDs of 100 random repetitions for each species and sample size. Note that some error bars for the GEV estimates under low sample sizes were very large and have been omitted.

threshold) ranged between 0% for four species to > 10% for *Ambrosia salsola*, *Rhus ovata*, *Viburnum tinus* and *Vitis vinifera* (Table S7). The latter values imply that Ψ_{PLC50} was higher (less negative) than the threshold used to fit the GPD, and hence the probability of reaching Ψ_{PLC50} could not be estimated with precision. The consistency of these probability estimates when comparing different GPD thresholds (75th and 95th percentiles) was quite variable across species (4.9–150% deviation depending on the species and threshold comparison; Table S7), and was clearly lower than for the Ψ_{min} estimates themselves (Table S3). For the two species for which GEV distributions could be fitted, the probability of Ψ_{min} reaching Ψ_{PLC50} (or lower) differed substantially (3% for *Q. ilex*, 19% for *P. halepensis*). This difference was qualitatively consistent with that obtained using GPD-based estimates (Table S7) and reflected differences in the corresponding HSM (Fig. 5).

We note that our focus here is on the effect on HSM of different approaches to estimate Ψ_{min}, not on the HSM values themselves, which should be interpreted with caution. In particular, HSM estimates for some species may be problematic owing to uncertainty in Ψ_{PLC50} values. For instance, the highly negative HSM in *Viburnum tinus*, and particularly *Vitis vinifera*, probably arise from the overestimation of embolism vulnerability for these species (Charrier *et al.*, 2018), for which only one Ψ_{PLC50} value was available in the XFT dataset. The fact that we are comparing leaf/twig Ψ_{min} with stem Ψ_{50} may introduce additional bias, even if water potential gradients in the xylem are expected to be low under extreme drought (Choat *et al.*, 2012).

Implications for global patterns of hydraulic safety margins

We found a significant, negative correlation between species-level Ψ_{PLC50} and the number of entries in the 'species-level' dataset ($r = -0.22$, $P < 0.001$, $N = 226$), suggesting that embolism-resistant species tended to be more intensively sampled. A similar relationship was found for the species in our 'raw' dataset, which was statistically significant when minimum Ψ_{PLC50} was considered ($r = -0.55$, $P = 0.02$, $N = 17$) but not when using mean Ψ_{PLC50} ($r = -0.35$, $P = 0.17$, $N = 17$). Consistent with the original analysis in Choat *et al.* (2012) we found a strong positive relationship between Ψ_{min} and mean Ψ_{PLC50}, regardless of whether Ψ_{min} was estimated from absolute minimum values or GPD-based estimates ($r = 0.65$, $P < 0.01$, $N = 17$ in both cases) (Fig. 6). The slope of the relationship was significantly < 1 in the two cases, but declined from 0.44 ± 0.13 (intercept = 0.94 ± 0.19) when Ψ_{min} was estimated from absolute minimum values to 0.38 ± 0.12 (intercept = 0.92 ± 0.16) when GPD-based estimates were used (Fig. 6). These results indicate that more resistant species had wider safety margins, and that this trend was exacerbated when more robust Ψ_{min} estimates that accounted for sample size effects were used. We fitted a single relationship for angiosperms and gymnosperms because, in our dataset, HSM

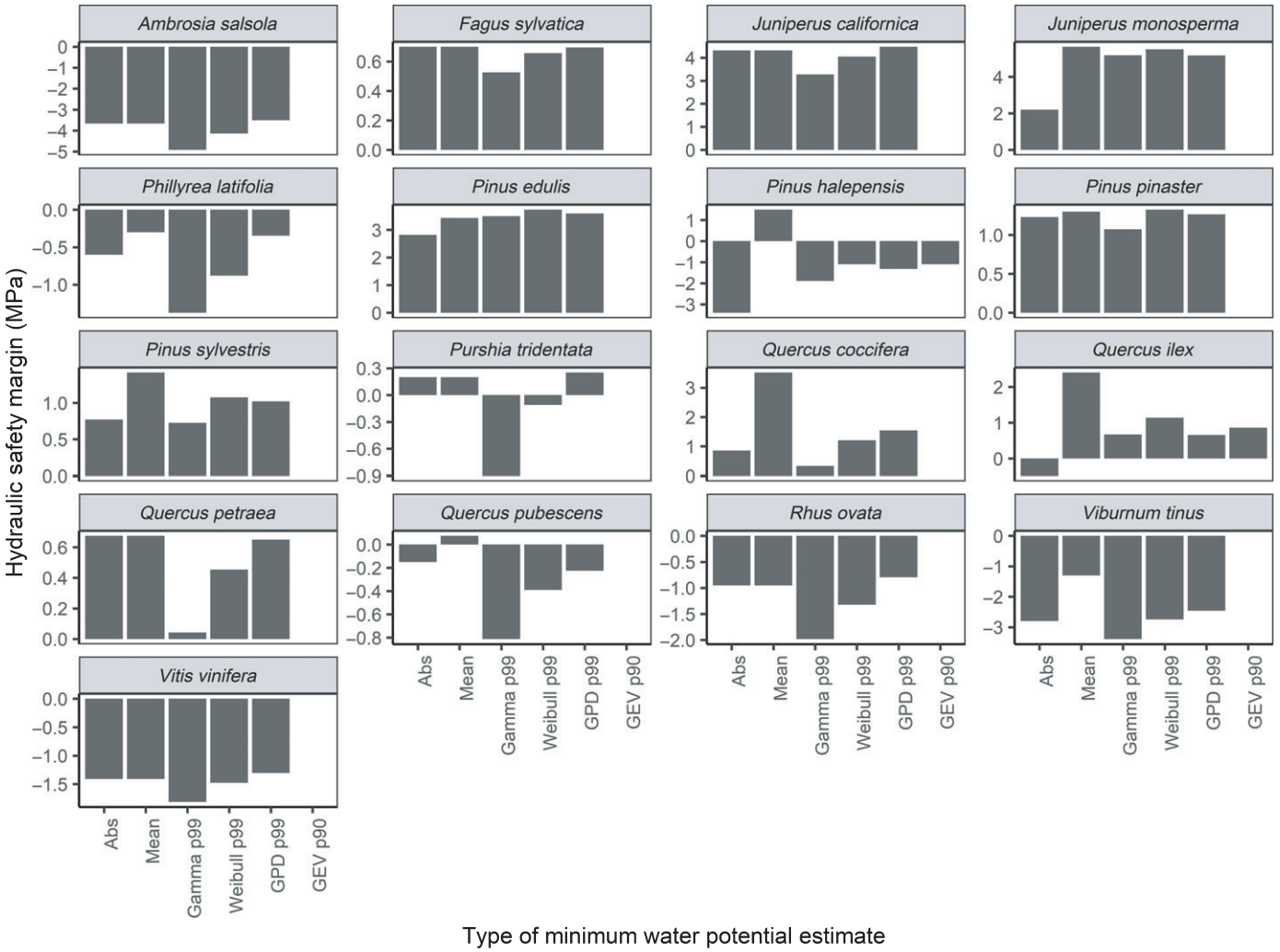

Fig. 5 Comparison of hydraulic safety margins (HSM, MPa) obtained using different estimates of minimum water potential (Ψ_{min}) on the species from the 'raw' dataset. Different estimation methods include the absolute lowest value from the sample ('Abs'), the mean across all the sites/years sampled ('Mean'), the 99th percentile from the fitted Gamma distribution ('Gamma p99'), the 99th percentile from the fitted Weibull distribution ('Weibull p99'), the 99th percentile of the fitted generalized Pareto distribution ('GPD p99'), and the 90th percentile of the generalized extreme value distribution fitted to site/year blocks, for the two species for which the number of combinations of site and year was > 10 ('GEV p90'). HSM is quantified as the difference between Ψ_{min} and stem Ψ_{PLC50}, and hence a lower (e.g. negative) value indicates higher hydraulic risk. Please note that HSM values reported in this figure are primarily intended for comparison of alternative Ψ_{min} estimation methods and should be interpreted with caution (see text).

were similar between these two group of species ($P = 0.23$, $N = 11$ angiosperm and $N = 6$ gymnosperm species).

Discussion

Current estimates of Ψ_{min} are biased

Minimum water potential (Ψ_{min}) is a key variable to characterize dehydration tolerance and hydraulic risk in plants (Choat *et al.*, 2012, 2018; Delzon & Cochard, 2014). Our results show that current Ψ_{min} values, estimated as the absolute minimum leaf water potential reported for a given species, are biased. This is because absolute extremes are a function of sample size and the underlying statistical distribution, which differ markedly among species (and among sites and years within species; Fig. 2). This bias can be large, of the order of several MPa (Figs 3, 5), and its magnitude is generally unknown. Moreover, this bias is not randomly distributed across species. Because species living in drier environments tend to be more intensively sampled, current Ψ_{min} estimates for these species would be necessarily more extreme (even if the underlying probability distribution was identical) than those for species living under milder conditions.

The fact that Ψ_{min} estimates are biased has significant implications for the determination of HSMs. Because HSM is usually calculated as the difference between Ψ_{min} and a measure of hydraulic resistance (e.g. Ψ_{PLC50}), estimated HSM would have the same absolute bias as the corresponding Ψ_{min} values (cf. Fig. 5). We also show that this bias can have important consequences for our understanding of the distribution of drought resistance and hydraulic risks at the global scale. Choat *et al.* (2012) showed that, for angiosperms, HSM were relatively invariant among biomes and only increased slightly for more

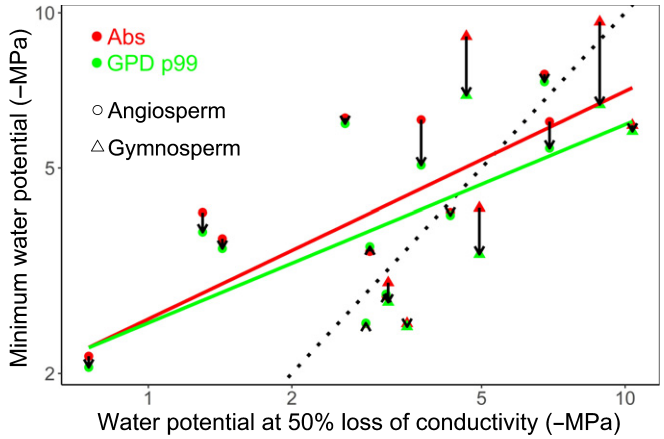

Fig. 6 Relationship between water potential at 50% loss of hydraulic conductivity (Ψ_{PLC50}) and minimum water potential (Ψ_{min}) across the species in the 'raw' dataset. Red dots correspond to the absolute lowest Ψ_{min} recorded for a given species, whereas green dots show generalized Pareto distribution (GPD)-based estimates of Ψ_{min}, which are robust to sample size effects. Observations for a given species are connected by a black arrow. Different symbols are used for angiosperm and gymnosperm species. Red and green lines show the corresponding linear fits. The 1 : 1 line is indicated by the dotted black line. Note the log scales in both axes.

embolism-resistant species. Our analyses show that this interpretation may need to be reassessed, as HSM estimates became wider when more robust, GPD-based estimates of Ψ_{min} were used, relative to using absolute lowest Ψ values, and this happened particularly in more embolism-resistant species as a result of their higher sampling effort (Fig. 6). Although it is impossible to quantify this effect precisely at present, because the sample size used to estimate Ψ_{min} in global hydraulic databases is unknown, these results suggest that safety margins may actually increase substantially with embolism resistance at the global scale. This is consistent with the higher difference between the water potential at stomatal closure and Ψ_{PLC50} (another measure of the safety margin) in embolism-resistant species (Martínez-Vilalta *et al.*, 2014; Mencuccini *et al.*, 2015; Martin-StPaul *et al.*, 2017). We note, however, that these patterns are likely to be scale-dependent, as increased hydraulic risk in embolism-resistant species has been reported in local communities (e.g. Venturas *et al.*, 2016).

The fact that the distribution of leaf water potentials differs greatly among species has been largely overlooked in the literature. Besides having substantial implications for the estimation of Ψ_{min}, this variation reflects important biological phenomena. In particular, the distribution of leaf water potentials is the end result of all the control mechanisms by which plants regulate their water status as a response to environmental drivers, and hence provides a synthetic, statistical description of water potential regulation (Martínez-Vilalta & Garcia-Forner, 2017). Under the same environmental conditions, relatively anisohydric species would have Ψ distributions that tend to be skewed to the right and show fatter tails, whereas more isohydric species would be characterized by more left-skewed distributions. Examples of this would be *J. monosperma* and *P. edulis* (or even more clearly *P. pinaster*; Fig. S3), which are paradigmatic examples of anisohydric and isohydric regulation of water potentials, respectively (Breshears *et al.*, 2008; McDowell *et al.*, 2008; Meinzer *et al.*, 2014). The Ψ distribution, however, is also determined by environmental conditions (Hochberg *et al.*, 2018), which probably explains part of the within-species variability we observed (Fig. 2). This effect could be accounted for by modelling leaf from soil Ψ (e.g. Martínez-Vilalta *et al.*, 2014) or by directly focusing on the difference between the two. This type of analysis should be facilitated by increased availability of environmental drivers and detailed soil and plant water content measures from remote sensing (cf. Wu *et al.*, 2021).

Guidelines for the estimation of Ψ_{min} and other ecophysiological extremes

Any proper estimation of extreme values from field measurements has to deal with the fact that these values are rare by definition, and thus their estimation is less straightforward than for typical measures of centrality (e.g. the mean) or dispersion (e.g. the standard deviation). Although the absolute minimum of any given sample is easy to calculate, this magnitude does not provide a sound basis to estimate Ψ_{min}, as it continuously declines (becomes more extreme) as sample size increases, and hence cannot be used as a reference to compare species (or populations) in practice. An operational definition of Ψ_{min} is therefore needed, and such definition should be probabilistic to make it independent of sampling effort. Although directly estimating (and comparing) a given sample percentile could be a reasonable solution in many cases, these estimates are more sensitive to low sample sizes (Fig. 4), and would be particularly problematic when the underlying probability distribution is heavy-tailed. We thus advocate here for the use of EVT (Coles, 2001; Reiss & Thomas, 2007) as a good basis for improved estimation.

Extreme value theory focuses on the tail of the probability distribution, as the tail conveys the most relevant information to estimate the magnitude and likelihood of extreme events. In general, the most efficient approach uses the 'peak over threshold' (POT) method (Coles, 2001). This approach involves identifying a threshold in the raw distribution of Ψ that characterizes extreme situations, fitting a GPD to values exceeding that threshold (the tail of the distribution), and then estimating the species' Ψ_{min} as a given (conventional) percentile of the fitted GPD. Because this method focuses on the most extreme values of the overall distribution, it has the advantage of being relatively insensitive to variability in Ψ among populations or years (cf. Fig. 2), which is particularly useful when the objective is to obtain species-level estimates of Ψ_{min}. We suggest that a 99.9^{th} percentile would be a reasonable choice, as it corresponds to the minimum value recorded, on average, every 1000 individual water potential measurements. Based on our analyses, Ψ_{min} estimates obtained this way are consistent when the number of individual Ψ measurements exceeds 20 (Fig. 3), and are robust to different threshold choices (see Coles (2001) for additional criteria to set the GPD threshold).

Using the GPD approach, however, requires fitting raw Ψ data, which are not always available. In those cases, and

particularly when midday Ψ measurements are available for several sites and years with a reasonably balanced experimental design, an alternative way to estimate Ψ_{min} involves grouping the data into blocks corresponding to combinations of site and year. Then, the absolute minimum Ψ is obtained for each block (analogous to data binning procedures), the GEV distribution is fitted to these block extremes (cf. Makkonen & Tikanmäki (2019) for a comparison of fitting methods), and species' Ψ_{min} is estimated as a given (conventional) percentile of the fitted distribution. In this case, the 90th percentile, corresponding to the minimum water potential that would be observed, on average, once every 10 yr (or 10 combinations of site and year), may be adequate in most cases. That implies, of course, that robust estimates of Ψ_{min} can only be obtained if data are available for at least 10 combinations of site and year (cf. Fig. 4), which unfortunately is quite rare at present (e.g. only two out of the 76 species recorded in our 'raw' dataset fulfilled this requirement). The GEV-based estimates may be particularly useful to capture biologically relevant differences among sites or years. It is noteworthy that, in our case, the resulting Ψ_{min} estimates were similar for GEV- and GPD-based methods for the two species in which both could be applied (Fig. 5).

Sampling issues and representativeness remain the main difficulty in obtaining robust estimates of Ψ_{min}. Water potential measurements tend to be heavily clustered around specific sites or years, and hence may not capture the relevant conditions (severe drought) and spatiotemporal variability to accurately determine extreme values. Our variance components analyses show that most of the Ψ variability occurs over time, and particularly at the daily scale (Table S4), and hence temporal replication is critical to characterize Ψ distributions. At the same time, it is clear that ecotypical variation may also be high (e.g. Atzmon *et al.*, 2004; see also our Fig. 2) and hence a proper characterization of Ψ_{min} at the species level requires some degree of spatial replication. Because water potential measurements are largely conducted to study vegetation responses to drought, sampling schemes tend to prioritize dry locations and dry periods, effectively biasing water potential distributions towards low values. Although, by focusing on the tail of the Ψ distribution, EVT-based methods are relatively insensitive to the impact of undersampling other parts of the distribution, they are not immune to bias, and extensions may be required to account for relevant sources of variation and to obtain accurate estimates. Psychrometers offer a powerful alternative to obtain quasi-continuous Ψ series and increase data representativeness, at least in the temporal dimension (e.g. Guo *et al.*, 2020), and have the additional advantage of measuring stem (as opposed to leaf) water potential, which is more directly comparable with most assessments of hydraulic vulnerability. However, psychrometer measurements are still rare and cannot be obtained for all taxa. In summary, even though sampling issues are minimized by the EVT approach we propose, they are certainly not solved.

Interestingly, for the seven species in the 'raw' dataset for which we could obtain Ψ_{lethal} values from the literature (Table S8), those compared well with our Ψ_{min} estimates ($r = 0.76$, average difference < 0.9 MPa). For these species, $\Psi_{lethal} < \Psi_{min}$ except for *J. monosperma*. The largest difference between Ψ_{lethal} and Ψ_{min} was observed for *P. edulis* (2.12 MPa), for which the absolute lowest Ψ after *c.* 1500 measurements (including an extreme drought treatment) was −4.19 MPa, still a long way from the estimated Ψ_{lethal} (−5.53 MPa on average). This discrepancy illustrates the fact that at least some species may effectively never reach, in the field, the Ψ_{lethal} values determined under controlled experimental conditions. Tagging and monitoring branches sampled under extreme drought to ensure they were alive and survived is a promising avenue to reconcile Ψ_{lethal} and Ψ_{min} estimates.

Towards a more robust determination of hydraulic risk

Characterizing the statistical distribution of water potentials and their minimum values opens new ground for the determination of plant hydraulic risk, as it allows the probability of reaching a certain, critical Ψ to be estimated. If enough data are available, deriving return intervals analogous to those used in hydrology and other applied sciences (i.e. the average time between events of a given magnitude) becomes straightforward (e.g. Reiss & Thomas, 2007). Because hydraulic vulnerability thresholds are relatively well understood, at least for the xylem, estimating the probability of a given amount of hydraulic damage (e.g. 50% or 88% loss of conductivity), or the corresponding return interval, also becomes possible. Conceptually, this is clearly superior to HSMs estimated as the difference between Ψ_{min} and embolism resistance, as currently done. Unfortunately, our results show that the quality of species-level, water potential data is rarely good enough at present to obtain robust probabilistic estimates of hydraulic risk. Therefore, even without considering the issues associated with the determination of hydraulic vulnerability, which are beyond the scope of this paper, a proper characterization of hydraulic risk remains a challenge (Martinez-Vilalta *et al.*, 2019).

In conclusion, we demonstrate that current Ψ_{min} estimates are biased and propose two EVT-based methods to obtain statistically robust estimates of Ψ_{min}. We validate these methods, which may also be applied to other ecophysiological extremes, using different empirical and synthetic datasets at different measurement levels. Our results show the potential of EVT-based methods to improve the determination of hydraulic risks in plants, and also point towards novel ways of characterizing water potential regulation. At the same time, our analyses highlight the large spatiotemporal variability in Ψ within species, particularly at daily scales, and the limitations of currently available data. We urge improved sampling designs to better characterize water potential variability within species and also better reporting of these data, if possible including individual-level measurements or at least a clear indication of the level of aggregation. Making raw data available together with relevant metadata (e.g. site coordinates, time stamps, local environmental data, occurrence of severe drought effects) would pave the way for building global, harmonized water potential databases and for a rigorous treatment of inclusion criteria for the determination of Ψ_{min} and associated measures

of hydraulic risk. Altogether, this effort should result in better understanding and prediction of plant responses to drought under ongoing climate change (Choat *et al.*, 2018; Brodribb *et al.*, 2020; Ruiz-Benito *et al.*, 2020).

Acknowledgements

This research was supported by the Spanish Ministry of Science and Innovation (MICINN) via competitive grants CGL2017-89149-C2-1-R and CGL2017-89149-C2-2-R. AV thanks MICINN for the project Inertia (PID201-111332RB-C22). JM-V benefited from an ICREA Academia award. We want to thank Sanna Sevanto, Nate McDowell and co-authors for making the SUMO water potential data publicly available (https://doi.org/10.15485/1439886).

Author contributions

JM-V, MM, LS, RP and MdC conceived the ideas and designed methodology; RP, LS, IA, SD, AV and JM-V provided data; JM-V and LB analysed the data; JM-V led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

ORCID

Ismael Aranda https://orcid.org/0000-0001-9086-7940
Llorenç Badiella https://orcid.org/0000-0002-9653-7421
Miquel Cáceres de https://orcid.org/0000-0001-7132-2080
Sylvain Delzon https://orcid.org/0000-0003-3442-1711
Jordi Martínez-Vilalta https://orcid.org/0000-0002-2332-7298
Maurizio Mencuccini https://orcid.org/0000-0003-0840-1477
Rafael Poyatos https://orcid.org/0000-0003-0521-2523
Louis S. Santiago https://orcid.org/0000-0001-5994-6122
Alberto Vilagrosa https://orcid.org/0000-0002-1432-1214

Data availability

All the data sources that support the findings of this study are available in the Supporting Information of this article. The full databases are available from the corresponding author upon reasonable request.

References

Adams HD, Zeppel MJB, Anderegg WRL, Hartmann H, Landhäusser SM, Tissue DT, Huxman TE, Hudson PJ, Franz TE, Allen CD *et al.* 2017. A multi-species synthesis of physiological mechanisms in drought-induced tree mortality. *Nature Ecology & Evolution* **1**: 1285.

Anderegg WRL, Klein T, Bartlett M, Sack L, Pellegrini AFA, Choat B, Jansen S. 2016. Meta-analysis reveals that hydraulic traits explain cross-species patterns of drought-induced tree mortality across the globe. *Proceedings of the National Academy of Sciences, USA* **113**: 5024–5029.

Atzmon N, Moshe Y, Schiller G. 2004. Ecophysiological response to severe drought in *Pinus halepensis* Mill. trees of two provenances. *Plant Ecology* **171**: 15–22.

Bartlett MK, Klein T, Jansen S, Choat B, Sack L. 2016. The correlations and sequence of plant stomatal, hydraulic, and wilting responses to drought. *Proceedings of the National Academy of Sciences, USA* **113**: 13098–13103.

Bartlett MK, Scoffoni C, Sack L. 2012. The determinants of leaf turgor loss point and prediction of drought tolerance of species and biomes: a global meta-analysis. *Ecology Letters* **15**: 393–405.

Benito Garzón MB, Muñoz NG, Wigneron J-P, Moisy C, Fernández-Manjarrés J, Delzon S, Jordan G. 2018. The legacy of water deficit on populations having experienced negative hydraulic safety margin. *Global Ecology and Biogeography* **27**: 346–356.

Bhaskar R, Ackerly DD. 2006. Ecological relevance of minimum seasonal water potentials. *Physiologia Plantarum* **127**: 353–359.

Bozinovic F, Calosi P, Spicer JI. 2011. Physiological correlates of geographic range in animals. *Annual Review of Ecology, Evolution, and Systematics* **42**: 155–179.

Breshears DD, Myers OB, Meyer CW, Barnes FJ, Zou CB, Allen CD, McDowell NG, Pockman WT. 2008. Tree die-off in response to global change-type drought: mortality insights from a decade of plant water potential measurements. *Frontiers in Ecology and the Environment* **7**: 185–189.

Brodribb TJ, Cochard H. 2009. Hydraulic failure defines the recovery and point of death in water-stressed conifers. *Plant Physiology* **149**: 575–584.

Brodribb TJ, Powers J, Cochard H, Choat B. 2020. Hanging by a thread? Forests and drought. *Science* **368**: 261–266.

Burgman M, Franklin J, Hayes KR, Hosack GR, Peters GW, Sisson SA. 2012. Modeling extreme risks in ecology. *Risk Analysis* **32**: 1956–1966.

Charrier G, Delzon S, Domec J-C, Zhang Li, Delmas CEL, Merlin I, Corso D, King A, Ojeda H, Ollat N *et al.* 2018. Drought will not leave your glass empty: Low risk of hydraulic failure revealed by long-term drought observations in world's top wine regions. *Science Advances* **4**: eaao6969.

Choat B, Brodribb TJ, Brodersen CR, Duursma RA, López R, Medlyn BE. 2018. Triggers of tree mortality under drought. *Nature* **558**: 531–539.

Choat B, Jansen S, Brodribb TJ, Cochard H, Delzon S, Bhaskar R, Bucci SJ, Feild TS, Gleason SM, Hacke UG *et al.* 2012. Global convergence in the vulnerability of forests to drought. *Nature* **491**: 752–755.

Cochard H, Badel E, Herbette S, Delzon S, Choat B, Jansen S. 2013. Methods for measuring plant vulnerability to cavitation: a critical review. *Journal of Experimental Botany* **64**: 4779–4791.

Coles S. 2001. *An introduction to statistical modeling of extreme values*. London, UK: Springer-Verlag.

Delzon S, Cochard H. 2014. Recent advances in tree hydraulics highlight the ecological significance of the hydraulic safety margin. *New Phytologist* **203**: 355–358.

Denny MW, Hunt LJH, Miller LP, Harley CDG. 2009. On the prediction of extreme ecological events. *Ecological Monographs* **79**: 397–421.

Duursma RA, Blackman CJ, Lopéz R, Martin-StPaul NK, Cochard H, Medlyn BE. 2019. On the minimum leaf conductance: its role in models of plant water use, and ecological and environmental controls. *New Phytologist* **221**: 693–705.

Edwards W, Moles AT, Chong C. 2015. Generalised extreme value distributions provide a natural hypothesis for the shape of seed mass distributions. *PLoS ONE* **10**: e0121724.

Embrechts P, Klüppelberg C, Mikosch T. 2013. *Modelling extremal events for insurance and finance*. Berlin, Germany: Springer-Verlag.

Guo JS, Hultine KR, Koch GW, Kropp H, Ogle K. 2020. Temporal shifts in iso/anisohydry revealed from daily observations of plant water potential in a dominant desert shrub. *New Phytologist* **225**: 713–726.

Gutschick VP, BassiriRad H. 2003. Extreme events as shaping physiology, ecology, and evolution of plants: toward a unified definition and evaluation of their consequences. *New Phytologist* **160**: 21–42.

Hammond WM, Yu K, Wilson LA, Will RE, Anderegg WRL, Adams HD. 2019. Dead or dying? Quantifying the point of no return from hydraulic failure in drought-induced tree mortality. *New Phytologist* **223**: 1834–1843.

Head AW, Hardin JS, Adolph SC. 2012. Methods for estimating peak physiological performance and correlating performance measures. *Environmental and Ecological Statistics* **19**: 127–137.

Hochberg U, Rockwell FE, Holbrook NM, Cochard H. 2018. Iso/Anisohydry: a plant-environment interaction rather than a simple hydraulic trait. *Trends in Plant Science* **23**: 112–120.

Hoshika Y, Osada Y, de Marco A, Peñuelas J, Paoletti E. 2018. Global diurnal and nocturnal parameters of stomatal conductance in woody plants and major crops. *Global Ecology and Biogeography* **27**: 257–275.
Katz RW, Brush GS, Parlange MB. 2005. Statistics of extremes: modeling ecological disturbances. *Ecology* **86**: 1124–1134.
Kingsolver JG, Buckley LB. 2017. Quantifying thermal extremes and biological variation to predict evolutionary responses to changing climate. *Philosophical Transactions of the Royal Society B: Biological Sciences* **372**: 20160147.
Kolb KJ, Davis SD. 1994. Drought tolerance and xylem embolism in co-occurring species of coastal sage and chaparral. *Ecology* **75**: 648–659.
Liang X, Ye Q, Liu H, Brodribb TJ. 2021. Wood density predicts mortality threshold for diverse trees. *New Phytologist* **229**: 3053–3057.
Makkonen L, Tikanmäki M. 2019. An improved method of extreme value analysis. *Journal of Hydrology X* **2**: 100012.
Martinez-Vilalta J, Anderegg WRL, Sapes G, Sala A. 2019. Greater focus on water pools may improve our ability to understand and anticipate drought-induced mortality in plants. *New Phytologist* **223**: 22–32.
Martínez-Vilalta J, Garcia-Forner N. 2017. Water potential regulation, stomatal behaviour and hydraulic transport under drought: deconstructing the iso/anisohydric concept. *Plant, Cell & Environment* **40**: 962–976.
Martínez-Vilalta J, Poyatos R, Aguadé D, Retana J, Mencuccini M. 2014. A new look at water transport regulation in plants. *New Phytologist* **204**: 105–115.
Martin-StPaul N, Delzon S, Cochard H. 2017. Plant resistance to drought depends on timely stomatal closure. *Ecology Letters* **20**: 1437–1447.
McDowell N, Pockman WT, Allen CD, Breshears DD, Cobb N, Kolb T, Plaut J, Sperry J, West A, Williams DG *et al.* 2008. Mechanisms of plant survival and mortality during drought: why do some plants survive while others succumb to drought? *New Phytologist* **178**: 719–739.
Meinzer FC, Johnson DM, Lachenbruch B, McCulloh KA, Woodruff DR. 2009. Xylem hydraulic safety margins in woody plants: coordination of stomatal control of xylem tension with hydraulic capacitance. *Functional Ecology* **23**: 922–930.
Meinzer FC, Woodruff DR, Marias DE, Mcculloh KA, Sevanto S. 2014. Dynamics of leaf water relations components in co-occurring iso- and anisohydric conifer species. *Plant, Cell & Environment* **37**: 2577–2586.
Mencuccini M, Minunno F, Salmon Y, Martínez-Vilalta J, Hölttä T. 2015. Coordination of physiological traits involved in drought-induced mortality of woody plants. *New Phytologist* **208**: 396–409.
Murray M, Soh WK, Yiotis C, Batke S, Parnell AC, Spicer RA, Lawson T, Caballero R, Wright IJ, Purcell C *et al.* 2019. Convergence in maximum stomatal conductance of C_3 woody angiosperms in natural ecosystems across bioclimatic zones. *Frontiers in Plant Science* **10**: 558.
Reiss R-D, Thomas M. 2007. *Statistical analysis of extreme values, with applications to insurance, finance, hydrology and other fields.* Basel, Switzerland: Birkhäuser.
Ruiz-Benito P, Vacchiano G, Lines ER, Reyer CPO, Ratcliffe S, Morin X, Hartig F, Mäkelä A, Yousefpour R, Chaves JE *et al.* 2020. Available and missing data to model impact of climate change on European forests. *Ecological Modelling* **416**: 108870.
Sanchez-Martinez P, Martínez-Vilalta J, Dexter KG, Segovia RA, Mencuccini M. 2020. Adaptation and coordinated evolution of plant hydraulic traits. *Ecology Letters* **23**: 1599–1610.
Sapes G, Roskilly B, Dobrowski S, Maneta M, Anderegg WRL, Martinez-Vilalta J, Sala A. 2019. Plant water content integrates hydraulics and carbon depletion to predict drought-induced seedling mortality. *Tree Physiology* **39**: 1300–1312.
Scholz FW, Stephens MA. 1987. K-sample anderson-darling tests. *Journal of the American Statistical Association* **82**: 918–924.
Tyree MT, Jarvis PG. 1982. Water in tissues and cells. In: Lange OL, Nobel PS, Osmond CB, Ziegler H, eds. *Encyclopedia of plant physiology. Physiological plant ecology II: Water relations and carbon assimilation.* Berlin/Heidelberg, Germany: Springer, 35–77.
Tyree MT, Sperry JS. 1988. Do woody plants operate near the point of catastrophic xylem dysfunction caused by dynamic water stress? Answers from a model. *Plant Physiology* **88**: 574–580.
Tyree MT, Zimmermann MH. 2002. *Xylem structure and the ascent of Sap.* Berlin, Germany: Springer-Verlag.
Urli M, Porté AJ, Cochard H, Guengant Y, Burlett R, Delzon S. 2013. Xylem embolism threshold for catastrophic hydraulic failure in angiosperm trees. *Tree Physiology* **33**: 672–683.
Venturas MD, MacKinnon ED, Dario HL, Jacobsen AL, Pratt RB, Davis SD. 2016. Chaparral shrub hydraulic traits, size, and life history types relate to species mortality during California's historic drought of 2014. *PLoS ONE* **11**: e0159145.
Venturas MD, Pratt RB, Jacobsen AL, Castro V, Fickle JC, Hacke UG. 2019. Direct comparison of four methods to construct xylem vulnerability curves: Differences among techniques are linked to vessel network characteristics. *Plant, Cell & Environment* **42**: 2422–2436.
Wu G, Guan K, Li Y, Novick KA, Feng X, McDowell NG, Konings AG, Thompson SE, Kimball JS, De Kauwe MG *et al.* 2021. Interannual variability of ecosystem iso/anisohydry is regulated by environmental dryness. *New Phytologist* **229**: 2562–2575.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information section at the end of the article.

Fig. S1 Histograms showing the water potential distribution for the two 'species' included in the synthetic datasets.

Fig. S2 Histograms showing the distribution of water potentials (absolute value in MPa) from the 17 species in the 'MV14' dataset.

Fig. S3 Histograms showing the distribution of water potentials (absolute value in MPa) from the 26 species in the 'raw' dataset.

Fig. S4 Effect of sample size on minimum water potential estimates from the 'MV14' dataset.

Fig. S5 Effect of sample size on minimum water potential estimates from the synthetic datasets, allowing variability in the underlying distribution between sites.

Notes S1 A very short introduction to the statistics of extreme values.

Table S1 Species selected from the 'MV14' dataset.

Table S2 Species selected from the 'raw' dataset.

Table S3 Comparison of different distribution fits to the tails of the water potential distribution data for all the species in the 'raw' dataset.

Table S4 Variance components analysis of Ψ for all the species in the 'raw' dataset and specifically for *Juniperus monosperma* and *Pinus edulis*.

Table S5 Comparison of different distribution fits for different combinations of site and year for six representative species from the 'raw' dataset.

Table S6 Comparison of Ψ_{min} values estimated using the GPD approach on different data subsets for *Juniperus monosperma* and *Pinus edulis*.

Table S7 Mean water potential causing 50% loss of hydraulic conductivity (Ψ_{PLC50}), minimum Ψ_{PLC50}, mean water potential causing 88% loss of hydraulic conductivity (Ψ_{PLC88}) and probability of Ψ_{min} reaching Ψ_{PLC50} for the species in the 'raw' dataset.

Table S8 Comparison of Ψ_{min} values estimated in this study with Ψ_{lethal} values obtained from literature sources.

Please note: Wiley Blackwell are not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.