Biogeosciences

Gap-filling a spatially explicit plant trait database: comparing imputation methods and different levels of environmental information

Rafael Poyatos[1,2], **Oliver Sus**[3], **Llorenç Badiella**[4], **Maurizio Mencuccini**[1,5], **and Jordi Martínez-Vilalta**[1,6]

[1]CREAF, E08193 Bellaterra (Cerdanyola del Vallès), Catalonia, Spain
[2]Laboratory of Plant Ecology, Faculty of Bioscience Engineering, Ghent University, Coupure links 653, 9000 Gent, Belgium
[3]EUMETSAT, Eumetsat Allee 1, 64295 Darmstad, Germany
[4]Servei d'Estadística Aplicada, Universitat Autònoma de Barcelona, Cerdanyola del Vallès 08193, Barcelona, Spain
[5]ICREA, Barcelona, Spain
[6]Universitat Autònoma de Barcelona, E08193 Bellaterra (Cerdanyola del Vallès), Catalonia, Spain

Correspondence: Rafael Poyatos (r.poyatos@creaf.uab.es)

Abstract. The ubiquity of missing data in plant trait databases may hinder trait-based analyses of ecological patterns and processes. Spatially explicit datasets with information on intraspecific trait variability are rare but offer great promise in improving our understanding of functional biogeography. At the same time, they offer specific challenges in terms of data imputation. Here we compare statistical imputation approaches, using varying levels of environmental information, for five plant traits (leaf biomass to sapwood area ratio, leaf nitrogen content, maximum tree height, leaf mass per area and wood density) in a spatially explicit plant trait dataset of temperate and Mediterranean tree species (Ecological and Forest Inventory of Catalonia, IEFC, dataset for Catalonia, north-east Iberian Peninsula, 31 900 km^2). We simulated gaps at different missingness levels (10–80 %) in a complete trait matrix, and we used overall trait means, species means, k nearest neighbours (kNN), ordinary and regression kriging, and multivariate imputation using chained equations (MICE) to impute missing trait values. We assessed these methods in terms of their accuracy and of their ability to preserve trait distributions, multi-trait correlation structure and bivariate trait relationships. The relatively good performance of mean and species mean imputations in terms of accuracy masked a poor representation of trait distributions and multivariate trait structure. Species identity improved MICE imputations for all traits, whereas forest structure and topography improved imputations for some traits.

No method performed best consistently for the five studied traits, but, considering all traits and performance metrics, MICE informed by relevant ecological variables gave the best results. However, at higher missingness (> 30 %), species mean imputations and regression kriging tended to outperform MICE for some traits. MICE informed by relevant ecological variables allowed us to fill the gaps in the IEFC incomplete dataset (5495 plots) and quantify imputation uncertainty. Resulting spatial patterns of the studied traits in Catalan forests were broadly similar when using species means, regression kriging or the best-performing MICE application, but some important discrepancies were observed at the local level. Our results highlight the need to assess imputation quality beyond just imputation accuracy and show that including environmental information in statistical imputation approaches yields more plausible imputations in spatially explicit plant trait datasets.

1 Introduction

Trait-based ecology has emerged in recent years as one of the most active ecological sub-disciplines, specially in plant ecology (Westoby and Wright, 2006; Violle et al., 2007). The move from a taxonomic perspective of biodiversity towards a focus on continuous axes of functional variation

holds promise for greater generalisation, synthesis and predictive ability in ecology (Funk et al., 2016; Shipley et al., 2016). As a result, plant ecologists have increasingly embraced trait-based approaches because they may be specially suited to study plant strategies (Reich, 2014), community assembly and dynamics (McGill et al., 2006), or ecosystem functioning, particularly in the context of global environmental change (Reichstein et al., 2014). But trait-based ecology is also unquestionably thriving because of the increasing availability and reliability of plant trait data (Kattge et al., 2011).

Plant trait databases compiled from multiple individual contributions lack a common design and inevitably result in sparse data matrices (e.g. Jetz et al., 2016). Complete-case analyses (i.e. data analyses using only sampling units with complete data availability) entail a reduced sampling size, which complicates community-level studies (Pakeman, 2014) and limits the spatial coverage of trait maps usable in trait-based models of ecosystem function. Data deletion may also bias parameter estimates (e.g. in trait relationships) if the data are not missing completely at random (MCAR; Little and Rubin, 2002; Nakagawa and Freckleton, 2008). Imputation (i.e. gap-filling) of missing data with plausible values has the potential to overcome some of these limitations, but has only relatively recently started to be widely advocated in ecology (Nakagawa and Freckleton, 2008). It should be noted, however, that imputation may not be recommended in certain studies (Blonder, 2016).

Single imputation methods replace a missing datum by one value and proceed with the analysis as if the imputed data had been observed (Nakagawa and Freckleton, 2008). Within these approaches, species mean or median imputation are probably the most widely used methods in ecology, but they ignore the variance of the imputed variables. Model-based imputation methods use other variables in the dataset to impute missing data, but they substantially alter the univariate trait distributions and the covariance structure of the dataset (Gelman and Hill, 2007). Approaches such as k nearest neighbour (kNN) or machine-learning methods (Stekhoven and Bühlmann, 2012) may be more appropriate to impute multivariate datasets, preserving their covariance structure (Eskelson et al., 2009; Penone et al., 2014). In a multiple imputation framework, m imputed datasets are obtained through simulation and may be jointly analysed to provide parameter estimates that take into account the uncertainty introduced by the imputations themselves (e.g. Fisher et al., 2003). Some multiple imputation techniques, such as multivariate imputation using chained equations (MICE) may be specially well suited to preserve the original structure and distribution of multivariate datasets (van Buuren and Groothuis-Oudshoorn, 2011; van Buuren, 2012).

While forest inventories have adopted statistical imputation methods for some time, as for example the kNN methods (Eskelson et al., 2009, and references therein), imputation methods have only recently started to be used in trait-based ecology (Baraloto et al., 2010; Pyšek et al., 2015).

Complex imputation methods such as kNN, MICE or random forests generally outperform overall mean or species mean imputations (Penone et al., 2014; Taugourdeau et al., 2014). In earlier applications of these methods, it has been common to assume that interspecific trait variability was dominant, compared to intraspecific variability. The strong phylogenetic signal may then be sufficient to impute species-averaged trait values using taxonomic information (Swenson, 2014). However, intraspecific variability in plant traits may be substantial (Siefert et al., 2015; Vilà-Cabrera et al., 2015) and imputation methods that use environmental information may be more appropriate when assessing trait relationships and trait–environment covariance in a spatially explicit context. Biotic or abiotic variables other than the trait matrix of interest can be included in imputation algorithms as auxiliary variables to reduce imputation bias (Azur et al., 2011; Rezvan et al., 2015). Geostatistical methods of spatial interpolation can also be used with (e.g. regression kriging) or without (e.g. ordinary kriging) auxiliary variables (e.g. Hengl et al., 2007).

Additional challenges occur in the imputation of traits in large databases. The expected declining performance of imputation methods with increasing missingness levels may be trait and dataset dependent (Penone et al., 2014; Taugourdeau et al., 2014). Moreover, the impact of imputations on altering bivariate trait relationships has only been assessed for single relationships (Penone et al., 2014; Schrodt et al., 2015) and not for the multiple relevant relationships within a plant trait dataset. Likewise, there are few studies quantifying how different imputation methods alter the multivariate covariance structure of plant trait datasets (Schrodt et al., 2015).

Our overarching aim here is to assess the performance of different imputation methods to fill simulated gaps at different missingness levels in a spatially explicit plant trait dataset (IEFC, Ecological and Forest Inventory of Catalonia, northeast Iberian Peninsula). We imputed these missing data using single imputation (kNN), multiple imputation (MICE) and geostatistical approaches (ordinary and regression kriging, OrdKrig and RegKrig, respectively) and compared the imputations with baseline scenarios of overall mean and species mean imputation. Imputation performance was assessed in terms of accuracy, univariate trait distributions, multivariate trait structure and deviations in trait relationships. Our specific objectives are (i) to test which imputation method (overall mean imputation, kNN, MICE, OrdKrig) performed best when relying only on plant trait data; (ii) to assess the impact of including additional predictors (i.e. environmental information such as species identity, climate, forest structure, topography, lithology and sampling date) in MICE and kNN imputations; (iii) to compare the performance of kNN, MICE and RegKrig using optimum levels of environmental information (i.e. the best set of predictors in objective ii); and, finally, (iv) to apply the best-performing method to fill the gaps in a major subset of the IEFC database to obtain contin-

uous maps of plant traits for the main forest species across a relatively large Mediterranean region.

2 Methods

2.1 Study area

The study area is the entire territory of Catalonia ($31\,900\,\mathrm{km}^2$), in the north-east Iberian peninsula. Catalonia has 38 % forest cover (1.2×10^6 ha) and forests are largely dominated by species belonging to the Pinaceae and Fagaceae families. We selected 13 tree species, including six *Pinus* spp., five deciduous and evergreen *Quercus* spp., *Abies alba* and *Fagus sylvatica*, which altogether cover > 90 % of the forested area in Catalonia (see Sect. S1, Fig. S1 in the Supplement).

2.2 Data

Plant trait and forest data were retrieved from the Ecological and Forest Inventory of Catalonia (IEFC), carried out between 1988 and 1998 (Gracia et al., 2000). A complete description of the sampling scheme and methods used to measure plant traits in the IEFC can be found in Sect. S1. The subset of the IEFC limited to the 13 study species, hereby called the IEFC incomplete dataset, included 5495 plots. Forest structure, lithology and sampling information for each plot were retrieved from the IEFC database, whereas climate data were obtained from the Climatic Digital Atlas of Catalonia, with a spatial resolution of 180 m (Ninyerola et al., 2000).

We selected five plant traits (leaf mass per area, LMA, $\mathrm{mg\,cm}^{-2}$; leaf nitrogen per unit mass, N_{mass}, %mass; maximum tree height, H_{max}, m; wood density, WD, $\mathrm{g\,cm}^{-3}$; leaf biomass to sapwood area ratio, $B_{\mathrm{L}} : A_{\mathrm{S}}$, $\mathrm{t\,m}^{-2}$) that are used to describe major plant functional strategies (Westoby et al., 2002; Wright et al., 2004; Chave et al., 2009; Laforest-Lapointe et al., 2014). In Catalan forests, four of these traits (LMA, N_{mass}, H_{max} WD) mostly vary between families (Pinaceae and Fagaceae) and within species (Vilà-Cabrera et al., 2015). The missing data patterns in this trait data matrix shows a much higher percentage of missing data (hereafter, missingness) for foliar traits, corresponding to a less intense sampling of these traits (Fig. 1). These intentional missing data (van Buuren, 2012) would correspond to a planned missing data design, where missingness at random (MAR) is deliberately applied (Nagakawa, 2015).

2.3 Experimental design

All data manipulations, imputations and statistical analyses were performed with the R programming language (R Core Team, 2015). We created a subset of the IEFC incomplete dataset only including those plots ($N = 630$) where all five

traits had been measured on the dominant species (IEFC complete dataset). In this dataset, we randomly deleted measured values at different probability levels (10, 20, 30, 50 and 80 %) and independently for each trait; thus, the missing data artificially introduced are missing completely at random. This data deletion was replicated to yield 30 simulated datasets for each missingness level (Fig. 1). Hence, the different imputation methods were assessed on 150 datasets (5 missingness levels × 30 replicates).

We ran different single and multiple imputation algorithms (see Sect. 2.4) to fill the gaps in the trait data of the simulated incomplete datasets. Single imputation methods yield $m = 1$ imputed dataset per simulated dataset and here we set the multiple imputation methods to yield $m = 5$ datasets per simulated dataset to incorporate imputation uncertainty. Prior to the calculation of different performance metrics for each dataset, trait values in multiply-imputed datasets were averaged (Penone et al., 2014). Performance metrics were assessed using the measured values of each trait in the IEFC complete dataset (see Sect. 2.6). Note that each imputed dataset contains both measured and gap-filled data, but the expression imputed values refers only to gap-filled data.

2.4 Imputation methods

We compared imputation methods with different degrees of complexity. We used two simple approaches to provide baseline imputations; mean imputation ("Mean") filled missing data using the overall mean value for each trait and species mean imputation (Spmean) replaced missing values with trait means computed for each species. Because of the spatial nature of the dataset, we also tested two geostatistical approaches – ordinary kriging (OrdKrig) and regression kriging (RegKrig). Lastly, we also used two methods designed to handle multivariate datasets: k nearest neighbour imputation and MICE.

Ordinary kriging calculates a weighted average of nearby observations to predict values of a target variable in an unmeasured location, with weights that minimise prediction error and depend on spatial structure of the target variable via a variogram model (Hengl et al., 2007). Regression kriging combines a deterministic model of the target variable as a function of auxiliary variables with kriging applied to fit the residuals (Hengl et al., 2007). We included climate and forest structure variables in the model used for regression kriging imputations (see Sect. 2.5), but not species identity, because there were not enough data to generate the experimental variograms for some of the less common species for all the simulations. We performed all kriging imputations with the "autoKrige" function in the automap R package. This function tests different variogram models and applies the best-fit variogram model for kriging (Hiemstra et al., 2009).

The kNN method calculates a multivariate distance using only non-missing variables, selects the k nearest plots with measured values for the target missing trait and aggregates
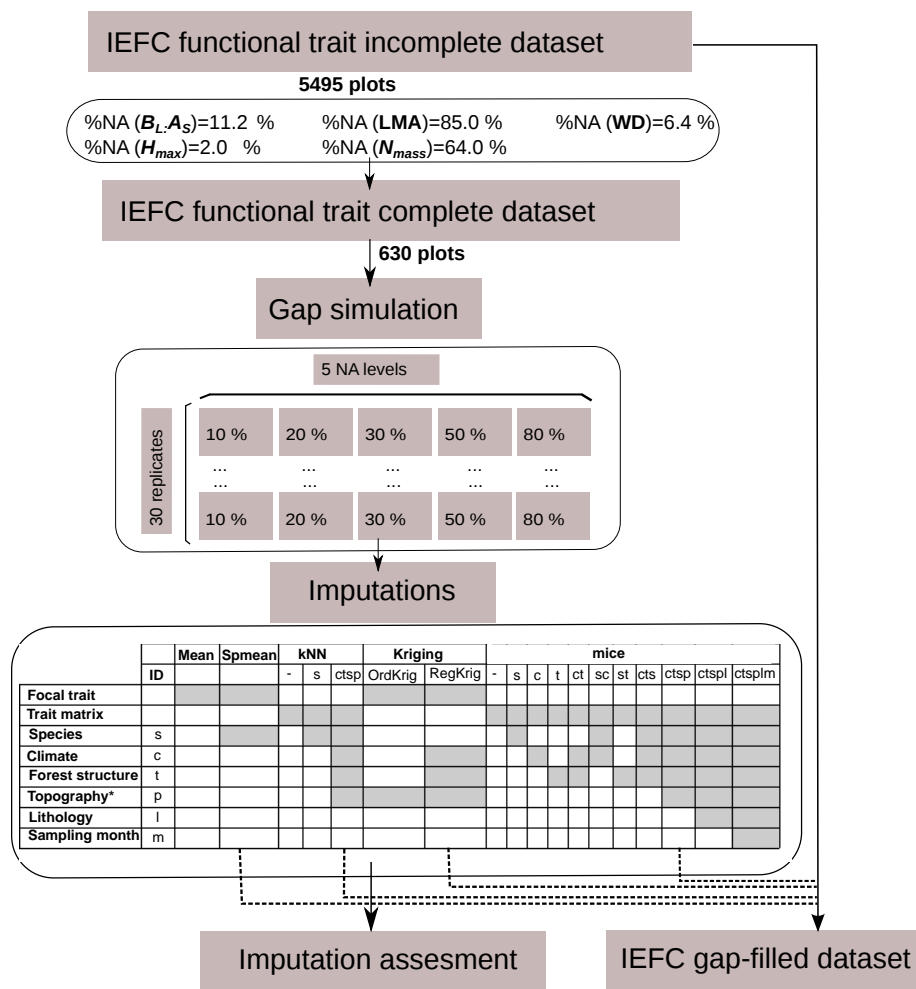
Figure 1. Description of the experimental design. A subset was obtained from the incomplete IEFC trait dataset containing only plots where all functional traits had been measured (complete dataset) to perform the gap simulations and the imputations. Imputation methods are described in terms of the input information used. The selected methods for the final application of imputation methods to obtain a gap-filled IEFC trait dataset are also shown.

these k neighbouring values to replace the missing value (R package VIM; Templ et al., 2013). We selected $k = 7$ and median aggregation after some preliminary tests (Sect. S2, Fig. S2 in the Supplement). We also analysed how the inclusion of auxiliary variables in the distance calculation affected imputation performance (see Sect. 2.5).

The MICE algorithm (van Buuren and Groothuis-Oudshoorn, 2011; van Buuren, 2012) sequentially and iteratively imputes incomplete data, variable by variable, using individual imputation models conditionally specified by the user. One cycle through all the imputed variables is one iteration, and MICE performs t iterations in m parallel streams, generating m multiple imputations (Sect. S3). We set $t = 20$ to ensure convergence and to minimise the effects of imputation order (van Buuren, 2012). Stochasticity is introduced in the imputation process because the parameters of the univariate imputation models are drawn from their posterior distri-

butions, obtained using a Gibbs sampler (van Buuren, 2012). Assessments of imputation methods in the ecological literature have not tested the impact of the choice of univariate imputation models within MICE (Penone et al., 2014; Taugourdeau et al., 2014). Here we showed that predictive mean matching (PMM), the default algorithm in the mice package, performed well compared to alternative methods (Sect. S3, Figs. S3 and S4). Therefore, we used MICE with PMM as the univariate imputation model, also because it is robust to non-normality and preserves non-linear relationships between variables (Morris et al., 2014). Several parameters must be tuned to specify the imputation models in the R implementation of MICE (mice package) to yield reliable imputations (van Buuren and Groothuis-Oudshoorn, 2011). The specific settings used in this study are assessed in Sect. S3 (Figs. S3 and S4). Please note that we will use the uppercase acronym "MICE" to refer to the technique in general and the

lowercase acronym "mice" to refer to a particular application in this study.

2.5 Comparative assessment of imputation methods

We conducted three methodological comparisons of imputation performance. A first exercise compared Mean, OrdKrig, kNN and mice imputations. "Mean" imputations used only the information on the target trait, OrdKrig additionally used the spatial coordinates, and mice and kNN included only the information in the trait matrix.

A second exercise assessed in detail the impact on trait imputation of including additional environmental information as auxiliary variables in MICE and kNN. We focused our detailed analysis on MICE only but we also made a simplified comparison between kNN and MICE (see next paragraph). The auxiliary variables we considered were species identity, a set of climatic variables (mean annual temperature, annual thermal amplitude, both in °C), a set of forest structure variables (total aboveground biomass [$T ha^{-1}$] and stem density [$stems ha^{-1}$]), a set of topographical variables (county, elevation [m a.s.l.], slope [°] and aspect), lithology (calcareous, non-calcareous or undetermined) and sampling month. These predictors were complete and they did not need to be imputed themselves. The selection of the specific variables describing climate and forest structure was based on a recent analysis of trait variation in the same IEFC dataset (Vilà-Cabrera et al., 2015). We further added topographical variables, lithology and sampling month given that they may influence some trait values (Niinemets, 2015; Simpson et al., 2016). Species identity (s), climate (c) and forest structure (t) were introduced in a factorial design to identify those combinations of variables leading to improved imputations. Because we expected them to play a secondary role in explaining trait variability, topography (p), lithology (l) and sampling month (m) were sequentially added to MICE and kNN imputations using species, climate and forest structure. Topography included spatial structure through the "county" variable; preliminary tests using coordinates instead of county did not show better results. Thus, mice_ctsplm was the MICE application with the highest level of environmental information (Fig. 1).

The third exercise compared species mean imputations (Spmean) with MICE and kNN using two different levels of auxiliary variables: (i) only species identity (mice_s and kNN_s) and (ii) the level of auxiliary variables which performed best overall in the second exercise. In this same exercise, we also compared the previous approaches with OrdKrig and regression kriging (RegKrig) imputations. This third exercise thus compares a baseline scenario of Spmean with imputation approaches informed either by species identity only or by an optimum level of environmental information.

2.6 Statistical evaluation of the imputations

Imputation performance was evaluated by comparing the imputed datasets with the complete, original dataset. A first set of metrics, normalised root mean square error (NRMSE) and Kling–Gupta efficiency (KGE), was calculated only for those values that had been randomly deleted and subsequently gap-filled. We tested whether the distribution of imputed and original trait values differed using a two-sample Kolmogorov–Smirnov test, which tests the null hypothesis that two samples are identically distributed.

For each simulated dataset and trait, we calculated the normalised root mean square error as a measure of accuracy:

$$\text{NRMSE} = \sqrt{\frac{\text{mean}\left[\left(y_{\text{imp}} - y_{\text{obs}}\right)^2\right]}{\text{var}\left(y_{\text{obs}}\right)}}, \qquad (1)$$

where y_{imp} and y_{obs} represent the vectors of imputed and observed values for a given trait, respectively. Values of NRMSE approaching zero denote a better performance of the imputation method. We also calculated a dataset-averaged NRMSE by averaging the values of NRMSE for all the traits.

We further assessed imputation performance for each trait by using KGE, a goodness-of-fit measure originally developed for hydrological models, as implemented in the R package hydroGOF (Zambrano-Bigiarini, 2014):

$$\text{KGE} = 1 - \sqrt{(r-1)^2 + (\text{vr}-1)^2 + (\beta-1)^2}, \qquad (2)$$

where r is the Pearson correlation coefficient between observed and imputed values, vr is the ratio of the standard deviations between imputed and observed values, and β is the ratio of imputed and observed means. The KGE range is $[-\infty, 1]$, with higher values indicating better imputation performance. KGE jointly assesses correlation, bias and difference in variability between imputed and observed values, and it is therefore a powerful, synthetic indicator of imputation quality in spatially explicit datasets.

A second set of metrics compared the whole complete trait dataset Y_{obs} with the whole imputed dataset Y_{imp} (i.e. including observed and gap-filled trait values). The deviations from the original multi-trait correlation structure of the trait dataset were quantified by comparing the correlation matrices of the original and imputed datasets using the following index:

$$\Delta \text{ cormat} = \sum |L[\text{cor}(Y_{\text{obs}})] - L[\text{cor}(Y_{\text{imp}})]|, \qquad (3)$$

where $L[\text{cor}(Y_{\text{obs}})]$ denotes the lower triangular part of the correlation matrix of the observed dataset and $L[\text{cor}(Y_{\text{imp}})]$ denotes the lower triangular part of the correlation matrix of the imputed dataset. Δ cormat is indicative of the aggregated absolute difference between correlation matrices. Note that some traits were log transformed before the calculation of the corresponding correlation matrix, following Vilà-Cabrera et al. (2015).
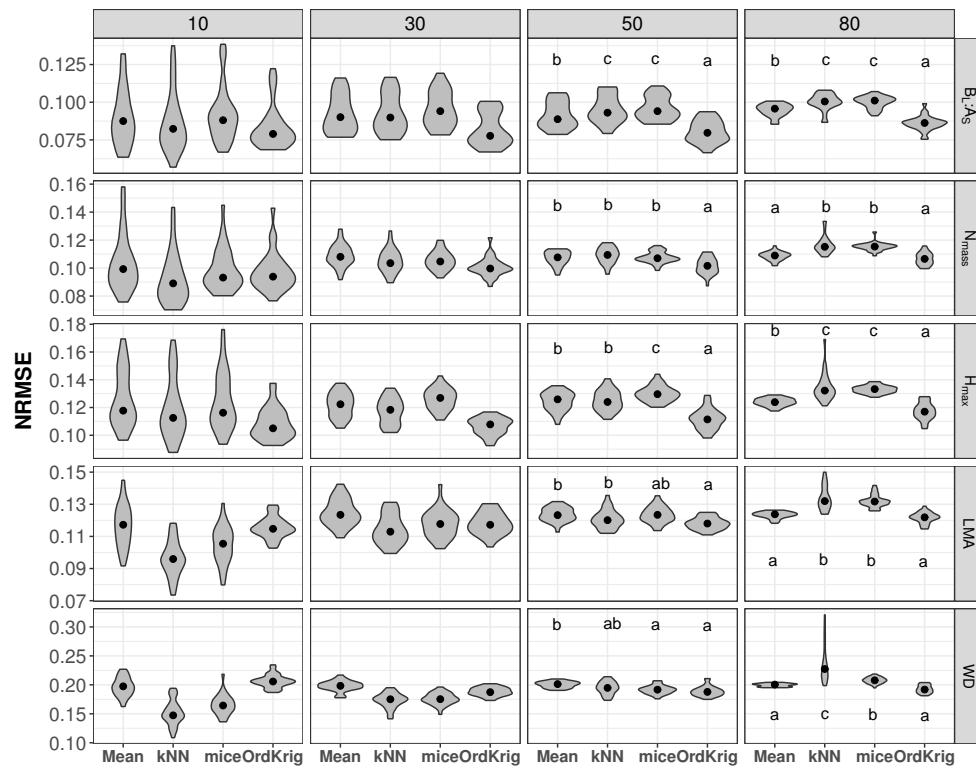
Figure 2. Trait-specific NRMSE at increasing missingness levels (10–80 %) for different imputation methods: overall trait mean (Mean), mice (using only the trait matrix in the predictor set), kNN (using only the trait matrix for the distance calculation) and ordinary kriging (OrdKrig). Traits: leaf biomass to sapwood area ratio, $B_L : A_S$ $(\mathrm{t\,m^{-2}})$; leaf nitrogen per unit mass, N_{mass} (%mass); maximum tree height, H_{max} (m); leaf mass per area, LMA $(\mathrm{mg\,cm^{-2}})$; wood density, WD $(\mathrm{g\,m\,cm^{-3}})$. Letters denote results of multiple comparisons, in alphabetical order from highest to lowest performance.

We also tested the impact of the imputation algorithms on selected bivariate trait relationships – H_{max}–WD and N_{mass}–LMA (log transformed when necessary) – as the correlation coefficients (r) of these relationships were > 0.3 in absolute value and were highly significant in the complete dataset. We quantified the relative difference between the complete and the imputed datasets by calculating

$$\%\Delta r = 100 \cdot |r_{obs} - r_{imp}| \big/ |r_{obs}| . \qquad (4)$$

Throughout the paper, we show violin plots representing the median and the distribution of each performance metric as a function of missingness levels, but we only graphically display the 10, 30, 50 and 80 % levels, for ease of visualisation. We modelled imputation metrics in a linear mixed-effects model (LME) as a function of the interaction between imputation method and missingness, with dataset replicate as the random effect. The LME model was fitted using the nlme package in R (Pinheiro et al., 2018) and pairwise comparisons of model coefficients were performed using the lsmeans and lstrends functions in the lsmeans package (Lenth, 2016).

2.7 Imputing traits for the main forest species in Catalonia

Finally, we applied three imputation methods to gap-fill and map the five traits across all the plots in the IEFC incomplete dataset. We chose Spmean as the most widely used imputation method in trait-based studies, RegKrig as a reference geostatistical approach including auxiliary variables and mice_ctsp as the best method overall, considering all traits and performance metrics (see Sect. 3). We ran mice_ctsp setting $m = 50$ (i.e. 50 imputations per missing value), a value closer to the missingness rate, as recommended for final MICE applications (van Buuren, 2012).

3 Results and discussion

3.1 Mean imputations compared to MICE and kNN imputations using only trait information

In general, mice and kNN imputations resulted in more accurate imputations in terms of NRMSE than Mean at low missingness rates (10 %). However, at moderate and high miss-
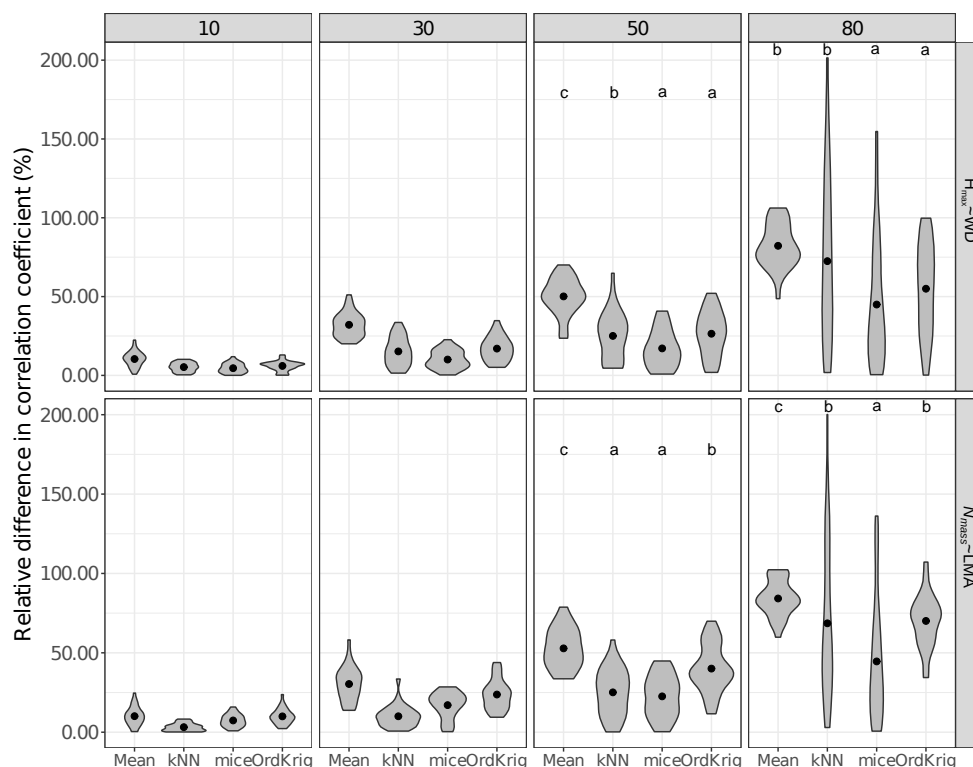
Figure 3. Errors in the correlation coefficient for two selected trait relationships at increasing missingness levels (10–80 %) and for different imputation methods: overall trait mean (Mean), mice (using only the trait matrix in the predictor set), kNN (using only the trait matrix for the distance calculation) and ordinary kriging (OrdKrig). Letters denote results of multiple comparisons, in alphabetical order from highest to lowest performance. Traits involved in the relationships are leaf nitrogen per unit mass, N_{mass} (%mass); maximum tree height, H_{max} (m); leaf mass per area, LMA (mg cm^{-2}); wood density, WD (gm cm^{-3}).

ingness both mice and kNN were comparable to or outperformed by Mean, and specially by OrdKrig (Figs. 2, S5). OrdKrig was the best-performing method, in terms of NRMSE, at missingness ≥ 50 % ($P < 0.05$), although for three traits its performance was indistinguishable from that of Mean imputations (N_{mass}, H_{max}, LMA; $P > 0.05$). Even if Mean imputations imply the rather naive assumption that species identity may be unknown in a given dataset, it is nonetheless useful to compare Mean imputations against mice and kNN, which use the full trait matrix for prediction. In this case, trait covariation did not improve imputations at high missingness; recent assessments also report that the performance of MICE and kNN notably declines when missingness is ≥ 30 % (Penone et al., 2014; Taugourdeau et al., 2014). Therefore, our results for OrdKrig, compared to those for mice and kNN, show that spatial structure, rather than trait covariation, may provide more accurate trait imputations when gaps are frequent (Fig. 2, Figs. S5 and S6).

As expected (Gelman and Hill, 2007), Mean imputation severely altered trait distributions (Fig. S6) and introduced larger errors in selected trait correlations (Fig. 3). Mean imputations also tended to cause larger deviations in the cor-

relation matrix (Fig. S5). kNN showed the lowest Δ cormat below 50 % missingness ($P > 0.05$) but its performance declined at high missingness (Fig. S5). In contrast, mice closely tracked observed trait distributions (Fig. S6), introduced the least error in trait correlations under high missingness levels (Fig. 3; $P < 0.05$) and yielded low Δ cormat at extreme missingness levels (Fig. S5). Recent results also show that kNN tends to introduce larger bias in bivariate trait relationships compared to MICE (Penone et al., 2014). OrdKrig imputations altered distributions and trait correlations more than mice (Fig. 3, Fig. S6), but they performed similarly in terms of Δ cormat at all missingness levels (Fig. S5).

3.2 MICE imputations using different levels of environmental information

Introducing auxiliary variables as predictors improved MICE performance substantially but these improvements were dependent on the specific predictor set and trait (Fig. 4). Species identity increased KGE for all traits (Fig. 4) and it was the major predictor for N_{mass}, LMA and WD, as all MICE applications with species identity performed significantly better than those not including it (Fig. 4; $P < 0.05$). Forest struc-
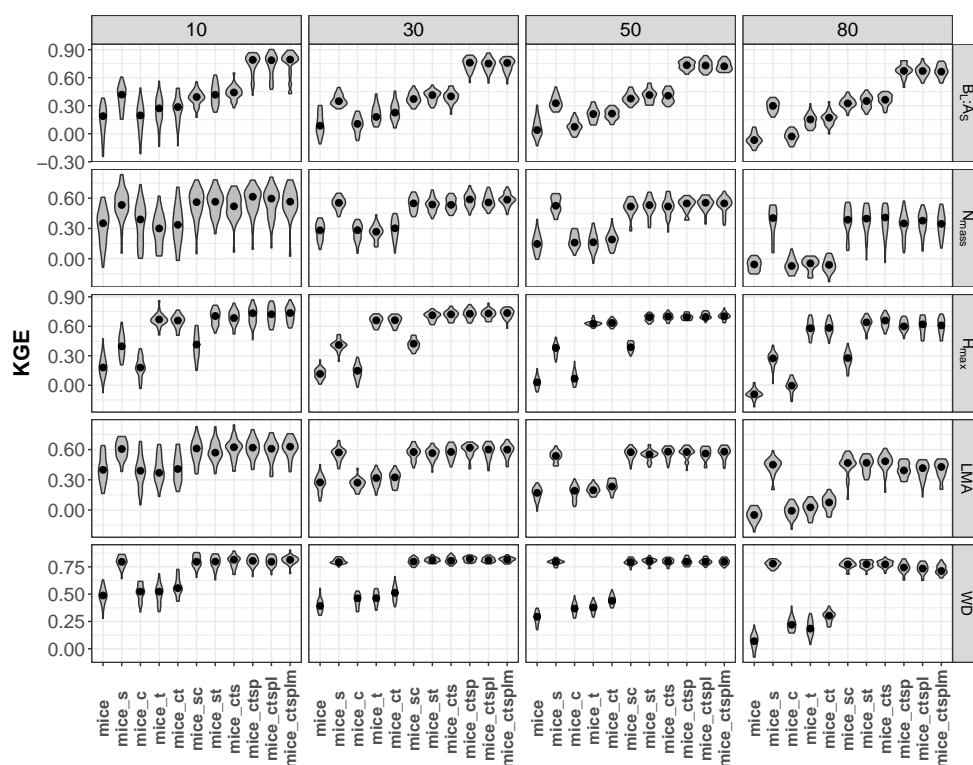
Figure 4. Trait-specific KGE at increasing missingness levels (10–80 %) and for different MICE imputations using different combinations of additional predictor sets: species identity (s), climate (c), forest structure (t), spatial structure (p), lithology (l) and sampling month (m). See Fig. 1 for an overall view of the experimental design and the methods section for a detailed description of the variables employed in each predictor set. Traits: leaf biomass to sapwood area ratio, $B_L : A_S$ (t m^{-2}); leaf nitrogen per unit mass, N_{mass} (%mass); maximum tree height, H_{max} (m); leaf mass per area, LMA (mg cm^{-2}); wood density, WD (g m cm^{-3}). Higher values of KGE imply higher performance.

ture notably improved imputations for H_{max} and for $B_L : A_S$ particularly at missingness ≥ 50 % ($P < 0.05$). Climate only produced significant increases in KGE (i.e. compare mice with mice_c in Fig. 4) for H_{max} and WD ($P < 0.05$). Our results are in line with the distinct role of phylogeny and environmental variables as drivers of trait variability recently observed for the same tree species using the IEFC (Laforest-Lapointe et al., 2014; Vilà-Cabrera et al., 2015). One of these studies shows that, after controlling for family (Pinaceae and Fagaceae), environmental variables only explained a substantial fraction of the variability for H_{max}; they explained very little variability for LMA and WD and played no role in explaining N_{mass} (Vilà-Cabrera et al., 2015).

Including topography in MICE imputations only substantially improved $B_L : A_S$ imputations (compare mice_cts with mice_ctsp, $P < 0.05$), probably because the leaf area used in $B_L : A_S$ calculations are obtained from county-level allometries, and county is one of the variables included in the topography predictor set (see Sects. 2 and S1). Nevertheless, introducing sampling month in the predictor sets did not appreciably improve MICE imputations in terms of KGE (Fig. 4), despite the fact that phenological variation has been reported

for some foliar traits (Niinemets, 2015; but see Fajardo and Siefert, 2016). Lithology did not appreciably improve MICE imputations, in contrast to the reported influence of soil pH on some foliar traits (Maire et al., 2015; Simpson et al., 2016).

At high missingness (≥ 50 %), mice_ctsp (including climate, forest structure, species and topography) was always within the best-performing methods ($P < 0.05$), except for LMA and WD at 80 % missingness, according to KGE results (Fig. 4). In terms of dataset-averaged NRMSE, Δcormat (data not shown) and preservation of trait distributions (Fig. S7), the inclusion of topography only produced a significant improvement for dataset NRMSE at 50 % missingness ($P < 0.05$).

Including auxiliary variables as predictors also decreased $\% \Delta r$ for selected trait relationships (Fig. 5). The best-performing MICE applications (lower $\% \Delta r$) always included species identity and forest structure; including other auxiliary variables did not lower $\% \Delta r$ significantly ($P > 0.05$).

Our results collectively suggest that, apart from species identity, different types of environmental information, par-
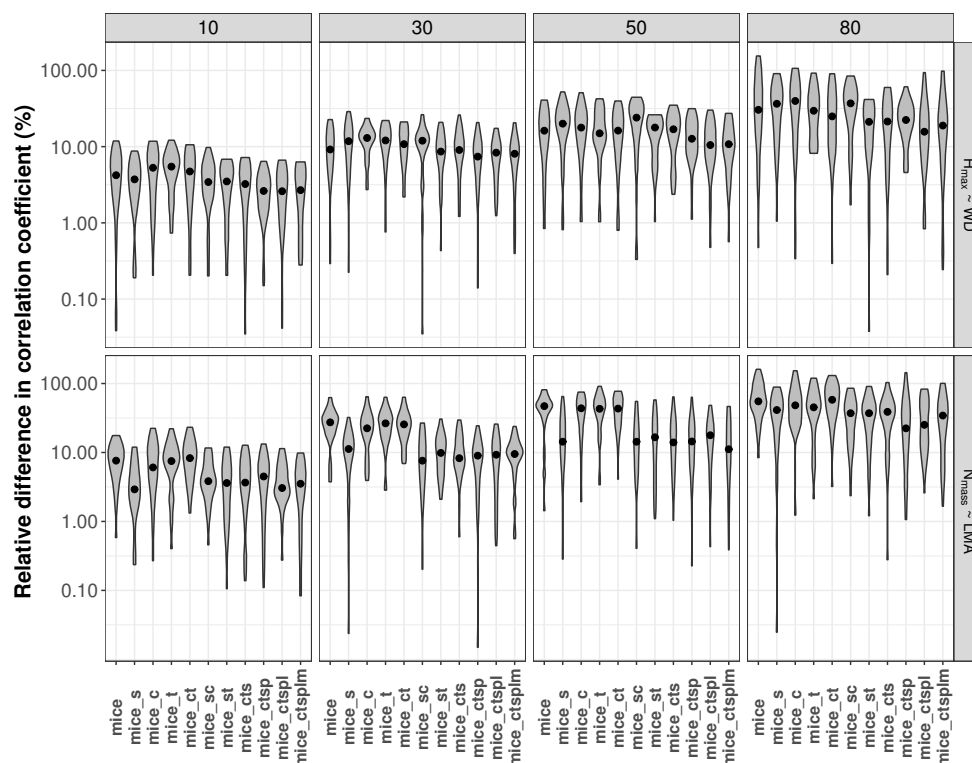
Figure 5. Errors in the correlation coefficient for two selected trait relationships at increasing missingness levels (10–80 %) and for different MICE imputations using different combinations of additional predictor sets: species identity (s), climate (c), forest structure (t), topography (p), lithology (l) and sampling month (m). See Fig. 1 for an overall view of the experimental design and the methods section for a detailed description of the variables employed in each predictor set. Note that the y axis is in the logarithmic scale. Traits involved in the relationships are leaf nitrogen per unit mass, N_{mass} (%mass); maximum tree height, H_{max} (m); leaf mass per area, LMA (mg cm^{-2}); wood density, WD (g m cm^{-3}).

ticularly forest structure and topography, may improve statistical imputation schemes. In contrast, the role of climate, lithology and sampling month in improving imputations was comparatively minor. However, we selected mice_ctsp as the method that performed best for all traits, because adding climate did not deteriorate imputation performance and not including topography would worsen $B_L : A_S$ imputations. A negligible influence of climate and soil data on trait imputation in the TRY database was also recently reported (Schrodt et al., 2015). It is unclear, however, to what extent these results simply reflect the relatively poor quality of the climate and soil data generally available at regional scales.

### 3.3 Comparing imputation methods using optimum levels of environmental information

Adding auxiliary variables to calculate the distance matrix also improved kNN imputations. Values of KGE for kNN_ctsp were much higher than those observed for kNN imputations (data not shown), which only included the trait data in the distance matrix. This improvement was largely driven by the inclusion of species identity; only for $B_L : A_S$

and H_{max} did kNN_ctsp perform significantly better than kNN_s (Fig. 6; $P < 0.05$). Likewise, adding climate and forest structure as auxiliary variables improved RegKrig performance compared to OrdKrig (Fig. 6, $P < 0.05$), except for N_{mass}. For both kNN and kriging methods, WD and H_{max} were the traits for which these improvements were largest.

In terms of KGE, mice_ctsp was the best-performing method at 50 % missingness for all traits, together with Spmean for N_{mass} and LMA and with RegKrig for H_{max} ($P > 0.05$ for comparisons between mice_ctsp, Spmean and RegKrig for these traits). However, at 80 % missingness, mice_ctsp only ranked first for $B_L : A_S$, whereas Spmean showed the highest KGE for N_{mass}, LMA and WD and RegKrig performed best for H_{max} (Fig. 6). These results are consistent with the prominent role of taxonomic identity in explaining variability in foliar traits and WD and with the higher predictive ability of environmental and spatial information in explaining H_{max} (Vilà-Cabrera et al., 2015). The LME model showed that the rate of increase in KGE with increasing missingness was lowest for 'Spmean' in four out of five traits (Table 1). Compared to Spmean and RegKrig,

Table 1. Tukey pairwise comparisons of the LME model coefficients (and 95 % lower and upper confidence limits) relating trait-specific KGE (higher values of KGE imply higher performance) to increasing missingness levels for different imputation methods: species mean (Spmean); mice and kNN with species as predictor (mice_s and kNN_s, respectively); mice and kNN with species, climate, forest structure and spatial variables as predictors (mice_ctsp and kNN_ctsp, respectively); ordinary kriging (OrdKrig); and regression kriging (RegKrig). Traits: leaf biomass to sapwood area ratio, $B_L : A_S$ (t m^{-2}); leaf nitrogen per unit mass, N_{mass} (%mass); maximum tree height, H_{max} (m); leaf mass per area LMA (mg cm^{-2}); wood density, WD (g cm^{-3}). Different letters, in alphabetical order following the increasing order of the model coefficient, denote significant differences ($P < 0.05$) in the results of multiple comparisons.

	Coefficient	SE	df	Lower CL	Upper CL	
$B_L : A_S$						
OrdKrig	5.27×10^{-4}	3.93×10^{-4}	148	2.50×10^{-4}	1.30×10^{-3}	a
RegKrig	9.25×10^{-4}	3.93×10^{-4}	148	1.49×10^{-4}	1.70×10^{-3}	ab
mice_s	1.16×10^{-3}	3.93×10^{-4}	148	3.79×10^{-4}	1.93×10^{-3}	ab
Spmean	1.29×10^{-3}	3.93×10^{-4}	148	5.17×10^{-4}	2.07×10^{-3}	ab
mice_ctsp	1.70×10^{-3}	3.93×10^{-4}	148	9.28×10^{-4}	2.48×10^{-3}	bc
kNN_s	2.29×10^{-3}	3.93×10^{-4}	148	1.52×10^{-3}	3.07×10^{-3}	c
kNN_ctsp	3.27×10^{-3}	3.93×10^{-4}	148	2.50×10^{-3}	4.05×10^{-3}	d
N_{mass}						
OrdKrig	1.14×10^{-3}	2.80×10^{-4}	148	5.88×10^{-4}	1.69×10^{-3}	a
Spmean	1.41×10^{-3}	2.80×10^{-4}	148	8.54×10^{-4}	1.96×10^{-3}	a
RegKrig	1.49×10^{-3}	2.80×10^{-4}	148	9.40×10^{-4}	2.04×10^{-3}	a
mice_s	1.50×10^{-3}	2.80×10^{-4}	148	9.49×10^{-4}	2.05×10^{-3}	a
mice_ctsp	2.33×10^{-3}	2.80×10^{-4}	148	1.77×10^{-3}	2.88×10^{-3}	b
kNN_ctsp	2.38×10^{-3}	2.80×10^{-4}	148	1.82×10^{-3}	2.93×10^{-3}	b
kNN_s	2.70×10^{-3}	2.80×10^{-4}	148	2.15×10^{-3}	3.26×10^{-3}	b
H_{max}						
RegKrig	7.11×10^{-4}	2.32×10^{-4}	148	2.53×10^{-4}	1.17×10^{-3}	a
Spmean	7.23×10^{-4}	2.32×10^{-4}	148	2.65×10^{-4}	1.18×10^{-3}	a
mice_s	1.14×10^{-3}	2.32×10^{-4}	148	6.84×10^{-4}	1.60×10^{-3}	ab
OrdKrig	1.22×10^{-3}	2.32×10^{-4}	148	7.62×10^{-4}	1.68×10^{-3}	abc
mice_ctsp	1.64×10^{-3}	2.32×10^{-4}	148	1.19×10^{-3}	2.10×10^{-3}	bc
kNN_s	1.75×10^{-3}	2.32×10^{-4}	148	1.30×10^{-3}	2.21×10^{-3}	bc
kNN_ctsp	1.77×10^{-3}	2.32×10^{-4}	148	1.32×10^{-3}	2.23×10^{-3}	c
LMA						
OrdKrig	7.36×10^{-4}	1.99×10^{-4}	148	3.43×10^{-4}	1.13×10^{-3}	a
RegKrig	8.36×10^{-4}	1.99×10^{-4}	148	4.43×10^{-4}	1.23×10^{-3}	a
Spmean	1.24×10^{-3}	1.99×10^{-4}	148	8.50×10^{-4}	1.64×10^{-3}	a
mice_s	1.82×10^{-3}	1.99×10^{-4}	148	1.43×10^{-3}	2.21×10^{-3}	b
kNN_ctsp	2.26×10^{-3}	1.99×10^{-4}	148	1.87×10^{-3}	2.65×10^{-3}	b
mice_ctsp	2.88×10^{-3}	1.99×10^{-4}	148	2.48×10^{-3}	3.27×10^{-3}	c
kNN_s	3.20×10^{-3}	1.99×10^{-4}	148	2.80×10^{-3}	3.59×10^{-3}	c
WD						
OrdKrig	-7.13×10^{-4}	1.41×10^{-4}	148	-9.92×10^{-4}	-4.35×10^{-4}	a
RegKrig	2.65×10^{-5}	1.41×10^{-4}	148	-2.52×10^{-4}	3.05×10^{-4}	b
Spmean	2.94×10^{-4}	1.41×10^{-4}	148	1.52×10^{-5}	5.72×10^{-4}	b
mice_s	7.70×10^{-4}	1.41×10^{-4}	148	4.91×10^{-4}	1.05×10^{-3}	c
kNN_s	1.08×10^{-3}	1.41×10^{-4}	148	8.05×10^{-4}	1.36×10^{-3}	c
kNN_ctsp	1.14×10^{-3}	1.41×10^{-4}	148	8.58×10^{-4}	1.41×10^{-3}	cd
mice_ctsp	1.56×10^{-3}	1.41×10^{-4}	148	1.28×10^{-3}	1.84×10^{-3}	d
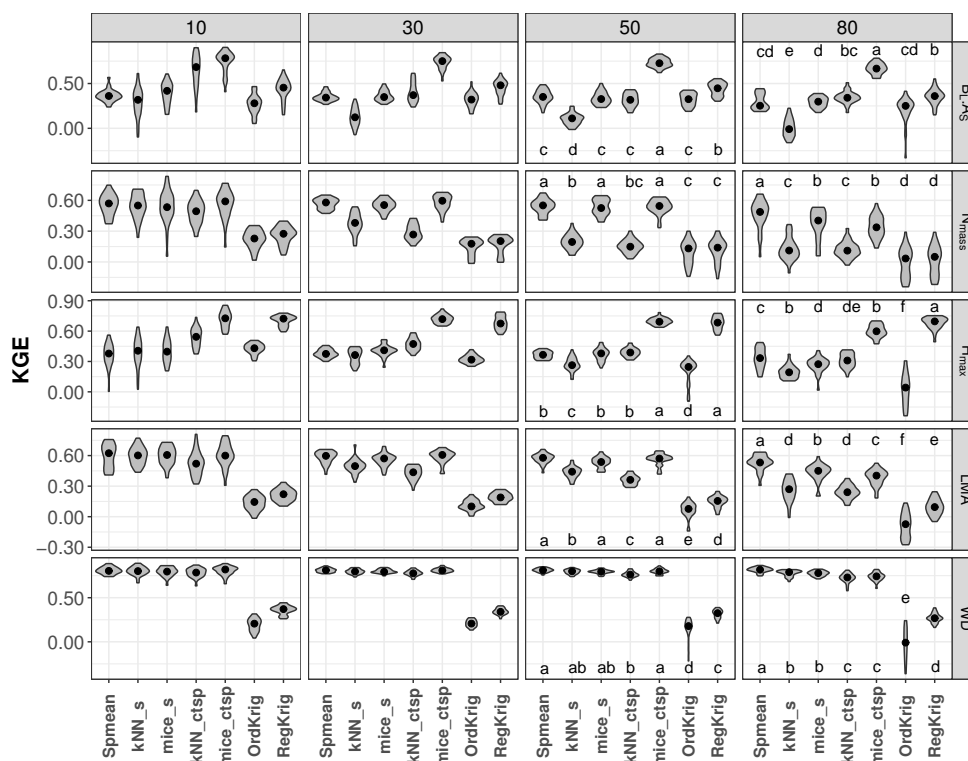
Figure 6. Trait-specific KGE at increasing missingness levels (10–80 %) for different imputation methods: species mean (Spmean); mice and kNN with species as predictor (mice_s and kNN_s, respectively); mice and kNN with species, climate, forest structure and spatial variables as predictors (mice_ctsp and kNN_ctsp, respectively); ordinary kriging (OrdKrig); and regression kriging (RegKrig). Higher values of KGE imply higher performance. Traits: leaf biomass to sapwood area ratio, $B_L : A_S$ (t m^{-2}); leaf nitrogen per unit mass, N_{mass} (%mass); maximum tree height, H_{max} (m); leaf mass per area, LMA (mg cm^{-2}); wood density, WD (g m cm^{-3}). Higher values of KGE imply higher performance. Letters denote results of multiple comparisons, in alphabetical order from highest to lowest performance.

performance of MICE and kNN declined more with increasing missingness (Table 1, Fig. S8 and Table S1 in the Supplement), but MICE generally outperformed kNN (Figs. 6 and 7), as already observed in a recent imputation assessment of species-level, life-history traits (Penone et al., 2014). In terms of dataset-averaged NRMSE, mice_ctsp and Spmean were the best-performing methods at 50 and 80 % missingness, respectively ($P < 0.05$).

Kernel density plots (Fig. S9) and Kolmogorov–Smirnov tests (Fig. S10) showed that MICE produced imputations (especially mice_ctsp) most consistent with observed distributions at all missingness levels (Figs. S9 and S10). Spmean and OrdKrig imputations modified trait distributions substantially, while kNN_ctsp and RegKrig showed an intermediate performance, but generally far from that of mice_ctsp (Table 1, Fig. S8, Table S1). Spmean and kriging imputations also yielded larger Δ cormat values compared to the rest of the methods ($P < 0.05$), reflecting their lower ability to maintain trait correlation structures.

For the selected trait relationships, mice_ctsp showed the lowest values of %Δr at ≥ 50 % missingness ($P < 0.05$), although for the N_{mass}–LMA relationship kNN_s and mice_s

performed equally well at 50 and 80 % missingness, respectively (Fig. 7; $P > 0.05$). Spmean imputations showed variable results, severely altering the H_{max}–WD relationship ($P < 0.05$, Fig. 7) but showing comparable performance to mice_ctsp for the N_{mass}–LMA relationship at 80 % missingness ($P > 0.05$, Fig. 7). Kriging imputations did not succeed in minimising changes in reproducing trait correlations (Fig. 7).

Using imputed or incomplete datasets did not lead to large differences in the studied trait relationships when missingness was < 50 % (Figs. S11 and S12). However, at high missingness, using imputed datasets led to comparatively larger departures from the relationships obtained with the complete dataset, especially for the N_{mass}–LMA relationship. No imputation method appeared to perform consistently better than others in preserving trait relationships at high missingness levels (Figs. S11 and S12) and, under these conditions, using incomplete datasets appeared to correctly reproduce the observed trait relationships in the complete dataset.
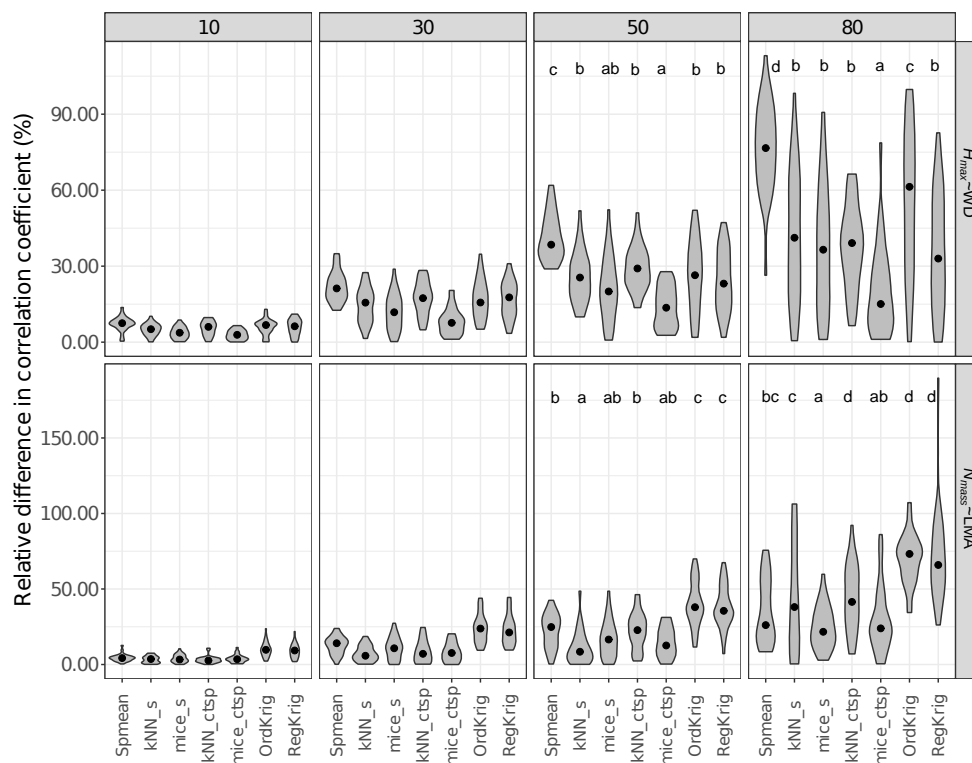
Figure 7. Errors in the correlation coefficient for two selected trait relationships at increasing missingness levels (10–80 %) and for different imputation methods: species mean (Spmean); mice and kNN with species as predictor (mice_s and kNN_s, respectively); mice and kNN with species, climate, forest structure and spatial variables as predictors (mice_ctsp and kNN_ctsp, respectively); ordinary kriging (OrdKrig); and regression kriging (RegKrig). Letters denote results of multiple comparisons, in alphabetical order from highest to lowest performance.

3.4 Imputing traits for the main forest species in Catalonia

The application of mice_ctsp allowed us to fill the gaps in the IEFC incomplete dataset and quantify the variation among the multiple imputations, providing an estimation of the level of confidence in the imputed values for specific traits (Fig. 8). Spmean and RegKrig show a broadly similar spatial pattern of trait variation compared to mice_ctsp, although some important discrepancies between Spmean and mice_ctsp can be observed in the north-eastern pre-coastal and coastal area for LMA (Fig. S13). Here, Spmean imputations tend to predict lower values compared to mice_ctsp. These areas are mostly dominated by *Quercus suber* forests (Fig. S1), and LMA was only measured in 5 out of the 149 plots of this species present in the IEFC incomplete dataset. Therefore, as there is little information on trait covariation for the imputation of LMA in *Q. suber* plots, MICE imputations are largely based on the auxiliary variables and they yield a distinct spatial pattern of trait variation, compared to Spmean. Imputations obtained using regression kriging result in more blurred spatial patterns, relative to other imputation methods (Fig. S13).

4 Implications and conclusions

The problem of missing data is ubiquitous in plant trait datasets of a regional to global scope. Recently, ecologists have made substantial progress in (i) the assessment of the best imputation methods in trait-based applications, (ii) how these methods perform with increasing missingness, (iii) which ecological covariates aid to improve imputations and (iv) how different imputation methods impact the results of trait-based analyses (Pakeman, 2014; Taugourdeau et al., 2014; Penone et al., 2014; Schrodt et al., 2015). Most effort thus far, however, has been directed at imputing species-level trait means and all the abovementioned questions have rarely been assessed on the same dataset. Here we deal with all the previous issues simultaneously and also deal with the spatial component of trait variability, where the intra-specific component cannot be neglected. We did not focus on differences in imputation errors across species because this issue is, to a large extent, related to the degree of trait variability explained by biotic and abiotic predictors across different taxa, which was recently reported by Vilà-Cabrera et al. (2015).

One limitation of this study is that we simulate data missing completely at random, whereas a missing at random as-
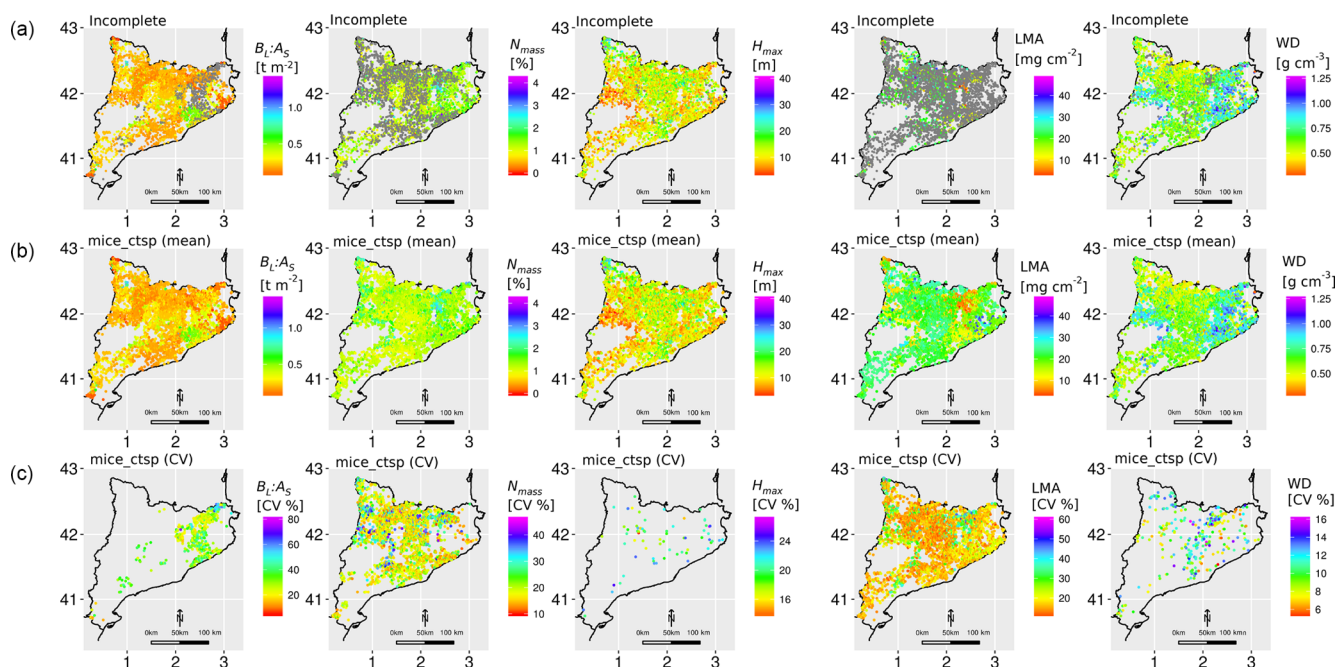
Figure 8. Maps with the distribution of functional traits across the selected plots in the IEFC. The first row shows the incomplete dataset, with missing values in grey. The second row shows the mean of 50 multiple imputations for each missing value using the mice_ctsp approach (MICE imputation using species identity, climate, forest structure and topography as predictors). The third row shows the corresponding coefficient of variation (CV) for these multiple imputations. Note that, for the third row, only imputed values are shown and that the colour scale varies across different traits. Traits: leaf biomass to sapwood area ratio, $B_L : A_S$ ($t\,m^{-2}$); leaf nitrogen per unit mass, N_{mass} (%mass); maximum tree height, H_{max} (m); leaf mass per area, LMA ($mg\,cm^{-2}$); wood density, WD ($g\,m\,cm^{-3}$). Higher values of KGE imply higher performance.

sumption would have probably been more realistic given the properties of the dataset (Nakagawa, 2015). However, a recent study did not show differences in trait imputation performance between these two missing data mechanisms (Penone et al., 2014). Our study assesses different imputation methods in spatially explicit datasets with multivariate missing data. Amongst the methods assessed here, MICE and kNN are the most adequate to impute multivariate datasets, as they can be used when predictors also include missing data. Kriging methods may be more difficult to apply when predictors are also missing, but we have shown that, at high missingness levels and when environmental information is lacking, they can outperform MICE and kNN. This implies that geostatistical methods may sometimes provide more accurate imputations than those using trait covariation.

Our results show that, in terms of trait prediction error, no imputation method performs best consistently for the five studied traits. However, when all performance metrics are jointly considered (i.e. errors in trait prediction, multivariate trait distribution and trait correlations), MICE informed by relevant ecological variables outperforms approaches based on trait averaging, geostatistical models and kNN methods, albeit this superiority of MICE tends to vanish at very high missingness levels. For kNN, MICE and kriging imputations we have highlighted the key role of auxiliary variables as

necessary covariates to yield reliable imputations in spatially explicit settings. This result calls for the inclusion of site-specific environmental variables associated with trait data in trait databases. The importance of covariates differed across traits, but, in addition to the expected influence of species, climate and topography in predicting trait values, we also showed a prominent role of forest structure for some traits. The ongoing development of global databases of vegetation structure (e.g. Dengler et al., 2014) will likely enable the incorporation of stand variables in trait imputation approaches using spatial and environmental information (Butler et al., 2017).

Given the limited number of species in our study, reflecting the relatively low richness of the studied communities, taxonomic information introduced as species identity was enough to improve imputations of all studied traits. However, in studies coping with a larger set of species, phylogeny may need to be considered in the imputation models (Schrodt et al., 2015; Swenson et al., 2017). For global trait datasets, a combination of imputation with data augmentation approaches (e.g. Nakagawa and Freckleton, 2008) has been proposed to minimise potential errors in trait-driven analyses caused by incomplete and biased species sampling (Sandel et al., 2015).

Compared to other imputation approaches, MICE is well suited to deal with multivariate missing data (i.e. MICE produces imputations when some predictors are also missing) and provides information to quantify the uncertainty associated with the imputed data (Fig. 8). MICE uses multivariate relationships in the dataset to impute missing data, and this may raise concerns about potential circularity in trait or trait–environment correlations. Despite these concerns, it has been argued that the full inference framework based on multiply-imputed datasets would minimise circularity (van Buuren, 2012). Because our comparative assessment of imputation methods is already complex, here we have focused on the initial imputation stage, the first step of the full process (e.g. Nakagawa and Freckleton, 2008). MICE produces multiple datasets, with imputed values drawn from distributions, and these datasets can be combined in the analysis and pooling steps. The analysis step refers to the estimation of the parameters of scientific interest (e.g. a regression coefficient) for each dataset. In MICE, parameters can be pooled across datasets to produce unbiased estimates and standard errors, providing a natural way to take into account the additional uncertainty introduced in the analysis by the presence of missing data, and to avoid circularity effects (van Buuren, 2012). However, ecological studies using multiple imputation approaches usually only apply the imputation step (Baraloto et al., 2010; Paine et al., 2011; Pyšek et al., 2015; Díaz et al., 2016) and do not take advantage of the multiple imputation framework to quantify the uncertainty resulting from the presence of missing data (but see Fisher et al., 2003).

Our results have important implications given that the demand for spatially explicit datasets is increasing rapidly and that species mean imputation and casewise data deletion are still widespread practices in trait-based ecology. We show that species mean imputation may result in substantial information loss that may hinder research development on important topics in functional biogeography, such as the ecological drivers and implications of intraspecific trait variability (e.g. Siefert et al., 2015). Gap-filled multivariate trait datasets may increase the robustness of syntheses of plant form and function (Díaz et al., 2016) and trait-driven modelling approaches (Yang et al., 2015). We also show that spatially distributed layers of environmental information may improve trait mapping, increasing spatial resolution and/or sample size in trait-driven ecosystem process models (Christoffersen et al., 2016).

The Supplement related to this article is available online at https://doi.org/10.5194/bg-15-2601-2018-supplement.

References

Azur, M. J., Stuart, E. A., Frangakis, C., and Leaf, P. J.: Multiple imputation by chained equations: what is it and how does it work?, Int. J. Meth. Psych. Res., 20, 40–49, https://doi.org/10.1002/mpr.329, 2011.

Baraloto, C., Timothy Paine, C. E., Poorter, L., Beauchene, J., Bonal, D., Domenach, A.-M., Hérault, B., Patiño, S., Roggy, J.-C., and Chave, J.: Decoupled leaf and stem economics in rain forest trees, Ecol. Lett., 13, 1338–1347, https://doi.org/10.1111/j.1461-0248.2010.01517.x, 2010.

Blonder, B.: Do hypervolumes have holes?, Am. Nat., 187, E93–E105, https://doi.org/10.1086/685444, 2016.

Butler, E. E., Datta, A., Flores-Moreno, H., Chen, M., Wythers, K. R., Fazayeli, F., Banerjee, A., Atkin, O. K., Kattge, J., Amiaud, B., Blonder, B., Boenisch, G., Bond-Lamberty, B., Brown, K. A., Byun, C., Campetella, G., Cerabolini, B. E. L., Cornelissen, J. H. C., Craine, J. M., Craven, D., Vries, F. T. de, Díaz, S., Domingues, T. F., Forey, E., González-Melo, A., Gross, N., Han, W., Hattingh, W. N., Hickler, T., Jansen, S., Kramer, K., Kraft, N. J. B., Kurokawa, H., Laughlin, D. C., Meir, P., Minden, V., Niinemets, Ü., Onoda, Y., Peñuelas, J., Read, Q., Sack, L., Schamp, B., Soudzilovskaia, N. A., Spasojevic, M. J., Sosinski, E., Thornton, P. E., Valladares, F., Bodegom, P. M. van, Williams, M., Wirth, C., and Reich, P. B.: Mapping local and global variability in plant trait distributions, P. Natl. Acad. Sci. USA, 114, E10937–E10946, https://doi.org/10.1073/pnas.1708984114, 2017.

Chave, J., Coomes, D., Jansen, S., Lewis, S. L., Swenson, N. G., and Zanne, A. E.: Towards a worldwide wood economics spectrum, Ecol. Lett., 12, 351–366, https://doi.org/10.1111/j.1461-0248.2009.01285.x, 2009.

Christoffersen, B. O., Gloor, M., Fauset, S., Fyllas, N. M., Galbraith, D. R., Baker, T. R., Kruijt, B., Rowland, L., Fisher, R. A., Binks, O. J., Sevanto, S., Xu, C., Jansen, S., Choat, B., Mencuccini, M., McDowell, N. G., and Meir, P.: Linking hydraulic traits to tropical forest function in a size-structured and trait-

driven model (TFS v.1-Hydro), Geosci. Model Dev., 9, 4227–4255, https://doi.org/10.5194/gmd-9-4227-2016, 2016.

Dengler, J., Bruelheide, H., Purschke, O., Chytr, M., Jansen, F., Hennekens, S. M., Jandt, U., Jiménez-Alfaro, B., Kattge, J., Pillar, V. D., Sandel, B., Winter, M., and sPlot Consortium: sPlot – the new global vegetation-plot database for addressing trait–environment relationships across the world's biomes, in: Biodiversity and Vegetation: Patterns, Processes, Conservation, edited by: Mucina, L., Price, J. N., and Kalwi, J. M., Kwongan Foundation, Perth, AU, p. 90, available at: https://www.bayceer.uni-bayreuth.de/pfloek/de/pub/pub/130464/JD196_Dengler_et_al_2014_IAVS_Proceedings.pdf (last access: 21 November 2016), 2014.

Díaz, S., Kattge, J., Cornelissen, J. H. C., Wright, I. J., Lavorel, S., Dray, S., Reu, B., Kleyer, M., Wirth, C., Prentice, I. C., Garnier, E., Bönisch, G., Westoby, M., Poorter, H., Reich, P. B., Moles, A. T., Dickie, J., Gillison, A. N., Zanne, A. E., Chave, J., Wright, S. J., Sheremet'ev, S. N., Jactel, H., Baraloto, C., Cerabolini, B., Pierce, S., Shipley, B., Kirkup, D., Casanoves, F., Joswig, J. S., Günther, A., Falczuk, V., Rüger, N., Mahecha, M. D. and Gorné, L. D.: The global spectrum of plant form and function, Nature, 529, 161–171, https://doi.org/10.1038/nature16489, 2016.

Eskelson, B. N. I., Temesgen, H., Lemay, V., Barrett, T. M., Crookston, N. L., and Hudak, A. T.: The roles of nearest neighbor methods in imputing missing data in forest inventory and monitoring databases, Scand. J. Forest Res., 24, 235–246, https://doi.org/10.1080/02827580902870490, 2009.

Fajardo, A. and Siefert, A.: Phenological variation of leaf functional traits within species, Oecologia, 180, 951–959, https://doi.org/10.1007/s00442-016-3545-1, 2016.

Fisher, D. O., Blomberg, S. P., and Owens, I. P. F.: Extrinsic versus intrinsic factors in the decline and extinction of Australian marsupials, P. Roy. Soc. B-Biol. Sci., 270, 1801–1808, https://doi.org/10.1098/rspb.2003.2447, 2003.

Funk, J. L., Larson, J. E., Ames, G. M., Butterfield, B. J., Cavender-Bares, J., Firn, J., Laughlin, D. C., Sutton-Grier, A. E., Williams, L., and Wright, J.: Revisiting the Holy Grail: using plant functional traits to understand ecological processes, Biol. Rev., 92, 1156–1173, https://doi.org/10.1111/brv.12275, 2016.

Gelman, A. and Hill, J.: Data Analysis Using Regression and Multilevel/Hierarchical Models, Cambridge University Press, Cambridge, UK, 2007.

Gracia, C., Burriel, J., Mata, T., and Vayreda, J.: Inventari Ecològic i Forestal de Catalunya: Centre de Recerca Ecològica i Aplicacions Forestals, 10 volumes, CREAF, Bellaterra, Spain, 2000.

Gracia, C., Vayreda, J., Burriel, J. A., Ibàñez, J. J., and Mata, T.: Ecological and Forest Inventory of Catalonia: Complete Plant Trait Dataset for Imputation Assessment, https://doi.org/10.5281/zenodo.1209812, 2018.

Hengl, T., Heuvelink, G. B. M., and Rossiter, D. G.: About regression-kriging: from equations to case studies, Comput. Geosci., 33, 1301–1315, https://doi.org/10.1016/j.cageo.2007.05.001, 2007.

Hiemstra, P. H., Pebesma, E. J., Twenhöfel, C. J. W., and Heuvelink, G. B. M.: Real-time automatic interpolation of ambient gamma dose rates from the Dutch radioactivity monitoring network, Comput. Geosci., 35, 1711–1721, https://doi.org/10.1016/j.cageo.2008.10.011, 2009.

Jetz, W., Cavender-Bares, J., Pavlick, R., Schimel, D., Davis, F. W., Asner, G. P., Guralnick, R., Kattge, J., Latimer, A. M., Moorcroft, P., Schaepman, M. E., Schildhauer, M. P., Schneider, F. D., Schrodt, F., Stahl, U., and Ustin, S. L.: Monitoring plant functional diversity from space, Nat. Plants, 2, 16024, https://doi.org/10.1038/nplants.2016.24, 2016.

Kattge, J., Díaz, S., Lavorel, S., Prentice, I. C., Leadley, P., Bönisch, G., Garnier, E., Westoby, M., Reich, P. B., Wright, I. J., Cornelissen, J. H. C., Violle, C., Harrison, S. P., Van Bodegom, P. M., Reichstein, M., Enquist, B. J., Soudzilovskaia, N. A., Ackerly, D. D., Anand, M., Atkin, O., Bahn, M., Baker, T. R., Baldocchi, D., Bekker, R., Blanco, C. C., Blonder, B., Bond, W. J., Bradstock, R., Bunker, D. E., Casanoves, F., Cavender-Bares, J., Chambers, J. Q., Chapin Iii, F. S., Chave, J., Coomes, D., Cornwell, W. K., Craine, J. M., Dobrin, B. H., Duarte, L., Durka, W., Elser, J., Esser, G., Estiarte, M., Fagan, W. F., Fang, J., Fernández-Méndez, F., Fidelis, A., Finegan, B., Flores, O., Ford, H., Frank, D., Freschet, G. T., Fyllas, N. M., Gallagher, R. V., Green, W. A., Gutierrez, A. G., Hickler, T., Higgins, S. I., Hodgson, J. G., Jalili, A., Jansen, S., Joly, C. A., Kerkhoff, A. J., Kirkup, D., Kitajima, K., Kleyer, M., Klotz, S., Knops, J. M. H., Kramer, K., Kühn, I., Kurokawa, H., Laughlin, D., Lee, T. D., Leishman, M., Lens, F., Lenz, T., Lewis, S. L., Lloyd, J., Llusià, J., Louault, F., Ma, S., Mahecha, M. D., Manning, P., Massad, T., Medlyn, B. E., Messier, J., Moles, A. T., Müller, S. C., Nadrowski, K., Naeem, S., Niinemets, Ü., Nöllert, S., Nüske, A., Ogaya, R., Oleksyn, J., Onipchenko, V. G., Onoda, Y., Ordoñez, J., Overbeck, G., Ozinga, W. A., Patiño, S., Paula, S., Pausas, J. G., Peñuelas, J., Phillips, O. L., Pillar, V., Poorter, H., Poorter, L., Poschlod, P., Prinzing, A., Proulx, R., Rammig, A., Reinsch, S., Reu, B., Sack, L., Salgado-Negret, B., Sardans, J., Shiodera, S., Shipley, B., Siefert, A., Sosinski, E., Soussana, J.-F., Swaine, E., Swenson, N., Thompson, K., Thornton, P., Waldram, M., Weiher, E., White, M., White, S., Wright, S. J., Yguel, B., Zaehle, S., Zanne, A. E., and Wirth, C.: TRY – a global database of plant traits, Glob. Change Biol., 17, 2905–2935, https://doi.org/10.1111/j.1365-2486.2011.02451.x, 2011.

Laforest-Lapointe, I., Martínez-Vilalta, J., and Retana, J.: Intraspecific variability in functional traits matters: case study of Scots pine, Oecologia, 175, 1337–1348, https://doi.org/10.1007/s00442-014-2967-x, 2014.

Lenth, R. V.: Least-squares means: the R package lsmeans, J. Stat. Softw., 69, 1–33, https://doi.org/10.18637/jss.v069.i01, 2016.

Little, R. J. A. and Rubin, D. B.: Statistical Analysis with Missing Data, Wiley, New York, USA, 2002.

Maire, V., Wright, I. J., Prentice, I. C., Batjes, N. H., Bhaskar, R., van Bodegom, P. M., Cornwell, W. K., Ellsworth, D., Niinemets, Ü., Ordonez, A., Reich, P. B., and Santiago, L. S.: Global effects of soil and climate on leaf photosynthetic traits and rates, Global Ecol. Biogeogr., 24, 706–717, https://doi.org/10.1111/geb.12296, 2015.

McGill, B. J., Enquist, B. J., Weiher, E., and Westoby, M.: Rebuilding community ecology from functional traits, Trends Ecol. Evol., 21, 178–185, https://doi.org/10.1016/j.tree.2006.02.002, 2006.

Morris, T. P., White, I. R., and Royston, P.: Tuning multiple imputation by predictive mean matching and local residual draws, BMC

Med Res Methodol, 14, 1–13, https://doi.org/10.1186/1471-2288-14-75, 2014.

Nakagawa, S.: Missing data: mechanisms, methods, and messages, in: Ecological Statistics: Contemporary Theory and Application, edited by: Fox, G., Negrete-Yankelevich, S., and Sosa, V. J., Oxford University Press, Oxford, UK, 81–105, 2015.

Nakagawa, S. and Freckleton, R. P.: Missing inaction: the dangers of ignoring missing data, Trends Ecol. Evol., 23, 592–596, https://doi.org/10.1016/j.tree.2008.06.014, 2008.

Niinemets, Ü.: Is there a species spectrum within the world-wide leaf economics spectrum? Major variations in leaf functional traits in the Mediterranean sclerophyll Quercus ilex, New Phytol., 205, 79–96, https://doi.org/10.1111/nph.13001, 2015.

Ninyerola, M., Pons, X., Roure, J. M., Ninyerola, M., Pons, X., and Roure, J. M.: A methodological approach of climatological modelling of air temperature and precipitation through GIS techniques, Int. J. Climatol., 20, 1823–1841, https://doi.org/10.1002/1097-0088(20001130)20:14<1823::AID-JOC566>3.0.CO;2-B, 2000.

Paine, C. E. T., Baraloto, C., Chave, J., and Hérault, B.: Functional traits of individual trees reveal ecological constraints on community assembly in tropical rain forests, Oikos, 120, 720–727, https://doi.org/10.1111/j.1600-0706.2010.19110.x, 2011.

Pakeman, R. J.: Functional trait metrics are sensitive to the completeness of the species' trait data?, Methods Ecol. Evol., 5, 9–15, https://doi.org/10.1111/2041-210X.12136, 2014.

Penone, C., Davidson, A. D., Shoemaker, K. T., Di Marco, M., Rondinini, C., Brooks, T. M., Young, B. E., Graham, C. H., and Costa, G. C.: Imputation of missing data in life-history trait datasets: which approach performs the best?, Methods Ecol. Evol., 5, 961–970, https://doi.org/10.1111/2041-210X.12232, 2014.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., and R Core Team: nlme: Linear and Nonlinear Mixed Effects Models, available at: https://CRAN.R-project.org/package=nlme (last access: 1 October 2016), 2018.

Pyšek, P., Manceur, A. M., Alba, C., McGregor, K. F., Pergl, J., Štajerová, K., Chytrý, M., Danihelka, J., Kartesz, J., Klimešová, J., Lučanová, M., Moravcová, L., Nishino, M., Sádlo, J., Suda, J., Tichý, L., and Kühn, I.: Naturalization of central European plants in North America: species traits, habitats, propagule pressure, residence time, Ecology, 96, 762–774, https://doi.org/10.1890/14-1005.1, 2015.

R Core Team: R: a Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, available at: https://www.R-project.org/ (last access: 1 October 2016), 2016.

Reich, P. B.: The world-wide "fast–slow" plant economics spectrum: a traits manifesto, J. Ecol., 102, 275–301, https://doi.org/10.1111/1365-2745.12211, 2014.

Reichstein, M., Bahn, M., Mahecha, M. D., Kattge, J., and Baldocchi, D. D.: Linking plant and ecosystem functional biogeography, P. Natl. Acad. Sci. USA, 111, 13697–13702, https://doi.org/10.1073/pnas.1216065111, 2014.

Rezvan, P. H., Lee, K. J., and Simpson, J. A.: The rise of multiple imputation: a review of the reporting and implementation of the method in medical research, BMC Med. Res. Methodol., 15, 30, https://doi.org/10.1186/s12874-015-0022-1, 2015.

Sandel, B., Gutiérrez, A. G., Reich, P. B., Schrodt, F., Dickie, J., and Kattge, J.: Estimating the missing species bias in plant trait measurements, J. Veg. Sci., 26, 828–838, https://doi.org/10.1111/jvs.12292, 2015.

Schrodt, F., Kattge, J., Shan, H., Fazayeli, F., Joswig, J., Banerjee, A., Reichstein, M., Bönisch, G., Díaz, S., Dickie, J., Gillison, A., Karpatne, A., Lavorel, S., Leadley, P., Wirth, C. B., Wright, I. J., Wright, S. J., and Reich, P. B.: BHPMF – a hierarchical Bayesian approach to gap-filling and trait prediction for macroecology and functional biogeography, Global Ecol. Biogeogr., 24, 1510–1521, https://doi.org/10.1111/geb.12335, 2015.

Shipley, B., Bello, F. D., Cornelissen, J. H. C., Laliberté, E., Laughlin, D. C., and Reich, P. B.: Reinforcing loose foundation stones in trait-based plant ecology, Oecologia, 180, 923–931, https://doi.org/10.1007/s00442-016-3549-x, 2016.

Siefert, A., Violle, C., Chalmandrier, L., Albert, C. H., Taudiere, A., Fajardo, A., Aarssen, L. W., Baraloto, C., Carlucci, M. B., Cianciaruso, M. V., de L. Dantas, V., de Bello, F., Duarte, L. D. S., Fonseca, C. R., Freschet, G. T., Gaucherand, S., Gross, N., Hikosaka, K., Jackson, B., Jung, V., Kamiyama, C., Katabuchi, M., Kembel, S. W., Kichenin, E., Kraft, N. J. B., Lagerström, A., Bagousse-Pinguet, Y. L., Li, Y., Mason, N., Messier, J., Nakashizuka, T., Overton, J. M., Peltzer, D. A., Pérez-Ramos, I. M., Pillar, V. D., Prentice, H. C., Richardson, S., Sasaki, T., Schamp, B. S., Schöb, C., Shipley, B., Sundqvist, M., Sykes, M. T., Vandewalle, M., and Wardle, D. A.: A global meta-analysis of the relative extent of intraspecific trait variation in plant communities, Ecol. Lett., 18, 1406–1419, https://doi.org/10.1111/ele.12508, 2015.

Simpson, A. H., Richardson, S. J., and Laughlin, D. C.: Soil–climate interactions explain variation in foliar, stem, root and reproductive traits across temperate forests, Global Ecol. Biogeogr., 25, 964–978, https://doi.org/10.1111/geb.12457, 2016.

Stekhoven, D. J. and Bühlmann, P.: MissForest – non-parametric missing value imputation for mixed-type data, Bioinformatics, 28, 112–118, https://doi.org/10.1093/bioinformatics/btr597, 2012.

Swenson, N. G.: Phylogenetic imputation of plant functional trait databases, Ecography, 37, 105–110, https://doi.org/10.1111/j.1600-0587.2013.00528.x, 2014.

Swenson, N. G., Weiser, M. D., Mao, L., Araújo, M. B., Diniz-Filho, J. A. F., Kollmann, J., Nogués-Bravo, D., Normand, S., Rodríguez, M. A., García-Valdés, R., Valladares, F., Zavala, M. A., and Svenning, J.-C.: Phylogeny and the prediction of tree functional diversity across novel continental settings, Global Ecol. Biogeogr., 26, 553–562, https://doi.org/10.1111/geb.12559, 2017.

Taugourdeau, S., Villerd, J., Plantureux, S., Huguenin-Elie, O., and Amiaud, B.: Filling the gap in functional trait databases: use of ecological hypotheses to replace missing data, Ecol. Evol., 4, 944–958, https://doi.org/10.1002/ece3.989, 2014.

Templ, M., Alfons, A., Kowarik, A., and Prantner, B.: VIM: Visualization and Imputation of Missing Values, R package version 4.0.0, available at: http://CRAN.R-project.org/package=VIM (last access: 15 January 2015), 2013.

van Buuren, S.: Flexible Imputation of Missing Data, CRC Press, Boca Raton, USA, 2012.

van Buuren, S. and Groothuis-Oudshoorn, K.: mice: multivariate Imputation by Chained Equations in R, J. Stat. Softw.,

45, available at: http://www.jstatsoft.org/v45/i03 (last access: 9 April 2015), 2011.

Vilà-Cabrera, A., Martínez-Vilalta, J., and Retana, J.: Functional trait variation along environmental gradients in temperate and Mediterranean trees, Global Ecol. Biogeogr., 24, 1377–1389, https://doi.org/10.1111/geb.12379, 2015.

Violle, C., Navas, M.-L., Vile, D., Kazakou, E., Fortunel, C., Hummel, I., and Garnier, E.: Let the concept of trait be functional!, Oikos, 116, 882–892, https://doi.org/10.1111/j.0030-1299.2007.15559.x, 2007.

Westoby, M. and Wright, I. J.: Land–plant ecology on the basis of functional traits, Trends Ecol. Evol., 21, 261–268, https://doi.org/10.1016/j.tree.2006.02.004, 2006.

Westoby, M., Falster, D. S., Moles, A. T., Vesk, P. A., and Wright, I. J.: Plant ecological strategies: some leading dimensions of variation between species, Annu. Rev. Ecol. Syst., 33, 125–159, https://doi.org/10.1146/annurev.ecolsys.33.010802.150452, 2002.

Wright, I. J., Reich, P. B., Westoby, M., Ackerly, D. D., Baruch, Z., Bongers, F., Cavender-Bares, J., Chapin, T., Cornelissen, J. H., and Diemer, M.: The worldwide leaf economics spectrum, Nature, 428, 821–827, 2004.

Yang, Y., Zhu, Q., Peng, C., Wang, H., and Chen, H.: From plant functional types to plant functional traits A new paradigm in modelling global vegetation dynamics, Prog. Phys. Geogr., 39, 514–535, https://doi.org/10.1177/0309133315582018, 2015.

Zambrano-Bigiarini, M.: hydroGOF: goodness-of-fit functions for comparison of simulated and observed hydrological time series, available at: https://cran.r-project.org/web/packages/hydroGOF/index.html (last access: 15 July 2016), 2014.