Forest Ecology and Management xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

Forest Ecology and Management

journal homepage: www.elsevier.com/locate/foreco

Tamm Review: On the strength of evidence when comparing ecosystem functions of mixtures with monocultures

David I. Forrester [a,*], Hans Pretzsch [b]

[a] Chair of Silviculture, Faculty of Environment and Natural Resources, Freiburg University, Tennenbacherstr. 4, 79108 Freiburg, Germany
[b] Chair of Forest Growth and Yield Science, Technische Universität München, Germany

ARTICLE INFO

Article history:
Received 3 June 2015
Received in revised form 10 August 2015
Accepted 13 August 2015
Available online xxxx

Keywords:
Allometry
Biodiversity
Complementarity
Mixed-species forest
Stand density

ABSTRACT

The diversity of dominant tree species in a forest might strongly influence ecosystem functions and services, but the current evidence is not strong enough to provide general insights on when and where these diversity effects will be large or small for a given combination of species. With this goal in mind, the aim of this study is to discuss some of the factors that may need to be considered when designing studies or judging the strength of evidence provided in studies about tree-species mixing effects in forests. While the focus is on productivity, other ecosystem functions relating to light, water and nutrients are also considered. Firstly we consider the implications of stand-level spatial replication, the effects of stand density and tracking mixing effects through time in the same stand or by using chronosequences. Mixing effects at single sites (or ages) can represent significant increases in productivity while the mean mixing effect for the same mixture across a wide range of sites (or a whole rotation) can be much smaller and insignificant. The use of tree- and neighbourhood-level analyses to expand the range of treatments compared with stand-level analyses is then discussed before examining upscaling issues relating to inter- and intra-specific variability in morphology, allometry, physiology and phenology. Ignoring intra-specific variability between individuals in monocultures and mixed-species stands when upscaling to the stand level can strongly distort mixing effects, resulting in very misleading conclusions. The difference between correlations and causality is then discussed using the production ecology equation and mass balance approaches. We also discuss some of the methodological considerations when calculating mixing effects. All of these factors can have significant implications for the calculation and interpretation of mixing effects in forests.

Contents

---

* Corresponding author.
E-mail address: david.forrester@waldbau.uni-freiburg.de (D.I. Forrester).

http://dx.doi.org/10.1016/j.foreco.2015.08.016

Please cite this article in press as: Forrester, D.I., Pretzsch, H. Tamm Review: On the strength of evidence when comparing ecosystem functions of mixtures with monocultures. Forest Ecol. Manage. (2015), http://dx.doi.org/10.1016/j.foreco.2015.08.016

1. Introduction

Forest growth depends not only on broad-scale drivers of climate and soil fertility, but also on species composition. Monocultures of different species often differ in productivity on a given site and interactions between species in mixtures can also influence forest dynamics. Most of the forests of the world are mixtures, so the overall patterns and processes of ecosystem functions and services in relation to species composition are fundamental in forest ecology and management.

Many of the processes and species interactions occurring in mixtures have been reviewed (Binkley, 1992; Kelty, 1992; Forrester et al., 2006b; Richards et al., 2010) and conceptual models have been developed to generalise when and where a given species may perform better or worse in mixtures than monocultures (Forrester, 2014). The development of such concepts requires many studies that are carefully designed, analysed and interpreted. Inappropriate assumptions made when designing or analysing such studies lead to biased calculations of mixing effects, which hamper progress to develop our understanding about these effects. This study combines insights from case studies with general approaches that are powerful tools for quantifying mixing effects.

The factors presented in this study are generally not new but they are sometimes forgotten or ignored. It is also important to note that many of the factors mentioned are only relevant under certain circumstances. Therefore it is important to know which factors, and under which circumstances, there could be significant consequences for mixing effect calculations when these factors are ignored, and these circumstances are also discussed.

2. Definitions and calculations

Since its beginnings, forest science has developed standards for the evaluation of experiments, standardized variables for reporting mean tree and stand characteristics, and standards for the transparency and documentation of calculation procedures (Pretzsch, 2009). However, such standards have focused on the analysis of monocultures, in terms of spacing, thinning and fertiliser application experiments. Comparable standards for the calculation and analysis of mixing effects have proven more difficult to achieve, not least due to the complexity of stand structural effects on estimating mixing proportions as well as definitions and algorithms for dominant height, site indices, stand density in mixed stands and approaches for upscaling from stem volume to tree mass.

This relatively slow development of standards contrasts with the long history of studies about mixing effects in forests. For example, studies by Schwappach (1909), Wimmenauer (1914), Dietrich (1928), Hofmann (1923), and Flury (1926, 1931) provided basic growth and yield data and highlighted the divergence of growth curves at the tree and stand level by mixing. They also showed a stabilising effect on productivity and stand structure in the event of disturbances. However, their comparisons with monocultures were questionable because they were based on yield tables and not neighbouring monocultures with equivalent site conditions. Given this long history, and the many questions that have been examined in relation to mixtures, it is not surprising that a wide range of approaches has been developed and used to calculate mixing effects and each calculation can result in a different mixing effect from a given data set. This section provides definitions of different levels of analyses, different sources of data, and different types of species interactions, as well as definitions and calculations of mixing effects and stand density, all of which will be referred to throughout this paper.

This study does not aim to review experimental designs or discuss statistical analyses. This has been the focus of many previous studies that present the advantages and disadvantages of designs such as replacement series, additive series, biodiversity experiments and many others (Vandermeer, 1989; Kelty and Cameron, 1995; Scherer-Lorenzen et al., 2005; Bruelheide et al., 2014) and in studies that discuss statistical analyses in ecology (e.g. Zuur et al., 2010).

2.1. Level of analysis – tree, neighbourhood, species, total stand and community

The consideration of different levels in a given study enables an examination of how changes at one level influences patterns at another level (Forrester, 2015; Pretzsch et al., 2015a). For example, a large change in the leaf-level physiology of a species in mixture compared with its monoculture may or may not result in a large change in growth or other functions at the stand level, depending on how other processes change, such as carbon partitioning. Combining tree- and stand-level analyses helps to determine which mixing effects are most important for forest functioning. They can also be used to indicate potential sources of error when comparing measurements of a given processes at different levels and when scaling up or down between each level (Pretzsch et al., 2015a). It is therefore important to define the main levels that are referred to in this study.

Tree-level analyses are those that examine individual trees e.g. when regression is used to examine whether the relationship between tree diameter and height varies between treatments. *Neighbourhood-level* analyses are a type of tree-level analysis that account for the characteristics of the trees' neighbourhood (e.g. in terms of basal area, species composition; Boyden et al., 2005; Vanclay, 2006a; Forrester et al., 2011; von Oheimb et al., 2011). This contrasts with typical tree-level analyses where the characteristics of the trees' neighbourhood are ignored or only considered in terms of the stand-level treatment, such that all trees within the plot have the same (mean plot) neighbourhood.

Stand-level analyses consider totals and means of all trees within the plot, such as total basal area (BA_T) or mean tree diameter. Stand-level analyses include *species-level* and *total stand-level* analyses. For species-level analyses the total stand is simply divided by species to provide the totals and means for each species within the stand. For example, in a two-species mixture, $BA_T = BA_{species1} + BA_{species2}$; a total stand-level analysis would consider BA_T, while a species-level analysis would consider $BA_{species1}$ or $BA_{species2}$. Total stand-level analyses are also sometimes referred to as *community-level* analyses because they consider the totals or means of the whole community. Many other levels exist, including finer scales such as leaf-level and organ-level (e.g. branches, roots) or coarser scales such as landscape-level, but these are not the focus of this synthesis.

Please cite this article in press as: Forrester, D.I., Pretzsch, H. Tamm Review: On the strength of evidence when comparing ecosystem functions of mixtures with monocultures. Forest Ecol. Manage. (2015), http://dx.doi.org/10.1016/j.foreco.2015.08.016

2.2. Empirical data sources – planted experiments, specific forest plots or inventory

Most empirical data used to examine patterns and processes in mixed-species forests come from three main sources; planted experiments, specific forest plots or inventories. The characteristics of these data sources influence analyses and their interpretation. With *planted experiments* all plots can be randomly located and it is relatively easy to control factors such as species compositions, initial stand density, ages and silvicultural treatments, all under a given set of soil and climatic conditions. However, planted experiments are expensive to establish and to maintain and it takes a long time for them to develop large trees. *Inventories* do not control for any of these factors or stand disturbances but potentially deliver the whole range in each of these factors within the inventoried forests. They are typically designed to provide unbiased data by systematically locating plots (e.g. based on grids or transects).

We refer to the third source of empirical data as *Specific Forest Plots* (SFPs). These are everything between planted experiments and inventories in terms of the control they offer regarding species composition, stand density, silviculture and site characteristics. SFPs differ from inventories because they are established in forests with specific species compositions, ages, densities, silviculture and site characteristics, rather than to achieve the (often much) wider range provided by inventories. This systematic selection of SFPs allows for more control of the stand conditions. However, they differ from planted experiments because many potentially important aspects will not have been controlled, at least for part of the stand development. For example, when SFPs are established in mature forests, the past stand conditions (e.g. silviculture, age structure) and the soil or climatic conditions may differ between plots (even of the same treatment) in ways that are unknown or that are difficult to control or account for without additional soil analyses or historic records. That is, it is often not possible to randomly locate the plots, even though they can be randomly selected. There can be many variations of SFPs that might be classified in terms of whether different treatments have been applied specifically for the study, or the plots were selected based on already existing site, age, composition or structural characteristics, or whether they are permanent sampling plots, etc. (Zhao et al., 2014) but these distinctions are less relevant to the mixing effects discussed in this synthesis. Generally, the number of variables measured and the intensity of measurement is highest in planted experiments and lowest in inventories (Baeten et al., 2013).

2.3. Types of interactions

Many terms have been used to describe different types of species interactions and their outcomes in forests. The types of interactions are often divided into three groups; facilitation, competitive reduction and competition (Vandermeer, 1989). *Competition* occurs when interactions between at least two plants (or species or populations) result in at least one exerting a negative effect on the growth or survival of the other. *Competitive reduction* occurs when inter-specific competition is less intense than intra-specific competition, usually due to an inter-specific differentiation in resource use (Vandermeer, 1989). This may also be referred to as the competitive production principle. Examples of competitive reduction are when there is a spatial, temporal or chemical stratification in the uptake of resources. *Facilitation* occurs when at least one of the species in a mixture positively influences the growth or survival of another (Vandermeer, 1989). Examples of facilitation are when the availability of a resource is increased such as by symbiotic nitrogen fixation or hydraulic redistribution of water. In reality, the contribution of facilitation, competitive reduction and competition to measured growth responses are very difficult to separate and are often collectively referred to as *complementarity*, which can be positive or negative depending on the balance of competition, competitive reduction and facilitation.

2.4. Mixing effects

The term "mixing effect" is used very broadly in this paper to describe any difference in the performance of mixtures compared with monocultures or the performance of a given species in a mixture compared to its monoculture, but there are several more specific terms that are also used. The outcomes of species interactions are often described in terms of overyielding, underyielding and transgressive overyielding. *Overyielding* occurs when the total mixture, or a species within the mixture, produces more than expected according to the monocultures. *Underyielding* occurs when the total mixture, or a given species, produces less than

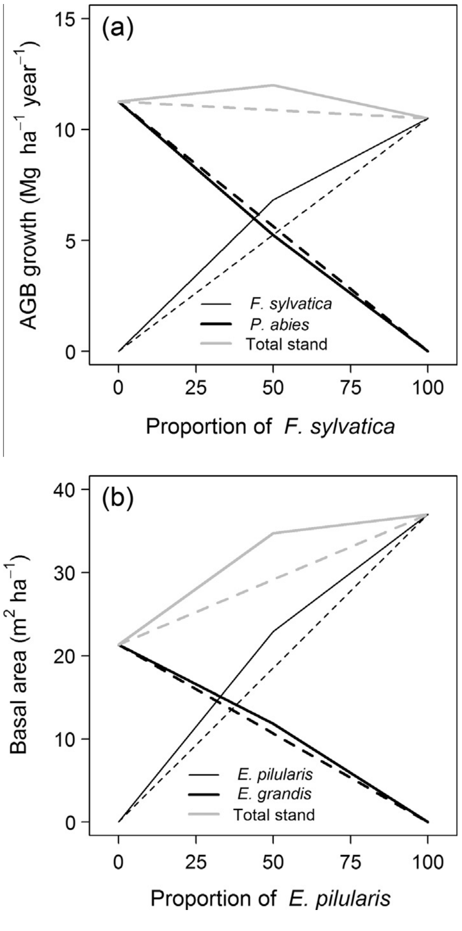

Fig. 1. The basal area or above-ground biomass (AGB) growth of monocultures and two-species mixtures in relation to species proportions. The solid lines indicate the measured responses and the dashed lines indicate the growth or yield that would be expected if there were no complementarity effects and each species grew exactly the same in mixture as in monoculture. In (a) *F. sylvatica* and the total mixture overyielded, but *P. abies* underyielded. The RP calculations (Eqs. (1) and (2)) in the 1:1 mixtures were 1.30 (*F. sylvatica*), 0.93 (*P. abies*) and 1.10 (total mixture), and this mixture showed transgressive overyielding. In (b) both species and the total mixture overyielded, with RP calculations in 1:1 mixtures of 1.24 (*E. pilularis*), 1.11 (*E. grandis*) and 1.19 (total stand), but the mixture did not show transgressive overyielding even though the RP of the mixture was much greater than that in (a). A total mixture RP calculation, of at least 1.27, would be required for transgressive overyielding to occur in (b) because the productivity difference between monocultures of *E. pilularis* and *E. grandis* is so large (73%) compared with the difference between monocultures of *F. sylvatica* and *P. abies* (only 7%). (a) modified from Pretzsch et al. (2010), where the site index = 28 m for *F. sylvatica* and 46 m for *P. abies* and (b) is modified from Forrester and Smith (2012).

Please cite this article in press as: Forrester, D.I., Pretzsch, H. Tamm Review: On the strength of evidence when comparing ecosystem functions of mixtures with monocultures. Forest Ecol. Manage. (2015), http://dx.doi.org/10.1016/j.foreco.2015.08.016

expected. These are often calculated using Relative Productivity (RP) equations (Kennel, 1965; Pretzsch et al., 2013a; Forrester, 2014) and can be applied to growth, yield or any other ecosystem functions and services. The RP equations vary depending on whether the RP is being calculated for the total stand (Eq. (1)), or for a given species using either stand totals of a given variable (Eq. (2)) or means of a given variable (Eq. (3)).

Eqs. 1–3 are shown for two-species mixtures (species 1 and 2) but can be applied to mixtures containing any number of species. The RP equation for the total stand calculates how much the mixed-species stand produces ($p_{1,2}$) in relation to the productivity that would be expected if there was no difference between the effect of inter- and intra-specific interactions. The expected productivity $\widehat{p}_{1,2}$ is calculated from that of neighbouring monospecific stands (p_1 and p_2) and the mixing proportions (m_1 and m_2) such that $\widehat{p}_{1,2} = m_1p_1 + m_2p_2$. Special consideration of stand density may be required when calculating m_1 and m_2, as discussed in Section 2.5.

$$\mathrm{RP}_{\text{total stand}} = \frac{p_{1,2}}{m_1p_1 + m_2p_2} \tag{1}$$

The RP of a given species is calculated with Eq. (2) when using totals of a given variable (e.g. stand biomass of species 1), but is calculated with Eq. (3) when using means of a given variable, such as mean tree diameter, or mean tree volume growth (Forrester, 2014).

$$\mathrm{RP}_{\text{species}} = \frac{p_{1,(2)}}{m_1p_1} \tag{2}$$

$$\mathrm{RP}_{\text{species}} = \frac{p_{1,(2)} - p_1}{p_1} + 1 \tag{3}$$

In Eqs. (2) and (3), $p_{1,(2)}$ is the total (for Eq. (2)) or mean (for Eq. (3)) productivity, or any other measure of performance, of species 1 in a mixture with species 2.

In Eqs. 1–3, when RP = 1 the performance of the mixtures is exactly as expected based on the monocultures. This is sometimes referred to as an *additive effect* and, in the case of $\mathrm{RP}_{\text{total stand}}$, indicates that there is currently either no complementarity effect or one species is performing better at the expense of another. In contrast, RP > 1 indicates overyielding and RP < 1 indicates underyielding. These indicate that there is currently a complementarity effect and are sometimes referred to as *non-linear* or *multiplicative* effects. Fig. 1 shows RP calculations for two-species mixtures. The RP minus 1 represents a proportional mixing effect, and is the vertical difference between the solid line and the expected (dashed) lines shown in Fig. 1.

Transgressive overyielding is the special case of overyielding where the mixture produces more than the most productive of all the monocultures. This often occurs when the monocultures of each species have similar productivities and there are facilitative and/or competitive reduction interactions that improve the growth of each species (Forrester et al., 2006a,b). The greater the difference in productivity between each monoculture, the greater the complementarity effects need to be for the mixture to outperform the most productive monoculture (Fig. 1). Transgressive overyielding is only applicable to the total stand-level. In contrast, over- and underyielding can apply at both the species-stand level and the total-stand level.

Most of the case studies and calculations described above were applied to two-species mixtures, although they can usually also be applied to > two-species mixtures and many studies also examine how mixing effects change as the diversity of the stands or neighbourhoods increase. These are broadly referred to here as *diversity* or *richness effects*. Diversity may be quantified in many different ways, such as in terms of tree species, functional types, phylogenetic diversity, and structural diversity. These variables are usually quantified as either counts, in which case they are richness effects (e.g. tree species richness) or indices, in which case they are diversity effects (e.g. tree species diversity, structural diversity). An example of a commonly used index is Shannon's index (Shannon, 1948), which considers the number of species (or functional types, etc.) and the mixing proportions. A special case of a diversity effect is an *identity effect*. This is where the explanatory variable is a measure of the contribution that a single species makes to the mixture. This could be a binary variable of presence/absence or perhaps the proportion of the mixture (in terms of trees per ha, basal area, etc.) that is contributed by the species (or functional type, etc.) of interest (Ratcliffe et al., 2015); thus any study containing only two species could also be considered to be showing identity effects. While diversity, richness and identity analyses can be performed at any level (tree, species, total stand) it is worth mentioning the additive partitioning calculations of Loreau and Hector (2001) that are designed for total-stand (e.g. community) level analyses. These calculations separate complementarity effects from selection effects. Selection effects are similar to identity effects and occur when a species dominants a community and influences the relative yield of the other species.

In addition to the commonly used RP calculations and additive partitioning equations, many other mathematical calculations have been developed to quantify, interpret and display overyielding and other aspects of interactions between plants. A review by Weigelt and Jolliffe (2003) included more than 50 of these calculations. This great variety has resulted from the many different experimental designs and research questions for which the calculations were specifically designed. Consequently, their selection has significant consequences for the way species interactions are quantified and therefore the interpretations and conclusions, and this has resulted in critical debates about which types of calculations are appropriate (e.g. Brooker et al., 2013). As a result many studies avoid or complement the use of such mathematical indices with other types of analyses, such as regression, when quantifying mixing effects (Sections 4 and 5).

Weigelt and Jolliffe (2003) suggested that it is worth considering the specificity and clarity of their meaning, mathematical properties (e.g. complexity in terms of numbers of variables and calculations to get the index), how stand density is considered, and freedom from bias in terms of initial plant sizes. To illustrate, the implications of using different types of calculations, the total stand overyielding and underyielding calculated using the RP equation (1) is compared with another calculation, Land Equivalent Ratio, LER, (Willey and Osiru, 1972) as shown in Table 1. The Land Equivalent Ratio is calculated using Eq. (4).

$$\text{Land Equivalent Ratio} = \frac{p_{1(2)}}{p_1} + \frac{p_{(1)2}}{p_2} \tag{4}$$

The $p_{1(2)}$ is the productivity of species 1 in a mixture with species 2, and p_1 is the productivity of species 1 in monoculture, and similarly for species 2. The Land Equivalent Ratio is a measure of the land area required for a mixture to be as productive as the same species grown in monocultures. Values >1 indicate a productivity advantage in mixture while <1 indicates the opposite. The Land Equivalent Ratio does not explicitly consider the mixing portions and is based on the two relative productivities. It is more common in herbaceous plant biology and agronomy where mixing proportions in terms of required growing space are difficult to measure and are not used when quantifying mixing effects (Harper, 1977; Vandermeer, 1989). In contrast, the RP equation is often used for forest stands, where the species-specific stand areas and mixing proportions can easily be quantified (Pretzsch, 2009, pp

Please cite this article in press as: Forrester, D.I., Pretzsch, H. Tamm Review: On the strength of evidence when comparing ecosystem functions of mixtures with monocultures. Forest Ecol. Manage. (2015), http://dx.doi.org/10.1016/j.foreco.2015.08.016

Table 1
Eight model triplets used to explain the ambiguity of the Land Equivalent Ratio compared with the Relative Productivity equation. A triplet is a set of three plots including the 2-species mixture and the monoculture of each species. When the productivity or mixing proportions of the two mixed species are similar (triplets 1 and 2) their Relative Productivity and Land Equivalent Ratio values are similar. The greater the difference in species-specific productivity and mixing portions, the more misleading the Land Equivalent Ratio values can be compared with the Relative Productivity equation. The mixed stand of triplet 3 considerably overyields the monospecific stands (Relative Productivity = 1.289) and triplet 4 represents underyielding (Relative Productivity = 0.945), but Land Equivalent Ratio = 1.1 for both stands. The difference becomes even worse for triplets 5, 6 and 8.

Model triplet	Monoculture growth species 1 (Mg ha^{-1} yr^{-1})	Monoculture growth species 2 (Mg ha^{-1} yr^{-1})	Mixing proportion of species 1	Species 1 growth in mixture (Mg ha^{-1} yr^{-1})	Mixing proportion of species 2	Species 2 growth in mixture (Mg ha^{-1} yr^{-1})	Relative Productivity	Land Equivalent Ratio
1	10.0	10.0	0.25	5.0	0.75	6.0	1.100	1.100
2	7.0	8.0	0.50	6.0	0.50	7.0	1.733	1.732
3	15.0	10.0	0.25	10.5	0.75	4.0	1.289	1.100
4	10.0	15.0	0.25	7.0	0.75	6.0	0.945	1.100
5	15.0	7.5	0.75	6.0	0.25	4.5	0.800	1.000
6	7.5	15.0	0.75	3.0	0.25	9.0	1.280	1.000
7	15.0	10.0	0.60	4.5	0.40	6.0	0.808	0.900
8	5.0	15.0	0.60	1.5	0.40	9.0	1.167	0.900

356-360). The differences in conclusions that can result from using Land Equivalent Ratio compared with the Relative Productivity equation are shown in Table 1.

2.5. Quantification of mixing proportions using stand density

It is usually necessary to quantify mixing proportions in order to calculate the mixing effects. However, calculating mixing proportions is not necessarily as straightforward as it sounds (Sterba et al., 2014). There are many ways to calculate mixing proportions depending on the objectives of the study (Bravo-Oviedo et al., 2014). For example, mixing proportions could be based on the number of trees, basal area, volume, biomass, biomass growth, crown areas, leaf areas, etc., and all are likely to give different proportions within a given mixture depending on how much the allometry of each species differs. For example, in a two-species mixture, the mixing proportions of species 1 (m_1) and species 2 (m_2) could be calculated from the number of trees per ha (N) using Eq. (5).

$$m_1(N) = \frac{N_1}{N_1 + N_2} \tag{5}$$

Proportions in terms of other variables could be calculated analogously. These calculations are often used in planted experiments where the starting conditions are controlled. The subsequent development of stand density and stand structure and differences between treatments or species in terms of tree size are part of the treatments. This contrasts with studies using plots in forests where the differences in species ages or sizes and stand density at the start of the study may have been caused by previous management activities unrelated to the treatments. In these cases in particular, Eq. (5) may not be adequate because it does not account for any initial species-specific differences in mean tree size or growing space (e.g. resource) requirements. For example, species 1 may occur as a large number of small trees, while species 2 may occur as a small number of large tall trees. In this case, the mixture proportion $m_1(N)$ would overestimate species 1's share of the stand area and its likely potential to obtain resources. In contrast, mixture proportions based on basal area or volume, take account of species-specific sizes. However, the stand density may still need to be accounted for in the calculation of mixing proportions because the resource requirements (growing space) for trees of a given size can vary significantly between species due to differences in physiology and allometry. In other words, the maximum stand density that each species can achieve in monocultures at the given site may vary.

Several studies based on inventory data have used Reineke's (1933) stand density index (SDI) or basal area to account for inter-specific differences in growing space occupancy (using SDI or basal area) and tree size (using SDI) when calculating mixing proportions (Río and Sterba, 2009; Condés et al., 2013; Sterba et al., 2014). Reineke's (1933) SDI is based on the allometric relationship between N and quadratic mean diameter (d_q) described by Eq. (6).

$$\mathrm{SDI} = N \times \left(\frac{25}{d_q}\right)^{-1.605} \tag{6}$$

If it is assumed that there is a maximum or potential SDI that can be achieved by monocultures of a given species at a site ($\mathrm{SDI}_{\max}$) then this can be used to calculate mixing proportions. Eq. (7) was used by Sterba et al. (2014) to adjust mixing proportions that account for inter-specific differences in carrying capacity at a given site ($m_{1\mathrm{adj}}$, $m_{2\mathrm{adj}}$).

$$m_{1\mathrm{adj}} = \frac{\frac{\mathrm{SDI}_1}{\mathrm{SDI}_{\max1}}}{\frac{\mathrm{SDI}_1}{\mathrm{SDI}_{\max1}} + \frac{\mathrm{SDI}_2}{\mathrm{SDI}_{\max2}}} \tag{7}$$

If the $\mathrm{SDI}_{\max}$ of each species is the same then the mixing proportion calculated using Eq. (5) (with SDI) will be the same as that using Eq. (7). The greater the differences between the two maximum SDI the greater the differences in mixing proportions calculated using Eqs. (5) and (7) (Sterba et al., 2014). For example, Sterba et al. (2014) found that maximum densities of *Pinus sylvestris* and *Fagus sylvatica* did not differ much and analyses required only minor adjustments in mixing proportions. However, there were much larger differences in the maximum densities of *P. sylvestris* and *Quercus pyrenaica*. This resulted in significant adjustments in mixing proportions, which if ignored, would have resulted in a distinct overestimation of the mixing effects.

Table 2
A scheme for rating the confidence in forestry studies modified from Binkley and Menyailo (2005).

Level	Type of evidence
1	Meta-analysis of many similar experiments showing consistent effects
2a	Experiment replicated at several sites in accordance with the population of interest
2b	Experiment replicated at several sites but without formal *a priori* plan for extrapolating to the population
2c	Simulation studies based on models parameterized and validated with data from mixtures
2d	Evaluation of forest inventories
3a	Experiment at a single site with replication
3b	Case studies across environmental gradients with potentially confounding spatial factors
4	Case studies at a single site with no replication
5	Expert opinion or inferences from expected first principles

Please cite this article in press as: Forrester, D.I., Pretzsch, H. Tamm Review: On the strength of evidence when comparing ecosystem functions of mixtures with monocultures. Forest Ecol. Manage. (2015), http://dx.doi.org/10.1016/j.foreco.2015.08.016

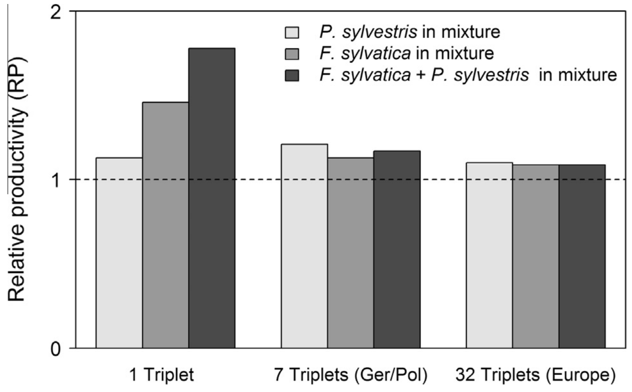

Fig. 2. Overyielding (Relative Productivity > 1) in terms of volume growth per hectare of *Pinus sylvestris* and *Fagus sylvatica* in mixed-species stands compared with neighbouring monospecific stands at the species level and the total stand level. Evaluation based on a long-term experiment (Pretzsch et al., 2013b), seven triplets in Germany and Poland (Pretzsch et al., 2015b), and 32 triplets along a productivity gradient through Europe (Pretzsch et al., 2015b) showed a decreasing mean effect size (from left to right). The Relative Productivity is calculated using Eqs. (1) and (2) and based on mixing proportions calculated using Eq. (7).

This approach enables analyses using inventory data to take stand density into account; however the approach used above should be used cautiously for three reasons. Firstly, the exponent 1.605 in Eq. (6) is not a constant and changes with environmental conditions and even in mixed compared with monospecific stands (Binkley, 1984; Woodall et al., 2005; Weiskittel et al., 2009; Reyes-Hernandez et al., 2013). Secondly, SDI_{max} may not be a static variable and could change with long- and short-term changes in climatic conditions at a site (e.g. during a drought), which could have implications for using this approach to study the effects of these factors in mixtures. Thirdly, the use of SDI_{max} in Eq. (7) does not account for the effects of vertical stratification or other complex stand structural influences on mixing effects e.g. two mixed-species stands containing the same pair of species, with the same SDI_1 and SDI_2, on the same site, could still perform differently if species 1 overtops species 2 in one of the mixtures, but not in the second mixture. See also Section 3.3.

3. Replication at the stand level

The replication within a study determines the range of conditions that the results apply to, so it is critical to consider where the replication should be when designing a study. On the other hand, the replication is also an important consideration when comparing and synthesising the results from different studies. Binkley and Menyailo (2005) provided a simple scheme to rate the strength of evidence from forest studies examining species–soil interactions (Table 2). The same scheme can be applied when considering mixing effects presented in different studies. Many studies are at levels 3 and 4 while far fewer are at the much more general levels of 1 and 2.

3.1. Spatial replication

The population of interest needs to be clearly defined (Binkley and Menyailo, 2005; Binkley, 2008), and the results should not need to be extrapolated in order to describe that population. This is illustrated using Fig. 2, which compares the volume growth of *P. sylvestris* and *F. sylvatica* in mixtures and monocultures. This set of three plots (one mixture and two monocultures) will be referred to as a “triplet”, but it could easily be extended to include >2 species. When the comparison is made at a single site (Fig. 2, left), the Relative Productivity calculation of the mixture is about 1.78, so their productivity is about 78% greater than expected based on the monocultures. Initial reports of the mixing effects for a given species combination might tend to be on the sites where the mixing effects were large. However, mixing effects can vary from site to site (Forrester, 2014), and the soil and climatic conditions might just happen to favour mixtures, such as Fig. 2-left. These results cannot be extrapolated to any other soil or climatic conditions, even when the triplet is replicated many times at this site, which would be like sub-sampling (not sampling) with respect to the German or European populations of *P. sylvestris* and *F. sylvatica* (Binkley and Menyailo, 2005; Binkley, 2008). The middle of Fig. 2 shows the same comparison of treatments using a series of seven triplets replicated at different sites within Germany and Poland. The mean mixing effect across these sites is only about 17% ($RP_{total\ stand} = 1.17$), and this result applies to the range of soils and climates that existed at these German and Polish sites. Extending this replication to 32 triplets across Europe, each at a different site, results in a mixing effect that is only about 9% ($RP_{total\ stand} = 1.09$), but this result applies to a large proportion of the natural distribution of this species combination. The $RP_{total\ stand}$ calculations, based on volume growth, across this European transect varied from about 0.5 to 2.5, but this variability was only weakly, and not significantly, correlated with site or climatic indices (Pretzsch et al., 2015b). Thus, Fig. 2 illustrates three important points. Firstly, that increasing the sample size leads to a gradual convergence on the population mean. Secondly, that biased samples (Fig. 2-left) can lead to misleading population inferences, and thirdly, that small samples can frustrate calculations of population mixing effects.

This use of triplets on different sites is similar to the “twin plot” approach that was used to examine fertiliser responses at 127 locations (each with a fertilised and unfertilised plot) in *Eucalyptus* plantations in Brazil (Stape et al., 2006). However, an important distinction is that in these case studies, the twin plot approach was applied to planted experiments whereas the triplet approach was applied to Specific Forest Plots, where many potentially important factors have not been controlled within a given triplet, such as site conditions, previous management, the age of each species etc (see Section 2.2).

It is also critical to consider the positioning of treatments within a site. At a given site that contains only a single replicate (triplet), even a random assignment of treatments could place the mixture on more fertile soil than the monocultures; the treatments are confounded by all other factors that vary between plots, such as the soil fertility, moisture availability, and aspect. This applies to all data sources (planted experiments, inventories and SFPs). To account for these confounding factors treatments are typically replicated within a given site. Using replicated blocks of plots (or a minimum distance between plots of the same treatment) reduces or removes the likelihood that all plots of a given treatment will by chance be placed close together or on more fertile positions than the other treatments (Bruelheide et al., 2014). Interestingly, in a planted tree-species diversity experiment in Panama, even random positioning of treatments resulted in a chance positioning of treatments with high diversities on more fertile positions than less diverse treatments, and this needs to be accounted for in analyses to avoid overestimating the true biodiversity effect (Healy et al., 2008; Bruelheide et al., 2014).

The triplet approach shown in Fig. 2 has often been applied to Specific Forest Plots and inventory data (Pretzsch and Schütze, 2009). In contrast to planting new experiments, inventory data and Specific Forest Plots enable the study of interactions between large mature trees and quickly provide urgently needed basic knowledge about mixing effects across whole landscapes. However, with Specific Forest Plots or inventory data it can be

Please cite this article in press as: Forrester, D.I., Pretzsch, H. Tamm Review: On the strength of evidence when comparing ecosystem functions of mixtures with monocultures. Forest Ecol. Manage. (2015), http://dx.doi.org/10.1016/j.foreco.2015.08.016

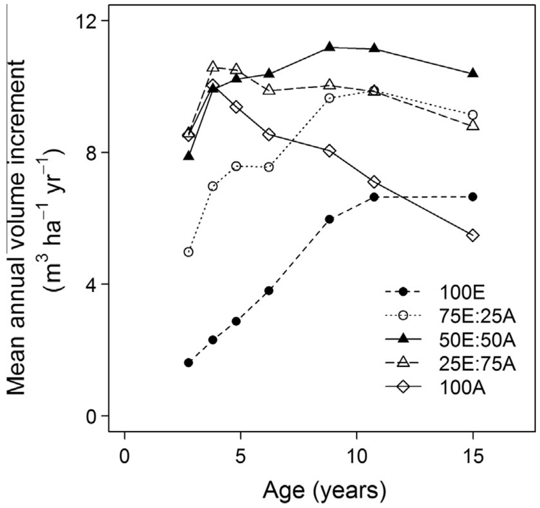

Fig. 3. The mean annual increment of mixtures and monocultures of *Eucalyptus globulus* and *Acacia mearnsii* planted at a spacing of 2 m × 3.3 m (Forrester et al., 2011). In the legend, the number represents the percent of planted trees per ha of *E. globulus* (E) and *A. mearnsii* (A).

difficult to ensure that each of the treatments experiences the same disturbance, soil and microclimatic conditions within a given location and that the treatments (mixtures or monocultures) are not themselves the result of different disturbances, soil or microclimatic conditions; this is much easier to control in planted experiments. Note that the "same soil and microclimatic conditions" are referred to loosely here because these conditions are well known to vary across very short distances (within metres) within single plots (Schume et al., 2004; Boyden et al., 2012). The comparability of growing conditions for each plot within a given location may need to be estimated using information that has not been influenced by the species, such as aspect, slopes and some soil properties. Without such confirmation, there will be less confidence in the results.

Previous silvicultural practices may also vary between plots included in SFPs and inventories, creating contrasting stand structures that are not the result of species interactions but that influence species interactions and tree growth. In contrast, planted experiments are typically more uniform and may lack some of these complex structural effects. The effects of different structures can possibly be accounted for using tree- and neighbourhood-level analyses (Section 4).

The quantification of soil, climatic and other plot or site characteristics is also important when quantifying the site gradients within the experiments. It is well known that the way a given combination of species interacts in forests changes with resource availability and climatic conditions (Forrester, 2014). To examine these patterns it is important to quantify which site variables change along the gradient and how much they change. Using general information such as site indices can make interpretation difficult because it does not show which resources or climatic conditions were responsible for the change in site indices and it is these characteristics that influence the species interactions, not the site index *per se* (Forrester, 2014). Alternatively, experiments could be established at a single site where a single resource is manipulated for different triplets. For example, water retention-irrigation experiments have been used to examine the effects of water availability on growth, carbon partitioning, light absorption, transpiration and other processes for entire rotations in *Eucalyptus* plantations (Stape et al., 2010), and similar experiments have recently begun in mixed-species stands (Pretzsch et al., 2014).

3.2. Temporal replication

In addition to spatial replication, consideration of the temporal range in measurements is important. Repeated measurements of the same plots, or chronosequences, provide information about how the species interactions change as stands develop, and also how the interactions change as climatic conditions change. For example, Fig. 3 shows the mean annual volume increment in monocultures and mixtures of *Eucalyptus globulus* and *Acacia mearnsii*. Initially *A. mearnsii* monocultures were the most productive and peaked very early at only about age 3 years. The peak in the mean annual increment of *E. globulus* occurred much later at about 10–15 years. The mixtures peaked at intermediate ages depending on the proportion of each species within the mixture and their peak mean annual increments were often higher than both of the monocultures. As a result of these temporal dynamics, measurements made at age 3 years would have shown that the *A. mearnsii* monoculture had a similar or higher productivity than the other treatments. However, measurements at age 15 years would have shown that *A. mearnsii* monocultures were the least productive treatment. This illustrates the value of data sets that span an entire rotation.

An alternative approach to long-term measurements is chronosequences that include plots of different ages on the same (or very similar) sites to avoid introducing confounding effects of site conditions. The disadvantage is that it can be hard to ensure that there are no differences in disturbances, soils and microclimates between plots of different ages that will confound the age effects. Assuming that there are no such confounding factors, an advantage of chronosequences is that all ages and development stages can be measured under the same climatic conditions (e.g. Forrester et al., 2010a). Conversely, the results will only be applicable to the climatic conditions that prevailed while the chronosequence was measured. With repeated measurements of the same plots there is a risk that the climatic conditions varied as the stand developed so that the age trend is distorted due to the occurrence of several years with favourable climates or several years of drought. This is more problematic in stands with short rotations (e.g. <20 years).

3.3. Stand density and structure

Stand density, in terms of variables such as basal area, volume and biomass is a strong determinant of stand growth, and may often be a more important determinant of productivity than species interactions and biodiversity (Vilà et al., 2013). Changes in stand density often modify the way species interact in terms of growth (Garber and Maguire, 2004; Boyden et al., 2005; Río and Sterba, 2009; Condés et al., 2013; Forrester et al., 2013), light absorption (Forrester and Albrecht, 2014) and transpiration (Forrester, 2015). Therefore, it is often necessary to take stand density into account when interpreting mixing effects. In acknowledgement of the interaction between species interactions and stand density, many experimental designs have been developed (e.g. additive series and replacement series) to examine how density influences the effects of intra- and inter-specific interactions and evenness of species proportions (Kelty and Cameron, 1995; Vanclay, 2006b). The potential effects of density are particularly important for foresters because modifying density by thinning is a major forest management tool used to control growth, mortality, water availability, pests, etc. (Allen et al., 2010; Hawthorne et al., 2013).

The manipulation of density either when planting or thinning also influences the structure of the stand, e.g. in terms of the diameter and height distributions or canopy stratification. Differences in structure between mixtures and monocultures (e.g. canopy

Please cite this article in press as: Forrester, D.I., Pretzsch, H. Tamm Review: On the strength of evidence when comparing ecosystem functions of mixtures with monocultures. Forest Ecol. Manage. (2015), http://dx.doi.org/10.1016/j.foreco.2015.08.016

structure, crown architecture) have been proposed as contributors to mixing effects (Kelty, 1992; Bauhus et al., 2004; Sapijanskas et al., 2014). It is therefore important to control for differences in structure as well as density. For example, the mixing effect resulting from natural canopy stratification might be very different to the mixing effect that occurs in a stand thinned to the same density but that develops a very different vertical structure.

When mixtures are more productive than monocultures, they will by definition develop a greater stand density and this difference is part of the treatment. Similarly, the vertical structure that develops is also part of the treatment. Therefore, after establishing planted experiments, the density is usually not manipulated again, or when it is it will be done systematically and acknowledged how the treatments themselves have been changed.

In contrast to planted experiments, differences in the stand density and stand structure of inventory plots and SFPs can often vary due to differences in past silvicultural practices or natural disturbances and not only due to the natural stand dynamics. One approach to reduce (but not avoid) the confounding effect of silviculture or natural disturbances is to select plots that are known not to have not been thinned or plots that have been thinned but have since reached a stand density close to the sites potential (e.g. Pretzsch et al., 2010). A second approach that has been applied to inventory data is to use stand density indices (Section 2.5), however, this approach does not account for vertical structure and requires assumptions described in Section 2.5. A third approach is to use tree- and neighbourhood-level analyses to account for the differences in stand density and stand structure (Section 4).

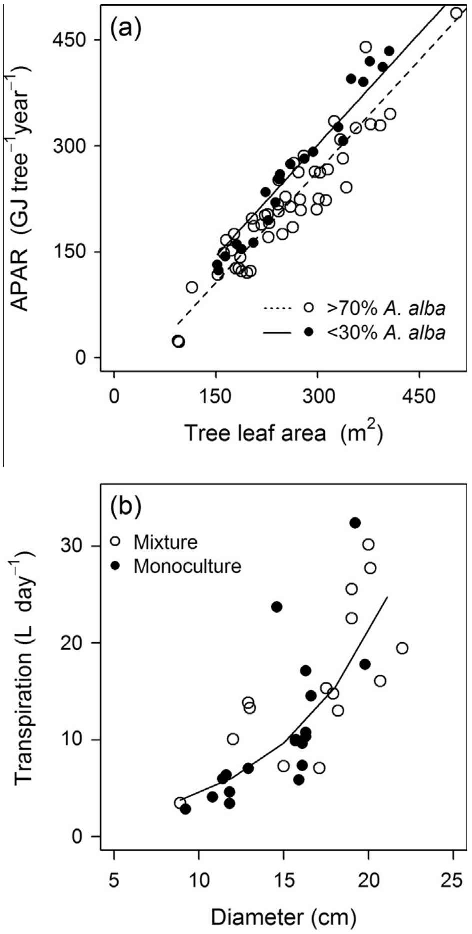

Fig. 4. The relationship between (a) individual tree leaf area and the absorption of photosynthetically active radiation (APAR) of *Abies alba* trees mixed with *Picea abies* trees, and (b) tree diameter (at 1.3 m) and individual tree transpiration of *Eucalyptus globulus* trees growing in monoculture or mixtures with *Acacia mearnsii* (modified from Forrester (2015)). (a) is modified from Forrester and Albrecht (2014) and contains data for *Abies alba* trees whose neighbourhoods (including all trees within a 10 m radius) contained >70% (by basal area) or <30% *A. alba*.

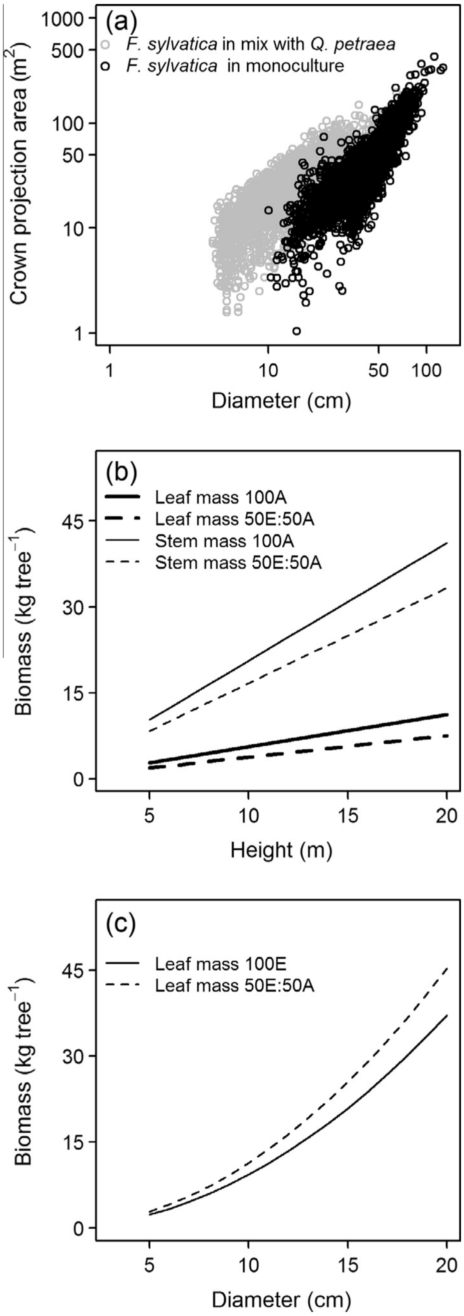

Fig. 5. The effects of species composition on the relationships (a) between diameter and crown projection area for *Fagus sylvatica* trees in mixtures with *Quercus petraea* or monocultures (modified from Pretzsch (2014)), (b) between height and leaf mass or stem mass of *Acacia mangium*, and (c) between diameter and leaf mass of *Eucalyptus grandis* (using equations from Laclau et al. (2008)). In the legend of (b) and (c), the number represents the percent of planted trees per ha of *E. grandis* (E) and *A. mangium* (A).

An example of how species composition can be confounded with silviculture can be seen in central European forests where there is a long tradition of systematic and regular thinning of monocultures that is well supported by quantitative guidelines. Thinning may differ in mixed-species stands when one species, e.

Please cite this article in press as: Forrester, D.I., Pretzsch, H. Tamm Review: On the strength of evidence when comparing ecosystem functions of mixtures with monocultures. Forest Ecol. Manage. (2015), http://dx.doi.org/10.1016/j.foreco.2015.08.016

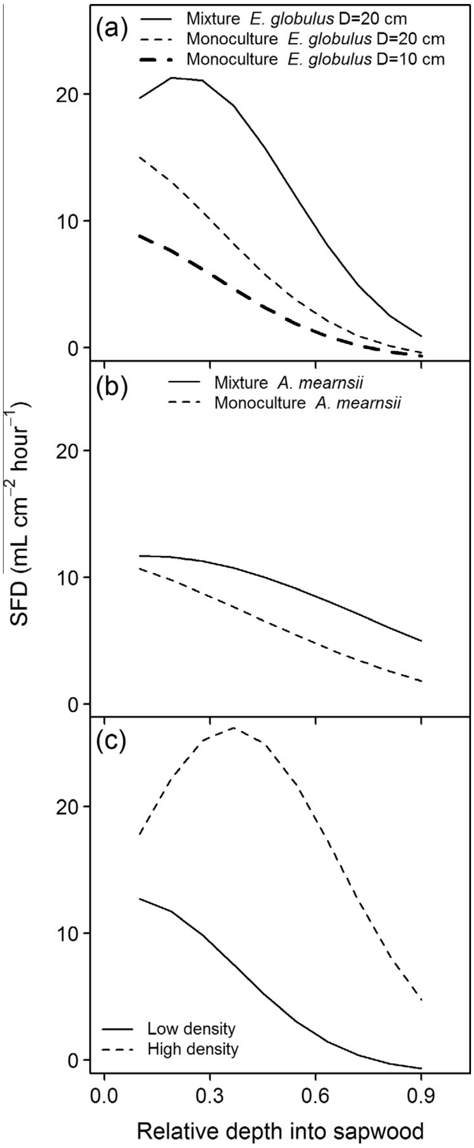

Fig. 6. The relationship between sap flux density (SFD) at 14:30 h and depth into the sapwood from the cambium (cambium = 0, sapwood-to-heartwood boundary = 1) of (a) *Eucalyptus globulus* with diameters (*D*) of 10 or 20 cm, (b) *Acacia mearnsii* with diameters of 15 cm, and (c) the effect of stand density on the same relationship for *E. globulus* with diameters of 15 cm. For each relationship, the effects of other factors have been accounted for such as tree diameter, species composition, climate, and stand density. Modified from Forrester (2015).

g. *F. sylvatica*, is at risk from being overgrown by a second species, e.g. *Picea abies*. The second species may then be thinned more frequently to give the other species a better chance, thereby reducing the stand density of the mixture compared with neighbouring monocultures (Lüpke and Spellmann, 1999). Conversely, the stand density of mixtures relative to monocultures may be increased when mixed stands are more resistant to windthrow, bark beetle attacks, or snow damage (Griess and Knoke, 2011; Griess et al., 2012). These differences in density are not directly related to species interactions but they are important because stand density can significantly influence growth. For example, moderately thinned monocultures in central Europe often grow 10–20% more than unthinned stands (Pretzsch, 2005). Therefore, using the thinned monocultures as a reference may result in an underestimation of the mixing effects because the mixtures are closer to the maximum density.

4. Tree- and neighbourhood-level replication

The previous section focused on the stand level. However, to get more detailed information about the processes driving stand-level patterns, it is often useful to examine tree-level relationships. Tree-level and neighbourhood-level analyses can show how relationships between tree size, neighbourhood density and growth or physiological variables vary between mixtures and monocultures and thereby help to explain the patterns observed at the stand level. Tree-level analyses can also be useful to separate the effects of tree size or age from other mixing effects where the relationships between tree size or age and the response variable actually vary between mixtures and monocultures. Examples of tree-level analyses are shown in Figs. 4–6. In Fig. 4 regression analysis was used to separate the effects of tree size from other mixing effects. The relationships between tree size and light absorption or transpiration were compared in mixtures and monocultures. In these examples there was a significant difference in the relationship between tree size and light absorption, such that *Abies alba* trees with a given leaf area in mixed-species neighbourhoods with *P. abies* absorbed more light than trees with the same leaf area in almost monospecific neighbourhoods (Fig. 4a). In contrast, there was no difference in this relationship for *P. abies* (data not shown). There was also no significant difference (in slopes or intercepts) in the relationship between tree size and transpiration in monocultures and mixtures of *E. globulus* and *A. mearnsii* (Fig. 4b). In this case the mixtures used more water because they contained larger trees.

Tree-level or neighbourhood-level analyses can also be used to increase the number and, to a lesser extent, the range of treatments. For example, at the plot level there may only be a low number of treatments (e.g. one two-species mixture and two monocultures) but within each plot there may be a lot of variation between the neighbourhoods of different individual trees in terms of species composition, species proportion, soil fertility, light and water availability, stand density, etc. To make use of this variability, neighbourhood indices have been widely used in forest research, including to examine the effects of species interactions and stand density on growth (e.g. Boyden et al., 2005; Vanclay, 2006a), light absorption (Forrester and Albrecht, 2014) and transpiration (Forrester, 2015). This allows a greater number of species proportions and stand densities to be examined than would otherwise be possible at the stand level. Nevertheless, this replication at the tree level is only pseudo replication of stand-level site and climate, and is complementary, not a replacement, for stand-level replication. Neighbourhood analyses can also be valuable tools for experiments with small plots. That is, stand-level analyses may not be practical after the trees become large relative to the plots such that the original plot-level treatments become redundant. In these cases, the buffers between plots can be ignored and the whole experiment, including buffer trees, can be examined using a neighbourhood approach.

5. Scaling issues relating to intra-specific variability

Forest research about allometry and physiological responses to resource availability and climatic conditions has largely focused on monocultures. This allometric and physiological information is often used to scale up from leaf or tree to stand level information, and in some studies, information from monocultures is used when upscaling for mixtures. However, when inter-specific interactions result in significant intra-specific variability (Figs. 4–6), using information from monocultures for mixtures, or simply applying the mean of all treatments, can have strong implications for upscaling and result in biased estimates of mixture effects. This is also a potential problem for forest growth models (Pretzsch et al., 2015a). This section discusses potential problems that can result from ignoring intra-specific variability in morphology, physiology and phenology.

5.1. Intra-specific variability in morphology/allometry

Tree diameter is often used to predict variables that are harder to measure, such as height, crown dimensions, volume, biomass, leaf area, sapwood area, transpiration etc. However, it is well known that for a given species and age these relationships can vary with resource availability, climatic conditions and stand density (Litton et al., 2007; Poorter et al., 2012). All of these factors can differ in mixtures compared with monocultures and it is not surprising that allometry has also been reported to differ. For example, Fig. 5a shows how the horizontal crown projection area (m^2) of *F. sylvatica* trees of a given diameter can vary significantly depending on the species composition of the forest. Fig. 4b and c shows how allometric biomass equations differed for *Eucalyptus grandis* and *Acacia mangium* in mixtures and monocultures. Using allometric equations that were developed in monocultures for trees in mixtures can lead to large errors in predictions. Similarly, using the mean equations for all treatments could reduce the predicted size of the actual mixing effects.

This can be shown using the study of *E. grandis* and *A. mangium* monocultures and mixtures where allometric biomass equations were carefully developed for each species in each treatment (Fig. 5b and c) (Laclau et al., 2008). Relationships between diameter, height and most biomass components differed between mixtures and monocultures for both species, such that *E. grandis* trees of a given diameter or height had a greater biomass in mixture than in monoculture, and the opposite for *A. mangium* (Laclau et al., 2008). When the treatment-specific allometric equations from Laclau et al. (2008) are used, the mean tree biomass of *E. grandis* in mixtures is calculated to be 100% larger than trees in monocultures. However, a mixture-monoculture difference of only 34% results when monospecific equations are used in the mixtures, which clearly underestimates the actual mixing effect. In contrast, when the correct equations are used for *A. mangium* the mean tree biomass is 27% lower in mixture than in monoculture, but is only 7% lower when the monospecific equations are used for the *A. mangium* in mixtures.

Some studies have found no differences in allometry between mixtures and monocultures (DeBell et al., 1997), and there may be species and allometric relationships that are more variable (e.g. crown dimensions and leaf area) than others (e.g. stem mass). Therefore, while it is not always necessary to use different allometric relationships for mixtures, it is worthwhile considering when this may be important.

5.2. Intra-specific variability in physiology

Physiological measurements are often made at high spatial and temporal resolutions using specialised equipment, which can make them very time consuming and expensive. This can encourage low replication, and in some studies no consideration is given to whether the physiological relationships used when upscaling vary between the mixtures and monocultures. For example, whole-tree transpiration can be calculated by measuring the sap flux density (SFD, $\mathrm{mL\,cm^{-2}\,h^{-1}}$) every few minutes or hours and then scaling this up to a transpiration value for the whole tree using the sapwood area of the tree. Depending on the method, information about the relationship between SFD and the depth into the sapwood (Fig. 6) is required when scaling up from the tree (and hour) to the stand (and day or month, etc.). It may also be required to correct for any discrepancy between the depth into the sapwood of the measurement probes and the depth of the sapwood-to-heartwood boundary. The SFD profile can differ between mixtures and monocultures, with stand density and with tree size (Fig. 6). This reflects differences in the vertical gradients of water uptake by roots and transpiration by leaves in response to vertical gradients in soil moisture and canopy micro-climates. That is, that outer sapwood is more likely to be used more by higher branches or lateral roots, whereas the inner sapwood is used more by lower branches and deeper roots (Dye et al., 1991; Lu et al., 2004; Nadezhdina et al., 2007; Fiora and Cescatti, 2008; Forrester et al., 2012).

Despite these potential treatment effects, about half of the studies that have compared the transpiration of mixtures and monocultures (reviewed in Forrester, 2015) measured the SFD profiles in each treatment, while the others did not. Those studies not only assumed that the profiles were the same in mixtures and monocultures, but sometimes used profiles measured in monocultures that may have even contained trees of different ages and diameters and different stand densities. The potential error resulting from these assumptions may not need to be high to have a significant influence on the relative differences between mixtures and monocultures.

5.3. Intra- and inter-specific variability in phenology

The magnitude and rhythm of physiological variables within a given growing season can vary between species, and can also vary for a given species depending on the environmental conditions where it is growing. This makes the timing of measurements critical and may require regular or continuous measurements for at least one whole growing season. An obvious example is the seasonality of light absorption and shading in mixtures of evergreen and deciduous species, or mixtures of multiple deciduous species that produce and shed leaves at different times. Similarly, the seasonality of nutrient uptake and cycling varies between species (Richards and Schmidt, 2010). Another example is the timing of the peak in transpiration, which can vary between species such that a species with the highest transpiration during one part of the season may have the lowest rates during another part of the season (Forrester et al., 2010b; Moore et al., 2011; Gebauer et al., 2012; Kunert et al., 2012). In addition to these inter-specific differences, the timing of the transpiration peak of a given species can change depending on whether it is in a monoculture or mixture, and in some cases the ranking of mixtures and monocultures or of different species in the mixture can actually reverse depending on the season (Forrester et al., 2010b; Kunert et al., 2012). Most studies that compare the transpiration of mixtures with monocultures therefore measure transpiration for at least one whole growing season, and these studies have all confirmed these strong seasonal trends (Forrester et al., 2010b; Moore et al., 2011; Gebauer et al., 2012; Kunert et al., 2012). Based on these studies, even if transpiration was measured at several times during the year, but only for a small proportion of the growing season, it may not reflect the actual rankings of species or treatments. While the examples above indicate that intra-specific variability can sometimes be important, this variability differed between species. It is important to consider when this intra-specific variability will be important.

6. Correlations and causality

Correlations do not mean causal mechanisms (Binkley and Menyailo, 2005). Some studies measure patterns in growth and assume that certain mechanisms were responsible. For example, large forest plot data sets have been used to clearly show that complementarity increases for *F. sylvatica* when mixed with *Quercus petraea* or *Q. robur* during drought years or on (drier) sites of lower

Please cite this article in press as: Forrester, D.I., Pretzsch, H. Tamm Review: On the strength of evidence when comparing ecosystem functions of mixtures with monocultures. Forest Ecol. Manage. (2015), http://dx.doi.org/10.1016/j.foreco.2015.08.016

productivity (Pretzsch et al., 2013a). One possible explanation is that the complementarity results from hydraulic redistribution by *Quercus* which increases water availability for *F. sylvatica*. However, a study that has examined this process found that *Q. petraea* did indeed redistribute water from deeper soil layers to shallow layers but *F. sylvatica* did not actually use it (Zapater et al., 2011). Other studies measure significant increases in a certain process in mixtures and assume that process is a main cause of growth responses without considering other processes that may be much more important. For example, taking the example above, even if hydraulic redistribution occurs it may, or may not, have a much smaller impact on water availability and growth than other water-related processes such as inter-specific differences in transpiration or canopy interception. Similarly, increases in growth have often been reported in mixtures containing nitrogen-fixing species and often this has been shown to result, at least in part, from improved nitrogen availability and uptake (Forrester et al., 2006b), however, in other studies there have been high rates of nitrogen fixation but no growth increase (Binkley et al., 1992) or there has been a large growth increase but relatively low rates of nitrogen fixation (Bouillet et al., 2008). These studies carefully showed that even though these processes occurred, they were not large enough to be the main cause of the growth responses that were observed.

An approach that has been used to link the growth responses with resource availability, uptake and efficiency of use is the application of the production ecology equation (Monteith, 1977).

$$\begin{aligned}\text{Gross primary production} = {} & \text{resource supply} \\ & \times \text{fraction of resource acquired} \\ & \times \text{resource use efficiency}\end{aligned} \tag{8}$$

This has been applied when comparing mixtures and monocultures in relation to nutrients (Richards et al., 2010), water (Forrester et al., 2010b; Moore et al., 2011; Kunert et al., 2012; Forrester, 2015) and light (Forrester and Albrecht, 2014). The magnitude of the change in resource availability or uptake or efficiency can be compared with the difference in the process to indicate whether it was large enough to cause the effect.

In addition to Eq. (8), the more comprehensive mass balance approach was recommended by Binkley and Menyailo (2005) when evaluating the strength of evidence of the effects of different tree species on soils. This can also be applied when comparing processes driving the dynamics of mixtures and monocultures. This approach considers all of the inputs and outputs (fluxes) of a given resource in an ecosystem and the changes in the size of all of the existing pools of that resource, which should all balance (e.g. Giardina and Ryan, 2002). This approach gives confidence in the values of the individual pools and fluxes. That is, there is always error associated with measurements of each individual pool or flux. However, while it can be difficult to determine the reliability of measurements of each individual pool or flux, when considered with other measurements in a mass balance, it is clearer whether or not there are any significant errors. Binkley and Menyailo (2005) give examples of studies where the size of pools and fluxes that were measured could not possibly be correct simply because there was not enough of the given resource in the ecosystem.

Mass balance approaches have been applied in mixed-species stands when examining carbon sequestration (Forrester et al., 2006a; Epron et al., 2013). While productivity of mixtures is often quantified in terms of above-ground production, this does not mean that the total biomass production is higher in mixtures; a higher above-ground biomass could also occur when increases in soil resource availability result in a shift in carbon partitioning more to aboveground. For example, fertiliser application has been shown to increase aboveground production in mature eucalypt forests but this resulted from a shift in partitioning rather than a change in gross primary production (Keith et al., 1997). Greater productivity in mixtures compared with monocultures has been associated with a combination of changes in partitioning and changes in net primary production (Forrester et al., 2006a; Epron et al., 2013).

7. Conclusions

Adapting forest ecosystems to global change is one of the largest current challenges in forestry, and mixed-species forests are viewed as one of the most important adaptation and risk-reduction strategies (Messier et al., 2013). Therefore, a good understanding of how species interactions influence growth and other ecosystem functions of mixed-species forests is critical. While many of the topics discussed in this synthesis are not new, they are sometimes forgotten or ignored even though they can have strong implications for the quantification and interpretation of mixing effects. Inappropriate replication through space and time, or assumptions that the allometry, physiology or phenology of a given species is the same in monocultures and mixtures can lead to biased predictions of mixing effects. Stand density and stand structure can also influence mixing effects but the processes driving these patterns are not well understood, which makes it difficult to account for stand density and structure in studies that use inventories or SFPs. This understanding may be improved by complementing commonly used stand-level analyses with neighbourhood-level analyses that help to account for the variability in density and structure within a given stand. Progress will also require a process-based approach, particularly within the framework of mass balances and the production ecology equation. This information will help to indicate which factors are the most important drivers of mixed-species forest dynamics and hence facilitate the development of process-based growth models that are simple enough to be used as management tools but that are still capable of reproducing the spatial and temporal dynamics of species interactions (Forrester and Tang, in press).

Acknowledgements

We would like to thank Dan Binkley for the invitation to write this paper. D. Forrester was funded by a Heisenberg Fellowship (FO 791/4-1) from the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG). Thank you also to four anonymous reviewers who provided very useful suggestions that improved the manuscript.

References

Allen, C.D., Macalady, A.K., Chenchouni, H., Bachelet, D., McDowell, N., Vennetier, M., Kitzberger, T., Rigling, A., Breshears, D.D., Hogg, E.H., Gonzalez, P., Fensham, R., Zhang, Z., Castro, J., Demidova, N., Lim, J.-H., Allard, G., Running, S.W., Semerci, A., Cobb, N., 2010. A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests. For. Ecol. Manage. 259, 660–684.

Baeten, L., Verheyen, K., Wirth, C., Bruelheide, H., Bussotti, F., Finér, L., Jaroszewicz, B., Selvi, F., Valladares, F., Allan, E., Ampoorter, E., Auge, H., Avăcăriei, D., Barbaro, L., Bărnoaiea, I., Bastias, C.C., Bauhus, J., Beinhoff, C., Benavides, R., Benneter, A., Berger, S., Berthold, F., Boberg, J., Bonal, D., Brüggemann, W., Carnol, M., Castagneyrol, B., Charbonnier, Y., Chécko, E., Coomes, D., Coppi, A., Dalmaris, E., Dănilă, G., Dawud, S.M., Vries, W.d., Wandeler, H.D., Deconchat, M., Domisch, T., Duduman, G., Fischer, M., Fotelli, M., Gessler, A., Gimeno, T.E., Granier, A., Grossiord, C., Guyot, V., Hantsch, L., Hättenschwiler, S., Hector, A., Hermy, M., Holland, V., Jactel, H., Joly, F.-X., Juckers, T., Kolb, S., Koricheva, J., Lexer, M.J., Liebergesell, M., Milligan, H., Müller, S., Muys, B., Nguyen, D., Nichiforel, L., Pollastrini, M., Proulx, R., Rabasa, S., Radoglou, K., Ratcliffe, S., Raulund-Rasmussen, K., Seiferling, I., Stenlid, J., Vesterdal, L., Wilpert, K.v., Zavala, M.A., Zielinski, D., Scherer-Lorenzen, M., 2013. A novel comparative research platform designed to determine the functional significance of tree species diversity in European forests. Perspect. Plant Ecol. Evol. Syst. 15, 281–291.

Bauhus, J., van Winden, A.P., Nicotra, A.B., 2004. Above-ground interactions and productivity in mixed-species plantations of *Acacia mearnsii* and *Eucalyptus globulus*. Can. J. For. Res. 34, 686–694.

Binkley, D., 1984. Importance of size-density relationships in mixed stands of Douglas-fir and Red alder. For. Ecol. Manage. 9, 81–85.

Binkley, D., 1992. Mixtures of nitrogen-fixing and non-nitrogen-fixing tree species. In: Cannell, M.G.R., Malcolm, D.C., Robertson, P.A. (Eds.), The Ecology of Mixed Species Stands of Trees. Blackwell Scientific, London, pp. 99–123.

Binkley, D., 2008. Three key points in the design of forest experiments. For. Ecol. Manage. 255, 2022–2023.

Binkley, D., Menyailo, O., 2005. Gaining insights on the effects of tree species on soils. In: Binkley, D., Menyailo, O. (Eds.), Tree Species Effects on Soils: Implications for Global Change. NATO Science Series, Kluwer Academic Publishers, Dordrecht, pp. 1–16.

Binkley, D., Sollins, P., Bell, R., Sachs, D., Myrold, D., 1992. Biogeochemistry of adjacent conifer and alder-conifer stands. Ecology 73, 2022–2033.

Bouillet, J.P., Laclau, J.P., Goncalves, J.L.M., Moreira, M.Z., Trivelin, P.C.O., Jourdan, C., Silva, E.V., Piccolo, M.C., Tsai, S.M., Galiana, A., 2008. Mixed-species plantations of *Acacia mangium* and *Eucalyptus grandis* in Brazil 2: Nitrogen accumulation in the stands and biological N_2 fixation. For. Ecol. Manage. 255, 3918–3930.

Boyden, S., Binkley, D., Senock, R., 2005. Competition and facilitation between *Eucalyptus* and nitrogen-fixing *Falcataria* in relation to soil fertility. Ecology 86, 992–1001.

Boyden, S., Montgomery, R., Reich, P.B., Palik, B., 2012. Seeing the forest for the heterogeneous trees: stand-scale resource distributions emerge from tree-scale structure. Ecol. Appl. 22, 1578–1588.

Bravo-Oviedo, A., Pretzsch, H., Ammer, C., Andenmatten, E., Barbati, A., Barreiro, S., Brang, P., Bravo, F., Coll, L., Corona, P., Ouden, J.d., Ducey, M.J., Forrester, D.I., Giergiczny, M., Jacobsen, J.B., Lesinski, J., Löf, M., Mason, B., Matovic, B., Metslaid, M., Morneau, F., Motiejunaite, J., O'Reilly, C., Pach, M., Ponette, Q., Rio, M.d., Short, I., Skovsgaard, J.P., Soliño, M., Spathelf, P., Sterba, H., Stojanovic, D., Strelcova, K., Svoboda, M., Verheyen, K., Lüpke, N.v., Zlatanov, T., 2014. European mixed forests: definition and research perspectives. For. Syst. 23, 518–533.

Brooker, R., Kikvidze, Z., Kunstler, G., Liancourt, P., Seifan, M., 2013. The concept and measurement of importance: a comment on Rees et al. (2012). J. Ecol. 101, 1369–1378.

Bruelheide, H., Nadrowski, K., Assmann, T., Bauhus, J., Both, S., Buscot, F., Chen, X.-Y., Ding, B., Durka, W., Erfmeier, A., Gutknecht, J.L.M., Guo, D., Guo, L.-D., Härdtle, W., He, J.-S., Klein, A.-M., Kühn, P., Liang, Y., Liu, X., Michalski, S., Niklaus, P.A., Pei, K., Scherer-Lorenzen, M., Scholten, T., Schuldt, A., Seidler, G., Trogisch, S., Oheimb, G.v., Welk, E., Wirth, C., Wubet, T., Yang, X., Yu, M., Zhang, S., Zhou, H., Fischer, M., Ma, K., Schmid, B., 2014. Designing forest biodiversity experiments: general considerations illustrated by a new large experiment in subtropical China. Methods Ecol. Evol. 5, 74–89.

Condés, S., Rio, M.D., Sterba, H., 2013. Mixing effect on volume growth of *Fagus sylvatica* and *Pinus sylvestris* is modulated by stand density. For. Ecol. Manage. 292, 86–95.

DeBell, D.S., Cole, T.C., Whitesell, C.D., 1997. Growth, development, and yield of pure and mixed stands of *Eucalyptus* and *Albizia*. For. Sci. 43, 286–298.

Dietrich, V., 1928. Untersuchungen in Mischbeständen. Mitt Württ Forstl Versuchsanst 1, 25–34.

Dye, P.J., Olbrich, B.W., Poulter, A.G., 1991. The influence of growth rings in *Pinus patula* on heat pulse velocity and sap flow measurements. J. Exp. Bot. 42, 867–870.

Epron, D., Nouvellon, Y., Mareschal, L., MoreiraeMoreira, R., Koutika, L.-S., Geneste, B., Delgado-Rojas, J.S., Laclau, J.-P., Sola, G., Gonçalves, J.L.d.M., Bouillet, J.-P., 2013. Partitioning of net primary production in Eucalyptus and Acacia stands and in mixed-species plantations: two case-studies in contrasting tropical environments. For. Ecol. Manage. 301, 102–111.

Fiora, A., Cescatti, A., 2008. Vertical foliage distribution determines the radial pattern of sap flux density in *Picea abies*. Tree Physiol. 28, 1317–1323.

Flury, P., 1926. Über Zuwachs und Ertrag reiner und gemischter Bestände. Schweiz Z Forstw 77, 337–342.

Flury, P., 1931. Untersuchungen über Zuwachs, Massen-und Geldertrag reiner und gemischter Bestände. Mitt Schweiz Anst Forstl Versuchswesen 16, 452–472.

Forrester, D.I., 2014. The spatial and temporal dynamics of species interactions in mixed-species forests: from pattern to process. For. Ecol. Manage. 312, 282–292.

Forrester, D.I., 2015. Transpiration and water-use efficiency in mixed-species forests versus monocultures: effects of tree size, stand density and season. Tree Physiol. 35, 289–304.

Forrester, D.I., Smith, R.G.B., 2012. Faster growth of *Eucalyptus grandis* and *Eucalyptus pilularis* in mixed-species stands than monocultures. For. Ecol. Manage. 286, 81–86.

Forrester, D.I., Albrecht, A.T., 2014. Light absorption and light-use efficiency in mixtures of *Abies alba* and *Picea abies* along a productivity gradient. For. Ecol. Manage. 328, 94–102.

Forrester, D.I., Tang, X., 2015. Analysing the spatial and temporal dynamics of species interactions in mixed-species forests and the effects of stand density using the 3-PG model. Ecol. Modell. (in press). http://dx.doi.org/10.1016/j.ecolmodel.2015.07.010.

Forrester, D.I., Bauhus, J., Cowie, A.L., 2006a. Carbon allocation in a mixed-species plantation of *Eucalyptus globulus* and *Acacia mearnsii*. For. Ecol. Manage. 233, 275–284.

Forrester, D.I., Collopy, J.J., Morris, J.D., 2010a. Transpiration along an age series of *Eucalyptus globulus* plantations in southeastern Australia. For. Ecol. Manage. 259, 1754–1760.

Forrester, D.I., Vanclay, J.K., Forrester, R.I., 2011. The balance between facilitation and competition in mixtures of *Eucalyptus* and *Acacia* changes as stands develop. Oecologia 166, 265–272.

Forrester, D.I., Bauhus, J., Cowie, A.L., Vanclay, J.K., 2006b. Mixed-species plantations of *Eucalyptus* with nitrogen fixing trees: a review. For. Ecol. Manage. 233, 211–230.

Forrester, D.I., Theiveyanathan, S., Collopy, J.J., Marcar, N.E., 2010b. Enhanced water use efficiency in a mixed *Eucalyptus globulus* and *Acacia mearnsii* plantation. For. Ecol. Manage. 259, 1761–1770.

Forrester, D.I., Kohnle, U., Albrecht, A.T., Bauhus, J., 2013. Complementarity in mixed-species stands of *Abies alba* and *Picea abies* varies with climate, site quality and stand density. For. Ecol. Manage. 304, 233–242.

Forrester, D.I., Collopy, J.J., Beadle, C.L., Warren, C.R., Baker, T.G., 2012. Effect of thinning, pruning and nitrogen fertiliser application on transpiration, photosynthesis and water-use efficiency in a young *Eucalyptus nitens* plantation. For. Ecol. Manage. 266, 286–300.

Garber, S.M., Maguire, D.A., 2004. Stand productivity and development in two mixed-species spacing trials in the Central Oregon Cascades. For. Sci. 50, 92–105.

Gebauer, T., Horna, V., Leuschner, C., 2012. Canopy transpiration of pure and mixed forest stands with variable abundance of European beech. J. Hydrol. 442–443, 2–14.

Giardina, C.P., Ryan, M.G., 2002. Total belowground carbon allocation in a fast-growing *Eucalyptus* plantation estimated using a carbon balance approach. Ecosystems 5, 487–499.

Griess, V.C., Knoke, T., 2011. Growth performance, windthrow, and insects: meta-analyses of parameters influencing performance of mixed-species stands in boreal and northern temperate biomes. Can. J. For. Res. 41, 1141–1159.

Griess, V.C., Acevedo, R., Härtl, F., Staupendahl, K., Knoke, T., 2012. Does mixing tree species enhance stand resistance against natural hazards? A case study for spruce. For. Ecol. Manage. 267, 284–296.

Harper, J.L., 1977. Population Biology of Plants. Academic Press, New York.

Hawthorne, S.N.D., Lane, P.N.J., Bren, L.J., Sims, N.C., 2013. The long term effects of thinning treatments on vegetation structure and water yield. For. Ecol. Manage. 310, 983–993.

Healy, C., Gotelli, N.J., Potvin, C., 2008. Partitioning the effects of biodiversity and environmental heterogeneity for productivity and mortality in a tropical tree plantation. J. Ecol. 96, 903–913.

Hofmann, F., 1923. Mischungen aus Buchen mit Nadelholz, insbesondere mit der Fichte und Tanne. Allgemeine Forst-und Jagdzeitung 99, 273–281.

Keith, H., Raison, R.J., Jacobsen, K.L., 1997. Allocation of carbon in a mature eucalypt forest and some effects of soil phosphorus availability. Plant Soil 196, 81–99.

Kelty, M.J., 1992. Comparative productivity of monocultures and mixed-species stands. In: Kelty, M.J., Larson, B.C., Oliver, C.D. (Eds.), The Ecology and Silviculture of Mixed-Species Forests. Kluwer Academic Publishers, Dordrecht, pp. 125–141.

Kelty, M.J., Cameron, I.R., 1995. Plot designs for the analysis of species interactions in mixed stands. Commonwealth For. Rev. 74, 322–332.

Kennel, R., 1965. Untersuchungen über die Leistung von Fichte und Buche im Rein- und Mischbestand. Sauerländer.

Kunert, N., Schwendenmann, L., Potvin, C., Hölscher, D., 2012. Tree diversity enhances tree transpiration in a Panamanian forest plantation. J. Appl. Ecol. 49, 135–144.

Laclau, J.-P., Bouillet, J.-P., Gonçalves, J.L.M., Silva, E.V., Jourdan, C., Cunha, M.C.S., Moreira, M.R., Saint-André, L., Maquère, V., Nouvellon, Y., Ranger, J., 2008. Mixed-species plantations of *Acacia mangium* and *Eucalyptus grandis* in Brazil 1. Growth dynamics and aboveground net primary production. For. Ecol. Manage. 255, 3905–3917.

Litton, C.M., Raich, J.W., Ryan, M.G., 2007. Carbon allocation in forest ecosystems. Glob. Change Biol. 13, 2089–2109.

Loreau, M., Hector, A., 2001. Partitioning selection and complementarity in biodiversity experiments. Nature 412, 72–76.

Lu, P., Urban, L., Zhao, P., 2004. Granier's thermal dissipation probe (TDP) method for measuring sap flow in trees: theory and practice. Acta Bot. Sin. 46, 631–646.

Lüpke, B.V., Spellmann, H., 1999. Aspects of stability, growth and natural regeneration in mixed Norway spruce-European beech stands as a basis of silvicultural decisions. Management of Mixedspecies Forest: Silviculture and Economics. IBN-DLO Scientific Contributions, Wageningen, pp. 245–267.

Messier, C., Puettmann, K.J., Coates, K.D. (Eds.), 2013. Managing Forests as Complex Adaptive Systems: Building Resilience to the Challenge of Global Change. Routledge Chapman & Hall, New York.

Monteith, J.L., 1977. Climate and the efficiency of crop production in Britain. Philos. Trans. R. Soc. B 281, 277–294.

Moore, G.W., Bond, B.J., Jones, J.A., 2011. A comparison of annual transpiration and productivity in monoculture and mixed-species Douglas-fir and red alder stands. For. Ecol. Manage. 262, 2263–2270.

Nadezhdina, N., Nadezhdin, V., Ferreira, M.I., Pitacco, A., 2007. Variability with xylem depth in sap flow in trunks and branches of mature olive trees. Tree Physiol. 27, 105–113.

Please cite this article in press as: Forrester, D.I., Pretzsch, H. Tamm Review: On the strength of evidence when comparing ecosystem functions of mixtures with monocultures. Forest Ecol. Manage. (2015), http://dx.doi.org/10.1016/j.foreco.2015.08.016

Poorter, H., Niklas, K.J., Reich, P.B., Oleksyn, J., Poot, P., Mommer, L., 2012. Biomass allocation to leaves, stems and roots: meta-analyses of interspecific variation and environmental control. New Phytol. 193, 30–50.

Pretzsch, H., 2005. Stand density and growth of Norway spruce (Picea abies (L.) Karst.) and European beech (Fagus sylvatica L.): evidence from long-term experimental plots. Eur. J. Forest Res. 124, 193–205.

Pretzsch, H., 2009. Forest Dynamics, Growth and Yield. From Measurement to Model. Springer, Berlin and Heidelberg.

Pretzsch, H., 2014. Canopy space filling and tree crown morphology in mixed-species stands compared with monocultures. For. Ecol. Manage. 327, 251–264.

Pretzsch, H., Schütze, G., 2009. Transgressive overyielding in mixed compared with pure stands of Norway spruce and European beech in Central Europe: evidence on stand level and explanation on individual tree level. Eur. J. Forest Res. 128, 183–204.

Pretzsch, H., Block, J., Dieler, J., Dong, P.H., Kohnle, U., Nagel, J., Spellmann, H., Zingg, A., 2010. Comparison between the productivity of pure and mixed stands of Norway spruce and European beech along an ecological gradient. Ann. For. Sci. 76, 712–723.

Pretzsch, H., Bielak, K., Block, J., Bruchwald, A., Dieler, J., Ehrhart, H.-P., Kohnle, U., Nagel, J., Spellmann, H., Zasada, M., Zingg, A., 2013a. Productivity of mixed versus pure stands of oak (Quercus petraea (MATT.) LIEBL. and *Quercus robur* L.) and European beech (*Fagus sylvatica* L.) along an ecological gradient. Eur. J. Forest Res. 132, 263–280.

Pretzsch, H., Bielak, K., Bruchwald, A., Dieler, J., Dudzińska, M., Erhart, H.-P., Jensen, A.M., Johannsen, V.K., Kohnle, U., Nagel, J., Spellmann, H., Zasada, M., Zingg, A., 2013b. Mischung und Produktivität von Waldbeständen. Ergebnisse langfristiger ertragskundlicher Versuche. Allg. Forst- u. J.-Ztg. 184, 177–196.

Pretzsch, H., Rötzer, T., Matyssek, R., Grams, T.E.E., Häberle, K.H., Pritsch, K., Kerner, R., Munch, J.C., 2014. Mixed Norway spruce (Picea abies [L.] Karst) and European beech (Fagus sylvatica [L.]) stands under drought: from reaction pattern to mechanism. Trees Struct. Funct. 28, 1305–1321.

Pretzsch, H., Forrester, D.I., Rötzer, T., 2015a. Representation of species mixing in forest growth models. A review and perspective. Ecol. Modell. 313, 276–292.

Pretzsch, H., del Río, M., Ammer, C., Avdagic, A., Barbeito, I., Bielak, K., Brazaitis, G., Coll, L., Dirnberger, G., Drössler, L., Fabrika, M., Forrester, D.I., Godvod, K., Heym, M., Hurt, V., Kurylyak, V., Löf, M., Lombardi, F., Matović, B., Mohren, F., Motta, R., den Ouden, J., Pach, M., Ponette, Q., Schütze, G., Schweig, J., Skrzyszewski, J., Sramek, V., Sterba, H., Stojanović, D., Svoboda, M., Vanhellemont, M., Verheyen, K., Wellhausen, K., Zlatanov, T., Bravo-Oviedo, A., 2015b. Growth and yield of mixed versus pure stands of Scots pine (*Pinus sylvestris* L.) and European beech (*Fagus sylvatica* L.) analysed along a productivity gradient through Europe. Eur. J. For. Res. 134, 927–947.

Ratcliffe, S., Holzwarth, F., Nadrowski, K., Levick, S., Wirth, C., 2015. Tree neighbourhood matters – tree species composition drives diversity-productivity patterns in a near-natural beech forest. For. Ecol. Manage. 335, 225–234.

Reineke, L.H., 1933. Perfecting a stand-density index for even aged forests. J. Agric. Res. 46, 627–638.

Reyes-Hernandez, V., Comeau, P.G., Bokalo, M., 2013. Static and dynamic maximum size-density relationships for mixed trembling aspen and white spruce stands in western Canada. For. Ecol. Manage. 289, 300–311.

Richards, A.E., Schmidt, S., 2010. Complementary resource use by tree species in a rain forest tree plantation. Ecol. Appl. 20, 1237–1254.

Richards, A.E., Forrester, D.I., Bauhus, J., Scherer-Lorenzen, M., 2010. The influence of mixed tree plantations on the nutrition of individual species: a review. Tree Physiol. 30, 1192–1208.

Río, M.d., Sterba, H., 2009. Comparing volume growth in pure and mixed stands of *Pinus sylvestris* and *Quercus pyrenaica*. Ann. For. Sci. 66, 502.

Sapijanskas, J., Paquette, A., Potvin, C., Kunert, N., Loreau, M., 2014. Tropical tree diversity enhances light capture through crown plasticity and spatial and temporal niche differences. Ecology 95, 2479–2492.

Scherer-Lorenzen, M., Potvon, C., Koricheva, J., Schmid, B., Hector, A., Bornik, Z., Reynolds, G., Schulze, E.-D., 2005. The design of experimental plantations for functional biodiversity research. In: Scherer-Lorenzen, M., Körner, C., Schulze, E.-D. (Eds.), Forest Diversity and Function: Temperate and Boreal Systems Ecological Studies, 176. Springer, Berlin, pp. 347–376.

Schume, H., Jost, G., Hager, H., 2004. Soil water depletion and recharge patterns in mixed and pure forest stands of European beech and Norway spruce. J. Hydrol. 289, 258–274.

Schwappach, A., 1909. Untersuchungen in Mischbeständen. Zeitschr Forst-u Jagdwesen 41, 313–332.

Shannon, C.E., 1948. A mathematical theory of communication. Bell Syst. Tech. J. 27 (379–423), 623–656.

Stape, J.L., Binkley, D., Jacob, W.S., Takahashi, E.N., 2006. A twin-plot approach to determine nutrient limitation and potential productivity in *Eucalyptus* plantations at landscape scales in Brazil. For. Ecol. Manage. 223, 358–362.

Stape, J.L., Binkley, D., Ryan, M.G., Fonseca, S., Loos, R.A., Takahashi, E.N., Silva, C.R., Silva, S.R., Hakamada, R.E., Ferreira, J.M.d.A., Lima, A.M.N., Gava, J.L., Leite, F.P., Andrade, H.B., Alves, J.M., Silva, G.G.C., Azevedo, M.R., 2010. The Brazil eucalyptus potential productivity project: influence of water, nutrients and stand uniformity on wood production. For. Ecol. Manage. 259, 1684–1694.

Sterba, H., Rio, M.d., Brunner, A., Condes, S., 2014. Effect of species proportion definition on the evaluation of growth in pure vs. mixed stands. For. Syst. 23, 547–559.

Vanclay, J.K., 2006a. Spatially-explicit competition indices and the analysis of mixed-species plantings with the Simile modelling environment. For. Ecol. Manage. 233, 295–302.

Vanclay, J.K., 2006b. Experiment designs to evaluate inter- and intra-specific interactions in mixed plantings of forest trees. For. Ecol. Manage. 233, 366–374.

Vandermeer, J., 1989. The Ecology of Intercropping. Cambridge University Press, New York.

Vilà, M., Carrillo-Gavilán, A., Vayreda, J., Bugmann, H., Fridman, J., Grodzki, W., Haase, J., Kunstler, G., Schelhaas, M., Trasobares, A., 2013. Disentangling biodiversity and climatic determinants of wood production. PLoS ONE 8, e53530.

von Oheimb, G., Lang, A.C., Bruelheide, H., Forrester, D.I., Wäsche, I., Yu, M., Härdtle, W., 2011. Individual-tree radial growth in a species-rich subtropical forest: the role of local neighbourhood competition. For. Ecol. Manage. 261, 499–507.

Weigelt, A., Jolliffe, P., 2003. Indices of plant competition. J. Ecol. 91, 707–720.

Weiskittel, A., Gould, P., Temesgen, H., 2009. Sources of variation in the self-thinning boundary line for three species with varying levels of shade tolerance. For. Sci. 55, 84–93.

Willey, R.W., Osiru, D.S.O., 1972. Studies on mixtures of maize and beans (*Phaseolus vulgaris*) with special reference to plant population. J. Agric. Sci. 79, 519–529.

Wimmenauer, K., 1914. Zur Frage der Mischbestände. Allgemeine Forst-und Jagdzeitung 90, 90–93.

Woodall, C.W., Miles, P.D., Vissage, J.S., 2005. Determining maximum stand density index in mixed species stands for strategic-scale stocking assessments. For. Ecol. Manage. 216, 367–377.

Zapater, M., Hossann, C., Bréda, N., Bréchet, C., Bonal, D., Granier, A., 2011. Evidence of hydraulic lift in a young beech and oak mixed forest using ^{18}O soil water labelling. Trees 25, 885–894.

Zhao, X., Corral-Rivas, J., Zhang, C., Temesgen, H., Gadow, K.v., 2014. Forest observational studies – an essential infrastructure for sustainable use of natural resources. For. Ecosyst. 1, 8.

Zuur, A.F., Ieno, E.N., Elphick, C.S., 2010. A protocol for data exploration to avoid common statistical problems. Methods Ecol. Evol. 1, 3–14.

Please cite this article in press as: Forrester, D.I., Pretzsch, H. Tamm Review: On the strength of evidence when comparing ecosystem functions of mixtures with monocultures. Forest Ecol. Manage. (2015), http://dx.doi.org/10.1016/j.foreco.2015.08.016