Unpunctual in diversity : the effect of stand species richness on spring phenology of deciduous tree stands varies among species and years

Reference:

Heinecke Thilo, De Frenne Pieter, Verheyen Kris, Nijs Ivan, Matthysen Erik, Campioli Matteo.- Unpunctual in diversity : the effect of stand species richness on spring phenology of deciduous tree stands varies among species and years
Global change biology - ISSN 1365-2486 - 30:4(2024), e17266
Full text (Publisher's DOI): https://doi.org/10.1111/GCB.17266
To cite this reference: https://hdl.handle.net/10067/2050320151162165141

Unpunctual in diversity: The effect of stand species richness on spring phenology of deciduous tree stands varies among species and years

Spring phenology and tree species richness

Thilo Heinecke[1], Pieter De Frenne[2], Kris Verheyen[2], Ivan Nijs[1], Erik Matthysen[3] and Matteo Campioli[1]

ORCIDs:

Thilo Heinecke[1]	https://orcid.org/0000-0001-9913-7901
Pieter DeFrenne[2]	https://orcid.org/0000-0002-8613-0943
Kris Verheyen[2]	https://orcid.org/0000-0002-2067-9108
Ivan Nijs[1]	https://orcid.org/0000-0003-3111-680X
Erik Matthysen[3]	https://orcid.org/0000-0002-7521-9248
Matteo Campioli[1]	https://orcid.org/0000-0002-3427-2368

[1] Plants and Ecosystems (PLECO), Department of Biology, University of Antwerp, Universiteitsplein 1, 2610 Wilrijk, Belgium; [2] Forest & Naturelab (ForNaLab), Department of Environment, Ghent University, Geraardsbergsesteenweg 267, 9090 Gontrode, Belgium; [3] Evolutionary Ecology (EVECO), Department of Biology, University of Antwerp, Universiteitsplein 1, 2610 Wilrijk, Belgium

Author for correspondence:

Thilo Heinecke

Tel: +49 170 590 35 85

Email: th-h@outlook.de

Abstract

Climatic drivers alone do not adequately explain the regional variation in budburst timing in deciduous forests across Europe. Stand-level factors, such as tree species richness, might affect budburst timing by creating different microclimates under the same site macroclimate. We assessed different phases of the spring phenology (start, midpoint, end, and overall duration of the budburst period) of four important European tree species (*Betula pendula*, *Fagus sylvatica*, *Quercus robur* and *Tilia cordata*) in monocultures and 4-species mixture stands of a common garden tree biodiversity experiment in Belgium (FORBIO) in 2021 and 2022. Microclimatic differences between the stands in terms of bud chilling, temperature forcing, and soil temperature were considerable, with 4-species mixtures being generally colder than monocultures in spring, but not in winter. In the colder spring of 2021, at the stand level, the end of the budburst period was advanced, and its overall duration shortened, in the 4-species mixtures. At species level, this response was significant for *F. sylvatica*. In the warmer spring of 2022, advances in spring phenology in 4-species stands were observed again in *F. sylvatica* and, less markedly, in *B. pendula* but without a general response at the stand level. *Q. robur* showed specific patterns with delayed budburst start in 2021 in the 4-species mixtures and very short budburst duration for all stands in 2022. Phenological differences between monocultures and 4-species mixtures were linked to microclimatic differences in light availability rather than in temperature as even comparatively colder microclimates showed an advanced phenology. Compared to weather conditions, tree species richness had a lower impact on budburst timing, but this impact can be of importance for key species like *F. sylvatica* and colder springs. These results indicate that forest biodiversity can affect budburst phenology, with wider implications, especially for forest- and land surface models.

Keywords: biodiversity, budburst, microclimate, common garden, forest ecology, FORBIO, phenology, tree ecology

Introduction

The spring phenology of temperate trees has received much attention in recent decades after it was demonstrated that it was advancing along with the rising temperatures induced by climate change (Menzel et al., 2006, 2020; Monahan et al., 2016). This owes itself to temperature being the major driver of leaf unfolding, with buds requiring initial chilling via an exposure to low temperatures during endodormancy, followed by higher temperatures forcing to initiate budburst during ecodormancy, in combination with a certain level of photoperiod (Basler & Körner, 2014; Baumgarten et al., 2021; Laube et al., 2014). Species differ in their individual requirements of these factors (i.e. chilling, forcing, photoperiod), often reflecting the genotype and thereby the evolutionary past of the species (e.g. climatic distribution limitations, Vitasse et al. 2013) as well as their ecological strategies (e.g. early and late flushing species, Baumgarten et al. 2021). Nevertheless, current budburst models that exclusively rely on these climatic drivers remain unable to predict budburst with reasonable precision (even when integrating species specific parameters), especially at larger spatial scales (Migliavacca et al., 2012; Peaucelle et al., 2019). This is problematic, as these phenological models form an important part of the land surface models, which are used to simulate atmosphere-biosphere fluxes of matter and energy within current climate models.

Studies show that 40-50% of the regional variability in budburst timing from 1970 until 2016 in Europe cannot be attributed to the variability of climatic drivers alone (Peaucelle et al., 2019). While part of this variability is expected to be due to intraspecific genetic differences (e.g. unaccounted phenological differences between different provenances of the same species), the increase of unexplained variability with increasing spatial scale could also be caused by non-climatic, spatially explicit co-drivers, either as a direct effect or in interaction with the known climatic drivers. For instance, recent studies have already demonstrated that tree spring phenology can be affected by dendrometric factors like tree size and diameter (Marchand et al., 2020) and soil characteristics (Arend et al., 2016; Vander Mijnsbrugge et al., 2016; Vander Mijnsbrugge & Moreels, 2020). Importantly, however, these studies only consider the general local weather conditions, the macroclimate (i.e. the overall air temperature that concerns all stands and trees), and not the tree microclimate (i.e. the temperature directly experienced by a tree within a stand).

We suggest that stand tree species richness also has the potential to affect tree spring phenology. The timing of budburst represents a risk-reward trade-off. The reward of an earlier budburst in form of earlier resource acquisition (and exclusive, unhindered access to sunlight), is balanced against the risk of damaging the unfolding leaves during late frost events. The reward side of this trade-off is, however, arguably subject to different forces in monospecific and multi-species stands. In monospecific stands, competition predominates with individuals competing (e.g. for access to sunlight) based on the same requirements (in chilling, temperature forcing and photoperiod) and therefore the benefit of an earlier budburst might outweigh the static risk of late frost. On the other hand, in multi-species stands, the larger complementarity might make an earlier, but risky, budburst less desirable (Forrester, 2014). This would lead to differences in budburst timing for the same species (and provenance) in monospecific and multi-species stands, despite the same macroclimatic spring weather conditions. A study in US wetlands linked plant species richness to a reduction in (satellite derived) phenological variability which indicated a complex relationship between phenology, climatic conditions and edaphic factors (Dronova et al., 2022). Also in grasslands, an experimental reduction in grass species diversity led to an earlier flowering phenology in most grass species (Wolf et al., 2017). Similar responses in deciduous trees have, to our knowledge, however, never been investigated.

Gaining insight into the impact of species diversity on phenology would thus be very important. The biodiversity of temperate broadleaved forests is projected to decline under current climate change scenarios if management practices are not adapted and improved (Thuiller et al., 2006). At the same time, multi-species stands store more carbon and can mitigate major consequences of climate change, such as increasing aridity, and buffer against climatic extremes (De Frenne et al., 2013; Rodríguez De Prado et al., 2022). Knowledge about the feedbacks between tree species diversity and tree phenology would enhance our ability to accurately model future climate changes. Studies in grasslands have already indicated shifts in community wide flowering phenology, as well as in community composition (based on phenological trait groups), when exposed to warming and cooling in a transplantation experiment along an elevational gradient (Meng et al., 2017). In deciduous forests, similar feedbacks between phenology, climate and species composition would have considerable consequences. Under decreased biodiversity and increased competition, for example, spring phenology could be advanced, leading to an increased risk of damage from late frost and greater water consumption(Richardson et al., 2018). On the other hand, an increase in biodiversity might delay spring phenology, leading to a reduction of growing season length, biomass production and carbon uptake (Peñuelas et al., 2009; Richardson et al., 2013).

Shifts in phenology in response to stand species richness could ultimately be facilitated by the different microclimatic conditions experienced by the trees in monospecific and multi-species stands, in particular differences in i) the temperature and/or ii) the light availability. Both these microclimatic factors have been demonstrated to affect the speed of leaf unfolding in individual buds (Vitasse et al., 2021). Microclimate also varies considerably between stands of different species richness configurations when within the same general macroclimatic conditions (De Frenne et al., 2021; Zhang et al., 2022). The same macroclimate can be modified by the stand species richness into different stand-specific microclimates, with the differences potentially being large enough to lead to differences in spring phenology.

Here, we investigate how species richness, as a stand-level characteristic, affects the average budburst timing of four common European tree species within a common-garden tree biodiversity experiment in western Belgium (FORBIO Zedelgem). The site consists of tree stands of varying species richness (ranging from monocultures to 4-species mixtures) but of equal tree age, tree density, soil type and macroclimatic conditions, thereby eliminating these otherwise confounding factors. As each tree stand is furthermore equipped with a separate weather station, the site allows us to connect stand microclimatic conditions to stand phenology. We compare the lowest (monospecific) and the highest (4-species) stand species richness levels available for the four studied species, namely *Betula pendula* Roth (further referred to as Betula for brevity), *Fagus sylvatica* L. (Fagus), *Tilia cordata* Mill. (Tilia) and

Quercus robur L. (Quercus). These species represent the most common, and often dominant, tree species in temperate European broadleaved forests and also differ considerably in their ecological strategies, e.g. early and late flushing species (Leuschner & Meier, 2018). We postulate that different local stand species richness configurations (monoculture vs. 4-species mixture) lead to potentially species-specific differences in average spring phenological timings at the stand level. Due to the higher competition, we hypothesize budburst timings in the monospecific stands to be consistently earlier when compared to the 4-species stands. We expect that these differences are related to stand specific differences in the microclimatic conditions of temperature and/or light availability. The results will indicate if species richness at the stand level can partly explain intra- and interspecific spatial variability in budburst timing not explained by macroclimatic factors alone.

Methods

Study site. The study was conducted in the spring of 2021 and 2022 at one of the FORBIO (FORest BIOdiversity and ecosystem functioning) tree diversity experimental sites, situated in Zedelgem (51°9’N, 3°7’E; Belgium). The site is characterized by a mild, temperate climate (mean annual temperature 9.4°C, mean annual precipitation 687 mm) and was until 2008 in agricultural use (Verheyen et al., 2013). Within the FORBIO framework, 2- to 3-year-old saplings measuring at least 100 cm and belonging to five common European tree species of local provenances (a single provenance per species was studied here) were planted in 2009-2011 in stands of varying species richness configurations, ranging from monospecific to 4-species stands. The multi-species stands are composed of monospecific blocks of 3x3 trees which were spatially randomly spread (Verheyen et al., 2013). Each stand measures 42 m x 42 m and originally contained 784 trees in a regular 1.5 m x 1.5 m grid, with the experimental site overall measuring approximately 1 km x 1 km. In spring 2021, the average tree height ranged from 5.49 m (Tilia) to 10.91 m (Betula), ensuring close contact between the individual trees. Furthermore, each stand has five designated observation zones. These zones are at fixed, regularly spaced points in monospecific stands, while they are placed at points in which four monospecific 3x3 blocks of different species border in the four-species stands (Supporting

Information, Figure S5). We studied the two available monospecific stands of Betula, Fagus, Quercus and Tilia as well as the four available 4-species mixture stands (12 stands in total). All four 4-species stands contained Betula and Fagus, while Quercus and Tilia were only present in three 4-species stands, the evergreen *Pinus sylvestris* L. being the fourth species in these stands. While this leads to three possible combinations of species constituting a 4-species stand, we treated these as replicates of '4-species environment' for simplicity and statistical soundness. Within each of the five observation zones of the observed stands, two representative and healthy individuals of each species were selected for observation in January 2021, leading to a total of 220 trees (ten trees per eight monospecific stands, ten trees per four species in two 4-species stands, ten trees per three species in two 4-species stands). Within each stand (close to its centre), and additionally in an open field outside, but adjacent, to the forest, the air and soil temperature was recorded hourly via Lascar Easylog EL-USB-1 and EL-USB-2 sensors. The air temperature sensors were attached to a wooden pole at 1 m height and shielded against solar radiation, while the soil temperature sensors were buried at a depth of 5 cm, 10 cm beside the pole (Zhang et al., 2022).

Sampling methodology. Every sampling consisted of observing five random buds in each the upper and lower half of the canopy of each study tree via binoculars (Swarovski EL 10x32), for a total of ten buds per tree. At every sampling, each random bud was categorized according to the following 5-stage system: 0 = bud dormant/closed; 1 = bud swollen and/or greening; 2 = budburst, bud shape disintegrating; 3 = leaf margin visible and 4 = leaf unfolded, petiole visible (Gričar et al., 2017; Vitasse et al., 2009). Sampling started weekly from the last week of February until the first bud was recorded at stage one (mid-March), then twice per week until almost all buds reached stage four (end of May). All observations were carried out by the same researcher (TH) except for Quercus in 2022 which was conducted by another researcher under the supervision of TH for feasibility reasons.

In April 2021, the height of each tree under observation was measured via a laser rangefinder (TruPulse 360 by Laser Technology Inc., USA), while the diameter was measured at breast height (130 cm; DBH) with a generic measurement tape. We tested differences in average tree height and DBH between the two species richness configurations via simple linear models that satisfied all model assumption. Furthermore, in 2022, the condition (‘normal’, ‘supressed’ = less than half as high than surrounding trees of the same species, ‘dead’ = tree developed and standing but dead, ‘non-present’ = non-existent or < 0.5 m high) of the two layers of trees around each observed tree (eight trees in the first layer, 16 in the second layer), i.e. their immediate neighbours and those then adjacent to them, were recorded to configure Area-Potentially-Available (APA) maps. These maps enable a weighted, raster based calculation of the area (in m²) that a tree has either already occupied or has available for further growth when considering the condition of its neighbours (R package ‘APAtree’ version 1.0.1, Glatthorn 2021). We configured the maps based on the originally planted 1.5 m x 1.5 m grid of the two surrounding layers of trees in a 10 cm² resolution and weighted via their condition (‘normal’ = 1; ‘supressed’ = 0.5; ‘dead’ = 0.25; ‘non-present’ = 0).

Analysis. Statistical analyses were conducted in R (version 4.1.2, R Core Team 2019) using various packages (see text), while graphs were produced with the ‘ggplot2’ package (version 3.3.5). We point out and acknowledge that conducting the analysis on a species richness treatment level can be considered a pseudo-replication within the experimental design. However, only two monospecific stands per species exist at the Zedelgem site, making the inclusion of the factor ‘stand’ statistically impossible without forfeiting the benefit of the common garden structure. We thereby treated the site as composed of one monospecific and one 4-species tree stand for each species while accepting that this removes any appropriate replication. Although this limits the generalization of the results, it is a common case in ecological observational studies.

Stand microclimate. The hourly temperature data of the individual stand weather stations was averaged for each stand species configuration of each tree species (n=2 for each monospecific stand, n=4 for 4-species stands of Betula and Fagus, n=3 for 4-species stands of Quercus and Tilia) to obtain an average daily air and soil temperature. From these we established: i) the sum of cold days (NCDair), i.e. days with an average air temperature below 5°C, from December 16th of the previous year (the earliest date measurements were available in 2020) to March 31st of the observation year, to characterize chilling (Baumgarten et al., 2021); (ii) the sum of degrees Celsius above 5°C in average air temperature (growing-degree-days, GDDair) from January 1st to June 1st of the observation year, to characterize temperature forcing (Peaucelle et al., 2019); and (iii) the soil temperature. After inspecting the soil temperature differences between the two species richness configurations, we extracted the difference in soil temperature of each configuration to the overall average soil temperature of all sensors for DOY 100 to 150 and statistically tested it via simple linear models that satisfied all model assumptions.

Spring phenology timing variables. We consider four response variables. Firstly, and most crucially, the tree budburst date (BBd), the day-of-year (DOY) at which 50% of the observed buds of a tree have opened (at or above stage two), the most common descriptor of tree budburst (Vitasse et al., 2009). To evaluate the gradual dynamics of bud opening further, inherently being more a process than a definite event in time, we established three more variables: the start (LUst), end (LUen) and overall duration (LUdur) of the budburst period. We compensate for random observation biases in sampling, induced by weather conditions or the observer, by deriving the four response variables from generalized linear models (GLMs), not from the direct observations. The BBd is thereby defined as the DOY in which a tree specific GLM estimate of the ratio of buds in or above stage two (against sampling DOY) crossed 0.5. Analogously, LUst and LUen are derived from a second tree specific GLM of the mean bud stage against the sampling DOY and defined as the DOY at which the model crossed the 0.05 (LUst) and 0.95 (Luen) threshold respectively. As many more observations were conducted before the initiation of bud development than after its completion, in combination with the inherent properties of a GLM, Luen was indirectly influenced by LUst, leading to a higher variability of the latter. The duration (Ludur in days) was then calculated as the tree specific interval between LUst and Luen.

Regression. The aim of the regression analysis was to establish if, and how much, the average spring phenology timings (as expressed via BBd, LUst, Luen and Ludur in separate models) differed between the two different stand species richness configurations (monospecific vs. 4-species stands), while accounting for potential differences between tree species and/or years. To focus on the effect of species richness configuration, we therefore created individual linear models for each combination of response variable (BBd, LUst, Luen, Ludur), species (Betula, Fagus, Quercus, Tilia) and year (2021, 2022) with 'Species richness' as the sole explanatory variable, resulting in 32 separate models. All models were ensured to fulfil linear model assumptions via diagnostic plots. To identify generalized effects and potential interactions, we moreover configured separate models for each response variable (BBd, LUst, Luen and Ludur), which included all aforementioned explanatory variables ('Species richness', 'Species' and 'Year') and initially all possible interactions. The factor 'Year' represents a proxy for the entirety of weather differences between the two years, as the overall macroclimatic weather conditions were the same for each stand given that they were all located in the same area (see 'Study site'). These models were constructed in the more flexible GAMLSS framework (Generalized Additive Models for Location, Scale and Shape; version 5.4.1; Rigby and Stasinopoulos 2005). Not only does GAMLSS allow the application of a wide range of distributions to complex data, it furthermore allows us to analyse potential effects on the variability, as well as on the mean, and was successfully applied to similar studies within the field (Mariën et al., 2022). These models were constructed and reduced until they contained only significant terms (minimal adequate models; terms reduced from the model due to non-significance are labelled "nS" in figures) following the recommendations by the GAMLSS creators (see details in supporting information, Methods M1; Rigby and Stasinopoulos 2005).

Variable importance. To further evaluate the results of the regression analysis, we investigated the relative importance of a set of stand factors that are known to affect phenology and additionally include species richness. For each response variable (BBd, LUst, Luen and Ludur), the relative importance of the following explanatory variables was evaluated via separate randomForest models for each species (Betula, Fagus, Quercus and Tilia). Beside the factors 'Species richness' and 'Year', which were also included in the regressions, this analysis additionally included parameters describing individual dendrometric characteristics, i.e. the continuous variables tree stem diameter at breast height DBH (in cm) and tree height (in cm), as well as the APA (in m^2). Missing measurements due to measurement complications for DBH (8/220 trees) and Height (2/220 trees) were replaced with the species' median. Sixteen separate randomForest models, one for each species and phenological timing combination, with 500 trees each, were constructed (R package 'randomForest', version 4.7.1.1). As these randomForest models were intended for the variable importance only, and not for classification purposes, all data was used without evaluation via a test dataset, thereby giving the algorithm the maximum of data and increasing the explained variance. The R^2 of the models ranged from 0.4 to 0.83 (the individual R^2 of each model can be found directly in the figure). The relative variable importance was extracted from each randomForest model via the *explain* function from each cycle of 100 permutations based on root-mean-square-loss to establish confidence intervals via the *model parts* function (R package 'DALEX' version 2.4.2, Biecek 2018).

Results

Tree dendrometry, macroclimate and microclimate. Both tree diameter and height showed species-specific differences between the monospecific and 4-species stands (Fig. 1a,b). The DBH of Betula was larger in the 4-species stands ($p< 0.001$), while the other three species showed no significant differences in DBH between species richness configurations (Fig. 1a). We found the inverse for tree height, with Betula showing no significant differences in height between species richness configurations, while Fagus, Quercus and Tilia were significantly higher in the 4-species configuration (all $p< 0.001$, Fig. 1b). The general macroclimatic conditions (established from the open field sensor) in spring were overall considerably warmer in 2022 than in 2021, especially at the start of January, the first half of February and from April to June (5-10°C difference, Fig. 3c). In accordance with this, the macroclimatic total number of chilling days (NCD) was considerably higher in 2021 (41 days) than in 2022 (22 days; Supporting Information Tab. S2). Under these shared macroclimatic conditions, the microclimatic conditions until budburst differed considerably between the species richness configurations (Fig. 2a, background bar graphs). The accumulated GDDair at budburst (BBd) was larger in the 4-species stands than in the monospecific stands only for Betula (4-5%), while the opposite (higher GDDair in the monospecific stands) was recorded for Quercus (2-4%), Tilia (5-8%) and especially Fagus (11-13%, Fig. 2a). For Quercus, however, GDDair was temporarily larger in the 4-species stands during January and February of both years (Fig. 2a). The difference in GDDair between the two species richness configurations was overall slightly larger in 2021 than in 2022 (Fig. 2a). Similarly, the soil temperature at BBd was also in all but two cases lower in the 4-species stands than in the monospecific stands but with even larger differences between species (2021/2022: Betula +2%/-9%, Fagus -29%/-7%, Quercus -16%/+2%, Tilia -12%/-11%), especially in 2021 and between DOY 100 and 150 (which encompasses the BBd of all species besides Betula). The difference in soil temperature to the overall average soil temperature from all stations during this period was significantly lower in the 4-species stands than in the monospecific stands for all species (all $p \leq 0.001$, Fig. 1c). This difference was considerably higher in 2021 for all species but Betula, for which the inverse was true (Fig. 1c). NCDair showed only minor differences between species richness configurations, with only the 4-species stands of Tilia in 2021 receiving two more chilling days before the 31st of March (Fig. 1d).

Species richness effects on individual species - BBd. Our analysis showed that the degree of difference in BBd between the monospecific and 4-species configurations depended on the species and the year (Fig. 2a, model evaluators in supporting information Table S3). We found no significant responses to stand species richness in Betula, Quercus and Tilia. For Fagus, BBd was significantly advanced in the 4-species stands, by ≈5 days in 2021 ($p= 0.004$) and by ≈4 days in 2022 ($p= 0.022$), indicating this effect to be stable enough to have a significant advancement in BBd in the 4-species stands independent of the year and thereby the weather conditions (Fig. 2a, see significant interaction ‘Species’:’Species richness’ in Tab. S1). An additional interaction ‘Species’:’Year’ shows the BBd of Fagus and Quercus (but not of Betula and Tilia) to be significantly advanced (independent of the stand species richness) in response to the spring conditions of 2022 in comparison to 2021 (Fig. 2a, Tab. S3).

Species richness effects on individual species - LUst, LUen and LUdur. For Fagus in 2021, LUst was marginally delayed ($p= 0.071$) while LUen was significantly advanced by ≈6 days ($p< 0.001$; Fig. 2a), leading to LUdur being overall ≈13 days shorter in the 4-species stands than in the monocultures ($p<$

0.001; Fig. 2b). In 2022, LUst in the 4-species stands switched to be significantly advanced (≈2 days, *p*= 0.045), while LUen was still slightly advanced (although marginally; *p*= 0.057, Fig. 2a). Because of this converse response of LUst in 2022, no significant difference in LUdur was observed in that year between species richness configurations for Fagus (LUdur *p*= 0.424; Fig. 2b). In summary, in 2021, Fagus in the 4-species stands showed a delayed onset, but much more rapid leaf unfolding than Fagus in the monoculture. In 2022, Fagus in the 4-species stands showed a uniform advancement in budburst timings, without differences in overall duration when compared to Fagus in the monoculture. In 2021, Quercus was similarly affected by the species richness configuration as Fagus, with LUdur in the 4-species stands being shortened (≈7 days, *p*< 0.001; Fig. 2b) and LUst being delayed (≈4 days, *p*= 0.003; Fig. 2a). However, no advancement of LUen was observed (Fig. 2a). In 2022, all effects of stand species richness on the budburst timings disappeared completely for Quercus (Fig. 2a). Tilia showed a variability about twice as high as all other species for LUst, LUen and LUdur, with no significant differences between species richness configurations (Fig. 2a, Tab. S1). Betula also showed little differences between species configurations besides a slightly advanced LUst in the 4-species stands in 2022 (≈3 days, *p*= 0.047; Fig. 2a). Furthermore, Betula was the only species that had a longer LUdur in 2022 (across species richness configuration), as indicated by the significant 'Species':'Year' interaction (Supporting Information, Tab. S1,S3).

Species richness effects across species. On average, and across species, LUen in the 4-species stands was advanced by ≈ 4 days (*p* = 0.002) and LUdur shortened by ≈ 6 days (*p* < 0.001) in the colder spring of 2021 (Fig. 3a,b). This stands in contrast to the warmer 2022, when the 4-species stands showed no significant advance in any budburst timing, although the average budburst timings were consistently earlier in the 4-species stands (LUst ≈ 2 days, BBd ≈ 3 days, LUen ≈ 4 days; Fig. 3d). The pattern in LUen and LUdur in 2021 was mostly driven by Fagus and Quercus (the interaction was non-significant when these two species were removed; Supporting Information, Tab. S4); this general trend can also be found (though not significant) in the other two species as well (Fig. 2a).

Relative variable importance. The factor 'Year', a proxy for the different spring weather conditions, as the general macroclimatic weather conditions were the same for each tree and stand, was by far the most important predictor for all considered response variables in Fagus and Quercus, with the notable exception of BBd in the former and LUst in the latter (Fig. 4a,b). In contrast, the factor 'Year' varied considerably in importance for Betula, from high in LUst to negligible for LUen, while being the universally least important factor for Tilia (Fig. 4). Beyond the factor 'Year', stand species richness was the most important variable for LUst in Quercus and for LUdur in both Quercus and Fagus (SpRich in Fig. 4a). Overall, 'Species richness' was generally less important than dendrometric parameters or APA in Betula and Tilia. For Fagus and Quercus, stand species richness was comparable in importance to both dendrometric parameters and APA and increased in importance with the progression of leaf unfolding for Fagus (LUst < BBd< LUen), while it decreased in importance for Quercus (LUst > BBd > LUen).

Discussion

We found that that stands of varying species richness configurations can indeed show differences in their spring phenology, albeit only in some species and with important variation between years. Contrarily to our hypothesis, when phenological timings differed between species richness configurations, they were generally advanced in the 4-species stands both within and across species (the budburst onset, LUst, of Quercus in 2021 being the only exception; Fig. 2a,3a). The effect of increased species richness was strongest in Fagus, which showed consistent advancement of budburst within the 4-species stands independent of the year (Fig. 2a). Quercus displayed a pattern comparable to Fagus in 2021, but the effect of species richness on budburst timing completely disappeared in 2022, when the overall budburst period was significantly shorter for the species (Fig. 2a,b). Except for the slight advancement of LUst in 2022 for Betula, we found no indications that stand species richness considerably affected the spring phenology of either Betula or Tilia.

We suggested two non-mutually exclusive microclimatic drivers that could lead to differences in the spring phenology of stands with varying species richness composition: microclimatic temperature and/or light availability. Our study strongly suggests temperature not to be the driver behind the differences we observed. Recent studies highlighted the importance of the temperature microclimate for the intra-individual bud development of seedlings with buds that experience higher temperature bursting earlier, and the leaves developing faster, than less warmed buds (Vitasse et al., 2021). We found a different picture in our considerably more mature trees in plantations. The differences in the chilling received (NCDair) was negligible between the two species richness configurations (only higher in the 4-species stands for Tilia in 2021) without any consequences for budburst timings (Fig. 1d,2a). Regarding the accumulated growing degree days (GDDair), Fagus and Tilia in the 4-species stands of received consistently between 5% to 13% less temperature forcing than in their respective monospecific stands. This nevertheless had no effect on the spring phenology of Tilia and translated into significantly earlier phenological timings (and even a significantly shorter overall budburst period) in Fagus (Fig. 2a,b). Furthermore, Betula (the only species accumulating more GDDair in the 4-species stands) only showed a significant advancement of LUst and only so in 2022 (Fig. 2a).

On the other hand, late-successional, shade-tolerant species like Fagus show an advanced budbreak when growing in shadowy conditions (Piper and Fajardo 2023). Therefore, more shadowy conditions in the 4-species configuration could have driven the advances in spring phenology presented here, and even be strong enough to overcome ≈12% less temperature forcing received by Fagus in the more shadowy 4-species stands (Fig. 2a). This could also serve as an explanation of the differences between Fagus and Quercus. The more shadowy conditions in the 4-species stands after leaf development starts speed up the development for Fagus in the later stages (earlier BBd, LUst). For more light-demanding Quercus, on the other hand, these more shadowy conditions delayed budburst onset in 2021 (Fig. 2a). We find several indirect indications for the differences in light conditions between the two species richness configurations. Firstly, the trees of Fagus, Quercus and Tilia were significantly higher in the 4-species stands, indicating a much higher competition for light there (Fig. 1b). Secondly, the soil temperature in the period from DOY 100 – 150, after the phenological timings of Betula but before the timings of the other three species, was consistently significantly lower in the 4-species stands, which can be an indication of less sunlight reaching the ground (Fig. 1c). Thirdly, point measurements of canopy closure at the temperature stations in August 2019 show ≈17% less canopy closure in the monospecific stands than in the 4-species stands (Zhang et al., 2022).

We also found considerable differences in budburst phenology between the two observed years, both across and between species. The macroclimatic conditions were considerably colder, and the temperature progression much less even, in 2021 than in 2022 (Fig. 3). While this had only minor effects on the budburst timings of Betula and Tilia, it dramatically shifted phenological timings and shortened the overall budburst period in both Fagus and Quercus (Fig. 2, Interactions Tab. S1). At the same time we observed, mostly driven by Fagus and Quercus, a significant advancement of LUen and a significant shortening of LUdur across species identity in 2021, but not in 2022. We suggest that in the generally warmer spring conditions of 2022, the overall speed of leaf unfolding was mainly driven by the superordinate macroclimatic temperature, with the differences induced by stand species richness being negligible in comparison (Fig. 2,3). When the macroclimatic warming is slower, however, as in 2021, the trend of budburst advancement in 4-species stands is stronger (for Fagus) or becomes significant (for Quercus) as the slower temperature forcing allows microclimatic differences based on species richness to be emphasized more. This can even be strong enough, at least for the completion of leaf development (LUen) and the overall duration (LUdur), to be valid across species (Fig. 3). An indication of the generality of this response can be found in a remote sensing based study of US wetlands that showed phenological variability to be lower in warmer and wetter climates (Dronova et al., 2022). It has also earlier been demonstrated that in rapidly warming springs, the variability of budburst timings is lower than in colder springs (Denéchère et al., 2019). Within our study, we find a significantly lower timing variability in 2022 for Tilia, but a higher variability in LUst and BBd for Betula, indicating species-specific differences not found by Denéchère *et al.* (sigma-interaction ‘Year’:’Species’, supplement Data, Tab. S2).

Regarding the relative importance of species richness on budburst timing, we found it to be rather moderate when compared to the main driver, i.e. annual macroclimatic weather conditions (see above), but mostly on par with other factors known to affect inter-individual budburst differences like DBH and height (Marchand et al., 2020), albeit with species-specific differences. While the species richness configuration was generally not very important for Betula and Tilia, it was relatively important for budburst initiation (LUst) in Quercus, the completion of leaf development in Fagus (LUen) and especially the overall budburst period duration (LUdur) in both species.

Even if species- and year-specific, the effect of species richness on spring phenology has important implications. For example, it implies that the maximum photosynthetic activity can be mobilized earlier in stands with a higher tree biodiversity, with multiple potential consequences on tree growth, C sequestration and water consumption (Lian et al., 2020; Piao et al., 2019). Moreover, evaluating the effects of non-climatic co-drivers of phenology opens different opportunities beyond increasing our understanding of phenology dynamics and its feedback to climate. Current large-scale predictive phenological models integrate mixed forests based on the relative species abundances and their phenology in a monospecific environment. Our study shows that this approach is flawed, as species differ in their response to a mixed species environment. Secondly, while the general macroclimatic conditions (e.g. site temperature) cannot be altered, non-climatic factors like species composition, tree density, etc. could be manipulated through selected forest management strategies. This would open a new dimension for forest management: modifying the growing season length and thereby potentially improving biomass production as well as atmospheric carbon removal.

Our results also highlight the importance of understanding and characterizing spring phenology as a process (thus considering start, peak and end of the process) rather than a sudden, date-bound occurrence. In fact, the different spring phenology timings, as well as the overall duration, provide more information than only the budburst date and, moreover, they might be differently affected by any given factor, as seen in our case for the stand species richness.

Based on our results, we expect species richer tree stands to have slightly advanced budburst timing, not consistently at species level but well at stand level during a cold spring. More research will be needed to establish how different stand biodiversity mixtures affect species specific budburst timings and their variability. Furthermore, it should be evaluated how the observed effects of tree biodiversity shape the phenological response pattern at a tree level in the long term and if it can be integrated into forest management, as well as restoration, strategies.

Supporting Information

Fig. S4: Exemplary planting scheme of monospecific and 4-species stands at FORBIO Zedelgem.

Methods M1: Description of the modelling procedure for the four GAMLSS interaction models for BBd, LUst, LUen and LUdur.

Table S1: Terms, coefficients and model evaluators of the final GAMLSS interaction models for BBd, LUst, LUen and LUdur.

Table S2: Total numbers of GDDacc Air and NCD from each species richness configuration and the open field sensor.

Table S3: Estimates and test statistics for the 32 individual linear models for each combination of species (*B. pendula*, *F. sylvatica*, *Q. robur*, *T. cordata*) and timing variable (LUst, BBd, LUen, LUdur).

Table S4: Estimates and test statistics for GAMLSS models regarding the 'Year':'Species richness' interaction for Betula and Tilia only.

Bibliography

Arend, M., Gessler, A., & Schaub, M. (2016). The influence of the soil on spring and autumn phenology in European beech. *Tree Physiology*, *36*(1), 78–85. https://doi.org/10.1093/treephys/tpv087

Basler, D., & Körner, C. (2014). Photoperiod and temperature responses of bud swelling and bud burst in four temperate forest tree species. *Tree Physiology*, *34*(4), 377–388. https://doi.org/10.1093/treephys/tpu021

Baumgarten, F., Zohner, C. M., Gessler, A., & Vitasse, Y. (2021). Chilled to be forced: The best dose to wake up buds from winter dormancy. *New Phytologist*, *230*(4), 1366–1377. https://doi.org/10.1111/nph.17270

Biecek, P. (2018). DALEX: Explainers for Complex Predictive Models in R. *Journal of Machine Learning Research*, *19*(84), 1–5.

De Frenne, P., Lenoir, J., Luoto, M., Scheffers, B. R., Zellweger, F., Aalto, J., Ashcroft, M. B., Christiansen, D. M., Decocq, G., De Pauw, K., Govaert, S., Greiser, C., Gril, E., Hampe, A., Jucker, T., Klinges, D. H., Koelemeijer, I. A., Lembrechts, J. J., Marrec, R., … Hylander, K. (2021). Forest microclimates and climate change: Importance, drivers and future research agenda. *Global Change Biology*, *27*(11), 2279–2297. https://doi.org/10.1111/gcb.15569

De Frenne, P., Rodríguez-Sánchez, F., Coomes, D. A., Baeten, L., Verstraeten, G., Vellend, M., Bernhardt-Römermann, M., Brown, C. D., Brunet, J., Cornelis, J., Decocq, G. M., Dierschke, H., Eriksson, O., Gilliam, F. S., Hédl, R., Heinken, T., Hermy, M., Hommel, P., Jenkins, M. A., … Verheyen, K. (2013). Microclimate moderates plant responses to macroclimate warming. *Proceedings of the National Academy of Sciences*, *110*(46), 18561–18565. https://doi.org/10.1073/pnas.1311190110

Denéchère, R., Delpierre, N., Apostol, E. N., Berveiller, D., Bonne, F., Cole, E., Delzon, S., Dufrêne, E., Gressler, E., Jean, F., Lebourgeois, F., Liu, G., Louvet, J.-M., Parmentier, J., Soudani, K., & Vincent, G. (2019). The within-population variability of leaf spring and autumn phenology is influenced by temperature in temperate deciduous trees. *International Journal of Biometeorology*. https://doi.org/10.1007/s00484-019-01762-6

Dronova, I., Taddeo, S., & Harris, K. (2022). Plant diversity reduces satellite-observed phenological variability in wetlands at a national scale. *Science Advances*, *8*(29), eabl8214. https://doi.org/10.1126/sciadv.abl8214

Forrester, D. I. (2014). The spatial and temporal dynamics of species interactions in mixed-species forests: From pattern to process. *Forest Ecology and Management*, *312*, 282–292. https://doi.org/10.1016/j.foreco.2013.10.003

Glatthorn, J. (2021). A spatially explicit index for tree species or trait diversity at neighborhood and stand level. *Ecological Indicators*, *130*, 108073. https://doi.org/10.1016/j.ecolind.2021.108073

Gričar, J., Lavrič, M., Ferlan, M., Vodnik, D., & Eler, K. (2017). Intra-annual leaf phenology, radial growth and structure of xylem and phloem in different tree parts of Quercus pubescens. *European Journal of Forest Research*, *136*(4), 625–637. https://doi.org/10.1007/s10342-017-1060-5

Laube, J., Sparks, T. H., Estrella, N., Höfler, J., Ankerst, D. P., & Menzel, A. (2014). Chilling outweighs photoperiod in preventing precocious spring development. *Global Change Biology*, *20*(1), 170–182. https://doi.org/10.1111/gcb.12360

Leuschner, C., & Meier, I. C. (2018). The ecology of Central European tree species: Trait spectra, functional trade-offs, and ecological classification of adult trees. *Perspectives in Plant Ecology, Evolution and Systematics*, *33*, 89–103. https://doi.org/10.1016/j.ppees.2018.05.003

Lian, X., Piao, S., Li, L. Z. X., Li, Y., Huntingford, C., Ciais, P., Cescatti, A., Janssens, I. A., Peñuelas, J., Buermann, W., Chen, A., Li, X., Myneni, R. B., Wang, X., Wang, Y., Yang, Y., Zeng, Z., Zhang, Y., & McVicar, T. R. (2020). Summer soil drying exacerbated by earlier spring greening of northern vegetation. *Science Advances*, *6*(1), eaax0255. https://doi.org/10.1126/sciadv.aax0255

Marchand, L. J., Dox, I., Gričar, J., Prislan, P., Leys, S., Van den Bulcke, J., Fonti, P., Lange, H., Matthysen, E., Peñuelas, J., Zuccarini, P., & Campioli, M. (2020). Inter-individual variability in spring phenology of temperate deciduous trees depends on species, tree size and previous year autumn phenology. *Agricultural and Forest Meteorology*, *290*, 108031. https://doi.org/10.1016/j.agrformet.2020.108031

Mariën, B., Papadimitriou, D., Kotilainen, T., Zuccarini, P., Dox, I., Verlinden, M., Heinecke, T., Mariën, J., Willems, P., Decoster, M., Gascó, A., Lange, H., Peñuelas, J., & Campioli, M. (2022). Timing leaf senescence: A generalized additive models for location, scale and shape approach. *Agricultural and Forest Meteorology*, *315*, 108823. https://doi.org/10.1016/j.agrformet.2022.108823

Meng, F. D., Jiang, L. L., Zhang, Z. H., Cui, S. J., Duan, J. C., Wang, S. P., Luo, C. Y., Wang, Q., Zhou, Y., Li, X. E., Zhang, L. R., Li, B. W., Dorji, T., Li, Y. N., & Du, M. Y. (2017). Changes in flowering functional group affect responses of community phenological sequences to temperature change. *Ecology*, *98*(3), 734–740. https://doi.org/10.1002/ecy.1685

Menzel, A., Sparks, T. H., Estrella, N., Koch, E., Aasa, A., Ahas, R., Alm-Kübler, K., Bissolli, P., Braslavská, O., Briede, A., Chmielewski, F. M., Crepinsek, Z., Curnel, Y., Dahl, Å., Defila, C., Donnelly, A., Filella, Y., Jatczak, K., Måge, F., … Zust, A. (2006). European phenological response to climate change matches the warming pattern. *Global Change Biology*, *12*(10), 1969–1976. https://doi.org/10.1111/j.1365-2486.2006.01193.x

Menzel, A., Yuan, Y., Matiu, M., Sparks, T., Scheifinger, H., Gehrig, R., & Estrella, N. (2020). Climate change fingerprints in recent European plant phenology. *Global Change Biology*, *26*(4), 2599–2612. https://doi.org/10.1111/gcb.15000

Migliavacca, M., Sonnentag, O., Keenan, T. F., Cescatti, A., O’Keefe, J., & Richardson, A. D. (2012). On the uncertainty of phenological responses to climate change, and implications for a terrestrial biosphere model. *Biogeosciences*, *9*(6), 2063–2083. https://doi.org/10.5194/bg-9-2063-2012

Monahan, W. B., Rosemartin, A., Gerst, K. L., Fisichelli, N. A., Ault, T., Schwartz, M. D., Gross, J. E., & Weltzin, J. F. (2016). Climate change is advancing spring onset across the U.S. national park system. *Ecosphere*, *7*(10). https://doi.org/10.1002/ecs2.1465

Peaucelle, M., Janssens, I. A., Stocker, B. D., Descals Ferrando, A., Fu, Y. H., Molowny-Horas, R., Ciais, P., & Peñuelas, J. (2019). Spatial variance of spring phenology in temperate deciduous forests is constrained by background climatic conditions. *Nature Communications*, *10*(1), 5388. https://doi.org/10.1038/s41467-019-13365-1

Peñuelas, J., Rutishauser, T., & Filella, I. (2009). Phenology Feedbacks on Climate Change. *Science*, *324*(5929), 887–888. https://doi.org/10.1126/science.1173004

Piao, S., Liu, Q., Chen, A., Janssens, I. A., Fu, Y., Dai, J., Liu, L., Lian, X., Shen, M., & Zhu, X. (2019). Plant phenology and global climate change: Current progresses and challenges. *Global Change Biology*, *25*(6), 1922–1940. https://doi.org/10.1111/gcb.14619

Piper, F. I., & Fajardo, A. (2023). Carbon stress causes earlier budbreak in shade-tolerant species and delays it in shade-intolerant species. *American Journal of Botany*. https://doi.org/10.1002/ajb2.16129

R Core Team. (2019). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Richardson, A. D., Hufkens, K., Milliman, T., Aubrecht, D. M., Furze, M. E., Seyednasrollah, B., Krassovski, M. B., Latimer, J. M., Nettles, W. R., Heiderman, R. R., Warren, J. M., & Hanson, P. J. (2018). Ecosystem warming extends vegetation activity but heightens vulnerability to cold temperatures. *Nature*, *560*(7718), 368–371. https://doi.org/10.1038/s41586-018-0399-1

Richardson, A. D., Keenan, T. F., Migliavacca, M., Ryu, Y., Sonnentag, O., & Toomey, M. (2013). Climate change, phenology, and phenological control of vegetation feedbacks to the climate system. *Agricultural and Forest Meteorology*, *169*, 156–173. https://doi.org/10.1016/j.agrformet.2012.09.012

Rigby, R. A., & Stasinopoulos, D. M. (2005). Generalized additive models for location, scale and shape,(with discussion). *Applied Statistics*, *54*, 507–554.

Rodríguez De Prado, D., Riofrío, J., Aldea, J., McDermott, J., Bravo, F., & Herrero De Aza, C. (2022). Species Mixing Proportion and Aridity Influence in the Height–Diameter Relationship for

Different Species Mixtures in Mediterranean Forests. *Forests*, *13*(1), 119. https://doi.org/10.3390/f13010119

Thuiller, W., Lavorel, S., Sykes, M. T., & Araújo, M. B. (2006). Using niche-based modelling to assess the impact of climate change on tree functional diversity in Europe. *Diversity and Distributions*, *12*(1), 49–60. https://doi.org/10.1111/j.1366-9516.2006.00216.x

Vander Mijnsbrugge, K., & Moreels, S. (2020). Varying Levels of Genetic Control and Phenotypic Plasticity in Timing of Bud Burst, Flower Opening, Leaf Senescence and Leaf Fall in Two Common Gardens of Prunus padus L. *Forests*, *11*(10), 1070. https://doi.org/10.3390/f11101070

Vander Mijnsbrugge, K., Turcsán, A., Depypere, L., & Steenackers, M. (2016). Variance, Genetic Control, and Spatial Phenotypic Plasticity of Morphological and Phenological Traits in Prunus spinosa and Its Large Fruited Forms (P. x fruticans). *Frontiers in Plant Science*, *7*. https://doi.org/10.3389/fpls.2016.01641

Verheyen, K., Ceunen, K., Ampoorter, E., Baeten, L., Bosman, B., Branquart, E., Carnol, M., De Wandeler, H., Grégoire, J.-C., Lhoir, P., Muys, B., Setiawan, N. N., Vanhellemont, M., & Ponette, Q. (2013). Assessment of the functional role of tree diversity: The multi-site FORBIO experiment. *Plant Ecology and Evolution*, *146*(1), 26–35. https://doi.org/10.5091/plecevo.2013.803

Vitasse, Y., Baumgarten, F., Zohner, C. M., Kaewthongrach, R., Fu, Y. H., Walde, M. G., & Moser, B. (2021). Impact of microclimatic conditions and resource availability on spring and autumn phenology of temperate tree seedlings. *New Phytologist*, *232*(2), 537–550. https://doi.org/10.1111/nph.17606

Vitasse, Y., Delzon, S., Dufrêne, E., Pontailler, J.-Y., Louvet, J.-M., Kremer, A., & Michalet, R. (2009). Leaf phenology sensitivity to temperature in European trees: Do within-species populations exhibit similar responses? *Agricultural and Forest Meteorology*, *149*(5), 735–744. https://doi.org/10.1016/j.agrformet.2008.10.019

Vitasse, Y., Hoch, G., Randin, C. F., Lenz, A., Kollas, C., Scheepens, J. F., & Körner, C. (2013). Elevational adaptation and plasticity in seedling phenology of temperate deciduous tree species. *Oecologia*, *171*(3), 663–678. https://doi.org/10.1007/s00442-012-2580-9

Wolf, A. A., Zavaleta, E. S., & Selmants, P. C. (2017). Flowering phenology shifts in response to biodiversity loss. *Proceedings of the National Academy of Sciences*, *114*(13), 3463–3468. https://doi.org/10.1073/pnas.1608357114

Zhang, S., Landuyt, D., Verheyen, K., & De Frenne, P. (2022). Tree species mixing can amplify microclimate offsets in young forest plantations. *Journal of Applied Ecology*, *59*(6), 1428–1439. https://doi.org/10.1111/1365-2664.14158

Figure legends

Figure 1: The differences between the two observed stand species richness configurations monospecific (Ms) and 4-species (4sp) for each species (*Betula pendula*, *Fagus sylvatica*, *Quercus robur*, *Tilia cordata*) in a: diameter at breast height (DBH, box plot in cm); b: tree height (Height, box plot in cm); c: difference in daily soil temperature to the overall average soil temperature from Day-Of-Year (DOY) 100 to DOY 150 (Δ Soil Temp. box plot in C°) and d: number of cold days as accumulated over each DOY (the sum of days with a mean air temperature below 5°C, NCDair, bar plot in number of days). The number in the box plots (a-c) shows the p-value of individual linear models with stand species richness configuration (Monospecific vs. 4-species) as the sole explanatory variable. In the bar graphs (d) the overlap between both species richness configurations is coloured grey, with a surplus shown at the fringe via colour (Ms vs. 4sp). Marks indicate the level of significance; ` , *p*= 0.1-0.05; * , *p*= 0.05-0.01; ** , *p*= 0.01-0.001; *** , *p*< 0.001.

Figure 2: Observed average timing variables and standard deviation of a: the start (LUst, in DOY), 50%-point (BBd, in DOY), end (LUen, in DOY) of the budburst period and b: overall duration (LUdur, in number of days), separated by tree species (*Betula pendula*, *Fagus sylvatica*, *Quercus robur*, *Tilia*

cordata), year (2021, 2022) and stand species richness configuration (Ms, 4sp). The number between the bars show the p-values of the difference between Monospecific (Ms) and 4-species (4sp) stands derived from individual linear models for each phenological timing variable (in colour) on subsets of the data split by species and year with stand species richness configuration (Monospecific vs. 4-species) as the sole explanatory variable. The bar graph in the background shows the percentual difference in accumulated growing degree days (GDDair) between the two species richness configurations in the given year. The dotted line (+10% GDDair) indicates the point at which one species richness configuration (colour coded; Ms, 4sp) has accumulated 10% more GDDair than the other configuration. Marks indicate the level of significance; ` , *p*= 0.1-0.05; * , *p*= 0.05-0.01; ** , *p*= 0.01-0.001; *** , *p*< 0.001.

Figure 3: The start (LUst), 50%-point (BBd) and end (LUen) of the budburst period in the monospecific (Ms, solid line) and 4-species (4sp, dashed line), including the associated p-value (as returned by the individual GAMLSS interaction models for each timing variable) in relation to the average macroclimatic air temperature measured in an open field close to the forest stands (in C°) in 2021 (upper panel a) and 2022 (lower panel d). The central panel (c) shows the net difference in air temperature (mirrored around 0 for better visibility, see y-axis markings) between the two years (dark-red, higher in 2022; light-blue, higher in 2021). The panels to the right (b,e) show the average overall duration (LUdur) of the budburst period in both species richness configurations (Ms, 4sp). The inserted pictures show the bud-stage associated with the timing variable (frame colour) in Betula (LUst = Stage 0-1; BBd = Stage 2; LUen = Stage 4). Marks indicate the level of significance; ` , *p*= 0.1-0.05; * , *p*= 0.05-0.01; ** , *p*= 0.01-0.001; *** , *p*< 0.001; nS, reduced from final model due to non-significance.

Figure 4: Relative variable importance for the a: start (LUst); b: 50%-point (BBd); c: end (LUen) and d: overall duration (LUdur) of the budburst period and the associated confidence interval (grey box plot bars) for each species (*Betula pendula*, *Fagus sylvatica*, *Quercus robur*, *Tilia cordata*) as returned by individual randomForest models (R^2 in first line) over 100 permutations (x-axis is the 1–AUC, Area-

under-curve, of the loss function). The variable included are diameter at breast height (DBH), tree height (Height), Area-potentially-available (APA) and species richness configuration (SpRich). The relative importance of the factor `Year` is not shown to improve visibility of the other factors but its value of 1-AUC is indicated in the upper right corner with a colour coding expressing the importance in relation to the highest scoring factor shown (red, >200%; black, 50-200%; blue, <50%).

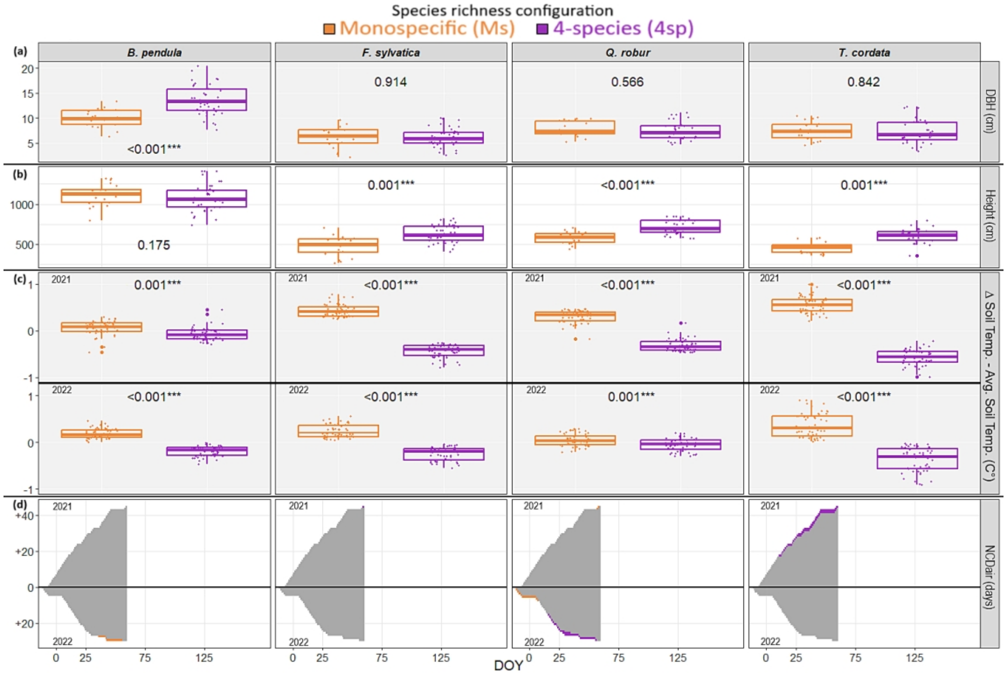

Figure 1: The differences between the two observed stand species richness configurations monospecific (Ms) and 4-species (4sp) for each species (Betula pendula, Fagus sylvatica, Quercus robur, Tilia cordata) in a: diameter at breast height (DBH, box plot in cm); b: tree height (Height, box plot in cm); c: difference in daily soil temperature to the overall average soil temperature from Day-Of-Year (DOY) 100 to DOY 150 (Δ Soil Temp. box plot in C°) and d: number of cold days as accumulated over each DOY (the sum of days with a mean air temperature below 5°C, NCDair, bar plot in number of days). The number in the box plots (a-c) shows the p-value of individual linear models with stand species richness configuration (Monospecific vs. 4-species) as the sole explanatory variable. In the bar graphs (d) the overlap between both species richness configurations is coloured grey, with a surplus shown at the fringe via colour (Ms vs. 4sp). Marks indicate the level of significance; ` , p= 0.1-0.05; * , p= 0.05-0.01; ** , p= 0.01-0.001; *** , p< 0.001.

126x85mm (300 x 300 DPI)

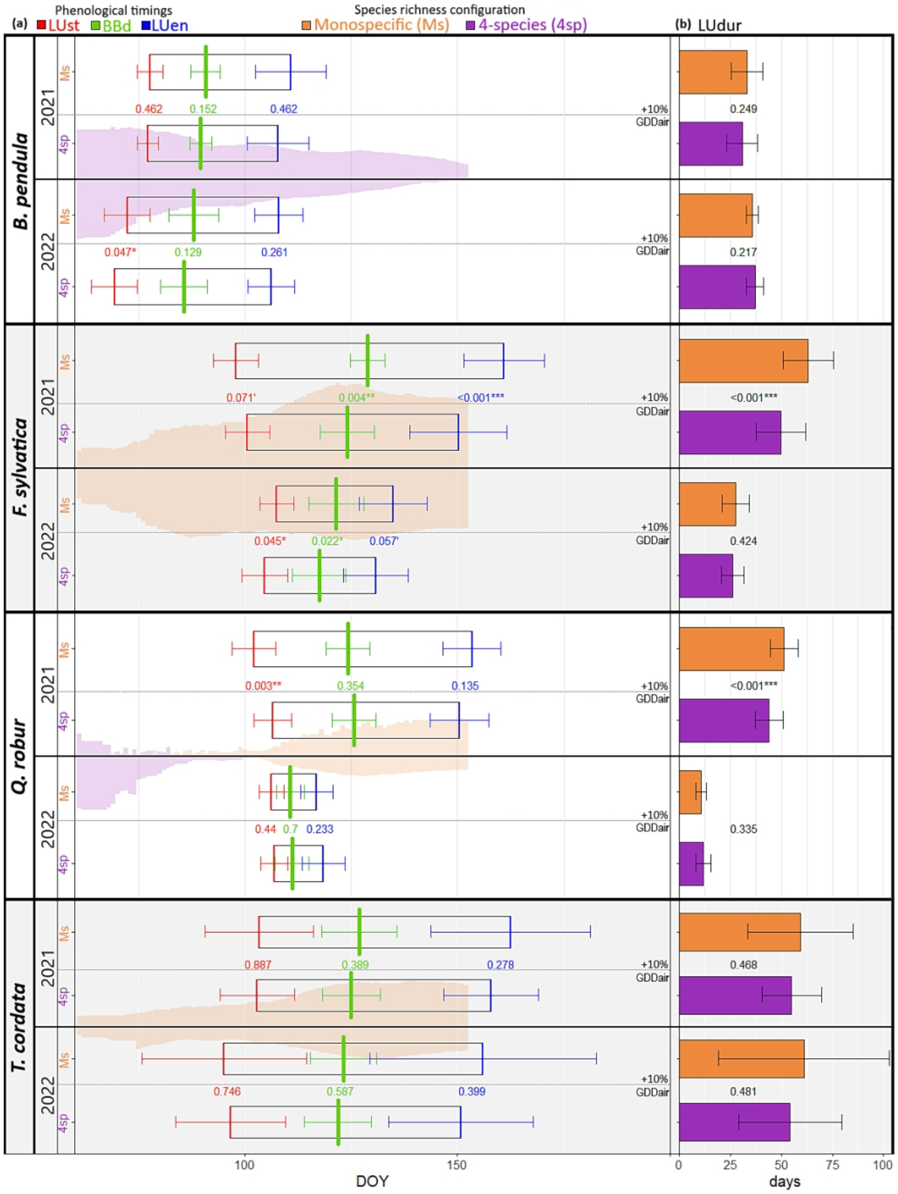

Figure 2: Observed average timing variables and standard deviation of a: the start (LUst, in DOY), 50%-point (BBd, in DOY), end (LUen, in DOY) of the budburst period and b: overall duration (LUdur, in number of days), separated by tree species (Betula pendula, Fagus sylvatica, Quercus robur, Tilia cordata), year (2021, 2022) and stand species richness configuration (Ms, 4sp). The number between the bars show the p-values of the difference between Monospecific (Ms) and 4-species (4sp) stands derived from individual linear models for each phenological timing variable (in colour) on subsets of the data split by species and year with stand species richness configuration (Monospecific vs. 4-species) as the sole explanatory variable. The bar graph in the background shows the percentual difference in accumulated growing degree days (GDDair) between the two species richness configurations in the given year. The dotted line (+10% GDDair) indicates the point at which one species richness configuration (colour coded; Ms, 4sp) has accumulated 10% more GDDair than the other configuration. Marks indicate the level of significance; ` , p= 0.1-0.05; * , p= 0.05-0.01; ** , p= 0.01-0.001; *** , p< 0.001.

101x134mm (300 x 300 DPI)

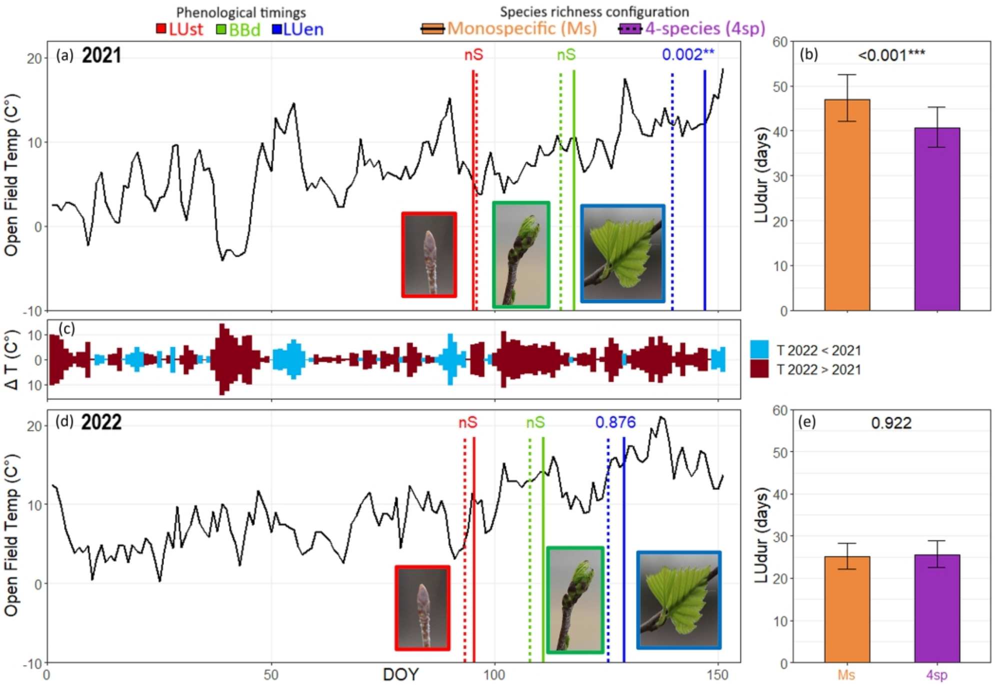

Figure 3: The start (LUst), 50%-point (BBd) and end (LUen) of the budburst period in the monospecific (Ms, solid line) and 4-species (4sp, dashed line), including the associated p-value (as returned by the individual GAMLSS interaction models for each timing variable) in relation to the average macroclimatic air temperature measured in an open field close to the forest stands (in C°) in 2021 (upper panel a) and 2022 (lower panel d). The central panel (c) shows the net difference in air temperature (mirrored around 0 for better visibility, see y-axis markings) between the two years (dark-red, higher in 2022; light-blue, higher in 2021). The panels to the right (b,e) show the average overall duration (LUdur) of the budburst period in both species richness configurations (Ms, 4sp). The inserted pictures show the bud-stage associated with the timing variable (frame colour) in Betula (LUst = Stage 0-1; BBd = Stage 2; LUen = Stage 4). Marks indicate the level of significance; ` , p= 0.1-0.05; * , p= 0.05-0.01; ** , p= 0.01-0.001; *** , p< 0.001; nS, reduced from final model due to non-significance.

244x166mm (300 x 300 DPI)

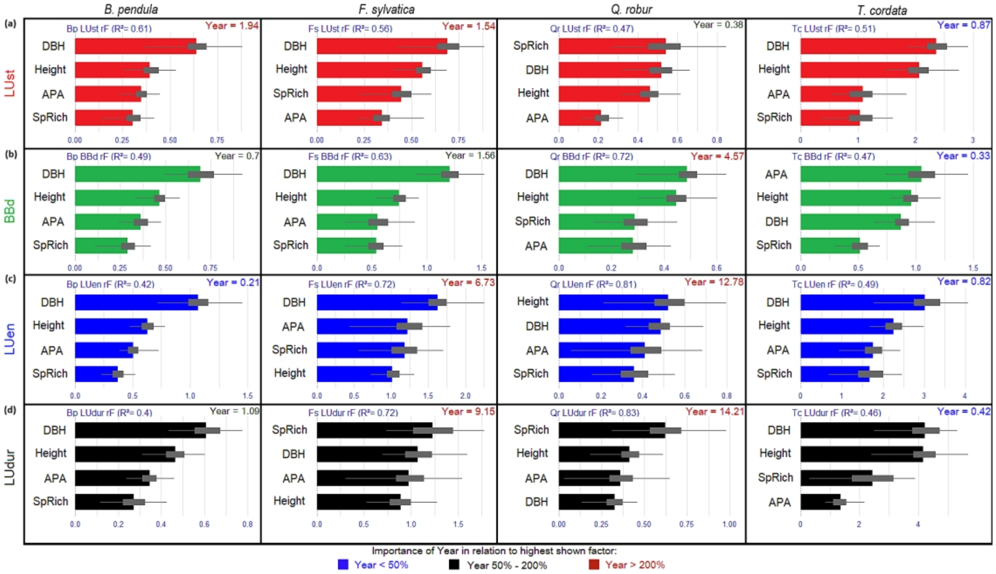

Figure 4: Relative variable importance for the a: start (LUst); b: 50%-point (BBd); c: end (LUen) and d: overall duration (LUdur) of the budburst period and the associated confidence interval (grey box plot bars) for each species (Betula pendula, Fagus sylvatica, Quercus robur, Tilia cordata) as returned by individual randomForest models (R² in first line) over 100 permutations (x-axis is the 1–AUC, Area-under-curve, of the loss function). The variable included are diameter at breast height (DBH), tree height (Height), Area-potentially-available (APA) and species richness configuration (SpRich). The relative importance of the factor `Year` is not shown to improve visibility of the other factors but its value of 1-AUC is indicated in the upper right corner with a colour coding expressing the importance in relation to the highest scoring factor shown (red, >200%; black, 50-200%; blue, <50%).

148x85mm (300 x 300 DPI)