Patron: Her Majesty The Queen

ROTHAMSTED RESEARCH

Rothamsted Research
Harpenden, Herts, AL5 2JQ

Telephone: +44 (0)1582 763133
Web: http://www.rothamsted.ac.uk/

Rothamsted Repository Download

A - Papers appearing in refereed journals

Metcalfe, H., Milne, A. E., Deledalle, F. and Storkey, J. 2020. Using Functional Traits to Model Annual Plant Community Dynamics. *Ecology.* https://doi.org/10.1002/ecy.3167

The publisher's version can be accessed at:

- https://doi.org/10.1002/ecy.3167

The output can be accessed at: https://repository.rothamsted.ac.uk/item/980x3/using-functional-traits-to-model-annual-plant-community-dynamics.

© 26 August 2020, Please contact library@rothamsted.ac.uk for copyright queries.

27/08/2020 11:08 repository.rothamsted.ac.uk library@rothamsted.ac.uk

Rothamsted Research is a Company Limited by Guarantee
Registered Office: as above. Registered in England No. 2393175.
Registered Charity No. 802038. VAT No. 197 4201 51.
Founded in 1843 by John Bennet Lawes.

Accepted Article

Running head: Using traits to model plant communities

Using functional traits to model annual plant community dynamics

HELEN METCALFE [1], ALICE E. MILNE, FLORENT DELEDALLE, JONATHAN STORKEY

Sustainable Agricultural Sciences, Rothamsted Research, West Common, Harpenden ,Hertfordshire, AL5 2JQ

[1]Corresponding author e-mail: helen.metcalfe@rothamsted.ac.uk

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/ECY.3167

This article is protected by copyright. All rights reserved

Accepted Article

Abstract

Predicting the response of biological communities to changes in the environment or management is a fundamental pursuit of community ecology. Meeting this challenge requires the integration of multiple processes: habitat filtering, niche differentiation, biotic interactions, competitive exclusion, and stochastic demographic events. Most approaches to this long-standing problem focus either on the role of the environment, using trait-based filtering approaches, or on quantifying biotic interactions with process-based community dynamics models. We introduce a novel approach that uses functional traits to parametrise a process-based model. By combining the two approaches we make use of the extensive literature on traits and community filtering as a convenient means of reducing the parametrisation requirements of a complex population dynamics model whilst retaining the power to capture the processes underlying community assembly. Using arable weed communities as a case study, we demonstrate that this approach results in predictions that show realistic distributions of traits and that trait selection predicted by our simulations is consistent with in-field observations. We demonstrate that trait-based filtering approaches can be combined with process-based models to derive the emergent distribution of traits. While initially developed to predict the impact of crop management on functional shifts in weed communities, our approach has the potential to be applied to other annual plant communities if the generality of relationships between traits and model parameters can be confirmed.

Key words: arable weeds, annual plants, community dynamics, ecological function, environmental filtering, functional diversity, functional traits, population dynamics modelling

This article is protected by copyright. All rights reserved

Accepted Article

Introduction

Predicting the assembly of biological communities and their resulting ecological function in different environments is a fundamental pursuit of community ecologists and has been characterised as the Holy Grail of ecology (Lavorel & Garnier, 2002). As society increasingly recognises the ecosystem services the biosphere contributes to human survival and well-being (Carpenter *et al.*, 2006) the need to understand the impact of changes in environment, land-use, or management on biological communities has become more urgent. Within this ecosystem service framework, it is more important to predict the impact of change on the functioning of the emergent biological community than on taxonomic composition (Fig. 1 A; Díaz *et al.*, 2007a). Meeting this challenge requires a unified approach that combines the theories of 1) habitat filtering and niche differentiation, 2) biotic interactions and competitive exclusion, and 3) stochastic demographic events (neutral theory). These processes, together with historical and evolutionary factors (which determine the regional species pool) all play a role in determining the local ecological community in a given environment (D'Amen *et al.*, 2017). Most approaches to this long-standing problem of predicting community composition at a given location focus either on the role of the environment, using trait-based filtering approaches (Fig. 1 B), or instead focus on quantifying biotic interactions with process-based community dynamics models (Fig. 1 C).

Trait-based filtering approaches that identify the abiotic and biotic filters acting on regionally available pools of species and determine those with favourable combinations of traits that can persist in a given habitat (Keddy, 1992) have now been applied across several taxa (e.g. plants (da Silveira Pontes *et al.*, 2010), arthropods (Braaker *et al.*, 2017), and bees (Hoiss *et al.*, 2012)), in a range of environments (e.g. tropics (Lebrija–Trejos *et al.*, 2010), streams (Poff, 1997), and rangelands (Bernard–Verdier *et al.*, 2012)) and across a number of different gradients (e.g. grazing (Díaz *et al.*, 2007b), geo-morphological (Gilardelli *et al.*, 2015), and aridity (Gross *et al.*, 2013)). However, all these studies rely on fitting statistical models to empirical relationships between environmental gradients and functional trait metrics and are, therefore, limited in their power to predict responses to environments with a novel combination of environmental variables. These models

This article is protected by copyright. All rights reserved

typically predict a convergence of trait attributes, as only species which are functionally similar will pass through successive filters on plant traits.

An alternative approach that avoids these limitations is to build process-based models of the responses of multiple interacting species to the environment. This more mechanistic approach involves describing key life-cycle processes mathematically, often from first principles, and can include spatially explicit individual based modelling approaches. Such process-based community dynamics models have also been widely developed to predict the community composition of a number of taxa (e.g. fish (Shin and Cury, 2001), coral (Langmead and Sheppard, 2004.), and trees (Purves *et al.*, 2008)), in a range of environments (e.g. tundra (Gilg *et al.*, 2003), freshwater lakes (van Nes *et al.*, 2002), and forests (Botkin, 1993)) and across a number of different environmental gradients (e.g. disturbance (Matsinos and Troumbis, 2002), fire (Thonickeet al., 2001), and nutrient limitation (Moore *et al.*, 2004)). In contrast to the trait-based filtering approach, these process-based community dynamics models often focus on biotic interactions which can be described mathematically and aim to predict relative species abundances in a more mechanistic way. By focussing on competitive processes, these models tend to select for species with divergent trait attributes in order to minimise overlapping resource use and competition, although practically this may not always be the observed outcome (Mayfield and Levine, 2010). Process-based community dynamics models often require extensive parametrisation to capture all the ecologically important processes. Each aspect of the life-cycle must be described mathematically for each simulated species, and where there is asymmetric competition for multiple resources this must also be quantified. As such, these models tend to be limited to a small pool of species and to a particular environment in which the parametrisation has been conducted (da Silveira Pontes *et al.*, 2010).

Ecological communities lie on a continuum: from those with strong biotic interactions to those where local interactions between individuals are weak and few (Cornell and Lawton, 1992) and models that aim to predict community dynamics should ideally avoid making prior assumptions on the dominant processes shaping that community. Several attempts have been made to include biotic processes into trait-based filtering models in order to simulate both the convergence and

This article is protected by copyright. All rights reserved

Accepted Article

Accepted Article

divergence of traits, and eliminate the need for *a-priori* knowledge of the dominant processes driving community dynamics at a given location. For example, Shipley *et al.* (Maxent, 2006) and later Laughlin *et al.* (Traitspace, 2012), developed generic models based on the trait-based filtering approach but limited convergence by selecting the community with the maximum Shannon index of all possible outcomes based on the environmental filtering step. Whilst these two models go some way to reconciling the role of trait-based filtering and competition in predicting community composition, they are both based on empirical relationships between observed trait distributions and environmental gradients. A valuable addition to these approaches would be to derive models that predict shifts in trait distributions in a changing environment from first principles (Laughlin & Laughlin, 2013).

Here we introduce a model which uses functional traits to parametrise a process-based model (Fig. 1 D), using arable weed communities as a case study. The immediate questions the model is designed to address are to do with an impact of a change in crop management on the functional composition of weed communities. However, the model structure is generic to any annual plant community. By combining the two approaches we make use of the extensive literature on traits and community filtering as a convenient means of reducing the parametrisation requirements of a complex population dynamics model whilst retaining the power to capture the processes underlying community assembly. In so doing, we aimed for the optimal balance between complexity and tractability. Weeds are dominated by annual species making the generic life cycle model more tractable and, because of their economic importance, are highly studied with a rich literature of population dynamics models parametrised at the species level. The parameters of the system are also clearly defined by the management operations in the arena of a cropped field. The arable species pool is also sufficiently large to demonstrate the usefulness of a trait-based approach for model parametrisation (including a range of ecological strategies (Bourgeois *et al.*, 2019)), and, because it is dominated by annual species, responds to change on relatively short time scales. In addition, the traits of arable weeds have been well-studied in recent years and trait-based approaches have quantified functional responses of weed communities to management filters (e.g. Fried *et al.*,

This article is protected by copyright. All rights reserved

2009, Gardarin *et al.*, 2010, Gunton *et al.*, 2011, Fried *et al.*, 2012, Colbach *et al.*, 2014, Armengot *et al.*, 2016).

We used functional traits and groups to parametrise the species specific mechanistic processes within our model (Fig. 1 D). We wanted to keep the model parsimonious and so chose only four continuous traits (*sensu* Violle *et al.*, 2007): seed mass, maximum height, date of first flowering and specific leaf area. These four traits are readily available for many annual plants and have been shown to relate to many life-cycle process (Table 1). For example increasing seed mass is known to be associated with decreased seed production (Henery & Westoby, 2003). In addition we also assigned species to functional groups according to i) Ellenberg N number to model the impact of soil fertility on community dynamics, ii) emergence periodicity to model responses to changes in management timings, iii) seedbank type to model persistence in the soil, and iv) phylogeny: whether they were grasses or broadleaves as many of the relationships between other traits and the model parameters varied between these two groups.

We selected these traits based not only on their relationship with various life cycle processes, making them suitable predictors of our model parameters but also due to their availability within the literature. We chose to use only 'soft traits' (*sensu* Díaz *et al.*, 2004) which are more easily measured than 'hard traits' (which may be more directly related to the life-cycle process) and are well documented for a large range of annual plant species across a number of databases (e.g. TRY plant trait database (Kattge *et al.*, 2020), Seed Information Database (SID, 2018), Ecoflora (Fitter and Peat, 1994), and LEDA traitbase (Kleyer *et al.*, 2008)).

The quantification of relationships between functional groups, traits and model parameters is based on a series of experiments screening ecophysiological parameters for 21 annual weed species summarised in Storkey (2006).

Methods

We developed a model of the annual plant life cycle based on transitions between seedlings, mature plants, fresh seed and seed in the seedbank (Fig. 1 C). Some of the processes governing the

This article is protected by copyright. All rights reserved

transitions between these four life stages are influenced by biotic interactions as well as habitat filtering. For each transition (except for fresh seed to seedbank) there are one or more response traits that we anticipate will be selected for or against by environmental or management filters (Table 1). These **response traits** (highlighted in bold throughout the methods section) are integrated into the simulation of mechanistic processes within the annual plant life cycle by quantifying relationships between traits and model parameters (see Appendix S1, Box S1 for a summary of the data sources used to parametrise our trait-response relationships). We fitted linear models to describe the relationships between the life cycle parameters of the simulation model and the weed traits (or functional groups) using GenSTAT. In each case this results in parameter estimates $\{a, b\}$ and an associated covariance function C that captures the uncertainty in the estimates. The data we used to fit the models came from a series of experiments screening ecophysiological parameters for 21 annual weed species summarised in Storkey (2006). In our simulation model, we explicitly account for the uncertainty in the relationships between the traits and life-cycle parameters by stochastically sampling the parameters values from multivariate normal distributions with mean $\{a, b\}$ and covariance C.

The weed life-cycle model proceeds as follows. For each weed species, the number of weed seedlings that emerge from the seedbank is calculated and this is converted to an initial estimate of green area. The green area increases as a function of thermal time up until the canopy reaches closure (which is defined as the total green area index, GAI, equalling 0.75). Thereafter the plants are assumed to grow in competition and both plant height and green area are monitored up until the crop matures to calculate partitioning of light in the canopy. At this stage, we calculate the total biomass for each species and use this to estimate seed production, a proportion of which is returned to the seedbank.

We integrated our model within an existing model of the agricultural landscape — the Rothamsted Landscape Model (RLM, Coleman *et al.*, 2017) to define the environmental and management context of the simulation arena. RLM simulates soil processes, including water and nutrient flows, and the growth of arable crops. We use the soil and crop variables generated by the RLM as inputs

This article is protected by copyright. All rights reserved

Accepted Article

into our life-cycle model. This allows us to simulate the response of the weed community to various environmental factors (such as light and nutrient availability) as well as management (timings of cultivation and application of herbicides). The model runs on a daily time step driven by daily weather variables.

Seedbank → Seedlings

Seedling Emergence The model is initialised by 'planting' a number of seeds per species in each of two layers of the seedbank (a shallow layer from which seeds can readily emerge and a deeper layer from which emergence is reduced). On the day on which the crop is 'sown' in RLM the weed-seedling emergence function is triggered. This function calculates the number of seedlings that emerge for each species. Firstly, we calculate the proportion of the total seedbank that can potentially emerge r_t. We model this as a generic, stochastic process across all species by drawing from a censored Weibull distribution (Eq. 1 with parameters $a = 1.52$ and $b = 0.21$). This distribution was chosen as it gave a good description of data on seedling emergence observed at 5 sites over three years for three contrasting weed species (Appendix S1 Fig. S1.). This is a pragmatic approach that deliberately avoids the need to model interactions between season, induced dormancy and soil microclimate in determining emergence in any given year.

$$\begin{aligned} f(r_t) &= \min\left(\frac{a}{b} {r_t}^{a-1} \exp\left(-\frac{{r_t}^a}{b}\right), 1\right) \text{ for } r_t > 0 \\ &= 0 \qquad\qquad \text{elsewhere.} \end{aligned} \tag{1}$$

Weeds are adapted to emerge at different times of the year. We use an emergence calendar for each species to describe this, and select the proportion of seeds (r_e) predicted to emerge in the time period between sowing and when germination is inhibited by the crop canopy (45 days and 30 days after sowing for autumn and spring sown crops respectively). It would be extremely costly to parameterise an emergence calendar for each species in turn. Instead, we use the functional

This article is protected by copyright. All rights reserved

groupings according to **emergence periodicity**. Here, each species is assigned to one of three groups: spring emergers, autumn emergers or generalists. We fitted bimodal normal probability distributions to each group using data from Storkey *et al.* (2015), see Fig. 2.

We assume that the seeds in the deep layer of the seedbank have a reduced probability of emergence. This is described by scaling the emergence of seeds from the bottom later by

$$r_d = \frac{1}{2}\left(\frac{D-2}{8}\right). \tag{2}$$

where D is the maximum depth from which seeds of that species can germinate. The maximum depth from which seeds of a given species can emerge (D) is estimated using the **seed mass** trait (S_m). The linear relationship

$$D = c\ln\left(S_m\right) + d \tag{3}$$

was derived using data from Storkey et al. (2015) for 18 weed species (see Appendix S1, Figure S2).

The number of seedlings that emerge for each species S_{em} is then given by:

$$S_{\mathrm{em}} = \left(S_B r_d + S_T\right) r_t r_e \tag{4}$$

where S_B and S_T are the seeds in the deep and shallow layers of the seedbank respectively.

Seedling mortality and seedbank decay The numbers of seeds that persist in the deep and shallow layers of the seedbank from one year (k) to the next ($k+1$) are given by

$$S_B\left(k+1\right) = \Delta\left[S_B\left(k\right) - S_B\left(k\right) r_d r_t \left(r_m + r_e\right) - l_g\right] \tag{5}$$

$$S_T\left(k+1\right) = \Delta\left[S_T\left(k\right) - S_T\left(k\right) r_t \left(r_m + r_e\right)\right], \tag{6}$$

where r_m is the proportion of seedlings that are removed by pre-emergence control methods (either through herbicides or cultivation). Here we use the emergence calendar for each species according

This article is protected by copyright. All rights reserved

Accepted Article

to its **emergence periodicity** and assume r_m is the proportion of seeds emerging between the 1st of January (September) and the date the crop is sown in the Spring (Autumn). We assume that 15% (Benvenuti *et al.*, 2001), of the seeds in the bottom layer that are above the maximum depth for emergence lethally germinate (l_g):

$$l_g = 0.15 S_B \left(\frac{D-2}{8} \right). \tag{7}$$

We also account for the fact that a certain proportion of seeds $1 - \Delta$ are lost due to seed-bank decay. The survival rate of seeds in the seedbank, Δ, is associated with the **seedbank type** functional grouping. Following Thompson et al (1997) each species is assigned to one of three seedbank types: transient, short-term persistent, or long-term persistent. Using data from Lutman et al (2002) on the seedbank survival rates for 20 species (3 transient, 11 short-term persistent, 6 long-term persistent) we calculated the average survival rate for each of the three groups: $\Delta_{\text{transient}}$ = 0.3, $\Delta_{\text{short}-\text{termpersistent}}$ = 0.6, and $\Delta_{\text{long}-\text{termpersistent}}$= 0.8.

Following emergence, a proportion of the seedlings are removed by post-emergence control methods. There is currently no known association between herbicide efficacy and plant traits and the response to different herbicides is species specific. To determine the proportion of seedlings of each species removed under different post-emergence herbicide programs we followed the method used by Benjamin *et al.* (2009) and categorised post-emergence herbicide control as either low, moderate, moderately-high or high cost. Expert knowledge was used to estimate the percentage kill of each weed in each crop, given the costing band of the herbicide programme. Cheap programmes were assumed to control weeds, which are easy to kill, whereas more expensive programmes are needed to kill more resilient weeds.

Seedlings $\rightarrow$ Mature plants

Early growth In the early part of the growing season, before the total green area index (GAI) of crop and weeds reaches 0.75, plant growth responds to thermal time (T) and we assume there is

This article is protected by copyright. All rights reserved

no competition between individuals. The GAI of a single plant grows according to

$$W_{\mathrm{GAI}}\left(T\left(j\right)\right) = A\exp\left(RT\left(j\right)\right) \tag{8}$$

where A is the initial value of the GAI when $T = 0$, the R is the seedling relative growth rate and $T\left(j\right)$ is the accumulated thermal time from sowing on day j (see Appendix S1 Box S2). The total GAI for a single species is obtained by multiplying the GAI of an individual by the number of seedlings of that species which emerged. This is calculated daily until canopy closure.

There is an allometric relationship between seed mass and relative growth rate (Shipley & Peters 1990) that we employ using the intermediate step of relating seed mass to initial green area. The initial value of the GAI for a single seedling (A) is estimated from the **seed mass** trait by

$$A = \alpha\ln\left(S_m\right) + \beta \tag{9}$$

where the parameters α and β vary according to two functional groupings; the **emergence periodicity** and the **phylogeny (grass/broadleaf)**. These parameters were derived for each combination of these functional groupings using data for 19 species (4 autumn-emerging grasses (AG), 11 autumn-emerging broadleaved weeds (AB), and 4 spring emergers (SE)) from Storkey (2004) (Appendix S1 Fig. S3).

The seedling relative growth rate R is then estimated from the initial green area (A):

$$R = \gamma A + \delta. \tag{10}$$

Here the parameters γ and δ vary according to the functional groupings of **emergence periodicity** and the **phylogeny** as well as the season in which the function is called. These parameters were derived for each combination of these functional groupings using data for 19 species (4 autumn-emerging grasses (AG), 11-autumn emerging broadleaved weeds (AB), and 4 spring emergers (SE)) from Storkey (2004) (Appendix S1, Figure S4).

This article is protected by copyright. All rights reserved

In RLM, crop relative growth rate is limited by nitrogen according to a scaling factor (N). We use this factor to also scale the growth rate of the weed species (R). When the **Ellenberg N number** is greater than or equal to that of the crop, the scaling factor N takes the same value used in the crop model (this is output from RLM). If the weed species is more sensitive to nitrogen than the crop (i.e. its **Ellenberg N number** is smaller than that of the crop) then N is scaled according to

$$N(q) = N(p)\frac{B(p)}{B(q)} \tag{11}$$

where q refers to the weed species in question and p refers to the crop. Here, B is the reduction in plant biomass (under nitrogen limitation) and is also related to **Ellenberg N**:

$$B = 6.5E_N - 14.4 \tag{12}$$

We derived this relationship using data (Storkey ,2010) on the difference in biomass for 7 weed species grown with and without nitrogen limitation (Appendix S1 Figure S5)

Growth Under Competition Once canopy closure has been achieved plants will compete for light. We used the method described in Kropff and van Laar (1993) to determine the share of light for each species, and to calculate growth rates using an estimate of light use efficiency (see Appendix S1 Box S3). The share of light (s) for a plant of species q on day j is calculated using information about its own height (W_H) as well as the height (W_H) and GAI (W_{GAI}) of the competing species (p of n species):

$$s(q) = \exp\left(-\sum_{p=1}^{n}\left[\varsigma(p)\,W_{\mathrm{GAI}}(p,j)\,\frac{W_H(p,j) - 0.5W_H(q,j)}{W_H(p,j)}\right]\right) \tag{13}$$
$$\text{when } W_H(p,j) - 0.5W_H(q,j) \geq 0$$

where ς is an extinction coefficient with a value of 0.9 for broadleaves and 0.6 for grasses (Kropff and van Laar, 1993).

This article is protected by copyright. All rights reserved

In order to determine s for each species we need to calculate the plant height. Crop height is provided by RLM. Weed height is assumed to grow according to

$$W_H(j) = H_I + \frac{H_M}{1 + \exp(-\zeta[P(j) - \tau])} \tag{14}$$

where $P(j)$ is the accumulated photo-thermal time on day j (see Appendix S1 Box S2). H_I is the initial plant height and $H_I + H_M$ is the **maximum height** for a plant of the given species. The initial plant height, H_I, depends on the **phylogenetic grouping** as in an analysis of initial plant heights (H_I) for 16 species (Storkey, 2006) we found significant differences between grasses and broadleaves with mean values of ($H_I(\text{grasses})$ = 5.204 (SEM = 2.133), $H_I(\text{broadleaves})$ = 0.89 (SEM = 0.738)) (Appendix S1, Figure S6). The ζ parameter describes the rate of growth; a common value across species was determined from data for 16 species (Storkey, 2006) to be 0.0106 (SEM = 0.0011).

The point of inflection, τ, is related to the **day of first flowering** trait (W_F):

$$\tau = \lambda W_F + \kappa \tag{15}$$

where λ = 1.354 (SE = 0.573), and κ = 501.8 (SE = 68.8) with a correlation between parameters of -0.955. These parameters were derived from plant height growth data for 16 species (Storkey, 2006).

Once we have calculated the share of light for a given species(s_q) and that of the competing species (s_p) we can calculate the proportion of intercepted light that each species (q) receives on a given day (j)

$$g(q) = \frac{\varsigma_q W_{\text{GAI}_q}(j)\, s_q}{\sum_{p=1}^{n} \varsigma_p W \text{GAI}_p(j)\, s_p} \left[1 - \exp\left(-\sum_{p=1}^{n} \varsigma_p W_{\text{GAI}_p}(j)\right)\right]. \tag{16}$$

In the case of the crop, we returned this parameter to RLM to adjust the PAR available for crop growth. Growth continues in this way until the weed species reaches maturity, the Julian date of

This article is protected by copyright. All rights reserved

which is predicted using the **day of first flowering** trait (W_F):

$$W_M = 0.314 W_F + 121.8 \tag{17}$$

We derived this relationship using data on weed maturation times for 15 weed species (Storkey, 2006). As data were only available for early flowering species we assumed a constant difference of 10 days between flowering and maturity for all later flowering species (flowering after Julian day 163). We would expect this relationship to be sigmoidal rather than linear however due to the lack of data and the fact that these species will often flower very close to harvest or even after harvest the additional biomass accumulation between flowering and maturity would be unimportant for our model.

During growth under competition, GAI also accumulates. The increase in GAI of a weed species from day j to day $j+1$ is given by

$$W_{\mathrm{GAI}}(j+1) = W_{\mathrm{GAI}}(j) + gI(j) E_m (1-\rho) \tag{18}$$

where I is the amount of incoming photosynthetically active radiation (PAR, given by RLM), E_m is the average light use efficacy (m^2 dry matter MJ^{-1}, see Appendix S1, Box S3) and ρ is a reflection coefficient (0.08 based on an average solar elevation of 45°; Kropff and van Laar,1993).

Mature plants → Fresh seed

Seed Production The number of seeds produced (S_d) by a given species are related to the plant biomass at maturity (W_{BM}, Lutman *et al.*, 2002). We assume this size dependency of reproductive allocation remains constant such that the slope of the relationship = 1 (Sugiyama and Bazzaz, 1998)

$$\ln(S_d) = \upsilon + \ln(W_{\mathrm{BM}}). \tag{19}$$

This article is protected by copyright. All rights reserved

The species dependent parameter υ is estimated from the **seed mass** trait:

$$\upsilon = -0.1177 \ln (S_m)^2 - 0.672 \ln S_m + 5.789. \tag{20}$$

We fitted this relationship to data on 14 weed species (Storkey *et al.*, 2015) (Appendix S1 Figure S7).

The weed biomass at maturity ($W_{\mathrm{BM}}(j)$) is related to the GAI on the day of maturation ($W_{\mathrm{GAI}}(j)$) and the **specific leaf area** trait(W_{SLA}):

$$W_{\mathrm{BM}}(j) = \varepsilon \frac{W_{\mathrm{GAI}}(j)}{W_{\mathrm{SLA}}(j)} \tag{21}$$

where ε = 6.121, SE = 0.363 and relates the leaf biomass (GAI/SLA) to total plant biomass on day j. We used data on measured green area and dry weights from Storkey (2006) to determine this relationship (Appndix S1, Figure S8).

Fresh seed $\rightarrow$ Seedbank

Seed Losses If the weed has not reached maturity on the Julian day when the crop is "harvested" in RLM then no seed is shed. A maximum of 100% seed shed is reached 38 days after maturity (mean of observed data from the UK for *Avena* spp (Barosso *et al.*, 2006) and *Alopecurus myosuroides* (R Hull unpublished data) and estimates for *Galium aparine* (Lutman, 2002) and we assume the response is linear. Any unshed seed is lost and not subsequently added to the seedbank.

Following a meta-analysis of post-harvest seed losses by Davis (2011), seed predation is randomly sampled from a normal distribution with mean 0.52 and standard deviation 0.05. This portion of the seed shed is not subsequently added to the seedbank.

Vertical Movement of Seed in the Soil Seeds are moved vertically between the shallow and deep soil layers following data described by Moss (1990). In years when the cultivation type is "plough" a proportion of seeds from the shallow soil layer are buried into the deep soil layer drawn

This article is protected by copyright. All rights reserved

Accepted Article

from a log-normal distribution with mean $= -0.0515$ and standard deviation $= 0.0191$, conversely some seeds are brought up to the shallow soil layer — this proportion is drawn from a log-normal distribution with mean $= -1.0570$ and standard deviation $= 0.1199$. For all other cultivation types there is no upward movement of seed (from the deep soil layer to the shallow soil layer). For "min till" data on cultivations at 10 cm were used to give the proportion of seeds that are buried taken from the distribution $\mathcal{N}(0.2, 0.051)$. In years where "direct drill" is chosen (data from <5 cm tine) no seeds move vertically.

If the seedbank for a species (in either the top or bottom soil layer) falls below 1 seed (m^{-2}) then that species is assumed to have gone extinct locally and is not included in subsequent years simulation.

Model Testing

To evaluate the performance of our trait-based community model we compared the community predicted by our model with the observed weed community (see Appendix S1 Box S4 for methods of data collection) in an arable field (Brome Pin, Brooms Barn, Suffolk, UK), for which the weather (e-RA, 2018), crops, tillage, and fertiliser input history was available for 30 years (1987–2016). We initialised our model with 100 weed seeds in each soil layer of each species in the regional pool (101 annual arable weeds - see Appendix S1 Box S5). We simulated the 30 years prior to seedbank collection (1987–2016) using the known management information for those years. As we did not know the level of herbicide input used in the field we ran the model 20 times for each level of herbicide input (none, low, medium, high, and very high) to determine whether this significantly altered the number of plants, seeds in the top layer of the seedbank, or seeds per plant in the final simulated community (One-Way ANOVA).

We calculated the functional diversity (*sensu* Petchey and Gaston, 2002) of each simulated community at the end of the 30 year simulation in R 3.5.0 (R Core Team, 2018). We first standardised the traits data and computed a dissimilarity matrix using the *vegan* package (Oksanen *et al.*, 2019). We then used hierarchical clustering to create a dendogram of the relations between species

This article is protected by copyright. All rights reserved

and computed the functional diversity (total branch length) using the *picante* package (Kembel *et al.*, 2010). We tested to see if the selected communities were functionally different under the different herbicide regimes (One-Way ANOVA) and also whether the functional diversity of the selected communities differed significantly from the regional species pool (One-Way ANOVA).

For each model realisation we also compared the resulting density distribution of each trait in the simulated community with the initial trait distribution of the regional pool and that of the observed weed community in Brome Pin.

Results

The weed community in Brome Pin comprised 23 species. The two most abundant species were volunteer crops of oats and oilseed rape. Of the remaining 21 weed species, 6 were perennials and 15 were annuals (Appendix S2 Table S1). In our subsequent analyses we only considered the community of 15 annuals to align with the scope of our model and excluded crop volunteers as their population dynamics are driven by repeated reintroduction.

In our simulations, the abundance of plants varied significantly at different levels of herbicide ($P<0.001$, One-way ANOVA). Plant abundance was highest when herbicide input was low and decreased with increasing herbicide input (Fig. 3a). The number of seeds in the top layer of the seedbank followed a different pattern. Following 30 years of simulation with no herbicide there were few seeds in the seedbank, yet significantly higher seed numbers were simulated at all levels of herbicide application ($P<0.001$, One-way ANOVA). Seed abundance also increased with increasing herbicide input (Fig. 3b). This had the interesting effect that the number of seeds per plant was significantly altered under different herbicide regimes ($P<0.001$, One-way ANOVA) with an exponential-like increase in seed production at increasing levels of herbicide input (Fig. 3c). We suggest this is a result of communities being dominated by species with high fecundity, allowing them to buffer the effects of herbicide, and a reduction in competition between weed individuals.

The community of weed species selected for by the model was fairly consistent across simula-

This article is protected by copyright. All rights reserved

tions. In the majority of our simulations *Sonchus asper* was the most abundant species but across all simulations there were only nine different species which were ever predicted to be the most abundant ($P<0.001$ compared to random selection of species; Appendix S2 Table S2). The species predicted to be the most abundant remained fairly consistent across higher levels of herbicide input, yet as herbicide input was reduced we saw a greater variety in the most abundant species predicted by the simulations (Appendix S2 Table S2). Our model very rarely predicted the local extinction of species, however, the abundances of most species remained very low. The species which did maintain high abundance were often similar across simulations, with eight species consistently ranking among the 20 most abundant species (across all 20 simulations for each herbicide scenario, see Appendix S2 Table S3). *Atriplex patula*, *Conyza cadensis*, *Fumaria officinalis* and *Veronica persica* were often found amongst the 20 most abundant species when no herbicide was applied but markedly less so at any level of herbicide application. Our model had mixed success at predicting the species found in Brome Pin with only 8 of the 15 annuals observed in Brome Pin ranking among the 20 most abundant species in any simulation. However, as it is not possible to separate out environmental filtering from founder effects, and these species were only found in small numbers in Brome Pin it could be that these species would not always be abundant at this field site given the environmental and management conditions. The model was more succesful at predicting the emergent distribution of functional traits.

There was a significant difference in functional diversity of the resulting simulated communities compared to the regional pool (One-way ANOVA, $P<0.001$), indicating that there has been directional selection of functional traits. However, the functional diversity of the simulated communities were not significantly different under different herbicide regimes (One-Way ANOVA, $P<0.05$) indicating that herbicide input is not a key driver of functional diversity in our model, so we only show the trait distributions for the medium herbicide level here.

For the continuous traits included in our model the distribution of traits observed in Brome Pin (yellow distributions in Fig. 4) was a subset of the full regional pool (blue distributions in Fig. 4). In our model simulations (black distributions in Fig. 4) we saw different levels of selection for

This article is protected by copyright. All rights reserved
Accepted Article

the various traits. In our simulations there was a strong selection according to seed mass (Fig. 4a) with the simulated communities all showing similar seedmass trait distributions to that observed in Brome Pin, whereas for maximum height the trait distribution of the simulated communities was not very dissimilar to the regional pool indicating that there is not a strong selection for maximum height in our model (Fig. 4b). The distribution of flowering times (Fig. 4c) observed in our simulations centres on later flowering species than we observed at Brome Pin, however the latest flowering species from our species pool, (first day of flowering in August, Julian day ≥ 213), are excluded following our simulations and so there is some limited evidence for directional selection based on flowering times in our model. There is little evidence for selection based on SLA in our model (Fig. 4d). However, this lack of selection based on SLA is also reflected in the observed community in Brome Pin.

For the discrete functional groups used in our model there was also a distinction between the composition of the regional pool (blue bars in Fig. 5) and the community in Brome Pin (yellow bars in Fig. 5). Again, the model simulations (black lines in Fig. 5) showed varying levels of selection for the different factors. The observed species in Brome Pin all had Ellenberg N values between 6 and 8 with most individuals having an Ellenberg N of 6. The regional pool instead shows a peak at Ellenberg N =7. Many of our simulated communities show a broad spectrum of Ellenberg N values taken from the full range present within the regional pool, however, in some simulations there is selection towards a peak at Ellenberg N = 6, although this is not consistent. In the regional pool there are a similar number of species with each type of emergence calendar (Fig. 5b). However, in Brome Pin we found very few Autumn-emergers and most individuals were generalist emergers. Our simulated weed communities reflected this with strong selection against autumn-emerging species. Whilst most species in our regional pool are broadleaves with fewer grasses (Fig. 5c) there is an even stronger bias toward broadleaves in the weed community observed in Brome Pin, with very few grasses found in the sampled seedbank. Our model simulations reflected this selection pressure and in all simulations the frequency of grasses was reduced compared to the regional pool. In Brome Pin we saw a prevalence of short-term persistent seedbank types and an

This article is protected by copyright. All rights reserved
Accepted Article

absence of species with a transient seedbank. Our model selects for both short-term and long-term persistent seedbank species but we also saw a removal of species with a transient seedbank in line with our observations from Brome Pin.

Discussion

Predicting the relative abundance of species along environmental gradients or following changes in management practices is a fundamental goal in community ecology. Our approach, which links trait-based environmental filtering with a process-based community model, allows both the divergent and convergent selection pressures of environmental filters and biotic interactions to be considered in combination. The observed data on functional traits from the study field generally reflected a convergence of traits, especially for seed mass and this was captured by the model. However two distinct peaks were observed in the density plots of observed data for maximum height and specific leaf area reflecting a divergence of traits in response to crop competition. The simulation output for maximum height also had two peaks, although underestimating the dominance of shorter species. As the functions modelling competition incorporate height, it is encouraging that biological interactions result in a degree of trait divergence. However, the effect of variation in SLA on competition for light is not currently included in the model and the results indicate that further development is required to reflect the observed divergence in this trait.

We demonstrated that by parameterising a process-based model using data from well-studied plant traits that we can effectively model the effect of environmental filters on plant communities at the level of functional traits. In all of our simulations the direction of selection was consistent with in-field observations. Although there was stronger selection for some traits than for others. We predicted different plant communities under different levels of herbicide indicating that this simple management filter does exert selection pressure at the trait level. We also demonstrated that stochasticity can play a role in community assembly as the inclusion of stochastic processes in our model resulted in different realisations of the final plant community, although the functional diversity of those communities remained similar.

This article is protected by copyright. All rights reserved

Accepted Article

Accepted Article

By combining the trait filtering approach with a process-based community model we revealed a number of emergent properties of the model which were not anticipated by the inputs alone. Our model predictions under varying levels of herbicide input predict the largest number of plants when herbicide input is low. This phenomenon has been observed in the field for a number of weed species (e.g. Buhler, 1999, Boström & Fogelfors, 2002) and is consistent with the intermediate disturbance hypothesis which states that at intermediate levels of disturbance (low herbicide) coexistence is more likely (Catford *et al.*, 2012). This unanticipated emergent property highlights the importance of including mechanistic processes in the model in addition to empirical relationships between traits and environmental filters as the synergistic effect of these processes may reveal interesting aspects of community dynamics such as these which can only be revealed when there is both convergent and divergent selection acting simultaneously.

Our model takes mean trait values as input for each species, yet within a species the value for that trait may vary along environmental gradients or change through time (Violle *et al.*, 2007). It is important that we recognise this intraspecific variation in models of this kind. For example, plant height is very dynamic, and depends strongly on disturbance regime (Garnier & Navas, 2012) meaning that the mean values reported in the literature and used here in our model may not be very accurate for plants of the same species growing in a highly disturbed arable field. A similar argument can be made for flowering time. However, as yet, these data are not readily available for arable systems in the UK and as such this may be a source of error in our model. Despite not explicitly incorporating intraspecific trait variation in our model we do include it implicitly by accounting for the uncertainty in each trait–parameter relationship and so by including this stochasticity within the model we account, to some extent, for variation between individuals.

The discrepancies between the ability of our model to successfully predict the correct species list for our studied field (limited success) and the ability to predict the correct distribution of functional traits (greater success) highlights an important question surrounding its utility in predicting community composition. If, as we state in the introduction to this paper, the primary objective of community ecology is to predict the impact of change (environment, land-use, and management)

This article is protected by copyright. All rights reserved

on the *function* of the emergent community then our model succeeds. This will be particularly pertinent where there are associations between the response traits included within the model and any effect traits that determine ecosystem function (Díaz *et al.*, 2007a). However, if the objective is to simply predict the composition of species then our model is of more limited use.

By demonstrating the ability of our model to predict changes in both weed abundance, and the distribution of functional traits we have shown that it will have utility in assessing the viability of various management scenarios. For example, if the aim of weed management is to reduce overall weed abundance we could use our model to assess the success of a number of hypothetical management regimes in achieving this for a given field. Similarly, if the aim of weed management is to provide a functionally diverse weed community which can support the provision of ecosystem services then this too could be assessed through simulation of various management options to determine the best approach for achieving this.

Whilst our model is intrinsically linked to the arable production system, the principle of combining trait-based filtering with process-based models could be easily extended to any ecosystem where the community composition of annual plants is of interest. The generic model of the annual plant life-cycle is broadly applicable and questions surrounding changes in management or environment can be easily addressed as demonstrated by our inclusion of different herbicide programs. For example, the effect of post-emergence herbicide on seedling mortality included in our model could be easily mapped to other management practices such as grazing or even natural disturbances such as burning — provided details are known about the proportion of the population removed by such disturbances. The main factor limiting the application of our modelling framework to habitats other than cultivated fields is the level of specificity of the relationships between functional traits and model parameters that have been quantified for a subset of arable weeds. There is evidence in the literature that some of these allometric relationships follow ecological rules and are conserved across functional groups (for example seed weight and seedling growth rate (Shipley & Peters, 1990). However, the extent to which the model can be applied to other annual plant communities without additional experimental parameterisation remains to be determined.

This article is protected by copyright. All rights reserved
Accepted Article

Acknowledgements

This research was funded by the Natural Environment Research Council (NERC) and the Biotechnology and Biological Sciences Research Council (BBSRC) Institute Strategic Programme NE/N018125/1 LTS-M ASSIST, the BBSRC Institute Strategic Programme BB/P012671/1 Soils 2 Nutrition (S2N) and the Horizon 2020 project IWMPraise grant agreement no 727321 using facilities funded by the BBSRC.

Literature Cited

Armengot, L., et al. 2016. Tillage as a driver of change in weed communities: a functional perspective. Agr ecosyst environ 222:276–285.

Barosso, J., L.Navarrete, M.J. S'anchez del Arco, C.FernandezQuintanilla, P.J.W. Lutman, N.H. Perry, and R.I. Hull. 2006. Dispersal of *Avena fatua* and *Avena sterilis* patches by natural dissemination, soil tillage and combine harvesters. Weed Res 46:118–128.

Benjamin, L.R., A.E. Milne, D.J. Parsons, J. Cussans and P.J.W. Lutman. 2009. Using stochastic dynamic programming to support weed management decisions over a rotation. Weed Res 49:207–216.

Benvenuti, S., M. Macchia, and S. Miele. 2001. Quantitative analysis of emergence of seedlings from buried weed seeds with increasing soil depth. *Weed Sci*, 49, 528–535.

BernardVerdier, M., M.L.Navas, M. Vellend,C. Violle, A. Fayolle and E. Garnier. 2012. Community assembly along a soil depth gradient: contrasting patterns of plant trait convergence and divergence in a Mediterranean rangeland. J Ecol 100:1422–1433.

Boström, U. and Fogelfors, H., 2002. Long-term effects of herbicide-application strategies on weeds and yield in spring-sown cereals. Weed Science, 50:196-203.

Botkin, D.B. 1993. Forest dynamics: an ecological model. Oxford University Press, New York, USA.

This article is protected by copyright. All rights reserved

Accepted Article

Accepted Article

Bourgeois, B. et al. 2019. What makes a weed a weed? A large-scale evaluation of arable weeds through a functional lens. Am. J. Bot. 106:90–100.

Braaker, S., M.K.Obrist, J. Ghazoul, and M. Moretti. 2017. Habitat connectivity and local conditions shape taxonomic and functional diversity of arthropods on green roofs. J Anim Ecol 86:521–531.

Buhler, D.D., 1999. Weed population responses to weed control practices. I. Seed bank, weed populations, and crop yields. Weed Science, 47:416-422.

Carpenter, S.R., R. DeFries, T. Dietz, H.A. Mooney, S. Polasky, W.V. Reid, W.V. and R.J. Scholes. 2006. Millennium ecosystem assessment: research needs. Science 314:257–258.

Catford, J.A., *et al.*, 2012. The intermediate disturbance hypothesis and plant invasions: Implications for species richness and management. Perspect Plant Ecol Syst 14:231–241.

Cornell, H.V. and J.H. Lawton. 1992. Species interactions, local and regional processes, and limits to the richness of ecological communities: a theoretical perspective. J Anim Ecol 61:1–12.

Colbach, N., S. Granger, S.H. Guyot, and D Mézière. 2014. A trait-based approach to explain weed species response to agricultural practices in a simulation study with a cropping system model. Agr ecosyst environ, 183: 197–204.

Coleman, K., S.E. Muhammed, A.E. Milne, L.C. Todman, A.G. Dailey, M.J. Glendining, and A.P. Whitmore. 2017. The Landscape Model: A model for exploring trade-offs between agricultural production and the environment. Sci Total Environ 609:1483–1499.

da Silveira Pontes, L., F. Louault, P. Carrere, V. Maire, D. Andueza, and J.F. Soussana. 2010. The role of plant traits and their plasticity in the response of pasture grasses to nutrients and cutting frequency. Ann Bot 105: 957–965.

This article is protected by copyright. All rights reserved

Accepted Article

D'Amen, M., C. Rahbek, N.E. Zimmermann,and A. Guisan. 2017. Spatial predictions at the community level: from current approaches to future frameworks. Biol Rev 92:169–187.

Davis, A.S., D. Daedlow, B.J. Schutte, and P.R. Westerman. 2011. Temporal scaling of episodic point estimates of seed predation to longterm predation rates. Methods Ecol Evol 2:682–890.

Díaz, S., Hodgson, J.G., Thompson, K., Cabido, M., Cornelissen, J.H., Jalili, A., Montserrat-Marti, G., Grime, J.P., Zarrinkamar, F., Asri, Y. and Band, S.R., 2004. The plant traits that drive ecosystems: evidence from three continents. Journal of vegetation science, 15(3), pp.295-304.

Díaz, S., S. Lavorel, F. de Bello, F. Quétier, K. Grigulis, K. and T.M. Robson. 2007a. Incorporating plant functional diversity effects in ecosystem service assessments. Proc Natl Acad Sci, 104:20684–20689.

Díaz, S., *et al.*, 2007b. Plant trait responses to grazinga global synthesis. Glob Chang Biol 13:313–341.

e-RA: the electronic Rothamsted Archive (2018). Available at http://www.era.rothamsted.ac.uk

Fitter, A. H. and Peat, H. J., 1994, The Ecological Flora Database, J. Ecol., 82, 415-425.

Fried, G., B. Chauvel, and X. Reboud. 2009. A functional analysis of largescale temporal shifts from 1970 to 2000 in weed assemblages of sunflower crops in France. J Veg Sci 20:49–58.

Fried, G., E. Kazakou, and S. Gaba. 2012. Trajectories of weed communities explained by traits associated with species response to management practices. Agr ecosyst environ 158:147–155.

Gardarin, A., C. Dürr, and N. Colbach, N. 2010. Effects of seed depth and soil aggregates on the emergence of weeds with contrasting seed traits. Weed Res 50:91–101.

This article is protected by copyright. All rights reserved

Accepted Article

Garnier, E. and M.L. Navas, .2012. A trait-based approach to comparative functional plant ecology: concepts, methods and applications for agroecology. A review. Agron Sustain Dev 32:365–399.

Gilardelli, F., S. Sgorbati, S. Armiraglio, S. Citterio, and R. Gentili. 2015. Ecological filtering and plant traits variation across quarry geomorphological surfaces: implication for restoration. Environ Manage 55:1147–1159.

Gilg, O., I. Hanski,and B. Sittler. 2003. Cyclic dynamics in a simple vertebrate predator–prey community. Science 302:866–868.

Gross, N., Börger, L., *et al.*, 2013. Uncovering multiscale effects of aridity and biotic interactions on the functional structure of Mediterranean shrublands. J Ecol 101:637–649.

Gunton, R.M., S. Petitand S. Gaba. 2011. Functional traits relating arable weed communities to crop characteristics. J Veg Sci 22:541–550.

Henery, M.L., and Westoby, M., 2001. Seed mass and seed nutrient content as predictors of seed output variation between species. Oikos, 92:479–490.

Hoiss, B., J. Krauss, S.G. Potts, S. Roberts and T. Steffan-Dewenter .2012. Altitude acts as an environmental filter on phylogenetic composition, traits and diversity in bee communities. Proc R Soc B 279:4447–4456.

Kattge, J., Bönisch, G., Díaz, S., Lavorel, S., Prentice, I.C., Leadley, P., Tautenhahn, S., Werner, G.D., Aakala, T., Abedi, M. and Acosta, A.T., 2020. TRY plant trait databaseenhanced coverage and open access. Global change biology.

Keddy, P.A. 1992. Assembly and response rules: two goals for predictive community ecology. J Veg Sci 3:157–164.

Kembel, S.W. *et al.*, 2010. Picante: R tools for integrating phylogenies and ecology. Bioinformatics 26:1463–1464.

This article is protected by copyright. All rights reserved

Accepted Article

Kleyer, M., *et al.*, 2008. The LEDA Traitbase: a database of lifehistory traits of the Northwest European flora. J Ecol 9:1266–1274.

Kropff, M.J. and H.H. Van Laar. 1993. Modelling crop–weed interactions. Int. Rice Res. Inst.

Langmead, O. and C. Sheppard. 2004. Coral reef community dynamics and disturbance: a simulation model. Ecol Modell 175:271–290.

Laughlin, D.C., C. Joshi, P.M. van Bodegom, Z.A. Bastow, and P.Z. Fulé. 2012. A predictive model of community assembly that incorporates intraspecific trait variation. Ecol Lett 15:1291–1299.

Laughlin, D.C. and D.E. Laughlin. 2013. Advances in modeling trait-based plant community assembly.Trends Plant Sci 18:584–593.

Lavorel, S. and E. Garnier. 2002. Predicting changes in community composition and ecosystem functioning from plant traits: revisiting the Holy Grail. Funct Ecol 16:545–556.

Lebrija-Trejos, E., E.A. Pérez-Garca, J.A. Meave, F. Bongers, and L. Poorter. 2010. Functional traits and environmental filtering drive community assembly in a speciesrich tropical system. Ecology 91:386–398.

Lutman, P.J.W., N.H. Perry, R.I.C. Hull, P.C.H. Miller, H.C. Wheeler, and R.O. Hale. 2002. Developing a weed patch spraying system for use in arable crops. HGCA Project Report.

Lutman, P.J.W. 2002. Estimation of seed production by Stellaria media, Sinapis arvensis and Tripleurospermum inodorum in arable crops. Weed Res 42:359–369.

Matsinos, Y.G. and A.Y. Troumbis. 2002. Modeling competition, dispersal and effects of disturbance in the dynamics of a grassland community using a cellular automaton model. Ecol Modell 149:71–83.

Mayfield, M. M., and J. M. Levine. 2010. Opposing effects of competitive exclusion on the phylogenetic structure of communities. Ecology Letters 13:10851093.

This article is protected by copyright. All rights reserved

Accepted Article

Moore, J.K., S.C. Doney, and K. Lindsay. 2004. Upper ocean ecosystem dynamics and iron cycling in a global threedimensional model. Global Biogeochem Cycles 18:1–21.

Moss S.R. 1990. The seed cycle of *Alopecurus myosuroides* in winter cereals: a quantitative analysis. EWRS (European Weed Research Society) symposium, Helsinki 1990, Finland.

Oksanen, J., *et al.*, 2019. vegan: Community Ecology Package. R package version 2.5-4. https://CRAN.R-project.org/package=vegan

Petchey, O.L. and K.J. Gaston .2002. Functional diversity (FD), species richness and community composition. Ecol Lett 5:402–411.

Poff, N.L. 1997. Landscape filters and species traits: towards mechanistic understanding and prediction in stream ecology. J North Am Benthol Soc 16:391–409.

Purves, D.W., J.W. Lichstein, N. Strigul, and S.W. Pacala .2008. Predicting and understanding forest dynamics using a simple tractable model. Proc Natl Acad Sci 105:17018–17022.

R Core Team. 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Royal Botanic Gardens Kew. (2020) Seed Information Database (SID). Version 7.1. Available from: http://data.kew.org/sid/ (2018)

Shin, Y.J. and P. Cury, 2001. Exploring fish community dynamics through size-dependent trophic interactions using a spatialized individual-based model. Aquat. Living Resour 14:65–80.

Shipley, B.and R.H. Peters. 1990. The allometry of seed weight and seedling relative growth rate. Funct. Ecol. 4:523–529.

Shipley, B., D. Vile,and E. Garnier. 2006. From plant traits to plant communities: a statistical mechanistic approach to biodiversity. Science 314:812–814.

This article is protected by copyright. All rights reserved

637 Storkey, J. 2004. Modelling seedling growth rates of 18 temperate arable weed species as a
638 function of the environment and plant traits. Ann Bot 93:681–689.

639 Storkey, J. 2006. A functional group approach to the management of UK arable weeds to support
640 biological diversity. Weed Res 46:513–522.

641 Storkey, J., S.R. Moss, and J.W. Cussans. 2010. Using assembly theory to explain changes in a
642 weed flora in response to agricultural intensification. Weed Sci 58:39–46.

643 Storkey, J., *et al.*, 2015. Combining a weed traits database with a population dynamics model
644 predicts shifts in weed communities. Weed Res 55: 206–218.

645 Sugiyama, S. and F.A. Bazzaz. 1998. Size dependence of reproductive allocation: the influence
646 of resource availability, competition and genetic identity. Funct Ecol 12:280–288.

647 Thompson, K., Bakker, J.P. and Bekker, R.M., 1997. The soil seed banks of North West Europe:
648 methodology, density and longevity. Cambridge university press.

649 Thonicke, K., S. Venevsky, S. Sitch, and W. Cramer. 2001. The role of fire disturbance for global
650 vegetation dynamics: coupling fire into a Dynamic Global Vegetation Model. Glob Ecol
651 Biogeogr 10:661–677.

652 van Nes, E.H., E.H. Lammens, and M. Scheffer. 2002. PISCATOR, an individual-based model
653 to analyze the dynamics of lake fish communities. Ecol Modell, 152:261–278.

654 Violle, C., M.L. Navas, D. Vile, E. Kazakou, C. Fortunel, I. Hummel, and E. Garnier 2007. Let
655 the concept of trait be functional!. Oikos 116:882–892.

Accepted Article

This article is protected by copyright. All rights reserved

Accepted Article

Table 1: Each transition from one stage of the annual plant life cycle to the next is the result of one or more mechanistic processes. The success of each of these processes is determined by the strength of certain environmental filters which act on various response traits.

Transition	Process	Filters	Response Traits
Seedbank → Seedlings	Seedling emergence	Date of crop sowing	Emergence periodicity*
		Depth of cultivation	Seed mass
	Seedling mortality	Herbicide timing	Emergence periodicity*
	Seedbank Decline	Frequency and depth of cultivation	Seedbank persistence*
Seedlings → Mature plants	Early Growth	Temperature	Seed mass, emergence periodicity*, grass / broad-leaf*
		Fertilization	Ellenberg N number*
	Competition	Crop canopy architecture	Maximum height, flowering time
		Fertilisation	Ellenberg N number*
Mature plants → Fresh seed	Seedbank decline	Timing of harvest	Flowering time, SLA, Seed mass

This article is protected by copyright. All rights reserved

Accepted Article

Figure Captions

Fig.1: Combining trait filtering and community dynamics modelling approaches allows us to predict changes in community composition. We use relationships between functional traits published in trait databases and parameters in the annual plant life cycle to parameterise a mechanistic model for multiple species.

Fig.2: Emergence calendars for Spring emergers, autumn emergers and generalist emergers. Bimodal normal probability distributions are fitted to each group using data from Storkey *et al.* (2015), (5 spring emergers, 2 autumn emergers, 3 generalist emergers.

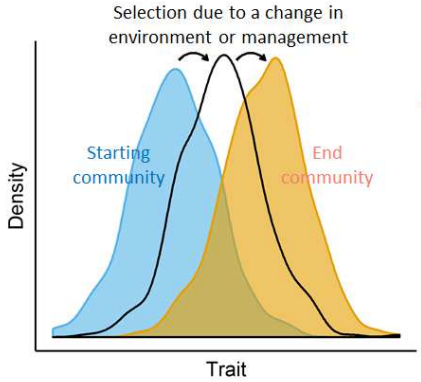
Fig. 3: Summary at all levels of herbicide input of the total a) plants, b) seeds in the top layer of the seedbank, and c) seeds per plant in the end community after 30 years of simulation. Bar height represent the means from 20 simulations at each level of herbicide input and error bars show the standard error of the mean.

Fig. 4: Density plots showing the frequency of the continuous traits a)seed mass, b) maximum height, c) flowering day, and d) specific leaf area. The green distribution shows the density function fitted to the observed data from the Brooms Barn field site, the blue distribution shows the full range of trait data included in the model and represents a density function fitted to an even community consisting of all species, the purple lines are the density functions fitted to each realisation of the field following the simulation of 30 years of management history at the brooms barn site.

Fig. 5: Density plots showing the frequency of the discrete groping factors a) Ellenberg N, b) emergence group, c) phylogeny, and d) seedbank longevity. The green bars show the density of the observed data from the Brooms Barn field site, the blue bars show the full range of trait data included in the model and represents the density of an even community consisting of all species, the purple lines are the density values of the group for each realisation of the field following the simulation of 30 years of management history at the brooms barn site.

This article is protected by copyright. All rights reserved

Accepted Article



A. Predicting community composition and the effects of environmental change on functional diversity is a key goal of community ecology.

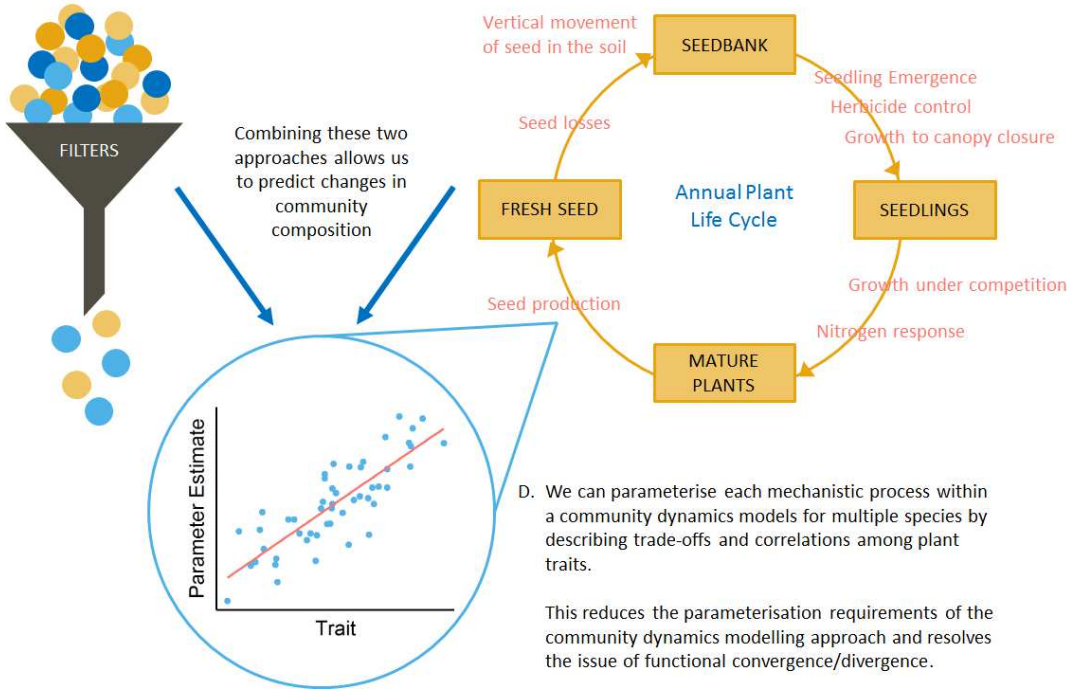
There are two main approaches to this.

B. **Trait-based filtering approach**
Species possessing traits within certain thresholds determined by filters are selected.

This leads to a *convergence* in traits

C. **Community dynamics modelling approach**
The life cycle of individuals is described mechanistically and species compete with one another for resources.

This leads to *divergence* in traits

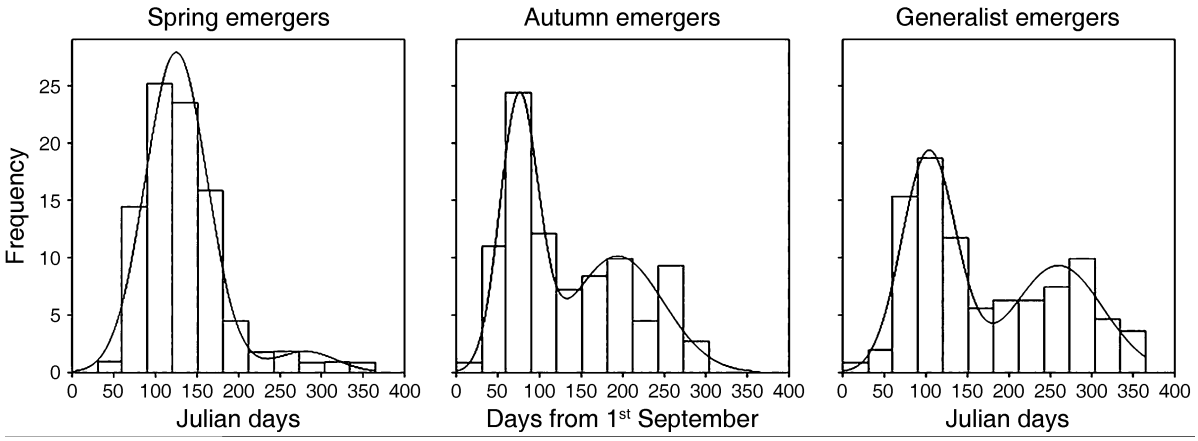


D. We can parameterise each mechanistic process within a community dynamics models for multiple species by describing trade-offs and correlations among plant traits.

This reduces the parameterisation requirements of the community dynamics modelling approach and resolves the issue of functional convergence/divergence.

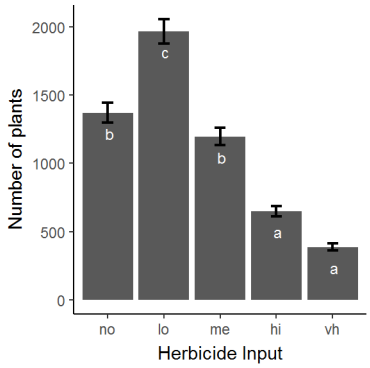
This article is protected by copyright. All rights reserved

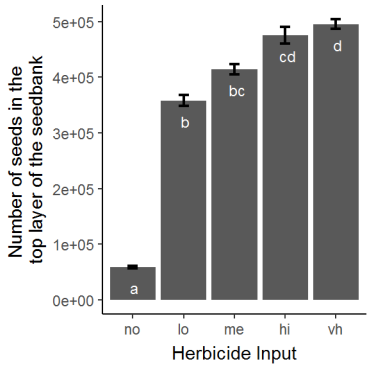
Accepted Article

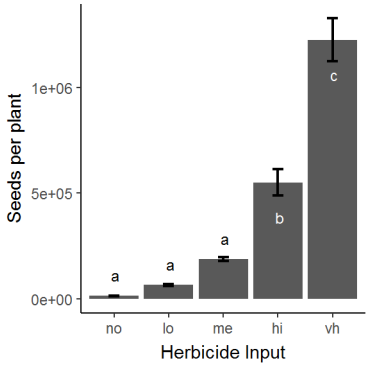


This article is protected by copyright. All rights reserved

Accepted Article

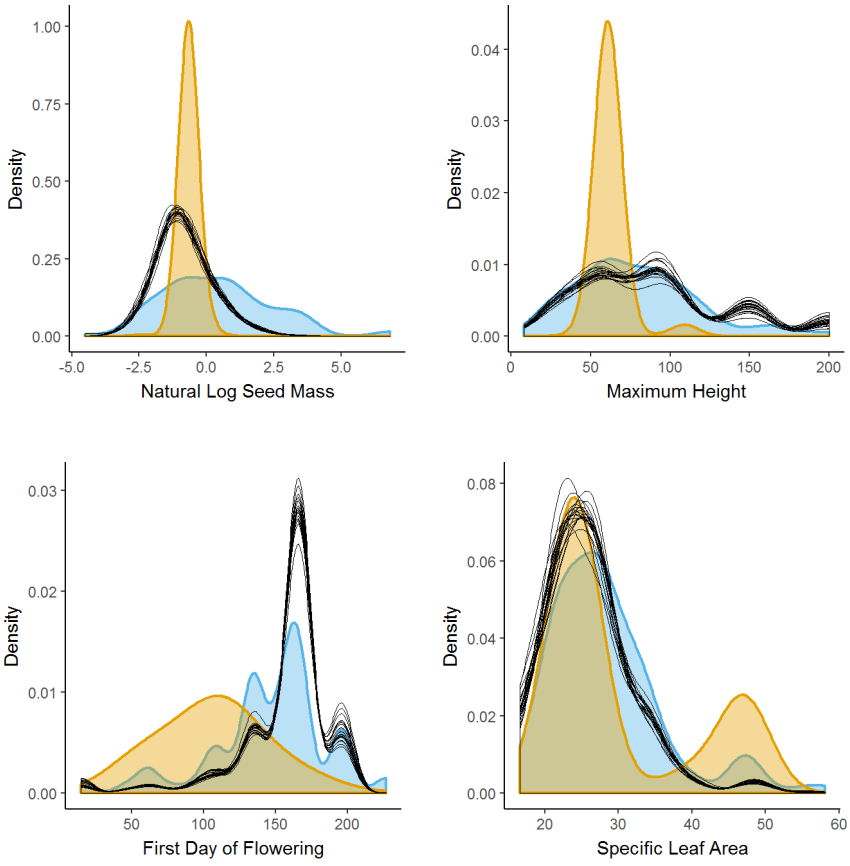






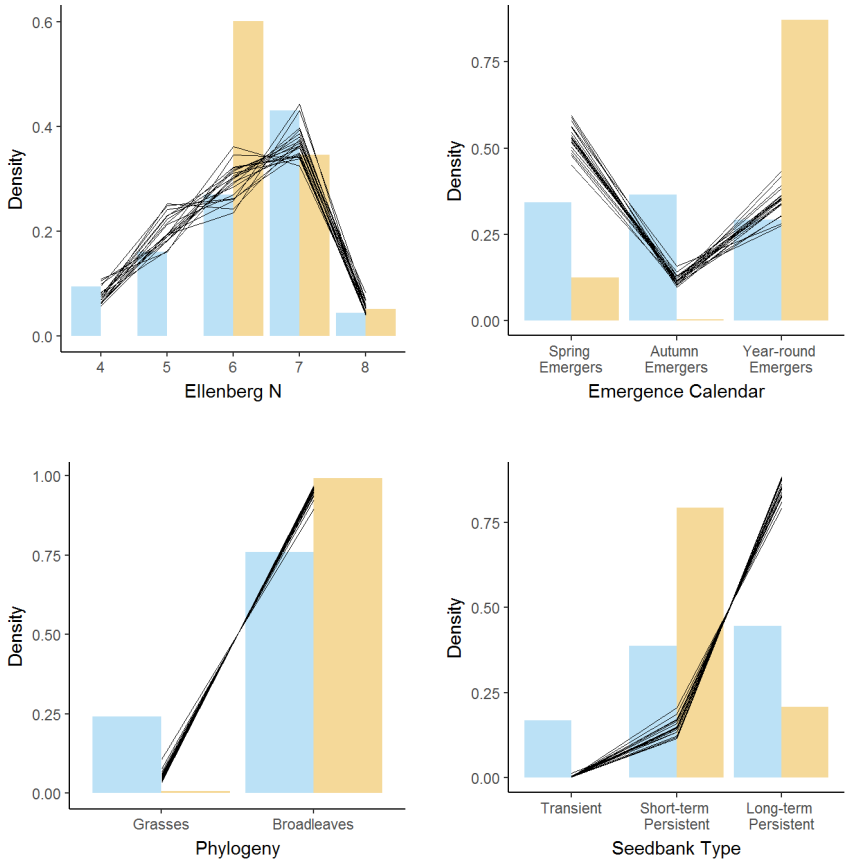
This article is protected by copyright. All rights reserved

Accepted Article



This article is protected by copyright. All rights reserved

Accepted Article



This article is protected by copyright. All rights reserved