Using functional traits to model annual plant community dynamics

HELEN METCALFE 1, 2, ALICE E. MILNE, FLORENT DELLEDALLE, AND JONATHAN STORKEY

Citation: Metcalfe, H., A. E. Milne, F. Deledalle, and J. Storkey. 2020. Using functional traits to model annual plant community dynamics. Ecology 101(11):e03167. 10.1002/ecy.3167

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 parameterize 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 parameterization 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: annual plants; arable weeds; community dynamics; ecological function; environmental filtering; functional diversity; functional traits; population dynamics modeling.

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 characterized as the Holy Grail of ecology (Lavorel and Garnier 2002). As society increasingly recognizes 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. 1A; Diaz 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. 1B), or instead focus on quantifying biotic interactions with process-based community dynamics models (Fig. 1C).

Trait-based filtering approaches that identify the abiotic and biotic filters acting on regionally available pools of species and determine those with favorable 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 [Braak 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 [Diaz 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 typically predict a convergence of trait attributes, as only species that 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 modeling 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 [Thonicke et 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 focusing on competitive processes, these models tend to select for species with divergent trait attributes in order to minimize 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 parameterization 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 parameterization 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 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. While 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 and Laughlin 2013).

Here we introduce a model that uses functional traits to parameterize a process-based model (Fig. 1D), 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 parameterization 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 parameterized 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 parameterization (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. 2009, 2012, Gardarin et al. 2010, Gunton et al. 2011, Colbach et al. 2014, Armengot et al. 2016).

We used functional traits and groups to parameterize the species specific mechanistic processes within our model (Fig. 1D). 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 and Westoby 2001). In addition, we also assigned species to functional groups according to (1) the Ellenberg N number (Ellenberg et al. 1991) to model the impact of soil fertility on community dynamics, (2) emergence periodicity to model responses to changes in management timings, (3) seedbank type to model persistence in the soil, and (3) phylogeny: whether they were grasses or broad-leaves 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 Diaz 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 [available online]. 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 summarized in Storkey (2006).
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.

| Transition                          | Process                  | Filters                        | Response traits                                                                 |
|-------------------------------------|--------------------------|--------------------------------|--------------------------------------------------------------------------------|
| Seedbank → Seedlings                | Seedling emergence       | Date of crop sowing            | Emergence periodicity*                                                          |
| Seedling mortality                  |                          | Depth of cultivation           | Seed mass                                                                       |
| Seedbank Decline                    | Seedling mortality       | Herbicide timing               | Emergence periodicity*                                                          |
|                                     | Seedbank to seedlings    | Frequency and depth of cultivation | Seedbank persistence*                                                      |
| Seedlings → Mature plants           | Early Growth             | Temperature                    | Seed mass, emergence periodicity*, grass/broad-leaf*                            |
|                                     | Competition              | Fertilization                  | Ellenberg N number*                                                             |
|                                     |                          | 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                                                 |

Note: The success of each of these processes is determined by the strength of certain environmental filters that act on various response traits. Items marked * are functional groups and not strictly response traits.

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. 1C). Some of the processes governing the 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 parameterize 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 summarized 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, equaling 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 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 to seedlings

Seedling emergence.—The model is initialized 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. First, we calculate the proportion of the total seedbank that can potentially emerge \( r \). 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 five sites over 3 yr 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.

\[ f(r_t) = \min \left( \frac{a}{B_t} r_t^{-a-1} \exp \left( -\frac{r_t^n}{B_t} \right), 1 \right) \text{ for } r_t > 0 \]
\[ = 0 \text{ elsewhere.} \]

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 predicted to emerge in the time period between sowing and when germination is inhibited by the crop canopy (45 and 30 d 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 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

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

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

\[ D = c \ln(S_m) + d \]

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

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

\[ S_{em} = (S_B r_d + S_T) r_e \]

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

\[ S_B(k + 1) = \Delta[S_B(k) - S_B(k) r_d r_t (r_m + r_e) - l_g] \]
\[ S_T(k + 1) = \Delta[S_T(k) - S_T(k) r_t (r_m + r_e)]. \]

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 to its emergence periodicity and assume \( r_m \) is the proportion of seeds emerging between 1 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 \( l_g \):

\[ l_g = 0.15 S_B \left( \frac{D - 2}{8} \right). \]

We also account for the fact that a certain proportion of seeds \( 1 - \Delta \) are lost due to seedbank decay. The survival rate of seeds in the seedbank, \( \Delta \), 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: $\Delta_{\text{transient}} = 0.6$ and $\Delta_{\text{long-term persistent}} = 0.8$.

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 categorized 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 program. Inexpensive programs were assumed to control weeds, which are easy to kill, whereas more expensive program are needed to kill more resilient weeds.

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

$$W_{\text{GAI}}(T(j)) = A \exp(RT(j)) \quad (8)$$

where $A$ is the initial value of the GAI when $T = 0$, the $R$ is the seedling relative growth rate and $T(j)$ 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 that emerged. This is calculated daily until canopy closure.

There is an allometric relationship between seed mass and relative growth rate (Shipley and 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

$$A = \alpha \ln(S_m) + \beta \quad (9)$$

where the parameters $\alpha$ and $\beta$ 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 ($A$)

$$R = \gamma A + \delta \quad (10)$$

Here the parameters $\gamma$ and $\delta$ 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: Fig. S4).

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

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

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 $N$

$$B = 6.5E_N - 14.4. \quad (12)$$

We derived this relationship using data (Storkey et al. 2010) on the difference in biomass for seven weed species grown with and without nitrogen limitation (Appendix S1: Fig. 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 ($W_{\text{GAI}}$) of the competing species ($p$ of $n$ species)

$$S(q) = \exp \left( -\sum_{p=1}^{n} \varsigma(p) \frac{W_{\text{GAI}}(p,j) W_H(p,j) - 0.5 W_H(q,j)}{W_H(p,j)} \right)$$

$$\text{when } W_H(p,j) - 0.5 W_H(q,j) \geq 0 \quad (13)$$

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

$$W_H(j) = H_i + \frac{H_M}{1 + \exp(-\xi[jP(j) - \tau])}$$  \hspace{1cm} (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 $H_M$ 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 ($H_i$(grasses) = 5.204 (SEM = 2.133), $H_i$(broadleaves) = 0.89 (SEM = 0.738)) (Appendix S1: Fig. S6). The $\xi$ 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, $\tau$, is related to the day of first flowering trait ($W_F$)

$$\tau = \lambda W_F + \kappa$$  \hspace{1cm} (15)

where $\lambda = 1.354$ (SE = 0.573), and $\kappa = 501.8$ (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 ($s_j$) 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 ($j$)

$$g(q) = \frac{\sum_{p=1}^{n} s_p W_{GAI}(j)s_p}{\sum_{p=1}^{n} \sum_{q=1}^{p} s_p W_{GAI}(j)s_q} \left[ 1 - \exp\left( -\sum_{p=1}^{n} \sum_{q=1}^{p} s_p W_{GAI}(j) \right) \right].$$  \hspace{1cm} (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 day of the year (DOY, 1 January = 1) of which is predicted using the day of first flowering trait ($W_F$)

$$W_M = 0.314 W_F + 121.8.$$  \hspace{1cm} (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 d between flowering and maturity for all later flowering species (flowering after DOY 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

$$W_{GAI}(j+1) = W_{GAI}(j) + g(j)E_m(1 - \rho)$$  \hspace{1cm} (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 per MJ, see Appendix S1: Box S3), and $\rho$ is a reflection coefficient (0.08 based on an average solar elevation of 45°; Kropff and van Laar 1993).

**Mature plants to 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)

$$\ln(S_d) = \nu + \ln(W_{BM}).$$  \hspace{1cm} (19)

The species dependent parameter $\nu$ is estimated from the seed mass trait

$$\nu = -0.1177\ln(S_m)^2 - 0.672\ln S_m + 5.789.$$  \hspace{1cm} (20)

We fitted this relationship to data on 14 weed species (Storkey et al. 2015; Appendix S1: Fig. S7).

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

$$W_{BM}(j) = \epsilon \frac{W_{GAI}(j)}{W_{SLA}(j)}$$  \hspace{1cm} (21)

where $\epsilon = 6.121$, 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 masses from Storkey (2006) to determine this relationship (Appendix S1: Fig. S8).

**Fresh seed to seedbank**

**Seed losses.**—If the weed has not reached maturity on the DOY when the crop is “harvested” in RLM then no seed is shed. A maximum of 100% seed shed is reached 38 d 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 et al. (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 “plow” a proportion of seeds from the shallow soil layer are buried into the deep soil layer drawn 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 $A \gamma(0.2, 0.051)$. 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/m² 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 (available online), crops, tillage, and fertilizer input history was available for 30 yr (1987–2016).3 We initialized 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 yr 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-yr simulation in R 3.5.0 (R Core Team 2018). We first standardized the traits data and computed a dissimilarity matrix using the vegan package (Oksanen et al. 2019). We then used hierarchical clustering to create a dendrogram of the relations between species 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 realization, 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 yr 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 simulations. In the majority of our simulations, *Sonchus asper* was the most abundant species but across all simulations there were only nine different species that 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 that 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 cadiensis, Fumaria officinalis,* and *Veronica persica* were often found among 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 successful 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 the various traits. In our simulations, there was a strong selection according to seed mass (Fig. 4A) with the simulated communities all showing similar seed mass 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 centers 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, DOY ≥ 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. While most species in our regional pool are broadleaves with fewer grasses (Fig. 5C) there is an even stronger bias toward broad-leaves 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 absence of species with a transient seedbank. Our model selects for both short-term and long-

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 orange 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 black lines are the density functions fitted to each realization of the field following the simulation of 30 yr of management history at the Brooms Barn site. DOY, day of year, with 1 January = 1.
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 modeling 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.

**Fig. 5.** Density plots showing the frequency of the discrete grouping factors (A) Ellenberg N, (B) emergence group, (C) phylogeny, and (D) seedbank longevity. The orange 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 black lines are the density values of the group for each realization of the field following the simulation of 30 yr of management history at the Brooms Barn site.
We demonstrated that by parameterizing 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 realizations of the final plant community, although the functional diversity of those communities remained similar.

By combining the trait filtering approach with a process-based community model we revealed a number of emergent properties of the model that 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 and 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 that 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 (Viöle et al. 2007). It is important that we recognize this intraspecific variation in models of this kind. For example, plant height is very dynamic, and depends strongly on disturbance regime (Garnier and Navas 1990). However, the extent to which the model can be applied to other annual plant communities without additional experimental parameterization remains to be determined.

Acknowledgments

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 IWM-Praise grant agreement no. 727321 using facilities funded by the BBSRC.

**LITERATURE CITED**

Armengot, J., L. Navarrete, M. J. Sanchez del arco, C. Fernandez-quintanilla, 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 Research 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 Research 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 Science 49:528–535.

Bernard-Verdier, 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. Journal of Ecology 100:1422–1433.

Boström, U., and H. Fogelfors. 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, New York, USA.

Garnier, E., and M. L. Navas. 2012. A trait-based approach to community assembly along a soil depth gradient: contrasting patterns of plant trait convergence and divergence in a Mediterranean rangeland. Journal of Ecology 100:1422–1433.

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. Journal of Animal Ecology 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, 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. Perspectives in Plant Ecology, Evolution and Systematics 14:231–241.

Colusch, 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. Agriculture, Ecosystems & Environment 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. Science of the Total Environment 609:1483–1499.

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. Journal of Animal Ecology 61:1–12.

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. Annals of Botany 105:957–965.

D’Amen, M., C. Rahbek, N. E. Zimmermann, and A. Guisan. 2017. Spatial predictions at the community level: from current approaches to future frameworks. Biological Reviews 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 in Ecology and Evolution 2:682–890.

Diaz, S., et al. 2004. The plant traits that drive ecosystems: evidence from three continents. Journal of Vegetation Science 15:295–304.

Diaz, S., et al. 2007b. Plant trait responses to grazing global synthesis. Global Change Biology 13:313–341.

Diaz, S., S. Lavoire, F. de Bello, F. Quétier, K. Grigulis, and T. M. Robson. 2007a. Incorporating plant functional diversity effects in ecosystem service assessments. Proceedings of the National Academy of Sciences USA 104:20684–20689.

Ellenberg, H., R. Dill, V. Wirth, W. Werner, and D. Paulissen. 1991. Zeigerwerte von Pflanzen in Mitteleuropa. Scripta Geobotanica. Second edition. Verlag Erich Goltze KG, Göttingen, Germany.

Fitter, A. H., and H. J. Peat. 1994. The ecological flora database. Journal of Ecology 82:415–425.

Fried, G., B. Chauvel, and X. Reboud. 2009. A functional analysis of large-scale temporal shifts from 1970 to 2000 in weed assemblages of sunflower crops in France. Journal of Vegetation Science 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. Agriculture, Ecosystems & Environment 158:147–155.

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

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

Gill, G., I. Hanksi, and B. Sittler. 2003. Cyclic dynamics in a simple vertebrate predator–prey community. Science 302:866–868.

Gross, N., L. Börger, S. I. Soriano-Morales, Y. Le Bagoussé-Pinguet, J. L. Quero, M. García-Gómez, E. Valencia-Gómez, and F. T. Maestre. 2013. Uncovering multiscale effects of aridity and biotic interactions on the functional structure of Mediterranean shrublands. Journal of Ecology 101:637–649.

Gunton, R. M., S. Petiand, and S. Gaba. 2011. Functional traits relating arable weed communities to crop characteristics. Journal of Vegetation Science 22:541–550.

Henery, M. L., and M. Westoby. 2001. Seed mass and seed nutrient content as predictors of seed output variation between species. Ökos 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. Proceedings of the Royal Society B 279:4447–4456.

Kattge, J., et al. 2020. TRY plant trait database-enhanced coverage and open access. Global Change Biology 26:119–188.
Keddy, P. A. 1992. Assembly and response rules: two goals for predictive community ecology. Journal of Vegetation Science 3:157–164.

Kemper, S. W., P. D. Cowan, M. R. Helmus, W. K. Cornell, H. Morlon, D. D. Ackerly, S. P. Blomberg, and C. O. Webb. 2010. Picante: R tools for integrating phylogenies and ecology. Bioinformatics 26:1463–1464.

Kleyer, M., et al. 2008. The LEDA Traitbase: a database of life history traits of the Northwest European flora. Journal of Ecology 9:1266–1274.

Kropff, M. J., and H. H. Van Laar. 1993. Modelling crop–weed interactions. CAB International, Wallingford, UK and International Rice Research Institute, Manila, Philippines.

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

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

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

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

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

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

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. Home Grown Cereals Authority, London, UK.

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. Ecological Modelling 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.

Moore, J. K., S. C. Doney, and K. Lindsay. 2004. Upper ocean ecosystem dynamics and iron cycling in a global three-dimensional model. Global Biogeochemical 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, 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. Ecology Letters 5:402–411.

Poff, N. L. 1997. Landscape filters and species traits: towards mechanistic understanding and prediction in stream ecology. Journal of the North American Benthological Society 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. Proceedings of the National Academy of Sciences USA 105:17018–17022.

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

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

Shipley, B., and R. H. Peters. 1990. The allometry of seed weight and seedling relative growth rate. Functional Ecology 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.

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

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

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

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

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

Thompson, K., J. P. Bakker, and R. M. Bekker. 1997. The soil seed banks of North West Europe: methodology, density and longevity. Cambridge University Press, Cambridge, UK.

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

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

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

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at http://onlinelibrary.wiley.com/doi/10.1002/ecy.3167/suppinfo