

PREDICTING THE YIELDS OF SPECIES OCCUPYING A SINGLE TROPHIC LEVEL WITH INCOMPLETE INFORMATION: TWO APPROXIMATIONS BASED ON THE LOTKA-VOLTERRA GENERALIZED EQUATIONS

Hugo Fort

Institute of Physics, Faculty of Science, Universidad de la República, Iguá 4225, Montevideo 11400 Uruguay;

Abstract

The linear Lotka-Volterra generalized equations (LLVGE) serve for describing the dynamics of communities of species connected by negative as well as positive interspecific interactions.

Here we particularize these LLVGE to the case of a single trophic level community with $S > 2$ species, either artificial or natural. In this case, by estimating the LLVGE parameters from the yields in monoculture and biculture experiments, the LLVGE are able to produce quite accurate predictions for species yields.

However, a common situation we face is that we don't know all the parameters appearing in the LLVGE. Indeed, for large values of S , only a fraction of the experiments necessary for estimating the model parameters is commonly carried out.

We then analyze which quantitative predictions are possible with an incomplete knowledge of the parameters. We discuss two approximations that allow using these LLVGE as a quantitative tool. First, when we only know a fraction of the model parameters, the mean field approximation allows making predictions on aggregate or average quantities. Second, for cases in which all the interaction parameters involving a particular species are available, we have the focal species approximation for predicting the yield of this focal species.

Key words: *Prediction of species yields, Quantitative ecology, Agroecology, Biomass forecasting, Community matrix*

1. THE LOTKA- VOLTERRA GENERALIZED LINEAR MODEL FOR MANY INTERACTING SPECIES

The *Lotka-Volterra equations* (Lotka 1925; Volterra 1926) constitute the classical mathematical model for describing the species abundances in terms of their interactions. Lotka and Volterra derived two different sets of equations: One set applies to situations involving competition for food or space and the other set to the *predator-prey* situation. In fact, the possible interactions between two species are richer and include the following cases:

- mutual competition $-/-$, each species has an inhibiting effect on the growth of the other;
- amensalism $-/0$, the growth of one species is negatively affected while the growth of the other is unaffected;
- predation $-/+$, the 'predator' ('prey'), has an inhibiting (accelerating) effect on the growth of the prey (predator);
- commensalism $0/+$, the growth of one species accelerates while the other is unaffected,
- mutual cooperation or mutualism $+/+$, each species has an accelerating effect on the growth of the other.

[Notice that according to these definitions, the host-parasite and the plant-herbivore interactions would be classified as 'predation'.]

A unified framework accommodating all kinds of interactions between pair of species are the so-called Lotka–Volterra *generalized* equations (Hofbauer & Sigmund 1998, Pastor 2008, Fort 2020a) Indeed, in general, interspecific interactions different than mutually negative $-/-$ interaction between species, like $-/+$ or $+/+$ have been considered in the past to occur only between different trophic levels (Morin 2011). For instance, the typical example of $-/+$ is predation and of $+/+$ is mutualism, like the plant-pollinator relationship. However, several examples of these types of interspecific interactions have been also found for species sharing the same trophic level; from viruses (Turner & Chao 1999; Arbiza *et al.* 2010) to natural plant communities (Holmgren *et al.* 1997) or artificial plant polycultures (Haltz *et al.* 2017), just to mention few examples.

This theory in its simplest formulation consists in a set of linear equations for the per capita growth rates $\frac{1}{n_i} \frac{dn_i}{dt}$ where n_i is a variable denoting the *yield* of species i (its density = abundance per unit of area, or its biomass density). A straightforward way to write these *linear* Lotka–Volterra generalized equations (LLVGE) is:

$$\frac{1}{n_i} \frac{dn_i}{dt} = r_i \left(1 + \sum_{j=1}^S I_{ij} n_j \right) \quad i = 1, \dots, S, \quad (1)$$

where r_i is the intrinsic growth rate of species i , with dimension of time^{-1} , and I_{ij} is an interaction coefficient quantifying the *per capita* effect of species j on species i through their pairwise interaction, which can be either negative (competition), positive (facilitation) or zero, with dimension of density^{-1} .

These LLVGE can be thought as the first order or linear approximation in a Taylor series expansion of the per-capita growth rates of species about the equilibrium points of a more complex and general theory (Volterra 1926, 1931):

$$\frac{1}{n_i} \frac{dn_i}{dt} = f_i(n_1, n_2, \dots, n_S) \quad i = 1, \dots, S, \quad (2)$$

where $f_i(n_1, n_2, \dots, n_S)$ are arbitrary functions of the S species making up the community, which we denote more compactly in terms of the S -dimensional vector of yields $\mathbf{n} = [n_1, n_2, \dots, n_S]$ as $f_i(\mathbf{n})$. Therefore, let $\delta \mathbf{n} = \mathbf{n} - \mathbf{n}^*$ be an small displacement from an equilibrium point \mathbf{n}^* , which by definition satisfies $f_i(\mathbf{n}^*) = 0$, and if we perform a Taylor expansion of $f_i(\mathbf{n})$ around \mathbf{n}^* we get:

$$\begin{aligned} f_i(\mathbf{n}) &= f_i(\mathbf{n}^*) + \sum_{j=1}^S \left. \frac{\partial f_i(\mathbf{n})}{\partial n_j} \right|_{\mathbf{n}^*} \delta n_j + \sum_{j=1}^S \sum_{k=1}^S \left. \frac{\partial^2 f_i(\mathbf{n})}{\partial n_j \partial n_k} \right|_{\mathbf{n}^*} \delta n_j \delta n_k + \sum_{j=1}^S \sum_{k=1}^S \sum_{l=1}^S \left. \frac{\partial^3 f_i(\mathbf{n})}{\partial n_j \partial n_k \partial n_l} \right|_{\mathbf{n}^*} \delta n_j \delta n_k \delta n_l + \dots \\ &= 0 + \sum_{j=1}^S \left. \frac{\partial f_i(\mathbf{n})}{\partial n_j} \right|_{\mathbf{n}^*} \delta n_j + O(\delta n^2) \\ &= - \sum_{j=1}^S \left. \frac{\partial f_i(\mathbf{n})}{\partial n_j} \right|_{\mathbf{n}^*} n_j^* + \sum_{j=1}^S \left. \frac{\partial f_i(\mathbf{n})}{\partial n_j} \right|_{\mathbf{n}^*} n_j + O(\delta n^2). \end{aligned} \quad (3)$$

Where $O(\delta n^2)$ involve quadratic terms given by $\sum_{j=1}^S \sum_{k=1}^S \left. \frac{\partial^2 f_i(\mathbf{n})}{\partial n_j \partial n_k} \right|_{\mathbf{n}^*} \delta n_j \delta n_k$, or terms of order greater than two, *i.e.* cubic terms given by $\sum_{j=1}^S \sum_{k=1}^S \sum_{l=1}^S \left. \frac{\partial^3 f_i(\mathbf{n})}{\partial n_j \partial n_k \partial n_l} \right|_{\mathbf{n}^*} \delta n_j \delta n_k \delta n_l$, in such a way that we obtain equations (1) by identifying

$$\boldsymbol{\tau} = - \sum_{j=1}^S \left. \frac{\partial f_i(\mathbf{n})}{\partial n_j} \right|_{\mathbf{n}^*} n_j^*, \quad (4a)$$

$$\boldsymbol{\tau} I_{ij} = \left. \frac{\partial f_i(\mathbf{n})}{\partial n_j} \right|_{\mathbf{n}^*}. \quad (4b)$$

Let us introduce at this point a widely used matrix in community ecology, which is closely related with the matrix $[I_{ij}]$ but different and that often has been used interchangeably, the *community matrix* J_{ij} (Levins 1968). The community matrix, is nothing but the mathematical Jacobian matrix (see Appendix I of Fort 2020a), defined by:

$$J_{ij} \equiv \left. \frac{\partial \left(\frac{dn_i}{dt} \right)}{\partial n_j} \right|_{\mathbf{n}^*}. \quad (5)$$

Thus eq. (4b) can be rewritten as connecting the matrix $[I_{ij}]$ with the community (*i.e.* Jacobian) matrix (see Appendix I of Fort 2020a):

$$I_{ij} = \frac{J_{ij}}{\boldsymbol{\tau}_i n_i^*}. \quad (4'b)$$

Even though the linear Lotka-Volterra equations have been criticized as being too simple for quantitative modeling real systems involving many interacting species¹, they can be regarded as *descriptive* or *phenomenological* models, *i.e.* they describe how the abundance of one species affects the abundance of another, without specifically including a particular mechanism for such interaction (Morin 2011). Neither the nature of the particular competition mechanism nor if it is direct or mediated by other(s) species, is specified. Rather, *effective* interaction coefficients summarize the per capita effects of one species on another.

2 THE LOTKA-VOLTERRA LINEAR MODEL PARTICULARIZED FOR SINGLE TROPHIC COMMUNITIES

As we mentioned in the previous section, the LLVGE serve to model even single trophic communities. In this section we will consider different versions of single-trophic equations. Firstly, the purely competitive case and secondly the more general case in which species, despite belonging to a single trophic level, can exert besides competition facilitative interactions one on another.

2.1 Purely competitive communities.

If the effect of all interactions between species has an inhibiting effect on the growth of the other species, *i.e.* interspecific coefficients are given by:

$$I_{ij} = \alpha_{ij} / \mathcal{K}_i, \quad (6)$$

where \mathcal{K}_i is a maximum sustainable population density, the *carrying capacity*, and the α_{ij} are negative competition coefficients measuring the per capita effect of species j on the abundance of species i . It is customary to take the intraspecific competition coefficients $\alpha_{ii} = -1$. Hence, we can write the Lotka-Volterra competition equations (LVCE) as:

$$\frac{dn_i}{dt} = r_i n_i \left(1 + \frac{\sum_{j=1}^S \alpha_{ij} n_j}{\mathcal{K}_i} \right) \quad i = 1, \dots, S \quad (\alpha_{ij} < 0 \text{ for all } i \text{ and } j, \alpha_{ii} = -1). \quad (7)$$

2.2 Single trophic communities with interspecific interactions of different signs.

The LLVGE for S interacting species within a single trophic level can be written as:

¹ For example, for $S > 2$ species the model would be valid only to the extent that higher-order interactions are null or negligible compared to pairwise interactions (Simberloff 1982). When a pair of species is entangled within a large and complex web of interacting species, like in food webs, it is possible to generate interactions with third or fourth parties that create mutually negative effects for the pair under consideration. This phenomenon is called 'apparent competition' (Keddy 2001).

$$\frac{dn_i}{dt} = r_i n_i \left(1 + \frac{\sum_{j=1}^S \alpha_{ij} n_j}{\mathcal{K}_i} \right) \quad i = 1, \dots, S \quad (\alpha_{ii} = -1), \quad (8)$$

i.e. identical to the competition equations (7) but without the restrictive condition that the interspecific α_{ij} coefficients be negative.

At this point it is worth summarizing the definitions and interpretations of the three alternative interaction strength matrices we have introduced (Table 1, adapted from Novak et al. 2016).

Table 1 Summary of the definitions & interpretations of the three alternative interaction strength matrices

	Community matrix	α_{ij} matrix	LV Interaction matrix
Mathematical definition	$J_{ij} \equiv \frac{\partial \left(\frac{dn_i}{dt} \right)}{\partial n_j} \Big _{\mathbf{n}^*}$	$I_{ij} \equiv \frac{\partial \left(\frac{1}{r_i n_i} \frac{dn_i}{dt} \right)}{\partial n_j} \Big _{\mathbf{n}^*}$	$\alpha_{ij} \equiv \frac{\partial \left(\frac{\mathcal{K}_i}{r_i n_i} \frac{dn_i}{dt} \right)}{\partial n_j} \Big _{\mathbf{n}^*}$
Defined for	any model of differentiable functions $f_i(n_1, n_2, \dots, n_S)$.	any model of differentiable functions $f_i(n_1, n_2, \dots, n_S)$.	the linear generalized Lotka-Volterra model.
Dimensions	time ⁻¹	time ⁻¹ N ⁻¹ L ² or time ⁻¹ M ⁻¹ L ²	non dimensional
Interpretation	Direct effect of the average species individual j on species i 's population growth rate	per-capita growth rate relative to its intrinsic growth rate	per-capita growth rate relative to its intrinsic growth rate

2.3 Obtaining the parameters of the linear Lotka-Volterra generalized model from monoculture and biculture experiments

To obtain the model parameters our starting point are the equilibrium abundances predicted by equation (7) or (8) which verify:

$$n_i^* \left(\mathcal{K}_i + \sum_{j=1}^S \alpha_{ij} n_j^* \right) = 0 \quad i = 1, \dots, S, \quad (9)$$

where the asterisks denote quantities at equilibrium. A main difficulty for obtaining from the set of equations (8) the equilibrium species yields n_i^* is how to get the set of parameters $\{\mathcal{K}_i, \alpha_{ij}\}$. A straightforward procedure is to perform, during sufficiently long enough periods (in order that the equilibrium state is reached): a) the S single species or monoculture experiments, and from each of them to estimate the carrying capacities as the yield of the species i in monoculture $m_i^{\text{ex}*}$ (we use m to emphasize that they are yields in monoculture, while the superscript 'ex' is for denoting experimentally measured quantities to distinguish them from the theoretical ones); b) the $S \times (S-1)/2$ pairwise experiments and for each of them, obtain the pair of the *biculture* (pairwise experiments) yields, $n_{i(j)}^{\text{ex}*}$ and $n_{j(i)}^{\text{ex}*}$ (the subscripts $i(j)$ and $j(i)$ stand for the relative yield of species i in presence of species j and *vice versa*). Using a) we obtain \mathcal{K}_i , and then from b) we obtain α_{ij} and α_{ji} by solving equation (2.9) for $S = 2$, as:

$$\mathcal{K}_i = m_i^{\text{ex}*}, \quad (10a)$$

$$\alpha_{ij} = (n_{i(j)}^{\text{ex}*} - m_i^{\text{ex}*}) / n_{j(i)}^{\text{ex}*}, \quad \alpha_{ji} = (n_{j(i)}^{\text{ex}*} - m_j^{\text{ex}*}) / n_{i(j)}^{\text{ex}*}. \quad (10b)$$

Therefore, if the yield of species i in biculture with species j is $> \mathcal{K}_i$ ($< \mathcal{K}_i$) then $\alpha_{ij} > 0$ (< 0) and the interaction of j on i is facilitative (competitive). This is the kind of approach followed by Vandermeer (1969) in a pioneering experimental study with protozoa.

It is worth stressing that the above procedure to obtain the model parameters works by assuming that the community is in a state of equilibrium. In the strict mathematical sense a community is at equilibrium only when the rate of change for all species, i.e. the left hand side of equations (8) is zero. But this theoretical ideal is rarely achieved in natural communities (Wiens 1984; Roxburgh & Wilson 2000). Moreover, in nature the intensity of interactions may vary with a range of environmental factors including climate, kind of resources, spatial distribution of resources and temporal variation in all of the foregoing. Thus, the coefficients themselves become variables.

2.4 Quantifying the Accuracy of the Linear Model for Predicting Species Yields

As we have mentioned, the minimality of the LLVGE often raises doubts on their ability to make quantitative predictions, like species abundances. Rather Lotka-Volterra models are regarded more as a qualitative than a quantitative tool in population or community ecology (Brown *et al.* 2001). More 'realistic' and complex theories, *e.g.* equations involving non-linear response functions (Vandermeer & Goldberg 2013) and higher order interactions (Abrams 1983) implying non-additive effects (Morin 2011), are often preferred because they are perceived as more reliable (although they can be as intractable as the real systems they aim to model). However, this is done at the price of including additional parameters which are very hard to measure. Actually, in many natural communities –like tropical forests, plankton or mutualistic networks– the species richness S is of the order of hundreds. Hence, estimating all the parameters of the LLVGE from empirical data is an unfeasible task, let alone estimating additional parameters of more complex models.

The viewpoint I adopt here is that models should be mainly evaluated not on the basis of the realism of their assumptions but on the basis of the accuracy of their predictions. Thus, in a recent study (Fort 2018b) I analyzed how well the LLVGE work as a *quantitative* tool for explaining/predicting the outcome of 33 experiments for single trophic species belonging to the same taxonomic group (plants, algae, etc.) assuming that a state of equilibrium was reached. With this aim, I used a dataset of experiments designed to measure the effects of intra and interspecific interactions in single-trophic communities with $S > 2$ species, including experimental treatments that measure all the equilibrium yields of:

- (i) the S coexisting species (one treatment),
- (ii) species in monoculture (S treatments),
- (iii) species in bicultures ($S \times (S-1) / 2$ treatments).

This allowed to obtain the polyculture experimental yields at equilibrium, $n_i^{\text{ex}*}$, and the monoculture yields $m_i^{\text{ex}*} = \mathcal{K}_i$, respectively, by taking averages over replicas of (i) and (ii). Similarly, the biculture yields $n_{i(j)}^{\text{ex}*}$ and $n_{j(i)}^{\text{ex}*}$ were obtained from replicas of (iii), and the

interaction coefficient between species i and j were in turn computed from $n_{i(j)}^{\text{ex}*}$ and $n_{j(i)}^{\text{ex}*}$ though eq. (10b)².

A commonly used metric to assess how well a model fits observed data is the familiar Pearson's correlation coefficient (r) or its square, the coefficient of determination (R^2). Nevertheless, a problem with these two statistics is that they actually describe the degree of *collinearity* between the observed and model-predicted values rather than their numerical agreement (Willmott 1984). In fact, by their very definition, both indices are insensitive to additive and proportional differences between the model predictions and observations (Willmott 1984). Thus, both suffer from limitations that make them poor measures of model performance. For further discussion on why r and R^2 are incorrect measures of predictive accuracy we refer the reader to Li 2017 and references therein. Therefore, to quantitative assess the degree to which the LLVGE match the observations, I resorted to four more appropriate indices (Fort 2018b, Fort 2020a). Indeed, some of these metrics are commonly used in atmospheric and hydrologic sciences. They are:

- The relative mean absolute error (*RMAE*) defined by

$$RMAE = \frac{\sum_{i=1}^S |n_i^{\text{ex}} - n_i| / S}{\overline{n^{\text{ex}}}} 100(\%), \quad (11)$$

where $\overline{n^{\text{ex}}}$ and SE denote, respectively, the mean and standard error for the experimental yields.

- Predictions within 95 % confidence intervals (*P95*), i.e.

$$P95 = \frac{\sum_{i=1}^S \delta(|n_i^{\text{ex}} - n_i|)}{S} 100(\%), \quad \delta(x) = \begin{cases} 1 & \text{if } x < 1.96 \text{ SE} \\ 0 & \text{if } x > 1.96 \text{ SE} \end{cases}, \quad (12)$$

- The modified coefficient of efficiency, defined by

$$E_1 = 1 - \frac{\sum_{i=1}^S |n_i^{\text{ex}} - n_i|}{\sum_{j=1}^S |n_j^{\text{ex}} - \overline{n^{\text{ex}}}|} \quad (13)$$

- The modified index of agreement, defined by

$$d_1 = 1 - \frac{\sum_{i=1}^S |n_i^{\text{ex}} - n_i|}{\sum_{j=1}^S \left(|n_j - \overline{n^{\text{ex}}}| + |n_j^{\text{ex}} - \overline{n^{\text{ex}}}| \right)} \quad (14)$$

Notice that all the above metrics are in terms of the absolute value instead of the square of differences. This is because absolute values are preferable over squares since by using absolute

² Thereafter, since we always assume an equilibrium or quasi-equilibrium state, we will omit the * for all yields in the understanding that they are yields in equilibrium.

values errors and differences are given a more appropriate weighting, not inflated by their squared values (Willmott 1981). Squaring in statistics is useful because squares are easier to manipulate mathematically than are absolute values, but use of squares forces an arbitrarily greater influence on the statistic by way of the larger values (Legates & McCabe1999).

Let us briefly comment these indices:

The relative mean absolute error (*RMAE*) is obtained from dividing the mean absolute error (*MAE*) between the mean of the species yields. *MAE* and the similar root mean square error (*RMSE*) are two commonly used measures for assessing the predictive accuracy in the environmental sciences (Li & Heap 2008). To avoid any dependence of *MAE* on *S* we will use the relative metric *RMAE*. In order to quantify the accuracy we need to introduce some reference point. Actually error measures, like *RMAE*, are not accuracy measures, so they can only tell which model produce less error but there are unable to tell how accurate a model is (Li 2017). At any rate, a very tolerant measure of model goodness would be $RMAE < 100\%$, *i.e.* every quantity is measured with an error smaller than the size of the quantity itself. We will consider here the more stringent threshold of $RMAE < 50\%$. A reference point of 50% might seem too high. Still, it is comparable with the typical SE of the experimental yields (as we shall see in the companion Application 2 chapter).

P95 measures the percentage of predictions which fall within the confidence intervals of 1.96σ (within the error bars shown in Figure 2). $P95 = 100$ (0) % means that all (none of) the yields predicted by the model fall within the error bars around the corresponding experimental values. For *P95* we will consider two thresholds: its maximum possible value of 100 % and the (arbitrary) 66.7 %, so that $P95 \geq 66.7$ % indicates the model does a decent job.

E_1 (Legates & McCabe1999) is a modified version of the coefficient of efficiency (Nash and Sutcliffe 1970) defined by $E = 1 - MSE/\sigma^2$ (MSE = mean square error), but in terms of absolute differences rather than square differences. It ranges from minus infinity to 1, the larger its value the better the agreement. In particular, $E_1 = 1$ indicates perfect match between model predictions and measures. For example, if the absolute differences between the model and the observation is as large as the variability in the observed data (measured by $\sum_{i=1}^S |n_i^{ex} - \overline{n^{ex}}|$), then $E_1 = 0.0$, and if it exceeds it, then $E_1 < 0.0$ (*i.e.*, the observed mean is a better predictor than n_i). In other words, a value of zero for E_1 indicates that the observed mean $\overline{n^{ex}}$ is as good a predictor as the model, while negative values indicate that the observed mean is a better predictor than the model (Wilcox *et al.* 1990).

d_1 (Legates & McCabe1999) is an index similar to E_1 . An advantage over E_1 is that it is bounded *i.e.* it varies from 0.0 to 1.0 (again the higher its value the better agreement between model and observations). By the same token it loses the meaningful reference point of 0.0 for the coefficient of efficiency, which serves to assess when the model is a better predictor than the observed mean. However, it is possible to introduce a reference point for d_1 by observing that for two completely uncorrelated random vectors drawn from a uniform distribution, d_1 is on average $=1/3$ (Fort 2018a, 2020a). Therefore if $d_1 \leq 1/3$ we conclude that the model is a poor predictor.

Using the metrics of Table 2 it was shown that the LLVGE can predict with reasonable accuracy not all but the majority of the species yields for most of the experiments.

3. WORKING WITH IMPERFECT INFORMATION

A common situation we face when trying to predict yields and final diversity of multi-crops in agriculture or of natural communities is that we don't know all the parameters appearing in the LLVGE. Indeed, a main limitation of the experimental procedure of performing all the monoculture and biculture treatments is that it is only feasible provided S is not too large. This is because the number of required experiments grows as S^2 . Thus, for large values of S , only a fraction of these experiments is commonly carried out and consequently we have an incomplete knowledge of the S^2 parameters required to compute the equilibrium yields from the LLVGE.

In this section we will consider two situations in which we can still predict some relevant quantities about species yields with an incomplete knowledge of the LLVGE parameters. Both cases correspond to different practical problems in agricultural sciences. To motivate the first case we can think about mixtures of perennial crops grown for biomass, forage, or food production, and the related phenomenon of *overyielding*, i.e. the increased biomass production in species mixtures relative to monoculture (Beckage & Gross 2006). Thus, let us suppose that we only know a fraction of the interaction coefficients. We will show that in this case of incomplete information the LLVGE, via the recently proposed *mean field* approximation (Fort 2018b), can still predict global or aggregate quantities which are important for farmers. The second situation is when one is interested in predicting the abundance of a *particular* species embedded in a community with other species belonging to the same trophic level. Imagine for example a crop coexisting with different species of weeds and that one is interested in designing rational weed management practices. For instance, by enhancing the competitive ability of crops and simultaneously avoiding environmental damage with the use of herbicides (Guglielmini *et al.* 2016). And, let us suppose that for this species we have quite detailed data, like all the interaction coefficients that involve this species of interest. Is it possible to predict the performance of this given particular species? We will see that the answer is yes; an approximation that works in such cases in which we know all the parameters corresponding to a single species of particular interest is the *focal species approximation* (Fort 2020a, 2020b).

By performing a bibliographic search, one can check that the number of published experimental studies that measured all the yields (i)-(iii) mentioned at the beginning of section 2.4 decreases quickly as S grows. For $S > 6$ is very hard to find an experiment in which the totality of the $S(S+1)/2$ treatments were carried out but only a percentage f_e of the them. Nevertheless, in such cases with $f_e < 100$ %, it is still possible to make quantitative predictions provided we work with relative yields $y_i = n_i / \mathcal{K}_i$ (i.e. the species yield in mixture normalized by its yield in monoculture) rather than yields n_i and use an approximate expression for the interaction matrix.

It is not difficult to show that the interaction matrix can be written in terms of these relative yields, y_i ,

$$y_i = n_i / \mathcal{K}_i \quad (15a)$$

as:

$$[a_{ij}] = [\alpha_{ij} \mathcal{K}_j / \mathcal{K}_i]. \quad (15b)$$

Notice thus that the diagonal elements of the matrix $[a_{ij}]$, corresponding to intra-specific competition, remain equal to -1 since $a_{ii} = \alpha_{ii} \mathcal{K}_i / \mathcal{K}_i = \alpha_{ii} = -1$. Thus, Eq. (9) for the equilibrium in which all species coexist can be written as (Fort 2020a):

$$\sum_{j=1}^S a_{ij} y_j = -1, \quad i = 1, \dots, S, \quad \text{or in matrix form: } \mathbf{A} \mathbf{y} = -\mathbf{1}, \quad (16)$$

where we use A to denote the matrix $[a_{ij}]$, while y and $\mathbf{1}$ are column vectors of S entries with, respectively, relative yields y_i and ones.

So now we are in position to consider the two above mentioned approximations for $A = [a_{ij}]$ which allow making predictions about the yields in cases with incomplete knowledge of the LLVGE parameters.

3.1 The 'Mean Field Matrix' (MFM) approximation for predicting global or aggregate quantities

The *Mean Field Matrix* (MFM) approximation (Fort 2018b, Fort & Segura 2018), is capable of predicting aggregate or mean quantities, expressed in terms of relative yields, with reasonable accuracy. The name alludes to the resemblance of this approximation with the **Mean Field approximation**, commonly used in physics, consisting in replacing spatially dependent variables by a constant equal to their mean value. Here the average will be taken over the values of the interspecific interaction coefficients *i.e.* the off-diagonal part of the interaction matrix $[a_{ij}]$ is replaced by a “mean field” competition coefficient.

The aggregate or mean quantities we will consider are the Relative Yield Total (de Wit 1970),

$$RYT = \sum_{i=1}^S y_i, \quad (17)$$

and the Mean Relative Yield,

$$MRY = RYT / S = \sum_{i=1}^S y_i / S. \quad (18)$$

Both these indices allow comparing community productivity on a relative basis. For instance, in agriculture science, the *RYT* is often used to quantify the overyielding of diverse plant mixtures relative to plant monocultures in studies of biodiversity effects on ecosystem function. A $RYT > 1$ implies that the yield performance will be better in polyculture than in monoculture, a phenomenon termed as *overyielding* (Vandermeer 1989).

The idea then is to replace the off-diagonal elements of the interaction matrix $[a_{ij}]$, corresponding to interspecific interactions, by their average value over the sample of available interspecific competition coefficients:

$$a = \overline{a_{i \neq j}}, \quad (19)$$

which thereafter will be called the *mean interspecific interaction strength parameter*. In such a way we get a mean field matrix (MFM):

$$A^{MF} \equiv [a_{ij}]^{MF} = \begin{pmatrix} -1 & a & \cdots & a & a \\ a & -1 & \cdots & a & a \\ & & \ddots & & \\ a & a & \cdots & -1 & a \\ a & a & \cdots & a & -1 \end{pmatrix}. \quad (20)$$

For the MFM Eq. (16) reduces to

$$-y_i + a \sum_{j \neq i}^{S-1} y_j = -1, \quad i = 1, \dots, S, \quad (21)$$

that by summing and subtracting ay_i can be written as:

$$-y_i(1+a) + a \sum_{j=1}^S y_j = -1, \quad i = 1, \dots, S, \quad (22)$$

that is,

$$-y_i(1+a) + aRYT = -1, \quad i = 1, \dots, S, \quad (23)$$

which we can solve for y_i as:

$$y_i = \frac{1+aRYT}{1+a}, \quad \text{for all } i. \quad (24)$$

And, if we sum from 1 to S both sides of (24), we get

$$RYT = S \frac{1+aRYT}{1+a},$$

which allows to express the RYT and MRY as a simple functions of a and S (Fort 2018b, Fort & Segura 2018):

$$RYT^{MF}(a, S) = \frac{S}{1-(S-1)a}, \quad (25)$$

$$MRY^{MF}(a, S) = \frac{1}{1-(S-1)a}, \quad (26)$$

where the superscript 'MF' is to emphasize that these expressions hold for the MFM approximation. The pair of Eqs. (25) and (26) tell us that the MFM approximation is particularly suited to address how the total or mean yield depends with the species richness and the mean intensity of competition. This is an interesting question in ecology.

A caveat of the method is that if facilitation is the dominant interspecific interaction, so that $a > 0$, the denominator of equations (25) and (26) could become negative for a or S sufficiently large. Since the RYT and MRY must be positive quantities, a requirement for this approximation to work properly is that facilitation is not too large.

Comparisons between predictions generated by formulas (25) and (26) against data from a dataset of 77 experiments (Fort 2018b) are shown, respectively, in Fig. 1 and Fig. 2 as a function of both S and a . In all these experiments the mean interspecific interaction strength parameter is negative and the number of species S varies from 2 to 32 (Fort 2018b).

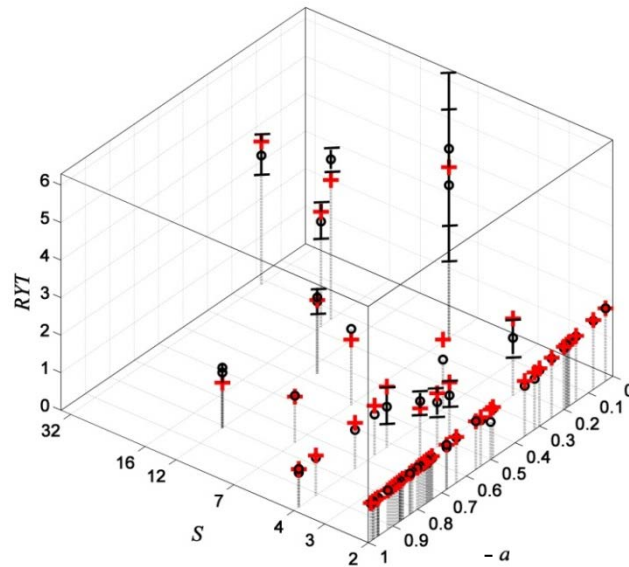


Figure 1. Empirical (o) and theoretical RYT , predicted by the MFM approximation Eq. (25) (red crosses), for a set of 77 experiments (see text) as a function of S (log scale) and the mean competition parameter a . The experimental error bars correspond to \pm the standard error (SE).

As mentioned, the MFM approximation is useful to analyze how the yield depends with the species richness and the mean intensity of competition. Notice that:

- First, regarding the dependence of the RYT as a function of S , Fig. 1 shows that the RYT increases with S , a phenomenon which has been observed in many empirical studies and interpreted as *niche partitioning* (Cardinale *et al.* 2011; 2012). That is, different species can use resources in a complementary way and therefore more diverse communities would lead to a higher RYT . The $RYT > 1$ for the great majority of the experiments considered here, especially for those with $S > 2$ interacting species (24 out of 25 as shown in Fig. 1), is a clear *quantitative* evidence of such species complementarity (Loreau & Hector 2001, de Wit 1970). Notice that for a fixed value of the mean interaction parameter a , Eq. (25) becomes a non-linear saturating relationship between diversity and yield. Non-linear saturating relationships have also been obtained in models by simulation or in other more complex models (Loreau 1998, Schnitzer *et al.* 2011, Downing, et al. 2012).
- Second, as expected, for fixed S the RYT and MRY (Fig. 2) both decrease with the absolute value of a . That is, the more intense the average interspecific competition the smaller the yield.

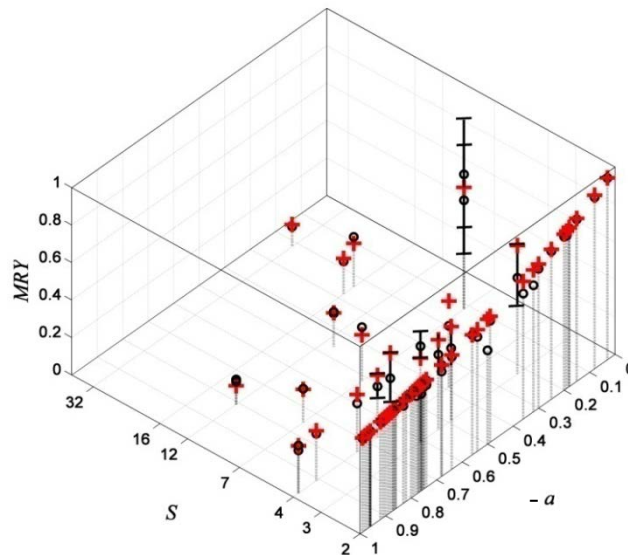


Figure 2 Theoretical *MRY* predicted by the MFM approximation Eq. (26) (red crosses) vs. empirical *MRY* (o) with error bars corresponding to \pm SE for a set of 77 experiments as a function of *S* (log scale) and *a*.

- Third, the relative absolute errors of the *RYT* and *MRY* obtained when using formulas provided by Eq. (25) and Eq. (26), in the case of experiments with incomplete knowledge of model parameters, range between 0.8% and 24 % (mean 12.4 % \pm 4.6 %) and are not significantly greater than those obtained by just solving the entire set of LLVGE, which was possible for experiments in which all the model parameters were available as the ones considered in Fort (2018a). Furthermore, in general the *RMAE* is proportional to the size of the experimental error bar and thus the failure of equations (25) and (26) can be mostly explained by a large variance among experimental replicas or by a low experimental precision.

Therefore:

Provided only competition interactions among species exist or at least they are dominant, we still can predict aggregate (*RYT*) or mean (*MRY*) relative competition indices with an accuracy inversely proportional to the experimental variance of the experimental measurements.

3.2 The 'Focal species' approximation for predicting the performance of a given species when our knowledge on the set of parameters is incomplete

Now, as mentioned at the beginning of this section, let us turn to a different situation and suppose we are interested in the performance of a particular or *focal* species that coexists with other species. For example, say a crop that coexists with several species of weeds. Since we are not interested in the yield of each individual weed, *we can treat all of them as a single species*. Thus in order to make more accurate specific predictions for this focal species than just the *MRY*, one possibility is to measure the interaction coefficients for this species k with all the others and use a *focal interaction matrix*, in terms of relative yields, of the form:

$$\mathbf{A}^{(k)} \equiv [a_{ij}]^{(k)} = \begin{pmatrix} -1 & -1 & \cdots & a_{1k} & \cdots & -1 & -1 \\ -1 & -1 & \cdots & a_{2k} & \cdots & -1 & -1 \\ \vdots & \vdots & \ddots & \vdots & \cdots & \vdots & \vdots \\ a_{k1} & a_{k2} & \cdots & -1 & \cdots & a_{kS-1} & a_{kS} \\ \vdots & \vdots & \cdots & \vdots & \ddots & \vdots & \vdots \\ -1 & -1 & \cdots & a_{S-1k} & \cdots & -1 & -1 \\ -1 & -1 & \cdots & a_{Sk} & \cdots & -1 & -1 \end{pmatrix}, \quad (27)$$

i.e. all the interspecific coefficients between species different from the focal species k are taken equal to -1 , while a_{ik} denotes the interaction coefficient of the focal species over species i and a_{ik} the interaction coefficient of species i over the focal one.

The recipe is then to use matrix (27) to solve Eq. (16) and only take into account the relative yield for the focal species $y_k^{(k)}$ (while neglecting the values $y_j^{(k)}$ for $j \neq k$). That is, from a matrix of the form (27) we obtain just one relative yield for the focal species ($y_k^{(k)}$). Let us call this approximation, which represents an improvement of the mean field one, the *focal species approximation* (Fort 2020b).

Obtaining the $y_j^{(k)}$ would be straightforward from equations if A) the theoretical coexisting equilibrium were *feasible*, *i.e.* $y_j^{(k)} > 0$ for all species (to warrant that Eq. (9) reduces to Eq. (16)), and B) the matrix $\mathbf{A}^{(k)}$ were invertible, by simple inversion of the matrix equation between parentheses in equation (16). However, it is easy for $S > 3$ species to prove that the focal matrix is non-invertible: the $\det \mathbf{A}^{(k)} = 0$. This determinant can be immediately computed by expanding the determinant of matrix (4) along its focal row or column. For $S = 3$ species, there is generally only one row or column with all elements equal to -1 and the determinant is not necessarily zero. However, in this case it is easy to show that at least one of the species must become extinct at equilibrium (Fort 2020b), and thus condition A) is not fulfilled, meaning that the equilibrium state with all species coexisting is unfeasible *i.e.* $y_j^{(k)} = 0$ for some species j .

Therefore, since either conditions A) or B) are violated, one cannot obtain the equilibrium relative yields by inverting Eq. (16) and has to obtain the equilibria of focal matrices by numerically integrating the dynamical equations:

$$\frac{dy_i^{(k)}}{dt} = r_i y_i^{(k)} \left(1 + \sum_{j=1}^S a_{ij}^{(k)} y_j^{(k)} \right), \quad i = 1, \dots, S. \quad (28)$$

To do this, one can use either a simple finite differences method or an ordinary differential equation solver. Once the set of S relative yields $y^{(k)}_j$ is obtained, the next step is to consider the relative yield for the focal species $y^{(k)}_k$ (neglecting the values $y^{(k)}_j$ for $j \neq k$).

To illustrate how this approximation works we will consider the experiment of four algal species of Huisman *et al.* (1999): *Aphanizomenon flos-aquae* (*Ap*, species #1), *Chlorella vulgaris* (*Ch*, species #2) a *Microcystis* strain (*Mi*, species #3) and and *Scenedesmus protuberans* (*Se*, species #3) and. The population densities, measured as millions per mL) of the four species in polyculture were:

$$\mathbf{n} = \begin{bmatrix} 0.002 \\ 18.600 \\ 2.150 \\ 0.001 \end{bmatrix}$$

The carrying capacity vector \mathbf{k} (millions per mL) and the interaction matrix $\alpha = [\alpha_{ij}]$ are given by (Fort 2018a):

$$\mathbf{k} = \begin{bmatrix} 0.180 \\ 20.40 \\ 17.90 \\ 0.531 \end{bmatrix}$$

$$\alpha = \begin{bmatrix} -1 & -0.0086 & -0.0060 & -0.4000 \\ 0 & -1 & 0 & 0 \\ -86.0000 & -0.8772 & -1 & -30.7692 \\ -2.9706 & -0.0256 & -0.0295 & -1 \end{bmatrix}$$

Therefore, using respectively Eq. (15a) and (15b), we get for the vector of experimental relative yields y^{ex} and for the matrix $A = [a_{ij}]$:

$$y^{\text{ex}} = \begin{bmatrix} 0.010 \\ 0.912 \\ 0.120 \\ 0.001 \end{bmatrix}$$

$$A = \begin{array}{cc} & \begin{array}{c} Ap \\ Ch \\ Mi \\ Se \end{array} \\ \begin{array}{c} Ap \\ Ch \\ Mi \\ Se \end{array} & \begin{bmatrix} -1 & -0.978 & -0.597 & -1.778 \\ 0 & -1 & 0 & 0 \\ -0.865 & -1.000 & -1 & -0.911 \\ -1.009 & -0.987 & -0.998 & -1 \end{bmatrix} \end{array}$$

If we consider that the focal species is for example *Ch*, then we obtain the focal matrix $\hat{A}^{(Ch)}$ (with boxed elements unchanged):

$$A^{(Ch)} = \begin{array}{c} \begin{array}{cccc} & Ap & Ch & Mi & Se \\ Ap & -1 & -1.133 & -1 & -1 \\ Ch & 0 & -1 & 0 & 0 \\ Mi & -1 & -1.003 & -1 & -1 \\ Se & -1 & -1.155 & -1 & -1 \end{array} \end{array}$$

The matrix $A^{(Ch)}$ is not invertible (three columns are equal and thus its determinant is 0). Therefore we cannot use Eq. (16) to get the equilibrium vector y . Nevertheless we can solve the differential equations (28) numerically, until the equilibrium state is reached, and we obtain:

$$y = \begin{bmatrix} 0.000 \\ 1.000 \\ 0.000 \\ 0.000 \end{bmatrix}$$

Thus $y_{Ap}^{(Ch)} = 1.000$, differs in 10 % respect to the experimental $y_{Ch}^{ex} = 0.912$ (second entry of Eq. (31)). This is not so bad. Similarly, one can check that $y_{Ap}^{(Ap)} = 0.00$ slightly differs with respect to its experimental value of 0.01.

Table 2 summarizes the *RMAE* and the three accuracy metrics $-P95$, E_1 and d_1 . It shows highlighted in gray when the indices are below thresholds, indicating poor accuracy.

Table 2. Error and accuracy metrics of the focal method for the 33 experimental studies of Fort (2018a). The first column provides the corresponding references. *S* is the number of species. Columns from *RMAE* % to d_1 are the four metrics to assess model error/accuracy used. Highlighted in gray are values that violate the required accuracy conditions: a) *RMAE* < 50%, b) *P95* \geq 66.7%, c) $E_1 > 0$ and d) $d_1 > 1/3$. Also listed for comparison the same metrics in the case of using the full interaction matrix as in Fort (2018a).

Taxon	Community	<i>S</i>	Focal species method				Full LVGE			Reference	
			<i>RMAE</i>	<i>P95</i>	E_1	d_1	<i>RMAE</i>	<i>P95</i>	E_1	d_1	
			as %	%			as %	%			
PLANTS	1939 pasture	4	182	25.0	-2.003	0.25	78	50	-0.286	0.542	Aarsen 1988
PLANTS	1993	3 *	21	66.7	0.655	0.837	19	100	0.675	0.858	Hooper & Duker 2004
	1998	3	49	33.3	-0.401	0.417	69	33.3	-0.948	0.452	
	1999	3	55	66.7	-0.321	0.539	68	0	-0.630	0.515	Huisman et al. 1999
ALGAE		4	21	75.0	0.832	0.923	21	75	0.862	0.936	Kastendiek 1982
ALGAE		3	25	100.0	0.513	0.804	32	100	0.376	0.762	Neill 1974
CRUSTACEAN		4	35	75.0	0.634	0.819	1	100	0.987	0.989	Picasso et al. 2008
PLANTS**	A-WC-IBF	3	17	100.0	0.863	0.926	13	66.7	0.893	0.945	
	A-OG-IW	3	22	100.0	0.720	0.872	29	66.7	0.631	0.844	
	A-OG-SW	3	17	100.0	0.811	0.914	21	100	0.763	0.894	
	A-IW-EG	3	19	100.0	0.712	0.853	16	100	0.755	0.878	
	A-SW-EG	3	20	100.0	0.676	0.860	23	100	0.638	0.847	
	WC-OG-IW	3	23	100.0	0.644	0.817	33	100	0.495	0.764	
	WC-OG-SW	3	36	66.7	0.629	0.772	31	100	0.680	0.817	
	WC-SW-EG	3	16	100.0	0.824	0.912	17	100	0.809	0.905	
	IBF-OG-IW	3	23	66.7	0.752	0.858	22	66.7	0.766	0.878	
	IBF-OG-SW	3	9	100.0	0.907	0.954	14	66.7	0.843	0.924	
	IBF-IW-EG	3	134	33.3	-0.328	0.374	74	33.3	0.260	0.440	
	IBF-SW-EG	3	63	66.7	0.290	0.601	50	66.7	0.445	0.637	
	A-WC-OG-IW	4	13	75.0	0.862	0.930	62	50	0.335	0.784	
	A-WC-SW-EG	4	17	75.0	0.845	0.920	42	100	0.607	0.857	
	A-IBF-SW-EG	4	35	50.0	0.690	0.833	20	100	0.822	0.927	
	A-IBF-OG-IW	4	31	75.0	0.691	0.847	27	75	0.731	0.906	
	WC-IBF-OG-IW	4	28	75.0	0.714	0.847	38	50	0.615	0.857	
	WC-IBF-SW-EG	4	108	50.0	-0.377	0.568	110	75	-0.413	0.651	Rees et al. 1996
PLANTS	area A, 1979	4	24	100.0	-0.110	0.101	28	100	-0.330	0.164	
	area B, 1979	4	30	75.0	0.324	0.595	33	75	0.253	0.748	Roxburgh & Wilson 2000
PLANTS		7	112	42.9	-0.519	0.527	43	57.1	0.549	0.795	Vandermeer 1969
PROTOZOA		4	22	100.0	0.792	0.900	18	75	0.826	0.935	Wilson & Keddy 1986
PLANTS		7	62	71.4	-0.494	0.308	64	28.6	0.206	0.760	Zarnteske et al. 2013
PLANTS	LOW SAND	3	20	66.7	-0.512	0.247	5	100	0.662	0.853	
	MID SAND	3	25	66.7	0.453	0.649	12	100	0.725	0.864	
	HIGH SAND	3	90	0.0	-1.049	0.328	6	100	0.857	0.923	
average			42	72.7	0.295	0.694	35	76.1	0.468	0.783	

*: Functional groups rather than species.

** : Averages of biomass over the three annual harvests (2004, 2005 and 2006), A = alfalfa, W = white clover, IBF = Illinois bundleflower, O = orchardgrass, IW = intermediate wheatgrass, SW = switchgrass, EG = eastern gamagrass.

The *RMAE* when using the focal interaction matrix and the full interaction matrix are very similar, though *P95* and E_1 indices show a greater number of experiments below the established accuracy value for the species focal approach. Overall, the accuracy shown for the proposed focal species approximation is similar to that of LLVGE using the full interaction matrix.

Figure 3 compares the results of using this approximation (open big circles) with those when using the full set of LLVGE (filled small circles) for the same four experiments shown in Fig. 2.3. We can see that the accuracy with which the focal species approximation predicts the focal species is comparable to the one obtained when using the full set of LLVGE.

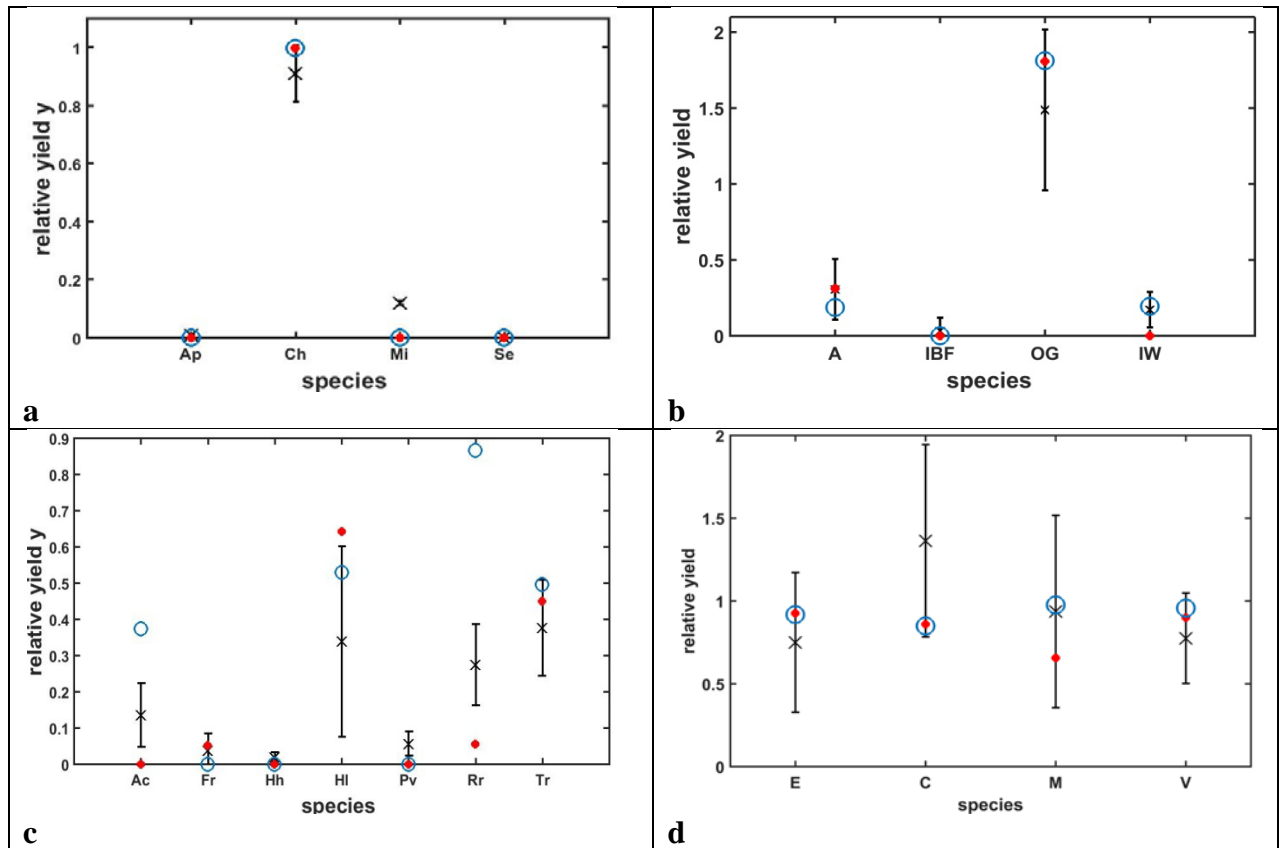


Figure 3. The focal species approximation compared with the full set of LLVGE. Experimental relative yields $R_i = Y_i/K_i$ with their error bars (crosses), predicted values ex ante by the LLVGE (red filled circles) and the corresponding values produced by the focal species approximation (open blue circles): a) Algae (Huisman *et al.* 1999). b) Plants A-IBF-OG-IW mixture (Picasso *et al.* 2008). c) Plants (Roxburgh & Wilson 2000). d) Plants, Area A (Rees *et al.*, 1996).

Therefore we conclude that:

It is possible to predict the yield of a given particular species when we just know the interaction coefficients involving this focal species, by means of the focal species approximation, with accuracy comparable to the one obtained when using the full set of LLVGE.

4 CONCLUSION

The linear Lotka-Volterra generalized equation (LLVGE) are the simplest system of equations to model a community of species with all possible types of interspecific interactions: $-/-$, $-/0$, $-/+$, $+/0$ and $+/+$. Particular cases of such equations are the competition model (all species interactions $-/-$) and the mutualistic model (all species interactions $+/+$).

These LLVGE can be used as a quantitative tool. If all the model parameters are known, the full set of LLVGE allows in general to predict the yields of all species making up the community with reasonable accuracy (Fort 2018a, 2020a). More elaborate models may be necessary for describing and predicting species abundances for particularly complex communities (Brown *et al.* 2001). However, it seems that such cases would be more the exception than the rule (Fort 2018a, 2020a).

However, we have also seen that for the majority of real communities we only have access to a fraction of the model parameters. In these circumstances with incomplete information, we still can make quantitative predictions. Firstly, we have the mean field approximation (MFA). The MFA allows making predictions on aggregate or average quantities, involving the whole community of coexisting species. This has a direct application in agriculture to predict metrics that serve to quantify the performance of mixtures of crops grown for biomass, forage, or food production. Secondly, for such cases in which all the interaction parameters involving a particular species are available, we have the focal species approximation which can be used for predicting the yield of this focal species. Again, agriculture offers an example of a field in which this approximation can be used in the form of a crop coexisting with several species of weeds.

REFERENCES

- Aarssen, L.W. 1988. Pecking order of species from pastures of different ages. *Oikos* **51**, 3-12.
- Abrams, P. A. 1983. Arguments in favor of higher order interactions. *Am. Nat.* **121**, 887-891.
- Arbiza, J., Mirazo, S. and Fort, H. 2010. Viral quasispecies profiles as the result of the interplay of competition and cooperation, *BMC Evolutionary Biology* **2010**, 10:137.
- Beckage, B., and Gross, L. J. 2006. Overyielding and species diversity: what should we expect? *New Phytologist*, **172**, 140–148.
- Brown, J.H., T.G. Whitham, S.K.M. Ernest, C.A. Gehring. 2001 Complex species interactions and the dynamics of ecological systems: long-term experiments. *Science* **293**, 643-650.
- Burton, T. A. 1969 On the construction of Lyapunov functions. *SIAM J. Appl. Math.* **17** 1078-1085.
- Cardinale, B. J. *et al.* 2012. Biodiversity loss and its impact on humanity. *Nature* **486**, 59-67.
- Cardinale, B. J. *et al.* 2011. The functional role of producer diversity in ecosystems. *Am. Jour. of Bot.* **98**, 572-592.
- Downing, *et al.* 2012. The resilience and resistance of an ecosystem to a collapse of diversity. *PloS one* **7**, e46135.
- Fort, H. 2020a *Ecological Modelling and Ecophysics: Agricultural and environmental applications* (IOP ebooks), IOP, Bristol, UK.
- Fort, H. 2020b Making quantitative predictions on the yield of a species immersed in a multispecies community: The focal species method. *Ecological Modelling* **430**, 109108.
- Fort, H. 2018a On predicting species yields in multispecies communities: Quantifying the accuracy of the linear Lotka-Volterra generalized model. *Ecological Modelling* **387**, 154–162.
- Fort, H. 2018b Quantitative predictions from competition theory with an incomplete knowledge of model parameters tested against experiments across diverse taxa. *Ecological Modelling* **368**, 104–110.
- Fort, H. & Segura, A. 2018 Competition across diverse taxa: quantitative integration of theory and empirical research using global indices of competition. *Oikos* **127**, 392-402.
- Goh, B. S. 1977 Global stability in many species systems. *The American Naturalist*, **111**, 135-143.
- Goh, B. S. 1980 *Management and analysis of biological populations*. Elsevier, Amsterdam.
- Guglielmini, A.C., Verdú, A.M.C., Satorre, E.H. 2017 Competitive ability of five common weed species in competition with soybean. *Int. J. of Pest Management* **63**, 30-36.

Gurel, O. and Lapidus, L. 1968 Stability via Liapunov's second method. *Ind. Eng. Chem.* **60**, 1-26.

Hafstein, S. F. 2007 An algorithm for constructing Lyapunov functions. *Electronic Journal of Differential Equations*, Monograph **08**.

URL: <http://ejde.math.txstate.edu> or <http://ejde.math.unt.edu>

Halty, V. *et al.* 2017 Modelling plant interspecific interactions from experiments of perennial crop mixtures to predict optimal combinations, *Ecological Applications* **27**, 2277-2289.

Hector, A. *et al.* 2010 Ecological Archives E091-155-S1.

<http://www.esapubs.org/archive/ecol/E091/155/>

Hofbauer M. and Sigmund, K. 1998 *Evolutionary Games and Population Dynamics*, Cambridge University Press, Cambridge, UK.

Holmgren, M., Scheffer, M. and Huston, M.A. 1997. The interplay of facilitation and competition in plant communities. *Ecology* **78**, 1966–1975.

Hooper, D. U. & Dukes, J. S. 2004 Overyielding among plant functional groups in a long-term experiment *Ecol. Lett.* **7**, 95-105.

Hubbel, S.P. 2002 *The Unified Neutral Theory of Biodiversity and Biogeography*. Princeton Univ. Press, Princeton, N.J.

Huisman, J., Jonker, R.R., Zonneveld, C. & Weissing, F.J. 1999 Competition for light between phytoplankton species: experimental tests of mechanistic theory. *Ecology* **80**, 211–222.

Kastendiek, J. 1982 Competitor-mediated coexistence: interactions among three species of macroalgae. *J. Exp. Mar. Biol. Ecol.*, **62**, 201-210.

Keddy, P.A. 2001. *Competition* (2nd edition). Kluwer Academic Publishers.

LaSalle, J. and Lefschetz, S. 1961 *Stability by Liapunov's Direct Method*. Academic Press, New York, N.Y.

Legates, D. R. & McCabe, G. J. 1999 Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation. *Water Resources Research*, **35**, 233-241.

Levins, R. 1968 *Evolution in changing environments: Some theoretical explorations*. Princeton Unive. Press, Princeton, NJ.

Li, J. 2017 Assessing the accuracy of predictive models for numerical data: Not r nor r^2 , why not? Then what? *PLoS ONE* **12**(8): e0183250.

Li, J. & Heap, A. 2008 A Review of Spatial Interpolation Methods for Environmental Scientists. *Geoscience Australia, Record 2008/23*, 137pp.

Loreau, M. 1998. Biodiversity and ecosystem functioning: A mechanistic model. *PNAS* **95**, 5632-5636.

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

Lotka, A. J. 1925 *Elements of Physical Biology*. Williams and Wilkins, Baltimore.

May, R. M. 1974 *Stability and Complexity in Model Ecosystems*. Princeton University, Princeton, NJ.

Morin, P.J. 2011 *Community Ecology*. Wiley-Blackwell, Chichester, UK.

Nash, J. E., and Sutcliffe, J. V. 1970 River flow forecasting through conceptual models, I, A discussion of principles. *J. Hydrol.*, **10**, 282-290.

Neill, W.E. 1974 The Community Matrix and Interdependence of the competition coefficients. *Am. Nat.* **108**, 399-408.

Novak, M. *et al.* 2016 Characterizing species interactions to understand press perturbations: What is the community matrix? *Annu. Rev. Ecol. Evol. Syst.* **47**, 409-432.

Pastor, J. 2008 *Mathematical ecology of populations and ecosystems*. Wiley-Blackwell.

Picasso, V., Brummer, E. C., Liebman, M., Dixon, P. M. & Wilsey, B. J. 2008 Crop species diversity affects productivity and weed suppression in perennial polycultures under two management strategies. *Crop. Sci.* **48**, 331.

Rees, M.P. *et al.* 1996. Quantifying the impact of competition and spatial heterogeneity on the structure and dynamics of a four-species guild of winter annuals. *Am. Nat.* **147**, 1-32.

Roxburgh, S.H. & Wilson, J.B. 2000 Stability and coexistence in a lawn community: mathematical prediction of stability using a community matrix with parameters derived from competition experiments. *Oikos* **88**, 395-408.

Schnitzer *et al.* 2011 Soil microbes drive the classic plant diversity-productivity pattern. *Ecology* **92**, 296-303.

Schultz, D.G. 1965 The generation of Liapunov functions. In: C T. Leondes (Editor), *Advances in Control Systems*, Vol. 2. Academic Press, New York, N.Y., pp. 1—64.

Simberloff, D. 1982 The status of competition theory in ecology. *Ann. Zool. Fenn.*, **19**, 241–253.

Strogatz, S. H. 1994 *Nonlinear dynamics and chaos. With Applications to Physics, Biology, Chemistry, and Engineering* Perseus Books, Reading, Massachusetts.

Turner, P. E. and Chao, L. 1999 Prisoner's dilemma in an RNA virus. *Nature* **398**, 441–443.

Vandermeer, J.H. 1969 The competitive structure of communities: an experimental approach with protozoa. *Ecology*, **50**, 362–371.

Vandermeer, J.H. 1989 *The Ecology of Intercropping*. Cambridge University Press, Cambridge.

Vandermeer, J. H., & Goldberg, D. E. 2013 *Population ecology: First principles*. 2nd Edition. Princeton: Princeton University Press.

Volterra, V. 1931 *Leçons sur la Théorie Mathématique de la Lutte pour la Vie*. Gauthier-Villars, Paris.

Volterra, V. 1926 Fluctuations in the abundance of a species considered mathematically. *Nature* **118**, 558-560.

Wiens, J. A. 1984 On understanding a non-equilibrium world: myth and reality in community patterns and processes. In: *Ecological communities: conceptual issues and the evidence*. Princeton Univ. Press, Princeton, NJ, pp. 439–457.

Wilcox, B. P., Rawls, W. J. , Brakensiek, D. L. and Wight, J. R. 1990 Predicting runoff from rangeland catchments A: comparison of two models, *Water Resour. Res.*, **26**, 2401-2410.

Willems, J. L. 1970 *Stability Theory of Dynamical Systems*. Nelson, London.

Willmott, C. J. 1984 On the evaluation of model performance in physical geography, in *Spatial Statistics and Models*, edited by G. L. Gaile, and C. J. Willmott, pp. 443-460, D. Reidel, Norwell, Mass.

Willmott, C. J. 1981 On the validation of models, *Phys. Geogr.*, **2**, 184-194.

Wilson, S. D. & Keddy, P. A. 1986 Species competitive ability and position along a natural stress/disturbance gradient. *Ecology*, **67**, 1236-1242.

de Wit, C. T. 1970 *Proc. Adv. Study Inst Dyn. Numbers Popul.* 269.

Zarnetske P. L., Gouhier T. C., Hacker S. D., Seabloom E.W., Bokil V. A. 2013 Indirect effects and facilitation among native and non-native species promote invasion success along an environmental stress gradient. *Journal of Ecology*, **101**, 905–915