TOWARDS QUANTITATIVE FORECASTING OF SPECIES YIELDS WITH INCOMPLETE INFORMATION ON MODEL PARAMETERS

Hugo Fort

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

Abstract

Predicting both the absolute and the relative abundance of species in a spatial patch is of paramount interest in areas like, agriculture, ecology and environmental science.

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 single trophic ecological communities, like mixtures of plants, with $S > 2$ species. Thus, by estimating the LLVGE parameters from the yields in monoculture and biculture experiments, the LLVGE are able to produce decently 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.

1 INTRODUCTION

The ability to predict the absolute and the relative abundance of species in a place is of paramount interest in areas like, agriculture, ecology and environmental science. Indeed, abundance is an ecological quantity of crucial importance when making management and conservation decisions (Andrewartha and Birch 1954; Krebs 1978; Gaston 1994). For species of value to humans, estimates of their future yields will be essential for regulating current and future harvests. For species of conservation concern, their future populations will be an important determinant of their extinction risk.

Interspecific interactions strongly influence the absolute and relative abundances of species within a patch (Sousa 1984). In the case of agricultural crops, the mathematical methods commonly employed to forecast the yield have been mostly linear statistical models (NASS-SMB 2012). However, the interest in phenomenological modeling of biomass production and yield of crops has increased recently (Marcelis et al. 1998; Kirwan et al. 2009; Halty et al. 2017). Phenomenological models have in general a short computing time and they usually contain few state variables, making them helpful tools for reliable predictions.

Here we start by briefly reviewing the linear Lotka-Volterra *generalized* equations (LLVGE) (Pastor 2008; Fort 2020a) which serve for describing the dynamics of communities of species connected by negative as well as positive interspecific interactions. More specifically, we particularize these LLVGE to the case of a single trophic level community with $S > 2$ species, either artificial or natural.

However, a common situation we face is that we don't know all the parameters appearing in the LLVGE. Therefore, we 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.

2 THE LOTKA- VOLTERRA GENERALIZED LINEAR MODEL

If n_i denotes the *yield* of species i (e.g. its density or its biomass density) the LLVCE for S coexisting species consist in a set of linear equations for the per capita growth rates $\frac{1}{n_i} \frac{dn_i}{dt}$ given by:

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

where r_i is the intrinsic growth rate of species i , with dimension of time^{-1} , and I_{ij} is an interaction coefficient, with dimension of yield^{-1} , quantifying the *per capita* effect of species j on species i through their pairwise interaction. This is a phenomenological model since it describes how the abundance of one species affects the abundance of another, without specifically including a particular mechanism for such interaction (Morin 2011). ‘Generalized’ means that all the combinations of pairs of signs for both species ($-/-$, $-/+$, $-/0$, $+/0$, $+/+$) are possible.

2.1 The linear Lotka-Volterra generalized equations for single trophic communities

In the case of a single trophic community, the trophic level below the one of the community (e.g. plants for a community of herbivore species), corresponding to the food required for individuals of a species i to thrive, is modeled through a parameter $\mathcal{K}_i$ known as the *carrying capacity* (Pastor 2008). $\mathcal{K}_i$ is interpreted as the maximum sustainable yield for species i in monoculture, *i.e.* when the species is isolated from the rest of the other species making up the community. It turns out that I_{ij} can be thus written in terms of a non-dimensional interaction coefficient α_{ij} (Fort 2020):

$$I_{ij} = \alpha_{ij} / \mathcal{K}_i. \quad (2)$$

It is customary to take the intraspecific competition coefficients $\alpha_{ii} = -1$. Hence, by (1) and (2) the LLVGE become:

$$\frac{dn_i}{dt} = r_i n_i \left(1 + \frac{\sum_{j=1}^S \alpha_{ij} n_j}{\mathcal{K}_i} \right) \quad i = 1, \dots, S \quad (\alpha_{ii} = -1). \quad (3)$$

2.2 Obtaining the model parameters from monoculture and biculture experiments

A straightforward procedure to get the set of model parameters $\{\mathcal{K}_i, \alpha_{ij}\}$ is to perform, until the equilibrium state is attained: 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 $\mathcal{K}_i = m_i^{\text{ex}*}$ (the superscript ‘ex’ is to emphasize that this is an experimentally measured quantity and the star denotes equilibrium); b) the $S \times (S-1)/2$ pairwise experiments producing the *biculture* yields, $n_{i(j)}^{\text{ex}*}$ and $n_{j(i)}^{\text{ex}*}$ (the subscripts $i(j)$ and $j(i)$ stand for the yield of species i in presence of species j and *vice versa*). We then obtain α_{ij} and α_{ji} , as (Fort 2020a): $\alpha_{ij} = (n_{i(j)}^{\text{ex}*} - m_i^{\text{ex}*}) / n_{j(i)}^{\text{ex}*}$, $\alpha_{ji} = (n_{j(i)}^{\text{ex}*} - m_j^{\text{ex}*}) / n_{i(j)}^{\text{ex}*}$.

3 WORKING WITH IMPERFECT INFORMATION

A main limitation of the above experimental procedure 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.

However, there are situations in which we can still predict some relevant quantities about species yields with an incomplete knowledge of the LLVGE parameters. With this aim, it is convenient to work with relative yields

$$y_i = n_i / \mathcal{K}_i, \quad (4)$$

i.e. the species yield in mixture normalized by its yield in monoculture. It turns out that the equilibrium¹ relative yields verify (Fort 2020a):

$$y_i \left(\sum_{j=1}^S a_{ij} y_j + 1 \right), \quad i = 1, \dots, S, \quad (5)$$

where

$$a_{ij} = \alpha_{ij} \mathcal{K}_j / \mathcal{K}_i. \quad (6)$$

Notice thus that the diagonal elements of the matrix $\mathbf{A}=[a_{ij}]$, corresponding to intra-specific competition, remain equal to -1 since $a_{ii}=\alpha_{ii}\mathcal{K}_i/\mathcal{K}_i=\alpha_{ii}=-1$.

Let us present a couple of recently introduced methods for making quantitative predictions with incomplete information on the matrix $\mathbf{A}$, motivating each method through a different practical problem in crop science.

First, we can think about mixtures of 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). E.g., a popular aggregate metric used to quantify the overyielding of diverse crop mixtures relative to crop monocultures is the Relative Yield Total (*RYT*) (de Wit 1970):

$$RYT = \sum_{i=1}^S y_i, \quad (7)$$

A $RYT > 1$ implies that the yield performance will be better in polyculture than in monoculture (Vandermeer 1989). Another index directly connected with the *RYT* is the Mean Relative Yield (*MRY*),

$$MRY = RYT / S = \sum_{i=1}^S y_i / S. \quad (8)$$

In this case we can use the recently proposed *mean field* approximation (MFA) (Fort 2018b; Fort & Segura 2018; Fort 2020a). The name alludes to the resemblance of this approximation with the one 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. That is, the off-diagonal part of the interaction matrix $[a_{ij}]$ is replaced by a “mean field” competition coefficient given by the average over the sample of available interspecific competition coefficients:

$$a = \overline{a_{i \neq j}}. \quad (9)$$

In such a way we get a *mean field matrix*:

¹ Thereafter, we will omit the * for all yields in the understanding that they are yields in equilibrium.

$$\mathbf{A}^{MF} \equiv [a_{ij}]^{MF} = \begin{pmatrix} -1 & a & \cdots & a & a \\ a & -1 & \cdots & a & a \\ & & \ddots & & \\ a & a & \cdots & -1 & a \\ a & a & \cdots & a & -1 \end{pmatrix}. \quad (10)$$

The MFA allows expressing the RYT and MRY as a simple functions of a and S (Fort 2018b, Fort 2020a):

$$RYT^{MF}(a, S) = \frac{S}{1 - (S-1)a}, \quad (11)$$

$$MRY^{MF}(a, S) \square \frac{1}{1 - (S-1)a}, \quad (12)$$

Thus the MFA, through Eqs. (11) and (12), is particularly suited to address an important question in ecology, namely, how the total or mean relative yield depends with the species richness and the mean intensity of competition. Comparisons between predictions generated by formula (11) against data from a dataset of 25 experiments involving $S > 2$ species and $a < 0$ are shown in Fig. 1 as a function of both S and a (Fort 2018b).

Figure 1. Empirical (o) and theoretical RYT , predicted by the MFA Eq. (11) (red crosses), for a set of 25 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).

Second, suppose one is interested in predicting the yield of a *particular* species embedded in a community with other species belonging to the same trophic level. *E.g.* a crop coexisting with different species of weeds. Often we have quite detailed data on the crop species, like the interaction coefficients that involve this crop (Guglielmini 2016). An approximation that works in such case is the *focal species approximation* (FSA) (Fort 2020a, 2020b).

Since we are not interested in the yield of each individual weed species, *we can treat all of them as a single species*, and this implies to take the interspecific coefficients between species different from the focal species k equal to -1 (remember that $a_{ii} = -1$). Thus, one uses the *focal interaction matrix*:

$$\mathbf{A}^{(k)} \equiv [a_{ij}]^{(k)} = \begin{pmatrix} -1 & -1 & \cdots & a_{1k} & \cdots & -1 & -1 \\ -1 & -1 & \cdots & a_{2k} & \cdots & -1 & -1 \\ \vdots & \vdots & \ddots & \vdots & \cdots & \vdots & \vdots \\ a_{k1} & a_{k2} & \cdots & -1 & \cdots & a_{kS-1} & a_{kS} \\ \vdots & \vdots & \cdots & \vdots & \ddots & \vdots & \vdots \\ -1 & -1 & \cdots & a_{S-1k} & \cdots & -1 & -1 \\ -1 & -1 & \cdots & a_{Sk} & \cdots & -1 & -1 \end{pmatrix}, \quad (13)$$

where the 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 (13) to solve Eq. (5) 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 (13) we obtain just one relative yield for the focal species ($y_k^{(k)}$).

Figure 2 compares the results of using this approximation with those when using the full set of LLVGE (filled small circles) for four different experiments, three of them involving plants and one for algae.

Figure 2. The focal species approximation compared with the full set of LLVGE. Experimental relative yields with their error bars (crosses), predicted values 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).

4 CONCLUSION

Therefore we conclude that in cases of incomplete information on the LLVGE parameters:

- I. **Provided competition interactions among species is dominant, we can predict *aggregate* or *mean* relative competition indices, via de MFA, with an accuracy inversely proportional to the experimental variance of the experimental measurements.**

- II. **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 FSA, with accuracy comparable to the one obtained when using the full set of LLVGE.**

REFERENCES

- Andrewartha, H. G., and Birch, L. C. (1954) *The distribution and abundance of animals*. University of Chicago Press, Chicago.
- Beckage, B., and Gross, L. J. 2006. Overyielding and species diversity: what should we expect? *New Phytologist*, **172**, 140–148.
- 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.
- Gaston, K. J. (1994) *Rarity*. Chapman & Hall, London.
- 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.
- Halt, V. *et al.* (2017) Modelling plant interspecific interactions from experiments of perennial crop mixtures to predict optimal combinations, *Ecological Applications* **27**, 2277–2289.
- 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.
- Kirwan, L. *et al.* (2009) Diversity-interaction modeling: estimating contributions of species identities and interactions to ecosystem function. *Ecology* **90**, 2032–2038.
- Krebs, C. J. (1978) *Ecology: the experimental analysis of distribution and abundance*. 2d ed. Harper & Row, New York.
- Lotka, A. J. 1925 *Elements of Physical Biology*. Williams and Wilkins, Baltimore.
- Marcelis, L.F.M., Heuvelink, E. and Goudriaanc, J. (1998) Modelling biomass production and yield of horticultural crops: a review. *Scientia Horticulturae* **74**, 83–111.
- Morin, P.J. (2011) *Community Ecology*. Wiley-Blackwell, Chichester, UK.

NASS-SMB (2012) The Yield Forecasting and Estimating Program of NASS, Statistical Methods Branch (SMB), Statistics Division, National Agricultural Statistics Service, U.S. Department of Agriculture, Washington, D.C., NASS Staff Report No. SMB 12-01.

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.

Sousa, W. P. 1984. The role of disturbance in natural communities. *Ann. Rev. Ecol. Syst.* **15**:353-391.

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

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

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.