

A measure of generalized soil fertility that is largely independent of species identity

Xavier Lamontagne and Bill Shipley*,[✉]

Département de biologie, Université de Sherbrooke, Sherbrooke (Qc), J1K 2R9, Canada

** For correspondence. E-mail Bill.Shipley@Usherbrooke.ca*

Received: 8 July 2021 Returned for revision: 24 August 2021 Editorial decision: 13 September 2021 Accepted: 17 September 2021
Electronically published: 20 September 2021

- **Background and Aims** In 2019, Daou and Shipley produced an operational definition of ‘generalized’ soil fertility (F_G) for plant community ecology and quantified F_G using a structural equation model (SEM) invoking a single latent variable. We evaluate a critical assumption of this model: that F_G is generalizable to any combination of plant species; i.e. that any combination of plant species will respond in the same direction to the soil ‘fertility’ gradient in terms of growth.
- **Methods** We grew nine widely different species singly in each of 25 soils from southern Quebec, Canada, whose F_G value had been previously quantified. The original SEM was tested using every possible combination involving from four to nine species.
- **Key Results** The assumption was rejected due to a subset of three species that responded to a second latent dimension. We then proposed an alternative model that includes F_G plus a second latent variable that measures species’ deviations from F_G due to specific adaptations to soil pH. This alternative model was consistent with every combination of up to eight species. The predictions of F_G when ignoring this second dimension and when using the new model were extremely correlated ($r=0.98$).
- **Conclusions** The initial unidimensional model of Daou and Shipley was successful in non-acid soils but not in soils with extreme pH and when species specifically adapted to such extreme soils were included. The alternative two-dimensional model takes into account these exceptions and is consistent with the notion of shared physiological niche responses along a gradient of generalized soil fertility.

Key words: Structural equation modelling, SEM, latent variable measurement model, plant communities, soil fertility gradient, relative growth rate, shared physiological preference, inclusive fundamental niche.

INTRODUCTION

Trait-based plant ecology has captured the attention of researchers because it promises generalizable predictions across ecological and spatial scales that are independent of taxonomic identity (Shipley *et al.*, 2016; Laliberté, 2017) although this promise has not yet been completely realized (Shipley *et al.*, 2016). Functional ecologists have developed standardized protocols to measure traits (Pérez-Harguindeguy *et al.*, 2013). This standardization of trait measurements is an essential element in the ability of trait-based ecology to generalize across species and environments and to create worldwide databases (Fraser, 2020) that allow researchers to compare trait measurements and reveal global patterns in trait variation.

However, the standardization of trait measurements is only half of the challenge since trait-based ecology also attempts to link variation in traits to variation in environmental variables. The current inability to define and measure environmental gradients using standardized methods has been a huge barrier for model development in plant community ecology (Austin, 1980; Austin and Smith, 1990; Garnier *et al.*, 2016). This is because an environmental gradient in community ecology is not synonymous with an environmental variable. To be useful and generalizable, the environmental variables that define an environmental gradient must be measured in a way that accords

with how plants, not physical or chemical measuring devices, respond to them and in a way that is comparable and generalizable across studies, geographical locations and species assemblages (Shipley *et al.*, 2016).

For example, the most influential non-neutral theories of plant community structure (Grime, 1977; Tilman, 1988) invoke the notion of a gradient of site ‘fertility’ (i.e. the ability of a local environment to produce plant growth) along which the assumed mechanisms of community assembly change systematically. Both Grime and Tilman agree that soil properties affecting the ability of a local environment to produce plant growth (hereafter, soil fertility) are key components of site fertility. Soil fertility is here defined as the amount of plant growth per unit time that a soil can support when holding constant all other components of site fertility. However, it is extremely difficult to define and measure soil fertility at the level of plant communities. Soil ‘fertility’ in ecology has been typically quantified by measuring its causes: concentrations or fluxes of various macro- or micro-nutrients, organic matter, pH, water-holding capacity and combinations of all these variables; see Daou and Shipley (2019) for a large but incomplete list. These cause-based soil ‘fertility’ metrics are not generalizable for two main reasons. First, different soils can sustain the same level of plant growth due to different combinations of causes. Second, the link between soil

properties and plant growth is species-dependent and different species in the same soil are often co-limited by different nutrients or nutrient ratios (Harpole *et al.*, 2011).

Daou and Shipley (2019) proposed a new method for quantifying soil fertility at the scale of plant communities. Their method focused on the *effect* of soils on plant growth of target species in controlled conditions rather than on soil properties that might *cause* differences in plant growth. They did this by combining the phytometer approach (Clements and Goldsmith, 1924) with a latent variable measurement model (Grace, 2006; Bollen, 2014; Shipley, 2016) in the context of structural equation modelling (SEM). They hypothesized that all plant species will experience the same ‘growth potential’ in the same soil, but each species will manifest this potential to a different degree. Thus, while each species will have a different growth rate in the same soil, all species will either increase or decrease growth together if this soil growth potential changes. They called this latent growth potential of a soil that is common to all species the ‘generalized’ fertility (F_G) of the soil. This hypothesis, translated into a causal diagram, is shown in Fig. 1 and, translated into an SEM, requires that all covariation between the growth responses of different species (relative growth rate, RGR) across different soils be caused by a single latent (unmeasured) causal property, i.e. F_G . If this assumption of a single latent cause is not true then the SEM will provide a poor fit to the empirical data and the SEM will be rejected. They tested and calibrated their model using four different species having large ecological differences but whose seeds are available worldwide (*Triticum aestivum*, *Festuca rubra*, *Arabidopsis thaliana*, *Trifolium pratense*) and that were grown in 76 soils in southern Quebec (Canada). The main result of their study is a method that quantifies soil fertility independently of the

underlying causes and this makes the model geographically generalizable. See Daou and Shipley (2019) for a more detailed explanation of the latent measurement model.

However, the Daou and Shipley (2019) model makes an important untested assumption: it assumes that any plant species, not just the four chosen phytometer species, will respond in the same direction to the soil ‘fertility’ gradient in terms of growth. If this is not true then the level of fertility of a soil will differ depending on the choice of phytometers, thus limiting the ‘generality’ of this fertility concept and its application to plant community ecology. The goal of the present paper is to provide a critical test of this untested assumption by asking three questions.

Question 1. Is the original model supported independently of the choice and of the number of phytometer species included in it? If yes, an SEM latent measurement model with only one latent variable will not be rejected independently of the species included within it. If the model is rejected then this leads to the following two questions.

Question 2. How many dimensions exist within our data and in which conditions of soils/species are those dimensions manifested?

Question 3. If there is more than one dimension, what is the quantitative effect on the predicted level of soil fertility if the original unidimensional Daou and Shipley (2019) model is used while ignoring the other dimensions?

METHODS

To evaluate our questions, we need to select plant species that maximize diverse responses across the hypothesized ‘fertility’ gradient while still being able to grow them in the same non-soil abiotic conditions (light intensity, temperature, humidity, water inputs). For practical reasons we selected nine herbaceous species of open grassland habitats and chose species that differed as much as possible in their field distributions with respect to soil resources (Appendix: Supplementary Data Table S1). These species were selected based on three criteria: (1) Ellenberg’s (1974) nitrogen (or nutrient) index, consisting of a subjective rank from 1 (*Festuca ovina* L.; ‘plant typical of extremely infertile sites’) to 9 (*Urtica dioica* L.; ‘plant typical of extremely rich sites such as cattle resting places or near polluted rivers’); Grime’s (1977) Competitor-Stress tolerator-Ruderal (CSR) strategies calculated using STRATEFY (Pierce *et al.*, 2017); and (3) specific adaptations to soil resource capture such as the intensity of mycorrhizal associations and *Rhizobium* symbiosis. Two of the species [*Urtica dioica* and *Capsella bursa-pastoris* (L.) Medik] lack mycorrhizal associations and one species (*Melilotus albus* Medik) is symbiotic with *Rhizobium*. The seeds of *Centaurea nigra*, *Melilotus albus* and *Chenopodium album* were collected from populations around the campus of the University of Sherbrooke and seeds of the other species came from commercial suppliers. More details on the ecology of the phytometers is available in Appendix 1: Table S1.

We used 23 of the 76 sites of Daou and Shipley (2019), which consisted of 76 soils from uncultivated ‘grasslands’ covering a range of ~4900 km² in southern Quebec (Canada). Those sites were chosen from pedological maps (Cann and Gérin-Lajoie, 1942; Lamontagne *et al.*, 2002) to be representative of the

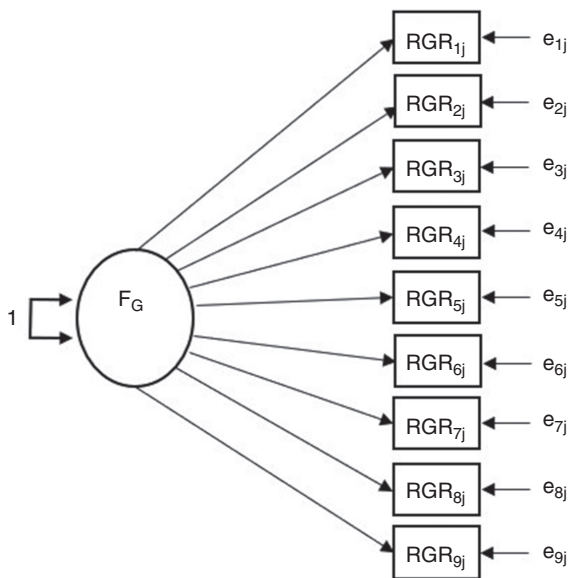


FIG. 1. Theoretical model illustrating the hypothesis of Daou and Shipley (2019). A latent variable named generalized fertility (F_G) causes variation in growth for multiple species (RGR_1 through RGR_9) in multiple soils ‘j’. The e ’s represent the mutually independent residual errors. The variance of F_G is fixed to 1 so that it is a standardized variable. Combinations of this model consist of using subsets of multiple RGRs (4 minimally due to degrees of freedom constraints).

diversity of soils within the region (Appendix 1: [Supplementary Data Table S2](#)). Our 23 sites, currently non-agricultural grasslands, wastelands and sand quarries due to anthropogenic activity, were chosen to cover the range of F_G values previously calculated by [Daou and Shipley \(2019\)](#) from the least to the most fertile. We added two soils to extend the two extremes of the gradient. A commercial horticultural soil (Black Earth, Fafard's, *Agro Mix*) produced to a maximally fertile soil (0.5 % total nitrogen, 0.07 % available P_2O_5 , 0.1 % soluble K_2O) and non-acidogenic mine tailings (essentially crushed bedrock) from northeastern Quebec produced a maximally infertile site making a total of 25 different soils for the study (see [Table S2](#) for more details). We collected soil in the top layer (0–15 cm) from each site. The soils were dried at room temperature, passed through a 2-mm sieve and then stored at 4 °C in pots (8 × 8 × 9 cm) prior to the phytometer experiment (below). One kilogram of each soil, stored in the same way, was kept to measure soil physical and chemical properties.

Each combination of species and soils was replicated four times giving 900 pots (9 species × 25 soils × 4 replicate pots). Because of limited space, we separated the experiment in three blocks each containing every combination of species and soils. The three blocks contained 306, 306 and 288 pots, respectively. Eighteen pots were randomly placed in each of 17 plastic boxes (60.5 × 40.95 × 34.5 cm) with lids and these boxes were randomly placed in a growth room. Each plastic box had two 5.5-cm-diameter holes to allow airflow and a perforated plastic platform suspended 5 cm from the bottom onto which the pots were placed. A small amount of free-standing water was maintained below the perforated platform to control humidity of the air. Within each box, temperature and relative humidity were monitored at 30-min intervals using logtag probes (logtagrecorders.com). Using these temperature and relative humidity values, we also calculated the vapour pressure deficit of the air following [Jones \(1992\)](#). Temperature was 28.7 ± 2.9 °C (mean ± s.d.), relative humidity was 90.2 ± 8.1 % and vapour pressure deficit was 0.41 ± 0.35 kPa. Two 1000-W high-pressure sodium lamps provided a mean daily quantum yield of $9.8 \text{ mol m}^{-2} \text{ d}^{-1}$ over a 16 : 8 h (day : night) cycle. The pots were watered with deionized water 1 week before planting the seeds to let the soils stabilize. Enough seeds were planted to produce four to six individuals per pot. During growth, excess seedlings were thinned to maximize the distance between individuals while still keeping four to six individuals per pot. All the pots were watered to field capacity with deionized water three times per week with sequential watering separated by an interval of 36–39 h of light. This spacing between watering intervals was chosen to allow the natural variation of soil water-holding capacity to be expressed while avoiding strong water limitation. Individuals of each species were harvested at one of two harvest dates; these two dates varied between species and were chosen to (1) allow sufficient growth between the two harvests dates, (2) allow sufficient time for seedlings to become independent of seed reserves while (3) minimizing the total growth period of each species. These harvest dates (days after germination) are (13, 26) for *Chenopodium album*, (14, 28) for *Capsella bursa-pastoris*, (15, 28) for *Centaurea nigra*, (25, 50) for *Festuca ovina*, (17, 28) for *Melilotus albus*, (22, 39) for *Pilosella aurantiaca*, (24, 45) for *Sedum acre*, (6,

20) for *Triticum aestivum* and (25, 40) for *Urtica dioica*. At each harvest date, the above-ground biomass of two or three individuals (two individuals for the naturally larger plants of *T. aestivum*, *M. albus* and *C. nigra*) were randomly harvested, dried at 60 °C for at least 2 weeks and weighed. We calculated RGR ($\text{mg g}^{-1} \text{ d}^{-1}$), as explained below, of each species–soil combination from these dried biomasses. Due to a COVID-19 interruption, the second harvests of *F. ovina*, *P. aurantiaca*, *S. acre* and *U. dioica* were not done in the first block.

We measured four soil variables from the 2-mm sieved soils used in the phytometer experiment. Particle size distribution (sand, silt and clay) was determined using the hydrometer method ([Kroetsch and Wang, 2007](#)). We used 40 g of mineral soil (burned at 400 °C for 16 h) with 50 mL of a dispersal solution (93 g of sodium hexametaphosphate and 7 g of sodium bicarbonate per litre of water). Organic matter content was measured based loss on ignition. The soils were first dried at 105 °C for 24 h and then we burned 2 g of soil at 400 °C for 16 h following [Ben-Dor and Banin \(1989\)](#). Soil water-holding capacity was determined using the gravimetric method (g H_2O /g dried soil) 24 h after saturation while preventing evaporative loss. Lastly, soil pH was measured in deionized water using a 1 : 2 (soil : water) ratio for mineral soils and 1 : 10 ratio for organic soils ([Hendershot et al. 2007](#)). Solutions were stirred for 30 min and pH measurements were taken 30 min after stabilization.

We estimated RGR in each soil by constructing a mixed-effects model using the lme4 package in R ([Bates et al., 2015](#)). We did this separately for each species to ensure statistical independence. The models were specified as $\ln(\text{biomass}) \sim \text{age} (\text{fixed effect, days}) + \text{soil} (\text{random effect})$. We also included the nested effect of the boxes, the pots and the blocks for each species if the model was convergent and not singular. These random effects were generally absent or very small compared to the effect of the soils. The slope associated with age was random across soils but the intercept (i.e. seedling mass at germination) was fixed since this parameter cannot vary between soils. RGR in each soil was quantified as the random slope estimate associated with each soil via the `coef()` function associated with the `lmer()` function. These random slope estimates are best linear unbiased predictors (BLUP).

Our first objective was to test the hypothesis that the patterns of covariation between the RGR values of the different species across the different soils are all caused by a single latent variable ('generalized' fertility, F_G , [Fig. 1](#)). We tested this by translating this graphical causal model into a maximum-likelihood SEM ([Shipley, 2016](#)) in which the RGR values of each species are caused by a single latent variable representing 'generalized soil fertility'. This model was tested using the `sem()` function of the lavaan package of R ([Rosseel 2012](#)). A null probability of less than 0.05 results in the rejection of our hypothesis. Four indicators (i.e. RGR values of four species) are the minimum number required in the SEM to have positive degrees of freedom. We therefore tested this hypothesis 382 times using all possible combinations of from four to nine species. All possible species combinations were generated with the `combinations()` function of `gtools` ([Warnes et al., 2015](#)). Next, because the [Daou and Shipley \(2019\)](#) hypothesis was rejected in certain circumstances, we proceeded to the other objectives mentioned in the Introduction.

The second objective consisted of two parts: (1) to determine the number of dimensions that existed in the RGR values of the phytometers in the different soils and (2) to determine the properties of the soils/species associated with these dimensions. To do this, we performed a principal component analysis (PCA) to explore the dimensionality of the RGR values of the nine species across the 25 different soils. We then searched for correlations between the axis scores of the PCA and the four measured soil properties. Next, we identified the combinations of species in the 382 different models that occurred in those models that were rejected ($P < 0.05$). Based on these results, we then proposed a second SEM. This model includes two latent variables influencing variation in RGR (Fig. 2). The generalized fertility not including the effects of pH (F_G^{-pH}) was the first latent variable whose variance was fixed to unity. We explicitly linked a second latent variable to soil pH by fixing the path coefficient from it to measured soil pH at unity. The covariance between the two latent variables was fixed to zero to ensure independence. The first latent was therefore a standardized variable having standard deviation units while the second latent was measured in pH units. We tested this alternative model using lavaan for all possible combinations of species. Because this new model has two latent variables, we could only test the combinations from five to nine species (252 models) due to degrees of freedom constraints.

Our third objective was to determine the amount of error introduced if the original model of Daou and Shipley is used even when it is statistically rejected; i.e. when ignoring any additional latent dimensions. To do so, we calculated the predicted value of the first latent variable, F_G in Fig. 1 and F_G^{-pH} in Fig. 2, using

the predict() function for each soil and for each species combination from five to nine species. Because we used many different models containing different indicator species and we required the latent variable to be measured on a comparable scale (standard deviation units), the latent variable was expressed in a common unit of standard deviations by fixing its variance to unity. We then measured the correlation between the two sets of predictions of this latent and the mean squared residual between the two sets.

All R code and the data files are available in [Supplementary Data S1.zip](#).

RESULTS

The variation in temperature, relative humidity and vapour pressure deficit of the air in the growth boxes were constant between blocks (Appendix 2: [Supplementary Data Figs S1–S3](#)). Although there were obvious differences in the average RGR value between species, there was just as much variation in RGR within each species across the different soils under the controlled conditions (Appendix 2: [Fig. S4](#)). The null probability associated with the hypothesized unidimensional SEM (Fig. 1) varied greatly depending on the species composition and the number of species (Appendix 2: [Fig. S5](#)). Most models involving only four species were not rejected but virtually all models involving six or more species were rejected. The patterns of covariations between the RGR values were not consistent with a unidimensional model involving a single latent variable as hypothesized by [Daou and Shipley \(2019\)](#). However, two of the models composed of seven species were accepted: the first model excludes *F. ovina* and *U. dioica* ($\chi^2 = 18.84$, 14

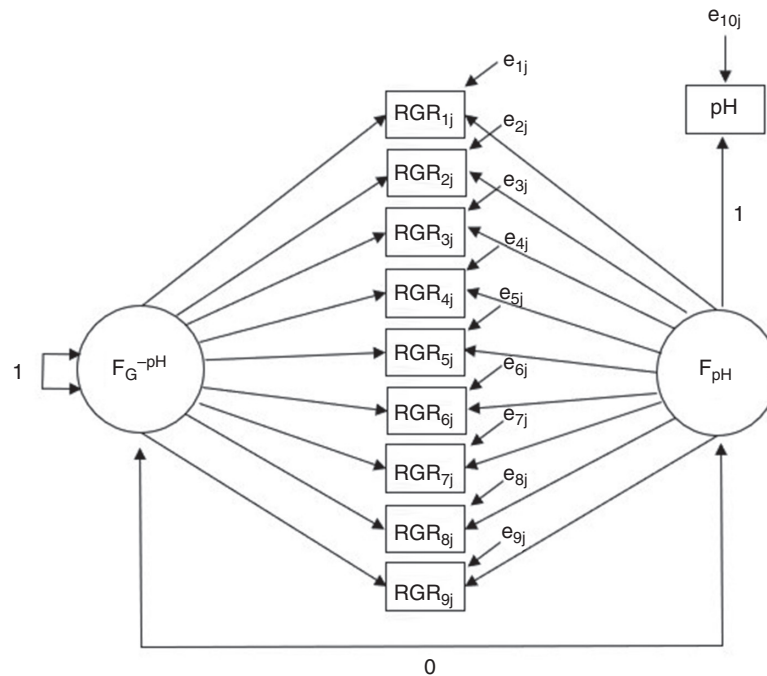


FIG. 2. Alternative theoretical model with two latent variables. F_G^{-pH} represents a first latent attribute, independent of soil pH, to which all species respond (RGR_{1j} to RGR_{9j}). F_{pH} represents a second latent attribute independent of F_G^{-pH} , which is the ‘true’ soil pH, which is measured with error from the laboratory pH, to which all species respond. The e ’s represent the errors e_{1j} to e_{10j} . Here the variance of F_G^{-pH} is fixed to 1 and the variance of F_{pH} is free while its slope to laboratory-measured soil pH is fixed to unity. Combinations of this alternative model consist of using a subset of multiple RGRs (five minimally due to degrees of freedom constraints).

d.f., $P = 0.171$) and the second model excludes *P. aurantiaca* and *U. dioica* ($\chi^2 = 19.124$, 14 d.f., $P = 0.160$).

The first dimension of the PCA accounted for 79.9 % of the variance in the RGR values of the phytometers across the different soils (Appendix 2: [Supplementary Data Fig. S6](#)). The second dimension accounted for an additional 7.8 % of the variance and the scores of the second dimension were highly correlated with the pH of the soils ($r = -0.758$, $P = 1.14e-05$). This second dimension was expressed primarily through the differential responses of three species. *Festuca ovina* and *P. aurantiaca* had higher RGR values than predicted using the model of [Daou and Shipley \(2019\)](#) in acidic soils and *S. acre* had higher RGR values than predicted in alkaline soils. The alternative model ([Fig. 2](#)) that includes a latent pH dimension fitted the data well (Appendix 2: [Fig. S7](#)) even with eight species ($\chi^2 = 20.321$, 19 d.f., $P = 0.376$, excluding *U. dioica*) although the alternative model was still rejected if *U. dioica* was included ($\chi^2 = 50.193$, 26 d.f., $P = 0.003$).

By how much would the predicted values of generalized fertility be wrong if one used the original model of [Daou and Shipley \(2019\)](#); that is, if one ignored the second latent variable associated with soil pH? The relationship between the predicted values of F_G in [Fig. 1](#) (i.e. when ignoring pH) and the predicted values of the first latent variable (F_G^{-pH}) in [Fig. 2](#) (i.e. when controlling for pH) is shown in [Fig. 3](#). The predicted values are extremely tightly correlated ($r = 0.98$, $t = 402.53$, $P = 2.2e-16$) and the residual sum of squares is low ($3.11e-15$).

DISCUSSION

A fundamental claim underlying the SEM measurement model by [Daou and Shipley \(2019\)](#) was that, in the absence

of interspecific competition, a soil that produces higher growth for one species would also produce higher growth for all other species when non-soil causes of growth are held constant. In other words, species must share a common linear response of their RGRs throughout the ‘fertility’ gradient. Such responses can differ quantitatively, but they need to be in the same direction. This is why [Daou and Shipley \(2019\)](#) called the single latent variable in their SEM measurement model ‘generalized’ fertility (F_G). F_G has a direct link to older theoretical niche concepts.

Hutchinson distinguished between ‘realized’ and ‘fundamental’ niches ([Hutchinson 1957](#)). The fundamental (or physiological) niche is the environmental space that a species would occupy in the absence of interspecific competition. The realized niche is the environmental space that a species would occupy in the presence of interspecific competition. It is empirically well established that the realized niche differs greatly between species and different species reach their fitness optima at a distinct point in this space. However, it is much less well known how fundamental niches are distributed in environmental space. The most common idea is that fundamental niches also have distinct fitness optima: i.e. each species has a different optimum in fundamental niche space and so there is no one point in fundamental niche space that is best for all species. This view directly contradicts the notion of a generalized soil fertility. A less common idea is the notion of fundamental niche inclusion, or ‘shared’ physiological responses ([Colwell and Fuentes, 1975](#)). Here, all species have highest fitness at the same point in the fundamental niche (a shared response) but there is a trade-off between the breadth of the fundamental niche (i.e. the range of suboptimal environmental conditions that still allow the species

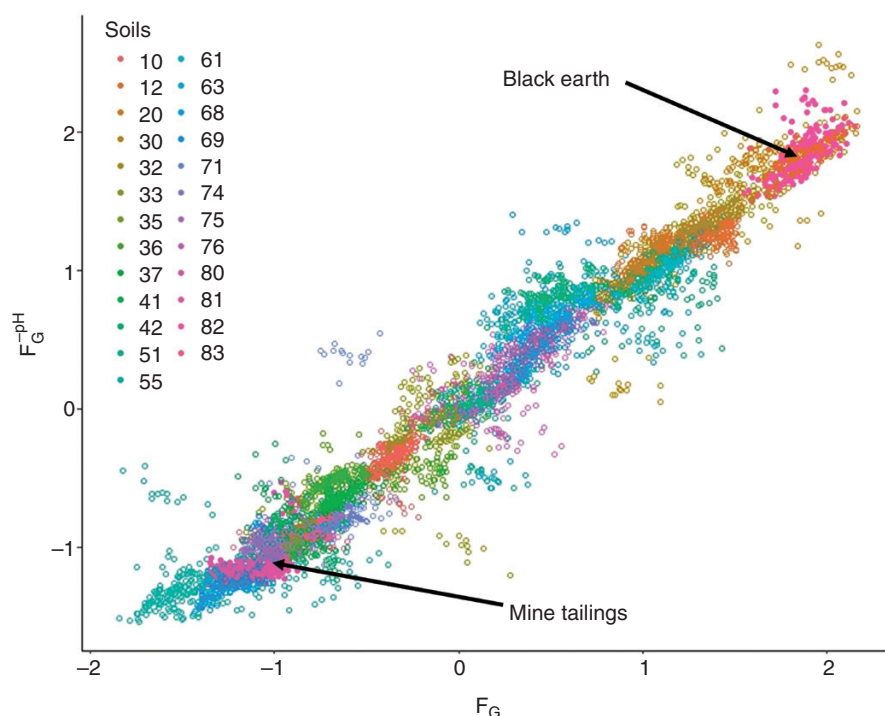


FIG. 3. Relationship between the predicted values of F_G (one-dimensional SEM) and F_G^{-pH} (first latent in a two-dimensional SEM) for 256 structural equation measurement models. Each colour represents a soil from which 256 predictions have been made. The two extreme soils are represented by solid circles (black earth = 82 and non-acidogenic mine tailings = 81).

to have positive fitness) and competitive ability to exclude other species from the preferred region of the fundamental niche. Species with the highest fitness at the preferred region along a shared fundamental niche axis will competitively exclude weaker species but such competitively dominant species, having a narrow tolerance for suboptimal points along this niche axis, are physiologically excluded at more suboptimal regions of environmental space. Species with lower fitness at the common preferred region of the niche axis have wider tolerance for suboptimal points and so are competitively displaced to these suboptimal points resulting in a competitive hierarchy (Keddy, 2001). The idea is also implicit in Grime's idea of the trade-off between a species position along stress and competition gradients. The notion of generalized fertility implicitly assumes that the various soil properties impacting plant growth (i.e. when all non-soil environmental variables are held constant) represent an example of shared responses to a unidimensional combination of fundamental niche axes of soils.

Researchers working with nutrient solutions (e.g. Austin and Austin, 1980; Fichtner and Schulze, 1992) have reported multiple species having a shared preference in their fundamental niche such that all species increase growth with increasing concentrations of limiting nutrients until each reaches a species-specific range of 'luxury' consumption, followed by decreases in growth due to nutrient toxicity. However, it is extremely difficult to apply such hydroponic experiments to real soils both because nutrient concentrations or fluxes in nutrient solutions are not comparable to those in real soils and because natural soils rarely, if ever, reach nutrient concentrations that are toxic. Reich *et al.* (2003), working in natural soils, found contrary results; four out of 32 species (two legumes and two C_4 grasses) decreased their RGR by 10–35 % following an N fertilization treatment while 20 species increased their RGR and eight species showed no, or only slight, increases in RGR. Raaimakers and Lambers (1996) documented one case of *Lecythis corrugata* with a low RGR_{max} not responding to increases in soil phosphorus while *Tapirira obtusa* with a high RGR_{max} strongly responding to the additions. As pointed out by Craine (2009) it is still unclear whether this result constitutes an experimental anomaly or an exception. If this exception were common then it would constitute a limitation of the fundamental shared niche preference of soil 'fertility'. Furthermore, even if these exceptions are more common, we still do not know if such results apply along natural gradients in soils since those studies used either applied single nutrient additions to a natural soil (N or P) or else used hydroponic nutrient solutions.

Daou and Shipley (2019), using 76 natural soils varying widely in above-ground biomass production, found that four species with contrasting ecology varied their RGR values in the same rank order across these soils, thus displaying a shared response to soil 'fertility'. However, those initial results were based on only four species. It is possible that a wider selection of species could manifest different, and distinct, responses. In our results, all nine species with contrasting ecology responded to the same soil 'fertility' by increasing their RGRs. However, this was not true in all soils and for all species, resulting in the model of Daou and Shipley (2019) being increasingly rejected with an increase in the number of species in the combination. Extreme soils in terms of pH generated a statistically detectable

bidimensional response for three species that previous authors (Tyler, 1996; Czarapata, 2005; Grime *et al.*, 2014) had identified as adapted to particularly acidic or alkaline soils. In such cases, i.e. when an 'extreme' soil is included in a model along with a phytometer that has a particular adaptation to this extreme soil, the SEM hypothesized by Daou and Shipley (2019) was rejected. This occurred because the same soil produced different amounts of increase in RGR between species depending on the soil pH. This added dimension of soil properties made it impossible to use a single latent variable to explain the pattern of covariances between the RGRs and explains why the null probability of the hypothesized SEM of Daou and Shipley fluctuated depending on the species composition.

The alternative model is composed of two latent variables. This alternative model fitted the data well with all combinations of up to eight species but was rejected when *U. dioica* was present. We can suggest two possible reasons for this result. First, it is possible that strongly nitrophilic species such as *U. dioica* might be unable to grow in the most 'infertile' soils since they require a higher minimum soil nitrogen supply. Other species would increase their RGR values when starting at the least fertile soils whereas *U. dioica* would not begin increasing its RGR until a sufficiently fertile soil was encountered to begin any growth. This would necessarily generate a non-linear relationship of over the entire range of fertility for *U. dioica*. The notion of inclusive fundamental niches, and of generalized soil fertility, does not require linearity, but linearity is a statistical limitation for the SEM measurement model. However, a more prosaic possibility is that the RGR values of *U. dioica* were underestimated in some soils due to the choice of the first harvest date. This possibility is supported by two observations. First, even the maximum RGR of *U. dioica* (Appendix 2: Supplementary Data Fig. S4) was lower than six of the phytometers even though this species is known to have a very high RGR_{max} (Grime *et al.*, 2014). Second, the RGR of *U. dioica* between the two harvest dates in some soils was virtually zero even though the plants had clearly grown before the first harvest. On the other hand, it is undeniable that *U. dioica* had superior growth in several soils, particularly the 'fertile' soils where several species achieved maximum growth, despite the underestimated measurements. It is therefore difficult to draw firm conclusions regarding the rejection of the alternative two-latent SEM in the presence of *U. dioica*. Further experimentation will be needed to test the hypothesis that a highly nitrophilic species would respond differently to the 'fertility' gradient. However, *C. album* is also a nitrophilic species (Ellenberg N-number 7) that exhibits similar traits to *U. dioica* but had a linear response to soil 'fertility' as with the other species. Moreover, Fichtner and Schulze (1992) found that species having different N-numbers (1 to 9) responded quantitatively differently but in the same directions to most levels of NH_4NO_3 additions in hydroponic sand culture. As pointed out earlier, such hydroponic nitrogen concentrations cannot be quantitatively related to available nitrogen flux rates in natural soils.

Species with larger seeds can subsidize growth in infertile soils for longer periods of time during initial seedling growth. Although we explicitly chose the initial harvest dates of each species in order to take into account this possibility and to exclude the initial period in which growth is subsidized by seed

reserves, we cannot exclude this possibility. However, if this effect still existed in our data then this would introduce an extra dimension related to seed size and would result in the rejection of an SEM that does not include this latent variable. Since this did not occur, we conclude that any remaining seed size effects, if present, are sufficiently minor as to be undetectable given our level of statistical power.

The hypothesis that there exists a single latent property of soils ('generalized' fertility, F_G) causing the correlated growth responses of plant species was rejected. We conclude that there are at least two underlying dimensions to soil fertility. However, such a statistical result does not mean that these added dimensions are equally important in terms of determining soil fertility. We found that the original model of Daou and Shipley (2019), i.e. when ignoring the second latent variable linked to soil pH, resulted in only very minor decreases in predictive ability. The correlation between the predicted values of F_G using the original model of Daou and Shipley (2019) and the predicted values of the first latent in the alternative model (i.e. after statistically holding constant variation in soil pH) was 0.98 and with a very small residual sum of squares.

Such results suggest that the notion of a unidimensional 'generalized' fertility is still valid for most soils and for most plant species although some conditions could limit this potential generalization. First, the model of Daou and Shipley (2019) will probably be less precise at the extremes of soil pH and when using phytometers that are adapted to such extreme conditions or possibly at the extremes of soil nitrogen flux when using strongly nitrophilous species. We can easily imagine hypothetical scenarios in which such errors would be more pronounced but we did not encounter such situations in our data. In fact, the predicted values of F_G from the model of Daou and Shipley (2019) were highly robust and largely insensitive to the species assemblage. None the less, we note that only three of our phytometers responded to pH and the majority of our soils (13/25) had close to neutral pH ranges (6.25–7.75) and so combinations of soils and species are certainly possible in which a one-dimensional model will provide poor predictions of soil fertility.

This conclusion is relevant to one of Grime's foundational assumptions of his CSR theory (Grime, 1977). Grime claimed that the growth of all plant species could be arranged along a gradient of stress, which he defined as any abiotic cause that decreases net primary productivity. When all abiotic causes of net primary productivity except those inherent in soils are kept constant, 'stress' becomes the inverse of generalized soil fertility. Based on our results, Grime's concept of stress, as applied to soil fertility, is not strictly valid but is a good approximation to most species and most soils.

We envisage two possible extensions of our results. First, we need to know how many dimensions (i.e. underlying latent variables) are required to completely characterize the growth responses of plants to soils and to determine the relative quantitative importance of these dimensions in different environments. Besides soil pH, adaptations to salinity, heavy metals or low oxygen concentration could express other dimensions in the soil 'fertility' concept. If such additional dimensions prove to be quantitatively important then they would complicate the quantification of soil fertility. If these additional dimensions

are, like soil pH, quantitatively minor relative to F_G , then the simpler unidimensional model of Daou and Shipley (2019) will still be a useful approximation in most situations. If this were the case, then it would be preferable to incorporate only phytometers that are not specifically adapted to these additional dimensions when quantifying F_G but to explicitly include such specifically adapted species when quantifying these additional dimensions. In this way, it might still be possible to produce a general metric of soil fertility that is applicable to plant species globally.

SUPPLEMENTARY DATA

Supplementary data are available online at <https://academic.oup.com/aob> and consist of the following. Table S1: Additional information on phytometer species. Table S2: Sampling sites. Figure S1: Temperature variation between experimental blocks for day, night and total periods. Figure S2: Relative humidity variation between experimental blocks for day, night and total periods. Figure S3: Vapour pressure deficit variation between experimental blocks for day, night and total periods. Figure S4: Measured variation in RGR across species under constant conditions in 25 soils from a 'fertility' gradient. Figure S5: Proportion of the SEM latent variable measurement models accepted as a function of the number of phytometer species for 382 possible models. Figure S6: Biplot of a principal component analysis of the RGR of nine species on 25 soils from a 'fertility' gradient. Figure S7: Proportion of SEM latent variable measurement models accepted as a function of the number of phytometer species for 256 possible models.

ACKNOWLEDGEMENTS

We thank Nicolas Grenier for help in the experimental setup.

FUNDING

The Natural Sciences and Engineering Research Council of Canada (NSERC) funded this research.

LITERATURE CITED

- Austin MP. 1980. Searching for a model for use in vegetation analysis. *Vegetatio* **42**: 11–21.
- Austin MP, Austin BO. 1980. Behaviour of experimental plant communities along a nutrient gradient. *Journal of Ecology* **68**: 891–918.
- Austin MP, Smith TM. 1990. A new model for the continuum concept. In: *Progress in theoretical vegetation science*. Berlin: Springer, 35–47.
- Bates D, Mächler M, Bolker B, Walker S. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* **67**: 1–48.
- Ben-Dor E, Banin A. 1989. Determination of organic matter content in arid-zone soils using a simple 'loss-on-ignition' method. *Communications in Soil Science and Plant Analysis* **20**: 1675–1695.
- Bollen KA. 2014. *Structural equations with latent variables*. Chichester: John Wiley & Sons.
- Cann DB, Gérin-Lajoie P. 1942. *Soil survey of Stanstead, Richmond, Sherbrooke, and Compton counties in the province of Quebec*. Quebec: Dominion Department of Agriculture.
- Clements FE, Goldsmith GW. 1924. *The phytometer method in ecology: the plant and community as instruments*. Washington: Carnegie Institution of Washington.

- Colwell RK, Fuentes ER. 1975. Experimental studies of the niche. *Annual Review of Ecology and Systematics* 6:281–310.
- Craine JM. 2009. *Resource strategies of wild plants*. Princeton: Princeton University Press.
- Czarapata EJ. 2005. *Invasive plants of the upper Midwest: an illustrated guide to their identification and control*. Madison: University of Wisconsin Press.
- Daou L, Shipley B. 2019. The measurement and quantification of generalized gradients of soil fertility relevant to plant community ecology. *Ecology* 100: e02549.
- Ellenberg H. 1974. Indicator values of vascular plants in central Europe. In: *Indicator values of vascular plants in central Europe*, 9.
- Fichtner K, Schulze E-D. 1992. The effect of nitrogen nutrition on growth and biomass partitioning of annual plants originating from habitats of different nitrogen availability. *Oecologia* 92: 236–241.
- Fraser LH. 2020. TRY—A plant trait database of databases. *Global change biology* 26: 189–190.
- Garnier E, Navas M-L, Grigulis K. 2016. *Plant functional diversity: organism traits, community structure, and ecosystem properties*. Oxford: Oxford University Press.
- Grace JB. 2006. *Structural equation modeling and natural systems*. Cambridge: Cambridge University Press.
- Grime JP. 1977. Evidence for the existence of three primary strategies in plants and its relevance to ecological and evolutionary theory. *The American Naturalist* 111: 1169–1194.
- Grime JP, Hodgson JG, Hunt R. 2014. *Comparative plant ecology: a functional approach to common British species*. Berlin: Springer.
- Harpole WS, Ngai JT, Cleland EE, et al. 2011. Nutrient co-limitation of primary producer communities. *Ecology Letters* 14:852–862.
- Hendershot WH, Lalonde H, Duquette M. 2007. *Soil reaction and exchangeable acidity*, 2nd edn. In Carter MR, Gregorich EG, eds. Boca Raton: CRC Press, Taylor & Francis Group, 173–174.
- Hutchinson GE. 1957. *Cold Spring Harbor symposium on quantitative biology. Concluding remarks* 22: 415–427.
- Jones HG. 1992. *Plants and microclimate: a quantitative approach to environmental plant physiology*. Cambridge: Cambridge Univ. Press.
- Keddy PA. 2001. *Competition*, 2nd edn. Dordrecht: Kluwer Academic Publishers.
- Kroetsch D, Wang C. 2007. Particle size distribution. In Carter MR, Gregorich EG, eds. *Soil sampling and methods of analysis*, 2nd edn. Boca Raton: CRC Press and Taylor and Francis Group, 713–725.
- Laliberté E. 2017. Below-ground frontiers in trait-based plant ecology. *New Phytologist* 213: 1597–1603.
- Lamontagne L, Martin A, Grenon L, Cossette J-M. 2002. *Étude pédologique du comté de Saint-Jean (Québec)*. Agriculture et Agroalimentaire Canada, Centre de recherche et de développement sur les sols et les grandes cultures.
- Pérez-Harguindeguy N, Díaz S, Gamier E, et al. 2013. New handbook for standardised measurement of plant functional traits worldwide. *Australian Journal of Botany* 61: 167–234.
- Pierce S, Negreiros D, Cerabolini BE, et al. 2017. A global method for calculating plant CSR ecological strategies applied across biomes world-wide. *Functional Ecology* 31: 444–457.
- Raaijmakers D, Lambers H. 1996. Response to phosphorus supply of tropical tree seedlings: a comparison between a pioneer species *Tapirira obtusa* and a climax species *Lecythis corrugata*. *New Phytologist* 132: 97–102.
- Reich PB, Buschena C, Tjoelker MG, et al. 2003. Variation in growth rate and ecophysiology among 34 grassland and savanna species under contrasting N supply: a test of functional group differences. *New Phytologist* 157: 617–631.
- Rosseel Y. 2012. Lavaan: an R package for structural equation modeling and more. Version 0.5–12 (BETA). *Journal of Statistical Software* 48: 1–36.
- Shipley B. 2016. *Cause and correlation in biology: a user's guide to path analysis, structural equations and causal inference with R*. Cambridge: Cambridge University Press.
- Shipley B, Bello FD, Cornelissen JHC, Laliberté E, Laughlin DC, Reich PB. 2016. Reinforcing loose foundation stones in trait-based plant ecology. *Oecologia* 180: 923–931.
- Tilman D. 1988. *Plant strategies and the dynamics and structure of plant communities*. Princeton: Princeton University Press.
- Tyler G. 1996. Soil chemistry and plant distributions in rock habitats of southern Sweden. *Nordic Journal of Botany* 16: 609–635.
- Warnes GR, Bolker B, Lumley T, Warnes MGR. 2015. *Package 'gtools'. R Package version 3*. <https://github.com/r-gregmisc/gtools>