Reaction norm fitness peak tracking applied on sparse time series in the fossil record

Rolf Ergon

University of South-Eastern Norway

13.11.2024

Abstract

Species that have existed over millions of years have done so because of sufficient mean fitness for survival through the ages, such that the mean phenotypic traits continuously have been tracking the optimal phenotypic values that maximize fitness. These optimal values are predominantly functions of the environment, for example the sea temperature. The mean phenotypic traits over time will thus predominantly be determined by the environment over time, and the trait history may be found in the fossil record. Here, I simulate such a tracking system by use of a univariate intercept-slope reaction norm model and show how both linear and nonlinear mean phenotype vs. environment functions can be found also from quite sparse and short time series from the fossil record, and I also discuss how this methodology can be extended to multivariate systems. The simulations also include cases with a constraint on the individual trait values and with other factors than environment influencing the optimal phenotypic values. The methodology is finally applied on two time-varying mean traits in a record of bryozoan *Microporella agonistes* fossils spanning 2.3 million years, using the $\partial^{18}O$ measure as environmental variable. From few irregular samples of phenotypic mean traits

found in the fossil record, it was possible to identify linear mean phenotype vs. environment functions, and thus to predict the continuous mean phenotypic traits as functions of time.

Keywords: bryozoans, fitness peak tracking, fossil record, optimal phenotypes, phenotypic plasticity, reaction norm.

1 Introduction

Evolutionary time series in the fossil record provide information on phenotypic changes over millions of years. Such time series may consist of quite few irregular samples, and they have traditionally been modeled by various forms of random walk processes (Hunt, 2006; Voje, 2023), or by Ornstein-Uhlenbeck (OU) processes describing evolution of phenotypic traits toward fixed or moving peaks in the adaptive landscape (Voje, 2023). In an example, Voje (2023) shows how an OU model describes the evolution in the radiolarian lineage *Eucyrtidium calvertense* from one optimal size to another (Kellogg,1975). Neither random walk processes nor OU processes have explicit means for handling of phenotypic plasticity, but one cannot exclude the possibility that plasticity plays an implicit part in such models. Such plasticity possibilities are discussed in Sætre et al. (2017), where an OU model was used in the analysis of changed body size measurements of individuals of the bird species reed warbler (*Acrocephalus scirpaceus)* over a 19-year period.

Reaction norm analysis of fossil data under the explicit assumption of phenotypic plasticity would in many cases appear to be more appropriate, especially for species that have adapted to environmental changes over millions of years with sufficient fitness for survival. In such cases the mean phenotypic traits have continuously been tracking the optimal values that maximize fitness as determined by moving fitness peaks in the adaptive landscape (Bell, 2013).

Phenotypic plasticity is ubiquitous in nature and generally regarded as a key mechanism for enabling organisms to survive in the face of environmental change (Williams et al., 2008).

Phenotypic changes in natural populations caused by climate change have been widely studied, and a fundamental difficulty is then to disentangle the effects of microevolution and phenotypic plasticity. Merilä & Hendry (2013) discussed this problem and summarized the results of 11 review articles. Among the review papers were Charmantier & Gienapp (2014) (birds), Boutin & Lane (2014) (mammals), and Crozier & Hutchings (2014) (fish), which summarize the findings in altogether 66 studies of responses on climate change. Merilä & Hendry (2013) wrote that "Summarizing and relying on the results of these reviews, we arrive at the conclusion that evidence for genetic adaptation to climate change has been found in some systems, but is still relatively scarce." They further wrote that "Indeed, a number of phenotypic differences originally thought to be genetically based were subsequently attributed to phenotypic plasticity (e.g., James 1983; Charmantier et al. 2008; Teplitsky et al. 2008). In the wake of such demonstrations phenotypic plasticity is increasingly adopted as a parsimonious model that is to be rejected only if direct evidence is obtained for genetic change: That is, plasticity is treated as a null model. The reality, however, is that inferences regarding phenotypic plasticity also should be supported by a specific set of evidentiary criteria". The disentanglement problem was also discussed in Lister (2021), who concluded that "The role of phenotypic plasticity is particularly challenging to identify in the fossil record as the environmental and genetic influences on morphological variation and change are difficult to separate".

In a more recent review regarding temporal variation in selection on breeding date in 39 populations of 21 wild animals, de Villemereuil et al. (2020) found that "plastic changes track movements of the optimum phenotype across years, especially in bird species, reducing temporal variation in directional selection. This suggests that phenological plasticity has evolved to cope with fluctuations in the optimum, despite their currently modest contribution to variation in selection".

Reaction norm models are in general multivariate, allowing for correlated traits (Ergon, 2023). Here, I will use a univariate reaction norm model in a study of fitness peak tracking systems, where mean phenotypic traits by means of microevolution and plasticity are tracking the optimal phenotypic values that maximize fitness. In such systems the location of fitness peaks over time will predominantly be determined by the changing environment, i.e., by fitness peak vs. environment functions for the optimal phenotypic values. Since the mean phenotypic traits are tracking the optimal values, these fitness peak vs. environment functions are approximately equal to mean phenotype vs. environment functions that can be found from data. This is so also if genetic factors, developmental constraints etc. affects the location of the fitness peak, although then the understanding of such functions as only affected by the environment is not quite correct. Provided that the tracking system keeps the mean fitness on a sufficiently high level for survival there is thus no need to know the details of the evolutionary system, neither reaction norm parameter values nor fitness functions. This also means that there is no need for disentanglement of the microevolutionary and plasticity contributions to the mean traits, as discussed by Merilä & Hendry (2013) and Lister (2021). What is needed is the mean phenotype vs. environment functions that describe how given mean phenotypic traits varies with the environment and possibly other factors, and parameters in such functions may be estimated also from very few fossil data points at irregular sampling times.

For simplicity I will in both a theoretical analysis and in simulations make use of a population with an individual intercept-slope reaction norm model with parameters $a(t)$ and $b(t)$ as given by

$$y(t) \;=\; a(t) + b(t)\big(u(t) - u_{ref}\big) + \varepsilon(t), \tag{1}$$

where $u(t)$ is the environmental cue, and where $\varepsilon(t)$ is assumed to be normally distributed with constant mean 0 and variance σ_ε^2 among individuals in every generation. Here, u_{ref} is a reference environment defined as the environment at which the phenotypic variance is at a

minimum (Lande, 2009). I will use simulations based on the model in Eq. (1) to see how the mean phenotypic trait $\bar{y}(t)$ in a population evolves in a changing environment under the fundamental assumption that the mean fitness is sufficiently high for survival. In the simulations I will use the $\partial^{18}O(t)$ measure as environmental cue, both because it is a widely used proxy for seawater temperature in paleontological studies (Lisiecki and Raymo, 2005; Liow et al., 2024), and because it gives a connection between the simulations and a real data case based on the fossil record. Since both $\partial^{18}O(t)$ and $\bar{y}(t)$ have large variances, and since data from the fossil record may give mean values over both populations and time (Liow et al., 2024), I will in the simulations also use moving average filtered versions of these variables. The simulations are intended to show that mean phenotype vs. environment functions can be estimated from sparse data also in cases where the population in shorter time periods have so low mean fitness that they are on the verge of extinction. Cases with a constraint on individual trait values and with other drivers than temperature involved are included.

In addition to simulations, the theory is extended to cover multivariate cases. It is also applied on recently presented data from a field study of the bryozoan species *Microporella agonistes*, based on a fossil record over 2.3 million years and with two time-varying mean phenotypic traits (Liow et al., 2024). Parameters in linear mean phenotype vs. environment functions are here estimated by as few as 8 data points at irregular sampling times.

Basic theory is given in Section 2, followed by simulations in Section 3 and a real data case in Section 4. Finally follows a discussion in Section 5.

2 Theory

2.1 A basic univariate case

The individual intercept-slope reaction norm model with parameters $a(t)$ and $b(t)$ in Eq. (1) will for a given population result in the mean reaction norm

$$\bar{y}(t) = \bar{a}(t) + \bar{b}(t)\big(u(t) - u_{ref}\big). \tag{2}$$

Provided that $\bar{a}(t)$ and $\bar{b}(t)$ evolve such that the mean fitness is kept at a sufficiently high level for survival, $\bar{y}(t)$ will be kept in the vicinity of the optimal phenotypic values that correspond to a moving fitness peak in the adaptive landscape. For simplicity and for use in simulations I will initially assume that mean fitness is determined by the single phenotypic mean trait $\bar{y}(t)$ according to Eq. (2), and that the adaptive landscape thus is two-dimensional. I will later extend some essential results to multivariate cases, and in the simulations I will also include examples where other factors than the physical environment are involved.

The optimal phenotypic value is in general a function of the environment,

$$\theta(u,t) = \alpha + f\big(u(t) - u_{ref}\big), \tag{3}$$

in the linear case $\theta(u,t) = \alpha + \beta\big(u(t) - u_{ref}\big)$, where α and β are constant parameters. Here, the $\theta(u,t)$ will as the notation shows vary with the environment $u(t)$, for example sea temperature. There will also be genetic factors, developmental constraints, external biotic interactions etc. that affects the location of the fitness peak in the adaptive landscape, such that the understanding of $\theta(t)$ as only affected by the environment is not quite correct.

In the simulations I will assume an individual fitness function

$$W(t) = W_{max} \exp\left(-\frac{\big(y(t) - \theta(u,t)\big)^2}{2\omega^2}\right), \tag{4a}$$

which gives the mean fitness in a population (Lande, 2009)

$$\bar{W}(t) = W_{max}\sqrt{\frac{\omega^2}{\omega^2 + \sigma_y^2(t)}}\exp\left(-\frac{\big(\bar{y}(t) - \theta(u,t)\big)^2}{2\big(\omega^2 + \sigma_y^2(t)\big)}\right), \tag{4b}$$

where $\sigma_y^2(t)$ is the variance of $y(t)$ according to Eq. (1). The mean fitness will thus be maximized when $\bar{y}(t) = \theta(u,t)$.

Under the assumptions of nonoverlapping generations and random mating in a large population the evolution of the mean reaction norm parameters is governed by (Lande, 1979)

$$\begin{bmatrix}\bar{a}(t+1)\\ \bar{b}(t+1)\end{bmatrix} = \begin{bmatrix}\bar{a}(t)\\ \bar{b}(t)\end{bmatrix} + \begin{bmatrix}G_{aa} & G_{ab}\\ G_{ab} & G_{bb}\end{bmatrix}\begin{bmatrix}G_{aa}+\sigma_\varepsilon^2 & G_{ab}\\ G_{ab} & G_{bb}\end{bmatrix}^{-1}\begin{bmatrix}cov(w(t),a(t)+\varepsilon(t))\\ cov(w(t),b(t))\end{bmatrix}, \quad (5)$$

where G_{aa}, G_{bb} and G_{ab} are variances and the covariance of the reaction norm parameters, and where $w(t) = W(t)/\overline{W}(t)$ is the relative fitness. In a stationary stochastic environment, the optimal reaction norm slope is $\bar{b}_{opt} = cov(\theta(u,t),u(t))/var(u(t))$ (Lande, 2009).

When the mean phenotypic trait in a population is tracking a moving value $\theta(u,t)$ according to Eq. (3) it will be given by $\bar{y}(t) = \theta(u,t) - e(t)$, where $e(t)$ is the tracking error, see Fig. 1 for visualization. This means that the parameter values $\bar{a}(t)$ and $\bar{b}(t)$ are forced to evolve such that $\bar{y}(t) \approx \theta(u,t) = \alpha + f\big(u(t) - u_{ref}\big)$, i.e., such that the mean fitness according to Eq. (4b) is kept close to a near maximum level and that the tracking error is small. In feedback control terminology it is thus the mean trait tracking error $e(t) = \theta(u,t) - \bar{y}(t)$ that via individual and mean fitness is the driving force in the adaptive process. As shown in simulations as well as in the real data case the parameters in the $\bar{y}(t) \approx \alpha + f\big(u(u,t) - u_{ref}\big)$ model can be estimated also from rather few fossil data points $\big[u_{filt}(t), \bar{y}_{filt}(t)\big]$ at irregular sampling times, where $\bar{y}_{filt}(t)$ and $u_{filt}(t)$ are moving average filtered versions of often quite noisy signals. This is so also for cases with constraints on the individual trait values. Note that even though $\bar{y}(t) \approx \alpha + f\big(u(t) - u_{ref}\big)$, such a model is not a reaction norm, i.e., the reaction to a sudden change Δu in $u(t)$ will still be a sudden change $\Delta\bar{y} = \bar{b}(t)\Delta u$.

The mean phenotype tracking system is described in Fig. 1. The details in the tracking process are described by Eqs. (2), (4) and (5), but also if the parameters in those equations were known they could not be used for computation of $\bar{y}(t)$ based on only a few irregular samples. The remaining possibility for prediction of $\bar{y}(t)$ is to assume a good enough tracking such that $\bar{y}(t) \approx \theta(u,t) = \alpha + f\big(u(t) - u_{ref}\big)$, which as shown in the simulations as well as in the real data case makes it possible to estimate the parameters in the $\theta(u,t)$ function from few $\big[u_{filt}(t), \bar{y}_{filt}(t)\big]$ data points.

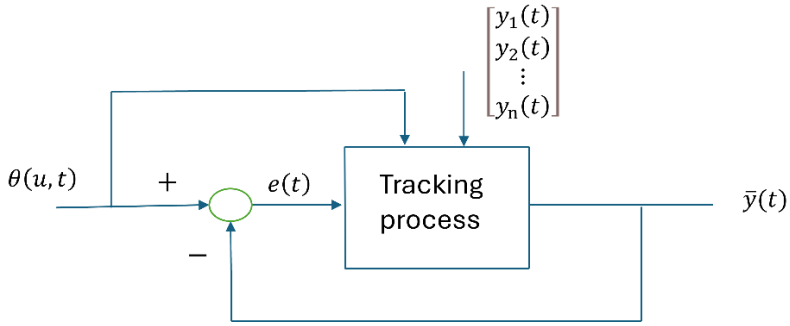

Figure 1. Block diagram for reaction norm tracking system, with the phenotypic value $\theta(u,t)$ that maximizes fitness and the individual phenotypic traits $y_1(t), y_2(t), \dots, y_n(t)$ as inputs and the mean phenotypic trait $\bar{y}(t)$ as output.

2.2 Influence of other factors than u(t)

Eqs. (2) and (3) assume that u(t) is the only factor that affects the mean phenotypic trait $\bar{y}(t)$ and the optimal phenotypic value $\theta(u,t)$, respectively. When other factors affect $\theta(u,t)$ the tracking system in Fig. 1 will force the parameters $\bar{a}(t)$ and $\bar{b}(t)$ in Eq. (2) to evolve such that these factors are compensated for, and the result will be seen in the mean phenotype vs. environment function that can be found from data. When other factors affect $\bar{y}(t)$ it is again the $\bar{a}(t)$ and $\bar{b}(t)$ parameter values that must compensate for that, as shown in a simulation example in Section 3.

2.3 Multivariate cases

When the mean phenotypic trait in a univariate population is tracking the optimal phenotypic value, the mean phenotypic trait can as shown above be described by $\bar{y}(t) \approx \theta(u,t) = \alpha + f\big(u(t) - u_{ref}\big)$, i.e., by the fitness peak vs. environment function. Here, it is again important to note that we do not need to know the details of the underlying reaction norm model and fitness function.

The univariate theory can be extended to multivariate cases, and although we in practice do not need to know the underlying fitness function, we may for a theoretical discussion use an

example with two phenotypic traits and the bivariate fitness function (Ch. 4, Johnson & Wichern, 2008)

$$W(t) = W_{max} \exp\left(-[y_1(t) - \theta_1(u,t) \quad y_2(t) - \theta_2(u,t)][\Sigma(\mathrm{t})]^{-1} \begin{bmatrix} y_1(t) - \theta_1(u,t) \\ y_2(t) - \theta_2(u,t) \end{bmatrix} / 2c\right), \tag{6}$$

where $\Sigma(\mathrm{t}) = \mathrm{cov}(y_1(t), y_2(t))$ is the variance-covariance matrix and where c is a scaling factor controlling the 'width' of the fitness function. In order to keep the mean fitness at a level necessary for survival, the mean traits must in this case continuously evolve such that $\bar{y}_1(t) \approx \theta_1(u,t)$ and $\bar{y}_2(t) \approx \theta_2(u,t)$, and we will thus find

$$\bar{y}_1(t) \approx \alpha_1 + f_1\big(u(t) - u_{ref}\big) \tag{7a}$$

$$\bar{y}_2(t) \approx \alpha_2 + f_2\big(u(t) - u_{ref}\big), \tag{7b}$$

where the moving fitness peak vs. environment functions may be linear or nonlinear, and where as in the univariate case the parameters can be estimated from rather few fossil data points at irregular sampling times. Here, the essential point is that $\bar{y}_1(t)$ and $\bar{y}_2(t)$ in this example are tracking a peak in a three-dimensional adaptive landscape, and that this tracking process can be decomposed into two univariate tracking processes. We may thus consider the projections of the bivariate fitness function onto orthogonal axes for the two phenotypic traits, and such a projection approach may be extended to systems with higher dimensions, as discussed in Goodnight (2013).

3 Simulations

3.1 Linear reaction norm example

A system with a linear mean reaction norm $\bar{y}(t) = \bar{a}(t) + \bar{b}(t)\big(u(t) - u_{ref}\big)$ was simulated with use of the 2,115 available samples of $u(t) = \partial^{18}O(t)$, and with population size 1,000. These samples go 5.3 million years back in time with various sampling intervals (Lisiecki and Raymo, 2005), but here I just assume 2,115 nonoverlapping generations (however, see Section 5 for a

discussion on relevance for fossil data). The reaction norm parameter variances and covariance were chosen as $G_{aa} = \sigma_{\varepsilon}^2 = 0.5$, $G_{bb} = 0.05$ and $G_{ab} = 0$, and at each new generation random individual samples around $\bar{a}(t)$, $\bar{\varepsilon}(t) = 0$ and $\bar{b}(t)$ were drawn from normal populations with the given variances. The fitness function in Eq. (4a) had parameter values $W_{max} = 1$ and $\omega^2 = 50$. The optimal phenotypic value was in a first linear case chosen as $\theta(u,t) = 0.5 + u(t) - 2.9$, where $\theta(u,t)$ thus and for clarity of presentation is assumed to be a deterministic function, and where 2.9 is the mean value of $\partial^{18}O(t)$ around $t = 0$. From this follows that the optimal plasticity slope in a stationary stochastic environment is $\bar{b}_{opt} = cov(\theta(t), u(t))/var(u(t)) = 1$. The initial intercept value was $\bar{a}(0) = 0$, and simulations were performed with three different initial values of $\bar{b}(0)$. Since $\bar{y}(t)$, $\theta(u,t)$ and $u(t) = \partial^{18}O(t)$ have large variances, I used moving average filtered signals $\bar{y}_{filt}(t)$, $\theta_{filt}(u,t)$ and $u_{filt}(t)$ in the plots. These were obtained by moving window filters with window size 100, in MATLAB notation $\bar{y}_{filt}(t) = movmean(\bar{y}(t), 100)$ etc.

Simulation results in the linear case are shown in Fig. 2. Note that $\bar{b}(0) = \bar{b}_{opt} = 1$ results in a near constant plasticity slope, and that also $\bar{a}(t)$ settles on a near constant value (panel E). With an initial mean plasticity slope $\bar{b}(0) \neq \bar{b}_{opt}$, the mean intercept evolves all the time, while the mean plasticity slope very slowly evolves towards $\bar{b}_{opt} = 1$ (panels A and I). Note that $\bar{y}_{filt}(t)$ is tracking $\theta_{filt}(u,t)$ quite well (panels C, G and K). Also note that $\bar{y}_{filt}(t)$ as functions of $u_{filt}(t)$ are very similar straight lines as predicted by theory, with seven discrete values marked by stars (panels D, H and L). Using $u_{ref} = u_{filt}(0) = 2.9$, these discrete values could be used for determination of the parameters $\hat{\alpha}$ and $\hat{\beta}$ of this line by an ordinary least squares method, and thus to find predictions $\hat{\bar{y}}_{filt}(t) = \hat{\alpha} + \hat{\beta}\left(u_{filt}(t) - u_{ref}\right)$, see example in the Real data case section. With $\bar{b}(0) = 1$ it appears that $\bar{a}(t)$ and $\bar{b}(t)$ are constant over time, once $\bar{a}(t)$ has settled (panel E), but a closer look reveals slight fluctuations around fixed values. This is

because the individual values around $\bar{a}(t)$, $\bar{\varepsilon}(t) = 0$ and $\bar{b}(t)$ at each time step are drawn from random Gaussian distributions.

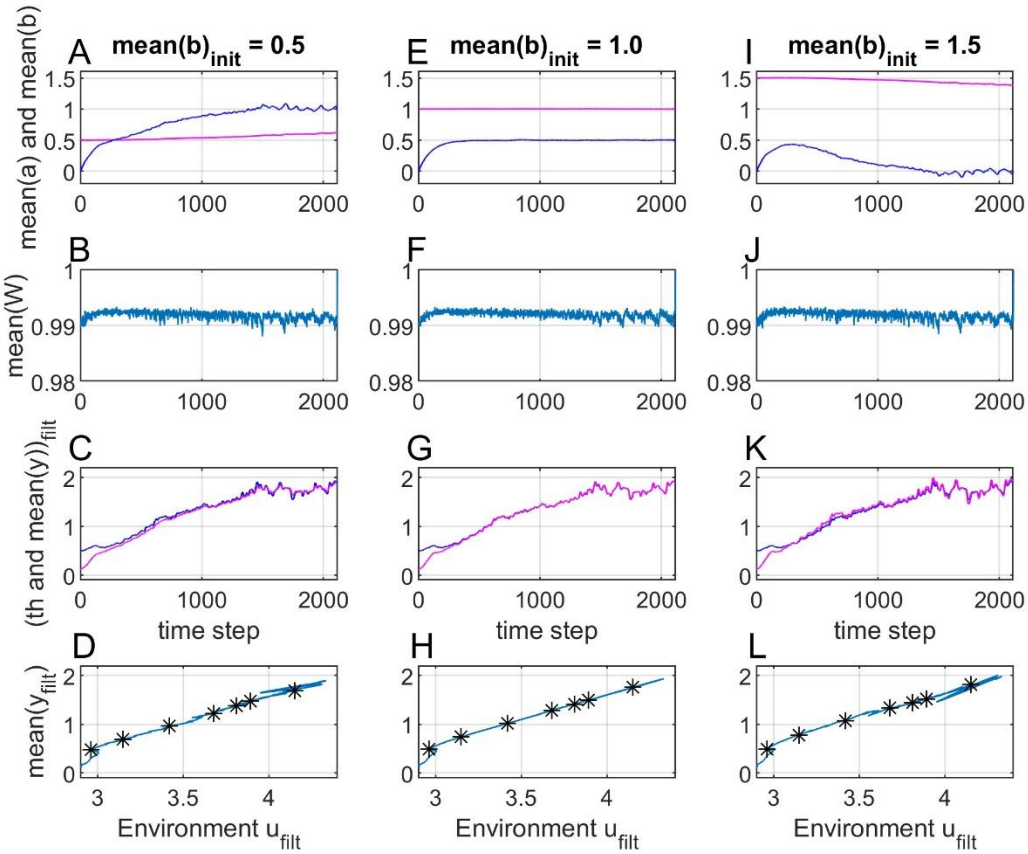

Figure 2. Simulation results with a linear $\theta(u,t) = 0.5 + u(t) - 2.9$ function and a linear reaction norm for three different initial values $\bar{b}(0)$, while $\bar{a}(0) = 0$. Panels A, E and I show $\bar{a}(t)$ (blue lines) and $\bar{b}(t)$ (red lines). Panels B, F and J show mean fitness $\bar{W}(t)$. Panels C, G and K show $\theta_{filt}(u,t)$ (blue lines) and $\bar{y}_{filt}(t)$ (red lines). Panels D, H and L show that $\bar{y}_{filt}(t)$ as function of $u_{filt}(t)$ in all cases is the same straight line, with values for samples 200, 400, 600, 800, 1000, 1200 and 1400 shown by stars.

3.2 Nonlinear reaction norm example

Fig. 3 shows results for a system with the nonlinear fitness peak vs. environment function $\theta(u,t) = 0.5 + (u(t) - 2.9) + 8(u(t) - 2.9)^2$, combined with the linear reaction norm

$\bar{y}(t) = \bar{a}(t) + \bar{b}(t)(u(t) - 2.9)$. Note that the overall mean fitness now is reduced to $mean(\bar{W}(t)) \approx 8.3$, and that the population for short time periods are on the verge of extinction because $\bar{W}(t) = 0$. The signal $\bar{y}_{filt}(t)$ is also in this case tracking $\theta_{filt}(u,t)$ well, although with a noticeable time lag (panels C, G and K). Here, $\bar{y}_{filt}(t)$ as function of $u_{filt}(t)$ is the same second order function in the different cases (panels D, H and L). Note that $\bar{b}(0) = 1$ in this case does not result in constant $\bar{a}(t)$ and $\bar{b}(t)$ values over time (panel E). See Discussion for results with a nonlinear reaction norm.

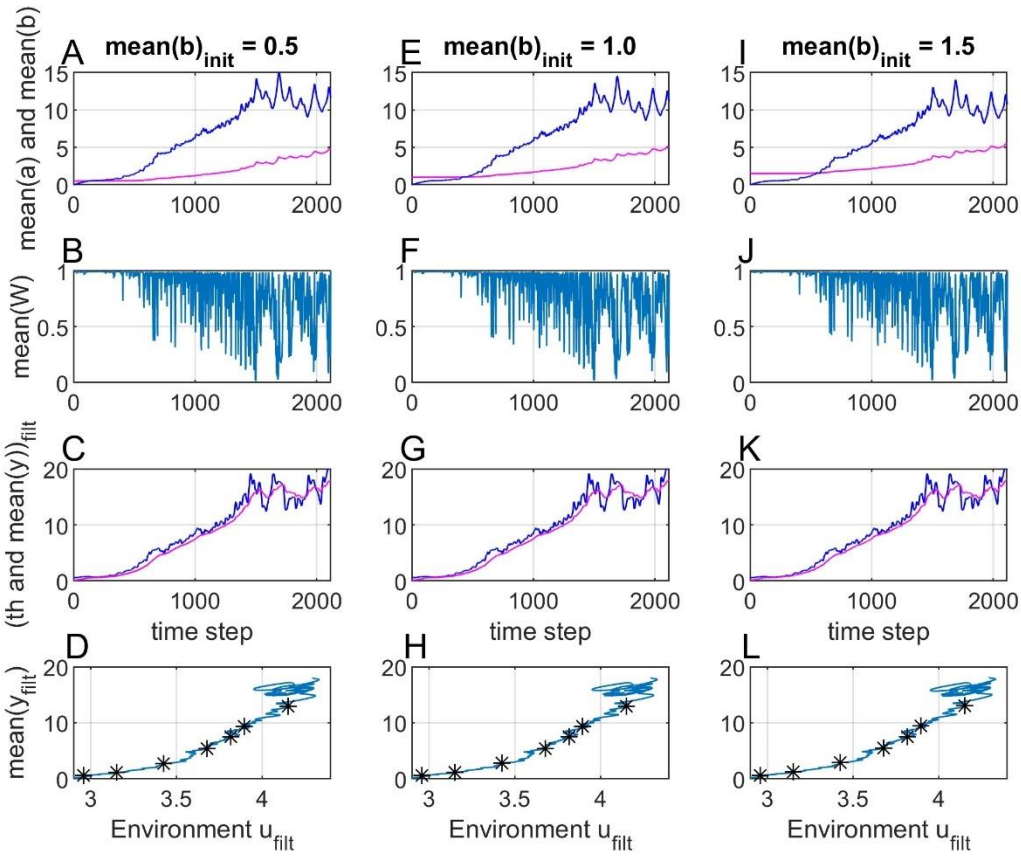

Figure 3. Simulation results as in Fig. 1, but with $\theta(u,t) = 0.5 + (u(t) - 2.9) + 8(u(t) - 2.9)^2$. The discrete values in panels D, H or L could here be used to determine the nonlinear function $\hat{\bar{y}}_{filt}(t) = f\left(u_{filt}(t)\right)$ by means of any suitable fitting method.

3.3 Example with phenotypic constraint

It follows from Eq. (4b) that the mean trait $\bar{y}(t)$ tracks the optimal value $\theta(u,t)$ that maximizes fitness, and this is so also in cases with constraints on the individual traits. When there for example is an upper limit for the individual phenotypic traits as shown in Fig. 4, the $\bar{a}(t)$ and $\bar{b}(t)$ parameters will still evolve such that $\bar{y}(t) \approx \theta(u,t)$, with results as shown in Fig. 5, panels A to D. The compensating effect is clearly seen by a comparison of panels A in Fig. 5 with panel E in Fig. 2, in that $\bar{a}(t)$ reaches a much higher value, such that $\bar{y}(t)$ without constraints on $y(t)$ would become larger than $\theta(u,t)$. Note that the mean fitness is increased compared to the results in Fig. 2 because the variance $\sigma_y^2(t)$ in Eq. (4b) is reduced. Also note that the mean phenotype vs. environment functions in panel D now is nonlinear in a way that with knowledge of the biological system possibly could be interpreted in a meaningful way. Note that individual trait values will have a larger variance than the mean trait value, such that the constraint in Fig. 4 has an effect before $\bar{y}(t) \approx \theta(u,t)$ reaches the upper limit.

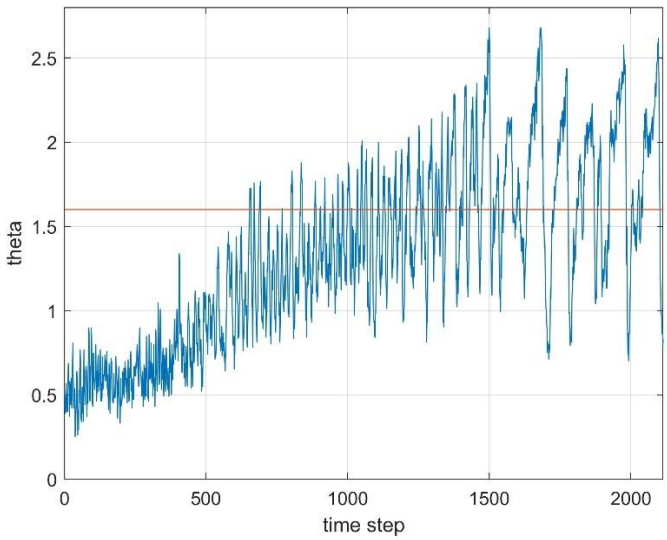

Figure 4. The linear function $\theta(u,t) = 0.5 + u(t) - 2.9$ as used for the results in Fig. 2 (blue line) and a constraint for the individual traits (red line).

3.5 Example with additional term in $\theta(u,t)$

The optimal value $\theta(u,t)$ that maximizes fitness may not be a function of only the environment, and Fig. 5, panels E to H, shows results as in Fig. 2, panels E to H, but with $\theta(u,t) = 0.5 + u(t) - 2.9 + (t/2115)^2$. Se Section 5 for a discussion of these results.

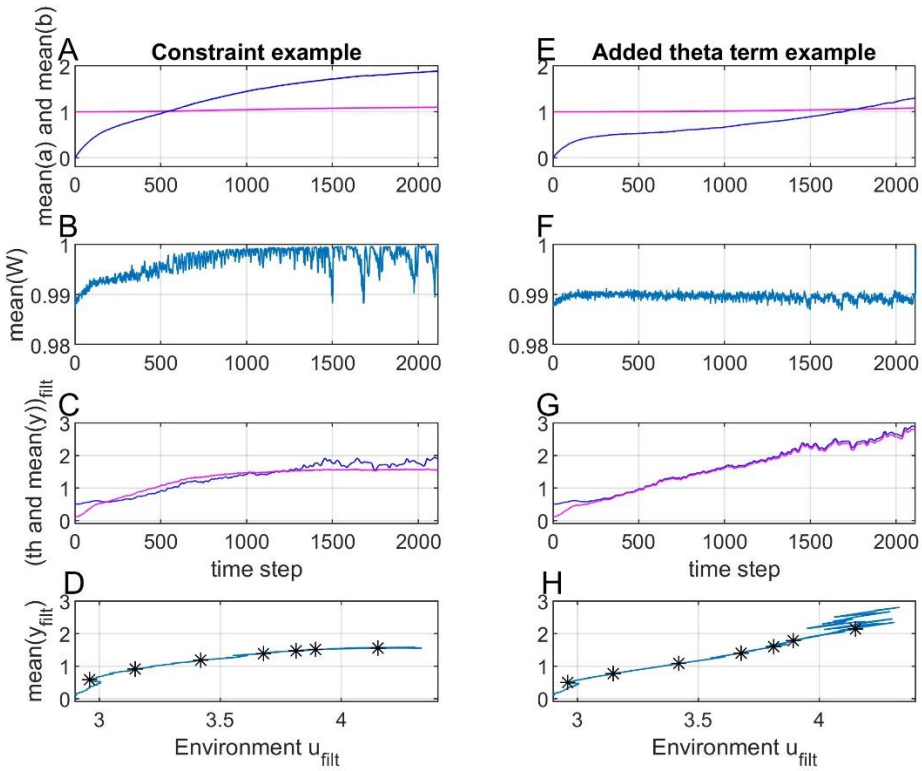

Figure 5. Panels A to D show simulation results as in Fig. 2 with $\bar{b}(0) = 1$, but with a constraint on the individual phenotypic trait values as shown in Fig. 4. Panels C and D show simulation results as in Fig. 2 with $\bar{b}(0) = 1$, but with $\theta(u,t) = 0.5 + u(t) - 2.9 + t/2115$.

4 Real data case

Liow et al. (2024) recently presented results from a field study of the bryozoan species *Microporella agonistes*, based on a fossil record over 2.3 million years. The fossils were

collected from colonies in New Zealand cliffs, and it is especially interesting that the same species still live in the waters near those cliffs. In a project running over several years, Liow and her coworkers collected 985 fossil colonies with a large total number of individuals and developed an artificial intelligence tool for data collection by use pictures from a scanning electron microscope. This image analysis system can rapidly assess three well defined phenotypic traits; autozooid (AZ) area, ovicell (OV) area and autozooid shape, and as seen from time series in Fig. S2 in Liow et al. (2024) the two AZ and OV area mean traits increase with time except for a drop in recent time owing to an extreme $\partial^{18}O(t)$ value. Liow et al. (2024) used average fecundity and competitive outcome as two different fitness components, where competitive outcome is a proxy for partial survival.

I will here use the log AZ and log OV area mean data in Fig. S2 in Liow et al. (2024) as examples. The last data points in these time series are based on an extreme $\partial^{18}O(t)$ value and are therefore omitted in the main analyses, see Fig. 6 in Liow et al. (2024). The reason is that these data points have a discrete $\partial^{18}O(t)$ value, while the analysis is based on average filtered $\partial^{18}O(t)$ values. Results with these points included will however also be given.

Fig. 4 shows $\partial^{18}O$ (panel A), $log\ AZ\ area\ mean$ (panel B) and $log\ OV\ area\ mean$ (panel C) as functions of time. Panel A also shows a moving average filtered version $u_{filt}(t)$ of $\partial^{18}O$, with moving average window size 100. Note that the values of $log\ AZ\ area\ mean$ and $log\ OV\ area\ mean$ are mean values of mean values in several colonies in time windows of various sizes (Liow et al., 2024).

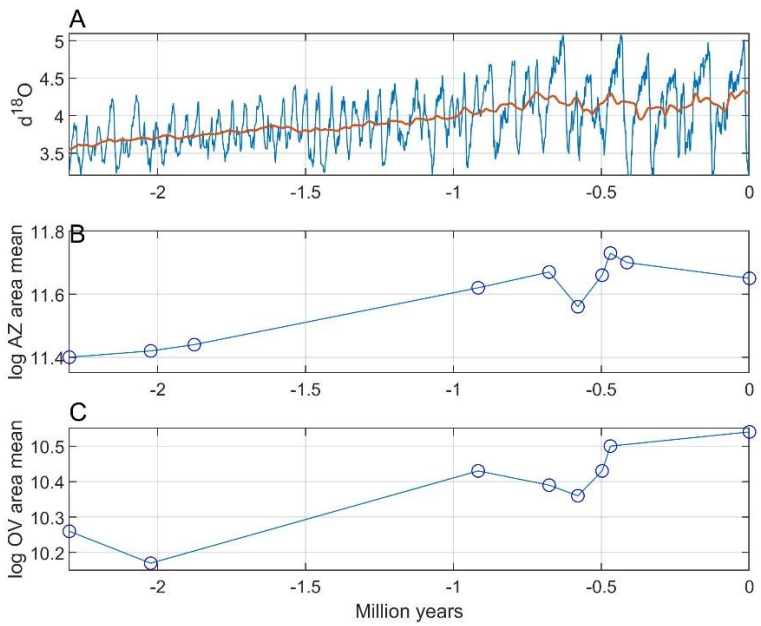

Figure 6. Raw $\partial^{18}O(t)$ data (panel A, blue line) and a moving average filtered version $u_{filt}(t)$ (panel A, red line), as well as *log AZ area mean* (panel B, circles) and *log OV area mean* (panel C, circles) as in Fig. S2 in Liow at al. (2024). Present time is used as zero point.

From Fig. 3 in Liow et al. (2024) follows that fitness measured as ovicell density (average fecundity) is fairly constant for all log OV area mean values, i.e., for the entire span of time. Fig. S8 shows that also competitive outcome is fairly constant for all of the three phenotypic traits. Since the bryozoan species under study have existed for at least 2.3 million years and still live it can be inferred that the mean fitness through the ages has been sufficient for tracking of the optimal phenotypic values that maximize fitness (compare with panels B, F and J, as well as panels C, G and K in Figs. 2 and 3, and note especially that the results in Fig. 3 are from a population that in shorter periods has zero mean fitness).

Fig. 7, panel A, shows *log AZ area mean* as function of $u_{filt}(i)$ for the 10 data points in Fig. 6, panel B. This plot should be compared with the corresponding plots in Figs. 1 and 2 (panels D, H and L), and considering the large standard errors indicated in Fig. S2 in Liow et al. (2024) it is rather obvious that it is not possible to fit more than a straight line

$log\ AZ\ area\ mean\ (i)\ = \hat{\alpha} + \hat{\beta}\big(u_{filt}(i) - u_{ref}\big)$. With $u_{ref}\ = u_{filt}(-2.3) = 3.6$ as found from Fig. 6, panel A, the ordinary least squares method results in $log\ AZ\ area\ mean\ (t) = 11.41 + 0.40\big(u_{filt}(t) - 3.6\big)$, as shown in Fig. 7, panel B. With unfiltered environmental data we thus find $log\ AZ\ area\ mean\ (t) = 11.41 + 0.40(\partial^{18}O(t)\ - 3.6)$ as shown in Fig. 7, panel C. With the omitted data point included, this changes into $log\ AZ\ area\ mean\ (t) = 11.47 + 0.29(\partial^{18}O(t)\ - 3.6)$.

When predictions for $log\ OV\ area\ mean$ are found in the same way as for $log\ AZ\ area\ mean$ the result with the last data point omitted is $log\ OV\ area\ mean\ (t) = 10.22 + 0.36\big(u_{filt}(t) - 3.6\big)$. As seen from Fig. 6, panel C, the number of data points is then reduced from 10 to 8 owing to a limited number of raw data points (Liow et al., 2024). A closer look reveals that the predictions of $log\ OV\ area\ mean\ (t)$ approximately are $0.9 \times log\ AZ\ area\ mean\ (t)$, and the two variables are thus linearly dependent. Details of such allometric relationships for *microporella* species are studied in Di Martino and Liow (2022).

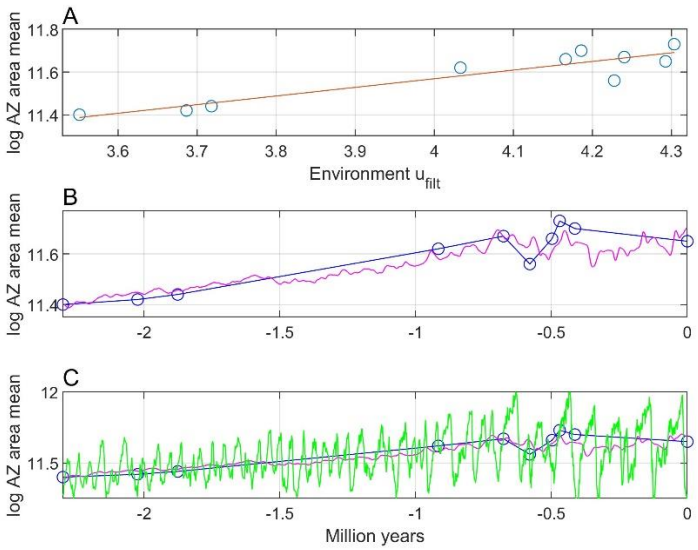

Figure 7. Panel A shows $log\ AZ\ area\ mean$ as function of u_{filt} for the ten data points in Fig. S2 in Liow at al. (2024) (circles), as well as a straight line fitted by the ordinary least squares method (red line). Panel B shows $log\ AZ\ area\ mean$ as function of time (blue

circles), and the fitted reaction norm response (red line). Panel C show predictions as in panel B (red line), and the prediction result with unfiltered $\partial^{18}O(t)$ data used as input (green line).

Fig. 7, panel C, illustrates the sampling difficulties involved, and thus underlines the need for filtering of the phenotypic information by finding mean values for time periods of up to 0.2 million years (Table S1, Liow et al., 2024). For the $\partial^{18}O(t)$ filtering, I here used a moving average window with a fixed width of 100 samples, but since the data is collected with various moving average window sizes a variable filtering window size could possibly improve the predictions.

Similar mean phenotype vs. environment results to those in Fig. 6 for several bryozoan species are discussed in Liow et al. (2017). Their results suggested that a common environmental factor drove the changes in zooid size, but that also overgrowth outcomes may play a role. The simulation results in Fig. 5, panels E to H, indicate that such additional terms to an otherwise linear fitness peak vs. environment function would result in nonlinear mean phenotype vs. functions, but whether a meaningful interpretation of such results found from the fossil record is possible is an open question.

5 Discussion

There are three main conclusions to be drawn from the results in this article. First, provided that the tracking system keeps the mean fitness at a sufficiently high level for survival, the mean phenotypic traits in a possibly multivariate system, $\bar{y}_1(t)$, $\bar{y}_2(t)$, ... $\bar{y}_m(t)$, will stay in the vicinity of a moving peak in the adaptive landscape defined by the optimal phenotypic values $\theta_1(u,t)$, $\theta_2(u,t)$, ... $\theta_m(u,t)$ that maximize fitness. Second, mean phenotypic vs. environment functions for $\bar{y}_1(t)$, $\bar{y}_2(t)$, ... $\bar{y}_m(t)$ can be found from sparse and short fossil data. Third, again provided that the mean fitness is kept at a sufficiently high level, predictions of phenotypic mean traits do not require knowledge of the details of the tracking system, i.e., neither the parameter values in an underlying reaction norm model nor the fitness function need to be known. Note

that the results in Fig. 3 for a system on the verge of extinction indicate that the mean fitness requirement is not very strict.

As indicated in the simulations mean phenotype vs. environment functions identified from fossil time series may be both linear and nonlinear, and additional factors like constraints on individual phenotypic values and additional terms to the $\theta_j(u,t)$ functions will contribute to the nonlinearities. Note that that the simulation results in Fig. 5 are only two special cases of factors that may affect fossil time series, and that other factors possibly also can be investigated by use of simulations. Also note that a period of poor fitness peak tracking will be seen as a dip in the mean phenotype vs. environment function, followed by either recovery or extinction, and that a dip after recovery may or may not be detected in the samples.

In the simulations with the linear fitness peak vs. environment function $\theta(u,t) = 0.5 + u(t) - 2.9$ and use of the linear reaction norm in Eq. (2) we find linear $\bar{y}_{filt}(t)$ vs. $f\big(u_{filt}(t)\big)$ plots. This is so with constant mean intercept and plasticity parameters as obtained in the case with initial mean plasticity $\bar{b}(0) = 1$ (Fig. 2, panel H), which is in line with the findings in de Villemereuil et al. (2020) regarding breeding date in populations of wild animals. But the same linear result is also found in cases with time varying parameters (Fig. 2, panels D and L), and there is thus no need for disentanglement of the effects of microevolution and plasticity, as discussed in Merilä & Hendry (2013) and Lister (2021). This is so because the function $\theta_{filt}(t) = f\big(u_{filt}(t)\big)$ only depends on how $\theta(u,t)$ varies with the environment, and that $\bar{y}_{filt}(t) \approx \theta_{filt}(u,t)$ (Figs. 2, panels C, G and K). From the $\bar{y}_{filt}(t)$ vs. $f\big(u_{filt}(t)\big)$ plots it is possible to fit a straight line also with quite few known values $u_{filt}(i)$ and $\bar{y}_{filt}(i)$ as done in Fig. 7, panel A, and which generally is a valuable property considering the effort needed to obtain fossil data. Fig. 2 also shows that the mean fitness in all three cases is slowly reduced over time. This is so because u_{ref} is chosen as $u_{filt}(0)$, such that the individual phenotypic traits owing to plasticity are increasingly spread around the fitness peak as time increases.

In the nonlinear case simulations with $\theta(u,t) = 0.5 + u(t) - 2.9 + 8(u(t) - 2.9)^2$ and use of the linear reaction norm in Eq. (2) we find that the mean fitness is clearly reduced (Fig. 3, panels B, F and J). We also find that $\bar{y}_{filt}(t)$ lags behind $\theta_{filt}(t)$ (Fig. 3, panels C, G and K), which is a well-known result for tracking systems, and which also is mentioned in de Villemereuil et al. (2020). We also find that the $\bar{y}_{filt}(t)$ vs. $f\left(u_{filt}(t)\right)$ plots indicate a nonlinear function, and as in the linear case we find very similar plots for the three different initial mean slope values (Fig. 3, panels D, H and L). In these plots it is here possible to fit a second order function from quite few known values $u_{filt}(i)$ and $\bar{y}_{filt}(i)$. It is interesting to note that $\bar{b}(0) = 1$ in this case does not result in constant $\bar{a}(t)$ and $\bar{b}(t)$ parameter values. As the combination of a linear reaction norm and a nonlinear $\theta(u,t)$ function leads to reduced mean fitness, it is natural to consider a nonlinear reaction norm, and the mean fitness is in fact considerably improved with a reaction norm $\bar{y}(t) = \bar{a}(t) + \bar{b}(t)(u(t) - 2.9 + 8(u(t) - 2.9)^2)$, where $\bar{a}(t)$ and $\bar{b}(t)$ as in the linear case are evolvable parameters. In this case an initial mean slope value $\bar{b}(0) = 1$ resulted in constant $\bar{a}(t)$ and $\bar{b}(t)$ parameter values.

The simulations with linear and nonlinear fitness peak vs. environment functions were repeated with the fitness function width parameter ω^2 increased from 50 to 100 and reduced to 10, without more than very minor changes in the resulting mean phenotype vs. environment functions.

It is interesting to note that constant $\bar{a}(t)$ and $\bar{b}(t)$ parameter values as found in Fig. 2, panel E, require a linear fitness peak vs. environment function. The main finding in de Villemereuil et al. (2020) that phenotypic plasticity may have evolved to cope with fluctuations in the optimum thus indicates that such a linear fitness peak vs. environment function is common in nature. The more intricate alternative is a nonlinear reaction norm that fits a nonlinear fitness peak vs. environment function.

In the simulations I assumed 2,115 nonoverlapping generations and found $\bar{a}(t)$ and $\bar{b}(t)$ for different initial values $\bar{b}(0)$. With a different timescale more relevant for fossil data, similar $\bar{a}(t)$

and $\bar{b}(t)$ values could plausibly be underlying samples of true and continuously varying parameter values, just as the $\partial^{18}O(t)$ data used as $u(t)$ are samples of a continuous variable.

For the real data case in Liow et al. (2024) I assume that the reference environment u_{ref} is the mean $\partial^{18}O(t)$ value 2.3 million years back in time. Since the data spans 2.3 million years there is a possibility that u_{ref} evolves, as alluded to in Lande (2009) and discussed in detail in Ergon and Ergon (2017), but since $\bar{y}_{filt}(t)$ as function of $u_{filt}(t)$ for the AZ and OV area mean traits are found to be straight lines such changes in u_{ref} have no effect on the results. The time series in this real data case have samples from around 2 million years ago and samples for the last 0.9 million years, and although it would have been nice with additional data from around 1.4 million years ago the available data indicate a linear fitness peak vs. environment function. Considering the large standard errors indicated in Fig. S2 in Liow et al. (2024) it is in any case no clear indications of a very nonlinear function, and it also seems reasonable to assume that long existing species in the quite variable $\partial^{18}O$ environment in general have fitness peak vs. environment functions that are not too difficult to track.

With knowledge of the time series in Fig. 6, panels B, it is obviously tempting to find the predictions in Fig. 7, panel B, without any consideration of the theoretic background, i.e., by use of an *ad hoc* approach. In a more sophisticated approach one could assume an underlying reaction norm model with constant $\bar{a}(t)$ and $\bar{b}(t)$ parameter values, as in Fig. 2, panels A to H. The simulation results in Fig. 2 show, however, that such an assumption is unnecessary, and that the prediction results for *log AZ area mean* and *log OV area mean* thus have a more general theoretical foundation. Note, however, the interpretational difficulties in relation to how other factors than the physical environment affect the results.

As discussed in Di Martino & Liow (2021) the AZ and OV mean traits vary with temperature according to Atkinsson's rule regarding larger body sizes in cooler environments. Or more precisely in my view, the fitness peak moves according to these rules, and the tracking process forces the mean traits to follow.

Liow et al. (2024) reports that plasticity plays a role in the real data case, although in a complex way owing to the clonal and colonial nature of the data. However, as pointed out above it is not necessary to know any details of the underlying mean reaction norm and fitness functions for the purpose of predicting the mean phenotypic traits over time, what is essential is that the fitness peak tracking process keeps the mean fitness at a sufficiently high level for survival. It may of course still be of interest to investigate details in the tracking process, as done in Liow et al. (2024), and the cumbersome data collection and mean trait identification and estimation process has in any case been necessary. If it in addition could be shown that the overall mean plasticity is constant, we would again have a situation in accordance with the findings in de Villemereuil et al. (2020) regarding breeding dates of 21 wild animals, but as shown in the simulations the trait tracking process does not depend on such a constant mean plasticity.

One could finally argue that there is an obvious difference between the studies of de Villemereuil et al. (2020) and Liow et al. (2024), in that de Villemereuil et al. rely on data over tens of years, while fossil data spans millions of years. But for species that have existed over millions of years with only minor size changes as seen in Fig. 6 it is reasonable to assume the quantitative genetic theory is applicable, and that results based on studies of the climate change thus are relevant for comparisons. This is consistent with the Liow et al. (2024) conclusion that “while controls on phenotypes are complex and varied in time, ecological and evolutionary processes that unfold on shorter timescales are not inconsistent with macroevolutionary patterns observed on longer timescales”.

Acknowledgement

I thank University of South-Eastern Norway for support and funding.

Author contribution

Rolf Ergon is the sole author of this article.

Data availability statement

MATLAB code for simulations is given in Supplementary Material archived on bioRxiv

bioRxiv 2024.10.30.621046; doi: https://doi.org/10.1101/2024.10.30.621046

Conflict of interest statement

There are no competing interests.

ORCID

Rolf Ergon https://orcid.org/0000-0003-2393-2997

References

Bell, M.A. (2013). Adaptive Landscapes, Evolution, and the Fossil Record. In The Adaptive Landscape in Evolutionary Biology, Eds. Svensson, E. and R. Calsbeek. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199595372.001.0001

Boutin, S., & Lane J. E. (2014). Climate change and mammals: evolutionary versus plastic responses. Evolutionary Applications 7: 29–41

Charmantier, A., McCleery R. H., Cole, L. R., Perrins C., Kruuk L. E. B., & Sheldon B. (2008). Adaptive phenotypic plasticity in response to climate change in a wild bird population. Science 320:800–803

Charmantier, A., & Gienapp, P. (2014). Climate change and timing of avian breeding and migration: evolutionary versus plastic changes. Evol. Appl. 7, 15–28. https://doi.org/10.1111/eva.12126

Crozier, L. G., & Hutchings J. A. (2014). Plastic and evolutionary responses to climate change in fish. Evolutionary Applications 7:68– 87

de Villemereuil, P. et al. (2020). Fluctuating optimum and temporally variable selection on breeding date in birds and mammals. Proc. Natl Acad. Sci. USA 117, 31 969–31 978. doi:10.1073/pnas.2009003117)

Di Martino, E., & Liow, L. H. (2021). Larger offspring associated with lower temperatures across species of Microporella, a widespread colonial invertebrate. Marine Ecology Progress Series 662:1–13

Di Martino, E. & Liow, L. H. (2022). Changing allometric relationships among fossil and Recent populations in two colonial species. *Evolution*. ISSN 0014-3820. 76(10), p. 2424–2435. doi: 10.1111/evo.14598

Ergon T., & Ergon, R. (2017). When three traits make a line: evolution of phenotypic plasticity and genetic assimilation through linear reaction norms in stochastic environments. J. Evol. Biol. 30:486-500. doi: 10.1111/jeb.13003

Ergon, R. (2023). Dynamical BLUP modeling of reaction norm evolution, accommodating changing environments, overlapping generations, and multivariate data. Ecology and Evolution. 2023;13:e10194. https://doi.org/10.1002/ece3.10194

Goodnight, C. J. (2013). Wright's Shifting Balance Theory and Factors Affecting the Probability of Peak Shifts. In The Adaptive Landscape in Evolutionary Biology, Eds. Svensson, E. and R. Calsbeek. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199595372.001.0001

Hunt, G. (2006). Fitting and comparing models of phyletic evolution: random walks and beyond. Paleobiology 32:578–601.

James, F. C. (1983). Environmental component of morphological differentiation in birds. Science 221:184–186

Johnson, R. A., & Wichern, D. W. (2008). Applied Multivariate Statistical Analysis.

Prentice-Hall, London

Kellogg, D. E. (1975). Character displacement in the radiolarian genus, Eucyrtidium. Evolution 4:736–749.

Lande, R. (1979). Quantitative genetic analysis of multivariate evolution, applied to brain:body size allometry. Evolution 33:402-416

Lande, R. (2009). Adaptation to an extraordinary environment by evolution of phenotypic plasticity and genetic assimilation. J. Evol. Biol. 22, 1435-1446. doi: 10.1111/j.1420-9101.2009.01754.x

Lisiecki, L. E., & Raymo, M. E. (2005). A Pliocene–Pleistocene stack of 57 globally distributed benthic δ18O records. Paleoceanography 20, A1003. doi:10.1029/ 2004PA001071

Liow, L. H., Porto, A. & Di Martino, E. (2024). Trait-Fitness Associations via Fecundity and Competition in a Two-Million-Year-Long Fossil Record. Am. Nat. 204/3. https://doi.org/10.1086/73133

Liow, L. H., Di Martino, E., Krzeminska, M., Ramsfjell, M. H., Rust, S., & Taylor, P. D. (2017). Relative size predicts competitive outcome through 2 million years. *Ecology Letters*. ISSN 1461-023X. 20(8), p. 981–988. doi: 10.1111/ele.12795.

Lister, A. M. (2021). Phenotypic Plasticity in the Fossil Record. In Phenotypic Plasticity & Evolution: Causes, Consequences, Controversies. Ed. *David W. Pfennig*.Imprint CRC Press https://doi.org/10.1201/9780429343001

Merilä, J. & Hendry, A. P. (2013). Climate change, adaptation, and phenotypic plasticity: the problem and the evidence. Evol. Appl. 7, 1-14. doi: 10.1111/eva. 12137.

Sætre, C. L. C., Coleiro, C.,Austad, M., Gauci, M., Sætre, G., Voje, K. L. & F. Eroukhmanoff. (2017). Rapid adaptive phenotypic change following colonization of a newly

restored habitat. *Nature Communication*, **8**, 14159. **https://doi.org/10.1038/ncomms14159**

Teplitsky, C., Mills, J. A., Alho, J. S., Yarall, J. W., & Merilä, J. (2008). Bergmann's Rule and climate change: disentangling environmental and genetic responses in a wild bird population. Proceedings of the National Academy of Sciences U. S. A 105:13492–13496

Voje, K. L. (2023). Fitting and evaluating univariate and multivariate models of within-lineage evolution. Paleobiology, 49(4), (2023), pp. 747–764 DOI: 10.1017/pab.2023.1

Williams, S. E., Shoo, L. P., Isaac, J. L., Hoffmann, A. A., & G. Langham. (2008). Towards an Integrated Framework for Assessing the Vulnerability of Species to Climate Change. PLoS Biol 6(12): e325. doi:10.1371/journal. pbio.0060325