Reaction norm tracking of fitness peaks may explain time series in the fossil record

Rolf Ergon

University of South-Eastern Norway

2.11.2024

Abstract

Species that have existed over millions of years have done so because of high mean fitness through the ages. This means that the mean phenotypic traits of such species continuously have been tracking the optimal phenotypic values that maximize fitness. These optimal values are functions of the environment, for example the sea temperature. The mean phenotypic traits over time will thus 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 short time series with irregular sampling, and I also discuss how this methodology can be extended to multivariate systems. 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 values found in the fossil record, it was possible to identify linear mean phenotype vs. environment functions, and thus to predict the continuous mean phenotypic values as functions of time.

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

Introduction

Evolutionary time series in the fossil record provide information on phenotypic changes over millions of years. Such time series may be the result of 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 no apparent loss of fitness. 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 and Hendry (2013) discussed this problem and summarized the

results of 11 review articles. Among the review papers were Charmantier and Gienapp (2014) (birds), Boutin and Lane (2014) (mammals), and Crozier and Hutchings (2014) (fish), which summarize the findings in altogether 66 studies of responses on climate change. Merilä and 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".

The plasticity contributions in adaptive evolutionary processes are often studied by means of reaction norm models, which in general are multivariate, allowing for correlated traits (Ergon, 2023). For simplicity I will here make use of a univariate 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 kept close to the top of a moving peak in an adaptive landscape. 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 filtered versions of these variables.

Basic theory is given in the Theory section, while simulations are presented in the Simulations section. In the Real data case section, I will apply the theoretical results on two phenotypic traits of the bryozoan species lineage *Microporella agonistes* with a 2.3 million long fossil record (Liow et al., 2024). Finally follows a Discussion section.

Theory

A 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 $\bar{b}(t)$ evolves such that the mean fitness is kept at a near maximum level, $\bar{y}(t)$ will be kept in the vicinity of the phenotypic values that correspond to a 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 $\bar{y}(t)$ in Eq. (2), and that the adaptive landscape thus is two-dimensional. I will later extend some essential results to multivariate cases.

The phenotypic value that maximizes fitness 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. 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).

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)

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

where G_{aa}, G_{bb} and G_{ab} are variances and the covariance of the reaction norm parameters, and where $w(t) = W(t)/\bar{W}(t)$ 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; Ergon, 2022).

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 a residual term. 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 at a near maximum level and the residual $e(t)$ at a low level. In feedback control terminology it is thus the mean trait following error $\theta(u,t) - \bar{y}(t)$ that via the 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 filtered versions of often quite noisy signals. Note that even though $\bar{y}(t) \approx \alpha + f\big(u(t) - u_{ref}\big)$, such a model is not a reaction norm. As verified by simulations, the reaction to a sudden change Δu in $u(t)$ will still be a sudden change $\Delta\bar{y} = \bar{b}(t)\Delta u$.

Multivariate cases

When the mean phenotypic trait in a univariate population is tracking the phenotypic value that maximizes fitness, the mean phenotypic trait will be determined 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 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 and Wichern, 2008)

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

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 near maximum level, 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).

Simulations

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 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 phenotypic value that maximizes fitness 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 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

$\bar{b}(0)$. Since $\bar{y}(t)$, $\theta(u,t)$ and $u(t) = \partial^{18}O(t)$ have large variances, I used 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 are shown in Fig. 1. 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 $\bar{b}_{opt} = 1$. 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 to determine the parameters $\hat{\alpha}$ and $\hat{\beta}$ of this line by an ordinary least squares method, and thus to find predictions $\hat{\bar{y}}_{filt}(t) = \hat{\alpha} + \hat{\beta}\big(u_{filt}(t) - u_{ref}\big)$, 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.

Fig. 2 shows result for a system with a nonlinear fitness peak vs. environment function $\theta(u,t) = 0.5 + (u(t) - 2.9) + 2(u(t) - 2.9)^2$, combined with the linear reaction norm $\bar{y}(t) = \bar{a}(t) + \bar{b}(t)(u(t) - 2.9)$. Note that $\bar{y}_{filt}(t)$ also now is tracking $\theta_{filt}(u,t)$ pretty well (panels C, G and K) and that $\bar{y}_{filt}(t)$ as function of $u_{filt}(t)$ are the same second order functions in the different cases (panels D, H and L). Also 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.

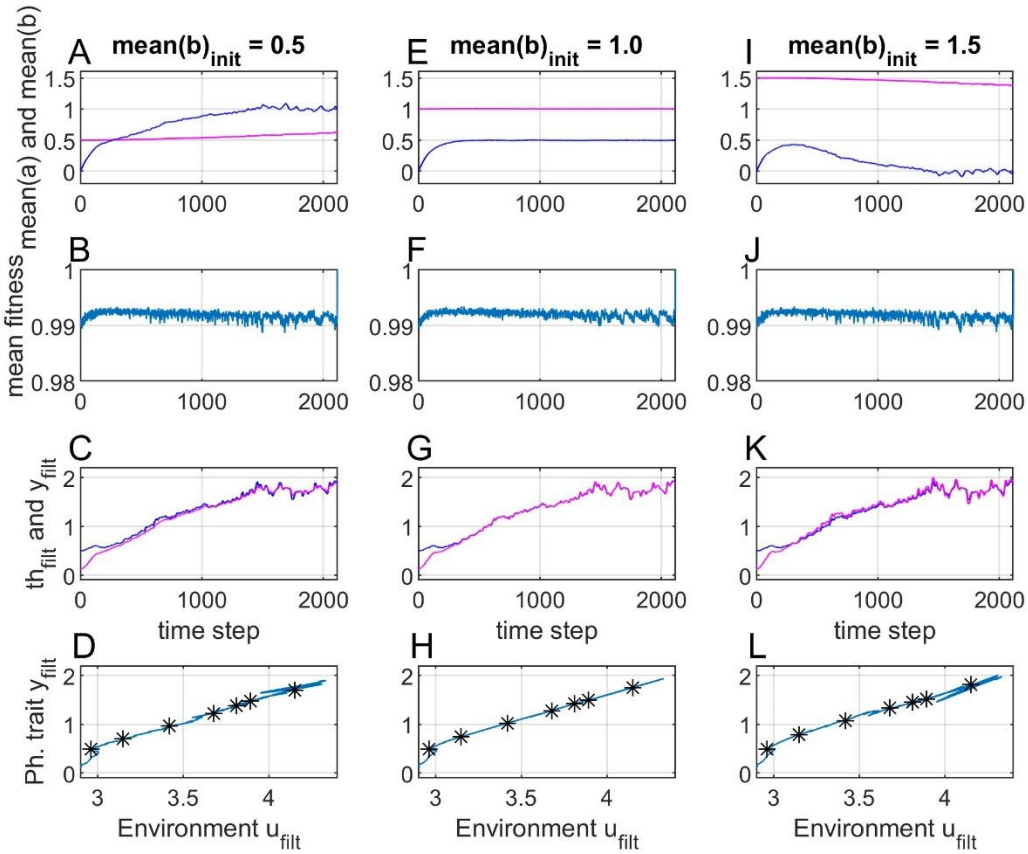

Figure 1. Simulation results with a linear $\theta(u,t) = \alpha + f(u(t))$ function and a linear reaction norm for three different initial values $\bar{b}(0)$, while $\bar{a}(0) = 0$. Panels A, E and I show $\bar{a}(t)$ (blue lines) and $\bar{b}(t)$ (red lines). Panels B, F and J show mean fitness $\bar{W}(t)$. 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.

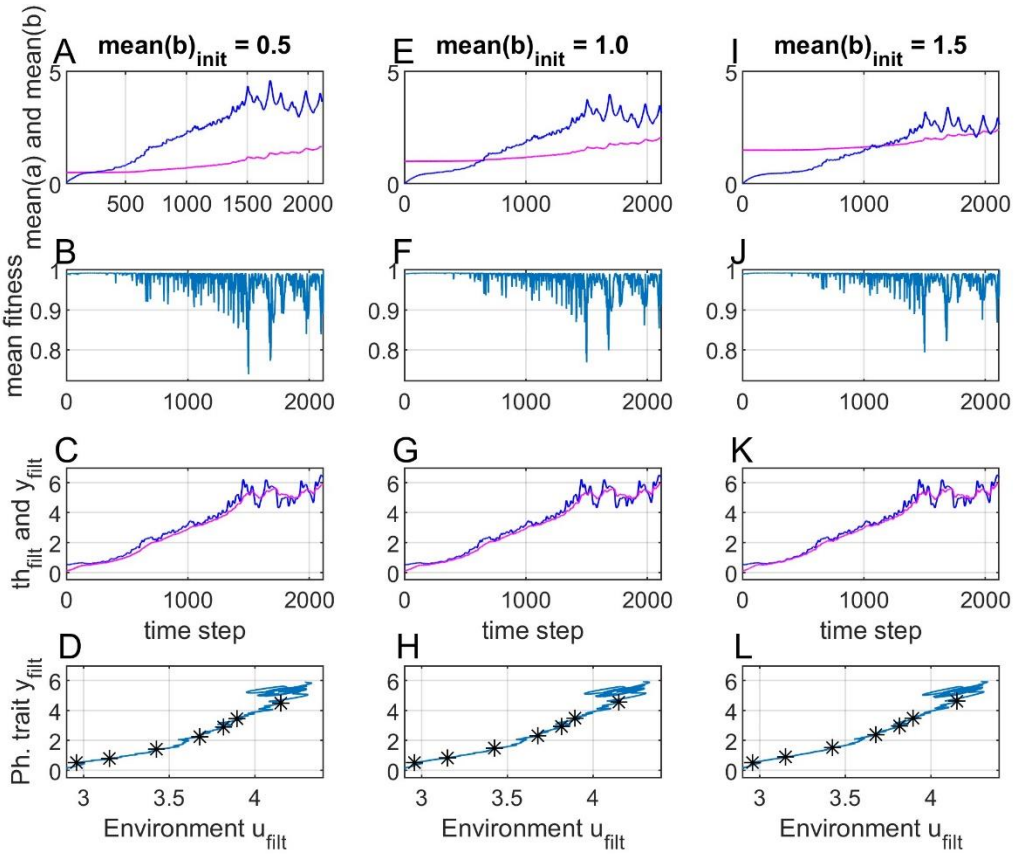

Figure 2. Simulation results as in Fig. 1, but with a nonlinear $\theta(u,t) = \alpha + f(u(t))$ function. The discrete values in panels D, H or L could here be used to determine the nonlinear function $\hat{y}_{filt}(t) = f\left(u_{filt}(t)\right)$ by means of any suitable fitting method.

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 called DeepBryo, and they then analyzed the data using scanning electron microscope pictures. 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 trait values increase with time. 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 Liew 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, see Fig. 6 in Liow et al. (2024). Fig. 3 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 window filtered version $u_{filt}(t)$ of $\partial^{18}O$, with window size 300. 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, i.e., highly filtered information (Liow et al., 2024).

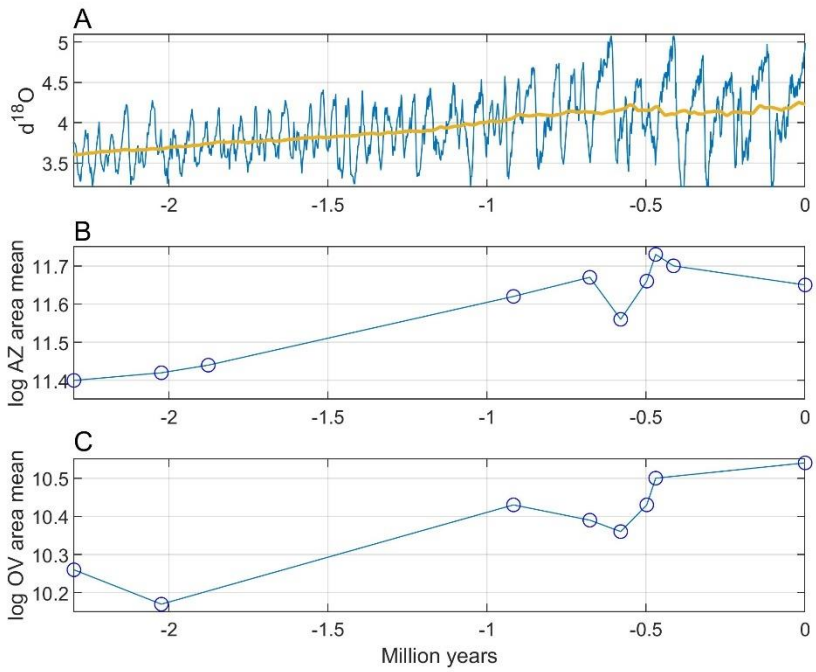

Figure 3. Raw $\partial^{18}O(t)$ data (panel A, blue line) and a filtered version $u_{filt}(t)$ (panel A, red line), as well as *log AZ area mean* (panel B, circles) and *log OV area mean* (panel C, circles) as in Fig. S2 in Liow at al. (2024). Present time is used as zero point.

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, and 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 high enough 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. 1 and 2).

Fig. 4, panel A, shows *log AZ area mean* as function of $u_{filt}(i)$ for the 10 data points in Fig. 3, panel B. This plot should be compared with the corresponding plots in Figs. 1 and 2 (panels D, H and L), and it is rather obvious that it is not possible to fit more than a straight line $log\ AZ\ area\ mean\ = \hat{\alpha} + \hat{\beta}\left(u_{filt}(i) - u_{ref}\right)$. With $u_{ref}\ = u_{filt}(-2.3) = 3.6$ as found from Fig. 3, panel A, the ordinary least squares method results in $log\ AZ\ area\ mean\ (t) = 11.39 + 0.48\left(u_{filt}(t) - 3.6\right)$, as shown in Fig. 4, panel B. With unfiltered environmental data we thus find $log\ AZ\ area\ \mathrm{mean}\ (\mathrm{t}) = 11.39 + 0.48(\partial^{18}O(t) - 3.6)$.

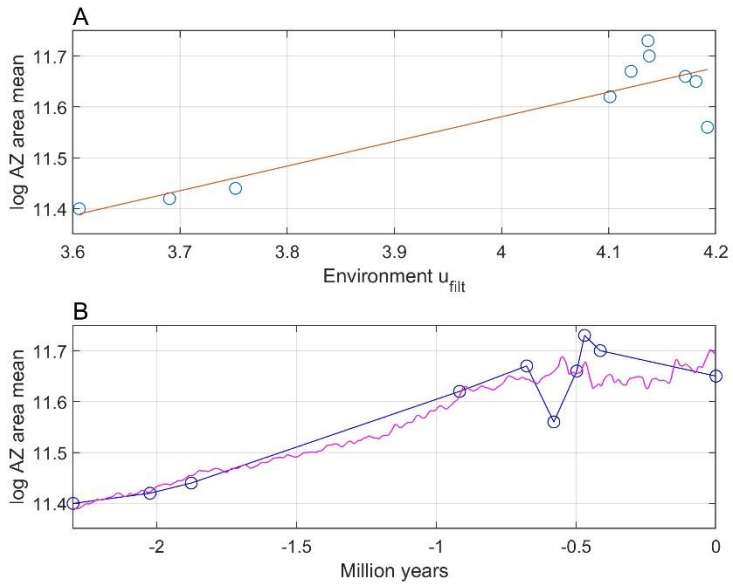

Figure 4. Panel A shows $log\ AZ\ area\ mean$ as function of $u_{filt}(i)$ for 10 data points as in Fig. S2 in Liow at al. (2024) (circles), as well as a straight line fitted by the ordinary least squares method (red line). Panel B shows $log\ AZ\ area\ mean$ as function of time (blue circles), and the fitted reaction norm response (red line).

Finally, Fig. 5 shows results as in Fig. 4, but with the fitted predictions $log\ OV\ area\ mean\ (t) = 10.20 + 0.43\big(u_{filt}(t) - 3.6\big)$. Note that compared with Fig. 4, the number of data points is 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 $log\ OV\ area\ mean\ (t)$ for all practical purposes are identical to $0.9 \times log\ AZ\ area\ mean\ (t)$, and the two variables are thus linearly dependent.

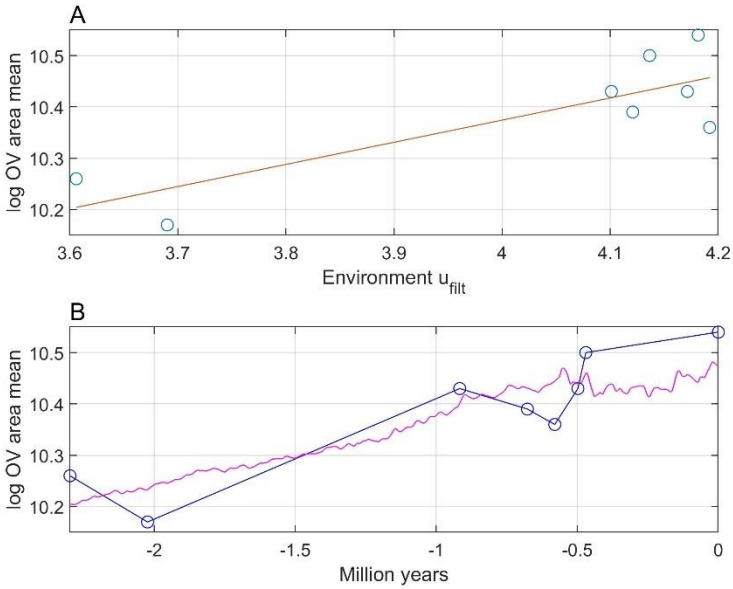

Figure 5. Panel A shows $log\ OV\ area\ mean$ as function of $u_{filt}(i)$ for 8 data points as in Fig. S2 in Liow at al. (2024) (circles), as well as a straight line fitted by the ordinary least squares method (red line). Panel B shows $log\ OV\ area\ mean$ as function of time (blue circles), and the fitted reaction norm response (red line).

If the unfiltered $\partial^{18}O(t)$ data are used as inputs to the model for $log\ AZ\ area\ mean\ (t)$ found by use of filtered $\partial^{18}O(t)$ data we would obtain the results in Fig. 6, which illustrates the sampling difficulties involved, and thus underlines the need for filtering. For the $\partial^{18}O(t)$ filtering, I here used a moving window with a fixed width of 300 samples, but since the data is collected with various window sizes a variable filtering window size could possibly improve the predictions.

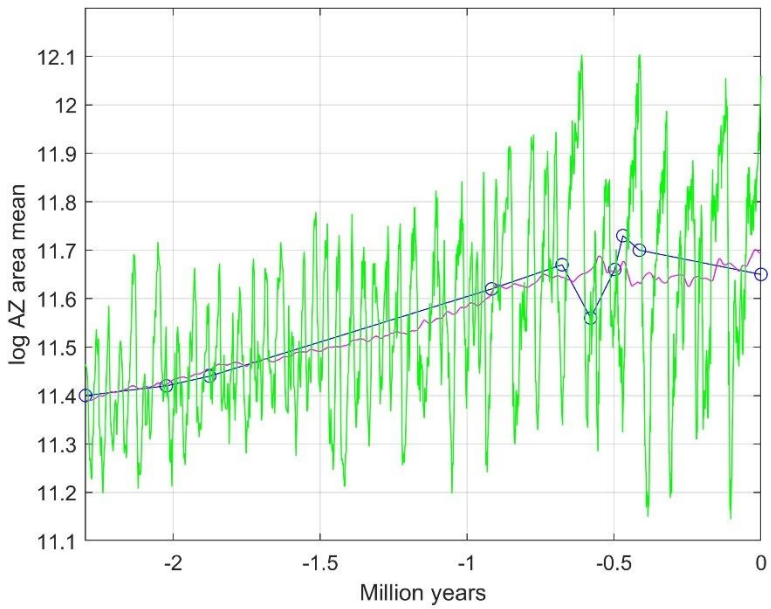

Figure 6. Predictions as in Fig. 4, panel B (red line), and with unfiltered $\partial^{18}O(t)$ data used as inputs to the same prediction model (green line).

Discussion

There are two main conclusions to draw from the results in this article. First, provided that the tracking system keeps the mean fitness at a high level throughout, the mean phenotypic traits in a possibly multivariate system, $\bar{y}_1(t), \bar{y}_2(t), \ldots \bar{y}_m(t)$, will stay in the vicinity of a moving peak in the adaptive landscape defined by the phenotypic values $\theta_1(u,t), \theta_2(u,t), \ldots \theta_m(u,t)$ that maximize fitness. Second, provided that the mean fitness is kept at a high level, it is not necessary to know the details of the tracking process that keeps $\bar{y}_j(t) \approx \theta_j(u,t)$, i.e., neither the parameter values in an underlying reaction norm model nor the fitness function need to be known. Under the assumption of a single environmental variable the quite simple result is thus that the function $\bar{y}_j(t) = f_j\big(u(t) - u_{ref}\big)$ for a given phenotypic trait approximately is equal to $\theta_j(u,t) = f_j\big(u(t) - u_{ref}\big)$ for the same trait, and that $\bar{y}_1(t), \bar{y}_2(t), \ldots, \bar{y}_m(t)$ thus can be predicted by fitting also quite short fossil data. With a dynamical adaptive landscape, one

must, however, expect that the mean phenotypic traits will be lagging behind the fitness peak, as also mentioned in de Villemereuil et al. (2020).

It should be noted that the conclusions above are not valid for time periods with so sudden changes in the adaptive landscape that they in a transition period markedly affect the mean fitness, as exemplified in Sætre et al. (2017). Such periods with poor mean trait tracking could possibly also be found in the fossil record used by Liow et al. (2024), although they might not be detected owing to the irregular sampling, or they cannot be detected owing to the large standard errors as seen in Fig. S2.

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. 1, 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. 1, panels D and L), and there is thus no need for disentanglement of the effects of microevolution and plasticity, as discussed in Merilä and Hendry (2013) and Lister (2021). This is so because the function $\bar{y}_{filt}(t) = f\big(u_{filt}(t)\big)$ only depends on how $\theta(u,t)$ varies with the environment, and that $\bar{y}_{filt}(t) \approx \theta_{filt}(u,t)$ (Figs. 1, 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 from quite few known values $u_{filt}(i)$ and $\bar{y}_{filt}(i)$, which is a valuable property considering the effort needed to obtain fossil data. Fig. 1 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 values owing to plasticity are increasingly spread around the fitness peak as time increases.

In the nonlinear case simulations with $\theta(u,t) = 0.5 + u(t) - 2.9 + 2(u(t) - 2.9)^2$ and use of the linear reaction norm in Eq. (2) we find that the mean fitness is somewhat reduced (Fig. 2), and that the $\bar{y}_{filt}(t)$ vs. $f\big(u_{filt}(t)\big)$ plots indicate a nonlinear function. As in the linear case we

find very similar plots for the three different initial mean slope values (Fig. 2, 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 $\bar{y}(t) = \bar{a}(t) + \bar{b}(t)(u(t) - 2.9 + 2(u(t) - 2.9)^2)$, 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. When the nonlinear $\theta(u,t)$ function was changed into $\theta(u,t) = 0.5 + u(t) - u_{ref} + 8\left(u(t) - u_{ref}\right)^2$ the mean fitness in simulations with a linear reaction norm turned out to be zero in several periods around $t = 1{,}500$, which illustrates that there are limits to how much nonlinearities in the $\theta(u,t)$ function an evolutionary process can track.

It is interesting to note that constant $\bar{a}(t)$ and $\bar{b}(t)$ parameter values as found in Fig. 1, panel E, require a linear fitness peak vs. environment function. The main finding in de Villemereuil et al. (2020) that phenological 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 time scale 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 $\bar{y}_{filt}(i)$ as function of $u_{filt}(i)$ 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 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 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.

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 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 high level. 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 a constant mean plasticity. One could argue that there is an obvious difference between the studies of de Villemereuil et al. (2020) and Liow et al. (2024), in that individual birds and small mammals can change breeding time from year to year, while bryozoans can change size only from generation to generation. But over time the effects will be comparable in that both microevolution and plasticity may be involved, as illustrated in Figs. 1 and 2.

Acknowledgement

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

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

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

Charmantier, A., and P. Gienapp. 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., and J. A. Hutchings. 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)

Ergon T. and R. Ergon. 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. 2022. The important choice of reference environment in microevolutionary

climate response predictions. Ecology and Evolution 12(4):e8836. https://doi.org/10.1002/ece3.8836

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. and D.W. Wichern. 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. and M. E. Raymo. 2005. A Pliocene–Pleistocene stack of 57 globally distributed benthic δ18O records. Paleoceanography 20, A1003. doi:10.1029/ 2004PA001071

Liow, L. H., A. Porto, and E. Di Martino. 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

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. and A. P. Hendry. 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., C. Coleiro, M.Austad, M. Gauci, G. Sætre, K. L. Voje and 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., J. A. Mills, J. S. Alho, J. W. Yarall, and J. Merilä. 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., L. P. Shoo, J. L. Isaac, A. A. Hoffmann, and 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