Adaptive peak tracking as explanation of sparse fossil data across fluctuating ancient environments

Rolf Ergon

rolf.ergon@usn.no

University of South-Eastern Norway

May 8. 2025

Abstract

Species that have existed over millions of years have done so because they have been able to track peaks in an adaptive landscape well enough to survive and reproduce. Such optima are defined by the mean phenotypic values that maximize mean fitness, and they are predominantly functions of the environment, for example the sea temperature. The mean phenotypic values 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, using both a basic non-plastic selection model and a univariate intercept-slope reaction norm model. I 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 discuss how this methodology can be extended to multivariate systems. The simulations include cases with a constraint on the individual trait values and with other factors than environment influencing the position of the adaptive peak. The methodology is finally applied on a time series of mean phenotypic values in a record of bryozoan *Microporella agonistes* fossils spanning 2.3 million years, using the $\partial^{18}O$ measure as proxy for sea water temperature.

From as few as eleven samples of mean phenotypic values found in the fossil record it was possible to identify a linear mean phenotype vs. environment function, and to predict the continuous mean phenotypic values as functions of time with prediction errors within or just outside the standard errors of the observations. Leave-one-out cross validation gave satisfactory results. It remains to verify predictions for longer time periods without known or investigated fossil data.

Keywords: adaptive landscape, adaptive peak tracking, bryozoans, fitness landscape, fossil record, optimal phenotypic values

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 sparse samples, and they are traditionally modeled by various forms of random walk processes (Hunt, 2006; Voje, 2023), or by Ornstein-Uhlenbeck (OU) processes describing evolution of mean phenotypic values toward fixed or moving peaks in the adaptive landscape (Voje, 2023). Here, the adaptive landscape and adaptive peak metaphors assume that the mean population fitness is a function of possibly multivariate mean phenotypic values, see Pigliucci (2013) for the historical background and a clarifying discussion.

As an alternative to the traditional fossil time series modeling, I propose that adaptive peak tracking may be a useful modeling approach, and that this requires two types of landscape metaphors. First, I assume a fitness landscape where individual fitness is a function of a possibly multivariate set of individual phenotypic values. Second, I assume an adaptive landscape where the mean population fitness is a function of a possibly multivariate set of mean phenotypic values. For simplicity I assume that the position of the individual fitness peak is a function of a dominating environmental factor, and that the adaptive peak may have a

somewhat different position owing to non-symmetrical individual fitness or trait distribution functions. I am thus not attempting to estimate the effect of various selective or other factors on the position of adaptive optima (Hansen, 1997), but I just assume that various selective factors over time keep the mean phenotypic values in the vicinity of the adaptive peak, and as an example I will use the relative size of feeding zooids (autozooids) in encrusting cheilostome bryozoans. Liow et al. (2017) studied communities of such organisms spanning more than 2 million years, and an interesting result was that six different species showed temporally coordinated changes in average zooid sizes, suggesting that the different species responded to a common external driver. Liow et al. pointed to sea temperature as a potential common driver, but they found that the available time series of changes in mean zooid size lacked sufficient length to rigorously test such a hypothesis. In my view, a study by means of an adaptive peak tracking model as introduced above is clarifying.

For an introduction, Fig. 1 shows data from Liow et al. (2024) (see a more detailed presentation in Section 4), where the $\partial^{18}O(t)$ measure is a widely used proxy for seawater temperature (Lisiecki and Raymo, 2005). A natural question is now which type of model we need for explanation of the *log AZ area mean* data in panel B given the $\partial^{18}O(t)$ data in panel A, and as already mentioned an adaptive peak tracking model is in my view a viable alternative.

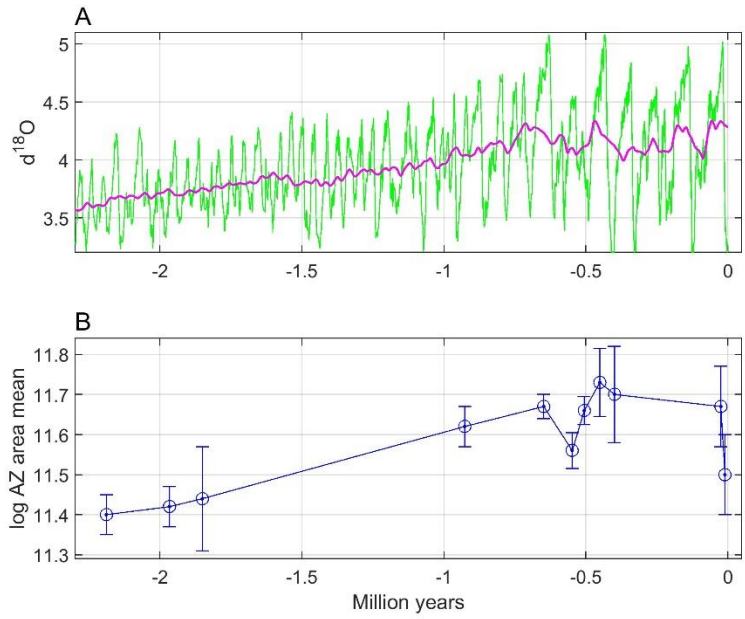

Figure 1. Raw $u(t) = \partial^{18}O(t)$ data (panel A, green line) and a centered moving average signal $u_{filt}(t)$ computed from a sliding window of length 100 samples along the time axis (panel A, red line, see Section 4 for details), as well as *log AZ area mean* (panel B, circles and dots) with standard errors as in Fig. S2 in Liow at al. (2024). Present time is used as zero point. Note that the moving average signal cannot capture the sharp decline of the $\partial^{18}O(t)$ value towards time zero.

A block diagram for the proposed adaptive peak tracking model is shown in Fig. 2, and as described above it is based on two types of landscape metaphors. The optimum phenotype that maximizes individual fitness is assumed to be a function $\theta_i(u,t) = f(u(t))$, where $u(t)$ is the dominating environmental variable (in the bryozoan example the $\partial^{18}O(t)$ measure). The optimum mean phenotypic value that maximizes mean fitness is assumed to be a function $\theta_m(u,t) = f(u(t))$, deviating from $\theta_i(u,t)$ when either the individual fitness function or the trait distribution is non-symmetrical. When both the fitness function and the trait distribution are symmetrical, we will thus have $\theta_m(u,t) = \theta_i(u,t)$ (see Section 2 for an example with Gaussian functions). The tracking process in Fig. 2 is attempting to keep the mean phenotypic value $\bar{y}(t)$ (in the bryozoan example the *log AZ area mean* value) in the vicinity of the optimal value

$\theta_m(u,t)$, such that the tracking error $e(t)$ is kept small. The basic assumption is thus that the tracking process keeps the tracking error small, regardless of the details of this process.

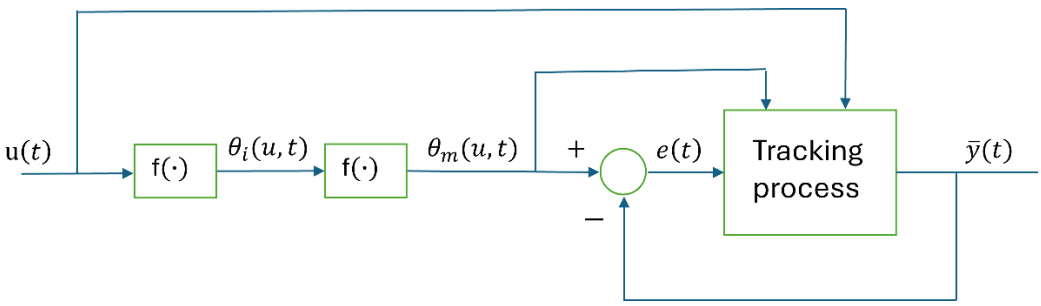

Figure 2. Block diagram for fitness peak tracking system, with the environmental variable $u(t)$ as input and the mean phenotypic value $\bar{y}(t)$ as output. The phenotypic values $\theta_i(u,t)$ and $\theta_m(u,t)$ that maximize individual and mean fitness, respectively, are together with the tracking error $e(t)$ internal signals in the system. See text for explanations of blocks marked by $f(\cdot)$ symbols.

A fluctuating environment as shown in Fig. 1 corresponds to an adaptive landscape where the adaptive peak position fluctuates, and how well the peak is tracked depends on the properties of the evolutionary system. For a population with phenotypic plasticity the tracking error may be kept small at all times, while a population without plasticity may only be able to track a moving average as indicated in Fig. 1. However, a population without plasticity may also keep the tracking error small, provided that additive genetic variances are large. Sparse fossil data will in any case not be informative enough for predictions of strongly fluctuating mean phenotypic values, and a remaining possibility is therefore to predict moving averages over a window size that may be used as prediction parameter. In cases where the fossil phenotypic data are mean values from given time periods, a natural choice is also to use mean environmental data from these time periods as input to the tracking model (see Section 4 for example).

The assumption that the mean phenotypic value stays in the vicinity of the adaptive peak is supported by findings regarding multiple species and traits indicating that “fossil lineages have

ascended adaptive peaks and remained at their summits as they shift position through time" (Bell, 2013). An important reference is here Estes and Arnold (2007), who based on a large data set found that "the underlying process causing phenotypic stasis is adaptation to an optimum that moves within an adaptive zone with stable boundaries", which applied to the data in Fig. 1, panel B, and considering the standard errors given, means that *log AZ area mean* moves between approximately 11 and 12. Using a database over many contemporary populations, Estes and Arnold also found that phenotypic means are typically close to the adaptive peak (46% within 1 phenotypic standard deviation of the optimum, and 65% within 2 standard deviations). Estes and Arnold (2007) based their results on tests with a single displacement of the optimum, using an extensive set of data (Gingerich, 2001), but they acknowledged that the actual optimum might move many times with a small total net displacement. That is a good description of the optimal mean phenotype corresponding to the environmental data in Fig. 1, panel A, with large short-term variations and a rather limited long-time net variation. I thus assume that short-term environmental fluctuations as in, e.g., Fig. 1, panel A, result in variations in optima within a stable adaptive zone (Uyeda et al., 2011). However, contrary to Uyeda et al. I assume a causal tracking system and not a purely stochastic model. In real cases there will of course be stochastic variations in the environmental variable and in the position of the individual fitness peak, and additional stochastic factors that forces the adaptive peak to deviate from the individual fitness peak. Such stochastic variations are included in simulations, but further discussion of this topic is beyond the scope of this paper.

A persistent system excitation in the form of short-term environmental fluctuations as exemplified in Fig. 1, panel A, makes it unlikely that a tracking system as shown in Fig. 2 gets stuck in a local optimum or in a valley in the adaptive landscape. As illustrated in simulations it is more likely that the mean phenotypic values fluctuate around global optima, see however an example of slow adaptation in subsection 3.5 and a discussion in Section 5.

In simulations I assume that the tracking process block in Fig. 2 can be described by use of Lande's selection theory for populations without plasticity (Lande, 1979), which was also used by Estes and Arnold (2007), but I also include examples with phenotypic plasticity (Lande, 2009). In real cases, however, other mechanisms such as, e.g., adaptive walks in response to mutations may also be involved (Ch. 27, Walsh and Lynch, 2018). It is as already mentioned essential to realize that although knowledge of the details of the tracking process may be interesting, such knowledge is not necessarily needed for finding continuous predictions $\hat{\bar{y}}(t)$ based on sparse measurements of mean phenotypic values $\bar{y}(t)$.

Under the assumption that the mean trait values have remained centered around the mean fitness peaks as they have shifted position through time, such that the following error $e(t)$ in Fig. 2 has remained small relative to the width of the underlying mean fitness function, the position of the mean fitness peak as function of the environment, $\theta_m(u,t) = f(u(t))$, can be reconstructed from fossil samples of $\bar{y}(t)$, and provided that this function is sufficiently smooth this can be done also when the fossil samples are sparse and at irregular time intervals. For the bryozoan data in Fig. 1, it turns for example out that the mean trait value $\bar{y} = log\ AZ\ area\ mean$ as function of the environment $u = \partial^{18}O$ can be modeled by a straight line, from which follows continuous predictions of $\hat{\bar{y}}$ as function of time (see Section 4 for details).

The prediction method outlined above is supported by simulations in Section 3, with the theoretical background given in Section 2. The simulations are based on a reaction norm model where I use the $\partial^{18}O(t)$ measure as environmental cue, and where the phenotypic plasticity may be set to zero. This gives a connection between the simulations and the real bryozoan data case in Section 4, in the sense that the same type of large but short-time fluctuations in the input signal is used. Since both $\partial^{18}O(t)$ and $\bar{y}(t)$ in the simulations have large variances, and since data from the fossil record often are mean values over both populations and time (Liow et al., 2024), I will in the simulations use moving average filtered versions of these variables. The

simulations are intended to show that mean phenotype vs. environment functions can be predicted from sparse data also in cases where the population in shorter time periods have quite low mean fitness. The simulations also show that $\bar{y}(t)$ fluctuates around the optimal value $\theta_m(u,t)$, as discussed above, although in some cases with a minor time lag. In order to show that $\theta_m(u,t)$ may be different from $\theta_i(u,t)$, a case with a constraint on individual trait values is included. In addition to theory necessary for the simulations, as given in Section 2, the theory is extended to cover multivariate cases.

In Section 4, the prediction method is applied on recently presented data from an already introduced field study of the bryozoan species *Microporella agonistes*, based on a fossil record over 2.3 million years (Liow et al., 2024). Parameters in a linear mean phenotype vs. environment function are here estimated from as few as 11 data points at irregular sampling times. This is done using mean values of the $\partial^{18}O(t)$ measure in the time periods of collected fossil data, but also, and with better results, using moving average mean environmental data. Since the moving average mean as shown in Fig. 1, panel A, cannot capture the sharp decline of the $\partial^{18}O(t)$ measure towards time zero, the best results are found by using a hybrid model with the moving average mean for the last 0.02 million years replaced by the mean value of $\partial^{18}O(t)$ in that time period. Finally follows a summary and discussion in Section 5.

2 Theory

2.1 A basic univariate reaction norm case

For simplicity and for use in simulations I will initially assume that mean fitness is determined by a univariate phenotypic mean value $\bar{y}(t)$, and that the adaptive landscape thus is two-dimensional. I will later extend some essential results to multivariate cases, and the simulations will also include examples where other factors than the physical environment are involved.

In the simulations I will make use of 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 $b(t)$ and the variance $G_{bb} = var(b(t))$ in some cases are set to zero. Here, $u(t)$ is the environmental cue, while $\varepsilon(t)$ is assumed to be normally distributed with mean 0 and variance σ_ε^2 among individuals in every generation. In Eq. (1), u_{ref} is a reference environment defined as the environment at which the phenotypic variance is at a minimum (Lande, 2009). For a given population this will result in the mean reaction norm

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

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

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

in the linear case $\theta_i(u,t) = \alpha + \beta\big(u(t) - u_{ref}\big)$, where α and β are constant parameters.

In the simulations I will assume an individual fitness function

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

which assuming that $y(t)$ is normally distributed gives the mean fitness in a population (Lande, 2009; Estes and Arnold, 2007)

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

where $\sigma_y^2(t)$ is the variance of $y(t)$ according to Eq. (1). The mean fitness will thus be maximized when $\bar{y}(t) = \theta_i(u,t)$. Note that this is an example with symmetrical individual fitness and trait distribution functions, such that optimal individual and mean phenotypic values are equal, i.e., $\theta_m(u,t) = \theta_i(u,t)$. In Subsection 3.3 I will, however, also include a case where $\theta_m(u,t) \neq \theta_i(u,t)$.

Under the assumptions of non-overlapping 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 $w(t) = W(t)/\bar{W}(t)$ is the relative fitness. In a stationary stochastic environment, the optimal reaction norm slope is $\bar{b}_{opt} = cov(\theta_i(u,t), u(t))/var(u(t))$ (Lande, 2009).

When the mean phenotypic trait in a population is tracking a moving optimum $\theta_m(u,t)$ it will be given by $\bar{y}(t) = \theta_m(u,t) - e(t)$, where $e(t)$ is the tracking error (see block diagram in Fig. 2). This means that the parameter values $\bar{a}(t)$ and $\bar{b}(t)$ are forced to evolve such that $\bar{y}(t) \approx \theta_m(u,t) = f\big(u(t) - u_{ref}\big)$, i.e., such that the mean fitness according to Eq. (4b) is kept close to the optimum and that the tracking error is small. In feedback control terminology it is thus the mean trait tracking error $e(t) = \theta_m(u,t) - \bar{y}(t)$ that via mean fitness is the driving force in the adaptive process. It follows from Eq. (5) that the tracking system in Fig. 2 has what in feedback control terminology is called integral action (Åström and Murray, 2008), which means that the tracking error in a stationary case with a fixed adaptive peak will asymptotically go to zero. As shown in simulations in Section 3 the parameters in the $\bar{y}(t) \approx \theta_m(u,t) = f\big(u(t) - u_{ref}\big)$ function 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. In the real data case in Section 4 the parameters in $\theta_m(u,t)$ are also found from data points $[u_{mean}(t), \bar{y}_{mean}(t)]$ for the time periods of collected fossil data. Note that even though $\bar{y}(t) \approx f\big(u(t) - u_{ref}\big)$, 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 details in the tracking process in Fig. 2 as used in the simulations 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 predictions $\hat{\bar{y}}(t)$ based on only a few fossil samples at irregular time points. The remaining possibility for prediction of $\bar{y}(t)$ is to assume good enough tracking such that $\bar{y}(t) \approx \theta_m(u,t) = f\big(u(t) - u_{ref}\big)$, which as shown in the simulations as well as in the real data case makes it possible to estimate the parameters in the $\theta_m(u,t)$ function also from few $\big[u_{filt}(t), \bar{y}_{filt}(t)\big]$ or $[u_{mean}(t), \bar{y}_{mean}(t)]$ data points. Based on those estimates and known values of $u(t)$ it is then possible to predict $\bar{y}(t)$.

2.2 Influence of other factors than the environment

Eqs. (2) and (3) assume that $\mathrm{u}(t)$ is the only factor that affects the mean phenotypic value $\bar{y}(t)$ and the optimal mean phenotypic value $\theta_m(u,t) = \theta_i(u,t)$, respectively. When other factors affect $\bar{y}(t)$, the tracking system in Fig. 2 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, and then most likely in the form of nonlinearities as shown in a simulation example in Subsection 3.3. When other factors affect $\theta_m(u,t)$ it is again the $\bar{a}(t)$ and $\bar{b}(t)$ parameter values that must compensate, also then most likely in the form of nonlinearities as shown in Subsection 3.4. A linear mean phenotype vs. environment function thus indicates that no other factors than the environment affect the fitness vs. environment function, although there may of course be several other factors that compensates the effects of each other. And in cases where $\theta_i(u,t)$ is a nonlinear function there may be factors that make $\theta_m(u,t)$ linear.

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 $\bar{y}(t) \approx \theta_m(u,t) = f\big(u(t) - u_{ref}\big)$, i.e., by the adaptive peak vs. environment function. Here, it is again important to note that we do not need to know the details of the tracking process in Fig. 2.

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)

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

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 traits in the vicinity of the adaptive peak, they must in this case continuously evolve such that $\bar{y}_1(t) \approx \theta_{m,1}(u,t)$ and $\bar{y}_2(t) \approx \theta_{m,2}(u,t)$, and we will thus find

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

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

where the moving adaptive peak vs. environment functions may be linear or nonlinear, and where 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). Note that we here are only interested in the $\theta_{m,1}(u,t)$ and $\theta_{m,2}(u,t)$ functions that define the position of the adaptive peak, and that we do not need to know the shape of the adaptive landscape around the peak.

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 400. 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). 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 mean phenotypic value was in a first linear case chosen as $\theta_m(u,t) = \theta_i(u,t) = u(t) - 2.9$, where $\theta_m(u,t)$ 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 $t = 0$ (however, see results with added white noise below).

As it is not realistic that $u(t)$ and $\theta_m(u,t)$ are purely deterministic signals, correlated white noise with variances $\sigma_u^2 = 0.05$ and $\sigma_\theta^2 = 0.2$, respectively, were added to $\theta_m(u,t)$ and $u(t)$. The covariance was $\sigma_{\theta,u} = 0.025$, see Lande (2009) for a possible explanation.

In a first case without phenotypic plasticity, the reaction norm parameter variances and covariance were chosen as $G_{aa} = 2$, $\sigma_\varepsilon^2 = 0.5G_{aa} = 1$, $G_{bb} = 0$ and $G_{ab} = 0$, while the initial parameter values were $\bar{a}(0) = \bar{b}(0) = 0$. 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 $u_{i,filt}(t)$ in some of the plots. These were obtained by moving window filters with window size 100, in MATLAB notation $\bar{y}_{filt}(t) = movmean(\bar{y}(t), 100)$, $\theta_{i,filt}(u,t) = movmean(\theta_i(u,t), 100)$ and

$u_{filt}(t) = movmean(u(t), 100)$. Simulation results in the linear case without plasticity are shown in Fig. 3, panels A to E. Note that $\bar{y}(0) = \bar{a}(t)$ follows the main trend in $u(t)$, and that the fluctuations are somewhat delayed compared to the fluctuations in $u(t)$ (panel A).

In a second case with phenotypic plasticity, the reaction norm parameter variances and covariance were chosen as $G_{aa} = 2$, $\sigma_{\varepsilon}^2 = 1$, $G_{bb} = 2$ and $G_{ab} = 0$, while the initial parameter values were $\bar{a}(0) = 0$ and $\bar{b}(0) = 0.8$. Simulation results in the linear case with plasticity are shown in Fig. 3, panels F to J. As follows from the reaction norm equations (1) and (2), the variance $\sigma_y^2(t)$ will increase with $u(t) - u_{ref}$, and as a result the mean fitness will be reduced (panel H).

In both cases (without and with plasticity) $\bar{y}_{filt}(t)$ tracks $\theta_{m,filt}(u,t) = \theta_{i,filt}(u,t)$ closely (panels D and I), such that the mean fitness is kept at a high level (panels C and H). This results in linear functions $\bar{y}_{filt}(t) = f\left(u_{filt}(t)\right)$ (panels E and J), with seven discrete values marked by circles. 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 examples in Section 4.

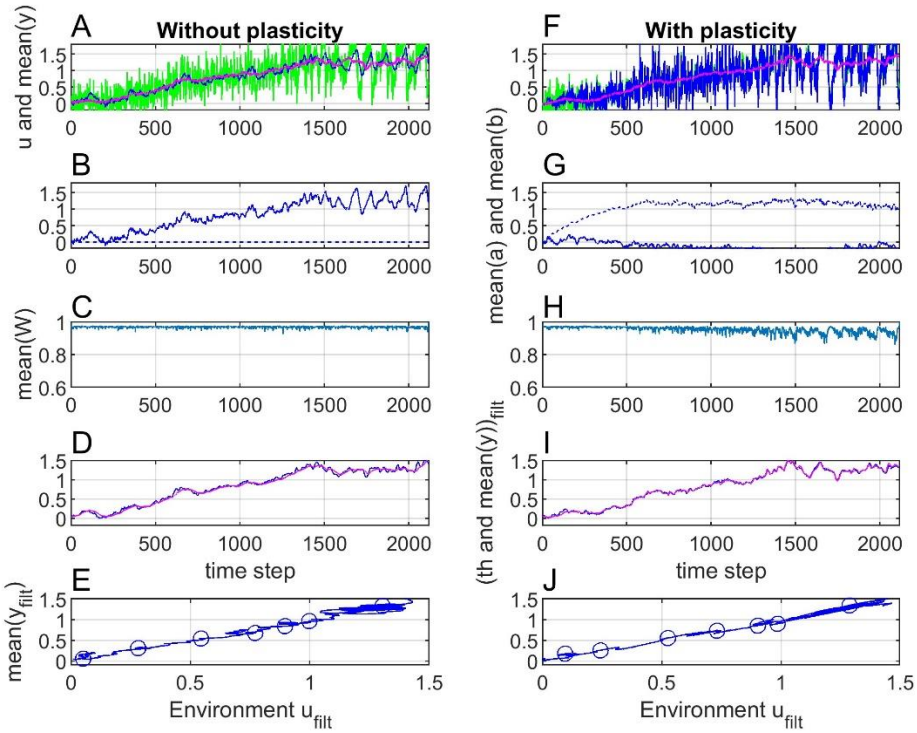

Figure 3. Simulation results with the linear $\theta_m(u,t) = \theta_i(u,t) = u(t) - 2.9$ function, without plasticity (panels A to E) and with plasticity (panels F to J). Panels A and F show $u(t)$ (green lines) and $\bar{y}(t)$ (blue lines), while the red line is $u_{filt}(t)$. Panels B and G show $\bar{a}(t)$ (solid lines) and $\bar{b}(t)$ (dashed lines), while panels C and H show mean fitness $\overline{W}(t)$. Panels D and I show $\theta_{i,filt}(u,t)$ (blue lines) and $\bar{y}_{filt}(t)$ (red lines). Panels E and J show that $\bar{y}_{filt}(t)$ as function of $u_{filt}(t)$ in both cases is the same straight line, with values for samples 200, 400, 600, 800, 1000, 1200 and 1400 shown by circles.

3.2 Nonlinear reaction norm example

Fig. 4 shows results as in Fig. 3, but with the nonlinear adaptive peak vs. environment function $\theta_m(u,t) = \theta_i(u,t) = (u(t) - 2.9) + 4(u(t) - 2.9)^2$. Note that the mean fitness now is reduced (panels B and F), especially in the case without plasticity. 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 and G). Here, $\bar{y}_{filt}(t)$ as function of $u_{filt}(t)$ is the same second order function in the different cases (panels D and H).

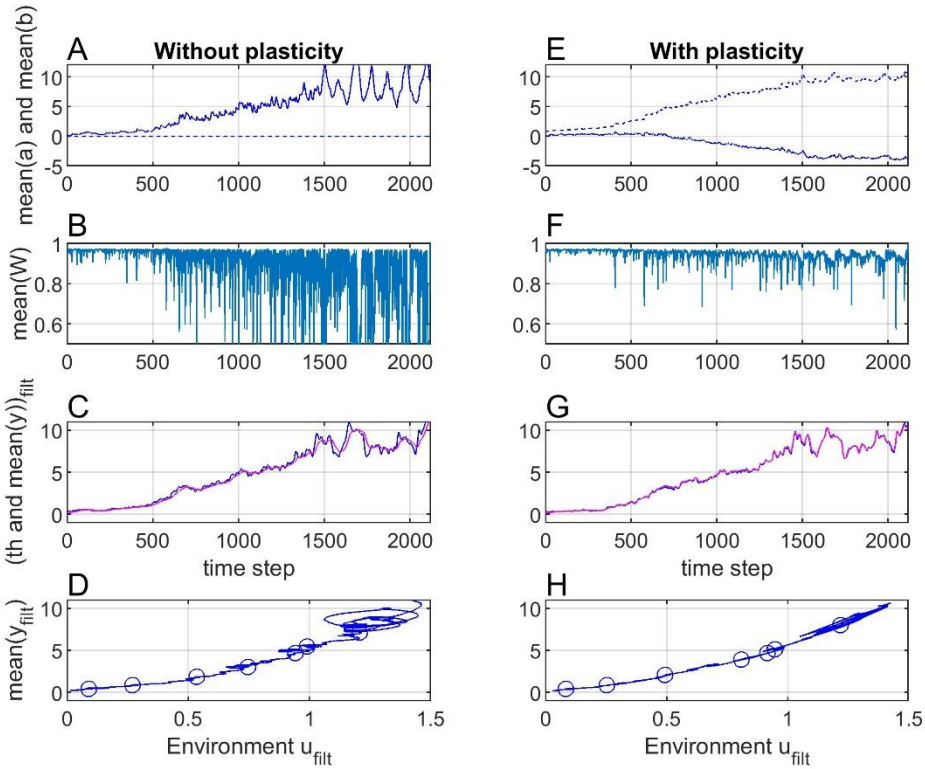

Figure 4. Simulation results as in Fig. 3, but with $\theta_m(u,t) = \theta_i(u,t) = (u(t) - 2.9) + 4(u(t) - 2.9)^2$. The discrete values in panels D and H could here be used to determine the nonlinear function $\hat{\bar{y}}_{filt}(t) = f\left(u_{filt}(t)\right)$ by means of any suitable fitting method.

3.3 Example with phenotypic constraint

With a constraint on the individual phenotypic values the distribution will no longer be normal, and the adaptive peak $\theta_m(u,t)$ will therefore deviate from the individual fitness peak such that Eq. (4b) no longer is correct. The mean trait $\bar{y}(t)$ will, however, still track the adaptive peak that maximizes mean fitness. When there for example is an upper limit for the individual phenotypic trait values as shown in Fig. 5, the $\bar{a}(t)$ and $\bar{b}(t)$ parameters will evolve such that $\bar{y}(t)$ deviates from $\theta_m(u,t)$ with results as shown in Fig. 6, panels A to D. The effect of the constraint is clearly seen in panel C in Fig. 6, in that $\bar{y}_{filt}(t)$ does not follow $\theta_{i,filt}(u,t)$, but instead follows $\theta_{m,filt}(u,t)$ (not shown) such that the mean fitness as shown in panel B is

maximized. The effect is also seen by a comparison of panels A in Fig. 6 and panel G in Fig. 3, in that $\bar{a}(t)$ evolves into a large value, such that $\bar{y}(t)$ without constraints on $y(t)$ would become larger than $\theta_m(u,t)$. Note that the mean fitness is increased compared to the results in Fig. 3, panel H, because the variance $\sigma_y^2(t)$ 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. 5 has an effect before $\bar{y}(t) \approx \theta(u,t)$ reaches the upper limit.

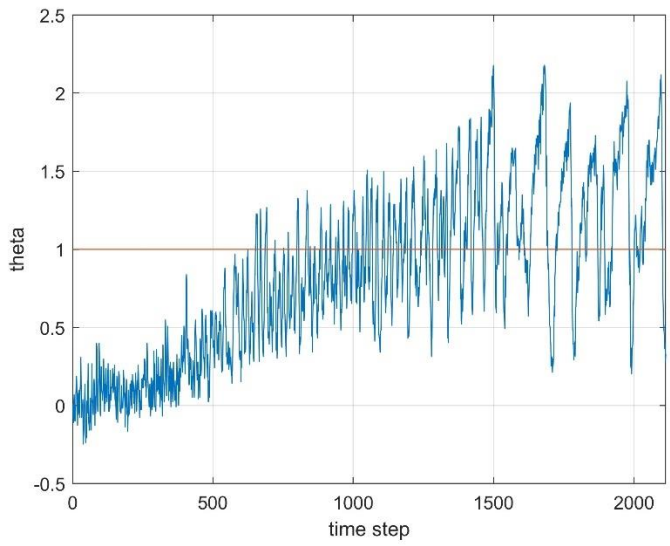

Figure 5. The linear function $\theta_m(u,t) = u(t) - 2.9$ as used for the results in Fig. 3 (blue line) and an upper constraint for the individual trait values (red line).

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

The optimal value $\theta_m(u,t)$ that maximizes mean fitness may not be a function of only the environment, and as an example Fig. 6, panels E to H, shows results as in Fig. 3, panels F to J, but with $\theta_m(u,t) = \theta_i(u,t) = u(t) - 2.9 + 2t/2115$. Note that panel H now shows a nonlinear function.

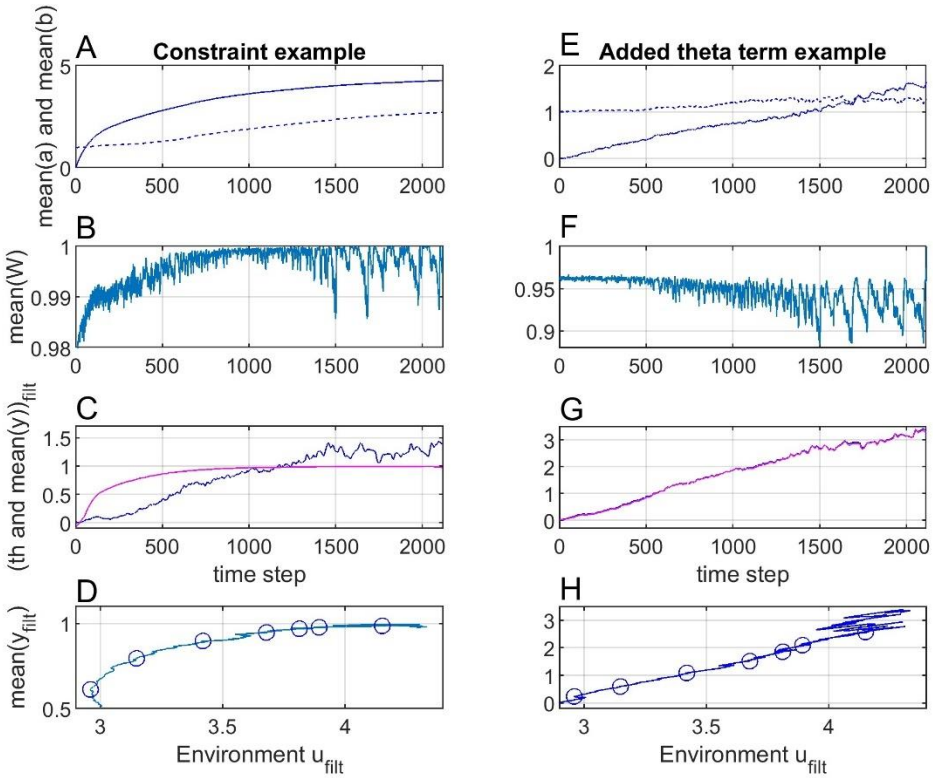

Figure 6. Panels A to D show simulation results as with plasticity in Fig. 3, but with a constraint on the individual phenotypic trait values as shown in Fig. 5. Panels E to H show simulation results as with plasticity in Fig. 3, but with $\theta_m(u,t) = \theta_i(u,t) = u(t) - 2.9 + 2t/2115$.

3.5 Example without plasticity, but with changed parameter values

Simulations as in Fig. 3, panels A to E, were repeated but with changed parameter values. First, G_{aa} and σ_ε^2 were reduced from 2 and 1 to 0.2 and 0.1, with results as in Fig. 7, panels A to E. As seen in panels A and F the mean phenotypic value $\bar{y}(t)$ cannot now follow the adaptive peak $\theta_m(u,t)$), and as a result the mean phenotype vs. environment function in panel E becomes non-linear. However, owing to the broad fitness function with $\omega^2 = 50$, the mean fitness is still close to one.

Second, with $G_{aa} = 2$ and $\sigma_\varepsilon^2 = 1$ as in Fig. 3, the width of the fitness function was reduced by reducing ω^2 from 50 to 5, with results as in Fig. 7, panels F to J. The mean phenotypic value still

follows the adaptive peak but owing to the narrow fitness function the mean fitness is clearly reduced (panel H).

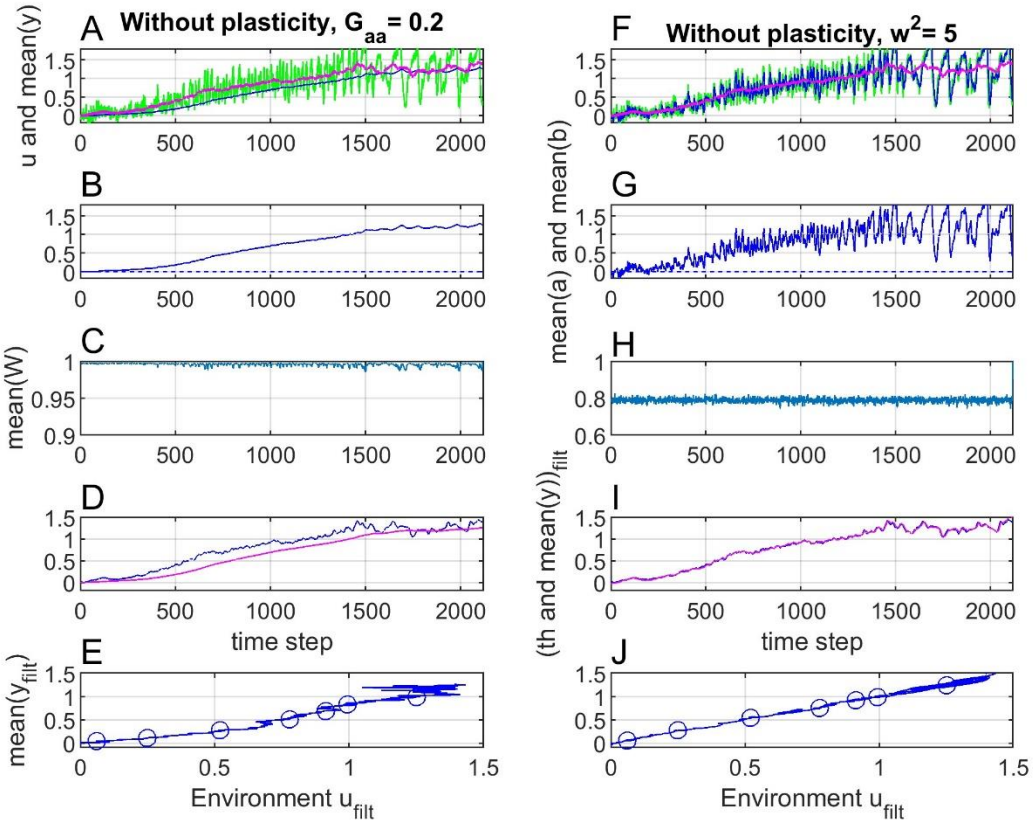

Figure 7. Panels A to E show simulation results as in Fig. 3 with plasticity, but with G_{aa} and σ_ε^2 reduced by a factor 10. Panels F to J show simulation results as in Fig. 3 with plasticity, but with the width of the fitness function reduced from $\omega^2 = 50$ to $\omega^2 = 5$.

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 (see Fig. 1 for some results). 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 retrieval by use of 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 trait values increase with time except for a drop in recent time owing to an extreme $\partial^{18}O(t)$ value.

Since the *log AZ area mean* and *log OV area mean* data in Fig. S2 are linearly dependent, I will here focus on the former. The last two data points in this time series are according to Table S1 found from fossils collected at time zero, with time span (duration) zero, but that is not consistent with Fig. 2, panel A, in Liow et al. (2024). After comparison with data from other time periods I therefore assume that the samples behind these two mean values are within time spans of approximately 0.05 and 0.02 million years. The fossil data are thus collected from eleven periods back in time according to Table 1, which for the purpose of comparison with moving average mean data also includes the time window size for 100 samples. The reason for use of the natural logarithm of the AZ area is that it gives normally distributed data (Di Martino and Liow, 2021).

Table 1. Fossil collection time data

sample number	mean sample time point	window size	window size for 100 samples
1	-2.1895	0.203	0.25
2	-1.9660	0.102	0.25
3	-1.8505	0.049	0.25
4	-0.9265	0.019	0.2
5	-0.6485	0.055	0.2
6	-0.5480	0.030	0.2
7	-0.5055	0.055	0.1
8	-0.4510	0.054	0.1
9	-0.3990	0.050	0.1
10	-0.025	0.050	0.1
11	-0.010	0.020	0.1

Since the *log AZ area mean* values are given as mean values from samples within time windows as given in Table 1, it is natural to use mean values $u_{mean}(t)$ of $\partial^{18}O$ in the same windows. Fig. 8 shows $u(t) = \partial^{18}O(t)$ and $y(t) = log\ AZ\ area\ mean\ (t)$, as well as predictions $\hat{\bar{y}}(t)$ based on least squares fitting of a straight line in a plot over $y(t)$ as function of $u_{mean}(t)$. The slope of the prediction line is 0.20, and the mean squares prediction error $MSE = \left(\bar{y}(t) - \hat{\bar{y}}(t)\right)^2/(number\ of\ samples)$ is $MSE = 0.0075$ (Table 2).

Fig. 9 shows the corresponding results based on centered moving average filtering of $\partial^{18}O(t)$. The optimal moving window size was 90, resulting in prediction line slope 0.39 and prediction error $MSE_{min} = 0.0051$. However, note that the way the $\partial^{18}O(t)$ data are organized in Lisiecki and Raymo (2005) means that a fixed window size given by number of samples corresponds to different time durations as shown in Table 1.

As seen in especially Fig. 9, it is the very last value $\bar{y}(t)$ that is difficult to predict (sample 11 in Table 1, population 19.2 in Liow et al., 2024). However, when this sample is removed the results with use of mean $u(t)$ for the various time periods (as in Fig. 8) changed very little because the last sample has little statistical leverage. The MSE value with use of moving average filtering of $u(t)$ (as in Fig. 9), on the other hand, changed a lot (Table 2) because the last sample has large statistical leverage.

A special problem with the last sample in Table 1 is that moving average filtering of $u(t)$ cannot capture the sharp decline in $u(t)$ towards time zero. If $u_{filt}(t)$ for the last 0.02 million years is replaced by the mean value of $u(t)$ in that period, i.e., a hybrid model, the last sample is predicted with the very small error 0.01, while MSE_{min} is improved from 0.0049 to 0.0022 (Table 2). Fig, 10 shows results for the hybrid model.

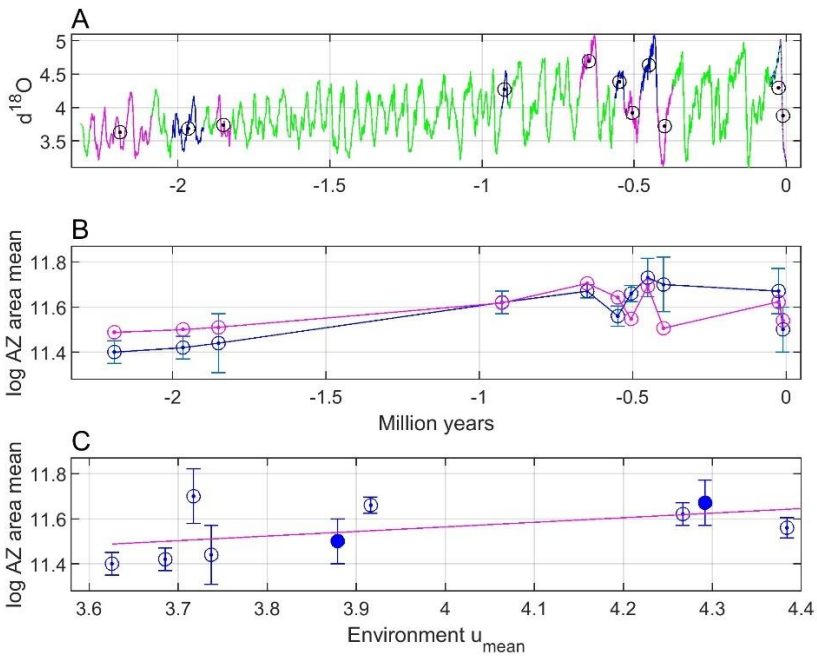

Figure 8. Panel A shows raw $\partial^{18}O(t)$ data (green line) with values in fossil collection time windows marked by red and blue, and with mean values marked by circles with dots. Panel B shows *log AZ area mean* with error bars as found from Fig. S2 in Liow at al. (2024) with corrections for the two last points as given in Table 1 (blue circles with dots and blue line), as well as predictions (red circles with dots and red line). Panel C shows

prediction line (red) computed by the ordinary least squares method, with the two last samples marked by solid circles. See prediction results in Table 2.

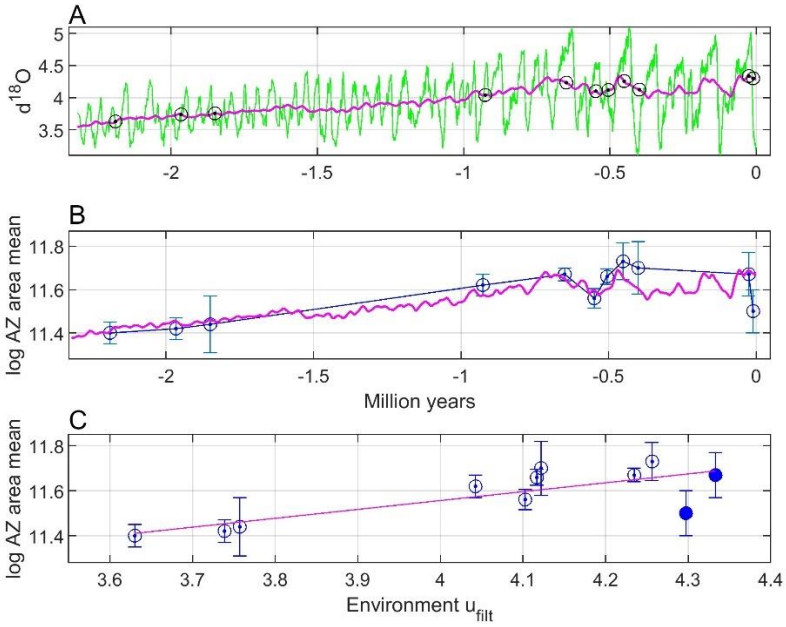

Figure 9. Results as in Fig. 8, but based on centered moving average filtering of $\partial^{18}O(t)$ and with continuous predictions of *log AZ mean area* values (red line in panel B). See prediction results in Table 2.

The prediction results for the hybrid model described above and shown in Fig. 10, i.e., with $u_{filt}(t)$ for the last 0.02 million years replaced by the mean value of $u(t)$ in that period, were validated by leave-one-out cross validation, with results as in Fig. 10 and Table 2.

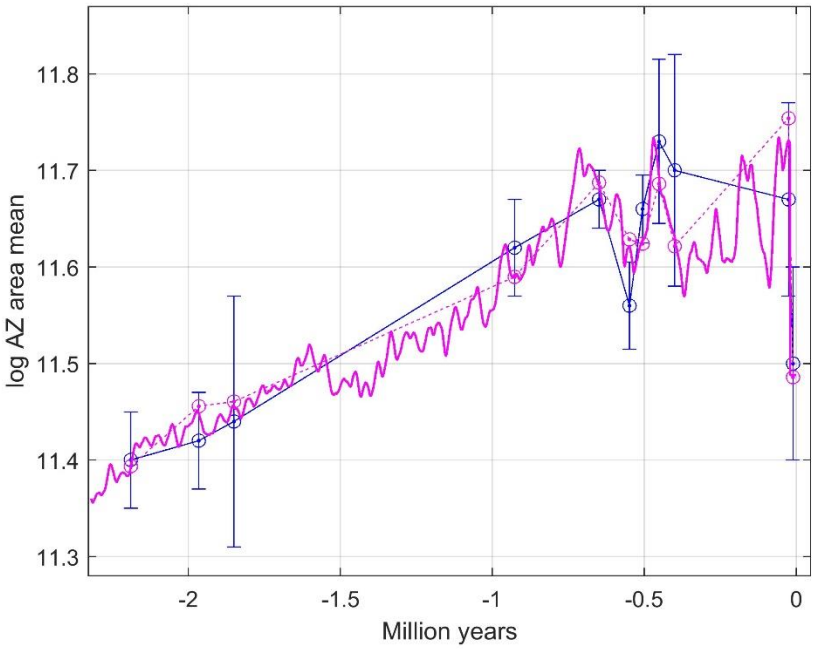

Figure 10. Results for hybrid model with $u_{filt}(t)$ for the last 0.02 million years replaced by the mean value of $u(t)$ in that period, and for leave-one-out cross validation. Solid red line shows prediction results, while dashed red line and red circled with dots show validation results for the left-out samples. See prediction and validation results in Table 2.

Table 2. Results with different prediction models.

	Model with use of u_{mean} Fig. 8	Model with use of u_{mean}, without sample 11	Model with use of u_{filt} Fig. 9	Model with use of u_{filt}, without sample 11	Hybrid Model, Fig. 10
b	0.20	0.20	0.39	0.51	0.48
MSE	0.0075	0.0081	0.0049	0.0015	0.0015
b_{CV}	-	-	-	-	0.48
MSE_{CV}	-	-	-	-	0.0022

As shown in Figs. 1, 8, 9 and 10 the measurements of *log AZ area mean* have various standard errors. Repeated simulations of the final hybrid model as used for Fig. 10, with discrete values drawn from normal distributions with standard deviations equal to the standard errors, gave results as in Fig. 11. As can be seen in the upper panel there is ample space for nonlinear

prediction functions, although they would be less likely than the fitted linear function marked by blue line. From the histogram in the lower panel follows that the prediction line slope value $b = 0.48$ has a standard error of approximately $\mathrm{SE} = 0.1$.

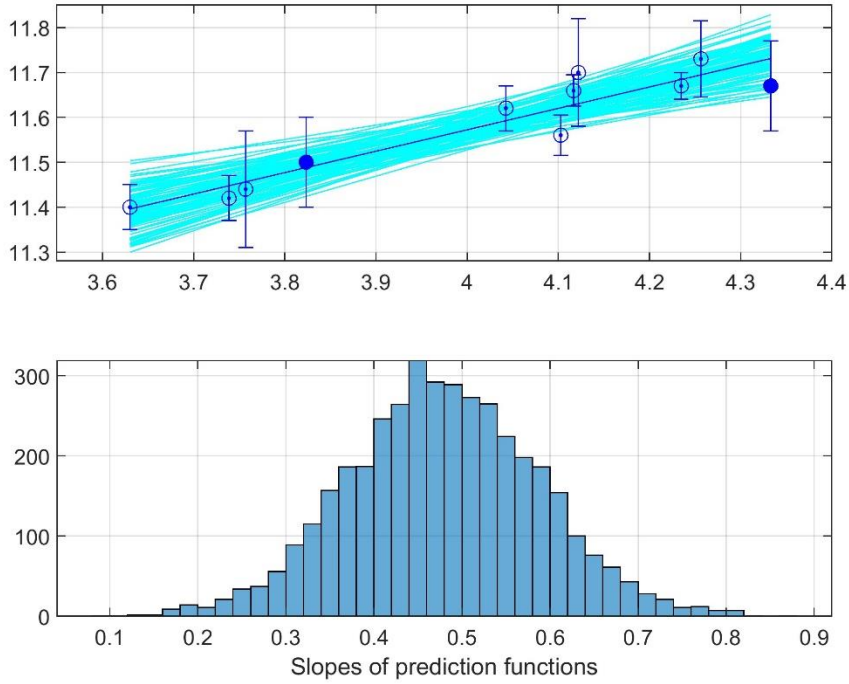

Figure 11. Upper panel shows linear prediction lines for *log AZ area mean* from 100 simulations with measurement values drawn from normal distributions of $\bar{y}(t)$ with the given standard errors (cyan lines), and for the mean prediction line corresponding to Fig. 10 (blue line with error bars). Lower panel shows histogram over slopes of fitted prediction lines as in upper panel, but for 4,000 simulations, indicating that the prediction line slope is $b = 0.48 \pm 0.1$.

5 Summary and Discussion

5.1 Main points

In this paper I make use of two landscape metaphors. First, an individual fitness landscape where the peak of the individual fitness as function of individual phenotypic values moves with

changes in the environment. Second, an adaptive landscape where the peak of mean fitness as function of mean phenotypic values moves with changes in the environment. The adaptive peak will essentially follow the individual fitness peak, but with deviations caused by non-symmetrical fitness and trait distribution functions and possibly also owing to other causes.

I also make some other fundamental assumptions. First, species that have existed for millions of years in a fluctuating environment have by means of various selection mechanisms been tracking a limited movement of an optimum in the adaptive landscape, and they have done so well enough to survive and reproduce. This assumption is supported by results in Estes and Arnold (2007) and other references in the introduction. Second, the location of the individual fitness peak is a simple continuous function of a dominating environmental variable, e.g., the temperature, such that the movement of the peak in the adaptive landscape is continuous without sudden jumps. Third, the dominating environmental variable fluctuates in a way that makes it unlikely that the mean phenotypic value gets stuck in local minima or valleys in the adaptive landscape.

Under the assumptions given, two main conclusions can be drawn from the results in this article: First, adaptive vs. environment functions, and thus approximate mean phenotype vs. environment functions can be found from sparse and short fossil data, provided that they are sufficiently spread over the time period of interest. As demonstrated in the real data example in Section 4, the environmental data used for prediction of mean phenotypic values may be means from time periods of fossil collection or a moving average mean. Second, prediction of mean phenotypic values does not require detailed knowledge of the tracking process in Fig. 2, i.e., neither parameter values in an underlying reaction norm or adaptive walk model nor the fitness functions need to be known. Also note that it is not necessary to find explicit mathematical adaptive peak vs. environment functions, predictions may instead be found by use of spline interpolation and similar methods.

Even if the position of the individual fitness peak is a linear or other simple function of the environment, the adaptive peak vs. environment function may be less simple. This is because non-symmetrical fitness functions and/or trait distributions may drive the adaptive peak away from the individual fitness peak. A closer look at the tracking process might thus be necessary for interpretations of nonlinear mean phenotype vs. environment results. A deeper understanding of the adaptive landscape may also be needed in cases with slow adaptation, where the mean phenotypic value can track the adaptive peak only with a large time lag.

5.2 Discussion of simulation results

The main purpose of the simulations is to show that the mean phenotypic values over time as functions of mean environment over time (panels E and J in Fig. 3 and panels D and H in Fig. 4) are determined by the adaptive peak vs. environment function, and thus independent of the details of the tracking process in Fig. 2. The variances of the reaction norm parameters are chosen such that there is a clear difference in the evolutionary responses between cases without and with plasticity. If these variances are increased by a factor ten (or a factor one hundred) this difference very much disappears in that $\bar{y}(t)$ in both cases appear to follow the adaptive peak $\theta_m(t) = f(u(t))$ instantaneously, but the mean phenotypic values over time as functions of mean environment over time are essentially unchanged.

Figs. 3 and 4 show that although a population without plasticity may not be able to track rapid changes in environment it may track a moving average quite well, and that the resulting mean phenotype prediction functions, linear or non-linear, will be the same as for a population with plasticity. Note that the mean plasticity slope with a linear fitness peak vs. environment function approaches an optimal value (Fig. 3, panel F), such that the linear mean phenotype vs. environment function seen in panel J can be interpreted as the result of adaptive phenotypic plasticity. However, a comparison with results without plasticity (panel E) shows that this interpretation may be wrong.

Fig. 6, left panels, illustrates that non-symmetrical trait distributions result in non-linear mean phenotype vs. environment functions also in cases where the individual fitness peak position is a linear function of a dominating environmental factor. Fig. 6, right panels, illustrates that other factors than the environment may contribute to non-linearities in the mean phenotype vs. environment function.

Panel 7, left panels, illustrates a case with slow adaptation as in, e.g., Toljagic et al. (2018), where a mean trait value lags behind changes in the adaptive peak. The result is a non-linear mean phenotype vs. environment function, but also in this case mean phenotypic values over time can be predicted from a limited number of samples. Panel 7, right panels, shows that a narrow fitness function, as must be expected results in reduced mean fitness.

5.3 Discussion of the real data case

The time series in the real data case in Section 4 has 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 adaptive peak vs. environment function. Considering the large standard errors in the *log AZ area mean* data it is in any case no clear indications of a very nonlinear function. It also seems reasonable to assume that long existing species in the quite variable $\partial^{18}O$ environment in general have adaptive peak vs. environment functions that are not too difficult to track, at least in the sense that the mean phenotypic value is fluctuating around a moving average as in Fig. 3, panel A.

The results in Fig. 9 show that the mean phenotypic values can be predicted quite well with use of a linear adaptive peak vs. environment function and moving average means of the $\partial^{18}O$ data. The exception is the very last point in the time series (population 19.2 in Liow et al., 2024), found after a minor correction of a data inconsistency in Liow et al. (2024). Note that Fig. 9 gives continuous predictions. Use of mean $\partial^{18}O$ values from the time periods of fossil collection

worked less well (Fig. 8), indicating that the raw data points in each sample period (as shown in Fig. 2 in Liow et al., 2024) should not all be given the same weight.

There is a special problem with the very last sample in Fig. 9 in that moving average filtering of $u(t)$ cannot capture the sharp decline in $u(t)$ towards time zero. If $u_{filt}(t)$ for the last 0.02 million years is replaced by the mean value of $u(t)$ in that period (a hybrid solution), the last sample is predicted with a very small error while MSE is very much improved. As shown in Fig. 10, this solution gives predictions that fall within or for one sample just outside of the error bars. The prediction results for the hybrid model were validated by leave-one-out cross validation, with results included in Fig. 10. Validation gave $MSE_{CV} = 0.0022$, which should be compared with $MSE_{min} = 0.0015$ for predictions. As shown in Fig. 10, the validation results are thus quite satisfactory in that all predictions of the left-out samples except two fall within the error bars.

Note that although the predictions in Fig. 10 are quite good as compared with the error bars there are other possible errors that are not covered by these bars, such as fluctuations around the moving adaptive peak due to imperfect tracking, and random fluctuations of the adaptive peak itself.

It is of course an underlying assumption that the continuous mean trait predictions in Fig. 10 not only give values that can be compared to mean values from collected data, but that also mean values in time periods without a known or investigated fossil record are predicted. That remains to be verified, and it would be especially interesting if the pronounced variations around 1.5 million years ago could be reproduced.

Liow et al. (2024) report 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. There is in fact a theoretical possibility that the apparently linear mean phenotype vs. environment function is caused by adaptive plasticity such that the mean phenotypic value tracks the fluctuations in sea temperature with

very small errors. However, as pointed out above, a linear mean phenotype vs. environment function does not necessarily imply plasticity.

It is finally worth noting that just by looking at Fig. 1 it might be tempting to approximate the mean phenotypic values as a linear function of the moving average of the environment. The results in Fig. 10 could thus be found by an *ad hoc* least squares approach, without any consideration of the theoretical background. However, the discussion of the tracking model in Fig. 2 shows that the prediction results have a more general theoretical foundation.

Acknowledgement

I thank two anonymous reviewers for constructive criticism, and 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 and data are archived on bioRxiv, doi.org/10.1101/2024.10.30.621046.

Conflict of interest statement

The first author of Liow et al. (2024) which is a main reference, Lee Hsiang Liow, is my daughter-in-law. Otherwise, there are no conflicts of interest.

ORCID

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

References

Åström, K. J., & Murray, R. M. (2008). Feedback systems: an introduction for scientists and

engineers. Princeton, NJ: Princeton University Press.

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

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

Estes, S. & Arnold, S. J. (2007). Resolving the paradox of stasis: models with stabilizing selection explain evolutionary divergence at all timescales. American Naturalist, 169, 227–244.

Gingerich, P. D. (2001). Rates of evolution on the time scale of the evolutionary process. Genetica 112/113:127–144.

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

Hansen, T. F. (1997). Stabilizing selection and the comparative analysis of adaptation. Evolution 51:1341–1351.

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

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

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

Pigliucci, M. (2013). Landscapes, Surfaces, and Morphospaces: What Are They Good For? 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

Uyeda, J. C., Hansen, T. F., Arnold, S. J., & Pienaar, J. (2011). The million-year wait for macroevolutionary bursts. Proceedings of the National Academy of Sciences. doi:10.5061/dryad.7d580

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

Walsh, B., & Lynch, M. (2018). Evolution and selection of quantitative traits. Oxford University Press