bioRxiv preprint doi: https://doi.org/10.1101/2022.01.07.475361; this version posted January 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

On the environmental zero-point problem in microevolutionary climate response predictions

Short Communication

Running head: On the environmental zero-point problem

Rolf Ergon

University of South-Eastern Norway

Campus Porsgrunn

Kjølnes ring 56, NO-3918 Porsgrunn, Norway

rolf.ergon@usn.no

https://orcid.org/**0000-0003-2393-2997**

Acknowledgement: The derivation of the prediction equations in the Supporting Information was inspired by a review comment by Michael Morrissey, although on a quite different manuscript with a derivation directly based on the Price equation.

The author has no conflict of interest to declare.

Word Count: Abstract 185, Main text 3996.

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.07.475361; this version posted January 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

On the environmental zero-point problem in microevolutionary climate response predictions

Abstract

It is well documented that populations adapt to climate change by means of phenotypic plasticity, but few reports on adaptation by means of genetically based microevolution caused by selection. Disentanglement of these separate effects requires that the environmental zero-point is defined, and this should not be done arbitrarily. Together with parameter values, the zero-point can be estimated from environmental, phenotypic and fitness data. A prediction error method for this purpose is described, with the feasibility shown by simulations. An estimated environmental zero-point may have large errors, especially for small populations, but may still be a better choice than use of an initial environmental value in a recorded time series, or the mean value, which is often used. Another alternative may be to use the mean value of a past and stationary stochastic environment, which the population is judged to have been fully adapted to, in the sense that the mean fitness was at a global maximum. An exception is here cases with constant phenotypic plasticity, where the microevolutionary change per generation follows directly from phenotypic and environmental data, independent of the chosen environmental zero-point.

Keywords: Climate response predictions; Environmental zero-point; Microevolution; Plasticity; Prediction error minimization

1 Introduction

Wild populations adapt to changing environments by phenotypic plasticity and microevolution, and especially climate change responses have been extensively studied. The aim is then to disentangle phenotypic changes owing to genetically based microevolution,

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.07.475361; this version posted January 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

caused by selection, and changes due to individual plasticity. Relying on 11 review articles, including reviews of altogether 66 field studies, Merilä & Hendry (2014) arrived at the conclusion that evidence for genetic adaptation to climate change has been found in some systems, but that such evidence is relatively scarce. They also concluded that more studies were needed, and that these must employ better inferential methods. The aim of the present short communication is to give a contribution in that respect.

It is obvious that for all evolutionary systems with interval-scaled environmental variables u_t, as for example temperature in ℃, a suitable zero-point u_{ref} must be chosen, and as argued in Section 2 this should not be done arbitrarily. In most cases where the environmental variable is for example a temperature, the zero-point should not, for example, be set to 0 ℃. Neither should it without further consideration be set to the initial or mean environmental value of a specific time series. It appears that the need for a proper environmental zero-point definition, and thus an environmental cue definition, has been largely ignored in the reviewed studies referred to in Merilä & Hendry (2014).

The present communication is an attempt to clarify some important questions relating to environmental zero-points, and for that purpose a method for model-based predictions of microevolutionary changes is also proposed. This method is based on parameter estimation by means of prediction error minimization, and it includes estimation of the environmental zero-point and initial mean trait values.

For a discussion of the general microevolution vs. plasticity disentanglement problem, we may for simplicity assume the two-trait individual reaction norm model

$$y_{i,t} = a_{i,t} + v_{i,t} + (b_{i,t} + \eta_{i,t})(u_t - u_{ref}), \tag{1}$$

where $u_t - u_{ref}$ and $y_{i,t}$ are the environmental cue and the individual phenotypic value, respectively, as functions of time t measured in generations. The generations are here assumed to be non-overlapping. The traits $a_{i,t}$ and $b_{i,t}$ are the additive genetic components of

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.07.475361; this version posted January 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

the individual reaction norm elevation and plasticity slope, respectively, while $v_{i,t}$ and $\eta_{i,t}$ are independent iid zero mean normal non-additive effects with variances σ_v^2 and σ_η^2, respectively. From eqn (1) follows the mean trait reaction norm model $\bar{y}_t = \bar{a}_t + \bar{b}_t(u_t - u_{ref})$.

The environmental zero-point u_{ref} is determined by the environment u_0 at which the phenotypic variance has its minimum, as defined in more detail in Section 2, and as discussed in Ergon & Ergon (2017) and Ergon (2018). In theoretical work it is often assumed that the population has fully adapted to a stationary stochastic environment with a given mean value, such that mean fitness is at a global maximum, and the environmental zero-point is then set to zero (Lande, 2009; Chevin & Lande, 2015). Although there is nothing wrong with this theoretical approach, it disguises the underlying problem discussed here, and u_{ref} is therefore included in eqn (1). This formulation also makes it possible to distinguish between the environment as such and the environmental cue. In a laboratory setting it may in some cases be possible to determine the environmental zero-point experimentally, see, e.g., Fossen *et al*. (2018), but that is obviously difficult for wild populations.

When the environmental cue $u_t - u_{ref}$ varies over time, the mean trait values $\bar{a}_t$ and $\bar{b}_t$ as follow from eqn (1) may evolve due to selection, and as a result also the mean phenotypic value $\bar{y}_t$ will evolve (Lande, 2009). Without changes due to selection, i.e., if the mean trait values $\bar{a}_t$ and $\bar{b}_t$ are constant, the value of $\bar{y}_t$ may still vary when $u_t - u_{ref}$ varies, as also follows from eqn (1).

Section 2 discusses several aspects of the general microevolution vs. plasticity disentanglement problem. First, a definition of the environmental zero-point u_{ref} is given. Second, it is shown how the mean trait values $\bar{a}_t$ and $\bar{b}_t$, and thus also $\bar{y}_t$, evolve as functions of the environmental cue $u_t - u_{ref}$ and the phenotypic selection gradient $\beta_{y,t}$. Third, it is shown how u_{ref} and $\beta_{y,t}$, as well as initial mean trait values and the parameter values in the

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.07.475361; this version posted January 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

$\boldsymbol{G}$ matrix, can be estimated by means of a prediction error minimization method (Ljung, 2002), using data from known time series of u_t and $y_{i,t}$, as well as of individual fitness values $W_{i,t}$. Forth, it is discussed why it may be difficult to estimate u_{ref}, as revealed by simulations, and which consequences errors in estimated values of u_{ref} will have. An exception is here cases with constant phenotypic plasticity, where the microevolutionary change per generation follows directly from phenotypic and environmental data, independent of the chosen environmental zero-point.

It must be underlined that the theory in Section 2 assumes that the phenotypic trait $y_{i,t}$ in eqn (1) is not correlated with other traits having causal effects on fitness, see Morrissey *et al*. (2010) for a discussion. Also note that the need for a proper environmental zero-point is not specific for the simple case according to eqn (1).

Simulations in Section 3 show that errors in the estimated or guessed value of u_{ref} may cause large mean trait prediction errors. They also show the feasibility of the proposed parameter estimation method. Finally follows a discussion in Section 4. Derivations of prediction equations, some additional simulation results, a short comparison with REML parameter estimation, and MATLAB code, are given in Supporting Information.

2 Theory and methods

2.1 Example system

For a study of the general environmental zero-point problem, and for a test of the proposed parameter estimation method, we may consider a true evolutionary system based on eqn (1),

$$\bar{y}_t = \bar{a}_t + \bar{b}_t(u_t - u_{ref}), \tag{2a}$$

$$\begin{bmatrix} \Delta\bar{a}_t \\ \Delta\bar{b}_t \end{bmatrix} = \frac{1}{\bar{W}_t}\boldsymbol{GP}^{-1}\begin{bmatrix} cov(W_{i,t}, a_{i,t} + v_{i,t}) \\ cov(W_{i,t}, b_{i,t} + \eta_{i,t}) \end{bmatrix}, \tag{2b}$$

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.07.475361; this version posted January 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

with $\boldsymbol{G} = \begin{bmatrix} G_{aa} & G_{ab} \\ G_{ab} & G_{bb} \end{bmatrix}$ and $\boldsymbol{P} = \begin{bmatrix} G_{aa} + \sigma_v^2 & G_{ab} \\ G_{ab} & G_{bb} + \sigma_\eta^2 \end{bmatrix}$. Here, eqn (2b) is the multivariate breeder's equation (Lande, 1979), where $W_{i,t}$ is found from any given fitness function. It is assumed that the phenotypic trait $y_{i,t}$ in eqn (1) is not correlated with other phenotypic traits having causal effects on fitness, and that generations are non-overlapping.

2.2 Environmental zero-point

As discussed in the Introduction, there is a need for a well-defined environmental zero-point:

Definition 1

Assuming a single environmental variable u_t, and given a reaction norm model, the environmental zero-point is

$$u_{ref} = u_0 + f(reaction\ norm\ parameter\ values), \tag{3}$$

where u_0 is the environment at which the phenotypic variance is at a minimum, and where the covariance between the plastic phenotypic value and reaction norm slope is zero. Here, $f(reaction\ norm\ parameter\ values)$ is a correction term that may be zero.

For the reaction norm model (1) we find for example (using $u' = u - u_{ref}$)

$$cov(y, b) = E\left[\left(a - \bar{a} + v + \left(b - \bar{b}\right)u' + \eta u'\right)\left(b - \bar{b} + \eta\right)\right] = G_{ab} + \left(G_{bb} + \sigma_\eta^2\right)u', \tag{4a}$$

which by setting $cov(y, b) = 0$ and $u' = u_0 - u_{ref}$ gives the environmental zero-point

$$u_{ref} = u_0 + \frac{G_{ab}}{G_{bb} + \sigma_\eta^2}. \tag{4b}$$

For $G_{ab} = 0$ the environmental zero-point is thus the environment where the phenotypic variance is minimized (see Fig. 1 for illustration). This is also the environment where the mean fitness has a global maximum, and thus the environment the population is fully adapted to. In this environment the environmental cue will be zero.

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.07.475361; this version posted January 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

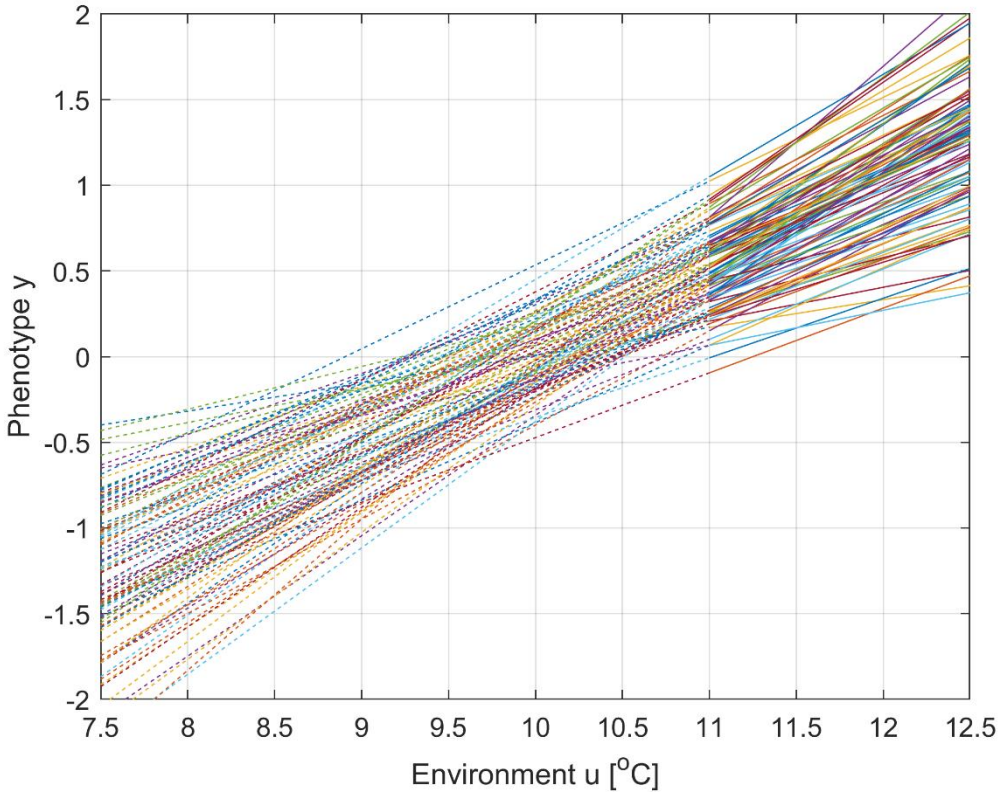


Fig. 1 Reaction norms for 100 individuals in a population according to eqn (1), with $G_{aa} + \sigma_v^2 = 0.05$, $G_{bb} + \sigma_\eta^2 = 0.02$ and $G_{ab} = 0$. The environmental zero-point is $u_{ref} = 10$ °C, which since $G_{ab} = 0$ also is the temperature u_0 to which the population is fully adapted. The mean trait values are $\bar{a}_t = 0$ and $\bar{b}_t = 0.5$. Solid lines indicate the range of environmental data used for parameter estimation and mean trait predictions in simulations. Note that $u_{ref} = u_0 = 10$ °C is not within that range.

2.3 Mean trait prediction equations

A fundamental equation for mean trait predictions follows from eqn (2a) as

$$\Delta\bar{y}_t = \Delta\bar{a}_t + \Delta\bar{b}_t\left(u_{t+1} - u_{ref}\right) + \bar{b}_t\Delta u_t, \tag{5}$$

where $\Delta u_t = u_{t+1} - u_t$, $\Delta\bar{a}_t = \bar{a}_{t+1} - \bar{a}_t$, $\Delta\bar{b}_t = \bar{b}_{t+1} - \bar{b}_t$ and $\Delta\bar{y}_t = \bar{y}_{t+1} - \bar{y}_t$ are changes per generation. From this follows that the value of u_{ref} has nothing to say in special cases

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.07.475361; this version posted January 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

with constant phenotypic plasticity slopes, i.e., when $\Delta\bar{b}_t = 0$. In such cases we simply have $\Delta\bar{y}_t = \Delta\bar{a}_t + \bar{b}\Delta u_t$, where $\bar{b}$ is constant, or only $\Delta\bar{y}_t = \bar{b}\Delta u_t$ if $\bar{a}_t$ does not evolve.

As shown in Appendix S1 in Supporting Information, eqn (5) leads to equations for $\Delta\bar{a}_t$ and $\Delta\bar{b}_t$ as functions of the phenotypic selection gradient $\beta_{y,t}$,

$$\Delta\bar{a}_t = \left(G_{aa} + G_{ab}\left(u_t - u_{ref}\right)\right)\beta_{y,t} \tag{6a}$$

and

$$\Delta\bar{b}_t = \left(G_{ab} + G_{bb}\left(u_t - u_{ref}\right)\right)\beta_{y,t}, \tag{6b}$$

where

$$\beta_{y,t} = \frac{1}{\bar{W}_t}\left(P_{aa} + 2G_{ab}\left(u_t - u_{ref}\right) + P_{bb}\left(u_t - u_{ref}\right)^2\right)^{-1} cov\left(W_{i,t}, y_{i,t}\right). \tag{6c}$$

In addition to time series of u_t and $y_{i,t}$, we thus need parameter values for u_{ref}, G_{aa}, G_{ab}, G_{bb}, σ_v^2 and σ_η^2, and a time series of individual fitness values $W_{i,t}$. For mean trait predictions we also need initial values. Note that these equations are valid only when the relationship matrix is a unity matrix (Lynch & Walsh, 1998).

2.4 Prediction error minimization method

From the prediction equations (6a,b) follows predicted values of $\bar{y}_t$ from eqn (2a). The prediction equations can thus be used for parameter estimation in a prediction error minimization method (PEM), as shown in Fig. 2. As follows from eqns (6a,b,c) we can then set G_{aa} to any value, and estimate G_{ab}, G_{bb}, σ_v^2 and σ_η^2 relative to that value.

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.07.475361; this version posted January 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

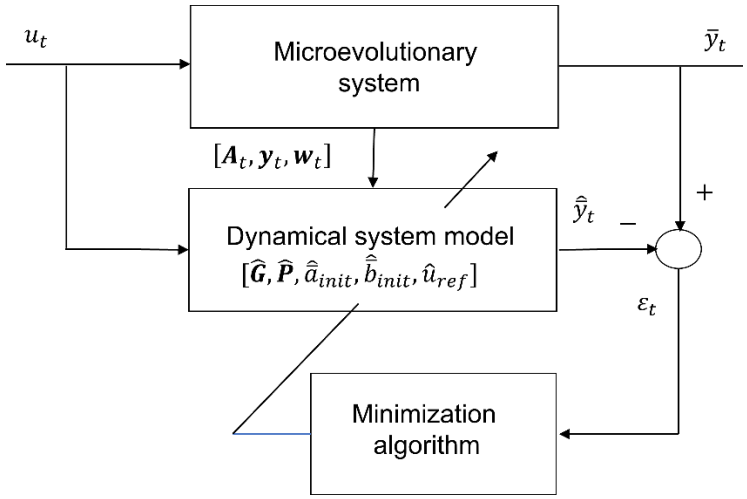


Figure 2. Block diagram of microevolutionary PEM, with dynamical tuning model based on a reaction norm model with mean traits $\bar{a}_t$ and $\bar{b}_t$. Here, u_t and $\bar{y}_t$ are the known environmental input and the known mean phenotypic value at time t, respectively. $\boldsymbol{A}_t$ is the additive genetic relationship matrix, which here is assumed to be $\boldsymbol{A}_t = \boldsymbol{I}$, while $\boldsymbol{y}_t$ and $\boldsymbol{w}_t$ are vectors of individual phenotypic and relative fitness values, respectively. The $\widehat{\boldsymbol{G}}$ and $\widehat{\boldsymbol{P}}$ matrices include the system parameters, while $\hat{\bar{a}}_{init}$, $\hat{\bar{b}}_{init}$ and $\hat{u}_{ref}$ are the initial mean trait values and the environmental zero-point, respectively. Assuming data over T generations, all these model parameters are tuned until $\sum_{t=1}^{T} \varepsilon_t^2 = \sum_{t=1}^{T} (\bar{y}_t - \hat{\bar{y}}_t)^2$ is minimized, with $\bar{y}_t = \hat{\bar{y}}_1 = 0$ and $\hat{\bar{a}}_1 = -\hat{\bar{b}}_1 (u_1 - \hat{u}_{ref})$.

2.5 Effects of environmental zero-point errors

With an environmental zero-point $\hat{u}_{ref}$ instead of u_{ref}, predictions based on eqn (2a) can be written

$$\hat{\bar{y}}_t = \hat{\bar{a}}_t + \hat{\bar{b}}_t(u_t - \hat{u}_{ref}) = \hat{\bar{a}}_t - \hat{\bar{b}}_t(\hat{u}_{ref} - u_{ref}) + \hat{\bar{b}}_t(u_t - u_{ref}), \tag{7}$$

where $\hat{\bar{a}}_t$ and $\hat{\bar{b}}_t$ are found from eqns (6a,b) with use of estimated parameter values.

For small values of G_{bb}, i.e., when $G_{bb} \to 0$ and $G_{ab} \to 0$, it follows from eqns (6a,b,c) that $\Delta\bar{a}_t$ is independent of u_{ref}, and that $\bar{b}_t$ is constant. This results in $\Delta\hat{\bar{a}}_t = G_{aa}/(G_{aa} + \hat{\sigma}_v^2)$,

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.07.475361; this version posted January 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

such that only $\hat{\sigma}_v^2$ must be tuned in order to minimize $\sum_{t=1}^{T}(\bar{y}_t - \hat{\bar{y}}_t)^2$. In this case an error in $\hat{u}_{ref}$ has very little effect on the change in $\hat{\bar{a}}_t$ over many generations (see Appendix S2 in Supporting information). This also follows from eqn (5).

For larger values of G_{bb}, $\Delta\hat{\bar{a}}_t$ will be affected by an error in $\hat{u}_{ref}$, and good predictions $\hat{\bar{y}}_t \approx \bar{y}_t$ for $t = 1$ to T can then only be obtained by parameter tuning such that $\hat{\bar{b}}_t \approx \bar{b}_t$ over all generations. That is possible because u_{ref} appears in both nominator and denominator of eqn (6b). According to eqn (7) we then find $\hat{\bar{a}}_t \approx \bar{a}_t + \hat{\bar{b}}_t(\hat{u}_{ref} - u_{ref})$, which as shown in Section 3 may result in large errors in predicted changes of $\bar{a}_t$ over time.

2.6 Effects of modeling errors

Modeling errors will obviously affect predictions of the mean traits. As an example, simulations with the true individual model

$$y_{i,t} = a_{i,t} + v_{i,t} + (b_{i,t} + \eta_{i,t})\big(u_t - c_{i,t} - \gamma_{i,t}\big), \tag{8}$$

are included in Appendix S3 in Supporting Information. Here, $c_{i,t}$ is a perception trait, as discussed in Ergon & Ergon (2017).

3 Simulation results

3.1 True model, fitness function, and environmental input signals

Assume that what we consider to be true mean responses, $\bar{y}_t$, $\bar{a}_t$ and $\bar{b}_t$, are generated by the state-space model (2a,b). Here, $G_{ab} = 0$ in the true system, but left as a free parameter in the tuning model in Fig. 2. The individual effects $a_{i,t}$, $b_{i,t}$, $v_{i,t}$ and $\eta_{i,t}$ are at each generation drawn from populations with normal distributions around $\bar{a}_t$, $\bar{b}_t$, 0 and 0, respectively.

The individual fitness function is assumed to be

$$W_{i,t} = exp\left(-\left(y_{i,t} - \theta_t\right)^2/2\omega^2\right), \tag{9}$$

where θ_t is the phenotypic value that maximizes fitness, while $\omega^2 = 10$.

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.07.475361; this version posted January 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Also assume a stationary or slowly varying mean $\mu_{U,t}$ of a stochastic environment, with added iid zero mean normal random variations $u_{n,t}$ with variance $\sigma_{U_n}^2$, i.e., $u_{\mathrm{t}} = \mu_{U,t} + u_{n,t}$, and that the population is fully adapted to a stationary environment with $\mu_{U,t} = u_{ref} = u_0 = 10$ °C (as in Fig. 1). In a corresponding way assume that $\theta_t = \mu_{\Theta,t} + \theta_{n,t}$, where $\theta_{n,t}$ is iid zero mean normal with variance $\sigma_{\Theta_n}^2$, and where $u_{n,t}$ and $\theta_{n,t}$ are correlated with covariance $\sigma_{\Theta_n U_n}$. Following Lande (2009), we may assume that juveniles of generation t are exposed to the environment $u_{t-\tau}$ during a critical period of development a fraction of a generation before the adult phenotype is expressed and subjected to natural selection. We will define $\theta_t = -2(u_t - 10)$, which implies a linear relationship $\mu_{\Theta,t} = -2(\mu_{U,t} - 10)$, variances $\sigma_{\Theta_n}^2 = 4\sigma_{U_n}^2$, and covariance $\sigma_{\Theta_n U_n} = -2\rho_\tau \sigma_{U_n}^2$, where ρ_τ is the autocorrelation of background environmental fluctuations. We will assume $\sigma_{U_n}^2 = 0.5$ and $\rho_\tau = 0.25$. The optimal value of the mean plasticity slope in a stationary stochastic environment is then $\bar{b}_{opt} = \sigma_{\Theta_n U_n}/\sigma_{U_n}^2 = -2\rho_\tau = -0.5$ (Ergon & Ergon, 2017).

Further assume that u_t and θ_t are noisy ramp functions as shown in Fig. 3, with $\mu_{U,t}$ starting from 10 °C at $t = 10$ generations. The choice of a negative trend in θ_t and thus in $\bar{y}_t$ is inspired by common cases where a positive temperature trend leads to earlier breeding dates for various natural populations.

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.07.475361; this version posted January 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

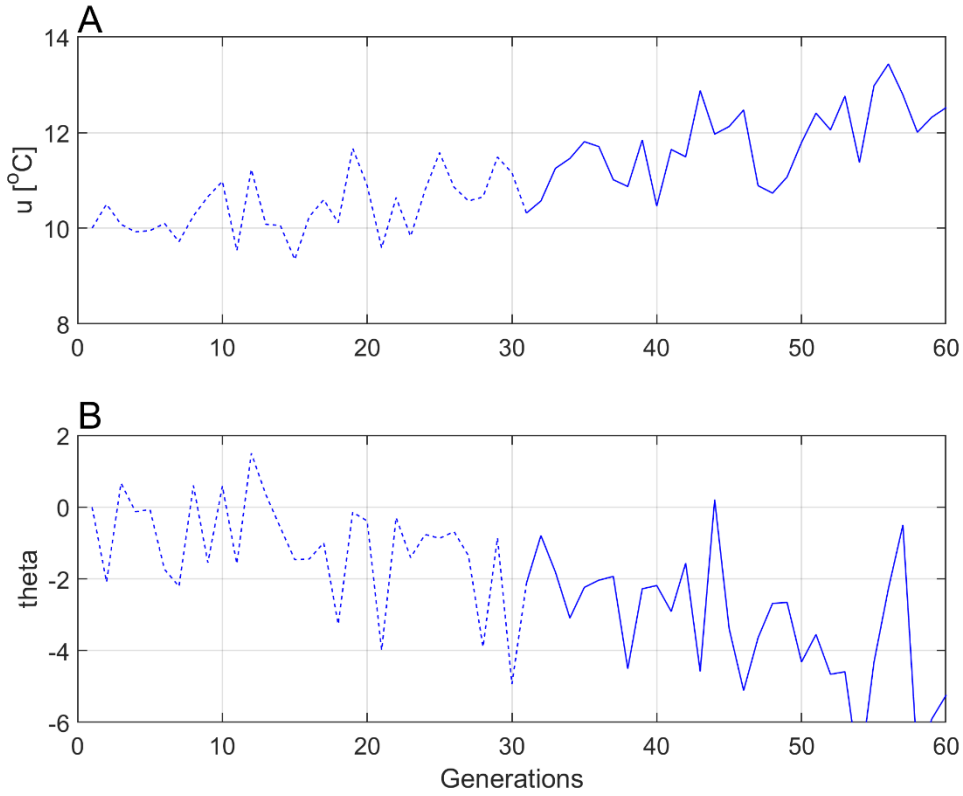


Figure 3. Noisy ramp functions u_t (panel A) and θ_t (panel B), with $u_{ref} = 10$, $\mu_{\Theta,t} = -2(\mu_{U,t} - 10)$, $\sigma^2_{\mathrm{U}_n} = 0.5$, $\sigma^2_{\Theta_n} = 2$ and $\sigma_{\mathrm{U}_n,\Theta_n} = -0.25$. The dashed parts of the curves indicate data that are not used for parameter estimation and mean trait predictions (compare with Fig. 1).

3.2 Parameter estimation and mean trait prediction results

Parameter estimation and mean trait prediction results were found by use of the MATLAB function *fmincon* in the PEM method in Fig. 2. Results with use of input-output data from $t = 31$ to 60 are given in Table 1. The relative errors in total change of predictions over 30 generations are included, computed as $\Delta_{30}^{error}\hat{\bar{a}}_t\% = 100(\Delta_{30}\hat{\bar{a}}_t - \Delta_{30}\bar{a}_t)/\Delta_{30}\bar{a}_t$ etc., where $\Delta_{30}\hat{\bar{a}}_t = \hat{\bar{a}}_{60} - \hat{\bar{a}}_{31}$ and $\Delta_{30}\bar{a}_t = \bar{a}_{60} - \bar{a}_{31}$. The final values $\sum\varepsilon^2_{t,final}$ of $\sum_{t=31}^{60}(\bar{y}_t - \hat{\bar{y}}_t)^2$ are also included, as they indicate the degree of optimization success. Results are presented as mean values and standard errors, $Mean \pm SE$, based on 100 repeated simulations with different realizations of random inputs.

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.07.475361; this version posted January 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Given the model in eqns (2a,b) and (9), there are six parameter values to be estimated (while $\hat{\bar{a}}_{31}$ follows from eqn (2a) with $\hat{\bar{y}}_{31}$ sat to zero). In the optimizations the initially guessed values of $\hat{G}_{bb}$, $\hat{G}_{ab}$, $\hat{\sigma}_v^2$, $\hat{\sigma}_\eta^2$ and $\hat{\bar{b}}_{31}$ were set to zero, while the initial value of $\hat{u}_{ref}$ was set to 10 (when $\hat{u}_{ref}$ was a free variable). The true value $\hat{G}_{aa} = 0.025$ was used, such that estimates of G_{bb}, G_{ab}, σ_v^2 and σ_η^2 are found relative to $G_{aa} = 0.025$. Table 1 presents results for three cases, first for $\hat{u}_{ref} = u_{ref} = 10$ (Case 1), second for $\hat{u}_{ref}$ as free variable (Case 2), and third for $\hat{u}_{ref} = 11$ (Case 3), which approximately is the initial value in the time series used. Note the fairly good estimates of G_{bb}, G_{ab}, σ_v^2 and σ_η^2 in Case 1. In Case 2, the estimates of these parameters have larger standard errors, and as a result also $\Delta_{30}^{error}\hat{\bar{a}}_t\%$ has a large standard error. With $\hat{u}_{ref} = 11$ (Case 3), $\Delta_{30}^{error}\hat{\bar{a}}_t\%$ has a large bias error. In this case also the estimates of G_{bb}, G_{ab} and σ_η^2 are biased. In all cases $\Delta_{30}^{error}\hat{\bar{b}}_t\%$ is close to the result with the correct value $\hat{u}_{ref} = 10$ (Case 1), as explained in Subsection 2.5.

Table 1 includes theoretical prediction error results based on eqn (7), $\hat{\bar{a}}_{t,corr} = \hat{\bar{a}}_t - \hat{\bar{b}}_t(\hat{u}_{ref} - u_{ref})$. These are in all cases close to the results with $\hat{u}_{ref} = u_{ref} = 10$ (Case 1).

Table 1. Estimation and prediction results with true system responses generated by means of eqns (2a,b) and (9). Results are for cases with population size $N = 100$ and perfect measurements $\boldsymbol{y}_t$ and $\boldsymbol{w}_t$, and they are based on 100 simulations with different realizations of all random input variables.

Case 1: $\hat{u}_{ref} = 10$ (the true value).

Case 2: $\hat{u}_{ref}$ as a free variable.

Case 3: $\hat{u}_{ref} = 11$ (approximate initial value in optimization data). Here, 9% of the simulations were discarded because $\sum\varepsilon_{t,final}^2 > 0.001$ (typically 40% with 10% measurement noise).

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.07.475361; this version posted January 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Parameter etc.	True value	Optimization results Case 1	Optimization results Case 2	Optimization results Case 3
$\hat{G}_{bb}$	0.01	0.0102 ± 0.0006	0.0102 ± 0.0029	0.0064 ± 0.0019
$\hat{G}_{ab}$	0	0.0005 ± 0.0037	0.0007 ± 0.0057	0.0069 ± 0.0018
$\hat{\sigma}_v^2$	0.025	0.0263 ± 0.0092	0.0285 ± 0.0254	0.0255 ± 0.0038
$\hat{\sigma}_\eta^2$	0.01	0.0104 ± 0.0019	0.0105 ± 0.0043	0.0164 ± 0.0038
$\hat{\bar{b}}_{31}$	–	-0.5028 ± 0.0093	-0.5036 ± 0.0069	-0.5033 ± 0.0081
$\hat{u}_{ref}$	10	10	-0.0525 ± 0.2854	11
$\sum \varepsilon_{t,final}^2$	–	$10^{-5}(1 \pm 3)$	$10^{-5}(1 \pm 3)$	$10^{-5}(10 \pm 13)$
$\Delta_{30}^{error} \hat{\bar{a}}_t \%$	–	0 ± 2	-3 ± 18	67 ± 10
$\Delta_{30}^{error} \hat{\bar{b}}_t \%$	–	0 ± 1	0 ± 2	4 ± 4
$\Delta_{30,corr}^{error} \hat{\bar{a}}_t \%$	–	0 ± 2	0 ± 2	-3 ± 3

The results in Table 1 will be altered with introduced measurement noise, and when the population size is increased. For a test, 10% random iid zero mean normal errors with standard deviations 0.03 and 0.1 were added to all values of $y_{i,t}$ and $W_{i,t}$, respectively. This is based on the facts that $y_{i,t}$ varies approximately ± 0.3 around the mean values $\bar{y}_t$, and that the maximum value of $W_{i,t}$ is one. Results for different combinations of measurement noise and population size are given in Table 2. Note that with the true value $\hat{u}_{ref} = 10$ (Cases 1 and 4), the standard errors of the predictions increase substantially when measurement noise is introduced, but that these errors are markedly reduced when the population size is increased from 100 to 1,000. With $\hat{u}_{ref}$ as a free variable (Cases 2 and 5), the results for $\Delta_{30}\bar{b}_t$ are very

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.07.475361; this version posted January 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

much the same as in Case 1, while the results for $\Delta_{30}\bar{a}_t$ are not improved with a larger population size.

Table 2. Mean trait prediction results with true system responses generated by means of eqns (2a,b) and (9), for two of the cases in Table 1, but with 10% random measurement errors in the individual phenotypic and fitness values.

Case 4: $\hat{u}_{ref} = 10$ (the true value).

Case 5: $\hat{u}_{ref}$ as a free variable.

Para-meter	Population size	Optimization results Case 4	Optimization results Case 5
$\Delta_{30}^{error}\hat{\bar{a}}_t\%$	100	-1 ± 8	0 ± 43
$\Delta_{30}^{error}\hat{\bar{b}}_t\%$	100	1 ± 6	1 ± 6
$\Delta_{30,corr}^{error}\hat{\bar{a}}_t\%$	100	-1 ± 8	-1 ± 7
$\Delta_{30}^{error}\hat{\bar{a}}_t\%$	1,000	0 ± 4	1 ± 39
$\Delta_{30}^{error}\hat{\bar{b}}_t\%$	1,000	0 ± 3	0 ± 3
$\Delta_{30,corr}^{error}\hat{\bar{a}}_t\%$	1,000	0 ± 4	1 ± 3

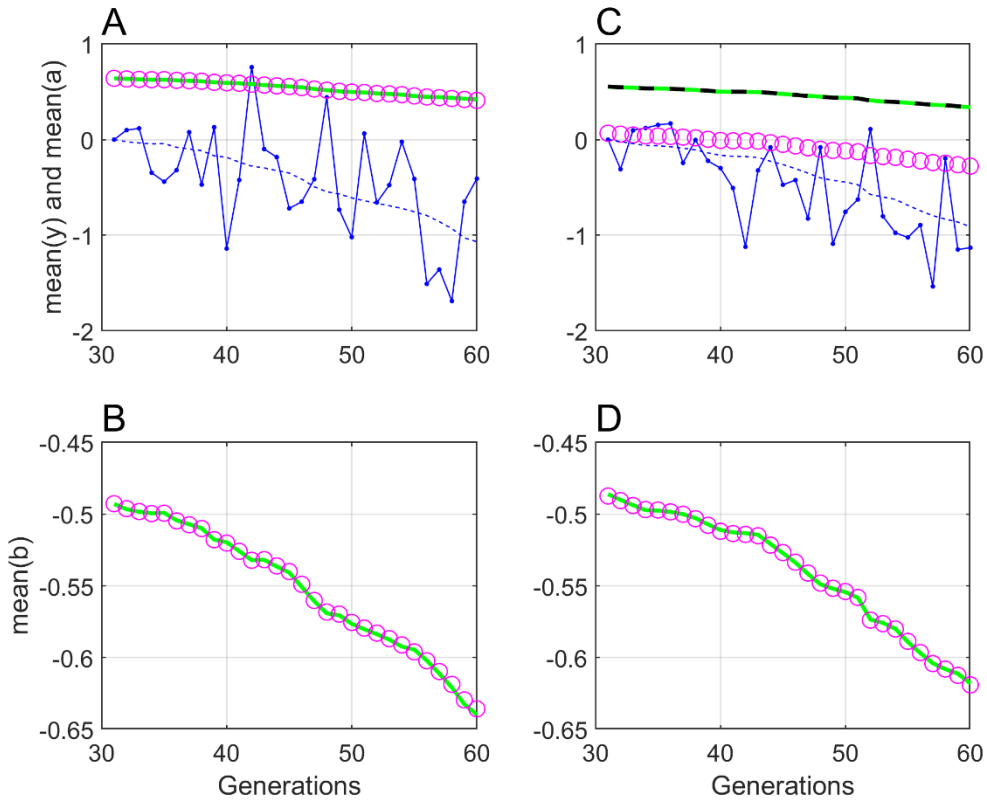
As shown in Tables 1 and 2, large errors in the assumed environmental zero-point $\hat{u}_{ref}$ result in large errors in predicted changes of in $\bar{a}_t$ over 30 generations (Case 3). Table 3 shows these errors for more moderate errors in $\hat{u}_{ref}$.

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.07.475361; this version posted January 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Table 3. Errors in predicted total relative change of $\bar{a}_t$ and $\bar{b}_t$ over 30 generations, without measurement noise, as functions of the environmental zero-point $\hat{u}_{ref}$ used in the optimization procedure.

$\hat{u}_{ref}$	$\Delta_{30}^{error}\hat{\bar{a}}_t\%$	$\Delta_{30}^{error}\hat{\bar{b}}_t\%$
-9.75	-16 ± 2	0 ± 1
10	0 ± 2	0 ± 1
10.25	17 ± 3	0 ± 1
10.5	31 ± 10	1 ± 1
11	67 ± 10	-3 ± 3

Fig. 4 shows predicted mean values $\hat{\bar{y}}_t$, $\hat{\bar{a}}_t$ and $\hat{\bar{b}}_t$, as compared to true mean values $\bar{y}_t$, $\bar{a}_t$ and $\bar{b}_t$, for Case 1 and Case 3 in Table 1.



bioRxiv preprint doi: https://doi.org/10.1101/2022.01.07.475361; this version posted January 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figure 4. Typical responses for Case 1 and Case 3 in Table 1. True $\bar{y}_t$ values are shown by solid blue lines. All parameter values except $\hat{G}_{aa} = 0.025$ were initially set to zero, which gave predictions $\hat{\bar{y}}_{t,start} = cov(W_{i,t}, y_{i,t})/\bar{W}_t$ as shown by dashed blue lines. Final predictions $\hat{\bar{y}}_t$ are shown by blue dots. True $\bar{a}_t$ and $\bar{b}_t$ responses are shown by green lines, while predictions $\hat{\bar{a}}_t$ and $\hat{\bar{b}}_t$ are shown by magenta circles. Panels A and B show results for Case 1 with $\hat{u}_{ref} = 10$ (true value), while panels C and D show results for Case 3 with $\hat{u}_{ref} = 11$. Here, the theoretical predictions $\hat{\bar{a}}_{t,corr} = \hat{\bar{a}}_t - \hat{\bar{b}}_t(\hat{u}_{ref} - u_{ref})$ are included as black dashed line. Note that $\bar{y}_{31} = \hat{\bar{y}}_{31}$ is set to zero, such that $\bar{a}_{31} = -\bar{b}_{31}u_{31}$ and $\hat{\bar{a}}_{31} = -\hat{\bar{b}}_{31}(u_{31} - \hat{u}_{ref})$, where u_{31} is not quite the same from simulation to simulation.

4 Discussion

It is well documented that populations adapt to climate change by means of plasticity, but few reports on adaptation by means of genetically based microevolution caused by selection (Merilä & Hendry, 2014). The main point in this communication is that disentanglement of these separate effects requires that the environmental zero-point u_{ref} is defined, and that this should not be done arbitrarily. Instead, it should be based on the environment u_0 where the phenotypic variance is at a minimum (Definition 1 and Fig. 1). This definition can be extended to multivariate cases. Another main point is that errors in u_{ref} may lead to large errors in predicted microevolutionary changes over time (Table 1 and Fig. 4).

In theoretical studies it is often assumed that the environmental variable is scaled such that $u_{ref} = u_0 = 0$ (Lande, 2009; Chevin & Lande, 2015). This can be done also in databased applications, provided that u_0 is known, and that the correction term in Definition 1 is zero.

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.07.475361; this version posted January 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

By use of a simple example it is verified that the environmental zero-point together with initial mean trait and parameter values can be estimated from environmental, phenotypic and fitness data, by use of the prediction error minimization method in Fig. 2. The simulations make use of an environmental trend, as in noisy temperature trends caused by climate change (Fig. 3), and a correct environmental zero-point then results in quite good parameter estimates and predicted changes in mean traits over time (Table 1, Case 1). The environmental zero-point can also be estimated, but with large standard errors, especially for small population sizes, and this results in a correspondingly large standard error in predicted change in mean elevation $\bar{a}_t$ over time (Table 1, Case 2). An estimated zero-point may still be a better choice than use of an initial environmental value in a recorded time series, or the mean value, which may give large bias errors in predicted changes in mean traits (Table 1, Case 3). Another alternative may be to use the mean value of a past stationary stochastic environment, which the population is judged to have been fully adapted to.

The standard errors in the predicted mean trait changes over time increase with random errors in the individual phenotypic and fitness data (Table 2). This is somewhat compensated by an increase in population size, especially with a correct value of the environmental zero-point.

It is here assumed that the genetic relationship matrix is an identity matrix, and the simulation results are obtained by use of a prediction error minimization method. However, the fact that errors in the environmental zero-point may cause large errors in predictions of microevolution, as discussed in Subsection 2.5, is a generic problem. Independent of prediction method and the complexity of the model, an error in the environmental zero-point implies that an erroneous model is fitted to the input-output data, and that must inevitably result in prediction errors. There is therefore no reason to believe that such errors will disappear in cases where the genetic relationship matrix is not a unity matrix, or when other

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.07.475361; this version posted January 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

parameter estimation and mean trait prediction methods are used. A more specific argument regarding REML parameter estimation is given in Appendix S4 in Supporting Information. It must thus be expected that predictions of microevolutionary changes over time depend on the chosen environmental zero-point, and such predictions cannot therefore be trusted unless the chosen environmental zero-point can be trusted. An exception is here cases with a constant mean plasticity slope, where the change in mean reaction norm elevation per generation according to eqn (5) is independent of the environmental zero-point. This implies that a nearly constant mean plasticity slope must be expected to give small errors in the predicted changes of the mean elevation, also if there is an error in the environmental zero-point.

References

Chevin, L.-M., & Lande, R. (2015). Evolution of environmental cues for phenotypic plasticity. Evolution, 46(2), 390–411. doi:10.1111/evo.12755

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

Ergon, R. (2018). The environmental zero-point problem in evolutionary reaction norm modeling. Ecol Evol. 8, 4031–4041. doi: 10.1002/ece3.3929

Fossen, E.I.F., Pélabon, C. & Einum, S. (2018). An empirical test for a zone of canalization in thermal reaction norms J. Evol. Biol. 31, 936-943. doi: 10.1111/jeb.13287.

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

Ljung, L. (2002). Prediction Error Estimation Methods. Circuits, Systems, and

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.07.475361; this version posted January 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Signal Processing 21/1, 11-21.

Lynch, M., & Walsh, B. (1998). Genetics and Analysis of Quantitatve Traits. Sinauer Associates, Mass.

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

Morrissey, M.B., Kruuk, L.E.B, & Wilson, A.J. (2010) The danger of applying the breeder's equation in observational studies of natural populations. J. Evol. Biol. 23, 2277-2288. doi: 0.1111/j.1420-9101.2010.02084.x