Phylogenetic Ornstein-Uhlenbeck regression curves

Dwueng-Chwuan Jhwueng
National Institute for Mathematical and Biological Synthesis
The University of Tennessee
Knoxville, TN 37996
e-mail: djhwueng@nimbios.org

and

Vasileios Maroulas
Department of Mathematics
The University of Tennessee
Knoxville, TN 37966
e-mail: maroulas@math.utk.edu

Abstract: A novel method is developed to jointly estimate regression curves applied to the evolutionary biology for studying the trait relationships. The adaptive evolution model is built on a coupled system of Ornstein-Uhlenbeck processes. Our method is then applied to a set of ecological data and it is compared with the recent regression method established in [9].

AMS 2000 subject classifications: Primary 60H30, 62J12; secondary 62P10.

Keywords and phrases: Coupled Ornstein-Uhlenbeck processes, generalized linear models: evolutionary regression analysis, phylogenetic comparative methods.

1. Introduction

Phylogenetic comparative methods are statistical methods for analyzing data of groups of related species, called comparative data in the ecology and evolution literature. Since the species are related by shared evolutionary history, it may not be reasonable to view such data as independent, identically distributed realizations of the same stochastic process. Instead, information about the shared evolutionary history, explained by the phylogeny for the species, is often incorporated into the analysis. The importance of incorporating the phylogenetic tree has been validated through various studies. For a partial list of such studies the reader may refer to [2, 3, 5, 8, 9, 11, 20] and references therein.

While the phylogenetic tree describes the evolutionary relationship, the trait evolution of n species is considered as an n-tuple of random variables which evolve as continuous Markovian processes with statistical dependency described by the phylogenetic tree. For a species, let \( y_t \) denote its trait value at time \( t \), for instance the body and/or brain mass, and it is described via an appropriate stochastic process along the evolution. In this paper, as in [7], we assume that
the response trait evolves toward an optimum, \( \theta \). More precisely, this paper considers that \( y_t \) is a solution of an Ornstein-Uhlenbeck (OU) stochastic differential equation (SDE) \([4, 13, 15]\).

\[
dy_t = -\alpha_1 (y_t - \theta) dt + \sigma_y dW^y_t, \tag{1.1}
\]

where \( \alpha_1 \) measures the rate of adaption toward the optimum \( \theta \), \( W^y_t \) is a white noise with mean zero and appropriate covariance, and \( \sigma_y \) is the standard deviation of the random change. Considering the OU dynamics of evolution as expressed in (1.1), we note that the deterministic part is responsible for a linearly increasing pull of the trait toward the primary optimum, and the stochastic part expresses an indirect change.

For the model as described in (1.1), the works \([2, 7]\) considered that the optimal value of the trait, \( \theta \), is constant. The study in \([9]\) relaxed this restriction by assuming that the optimum \( \theta_t \) at any point on the phylogeny is a linear function of some stochastic variable \( x \) which evolved according to a Brownian motion.

In this paper, considering that the optimal trait, \( \theta_t \), expressed via a linear regression on the predictor \( x \), we assume that the predictor \( x \) evolves as an OU stochastic process. In other words, our phylogenetic model is given, now, via the following coupled equations.

\[
dy_t = -\alpha_1 (y_t - \theta_t(x)) dt + \sigma_y dW^y_t \tag{1.2}
\]
\[
dx_t = -\alpha_2 (x_t - \gamma) dt + \sigma_x dW^x_t \tag{1.3}
\]

where \( \alpha_1, \alpha_2 \) measure forces for pulling the trait \( y \) and \( x \) back to their own optima, \( \theta_t(x), \gamma \), respectively; the white noise \( W^x_t \) is correlated with the white noise \( W^y_t \); \( \gamma \) is considered constant and thus the covariance of traits does not change (see \((2.14)\)). For this reason, without loss of generality, \( \gamma \) will be ignored from our analysis. Expressing the trait evolution using (1.2) and (1.3) enables both the response and predictor evolved under the same type of Markovian processes while the Brownian motion considered in \([9]\) is a special case for the predictor.

The linear dependency of the optimal trait value \( \theta_t \) on the predictor \( x \) implies that \( \theta_t \) is also a solution of an appropriate OU stochastic differential equation. Having described the trait value and its corresponding optimal as solutions of two appropriate OU SDEs, we establish the evolutionary regression curve between the predictor \( x_t \) and the trait \( y_t \) (see Theorem 2.1). The key step for finding the regression curve is Proposition 2.1 which demonstrates the regression estimates on trait values, \( y_t \), given the optimal \( \theta_t \). To estimate the regression parameters, generalized least squares estimates (GLS) \([12]\) are used. The GLS estimates though depend on several parameters, e.g. the rate of adaptations and the variance of the trait values, which appear in the design and variance matrices of the regression. Overpassing this hurdle, we develop an algorithm different from \([9]\) which identifies the maximum likelihood estimators by using Powell’s
method [17, 18]. Briefly, Powell’s method relies on derivative-free univariate line optimization techniques, making the computations feasible.

Our methodology is then applied to an ecological data set of the evolution of the woodcreepers (see Figure 4). As woodcreepers use their tail for body support, we study how the tip width (the width of the rachis at the base of the medial rectrix) would adapt to the base width (the width of the rachis at the tip of the rectrix). Given this data set, our model is compared with the recent regression methods established in [9].

Our paper is organized in the following way. Section 2 describes the establishment of the OU regression curves. Section 3 provides an algorithmic implementation for estimating the various parameters of the regression curves of Section 2, and Section 4 considers an ecological data set of woodcreepers as a showcase. Finally, Section 5 offers a discussion on the subject.

2. Methodology

Let the response trait at time \( t \), \( y_t \), be a solution of the SDE as in (1.2). Let further assume that the optimum is linearly changing according to the predictor \( x_t \), i.e. \( \theta_t = b_0 + b_1 x_t \). The predictor variable \( x_t \) is presumed to be a solution of an OU stochastic differential equation given in (1.3). The linear relationship between the optimum, \( \theta_t \), and the predictor, \( x_t \), implies that \( \theta_t \) will be a solution of an OU process. At this point, we will write \( \theta_t \equiv \theta_t(x) \) by suppressing the variable \( x \). Let consider \( \theta_t \) be a solution of the SDE below.

\[
\frac{d\theta_t}{\theta_t} = -\alpha_3 \theta_t dt + \sigma_\theta dW^\theta_t, \tag{2.1}
\]

where \( W^\theta_t \) is correlated with the Brownian motion, \( W^y_t \), considered in (1.2), \( \alpha_3 \) measures the force of adaptation to the optimum, and \( \sigma_\theta = b_1 \sigma_x \) measures the magnitude of stochastic perturbation of \( \theta_t \). Without loss of generality we will assume that the rates of adaptation in (1.2) and (2.1), respectively, are equal, i.e. \( \alpha_1 = \alpha_3 = \alpha \). This assumption is not essential and it is made just to simplify the established formulas of the moments of the trait, \( y_t \), and the optimum, \( \theta_t \), verified in Lemmas 2.1-2.4.

Theorem 2.1 is the central theorem of this paper and establishes theoretically the evolutionary regression curve under the assumption that the trait and the predictor are solutions of OU-type SDEs.

**Theorem 2.1.** *Evolutionary Regression Curve* Let assume that the trait value \( y_t \) and its predictor are solutions of the coupled stochastic differential equations defined in (1.2) and (1.3), respectively. Let further assume that the optimum process \( \theta_t \) is related to the predictor \( x_t \) via a linear regression of the form, \( \theta_t = b_0 + b_1 x_t \). Then the evolutionary regression curve of the trait \( y_t \) on \( x_t \) is given by,

\[
E(y_t|x_t) = q(\alpha t) + p(\alpha t)b_1(x_t - x_a), \tag{2.2}
\]

where for \( t > a, x_a \) is the ancestral value of \( x \), \( p(\alpha t) = \frac{1 - \exp(-2\alpha t) - \alpha t \exp(-2\alpha t)}{2(1 - \exp(-2\alpha t))} \), and \( q(\alpha t) = (b_0 + b_1 x_a)(1 - \exp(-\alpha t)) + y_a \exp(-\alpha t) \).
The proof of Theorem 2.1 is based on the following proposition.

**Proposition 2.1.** Let consider the stochastic differential equations, (1.2) and (2.1), assuming that \( \alpha_1 = \alpha_3 = \alpha \). Let \( y_t, \theta_t \) be their corresponding solutions with initial conditions \( \theta_0, y_0 \), respectively. The regression of the trait on the optimum is given via \( \mathbb{E}[y_t|\theta_t] = \hat{\beta}_0 + \hat{\beta}_1 \theta_t \), where

\[
\hat{\beta}_1 = \frac{1 - \exp(-2\alpha t) - \alpha t \exp(-2\alpha t)}{2(1 - \exp(-2\alpha t))}, \tag{2.3}
\]

\[
\hat{\beta}_0 = \alpha \theta_0 t \exp(-\alpha t) + y_0 \exp(-\alpha t) - \hat{\beta}_1 \theta_0, \tag{2.4}
\]

**Proof of Theorem 2.1.** We have shown in Proposition 2.1 that the regression line

\[
\mathbb{E}[y_t|\theta_t] = \alpha \theta_0 t \exp(-\alpha t) + y_0 \exp(-\alpha t) + \frac{1 - \exp(-2\alpha t) - \alpha t \exp(-2\alpha t)}{2(1 - \exp(-2\alpha t))} (\theta_t - \theta_0). \tag{2.5}
\]

We substitute the optimal value \( \theta_t \) in (2.5) with its linear dependency on \( x_t \), i.e. \( b_0 + b_1 x_t \). Then,

\[
\mathbb{E}[y_t|x_t] = \alpha (b_0 + b_1 x_a) t \exp(-\alpha t) + y_a \exp(-\alpha t) + \frac{1 - \exp(-2\alpha t) - \alpha t \exp(-2\alpha t)}{2(1 - \exp(-2\alpha t))} b_1 (x_t - x_a),
\]

where \( x_a, y_a \) are the ancestral values of \( x_t \) and \( y_t \), respectively. From there it is clear what \( q(\alpha t) \) and \( p(\alpha t) \) are. ■

The proof of the Proposition 2.1 is a direct application of the definition of the regression line. For this reason the moments of the trait, \( y_t \), and the corresponding optimum, \( \theta_t \) need to be established. Lemma 2.1 and Lemma 2.3 demonstrate the first and second moments of the optimum \( \theta_t \), and the corresponding ones of the trait, \( y_t \), respectively. Also, the expected value of their product is given in Lemma 2.2. Their proofs blend ordinary differential equations derivations together with stochastic analysis, and they can be found to the appendix.

**Lemma 2.1.** Let consider the stochastic differential equation (2.1), assuming that \( \alpha_3 = \alpha \). Let \( \theta_t \) be its corresponding solution, i.e. an OU process, with initial condition \( \theta_0 \). Then,

\[
\mathbb{E}[\theta_t] = \theta_0 \exp(-\alpha t), \tag{2.6}
\]

\[
\mathbb{E}[\theta_t^2] = \frac{\sigma^2}{2\alpha} (1 - \exp(-2\alpha t)) + \theta_0^2 \exp(-2\alpha t). \tag{2.7}
\]

**Lemma 2.2.** Let consider the stochastic differential equations, (1.2) and (2.1), assuming that \( \alpha_1 = \alpha_3 = \alpha \). Let \( y_t, \theta_t \) be their corresponding solutions with initial conditions \( \theta_0, y_0 \). Then,

\[
\mathbb{E}[y_t \theta_t] = \frac{\sigma^2}{4\alpha} + (y_0 \theta_0 - \frac{\sigma^2}{4\alpha}) \exp(-2\alpha t) + (\theta_0^2 \alpha - \frac{\sigma^2}{2\alpha}) t \exp(-2\alpha t). \tag{2.8}
\]
Lemma 2.3. Let consider the stochastic differential equation (1.2), assuming that $\alpha_1 = \alpha$. Let $y_t$ be its corresponding solution with initial conditions $y_0$. Then,

$$\mathbb{E}[y_t] = \alpha \theta_0 t \exp(-\alpha t) + y_0 \exp(-\alpha t)$$  (2.9)

$$\mathbb{E}[y_t^2] = \left(\frac{\sigma_y^2}{2\alpha} + \frac{\sigma_\theta^2}{4\alpha} \right) + \left(\alpha^2 \theta_0^2 - \alpha \sigma_\theta / 2\right) t^2 \exp(-2\alpha t)$$

$$+ \left(2\alpha y_0 \theta_0 - \sigma_\theta^2 / 2\right) t \exp(-2\alpha t) + \left[y_0^2 - \left(\frac{\sigma_y^2}{2\alpha} + \frac{\sigma_\theta^2}{4\alpha}\right)\right] \exp(-2\alpha t)$$  (2.10)

Now, using the definitions of the variance and the covariance of a random variable, together with Lemma 2.1, Lemma 2.2, and Lemma 2.3, one establishes Lemma 2.4.

Lemma 2.4. Let consider the stochastic differential equations, (1.2) and (2.1), assuming that $\alpha_1 = \alpha_3 = \alpha$. Let $y_t, \theta_t$ be their corresponding solutions with initial conditions $\theta_0, y_0$. Then,

$$\text{Var}[\theta_t] = \sigma_\theta^2 \left(1 - \exp(-2\alpha t)\right) \quad (2.11)$$

$$\text{Cov}[y_t, \theta_t] = \sigma_y^2 \left(\frac{1 - \exp(-2\alpha t)}{4\alpha} - \frac{t}{2} \exp(-2\alpha t)\right)$$  (2.12)

$$\text{Var}[y_t] = \left(\frac{\sigma_y^2}{2\alpha} + \frac{\sigma_\theta^2}{4\alpha}\right) \left(1 - \exp(-2\alpha t)\right) - \sigma_\theta^2 \left(1 + \alpha t\right) \exp(-2\alpha t) / 2$$  (2.13)

To estimate the regression parameters, (2.3) and (2.4), we will use generalized least squares methods. For this we need to find the variance and the covariance of the residuals. Let $r_i = y_i - \mathbb{E}[y_i|\theta_i]$ be the $i$th residual associated with the $i$th prediction from the regression curve as verified in Proposition 2.1. The covariance between traits can be obtained by utilizing the definition in [8],

$$\text{Cov}[y_i, y_j] = \text{Cov}[\mathbb{E}[y_i|y_a], \mathbb{E}[y_j|y_a]]$$

where $y_a$ is the ancestral value of the $y$. Let consider a species which at time $t_a$ the pair of the trait and its corresponding optimum was $(y_a, \theta_a)$. Furthermore, let assume that at time $t_a$ the species diverges, and it takes $t_i, t_j$ time for the $i, j$ offspring respectively to evolve until the current time. In this paper, we assume that $t_i = t_j = \frac{t_{ij}}{2}$, where $t_{ij}$ is the total time of the evolution of the $i$ and $j$ offsprings until current time. When we proceed with the phylogenetic data analysis, the time $t$ is equal to $t_{ij}$ for the processes $y_t, \theta_t$. Thus, the index $t$ standing for the time will be suppressed.

The conditional expectation of the trait $y_i$ given its ancestral value $y_a$, $\mathbb{E}[y_i|y_a]$, is given below.

$$\mathbb{E}[y_i|y_a] = \alpha \theta_n \frac{t_{ij}}{2} \exp(-\alpha t_{ij} / 2) + y_a \exp(-\alpha t_{ij} / 2).$$
where $\theta_a$ are the ancestral values of the trait $y$, and $a$, $\sigma_a$ are the maximum likelihood estimator (MLE) of $\theta_a$, and its corresponding optimal $\alpha$. One could estimate the MLEs for the mean and the variance of the predictor $\hat{x}$, multiplied by $a$, via the following equations. 

$$
\hat{x} = \left( I' T_a I \right)^{-1} I' T_a^{-1} x, \\
\hat{\sigma}_x^2 = \frac{(x - E[x])' T_a^{-1} (x - E[x])}{n - 1}, 
$$

where $I \in \mathbb{R}^{n \times 1}$ is a vector with all entries 1; $T_a$ is the covariance matrix under the OU process for the trait evolution, i.e. $T_a[i, j] = \exp(-2\hat{\alpha} t_{ij}) \frac{1 - \exp(-2\hat{\alpha} t_{ij})}{2\hat{\alpha}}$. In order to estimate the parameters $\alpha$, $\sigma_y^2$, we consider the log-likelihood function $\ell$ as following, 

$$
\ell(\alpha, \sigma_y^2) = -\frac{n}{2} \log(2\pi) - \frac{1}{2} \log(\det(V)) - \frac{1}{2} (y - Xb) V^{-1} (y - Xb). 
$$

(3.1)
Table 1

| Model   | Method | Regression Line | $r^2$ | AICc |
|---------|--------|----------------|-------|------|
| M1: OUOU | Sections 2 & 3 | $y = 0.075 + 0.227x$ | 23.7 % | -31.83 |
| M2: OUBM | [9] | $y = 0.059 + 0.284x$ | 22.2 % | -32.72 |

Regression curves’ comparison

The regressors’ vector $b$ is estimated by

$$\hat{b} = (X^TV^{-1}X)^{-1}X^TV^{-1}y.$$  

We observe that the GLS estimates of regression parameters depend on the design and covariance matrices, $V$, $X$, respectively; and in turn they depend on $\alpha, \sigma_y^2$. Therefore, our algorithm starts with an ordinary least-squares estimate, $b_0 = (X^TX)^{-1}X^Ty$, which is independent of the rate of adaptation, $\alpha$, and standard deviation, $\sigma_y$. Then, the MLE pair $(\hat{\alpha}, \hat{\sigma}_y^2)$ is estimated through optimizing $\ell$ on a domain appropriate for $(\alpha, \sigma_y^2)$. The upper bound for $\alpha > 0$ is arbitrary. The variation of data should not exceed a range as much as the difference between the maximum and the minimum observation, in other words we set the domain of $\sigma_y^2$ to be $[0, y(n) - y(1)]$. After establishing the MLEs $(\hat{\alpha}, \hat{\sigma}_y^2)$, the regression estimator $\hat{b}$ is updated. The process is ongoing until the error measured, $err \doteq ||\hat{b} - \hat{b}_0||$, is less than some threshold $\delta$. We describe our algorithm below:

Step 1: Set $\hat{b}_0 = (X^TX)^{-1}X^Ty$.

Step 2: Search the MLEs $(\hat{\alpha}, \hat{\sigma}_y^2)$ for $\ell_0(\alpha, \sigma_y^2)$ through Powell’s method [17].

Step 3: Calculate $b_{est} = (X^TV^{-1}X)^{-1}X^TV^{-1}y$.

Step 4: Set $\delta = ||\hat{b}_{est} - \hat{b}_0||$

- If $\delta < err$, return $b_{est}$
- else set $b_0 = b_{est}$, Go to Step 2.

4. Data Analysis

Using the algorithm explained in Section 3, we establish the evolutionary regression curve for the woodcreepers data set of $n = 39$ species [22]. The explanatory variable $x$ is considered the rachis width at the tip and the corresponding response, $y$, the rachis width at the base. We further consider two regressors ($q = 2$). We also set the threshold for the error, $err$, not to exceed $\delta = 10^{-5}$ for identifying the convergent regressors. We also compare our method with the recent development of [9]. The regression output and the results from the regression comparison are given in the Table 1. Figure 4 displays the two evolutionary curves along with the data. The reader should remark at this point that for the woodcreepers data set, the $r^2$ calculated using the algorithm of this
paper is slightly larger than the corresponding $r^2$ value of the model in [9]. This signifies an improvement on the fit of the data. On the other hand, the AICc value,[10, 21], of the OUOU model is lower than the corresponding AICc of the OUBM model, the difference between the two values is $\Delta_{AICc} = 0.38$. [1] indicates that if the AICc values of two models differ no more than 2 units, then the model with the higher AICc within these two units should receive consideration for making inference.

5. Conclusion

A novel approach to the phylogenetic regression analysis was presented. Our method considered that the trait value of the species and the corresponding optimum evolve with OU-type dynamics. Our method was then implemented.
Fig 2. Evolutionary regression curves for OUOU model (red dashed line) and BMOU (blue solid line).
algorithmically using generalized least squares methods and it was further applied to a real ecological data set of the woodcreepers’ evolution. The established regression curve demonstrated a better fit in comparison with the model of [9]. Our model can be further generalized, and perhaps approach more realistic scenarios, by considering that the rate of adaptation is another stochastic process.

Appendix

Proof of Lemma 2.1. Consider expectations on both sides of the SDE (2.1). Given that the mean of the Brownian motion, $W^0$, vanishes, i.e. $\mathbb{E}(W^0) = 0$, we have the following ordinary differential equation (ODE), $\frac{d\mathbb{E}[\theta_t]}{dt} = -\alpha \mathbb{E}[\theta_t]$ with initial condition $\mathbb{E}[\theta(0)] = \mathbb{E}[\theta_0] = \theta_0$. Its solution is obviously the right hand side (RHS) of (2.6). Next, applying Itô’s formula; it yields $d\theta_t^2 = (\sigma^2 - 2\alpha \theta_t^2)dt + 2\sigma \theta_t dW_t^0$. After considering expectations, the ODE $\frac{d\mathbb{E}[\theta_t^2]}{dt} = \sigma^2 - 2\alpha \mathbb{E}[\theta_t^2]$ with initial condition $\mathbb{E}[\theta^2(0)] = \theta_0^2$ leads to (2.7).

Proof of Lemma 2.2. The stochastic integration by parts formula [13] implies $d(y_t \theta_t) = y_t d\theta_t + \theta_t dy_t + dy_t d\theta_t$, where $dy_t d\theta_t \equiv d[y_t \theta_t]$ is the quadratic variation. Therefore, $d(y_t \theta_t) = -\alpha \theta_t (2y_t - \theta_t)dt + \sigma y_t dW_t^0 + \sigma \theta_t dW_t^0$. Applying expectations, and recalling that $y_t, \theta_t$ are adapted processes, [13], one has $\frac{d\mathbb{E}[y_t \theta_t]}{dt} = -2\alpha \mathbb{E}[y_t \theta_t] + \alpha \mathbb{E}[\theta_t^2]$. The solution of this ODE is the RHS of (2.8).

Proof of Lemma 2.3. Let apply to the SDE, (1.2), expectations and substitute the expected value of the optimum $\theta_t$ as given in Proposition 2.1. One then gets the ODE $\frac{dy_t^2}{dt} + \alpha \mathbb{E}[y_t] = \alpha \theta_0 \exp(-\alpha t)$ with initial condition $\mathbb{E}[y(0)] = y_0$ whose solution is (2.9). Next, applying Itô’s formula; it yields $dy^2_t = \{\sigma^2 - 2\alpha y_t (y_t - \theta_t)\}dt + 2\sigma y_t dW^0_t$. After considering expectations and substituting the $\mathbb{E}[y_t \theta_t]$ according to Proposition 2.2, the ODE $\frac{d\mathbb{E}[y^2_t]}{dt} = \sigma^2 - 2\alpha \mathbb{E}[y_t^2] + 2\alpha \mathbb{E}[y_t \theta_t]$, with initial condition $\mathbb{E}[y^2(0)] = y_0^2$ leads to (2.10).

Acknowledgements

This work was conducted while the first author was a Postdoctoral Fellow and the second author was a Postdoctoral Mentor at the National Institute for Mathematical and Biological Synthesis (NIMBioS), an Institute sponsored by the National Science Foundation, the U.S. Department of Homeland Security, and the U.S. Department of Agriculture through NSF Award #EF-0832858, with additional support from The University of Tennessee, Knoxville.

References

[1] Anderson, D. R. and Burnham, K. P. (2002). Model selection and multimodel inference. Springer-Verlag New York.
[2] Butler, M. A., and King, A.A. (2004). Phylogenetic comparative analysis: a modeling approach for adaptive evolution. The American Naturalist, 164, 683-695.

[3] Felsenstein, J. (1985) Phylogenies and the comparative method. The American Naturalist, 125, 1-15.

[4] Gillespie, D.T. (1996) Exact numerical simulation of the Ornstein-Uhlenbeck process and its integral. Phys. Rev. E, 54, 2084-91.

[5] Gittleman, J. L. and Kot, M. (1990). Adaptation: statistics and a null model for estimating phylogenetic effects. Systematic Zoology, 39, 227-241.

[6] Grafen, A. (1989). The phylogenetic regression. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 326, 119-157.

[7] Hansen, T.F. (1997). Stabilizing selection and the comparative analysis of adaptation. Evolution, 51, 1341-1351.

[8] Hansen, T.F., and Martins, E.P. (1996). Translating between microevolutionary process and macroevolutionary patterns: the correlation structure of interspecific data. Evolution, 50, 1404-1417.

[9] Hansen, T.F., Pienaar, J., and Orzack, S. H. (2008). A comparative method for studying adaptation to a randomly evolving environment. Evolution, 62, 1965-1977.

[10] Hurvich, C. M., and Tsai, C.-L. (1989). Regression and time series model selection in small samples. Biometrika, 76, 297-307.

[11] Jhwueng, D.-C. (2011) Model selection and goodness of fit for phylogenetic comparative method. Submitted.

[12] Norman R. D. and Smith, H. (1998) Applied Regression Analysis, Wiley Series in Probability and Statistics.

[13] Øksendal, B. (2000). Stochastic differential equations: an introduction with applications. Springer-Verlag New York.

[14] O’Meara, B. C., Ané, C., Sanderson, M. J. and Wainwright, P. C. (2006). Testing for different rates of continuous trait evolution using likelihood. Evolution, 60, 922-933.

[15] Ornstein, L.S. and Uhlenbeck, G.E. (1930). On the theory of Brownian Motion. Phys.Rev., 36, 823-841.

[16] Paradis E., Claude, J. and Strimmer, K. (2004). APE: analyses of phylogenetics and evolution in R language. Bioinformatics, 20, 289-290.

[17] Powell, M. J. D. (1964). An efficient method for finding the minimum of a function of several variables without calculating derivatives. Computation Journal, 7, 155-162.

[18] Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. (2007). Numerical Recipes: the art of scientific computing. Cambridge University Press, Port Chester, New York, third edition.

[19] Raikowr, J. (1994). A phylogeny of the woodcreepers (Dendrocolaptinae). The Condor, 111, 104-114.

[20] Revell, L. J. (2008). On the analysis of evolutionary change along single branches in a phylogeny. The American Naturalist, 172, 140-147.

[21] Sugiuira N. (1978). Further analysis of the data by Akaike’s information criterion and the finite corrections. Communications in Statistics - Theory
and Methods, A7, 1326.

[22] Tubaro, P. L., Lijtmaer, D. A., Palacios, M. G. and Kopuchian, C. (2002). Adaptive Modification of Trail Structure in Relation to Body Mass and Buckling in Woodcreepers. The Condor, 14, 281-296.