

Calculation Update

Priscilla Lau, Bjørn Kopperud, Sebastian Höhna, Jeremy M. Beaulieu, and Brian C. O’Meara

04/21/2026

There was an issue in OUMA and OUMVA models in OUwie versions from its origin through 2.17. This has now been fixed in OUwie versions 3.00 and up. Jeremy and Brian deeply apologize for this error. Details below:

Timeline

Priscilla Lau reached out to Jeremy Beaulieu in Aug, 2025, at a conference about an issue she and her collaborators (Bjørn Kopperud and Sebastian Höhna) found in how **OUwie** calculates the variance-covariance matrix under the OU model when there are different α parameters, as well as an error in one of the equations in the original Beaulieu et al. (2012) paper. She followed up with an email on Sept 3, 2025; after a discussion of how best to handle this, we put a warning on the use of models with different α in **OUwie** (on Sept. 6, 2025). A further email from Priscilla and colleagues had additional details on the issues.

Brian spent some time trying to fix the code; Jeremy swooped in and *actually* fixed the code using the approaches from Priscilla, Bjørn, and Sebastian.

We are posting this vignette; we will soon reach out to *Evolution* for a correction to Beaulieu et al. (2012).

Rendered equation error in original Beaulieu et al. (2012)

In the appendix on page 2383, when it describes the covariation under an OUMVA model, the right-most calculation in the final term is rendered as

$$\left(\sum_{\gamma=1}^{k(i,j)} \sigma_{\gamma}^2 \frac{e^{2s_{ij,\gamma}} - e^{2s_{ij,\gamma-1}}}{2\alpha_{\gamma}} \right),$$

when it should be

$$\left(\sum_{\gamma=1}^{k(i,j)} \sigma_{\gamma}^2 \frac{e^{2\alpha s_{ij,\gamma}} - e^{2\alpha s_{ij,\gamma-1}}}{2\alpha_{\gamma}} \right),$$

as it is correctly written in Eq. A.26. This is unrelated to the issue below but still worth noting.

Variance-covariance matrix calculation error in OUwie

For model inference the way **OUwie** calculated the variance-covariance matrix under the OU model with multiple α parameters was indeed incorrect. In an Ornstein–Uhlenbeck process, shared variance decays through time as a function of the α parameter: higher values of α cause trait values to lose memory of their

shared history more rapidly, making the optima toward which lineages are pulled more important than their ancestral states.

The key issue is how variance decay is handled. The recursive approach of Lau et al. (in review – we will update this vignette when it becomes available) builds variance step-by-step along each branch, correctly applying decay at each stage. In contrast, Eq. 13 on page 2371 gives a closed-form expression, but does not fully account for how decay accumulates across segments when α varies. As a result, it is not equivalent in the general multi- α case, although the two agree when α is constant (i.e., the BM1, BMS, OU1, OUM, and OUMV cases).

Adopting the corrections proposed by Priscilla Lau and colleagues, we have updated the code in `OUwie` version 3.01 (released March 6, 2026) to correctly calculate the variance-covariance matrix under the OU model with multiple α parameters. We thank Priscilla Lau, Bjørn Kopperud, and Sebastian Höhna for their careful work in identifying this issue and proposing a solution.

Details of the issue

This part was written by Lau, Kopperud, and Höhna (though Beaulieu & O’Meara agree with it – we just don’t want to take credit for the discovery).

General equations

Expectation of the trait value for species i

The expectation of the trait value y for species i is a sum over the vector θ , weighted by W_{ik}

$$E[y_i] = \sum_{k=0}^r W_{ik} \theta_k \quad , \quad (1)$$

where θ_0 is the parameter for the root value, and $\theta_1, \theta_2, \dots$ are the state-dependent optima for regimes 1, 2, and so on. Thus, each W_{ik} represents the relative contribution of either the root value or the optima on the expected value at the tip. As described in Beaulieu et al. (2012), the weights are as follows

$$W_{i0} = \exp^{-\sum_{\gamma=1}^{\kappa(i)} ((\sum_{k=1}^r (\beta_{ik,\gamma} \tilde{\alpha}_k)(s_{i,\gamma} - s_{i,\gamma-1}))} \\ W_{ik} = e^{-\sum_{\gamma=1}^{\kappa(i)} (\sum_{k=1}^r \beta_{ik,\gamma} \tilde{\alpha}_k)(s_{i,\gamma} - s_{i,\gamma-1})} \times \sum_{\gamma=1}^{\kappa(i)} \beta_{ik,\gamma} \left(e^{(\sum_{k=1}^r \beta_{ik,\gamma} \tilde{\alpha}_k)s_{i,\gamma}} - e^{(\sum_{k=1}^r \beta_{ik,\gamma} \tilde{\alpha}_k)s_{i,\gamma-1}} \right) \quad . \quad (2)$$

In the new (Lau et al.) derivation, the general equations are,

$$W_{i0} = e^{-\sum_{\gamma=1}^{\kappa(i)} ((\sum_{k=1}^r (\beta_{ik,\gamma} \tilde{\alpha}_k)(s_{i,\gamma} - s_{i,\gamma-1}))} \\ W_{ik} = \sum_{\gamma=1}^{\kappa(i)} \left[\overbrace{e^{-\sum_{\gamma' > \gamma} ((\sum_{k=1}^r \beta_{ik,\gamma'} \tilde{\alpha}_k)(s_{i,\gamma'} - s_{i,\gamma'-1}))}}^{\text{decay of epoch } \gamma} \times \beta_{ik,\gamma} \left(1 - e^{(-\sum_{k=1}^r \beta_{ik,\gamma} \tilde{\alpha}_k)(s_{i,\gamma} - s_{i,\gamma-1})} \right) \right] \quad . \quad (3)$$

There are two differences in the way Beaulieu et al. (2012) construct W_{ik} when α takes different values across epochs compared with the new derivation. First—in Beaulieu et al. (2012)’s derivation, the decay variable is outside of the summation function, and their summation function starts with epoch index $\gamma = 1$.

This implies that the decay always starts from the root, regardless of when each epoch ends. However, the contribution of epoch γ on the expected value should only begin decaying when the epoch γ the ends. Therefore, in the new equations, the decay variable is inside the summation function, and is dependent on the time when each epoch ends (see the sum that iterates over all epochs $\gamma' > \gamma$).

The second factor of W_{ik} in Beaulieu et al. (2012)'s derivation is also different. The factor is taken from Hansen (1997). However, Hansen (1997) assumes that α takes the same value in all epochs, and the factor already incorporates how the weight of each epoch-specific optimum decays after the respective epoch. Since in Beaulieu et al. (2012)'s models α can take multiple values, the decay variable has to be extracted out of the factor.

Furthermore, Hansen (1997) stated that the weights should sum to one, which is the case using our derivation. The extra step where “each row entry in W is divided by the sum of its row to ensure that the weights for each species sum to 1” [Beaulieu et al. (2012)] is not required when the weights sum to one.

Covariance between species i and j

Here, we calculate the covariance between the trait value for species i and the trait value for species j , i.e., $\text{Cov}[y_i, y_j] = V_{ij}$. As described in Beaulieu et al. (2012),

$$V_{ij} = e^{-\left(\sum_{\gamma=1}^{\kappa(i)} \alpha_{i,\gamma}(s_{i,\gamma} - s_{i,\gamma-1}) + \sum_{\gamma=1}^{\kappa(j)} \alpha_{j,\gamma}(s_{j,\gamma} - s_{j,\gamma-1})\right)} \times \left(\sum_{\gamma=1}^{\kappa(i,j)} \sigma_{\gamma}^2 \frac{e^{2\alpha_{\gamma} s_{ij,\gamma}} - e^{2\alpha_{\gamma} s_{ij,\gamma-1}}}{2\alpha_{\gamma}} \right) . \quad (4)$$

In our derivation, the general equations are

$$V_{ij} = e^{-\left(\sum_{\gamma=\kappa(ij)}^{\kappa(i)} \alpha_{i,\gamma}(s_{i,\gamma} - s_{i,\gamma-1}) + \sum_{\gamma=\kappa(ij)}^{\kappa(j)} \alpha_{j,\gamma}(s_{j,\gamma} - s_{j,\gamma-1})\right)} \times \left(\sum_{\gamma=1}^{\kappa(i,j)} \sigma_{\gamma}^2 \frac{e^{2\alpha_{\gamma} s_{ij,\gamma}} - e^{2\alpha_{\gamma} s_{ij,\gamma-1}}}{2\alpha_{\gamma}} \right) . \quad (5)$$

The difference lies in the first factor, which describes the exponential decay of the variance since the divergence of species i and j . In @Beaulieu2012's derivation, the exponent in the first factor always calculates from the root ($\gamma = 1$), which means the decay always begins at the root. In our derivation, the decay starts from the time when species i and j diverge ($\gamma = \kappa(ij)$).

A simple example

We use a simple character history of two species for demonstration. Epochs in black are in regime A and epochs in red are in regime B. Thus, i th lineage has $\kappa(i) = 3$ epochs, j th lineage has $\kappa(j) = 2$ epochs, while the ancestral lineage ij until the divergence has $\kappa(ij) = 1$ epoch. The time variable s is understood as “time elapsed since the root of the phylogeny”. For simplification, we also set each epoch in lineage i to have equal lengths (Δt).

Therefore,

$$\begin{aligned}
\alpha_{ij=1} &= \alpha_{i,1} = \alpha_{i,2} \\
&= \alpha_{j,1} = \alpha_{j,2} \\
&= \alpha_A \\
\alpha_{i,3} &= \alpha_B \\
\sigma_{ij,1}^2 &= \sigma_A^2 \\
s_{i,1} - s_{i,0} &= s_{i,2} - s_{i,1} = s_{i,3} - s_{i,2} = \Delta t
\end{aligned} \tag{6}$$

Expectation of the trait value for species i

The respective weights of the root value and state-dependent optima calculated using equations in Beaulieu et al. (2012) are

$$\begin{aligned}
W_{i0} &= e^{-(\alpha_{i,1}(s_{i,1}-s_{i,0})+\alpha_{i,2}(s_{i,2}-s_{i,1})+\alpha_{i,3}(s_{i,3}-s_{i,2}))} \\
&= e^{-(2\alpha_A+\alpha_B)\Delta t} \\
W_{iA} &= e^{-(\alpha_A(s_{i,1}-s_{i,0})+\alpha_A(s_{i,2}-s_{i,1})+\alpha_B(s_{i,3}-s_{i,2}))} \times [(e^{\alpha_A s_{i,1}} - e^{\alpha_A s_{i,0}}) + (e^{\alpha_A s_{i,2}} - e^{\alpha_A s_{i,1}})] \\
&= e^{-(2\alpha_A+\alpha_B)\Delta t} \times (e^{2\alpha_A \Delta t} - 1) \\
&= (1 - e^{-2\alpha_A \Delta t}) \times e^{-\alpha_B \Delta t} \\
W_{iB} &= e^{-(\alpha_A(s_{i,1}-s_{i,0})+\alpha_A(s_{i,2}-s_{i,1})+\alpha_B(s_{i,3}-s_{i,2}))} \times (e^{\alpha_B s_{i,3}} - e^{\alpha_B s_{i,2}}) \\
&= e^{-(2\alpha_A+\alpha_B)\Delta t} \times (e^{3\alpha_B \Delta t} - e^{2\alpha_B \Delta t}) \\
&= e^{-(2\alpha_A+\alpha_B)\Delta t} \times (e^{\alpha_B \Delta t} - 1) \quad ,
\end{aligned} \tag{7}$$

which do not sum to 1.

The weights using our equations are

$$\begin{aligned}
W_{i0} &= e^{-(\alpha_{i,1}(s_{i,1}-s_{i,0})+\alpha_{i,2}(s_{i,2}-s_{i,1})+\alpha_{i,3}(s_{i,3}-s_{i,2}))} \\
&= e^{-(2\alpha_A \Delta t + \alpha_B \Delta t)} \\
W_{iA} &= (1 - e^{-\alpha_A(s_{i,2}-s_{i,0})}) \times e^{-(\alpha_B(s_{i,3}-s_{i,2}))} \\
&= (1 - e^{-2\alpha_A \Delta t}) \times e^{-\alpha_B \Delta t} \\
W_{iB} &= (1 - e^{-\alpha_B(s_{i,3}-s_{i,2})}) \\
&= 1 - e^{-\alpha_B \Delta t} \quad ,
\end{aligned} \tag{8}$$

which sum to 1.

Covariance between species i and j

Using equations in Beaulieu et al. (2012), the covariance between species i and j is

$$\begin{aligned}
V_{ij} &= e^{\alpha_{i,1}(s_{i,1}-s_{i,0})+\alpha_{i,2}(s_{i,2}-s_{i,1})+\alpha_{i,3}(s_{i,2}-s_{i,3})+\alpha_{j,1}(s_{j,1}-s_{j,0})+\alpha_{j,2}(s_{j,2}-s_{j,1})} \\
&\times \sigma_{ij,1}^2 \frac{e^{2\alpha_{ij,1}s_{ij,1}} - e^{2\alpha_{ij,1}s_{ij,0}}}{2\alpha_{ij,1}} \\
&= e^{\alpha_A \Delta t + \alpha_A \Delta t + \alpha_B \Delta t + \alpha_A \Delta t + 2\alpha_A \Delta t} \times \sigma_A^2 \frac{e^{2\alpha_A \Delta t} - 1}{2\alpha_A} \\
&= e^{(5\alpha_A + \alpha_B) \Delta t} \times \sigma_A^2 \frac{e^{2\alpha_A \Delta t} - 1}{2\alpha_A}
\end{aligned} \tag{9}$$

Using the new equations, the covariance is

$$\begin{aligned}
V_{ij} &= e^{\alpha_{i,2}(s_{i,2}-s_{i,1})+\alpha_{i,3}(s_{i,3}-s_{i,2})+\alpha_{j,2}(s_{j,2}-s_{j,1})} \\
&\times \sigma_{ij,1}^2 \frac{e^{2\alpha_{ij,1}s_{ij,1}} - 1}{2\alpha_{ij,1}} \\
&= e^{(\alpha_A+\alpha_B+2\alpha_A)\Delta t} \times \sigma_A^2 \frac{e^{2\alpha_A\Delta t} - 1}{2\alpha_A} \\
&= e^{(3\alpha_A+\alpha_B)\Delta t} \times \sigma_A^2 \frac{e^{2\alpha_A\Delta t} - 1}{2\alpha_A}
\end{aligned} \tag{10}$$

FAQ

How big of a difference does this make?

The exact impact will depend on the tree, the α values, and the data. In some cases it could be substantial. There is no predicted effect of whether this makes OUMVA or OUMA models seem to fit better or worse than they should otherwise.

Can I verify how this affects my analyses?

Yes, with `OUwie()` and similar functions we have added a `revert.old` argument that defaults to `FALSE`. If you set `revert.old=TRUE`, it will always use the old, *incorrect* method for calculating the variance-covariance matrix. You can compare results with `revert.old=TRUE` and `revert.old=FALSE` to see how much of a difference it makes for your specific analysis. We wanted to make sure that any problems are transparent and reproducible.

Was hOUwie affected?

Yes, in the same way for the multiple α models only, and corrected in the same way.

Can I see more details on the changes?

Yes, compare, for example, the code around line 56 of `varcov.ou.R` with the code around line 166 of the original calculation for trees with stochastic maps

```

for(regimeindex in 1:length(currentmap)){
  dt <- as.numeric(currentmap[regimeindex])
  regimenum <- which(colnames(phy$mapped.edge) == names(currentmap)[regimeindex])
  a <- alpha[regimenum]
  s2 <- sigma[regimenum]
  nodevar1[i] <- nodevar1[i] + a * dt
  nodevar2[i] <- nodevar2[i] + s2 * exp(2 * Acur) * (exp(2 * a * dt) - 1) / (2 * a)
  Acur <- Acur + a * dt
}
Ato[desc] <- Acur

for (regimeindex in 1:length(currentmap)){
  regimeduration <- currentmap[regimeindex]
  newtime <- oldtime+regimeduration
  regimenum <- which(colnames(phy$mapped.edge)==names(currentmap)[regimeindex])
  nodevar1[i] <- nodevar1[i]+alpha[regimenum]*(newtime-oldtime)
  nodevar2[i] <- nodevar2[i]+sigma[regimenum]*((exp(2*alpha[regimenum]*newtime)-exp(2*alpha[regimenum]*oldtime)) / (2*alpha[regimenum]))
  oldtime <- newtime
  newregime <- regimenum
}

```

Some of the variables have changed names, but a key change is in how the time accumulates along the tree, with it being scaled by alpha in the new code.

References

Lau P.L., B.T. Kopperud, J.T. Clarke, D. Metzler, and S. Höhna. 2026. An efficient Bayesian phylogenetic approach for joint inference of continuous and discrete trait evolution under a state-dependent Ornstein–Uhlenbeck model. *Methods in Ecology and Evolution*, In press.