

On wavelet analyses with the `mvwt` package, issues, and technical fixes

August 7, 2020

To study synchrony, Keitt (2008) presents the wavelet modulus ratio, i.e. $\rho(s, t) = \frac{\Lambda_{s,t}(|\sum_k w_k(s, t)|)}{\Lambda_{s,t}(\sum_k |w_k(s, t)|)}$ where $\Lambda_{s,t}(y) = \int_{-\infty}^{+\infty} e^{-\frac{1}{2}(\frac{t-\tau}{s})^2} (y) d\tau$, s is the scale (“frequency”) and t is the location (“date”) for which the wavelet w_k is computed for species k . The package `mvwt` is meant to compute this index.

The package works as follows:

1. the function `mvwt` takes a table of dates and time series and computes the corresponding Morlet wavelets for each scale and location required. By default, the number of locations L is the number of dates, and is equal to the number of scales S .
2. the function `wmr.boot` takes the wavelets and compute the wavelet modulus ratio (`wmr`) of the observed time-series. It then computes R surrogates to provide the quantiles corresponding to the observed value at each scale and location.
 - (a) it first computes the observed `wmr` (called z)
 - (b) for each scale, it shifts the phase of the observed wavelets for the number of surrogates required. This can be compared to surrogates produced by shifting in the time-domain. In total, there is therefore $S \times R$ bootstrapped wavelets.
 - (c) for each scale, it compares the value of the observed `wmr` at location t to the distribution of the index obtained with the surrogates. There is therefore $S \times L$ comparisons to a distribution made of R samples. This is performed with the `ecdf` function (Fig. 2) which returns $Pr(X \leq x)$. The table made of $S \times L$ $Pr(X \leq x)$ is called `z.boot`.
3. the function `image.mvwt` takes an object from `wmr.boot` (or `wmr`, which contains only the observed values of the index, z). It first plots the values in z as an image ; then uses `z.boot`. We describe first the classical way it does this, and will then explain the modifications we made.
 - (a) it first computes $1 - |1 - 2 * Pr(X \leq x)|$. This formulation corresponds to the usual two-sided p-value, that is $p_{val} = 2 \min(Pr(X \leq x), Pr(X \geq x))$.
 - (b) it computes the adjusted p-value $p_{val,adj}$ with a Benjamini-Yekutieli false discovery rate.
 - (c) it plots the contour for p_{val} and $p_{val,adj}$ at the chosen α level (0.05 by default).

What we changed

There were several adjustments made to the `image.mvwt` function. The first ones were only cosmetic. Panels were created to be able to plot a `wmr` image as well as other plots on the same figure, add a colorbar and increase the

size of certain labels. We also changed the Benjamini-Yekutieli False Discovery Correction (FDR) to Benjamini-Hochberg following the suggestion of Referee 2 in Revision 1 (who actually only suggested that we changed from Bonferroni to something else).

We also differentiated between low values and high values of the observed wmr. To do so:

1. we first compute p_{val} and $p_{val,adj}$
2. for each z.boot value, we check if $Pr(X \leq x) < Pr(X > x)$ (x has a low value and might therefore indicate compensation) or $Pr(X \leq x) > Pr(X > x)$ (x has a high value and might therefore indicate synchrony).
3. we draw a contour at the alpha level with the following code: red for each p_{val} (or $p_{val,adj}$) where $Pr(X \leq x) > Pr(X > x)$ and blue for each p_{val} (or $p_{val,adj}$) where $Pr(X \leq x) < Pr(X > x)$.

What were the issues?

There were several issues (theoretical, probably not due to the package itself) that we discovered after implementing the suggestions of Referee 1 and 3 in Revision 2.

- We first implemented, using the changes mentioned above, a visualization of significantly low (blue) and significantly high (red) synchrony zones, as suggested by referee 1 (Fig. 1).
- increasing the number of surrogates from 100 (Fig. 1) to 1000, as suggested by referee 1, most of the significant zones disappeared. As the number of surrogates increased, we could assume that the distribution of actually sampled values increased in breadth, with more and more observed values closer to the tail of the distribution, causing this behaviour (i.e., a sampling effect on the tails due a too low number of surrogates). This, however, was not clear when empirical cumulative distributions are plotted (Fig. 2). We actually do not know what causes this discrepancy, as the same technique seemed to be less sensitive to the number of surrogates for simulated data (not shown).

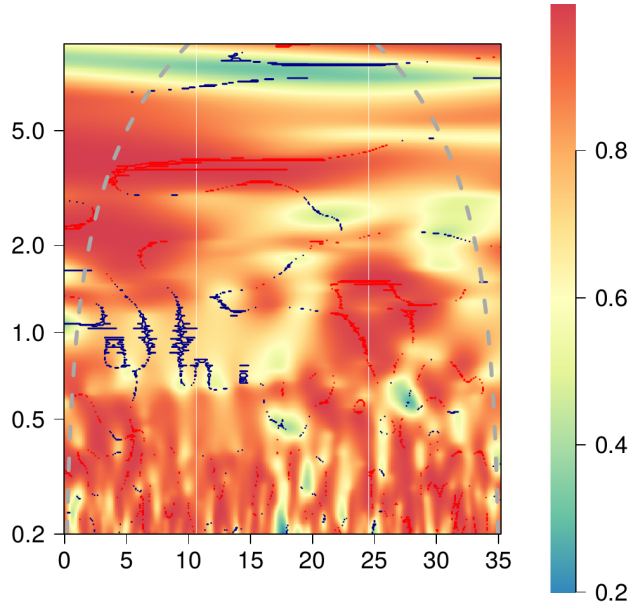


Figure 1: Reference wavelet analysis for the Cormoran/Heron/Egret case with 100 surrogates, BH correction and a fine grid (423x423)

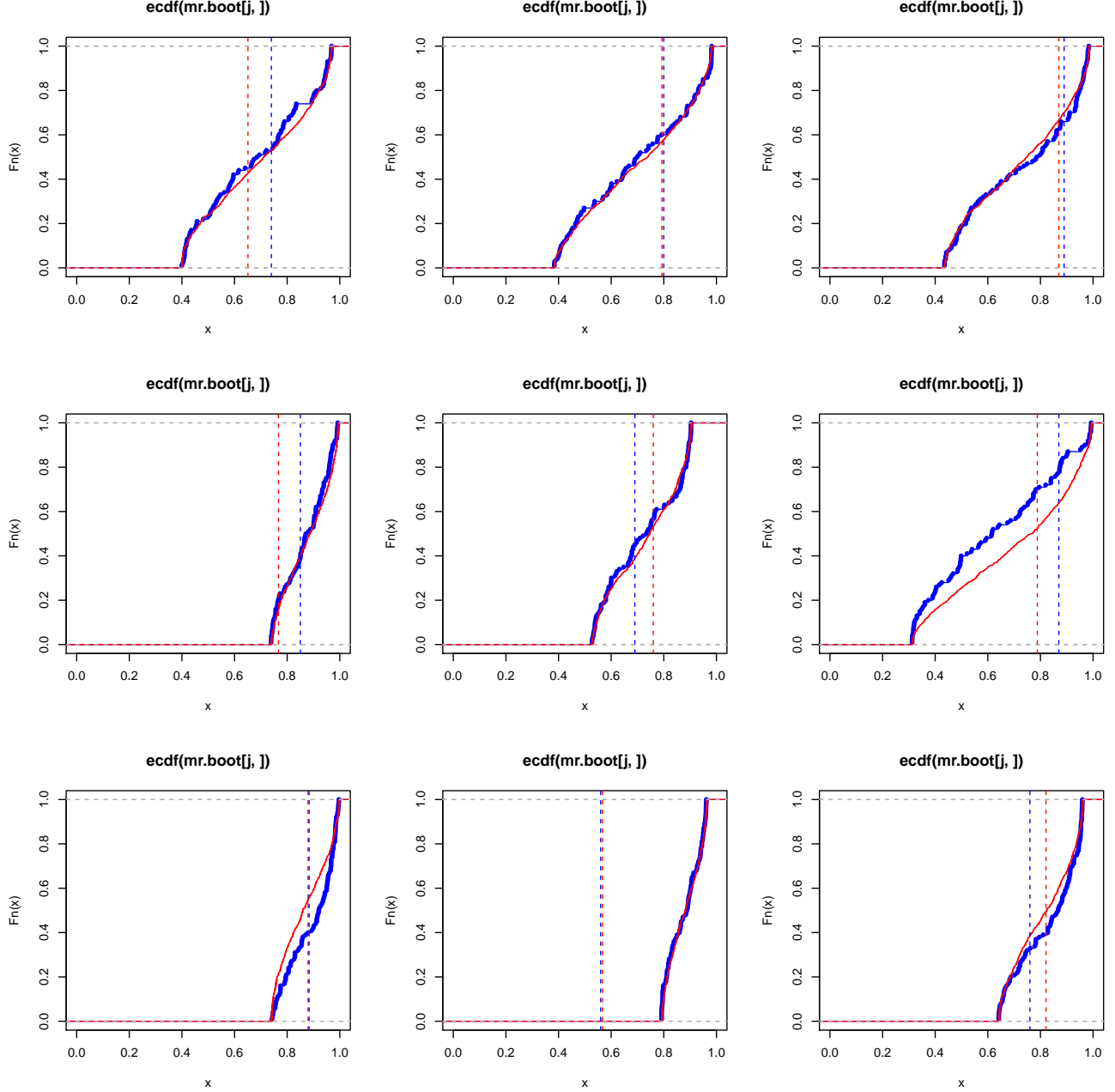


Figure 2: Empirical cumulative distribution of $\rho(s, t)$ for several randomly chosen s and t with 100 (blue) and 1000 (red) surrogates. Observed values are indicated by vertical dashed lines.

- the Benjamini-Hochberg-adjustment of p-values led from 12% $p\text{-values} < 0.1$ to 2% $p\text{-values} < 0.1$ in the Cormoran/Heron/Egret case with 100 surrogates and from 11% to 0.2% with 1000 surrogates. The important reduction in significant values after correction could perhaps come from the large number of comparisons due to the size of the grid. Indeed, when using a fine-scaled image grid, there was 178929 (423 scales \times 423 locations) comparisons. It appears that only the p-values equal to 0 when not adjusted appear as eventually significant with a correction (Fig. 3), which raises some questions as to the whole process. The False Discovery Rate correction therefore appears of dubious utility here, since this is essentially picking exactly zero p-values rather than highlighting zones with low p-values.

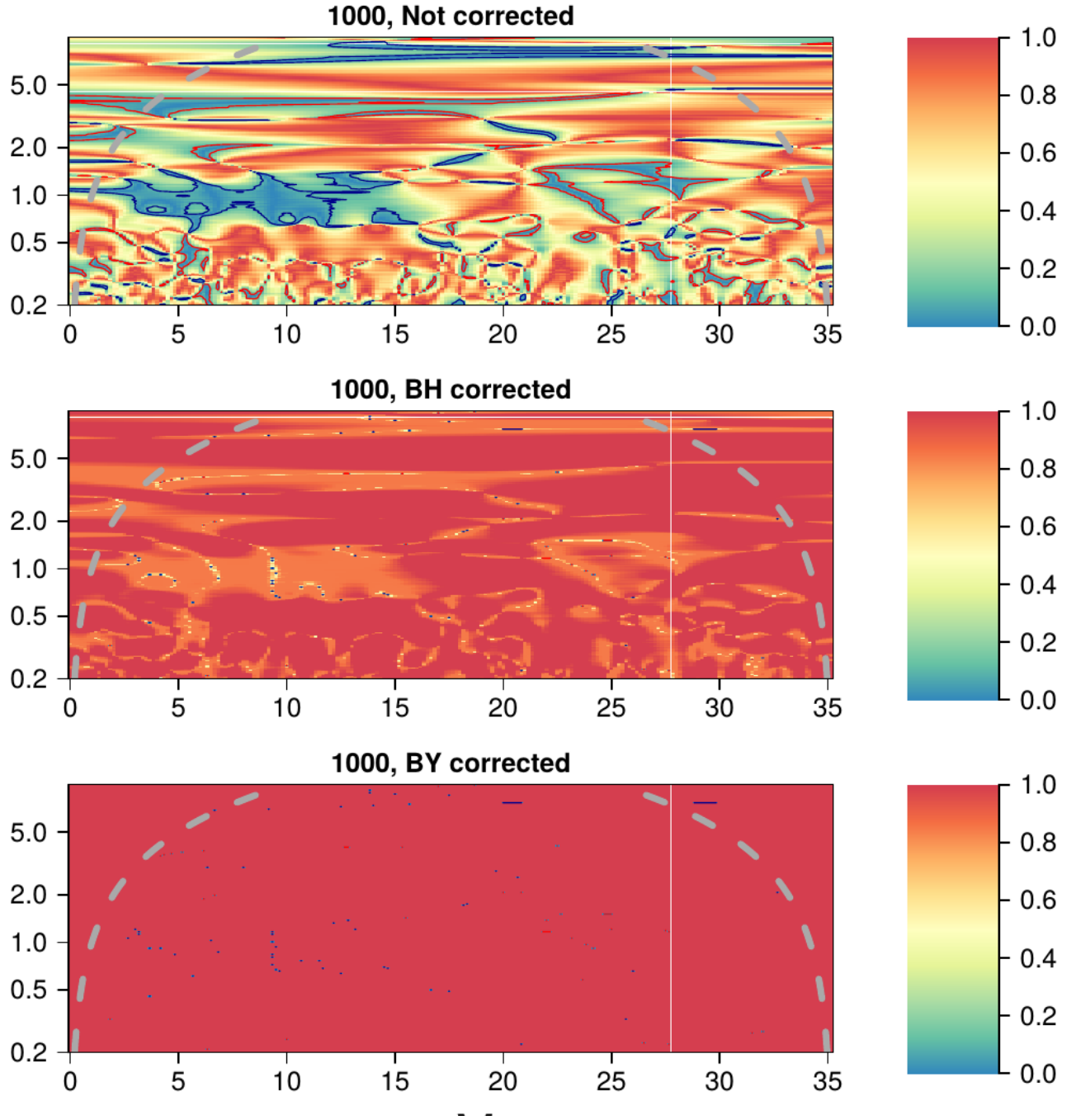


Figure 3: P-value images (i.e., each pixel in the grid is a p-value) for 1000 surrogates without FDR correction (top), with the BH correction (middle), and with BY correction (bottom) as recommended by Keitt (2008). The bottom and middle panel show “scattered significance” through the blue dots, rather than contours of high- and low-value wmr.

- Standardizing $((x - \text{mean}) / \text{sd})$ the time-series also led to a change in observed wmr values (see for instance the wavelet/wader wavelet plots in which the synchronizing effect of 2006 is much stronger with the scaled time-series). However, this suggestion was not a mere technical fix but rather implied a change in the meaning of the index. Normalised wavelet modulus ratios are not comparable to the Gross index and the notion of compensation they examine has no ecosystem-level meaning (i.e., rare species have equal weight than common species in normalised analyses, while common species are actually the species that matters most for compensation in real systems). Thus these normalised analyses have not been retained.

What we also checked and eventually decided

- The line contours that are drawn in the wmr plots do correspond well to small p-values; see Fig. 3: these are not a mere plotting bug.
- While removing the FDR correction (but still using Keitt 2008’s method), we changed the number of scales (from 423 to 100) and locations (from 423 to 212) to limit the number of computations. This had the benefit of reducing the noise in significance contours (as well as the size of the images), see Fig. 4. However, illogical contours still appeared in Fig. 4, in the sense that a zone with index around 0.5 was found to be of significantly low synchrony, which essentially suggested that was is an issue with the null hypothesis (i.e., we can reject too easily the null in places we should not be able to).

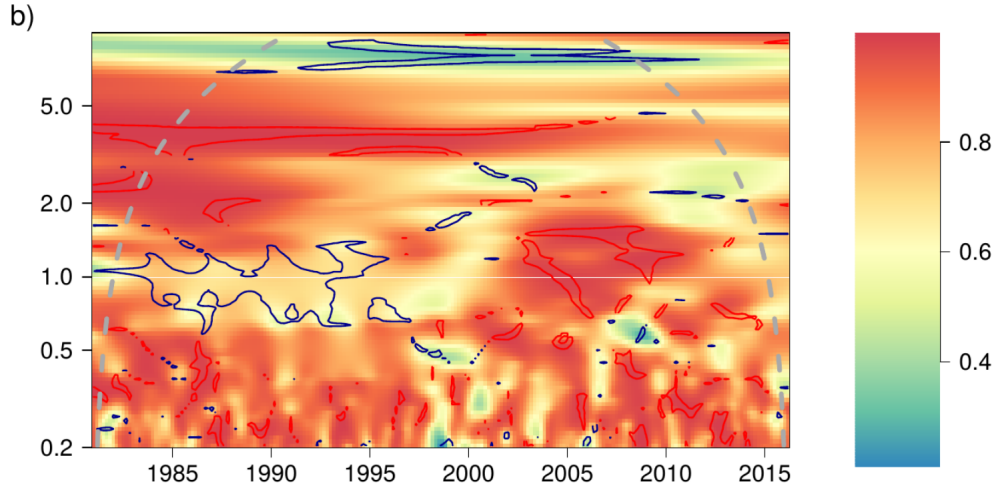


Figure 4: Wavelet analysis on a coarse grid, with 1000 surrogates from Keitt (2008) and no FDR-correction

- To fix the problems of null hypotheses and associated p-values, that could not be fixed by simply changing the grid size or the FDR correction, we constructed a new null hypothesis. Instead of phase-randomization of the wavelets sensu Keitt (2008), we built new time series (that is, in the time domain) with the IAAFT method and then built the p-values using each image pixel’s surrogate distribution, each surrogate $\rho_j(s, t)$ being obtained through wavelet transformation and wavelet modulus ratio computation of the IAAFT-obtained time series. This could lead to less significant zones than alternative methods (Cazelles et al. 2014, Journal of the Royal Society Interface), but provides a suitable and understandable null hypothesis: the Fourier spectrum is conserved while all other cross-correlations are lost. In Fig. 5, the number of significant pixels has not changed markedly, but all significant $\rho(s, t)$ are now in one block of obvious compensation instead of being spread out, which is desirable behaviour. In Fig. 5, we kept the Benjamini-Hochberg FDR correction but we eventually removed that correction, and the obtained graphs, on both real and simulated data were found to be satisfactory when using 1000 surrogates.

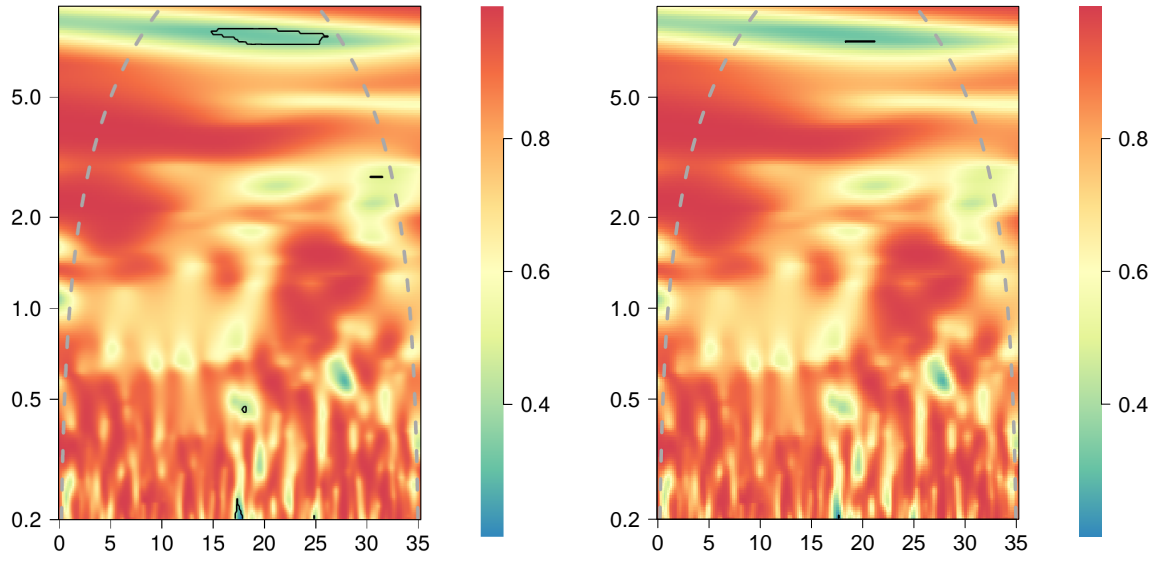


Figure 5: Wavelet analysis with 100 IAAFT surrogates (left) and 1000 IAAFT surrogates (right). These graphs include a BH-correction for the false discovery rate.

Conclusion

Although the null hypothesis testing proposed by Keitt (2008) does seem to provide sensible results in some cases, it was both difficult to

- understand theoretically the null hypothesis
- make it work in practice on our data

We have therefore used instead the well-known IAAFT procedure which has a clearly defined null hypothesis and provided sensible results on both simulated and real data.