POINT BY POINT RESPONSE RJournal 2021-156

"FMM: An R Package for Modeling Rhythmic Patterns in Oscillatory Systems"

Dear Dr. Catherine Hurley,

Thank you for giving us the opportunity to submit a revised draft of our manuscript titled “FMM: An R Package for Modeling Rhythmic Patterns in Oscillatory Systems” to The R Journal. We appreciate the time that you, the associate editor and the reviewers have dedicated to assess our manuscript. We are grateful for your constructive comments and suggestions. Bellow we have addressed all the concerns you raised. Changes to the text are highlighted in the revised manuscript.

Editors' comments:

Comment:

Along with the Associate Editor, I strongly urge you to move some of the examples to a package vignette, and shorten the current paper.

Response:

A vignette has been added in the latest version of our package. The example code in the “Basic usage of the FMM package” section has been reduced to a single example that illustrates the basic use of the package.

Reviewers' comments:

Reviewer #1:

Comment:

Overview

This paper and package describes a class of models for flexibly fitting rhythmic patterns. The paper is well-written and introduces some good illustrative examples. The methods implemented here are novel and are a useful tool in the R community. Below are my comments on the article and code.

Response:

Thank you very much for taking the time to review our manuscript.

Comment:

Comments on Article

1. The characterization of the class of FMM models at the start of the Section called "FMM Model" is a bit unclear. It would be preferred to have a concise and self-contained definition what is an FMM model in the paper rather than referencing it away.

Response:

Although the purpose of the manuscript is not to define the FMM models, we agree with reviewer #1 that it is important to characterize them to understand the work. Following reviewer #1's suggestion, we have included in the first paragraph of the "FMM Model" section of the revised manuscript, a brief summary sentence to introduce the FMM models.

Comment:

2. Step 2 of the estimation algorithm on page 3 should be described in more detail. For instance, what is the objective function in this context? Is the algorithm guaranteed to converge to a global optimum?

Response:

Following reviewer #1's suggestion, the step 2 of the estimation algorithm has been rewritten in the revised manuscript, including the objective function.

Although the convergence to a global optimum is not guaranteed, our experience both addressing problems from very disparate fields and demanding simulation experiments, indicates that the algorithm successfully avoids local optimums, most of the times.

This step is crucial for the optimization of the alpha and omega parameters, since the estimates for them are obtained by a grid-search, while all other parameters are obtained by least squares. Precisely, for challenging situations, we allow the user to modify the grid resolution and make a more exhaustive search. In particular, lengthAlphaGrid and lengthOmegaGrid arguments of fitFMM() function can be modified to specify the grid resolution and numReps to refine the search.

Comment:

Comments on Package

- No testing apparent as part of the package

Response:

Up to now, we have always tested our functions in an ad hoc way every time we made changes. In the latest version of the package, we have included a formal automated tests for the fitFMM() function because we agree with reviewer #1 that the testing process is more efficient using unit tests.

Comment:

- Documentation is good, but there is no vignette, no webpage, and the README is very sparse. I'd encourage you to take some of the content from this paper and include a vignette/more extensive readme.

Response:

Following reviewer #1's suggestion, we have added a vignette in the latest version of our package.

Comment:

- The code itself is easy to read and commented appropriately and in a useful way.

Response:

Thank you very much for your comment.

Comment:

1. The main package functions use the S4 class system. It seems that the fitting functions create the objects and shove them into the S4 slots, and the main package methods simply return those slots. Can you comment on what the motivation is for using the S4 system?

Response:

The S4 system allows the formal definition of classes, exactly typing the slots of the class. A more formal object-oriented programming paradigm has several benefits, such as the slots type validation or class inheritance. Furthermore, the use of an S4 class is closely related to the future developments of the FMM package. On the one hand, class inheritance will be key to implementing new specific FMM models that require the definition of new classes. On the other

hand, the S4 object system facilitates the connection between R and C++ with Rcpp, which could be essential to improve the package's performance.

Comment:

2. Why not use the plot generic function instead of plotFMM? This was surprising to me when I was testing out the package as many of the other generics were defined.

Response:

The definition of plotFMM() rather than using the generic plot is a programming style decision. We consider that this function adds much functionality, such as showing the FMM components, plotting the model over several periods or using the ggplot2 graphics engine, to be considered just an implementation of a generic.

Comment:

3. The showTime = TRUE default is a bit annoying (e.g., if running simulations) and there is no message saying what the time is.

Response:

At previous stages of the package development and implementation process it was important for us to control the computation time. We agree with the review #1 that in some cases, it may be unnecessary and also inconvenient, so the default value of the argument showTime of fitFMM() function has been changed to FALSE. This involves some minor changes to the revised manuscript, specifically in Table 2; in the example code in the "Basic usage of the FMM package"; and in the "Real data analysis using the FMM package" sections, where this argument has been removed.

Comment:

4. The underlying fit methods use optim, but there does not appear to be any way for a user to modify or query those optim calls. With these models, I guess that convergence can sometimes be an issue, so it would be useful to be able to more directly work with parameters of the optimization part of the estimation.

Response:

This observation is correct. Although we agree with reviewer #1 that the convergence can be an important issue, in our experience, optim() works reasonably well in the wide range of explored applications. As already we have mentioned in the “Comment on Article” #2, the most important step for the convergence issue is the grid-search of the alpha and omega parameters which, at this time, the user can controlled by three arguments of the fitFMM() function. We plan to improve our package by adding more effective solutions for this point, both to enable the user-defined optimization functions and to increase the convergence speed. We have added a short sentence in the last paragraph of the “Summary” section to address this concern.

Comment:

5. Is it possible to plot the figures with the x axis as the original time scale?

Response:

The plotFMM() function allows to create ggplot objects by specifying use_ggplot2 = TRUE. In order to modify the features of a ggplot object, the ‘+’ operator can be conveniently used to draw additional layers such as labels, annotations, scales, axis ticks, legends, themes, facets, and more. In particular, to plot the original time scale on the x-axis, the default scale aesthetics can be modified using scale_x_continuous(). We agree with the reviewer #1 that further elaborating on this point would be helpful, so the required code to make this transformation has been added to the plotFMM() function documentation as an example.

Reviewer #2:

Comment:

Overview:

This is a very well written paper.

I am not a statistics expert, and trust that the Rueda 2019 article has been a good benchmark for the statistical side of the. However, as an applied time series researcher and R package maintainer myself I can see the value of this package to the R community.

Response:

Thank you very much for taking the time to review our manuscript.

Comment:

Article:

The article is clearly written and I was able to run and understand the various examples. A few minor comments:

Response:

Thank you very much for your comment.

Comment:

1. Please cite the GitHub repository in the text such that users are encouraged to report bugs in the public space, rather than via e-mail. Further, it would help users find out about the implementation.

Response:

Following reviewer #2's suggestion, we have added a brief paragraph to the "Summary" section citing the GitHub repository.

Comment:

2. In case FMM has already been used by others outside the development team then consider mentioning in the text. There is no better way to showcase the value of a package to the community then to refer to other people in the community using FMM. I understand that for a new package it takes time before this use-cases emerge, so no problem if the authors have no such examples.

Response:

We fully agree with the reviewer #2's comment. Although the FMM methodology is of recent development, as far as we know other research groups have successfully started to use both the methodology and the package in their ongoing research. This is the case, for example of the Division of Sleep and Circadian Disorders of Harvard Medical School (USA); Cronolab at the University of Murcia (Spain); or the Biomedical Engineering Research Group from the University of Valladolid (Spain). We are certain that new use-case works will be published soon, but at this moment it is too early to cite them in our manuscript.

Comment:

Package:

The source code which I was able to find here: https://github.com/alexARC26/FMM, looks overall good. It has documentation for all the functions, the object names appear to be chosen with care, and code is well structured in functions that are not too long. A few comments:

Response:

Thank you very much for your comment.

Comment:

1. In line with my comments above it would be good to have a link to the GitHub repository in your DESCRIPTION file

Response:

Following reviewer #2's suggestion, we have added a link to the GitHub repository in the DESCRIPTION file.

Comment:

2. I miss a changelog file (./inst/NEW.Rd), which would help to clarify on CRAN how the package has changed across version.

Response:

We agree with the reviewer #2 that it is very important to document the changes that occurred across the package versions. Even with a clear commit history through the GitHub website, having a text version of changes can be very useful. For this reason, we are maintaining our package changelog in a NEWS.md file in the top level folder of our package.

Comment:

3. I miss an integration test, e.g. with GitHub Actions, to ensure that it is always clear that the latest version of FMM can be installed.

Response:

Following reviewer #2's suggestion, we have added an integration test with GitHub Actions.

Comment:

4. I miss unit-tests to help protect functional consistency. You can do this with R package testthat. As a software developer I would be anxious to contribute to this

repository as there is currently no way to verify whether my contribution breaks existing functionality.

Response:

We agree with reviewer #2 that making the testing process more formal would facilitate contributions from other users and developers. Therefore, following the suggestion, the latest version of the package includes unit tests for the fitFMM() function.

Thank you for pointing this out. We will be happy to have your valuable contributions to improve the functionality of our package.

Comment:

5. Once this paper is accepted, you may want to add a citation file to your R package.

Response:

We thank the reviewer #2 for this suggestion. We will add the citation as soon as our paper would be accepted.