

The paper describes a novel regression method to perform parameter inference for Approximate Bayesian Computation (ABC). Several regression approaches have been developed for ABC including neural networks and various forms of linear regression. The authors propose to consider a random forest approach, which has been proposed by Meinshausen (2016) to perform quantile regression. The numerical implementation of ABC with random forest is based on the RANGER R package that provides a fast implementation of random forest for high-dimensional data.

According to my reading of the manuscript, there are 3 main advantages when using random forest (RF) for parameter inference with ABC. The first advantage is that RF can handle many summary statistics and that dimension reduction is not needed because of attractive properties of RF. I perfectly understand this argument. However, I think that it is purely theoretical at the moment because the results do not provide support for this. Simulation results where uninformative summary statistics are added and where the properties of different ABC approaches (without dimension reduction) are compared may be worthwhile to report.

The second advantage concerns the coverage properties obtained with random forests. If correct, 95% posterior credibility intervals (C.I.) should contain 95% of the parameter values used in simulations. Figure 5 shows that posterior C.I. obtained with rejection are too large compared to other methods. By contrast, C.I. obtained with regression methods have been shrunken. However, the shrinkage can be excessive for the smallest tolerance rates, with coverage values that can be equal to 85% instead of the expected 95% value. The attractive property of RF is that C.I. have been shrunken but the coverage is of 100% resulting in a conservative decision about parameter values.

The last advantage is that no hyperparameter should be chosen. It is a parameter free approach, which is desirable because of the difficulty of choosing an appropriate acceptance rate.

The main drawback of the proposed approach concerns joint parameter inference. There are many settings where the joint parameter distribution is of interest and the proposed approach cannot provide joint posterior distribution. Estimating the covariance is not enough information because the joint posterior distribution may not be Gaussian.

Except for my remark about dimension reduction, I have no other major remarks that should be addressed. The paper nicely shows that modern machine learning approaches are useful for ABC and more generally for parameter inference in ecology and evolution. My other remarks listed below are minor.

Minor remarks

*Please indicate, if you consider homoscedastic or heteroscedastic adjustments when considering regression adjustment approaches

*It would nice in the R package to provide histograms or density distributions to visualize posterior distribution. Visualization routines are useful when performing data analysis.

The paper "Saulnier, Emma, Olivier Gascuel, and Samuel Alizon. "Inferring epidemiological parameters from phylogenies using regression-ABC: A comparative study." PLoS computational biology 13.3 (2017): e1005416." contains nice comparison results about accuracy of ABC estimators with regression adjustment. It could be discussed or cited in the present ms.

*Introduction

"is based on the tool of random forests" ->"is based on random forests"

I do not know what is the calibration side of RF and readers might not know either.

*ABC for parameter inference section

"calibration of the ABC algorithm" should be defined. It seems to me but I am not sure that the word calibration here refers to something different than the calibration of the posterior distribution. The same word should not be used with different meanings to avoid confusion.

*Alternative variance approximation

I do not understand the new approximation because I do not know what τ_{oob} is.

*A new R package

"Out R library" -> "Our R library" ?

*Results section

"Ridge regression" ->"ridge regression"

Figure S3 and corresponding text. Plotting cumulative distribution function is not suited to visualize differences between distributions. Density distribution functions should be preferred.

Figure 3: The caption should contain more details to understand what boxplots represent.

"greatly suffers" is not colloquial enough.

*Discussion

I would not mention the covariance approximation because it seems that it is still preliminary work.

Uniformly optimal is difficult to understand without being defined.

What are ancillary covariates?

*References

Please check the references

Csilléry should be Csilléry, Blum M should be BLUM MGB, Maples is not the only author of his paper, Wright and Ziegler paper has now been published in Journal of Stat Software...