Authors' Responses to Reviewer's Comments

Manuscript id 2021-162
*OTrecod: An R Package For Data Fusion Using Optimal
Transportation Theory*

Article status reminder :

1st submission: September 23, 2021
2nd submission (major corrections): July 25, 2022
3rd submission (minor corrections): October, 2022

Authors thank the Editor and associated reviewers for their global appreciations and for giving them the opportunity to submit a revised draft of the manuscript called *"OTrecod: An R Package For Data Fusion Using Optimal Transportation Theory"* for publication in R-Journal.

Following the minor corrections raised in the email dated 3 October 2022, please find below our point-by-point responses. In this document, the two reviewer's comments are recalled in bold font while our answers are in normal font.

The corresponding revisions in the manuscript are highlighted with a yellow font in a joint pdf document called "revised_article.pdf" (section *bibliography*).

Reviewer's comments and related answers

1st Reviewer's comment —

Paper revision looks good.
The code fails here:

```
> verif_outc1 = verif_OT(outc1, group.class = TRUE,
                            ordinal = FALSE,
# Error in verif_OT(outc1, group.class = TRUE,
#                   ordinal = FALSE, stab.prob = TRUE,
#        : unused argument (group.class = TRUE)
```

Using the CRAN version of the package. RJ policy is that the package has to be on CRAN.

Our response —

The new version of the OTrecod package (version 0.1.2) is now available on CRAN since October 5, 2022 as described in:
https://cran.r-project.org/web/packages/OTrecod/index.html
This update takes into account all the modifications proposed to provide positive issues to reviewers' comments in July 2022 (date of the 2nd submission of the paper).

Until October 2022, this last version was only available on GitHub at:
https://github.com/otrecoding/OTrecod/tree/master/R, the development web site of the package.

As mentioned earlier, the update of the package to CRAN was voluntarily delayed during the review of the second submission of the article (to avoid too frequent updates of the package to CRAN).

Installing and loading this new version of the package from CRAN now solve the mentioned error met on the verif_OT function, as described below:

```
> install.packages("OTrecod")
> libname(OTrecod)
>
> ...
>
> # Modele 1 (p15)
> outc1 = OT_outcome(bdd_ex, quanti = 1, ordinal = 2:6,
+                    dist.choice = "E",
+                    indiv.method = "sequential",
```

```
+                        which.DB = "B")
#----------------------------------------
# OT PROCEDURE in progress ...
#----------------------------------------
# Type                     = OUTCOME
# Distance                 = Euclidean
# Percent closest knn      = 100%
# Relaxation parameter     = NO
# Relaxation value         = 0
# Individual pred process  = Sequential
#DB imputed                = B
# ----------------------------------------
>
> ...
>
> ## Validation of the data fusion using verif_OT (p18)
>
> # Quality criteria for outc1:
> verif_outc1 = verif_OT(outc1, group.class = TRUE,
+                        ordinal = FALSE,
+                        stab.prob = TRUE,
+                        min.neigb = 5)
>
>   verif_outc1$nb.profil
#[1] 27
> ...
```

2nd Reviewer's comment —

And the paper does not latex. There appears to be an error somewhere in these two references:

```
\bibitem[Vantaggi(200]{Vantaggi08}
B.~Vantaggi.
\newblock Statistical matching of multiple sources:
          A look through coherence.
\newblock \emph{Int. J. Approx. Reason.},
            49\penalty0 (3):
                \penalty0 701--711, 2008.
\newblock URL
  \url{http://dblp.uni-trier.de/db/journals/ijar/
                            ijar49.html#Vantaggi08}.

\bibitem[Zeileis et~al.(200Zeileis, Hothorn, and
                                   Hornik]{RF2}
A.~Zeileis, T.~Hothorn, and K.~Hornik.
\newblock Model-based recursive partitioning.
\newblock \emph{Journal of Computational and Graphical
              Statistics},
17\penalty0
(2):\penalty0 492--514, 2008.
```

Our response —

These two references have been verified and new related bibtexs have replaced the previous ones in the new version of the *RJwrapper.pdf* document: these are the only changes made between the two versions.

We confirm that the pdf document compiles from a LaTeX document via the overleaf platform. Here are the old and new references related to these two articles:

OLD REFERENCES:

1. Vantaggi. Statistical matching of multiple sources: A look through coherence. Int. J. Approx. Reason., 49 (3):701–711, 2008.
   URL `http://dblp.uni-trier.de/db/journals/ijar/ijar49.html#V antaggi08`. [p1]

2. Zeileis, T. Hothorn, and K. Hornik. Model-based recursive partitioning. Journal of Computational and Graphical Statistics, 17(2):492–514, 2008. [p12]

NEW REFERENCES:

1. Vantaggi. Statistical matching of multiple sources: A look through coherence. Int. J. Approx. Reasoning, 49:701–711, 11 2008.
   doi: 10.1016/j.ijar.2008.07.005. [p1]

2. Zeileis, T. Hothorn, and K. Hornik. Model-based recursive partitioning. Journal of Computational and Graphical Statistics, 17(2):492–514, 2008.
   doi: 10.1198/106186008X319331.
   URL https://doi.org/10.1198/106186008X319331. [p12]