Bad smells in reviewers' reports?
Text-mining the MDPI Open Peer Review Corpus

Gilles Hubert, Guillaume Cabanac, Cyril Labbé

Paper Mills do exist!

What about Review Mills?

Review Mill at MDPI, by María de los Ángeles Oviedo-García, January 12, 2024

`https://predatoryjournals.org/news/f/review-mill-at-mdpi`

> *[…] a set of 85 very similar review reports in 23 journals published by MDPI [...] a standard text was copied and pasted to every manuscript regardless its content, following two patterns (type A and type B).*

The manuscript "**Analysis of Alkylphenol Ethoxylates in Tea by SPE Coupled to UPLC-MS/MS**" is a very interesting work. This work describes the Alkylphenol ethoxylates (APEOs) is a kind of widely used non-ionic surfactant, as an adjuvant in pesticide formulation; which is considered to cause endocrine disrupting effect. In the current study, detection method for APEOs residue in tea was established, on the basis of SPE (sol-id-phase extraction) for simultaneous analysis of nonylphenol ethoxylates (NPEOs) and oc-tylphenol ethoxylates (OPEOs) by UPLC-MS/MS. In the spiked concentrations from 0.024 µg/kg to 125.38 µg/kg for 36 monomers of APEOs (nEO = 3-20), the recoveries of APEOs ranged from 70.3-110.7% with RSD ≤16.9%, except OPEO20 (61.8%) and NPEO20 (62.9%). The results are consistent with the data and figures presented in the manuscript. While I believe this topic is of great interest to our readers, I think it needs major revision before it is ready for publication. So, I recommend this manuscript for publication with major revisions.

1. In this manuscript, the authors did not explain the importance of the Alkylphenol Ethoxylates in the introduction part. The authors should explain the importance of Alkylphenol Ethoxylates.

2) Title: The title of the manuscript is not impressive. It should be modified or rewritten it.

3) Correct the following statement "The detection rate and the range of monomers of NPEOs was generally higher than that of OPEOs. The current study would provide a theoretical basis to guide the ra-tional use of APEOs as adjuvant in commercial pesticide production".

4) Keywords: There so many keywords and reduce them up to 5. So, modify the keywords.

5) Introduction part is not impressive. The references cited are very old. So, Improve it with some latest literature like 10.3390/biom12010083, 10.3390/ph15101164

6) The authors should explain the following statement with recent references, "The total concentration of OPEOs was between 0.057 µg/kg and 12.9 µg/kg. The monomers with the highest detection concentration were OPEO9, OPEO10, and OPEO11".

7) Add space betwen magnitude and unit. For example, in synthesis "21.96g" should be 21.96 g. Make the corrections throughout the manuscript regarding values and units.

8) The author should provide reason about this statement "50 tea samples in total were randomly collected from the market, including 24 white tea, 12 black tea, 5 dark tea, 4 green tea, 3 scented tea, and 3 oolong tea".

9) Comparison of the present results with other similar findings in the literature should be discussed in more detail. This is necessary in order to place this work together with other work in the field and to give more credibility to the present results.

10) Conclusion part is very long. Make it brief and improve by adding the results of your studies.

11) There are many grammatic mistakes. Improve the English grammar of the manuscript.

The manuscript "**Green dynamic kinetic resolution-stereoselective acylation of secondary alcohols by enzyme-assisted ruthenium complexes**" is very interesting work. In this study, The resulting catalytic system of the ruthenium racemization catalysts and enzymatic acylation led to chiral esters obtained by dynamic kinetic resolution. The immobilized catalytic system in the ionic liquid expresses the same activity and selectivity as the homogeneous system while allowing for convenient separation of the desired products. Additionally, the efficient reuse of the catalytic system has been demon strated fulfilling the paradigm of green processes. I believe this topic is of great interest to our reader, I think it needs minor revision before it is ready for publishing. So, I strongly recommend this manuscript for publication in the Journal of Catalysts with major revisions.

1. In this manuscript, the authors did not explain the importance of ruthenium complexes in the introduction part. The authors should explain the importance of ruthenium complexes.

2. The author should provide reason about this statement with recent references "The immobilized catalytic system in the ionic liquid expresses the same activity and selectivity as the homogeneous system while allowing for convenient separation of the desired products".

3. Introduction part is not impressive and systematic. Cite the following articles in the introduction part. (i) 10.3390/molecules27196580 (ii) 10.3389/fchem.2022.995820

4. The authors should justify the following statement "The possibility of creating an efficient heterogeneous catalyst in which the ruthenium complex is non-covalently attached to support was also sought due to the simplicity of the system and the least influence of the catalytic properties".

5. The authors should explain regarding the recent literature why "The studies showed that neither the MWCNT support nor the IL affects the activity of the ruthenium catalyst".

6. The authors should explain the statement with reference to the literature. "This is because the catalyst is confined in a small volume of the IL as opposed to being dispersed throughout the entire volume in the absence of the addition of an IL".

7. Comparison of the present results with other similar findings in the literature should be discussed in more detail. This is necessary in order to place this work together with other work in the field and to give more credibility to the present results.

8. The conclusion part is very weak. Improve by adding the results of your studies.

9. The authors should pay more attention to the English grammar, and the abbreviation of journal names in Ref.

Is it possible to automatically detect *bad smells*?

Analysed dataset

- MDPI Open Peer Review Corpus 2 - Version 2.0
- Scraped by Miłkowski et al. (2023, https://doi.org/10.18150/shkp7b)
- 170 GB
- 135,437 articles and their associated reports (between 2011 and 2022)

Chosen approach

- Focus on *Round 1* reports in plain text (i.e., excluding those uploaded as attached files)
  → 320,380 reports (/ 353,131)
- Statistics on report length
- « Strong » inter-textual similarity between reports
- Common word sequences (10+)
- Identified references (DOIs) suggested by reviewers with regular expressions

Result 1: Report length
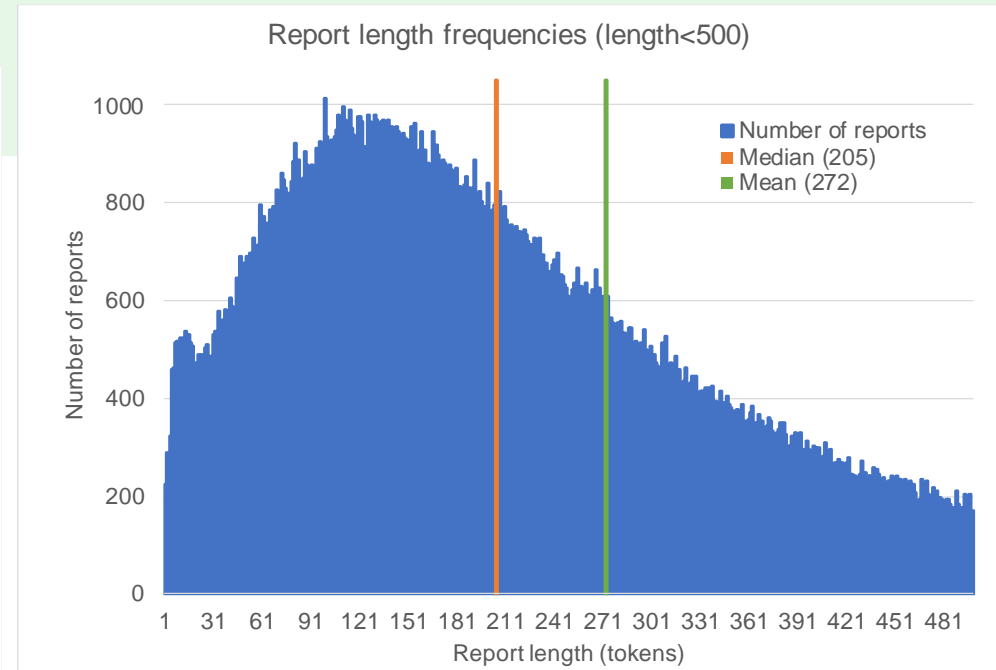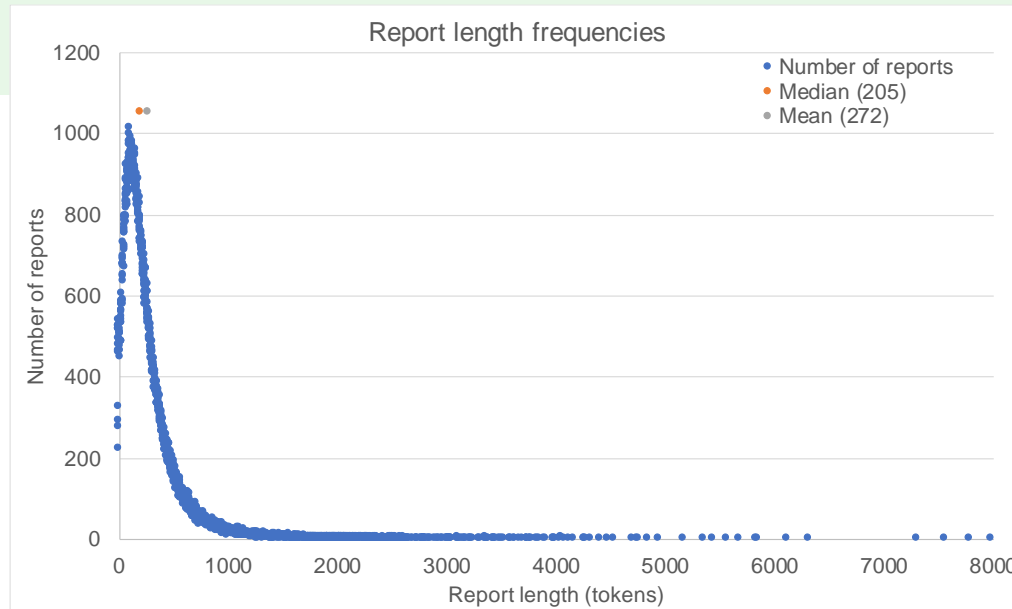
Some very very long reports

- For example, Report 3 for paper `pr10051002` is the same text pasted 50 times.

Quite a lot of short reports

- 223 'nanoscale' reports consist of 1 word only (`'accept','none','Nil' or 'N.A.'`)
- 'Micro' reports of less than 20 words account for 2.7% of the dataset.

'Micro' reports

Paper admsci12030097

Reviewer 1 Report

Update the literature review/references.

Author Response

The author had updated literature related to the [30]

Paper aerospace8070179

Reviewer 1 Report

This paper is good and acceptable quality.

Author Response

Thank you for your support.

Paper adolescents1010001

Reviewer 1 Report

Authors should be congratulated.

Author Response

Many thanks for yor kind comments

Technical issues: bogus scraping and editorial management

Paper aerospace9100612

In the dataset

Reviewer 1:

Summary:

On the website

Round 1

Reviewer 1 Report

Summary:

The paper proposed a new complexity metric, which is an adaptation of Sinh
and changes are supposed to make the metric suitable for the use with MBS
in three case studies and the results are discussed.
Feedback:
Positive: + the paper addresses a timely topic and uses the right/suitable app
provides a comprehensive overview and the case studies add to the substan
addressed: - the differences between the new metric and the one it is based
advantageous and necessary - a general lack of comparison to recent literatu
distinction regarding the assumptions made is not clear regarding the sours (
lack of explanation of disadvantages/boundaries - format errors, missing cap
pg. 4,7, ...), overlapping text (e.g. pg. 6) - very very wordy case study explana
the results; the setups do not need to be explained in such an extensive way

Paper agriculture11080744

In the dataset

Reviewer 2:

On the website

Round 1

Reviewer 1 Report

Good work addressing comments

Reviewer 2 Report

This manuscript is a resubmission of an earlier submission. The following is a list of the peer review reports and author responses from that submission.

Round 1

Reviewer 1 Report

The work is prepared quite diligently. I have no major remarks. However, I have a few observations:

Result 2: Inter-textual similarities between reports

Regarding Report-report similarities:

- 48,628 pairs of reports with inter-textual similarity >= 90%, 0.8 % of reports

- 74,179 pairs of reports with inter-textual similarity >= 75%, 1.6 % of reports

Very short reports

*Only first 300 nodes are displayed
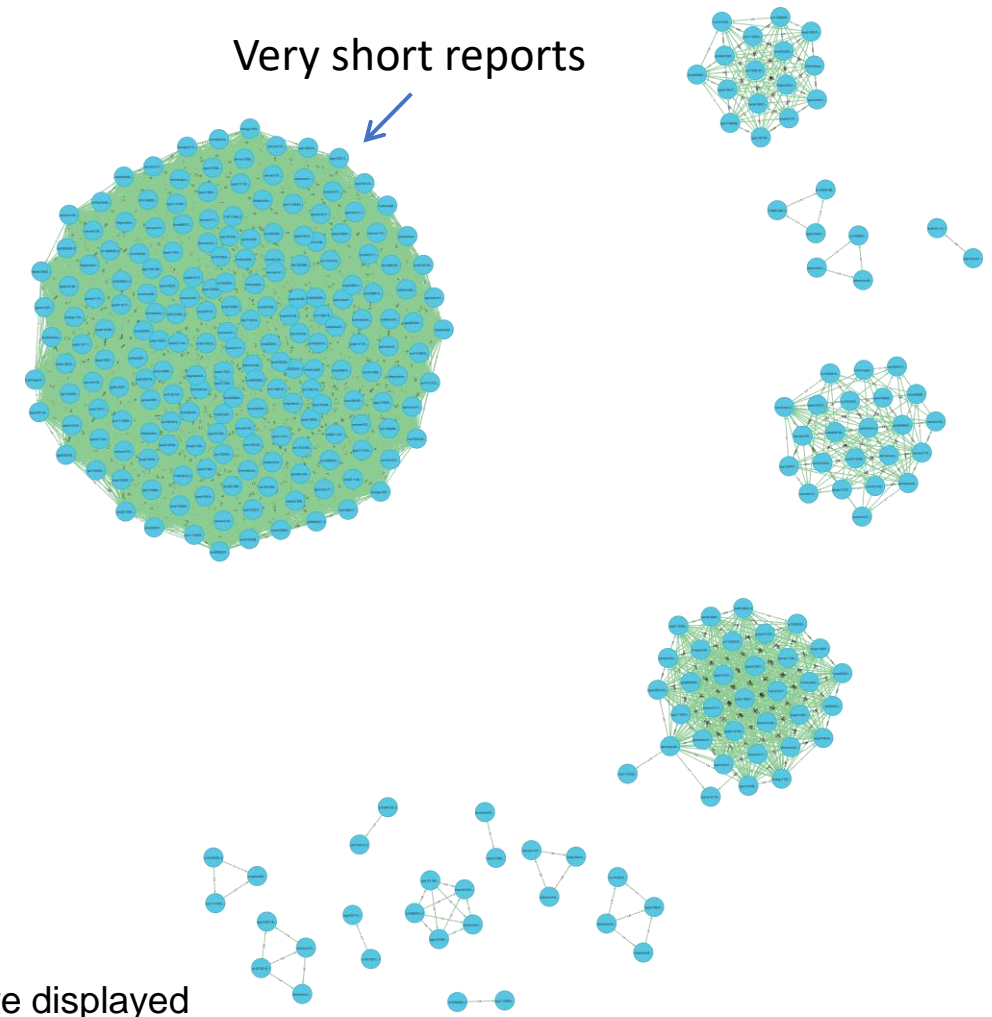
Report–report similarities show identical reports for the same paper

ijms232112760

Reviewer 2 Report

In this article, Washburn et al. assess the activation of the complement system in response to the allograft of Sertoli cells. They observed that the Sertoli cells showed substantial protection against humoral immunity. The microarray experiment with the mouse Sertoli cells showed enhanced expression of complement inhibitory proteins (CIPs). They conclude by suggesting that since Sertoli cells exhibit protection from complement-mediated tissue rejection, it could be a novel strategy in diabetes for enhancing the success of islet grafts.

Overall, the study seems to highlight the protective mechanism exhibited by Sertoli cells' graft in treating diabetes. The study is in line with a previous study (Fallarino et al., 2009) and one from the same group (Kaur et al., 2018). Although the findings are interesting and certainly pave the path for exciting avenues, at the current stage, several factors limit the enthusiasm in this study:

…

Reviewer 3 Report

In this article, Washburn et al. assess the activation of the complement system in response to the allograft of Sertoli cells. They observed that the Sertoli cells showed substantial protection against humoral immunity. The microarray experiment with the mouse Sertoli cells showed enhanced expression of complement inhibitory proteins (CIPs). They conclude by suggesting that since Sertoli cells exhibit protection from complement-mediated tissue rejection, it could be a novel strategy in diabetes for enhancing the success of islet grafts.

Overall, the study seems to highlight the protective mechanism exhibited by Sertoli cells' graft in treating diabetes. The study is in line with a previous study (Fallarino et al., 2009) and one from the same group (Kaur et al., 2018). Although the findings are interesting and certainly pave the path for exciting avenues, at the current stage, several factors limit the enthusiasm in this study:

Report–report similarities show groups of reports almost identical

a15050163

Reviewer 1 Report

1. **Results:** Recommend to be **Major revisions**

This paper bridges the gap between routing networks and spatial geometry by a global matching of geometric points to routing networks. Although graph theory has already been introduced to spatial reasoning, current spatial database systems do not provide out-of-the-box routing on geometric points that are not matched on the graph. Methods that connect new reference locations to the graph render different routing results. Moreover, current solutions break reasoning down to local analysis.

This paper is with minor merits for *Algorithms,* i.e., poor writing skills and lacking of insight analysis, it requires major revisions.

Firstly, for Section 1, authors should provide the comments of the cited papers after introducing each relevant work. What readers require is, by convinced literature review, to understand the clear thinking/consideration why the proposed approach can reach more convinced results. This is the very contribution from authors. In addition, authors also should provide more sufficient critical literature review to indicate the drawbacks of existed approaches, then, well define the main stream of research direction, how did those previous studies perform? Employ which methodologies? Which problem still requires to be solved? Why is the proposed approach suitable to be used to solve the critical problem? We need more convinced literature reviews to indicate clearly the state-of-the-art development.

For Section 2, authors should introduce their proposed research framework more effective, i.e., some essential brief explanation vis-à-vis the text with a total research flowchart or framework diagram for each proposed algorithm to indicate how these employed models are working to receive the experimental results. It is difficult to understand how the proposed approaches are working.

For Sections 3 and 4, authors should use more alternative models as the benchmarking models, authors should also conduct some statistical test to ensure the superiority of the proposed approach, i.e., how could authors ensure that their results are superior to others? Meanwhile, authors also have to provide some insight discussion of the results.

a14050127

Reviewer 2 Report

1. **Results:** Recommend to be **Major revised**

This paper proposes the novel search operation for the neuroevolution of augmented topologies, namely the difference-based mutation. The implemented neuroevolution algorithm allows backward connections and loops in the topology, and uses a set of mutation operators, including connections merging and deletion. The experimental results prove that the newly developed operator delivers significant improvements to the classification quality in several cases, and allow finding better control algorithms.

It is with some merits for *Algorithms*, however, it requires some major revisions.

Firstly, for Section 1, authors should provide the comments of the cited papers after introducing each relevant work. What readers require is, by convinced literature review, to understand the clear thinking/consideration why the proposed approach can reach more convinced results. This is the very contribution from authors. In addition, authors also should provide more sufficient critical literature review to indicate the drawbacks of existed approaches, then, well define the main stream of research direction, how did those previous studies perform? Employ which methodologies? Which problem still requires to be solved? Why is the proposed approach suitable to be used to solve the critical problem? We need more convinced literature reviews to indicate clearly the state-of-the-art development.

For Section 2, authors should introduce their proposed research framework more effective, i.e., some essential brief explanation vis-à-vis the text with a total research flowchart or framework diagram for each proposed algorithm to indicate how these employed models are working to receive the experimental results. It is difficult to understand how the proposed approaches are working. For the employed data set, please provide more details illustration.

For Section 3, authors should use more alternative models as the benchmarking models, authors should also conduct some statistical test to ensure the superiority of the proposed approach, i.e., how could authors ensure that their results are superior to others? Meanwhile, authors also have to provide some insight discussion of the results. Authors can refer the following references for conducting statistical test.

a15050154

Reviewer 1 Report

1. **Results:** Recommend to be **Major revisions**

This paper proposes the modification of the quantum multi-swarm optimization algorithm for dynamic optimization problems. The modification implies using the search operators from differential evolution algorithm with certain probability to improve the algorithms search capabilities in dynamically changing environments. For algorithms testing the Generalized Moving Peaks Benchmark was used. The experiments performed in two benchmarks with different changes frequency have shown that the proposed modification with dynamically changing probability of applying differential operators allows better tracking of the changing environment of the dynamic optimization problems.

This paper is with none merits for *Algorithms,* i.e., poor writing skills and lacking of insight analysis, it requires major revisions.

Firstly, for Section 1, authors should provide the comments of the cited papers after introducing each relevant work. What readers require is, by convinced literature review, to understand the clear thinking/consideration why the proposed approach can reach more convinced results. This is the very contribution from authors. In addition, authors also should provide more sufficient critical literature review to indicate the drawbacks of existed approaches, then, well define the main stream of research direction, how did those previous studies perform? Employ which methodologies? Which problem still requires to be solved? Why is the proposed approach suitable to be used to solve the critical problem? We need more convinced literature reviews to indicate clearly the state-of-the-art development.

For Section 2, authors should introduce their proposed research framework more effective, i.e., some essential brief explanation vis-à-vis the text with a total research flowchart or framework diagram for each proposed algorithm to indicate how these employed models are working to receive the experimental results. It is difficult to understand how the proposed approaches are working.

For Section 3, authors should use more alternative models as the benchmarking models, authors should also conduct some statistical test to ensure the superiority of the proposed approach, i.e., how could authors ensure that their results are superior to others? Meanwhile, authors also have to provide some insight discussion of the results. Authors can refer the following references for conducting statistical test.

Result 3: Graph of reports sharing long sequence of words (10+)

*Only first 300 nodes are displayed
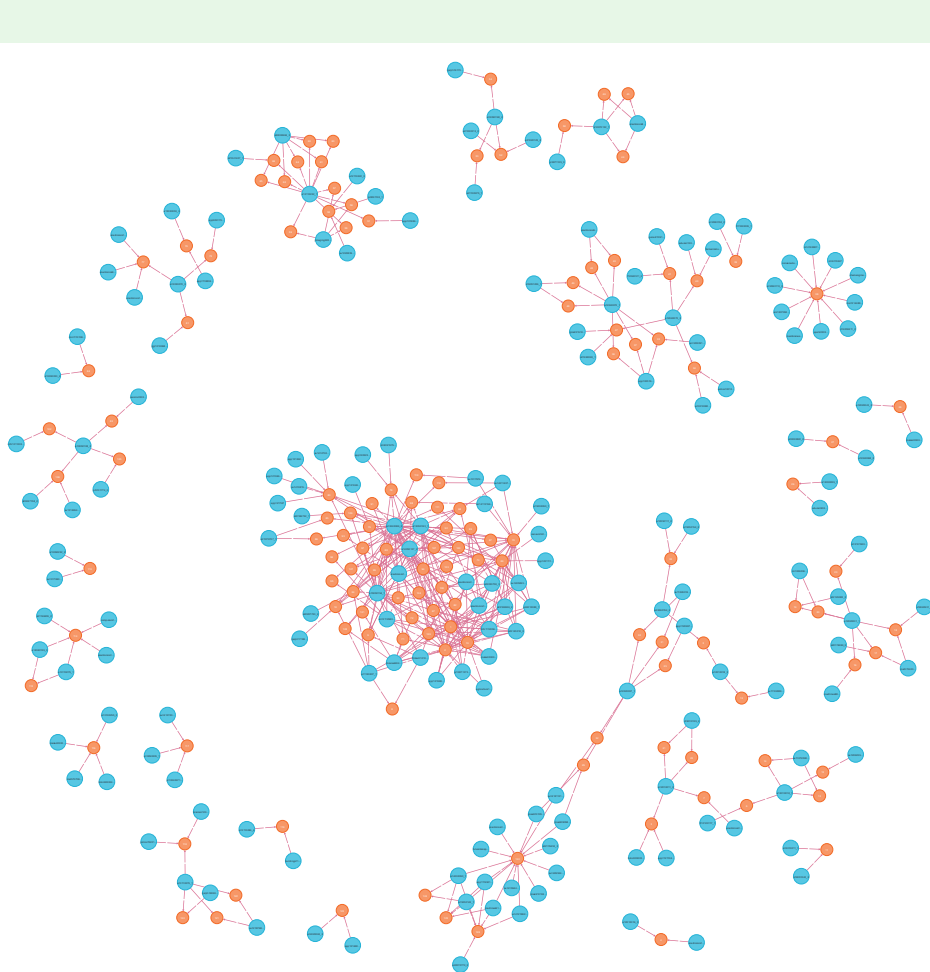
s20174734

Reviewer 1
This work is very interesting and presents a good scientific quality and could be relevant. It is well presented from the point of view of the Methodology and Results. Its publication is recommended. Just some aspects of improvement that the authors could consider: The abstract must be rewritten *[…]* more clearly. It is recommended *[…]* more clearly represented in figure 6 Bibliography is scarce and should be reviewed and updated to improve the quality and interest of readers and researchers Once these considerations have been made as a recommendation, the work could be published with the approval of the editors.

s20174734

Reviewer 2
Summary: The authors proposed *[…]* to detect the photoplethysmographic signals (PPG). The idea is *[…]* the authors should extend the description of mOEPS sensor. The reader does not *[…]* your case is a self-citation but personally I think that is necessary for this paper that the authors quote the prior works. Moreover, in row 54 the authors write "mOEPS" without specifying whats mean. In row 59 the authors write "previous published work" it is necessary the reference. Some abbreviations are also missing. Furthermore there are many abbreviations, for this reason, I suggest to the authors to report a short table with a short description of each term (mOEPS, PCB, PD, VDD, MCU, LPF...) The following typographical error was detected in line *[…]* like the previous paragraph (The same things for line 203 and 204). Moreover about this section, Which type of PPG sampling device has been used? Add details. Which kind of pipeline has been used to stabilize/filter the raw PPG signals with respect to artefacts (body movements, etc..) or electronic noise? Add more details about the above questions if the authors consider it would be useful. Furthermore, in line 208 *[…]*

app12136469

Reviewer 2
This paper starts by presenting state-of-the-art techniques *[…]* by short-term photoplethysmography (PPG).This work is interesting and presents good scientific quality and could be relevant. It is well presented from point of view of the Materials and Method. Its publication is recommended. Just some *[…]* the results that were briefly described in Section 3 "Results". The bibliography is scarce and should be reviewed and updated to improve the quality and interest of readers and researchers Once these considerations have been made as a recommendation, the work could be published with the approval of the editors.1) There are many abbreviations and several markers, for this reason, I suggest to the authors report a short table with a short description of each parameter.2) In section 1 (Introduction) the authors should extend the description of HRV and PRV. I think that is necessary for this paper that the authors quote the prior works. Moreover, in row 35 the authors write "SDNN, pNN50" without specifying what means. 3) Which type of PPG sampling device has been used? Add details. Which kind of pipeline has been used to stabilize/filter the raw PPG signals with respect to artefacts (body movements, etc..) or electronic noise? Add more details about the above questions if the authors consider it would be useful. 4) The figures must be *[…]*

Conclusions & Future work

- We found evidence of questionable practices in MDPI <u>open</u> reports… But few compared to the processed volume…

- There are strong limitations mainly due to a not so clean and very incomplete dataset.

- More research is needed to improve malpractice detection and assess its prevalence in (open) peer review reports.

https://www.mdpi.com/about/announcements/1405

12 October 2018

Open Peer Review for all MDPI Journals

It is now a little over four years since MDPI first started to offer open peer review. Given the popularity of this feature in the 14 journals operating it until now, we have decided to extend options for open reports and open identities to all MDPI journals. This means that authors have the option for the review reports and author responses to be published alongside their article, and reviewers have the option to add their name to the published review report. For further background and information, see our **blog post**.