Reproducible Research Platform for Electric Power Quality Algorithms

Arthur C. Piazzi, Augusto S. Cerqueira, Leandro R. Manso, Carlos A. Duque
Universidade Federal de Juiz de Fora, Brasil
E-mail: arthur.piazzi@engenharia.ufjf.br

Abstract—Scientific computation is emerging as the central pillar of science. However, loosed research practices and the own nature of the work is leading to a credibility crisis. Reproducible computational research is gaining attraction in the last decades not only because it benefits the validation and attestation of findings, making a more trustworthy publication, but also lowering the barrier of entry, making it more understandable for the target audience and, therefore, increasing its impact. In this article, an online platform where it is possible to store, share and execute linked code and data, from signal processing, applied to electric power quality, is presented. In such manner, this article presents an alternative for researchers in the field to conduct and share their work satisfying the requirements of reproducible research.

Index Terms—Reproducible Research, Signal Processing, Power Quality, Web Platform.

I. INTRODUCTION

The rise of computational science has led to exciting and fast developments in many scientific areas. New technologies, the increase of computing power, and methodological advances have dramatically improved our ability to collect and analyze complex high-dimensional data, but the nature of the work has exposed limitations in our ability to evaluate published findings [1].

The replication of scientific findings using independent investigators, methods, data, equipment, and protocols has long been and will continue to be, the standard by which scientific claims are evaluated. However, in many fields of study, there are examples of scientific investigations that cannot be fully replicated because of a lack of time and/or resources. In such situation, there is a need for a minimum standard that can help the researchers to replicate the results from others. A candidate for this minimum standard is the idea of Reproducible Research (RR), which requires that data sets and computer codes be available to others for verifying published results and conducting alternative analyses [2].

In [3], it is said that an article about computational science in a scientific publication is not the scholarship itself, it is merely advertising of the scholarship. The actual scholarship is the complete software development environment and the complete set of instructions which generated the figures. It points out to a fact that in the traditional research pipeline, represented in Fig.1, the authors do not always completely fulfill their purpose of propagating their discoveries to the readers.

Conducting a research in a reproducible manner, and making it available, not only ensure the capability of easier evaluation of scientific studies, it also lowers the barrier of entry to a publication and, therefore, increases its impact. Readers can assimilate the research work and use its results more easily, so they can focus on the implementation itself not only in the parts pointed out in the article.

Researchers and institutions across a range of computational science disciplines have been calling for reproducibility [4]-[9], and several initiatives have been created in order to facilitate this process, such as WaveLab[10], BEAT[11] and VisTrails[12]. Despite that, most journals and conferences do not demand that researchers adopt the culture of replication, discouraging them to take this extra effort.

Nowadays the panorama of RR is very heterogeneous, with researchers in some fields having access to sophisticated central data repositories and researchers in other fields having few useful resources for sharing code and data. In many cases, a researcher does not have an obvious place to make sure their work is reproducible and accessible by other. Until a definitive solution is created, incremental steps
towards RR would be a vast improvement over the current situation.

This paper presents one of these incremental steps, which is a web-based platform for signal processing algorithms, focused on power quality (PQ) which was first proposed in [13]. In the current scenario, PQ studies had become crucial for the society, due to the growth of disturbances sources, for instance, use of high power nonlinear loads and distributed generation allied to the increase of the sensibility of electronic devices. In this platform, it is possible to achieve reproduction of an experiment, whereas it encapsulates code, data and parameters settings. The execution of the experiment is carried out on the server-side, avoiding cross-platform incompatibility and the need for setting up environments.

The paper is divided as follows: Section II presents the web platform for PQ algorithms; in Section III, an example will be shown, focused on power quality novelty detection; and, the conclusions are in Section IV.

II. WEB PLATFORM FOR POWER QUALITY

In the process of scientific publication, there are basically two parties involved, the writer and the reader. When just the article is at reader’s disposable, the writer did not have any additional work to publish his findings and although the readers could still reproduce the results in some cases, a considerable effort is required. In the other hand, one author concerned in doing his research more reliable and useful for the community can take a series of steps to make all aspects of the publication transparent and available. In this scenario, the reproduction is simple and fast. There will always be a trade-off between the effort required by the parties and the proposed platform tries to improve this balance achieving the “one-click” reproduction for the reader without the necessity for too much additional work from the author.

As it was stated before, full replication is not always possible, mainly because it requires to replicate also the data acquisition, which raises difficulties in real-world data. Luckily, for understanding and attestation of proposed procedures, this is not always necessary. In fact, there is a whole spectrum of reproducibility [1] as shown in the Fig 2. The most basic level is publishing only the code and the step closest to full replication is reached when besides the code, the input data and thee used parameters settings are linked and also shared.

To achieve reproducibility, the proposed platform provides a code, data and the experiment repository as described below.

A. Code Repository

The core functionality of the proposed platform is archive and share scripts. Keeping in mind the sensitive of code sharing, the user has some options. Sharing publicly, the code will be available for all users of the platform. Sharing with a group, the access is restricted to people chosen by the creator of the algorithm (this option is especially useful for making peer-review maintaining the source code unavailable for the public). And last, making it private, only the owner can access the script. The last option hurts the purpose of RR but the code can still be used as a black box to generate results and to make comparisons.

Although it is possible to submit code from different languages to the repository, at this moment, only Python scripts can be executed in the platform. The choice for Python as the programming language is based not only on its wide use in the scientific community but also for being free and open-source. The use of freely available tools in RR is very important, in a scale of reproducibility presented in [14], only those get the highest score as shown in Table I.

To make possible algorithm execution, some precautions have to be followed. First of them is to implement a method named process. When an experiment that uses this algorithm is executed, this method will be called as the main process. For didactic reasons, it is advisable not using hard-coded variables, placing those variables as inputs of the method process and the platform will assign it as parameters for the experiments which uses the referred algorithm.

B. Experiment Repository

Code sharing is the first and more crucial step for making our discoveries more clear and reliable, but it is only the first. In fact, there are many state-of-art tools for code sharing available like GitHub\(^1\) and GitLab\(^2\). To accomplish our intention, the platform also integrates an experiment storage. A saved experiment binds the code with the data used on it and for that matter, the platform also has a data storage. The data is stored in data sets and each dataset contains several realizations. When creating an experiment, after the algorithm selection, the user will be asked to name the experiment and chose a data set Fig.3 already incorporated in the platform.

However, code and data linked are not always sufficient for conducting RR. As demonstrated in [15], a study that tried to reproduce several experiments in the field of bioinformatics, even with access to code and data, 30% of the results was not successfully attested. This low success rate

\(^1\)https://github.com/
\(^2\)https://gitlab.com/
TABLE I: Scores of Reproducibility [14].

<table>
<thead>
<tr>
<th>Score</th>
<th>Requirements</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>The results can be easily reproduced by an independent researcher with at most 15 minutes of user effort, requiring only standard, free available tools (C compiler, Python interpreter, etc.)</td>
</tr>
<tr>
<td>4</td>
<td>The results can be easily reproduced by an independent researcher with at most 15 minutes of user effort, requiring some proprietary source packages (MATLAB, etc.)</td>
</tr>
<tr>
<td>3</td>
<td>The results can be reproduced by an independent researcher, requiring considerable effort</td>
</tr>
<tr>
<td>2</td>
<td>The results could be reproduced by an independent researcher, requiring extreme effort</td>
</tr>
<tr>
<td>1</td>
<td>The results cannot seem to be reproduced by an independent researcher</td>
</tr>
<tr>
<td>0</td>
<td>The results cannot be reproduced by an independent researcher</td>
</tr>
</tbody>
</table>

Fig. 3: Database selection.

Fig. 4: Experiment schematic.

can be regarded by a variety of factors, including even dishonest result allegation. Nevertheless, one of the most probable source of inconstancy is due to the lack of proper parameterization.

Often, signal processing techniques are very dependent on parameter tuning. Those attributes are generally set up by trial and error, whether did manually, by automatic grid search or by optimization techniques, keeping track of which set of inputs is related to which result demands good research practices, even more, when it comes to large and lasting projects. Considering the limited length of a paper, and the tedious reading of a full documentation of variables setting would engender, it is not practical to disclose it, even when it is possible. Taking that into account, every experiment stored has its parameters values discriminated.

The last element that composes a saved experiment in the platform is the actual results. Although this last element differs from the others because it can not be saved or altered by the author, all the results are generated by executing the scripts on the server, ensuring the reliability of the outputs. The result can be returned in any Python native data format, in their numpy counterparts or as an image. For the last one, it just ought to create pyplot figure within the algorithm, no explicit return is needed, although image results can only be compared visually. Any user of the platform can request the re-execution of the experiment in other to attest it. Fig. 4 shows a schematic of one experiment in the platform.

In that way, it is possible to pack one research on the website, making it reproducible and available. Users can check the results, study the source-code, use the same data for different analyses and even produce new results applying the method to other signals or changing parameters. The importance of this interaction is fundamental because the advances in science over the last several hundred years have not come from individual scientists, they have come from scholars and scientists working in concert with each other and the community [16].

Virtually, every problem to be solved has a vast multiplicity of solutions and in signal processing, it is no different. Comparison between methods which has the same objective is the most efficient way to determine the validity of a new work. Only testing new methods against other in the literature, enables to establish benchmarks and consequently achieve scientific progress. Knowing that its possible for the users create routines that compares the results from various algorithms. Fig 5 exemplify this procedure.

III. CASE OF STUDY

The last decades of the 20th century experiment an enormous increase in electronic sensitive equipment due to the huge growth in communication, computer and automation industries. These new generations of electronic equipment are becoming progressively sensitive to power quality disturbances. Any variation in the supply voltage magnitude or frequency may have detrimental effects on the equipment. Electric utilities must assess the state of the power system signal to assess the quality of power. Therefore, detection of power quality disturbances has become a significant issue.
In order to illustrate the functionality of the platform, it was partially recreated the work presented in [18]. It consists of a system for detecting and compressing disturbances found in electrical signals. Perform such compression is vital for monitoring power quality events, seeing that is impractical and inefficient to store all raw data. An efficient compression rate is obtained in this system in three stages: the novelty stage, the wavelet stage and the byte stage.

The results exhibited here are regarded to the novelty detection step. This monitoring system in its first step consists in comparing frames of the signal with a reference one, in the article a frame was established as 4 cycles of the fundamental frequency component. Before the comparison, the signal is filtered by a high pass filter with cutoff frequency of 720 Hz to avoid detection due to frequency variations. To measure the similarity between frames the energy of the frame was used, if the difference between the energy of two consecutive frames is greater than a threshold those frames are said to be different and the actual frame is said to be a novelty frame and it is stored for further studies.

The task of indicating the novelty frames can be performed by a number of different ways. For instance, the technique proposed in [18] works in the time domain, but the novelty can also be detected in the frequency domain. In that way, the coefficients of the Discrete Fourier Transform can also be used to perform novelty detection. When any frequency sample from two consecutive frames, also adopted as 4 cycles from the fundamental frequency, differs from the reference one beyond a previously established threshold, the novelty is accused.

In both cases, the algorithms have two outputs, one image that shows the analyzed signal divided by frames, and the actual result, which consists of the frames which the novelty was detected. The comparison between these two methods is fairly straightforward and the results for each novelty can be considered equal when all the accessed frames are the same. Results from different methods in the platform can be compared as long as their results were generated from the same data set and the output has the same format. The signals used in this section were synthetically generated with a fundamental frequency of 60 Hz, 7680 Hz sample frequency and length of 50 cycles, several PQ disturbances were added in each signal, in order to introduce signal novelties.

Two comparisons were written, one that outputs a list of strings containing the data event index and a Boolean flag representing if the results of the algorithms are equal or not. This analysis is especially useful when regarding small data set, when individual results are meaningful. Fig.6 shows the output for a small data set containing 8 frames of the signal. As the number of events in the dataset scales, individual comparison loses its usefulness, therefore, Fig.7 shows a broader parallel between the methods, discriminating only percentages of parities.

Figures 8 and 9 shows a screenshot from the experiment page from the website, regarding the results generated by energy comparison and frequency comparison, respectively. It was selected a specific event where the results from both methods diverged. In the sidebar it can be noted the other elements which constitute an experiment, navigating through it, one can visualize the source code, a list of comparison algorithm that uses that experiment’s results, the input parameters and the database used in the analyses.

Through this example, it can be noted that both methods give similar results in most cases, suggesting that both are suitable for the task. However, as it can be seen in figures 8 and 9, where it also shows the average execution time of each method, the routine based on frequency coefficients is more

Fig. 5: Comparison schematic.

Fig. 6: Comparison Result.
Most of the advances in the field rely on computer science, nevertheless, the prevalence of loose practices is leading to a credibility crisis. To increase the reliability and the usefulness of the scientific discovery, reproducible research is the evident candidate to achieve it.

Without a standard set of tools for RR and the lack of demand from journals and conferences for this practice, the culture of reproducibility is not yet as strong as it should be. Initiatives that enable and facilitate RR is crucial for scientific progress. The present platform brings a way of sharing power quality studies, fulfilling all the requisites of RR (data, code, setting up parameters), with low additional effort for the authors. The cloud execution of the methods incorporated in the platform makes possible to verify results with zero effort from the readers, eliminating the need for acquiring files, software installation, and data download.

Testing different approaches with the same purpose against each other is the most efficient way of consolidating and standardize techniques. The capability of doing it so easily, quickly and reliably, without tiresome research, promotes the scientific progress.

REFERENCES