IS-ENES3 Milestone M5.3
Requirements for technical standards for diagnostic tools

Reporting period: 01/07/2020 – 31/12/2021

Authors: Javier Vegas-Regidor, Saskia Loosveldt-Tomas, Kim Serradell
Release date: 18/06/2021

Abstract:

This report contains a list of requirements for diagnostic tools that has been compiled from the results of the user survey conducted by Assimila for D3.1 and the discussions held at the Virtual workshop on requirements for a fast and scalable evaluation workflow held in May 18-19. Those technical requirements cover both infrastructure and software needs but also mention some dissemination and user engagement needs that arose during the process. This milestone will be one of the main inputs for D5.2 (Technical standards for diagnostic tools).

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824084

Table of contents

1. Objectives

The main objective of this milestone is to capture the current requirements from the model evaluation community so they can be used to inform the further development of the evaluation tools and, in particular, the development of the technical standards for diagnostic tools that will be delivered in D5.2.

Although the main focus of this deliverable is technical, requirements for non-technical aspects of the tools are also added as they may have an impact on some technical decisions.

2. Description of work: Methodology and Results

The users' requirements captured on this document came from two fundamental sources: the results of the survey conducted by Assimila for D3.1 and the IS-ENES3 Virtual workshop on requirements for a fast and scalable evaluation workflow held on May 18-19 (https://is.enes.org/events/workshops/is-enes3-virtual-workshop-on-requirements-for-a-fast-and-scalable-evaluation-workflow). The workshop gathered 30 experts during one day and half to discuss those topics. Those actions are complemented with the experience IS-ENES partners have on the topic after years of experience on evaluation activities both at group and community levels.

Those inputs lead to the following list of requirements. It is divided in three sections for better clarity: dissemination and user engagement; infrastructure and software.

2.1. Dissemination and user engagement

Unintuitively, the first requisite we have to mention is non-technical: both the survey and the workshop made clear that a big part of the evaluation community is not aware of all the tools they have at their disposal, including both infrastructure and software tools. This is the first and foremost issue that must be tackled in the future as the impact of any other work will be minimal if potential users are not reached out. This also makes capturing new requirements from users harder as most of the requests have, at least, a potential solution already in place.

As a side note, this problem may have been exacerbated by the lack of in person meetings in the last 18 months due to the Covid-19 pandemic. Evaluation tools usually leave much of their

dissemination to the word-of-mouth in congress and other community meetings, and online events do not favour this kind of interaction.

2.2. Infrastructure

From an infrastructure viewpoint, easy access to data is the key requirement from the users. That implies not only being able to access huge quantities of data but also that they are in a consistent format. The ESGF is a huge step in that direction, with a special remark for the compute facilities close to the data like CEDA Jasmin and DKRZ Mistral.

Nevertheless, the need for local data download is still there and needs to be taken into account. Users also report a need for the ability to download preprocessed data directly in order to keep the download sizes manageable.

The standardization of modelling data formats taken into CMIP is a huge success, even if minor departures of the format are a source of continuous issues for the users. A push for further standardization in the observational and reanalysis data will be of enormous help for both users and tool developers.

Regarding computing resources, the most critical resource usually is RAM memory, especially when dealing with big model ensembles or high-resolution experiments. This importance is due to the fact that its effect is binary: if there is enough RAM, you can evaluate the models; if there is not enough, you cannot.

Users are usually not very concerned with analysis time so tools have not really focused on exploiting the computational power of the current systems, making further improvements welcome but not really required. Another way of improving computing times that will potentially be more impactful is to get faster storage solutions, as a meaningful part of analysis time is spent on I/O tasks.

2.3. Software

From a software perspective, users are heavily interested (and invested) in interoperability: most of them currently use several tools written in different languages to leverage the different strengths that each one possess. Any analysis workflow that pretends to be widely used must allow this interoperability to happen easily.

This focus on interoperability must be addressed in a way that does not harm reproducibility and shareability of the results, as this is a rising need in the community. Developers should then focus on promoting (but not forcing) tools that are freely available and filling the gaps that those tools currently have.

A good (but not perfect) example of how to meet those two apparently contradictory requirements can be seen in the ESMValTool:

- Complete evaluation workflows can be integrated and distributed by the tool, but only if all parts are freely available. This includes installation of all the dependencies.
- It has direct support for Python, R, NCL and Julia diagnostics, but any executable can potentially be used.
- Users can also use the tool to preprocess the data to a CMOR-like format and then feed those outputs in the tools of their choice.

Another important request from the users is to ease the procedure for adoption of any tool. This implies having as easy installations as possible, good documentation and tutorials (both taught and on-demand) and meaningful examples that can be used as a starting point for the users' work. A good rule of thumb is that a completely new user should have a new tool working and assess its potential in just one day, as that is the time most of them feel that they can devote themselves to exploring a new tool.

Users also require tools that are able to incorporate new functionality with ease and at a strong pace. This need for quick evolution is inherent to scientific work, as it consists in constantly pushing the boundaries of what has been done. Nevertheless, a strong effort is required to make this quick evolution pace compatible with having highly reliable tools.

From a purely computational perspective, users are heavily interested in keeping memory consumption low, as this is sometimes the difference between being able to perform an analysis or not. Developers will need to focus on keeping the memory footprint of their tools to a reasonable level and to be especially cautious in their use of memory when dealing with model ensembles and high-resolution data. Execution time requirements are usually not as important, but any improvements are welcome.

3. Difficulties overcome

As one must expect, the Covid-19 pandemic has been an issue for all work as centered on user interaction as this milestone. The lack of in-person meetings does not favour engaging with the users. Specifically, it impacts our ability to have informal discussions on breaks and social events that are usually a great source of information about users' needs and current difficulties, as they tend to be more relaxed and speak more freely than in bigger sessions.

The work has also been affected by the delay on Assimila's survey for D3.1, as this was one of the main inputs for this work.

4. Next steps

The conclusion of this milestone will be one of the inputs for deliverables 5.2 (Technical standards for diagnostic tools) and 5.5 (Style guide on coding standards). This way we will try to ensure that both the technical and coding standards respond to the needs of the community.