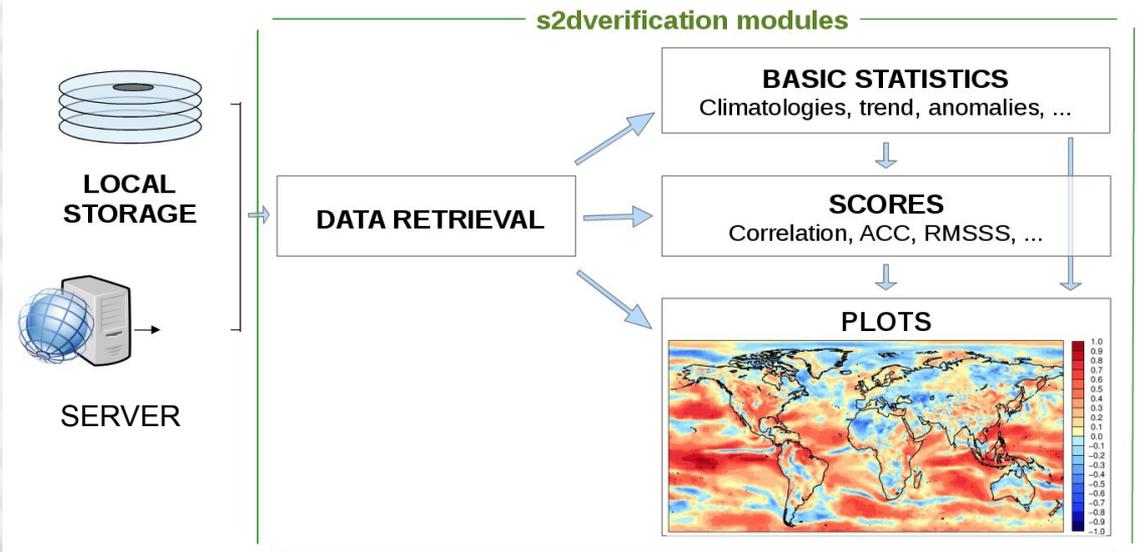


Package developed for verification



FORMATTING:

Regrid, Reference to seasonal forecast data structure, Means, Insert dimension, Seasonal means and Subsetting.

BASIC STATISTICS:

Climatologies with different methodologies for ensemble means or individual ensembles, Anomalies with our without cross validation, Spread, Trends, Regression, Cluster, Composites, EOF, NAO, Autocorrelation, Sample Size, Filter Probability bins, Single Value Decomposition, estimate frequency spectrum and Smoothing.

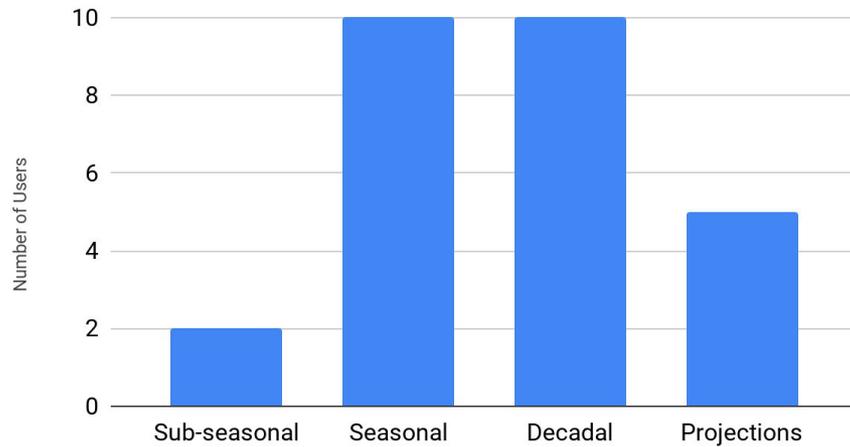
SCORES:

Correlation, Anomaly Correlation Coefficients, Brier Skill Score, Brier skill score, decomposition of Brier skill score, RMS, RMSSS, Ratio RMS between two forecasts or the ensemble spread,

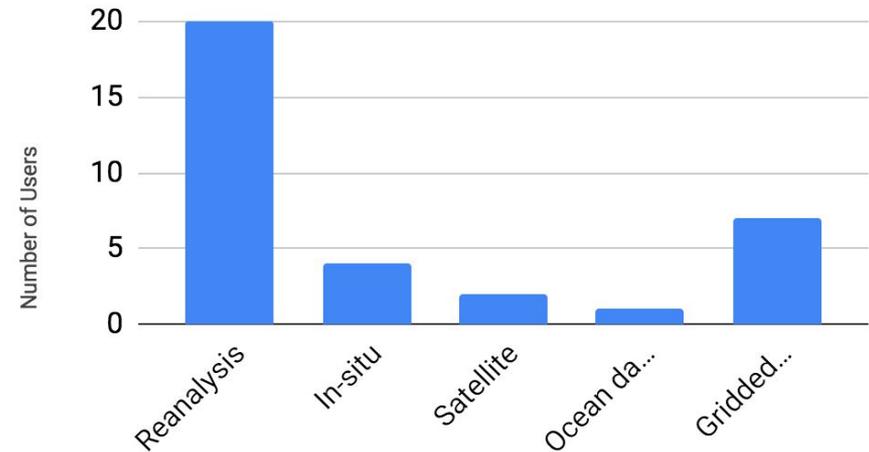
Users

20 Earth Science department members answered a poll at the beginning of 2020.

Forecast horizon

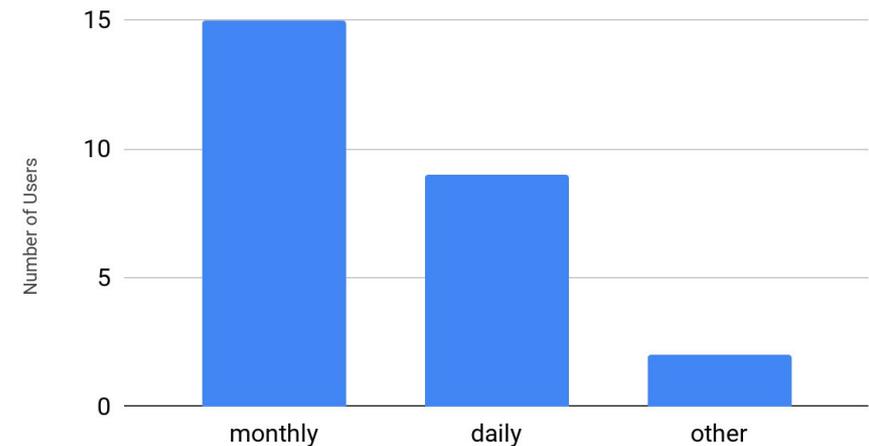


Reference Datasets



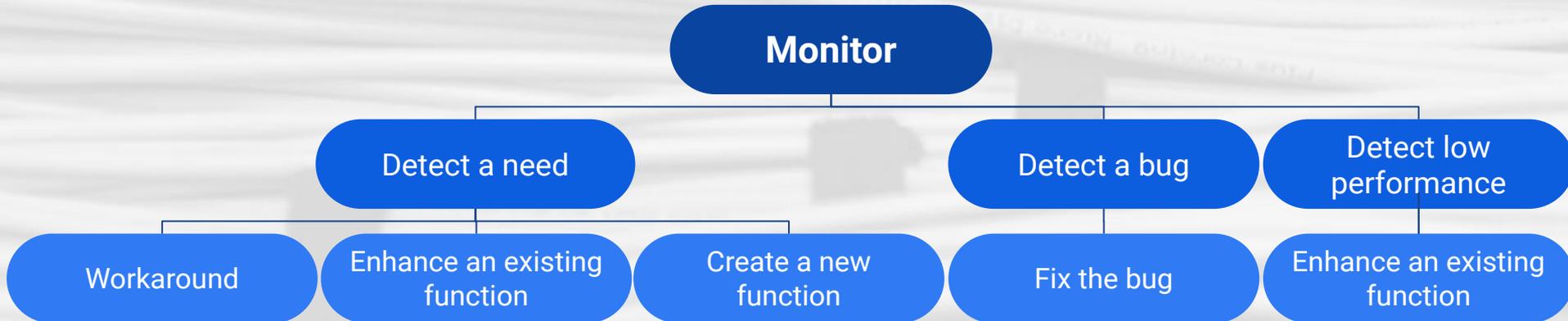
- The diversity in research analysis should be translated in flexibility of the tools
- 15 people ask for a tutorial in that poll

Data frequency



Development and Maintenance

Objective 1: Useful package



- **Balanced design to provide maximum flexibility while ensuring computational performance**
- **Avoid functions repetition to reduce maintenance cost**

Example of new function creation:

- new code to compute a skill score to be converted as a function on the package

Example of workaround:

- preprocessing of the data could allow using an existing function for a new purpose

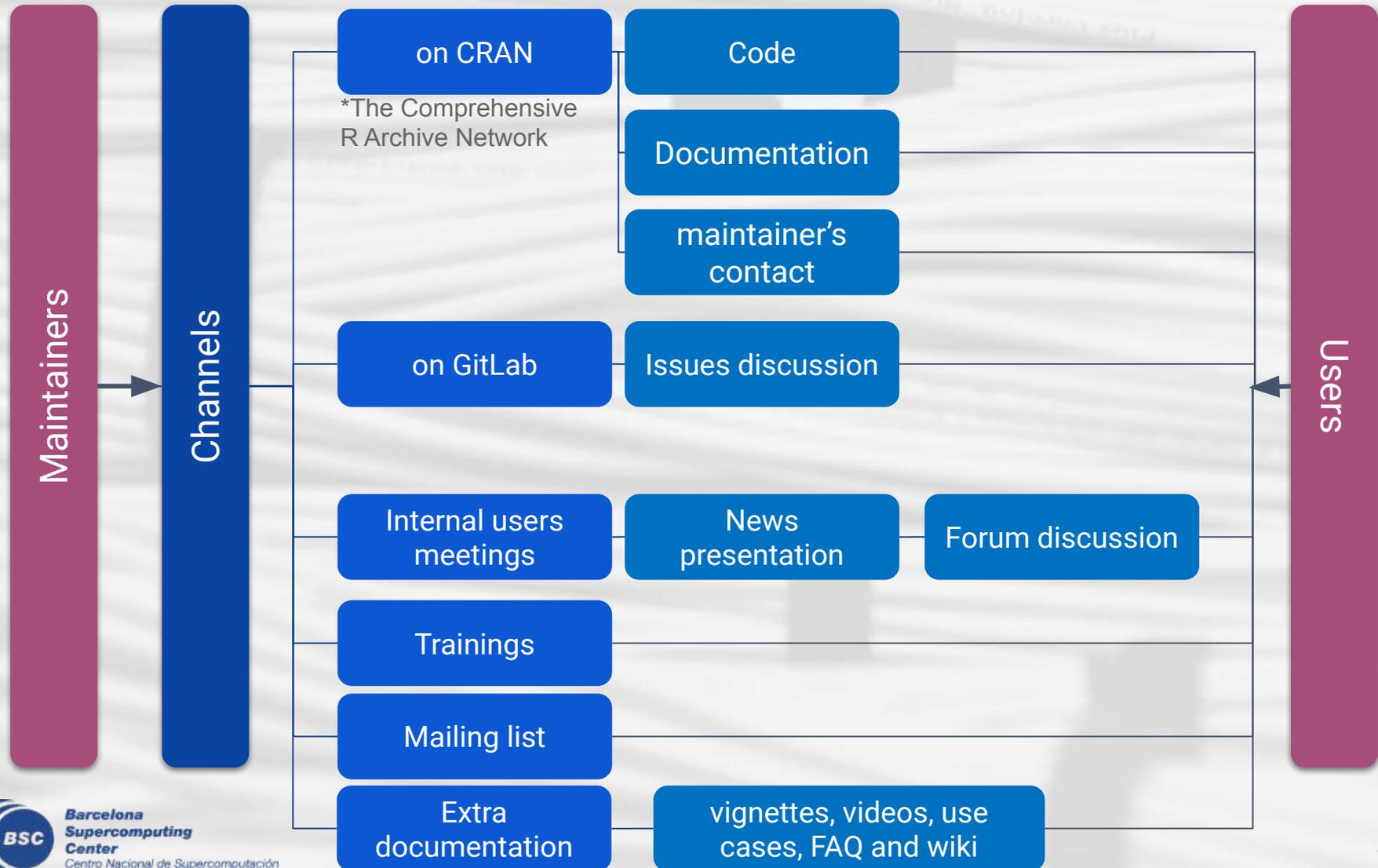
Example of function enhancement:

- add a new parameter to existing function

Development and Maintenance

Objective 2: Share the package and updates

Objective 3: Get feedback



Development and Maintenance

Objective 4: Safe development

CRAN

The Comprehensive R Archive Network

- Package shared in ftp and web servers around the world
- Common standards for R users
 - easy installation
 - typical R documentation
 - typical warnings and error messages
- Checks
 - to work on different OS
 - structure and documentation
 - package size
 - installation time
 - ...

Performance evaluation

Profiling tools

- ProfVis and Rprof

Control Version System

GitLab project

- track changes
- issues discussion
- tag versions
- branching strategy
- unit testing
- continuous integration

Guidelines

How to contribute

- Open a new issue
- Work in a new branch
- Follow function standards
- Review process
- Merge into the master branch
- Plan next release

Development and Maintenance

Objective 5: Functions design to overcome current challenges

Origin

s2dverification functions are originally tailored for ensemble forecasts and tested in seasonal forecast application.

Challenge

Data size: The increase in number of models, number of members and resolution and, functions applications in new horizon predictions (e.g.: subseasonal forecast and decadal predictions) bring new challenge to use the package under evolving conditions.

Solutions

Two approaches are considered that can be combined:

- 1) Functions working on **multidimensional arrays with named dimensions** thanks to **multiApply package**. This is translated in flexible functions that allows analyzing extra number of dimensions (e.g.: analyze multiple vertical levels simultaneously) and **parallelization** of the computation by defining a single parameter indicating the number of cores. **s2dv is the new version of s2dverification.**
- 2) **startR package:** when functions works with arrays with named dimensions they can easily be used in the startR workflow that allows users to **chunk the data** in pieces to perform analysis in local machines or HPC clusters.

Development and Maintenance

Difficulties

Monitor

- 1) Are users aware of the new developments?
- 2) Are they getting the best of the tools?
- 3) How to get more feedback from them?

Enhancement

- 1) Are there lacks in the guidelines or maintenance process?
- 2) Before developing, is it a need of 1 user or it is a need for several users? Is it worth it to do the development?

Performance

- 1) Is the tools performance good enough?
- 2) Are the maintainers in continuous learning process from the state-of-the-arts technologies?