Deliverable 7.1

D7.1 Distribute new CSEP 2.0 software code

Deliverable information	
Work package	[WP 7: Testing: Rigorous testing and validation of dynamic risk components]
Lead	[GFZ]
Authors	[William Savran (USC), Danijel Schorlemmer (GFZ), Pablo Iturrieta (GFZ), Marcus Herrmann (UNINA), Max Werner (UBRIS)]
Reviewers	[Management Board]
Approval	[Management Board]
Status	[Final]
Dissemination level	[Public]
Will the data supporting this document be made open access? (Y/N)	N/A
If No Open Access, provide reasons	
Delivery deadline	[28.02.2022]
Submission date	[17.02.2022]
Intranet path	[DOCUMENTS/DELIVERABLES/]

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 821115

Introduction

Over the last decade, the Collaboratory of the Study of Earthquake Predictability (CSEP) has led numerous prospective earthquake forecasting experiments (see, e.g., Michael & Werner, 2018). These experiments are formally conducted within testing centers (Schorlemmer & Gerstenberger, 2007) that contain the software required to autonomously run and evaluate earthquake forecasts. Such testing centers were installed at USC, ERI, GNS, and ETH, covering a variety of testing regions, e.g. California, Japan, New Zealand, Italy, and a global experiment. They are all operated by the same CSEP software stack as mainly developed at USC. The software design emphasized a carefully controlled computing and software environment which ensured integrity of testing results (Zechar et al., 2009). However, its monolithic design made it difficult for researchers to utilize various routines in the testing centers in their own work without replicating the entire testing center configuration on their own system. In addition, software was developed by a single developer, leading to personnel risk and a lack of opportunities for others to contribute directly.

As a consequence, the CSEP group decided to fundamentally change the design paradigm of the CSEP software to address these problems. The new software stack, formerly referred to as CSEP 2.0, is designed as a Python toolbox (called pyCSEP) for easy use by modelers but also for the assembly of readily deployable fully-reproducible earthquake forecasting experiments (see MS47). pyCSEP was designed to provide vetted methods to evaluate earthquake forecasts that researchers can include directly in their research. The statistical tests and tools to evaluate earthquake forecasts are required by all model developers, and greatly benefit from open-source development practices by providing standardized, well-tested, and community-reviewed software tools. As of now, pyCSEP has been used for three published articles (Bayona et al., 2020, 2022; Savran et al., 2020), and is being used by several research groups participating in RISE and other projects.

pyCSEP Overview

pyCSEP provides an open-source implementation of peer-reviewed statistical tests developed for evaluating probabilistic earthquake forecasts (Rhoades et al., 2011; Savran et al., 2020; Schorlemmer et al., 2007; Werner et al., 2011; Zechar et al., 2013). In addition, pyCSEP provides routines for working with earthquake catalogs and visualizations. The core design of pyCSEP includes classes that represent earthquake forecasts, catalogs, and various spatial regions, see Figure 1. Higher level functions are implemented using these classes to provide routines for common tasks in evaluating earthquake forecasts.

Earthquake forecasts can either be specified as expected earthquake rates over discrete space-magnitude-time regions (Schorlemmer et al., 2007) or as families of synthetic earthquake catalogs with each catalog representing a realization from the underlying stochastic model (e.g., Savran et al., 2020). Earthquake catalogs are row-based data sets that contain features of an earthquake. At a minimum, an earthquake must be defined by its geographical location (latitude, longitude), origin time, and magnitude. In addition, pyCSEP provides classes for working directly with forecasts from the Uniform California Earthquake Rupture Forecast with Epidemic-type Aftershock Sequences Version 3 (Field et al., 2017). pyCSEP also provides classes for interacting with earthquake catalogs and performing operations on them, such as filtering and binning events on the space-magnitude grids needed for evaluation. pyCSEP includes numerous flexible plotting utilities that interface directly with `matplotlib` and `Cartopy` (Hunter, 2007; Met Office, 2010–2015), allowing users to quickly visualize forecasts and test results, or to create publication-ready figures. Space-magnitude regions facilitate grid operations that are necessary for evaluating earthquake forecasts. These objects model regular latitude, longitude cells where earthquakes can be aggregated for evaluation and visualization purposes. pyCSEP provides pre-defined spatial regions that have been used in previous experiments (Field, 2007; Taroni et al., 2018).

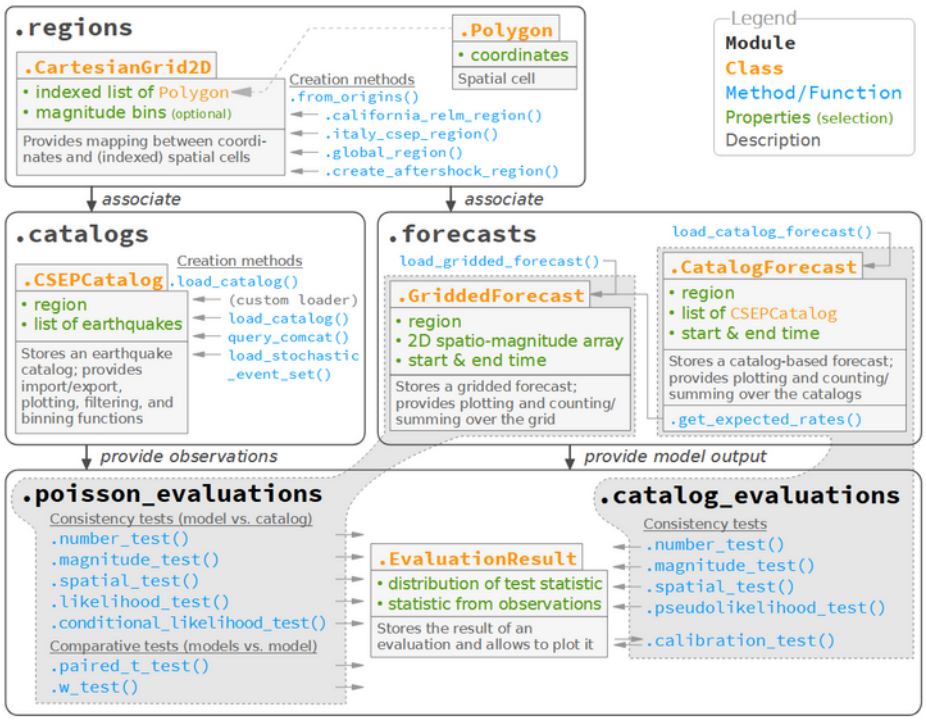

Figure 1: pyCSEP class structure. Taken from Savran et al. (submitted)

pyCSEP interfaces directly with popular numerical and plotting libraries such as `Numpy`, `matplotlib`, and `pandas` (Harris et al., 2020; Hunter, 2007; McKinney, 2010). Users already familiar with these libraries can adapt pyCSEP directly into their code. pyCSEP provides file-formats for forecasts and earthquake catalogs, and can allow users to specify custom filetypes. Along with providing modelers with quickly deployable tools, pyCSEP is flexible and powerful enough to become the backbone of future testing experiments, by abstracting forecasts, tests, catalogs objects in a simple fashion. As open-source software, new forecasting experiments can be easily implemented, shared and reproduced.

Parts of this text were taken from the first publication about pyCSEP (Savran et al., 2022). Another paper with a more detailed description about pyCSEP has been submitted (Savran et al., submitted).

The pyCSEP source code can be obtained at:

> https://github.com/SCECcode/pycsep

The Zenodo DOI is:

> https://doi.org/10.5281/zenodo.5659928

pyCSEP can be installed using `pip` or `conda`. Full installation instructions are found at:

> https://docs.cseptesting.org/getting_started/installing.html

The pyCSEP documentation with tutorials and examples can be found at:

> https://docs.cseptesting.org/

pyCSEP is released under the BSD 3-Clause "New" or "Revised" License.

References

Bayona, J. A., Savran, W. H., Strader, A., Hainzl, S., Cotton, F., & Schorlemmer, D. (2020). Two global ensemble seismicity models obtained from the combination of interseismic strain measurements and earthquake-catalogue information. *Geophys. J. Int.*, 224(3), 1945–1955. https://doi.org/10.1093/gji/ggaa554

Bayona, J. A., Savran, W. H., Rhoades, D. A., & Werner, M. J. (2022). Prospective evaluation of multiplicative hybrid earthquake forecasting models in California, *Geophys. J. Int.*, ggac018, https://doi.org/10.1093/gji/ggac018

Field, E. H. (2007). Overview of the working group for the development of Regional Earthquake Likelihood Models (RELM). *Seismol. Res. Lett.*, 78(1), 1–10. https://doi.org/10.1785/gssrl.78.1.7

Field, E. H., Jordan, T. H., Page, M. T., Milner, K. R., Shaw, B. E., Dawson, T. E., Biasi, G. P., Parsons, T., Hardebeck, J. L., Michael, A. J., Weldon, R. J., Powers, P. M., Johnson, K. M., Zeng, Y. H., Felzer, K. R., Elst, N. van der, Madden, C., Arrowsmith, R., Werner, M. J., & Thatcher, W. R. (2017). A synoptic view of the third Uniform California Earthquake Rupture Forecast (UCERF3). *Seismol. Res. Lett.*, 88(5), 1259–1267. https://doi.org/10.1785/0220170045

Harris, C. R., Millman, K. J., Walt, S. J. van der, Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., Kerkwijk, M. H. van, Brett, M., Haldane, A., Río, J. F. del, Wiebe, M., Peterson, P., ... Oliphant, T. E. (2020). Array programming with NumPy. *Nature*, 585(7825), 357–362. https://doi.org/10.1038/s41586-020-2649-2

Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science & Engineering*, 9(03), 90–95. https://doi.org/10.1109/mcse.2007.55

McKinney, Wes. (2010). Data Structures for Statistical Computing in Python. In Stéfan van der Walt & Jarrod Millman (Eds.), *Proceedings of the 9th Python in Science Conference* (pp. 56–61). https://doi.org/10.25080/Majora-92bf1922-00a

Met Office. (2010 - 2015). Cartopy: A cartographic python library with a matplotlib interface. https://scitools.org.uk/cartopy

Michael, A. J., & Werner, M. J. (2018). Preface to the focus section on the Collaboratory for the Study of Earthquake Predictability (CSEP): New results and future directions. *Seismol. Res. Lett.*, 89(4), 1226–1228. https://doi.org/10.1785/0220180161

Rhoades, D. A., Schorlemmer, D., Gerstenberger, M. C., Christophersen, A., Zechar, J. D., & Imoto, M. (2011). Efficient testing of earthquake forecasting models. *Acta Geophysica*, 59(4), 728–747. https://doi.org/10.2478/s11600-011-0013-5

Savran, W. H., Werner, M. J., Schorlemmer, D., & Maechling, P. J., (2022). pyCSEP: A Python Toolkit For Earthquake Forecast Developers. *Journal of Open Source Software*, 7(69), 3658. https://doi.org/10.21105/joss.03658

Savran, W. H., Werner, M. J., Marzocchi, W., Rhoades, D. A., Jackson, D. D., Milner, K., Field, E. H., & Michael, A. J. (2020). Pseudoprospective evaluation of UCERF3-ETAS forecasts during the 2019 Ridgecrest sequence. *Bull. Seismol. Soc. Am.*, 110(4), 1799–1817. https://doi.org/10.1785/0120200026

Savran, W. H., Bayona, J. A., Iturrieta, P., Asim, K. M., Bao, H., Bayliss, K., Herrmann, M., Schorlemmer, D., Maechling, P. J., & Werner, M. J. (2022). pyCSEP: A Python Toolkit for Earthquake Forecast Developers. *Seismol. Res. Lett.*: Electronic Seismologist, (submitted).

Schorlemmer, D., & Gerstenberger, M. C. (2007). RELM testing center. *Seismol. Res. Lett.*, 78(1), 30–36. https://doi.org/10.1785/gssrl.78.1.30

Schorlemmer, D., Gerstenberger, M. C., Wiemer, S., Jackson, D. D., & Rhoades, D. A. (2007). Earthquake likelihood model testing. *Seismol. Res. Lett.*, 78(1), 17–29. https://doi.org/10.1785/gssrl.78.1.17

Taroni, M., Marzocchi, W., Schorlemmer, D., Werner, M. J., Wiemer, S., Zechar, J. D., Heiniger, L., & Euchner, F. (2018). Prospective CSEP evaluation of 1-day, 3-month, and 5-yr earthquake forecasts for Italy. *Seismol. Res. Lett.*, 89(4), 1251–1261. https://doi.org/10.1785/0220180031

Werner, M. J., Helmstetter, A., Jackson, D. D., & Kagan, Y. Y. (2011). High-resolution long-term and short-term earthquake forecasts for California. *Bull. Seismol. Soc. Am.*, 101(4), 1630–1648. https://doi.org/10.1785/0120090340

Zechar, J., Schorlemmer, D., Liukis, M., Yu, J., Euchner, F., Maechling, P. J., & Jordan, T. H. (2009). The Collaboratory for the Study of Earthquake Predictability perspective on computational earthquake science. *Concurrency and Computation: Practice and Experience*, 22(12), 1836–1847. https://doi.org/10.1002/cpe.1519

Zechar, J., Schorlemmer, D., Werner, M. J., Gerstenberger, M. C., Rhoades, D. A., & Jordan, T. H. (2013). Regional earthquake likelihood models I: First-order results. *Bull. Seismol. Soc. Am.*, 103(2A), 787–798. https://doi.org/10.1785/0120120186