Madagascar software package and reproducible research

Sergey Fomel[*], **Paul Sava**[†], **Ioan Vlad**[‡], **Yang Liu**[*], **James Jennings**[§], **Jules Browaeys**[¶], **Vladimir Bashkardin**[*], **Jeff Godwin**[†], **Xiaolei Song**[*], and **Gilles Hennenfent**[‖]

[*]*Bureau of Economic Geology,*

Jackson School of Geosciences

The University of Texas at Austin

University Station, Box X

Austin, TX 78713-8924

[†] *Department of Geophysics*

Colorado School of Mines

Golden, CO 80401

[‡] *Statoil*

[§] *Shell*

[¶] *Total*

[‖] *Chevron*

(July 16, 2012)

GEO-2010

Running head: **Madagascar**

ABSTRACT

MADAGASCAR is a software package...

INTRODUCTION

The goal of this paper is to introduce the MADAGASCAR software package. MADAGASCAR provides a collection of programs for multidimensional data analysis, including seismic imaging and seismic data processing, and an integrated environment for arranging data processing workflows and computational experiments. MADAGASCAR follows the philosophy of *reproducible research*, introduced by Claerbout (1992). Reproducible research implies that the results of computational experiments are published along with the software code and, if possible, the input data, so that the reader of an article is able to reproduce and verify the experiment (Schwab et al., 2000; Fomel and Claerbout, 2010).

Why is reproducible research important? A scientific communication is valuable only if the result is communicated to a skeptic, someone who does not take the author's conclusions for granted unless they are confirmed by sufficient evidence. The ability to confirm and verify the published result allows the recepient of a scientific communtion to build on top of previous result and to extend scientific and technological progress. This is the case... computations...

It is the need for an integrated and open environment for geophysical research that motivated the development of the MADAGASCAR software package. reproducible research... ..similarities and differences with MATLAB... with SU...

To make the presentation more entertaining, we organize the paper as a story. In this story, a young geophysicist discovers Madagascar, uses it to perform data processing tasks and to develop new algorithms, and eventually contributes to the Madagascar development. Through this story, we introduce different features of the package that the reader may find useful.

JORDAN DISCOVERS MADAGASCAR

This is a story of Jordan, a young geophysicist who works on geophysical data analysis. One day, Jordan is thinking about...

The design of MADAGASCAR is based on several fundamental principles.

One of the principles is *modularity* and *encapsulation.* Similarly to Seismic Unix (Stockwell, 1997), the core of MADAGASCAR consists of a collection of separate modules, connected in data analysis workflows...

reproducible research and *test-driven development.*

file format

JORDAN WRITES A MADAGASCAR PROGRAM

JORDAN CONDUCTS NUMERICAL EXPERIMENTS

JORDAN WRITES A REPRODUCIBLE PAPER

JORDAN CONTRIBUTES TO MADAGASCAR

HOW TO FIND YOUR WAY AROUND MADAGASCAR

An easy way to start exploring MADAGASCAR is checking the list of reproducible papers at `http://www.ahay.org/wiki/Reproducible_Documents`. If any of these papers looks close to your interests, follow the links until you find a figure with a "wrench" button under it. Click on the wrench, and it will open a computational recipe used for generating the figure (the `SConstruct` file). Here is an simple example[*]:

---
[*]`http://www.reproducibility.org/RSF/book/rsf/rsf/sfmath.html`

Madagascar components

The MADAGASCAR installation consists of several components:

Main programs. Hundreds of main programs...

Processing scripts. Processing scripts (`SConstruct` files)

Reproducible papers. Scientific publications...

Reproducible figures. Figures from reproducible papers are saved...

APIs. The application programmer interfaces provide...

Installation

MADAGASCAR has been installed and tested on a variety of platforms...

HOW TO CONTRIBUTE YOUR OWN PROGRAMS TO MADAGASCAR

The MADAGASCAR community is open. Developers take a joint responsibility of maintaining and extending the package. It is easy to become a developer:

1. Follow the Guide on Adding new programs to Madagascar.

2. If your employer holds the copyright for your software, obtain a permission to distribute under a GPL license and place a copyright and GPL license notice in each file with code.

3. Register at SourceForge.

4. Find out who the current project administrators are and pass your SourceForge user name to one of them.

5. Upload your directory to the Subversion repository at SourceForge.

6. Make sure to add reproducible examples of using your program. As a rule, programs without examples are not included in the release version.

CONCLUSIONS

We have introduced MADAGASCAR, a software package for multidimensional data analysis...

ACKNOWLEDGMENTS

In addition to the authors of this paper, many people contributed to the development of MADAGASCAR. We would like to thank...

APPENDIX A

BASIC FACTS ABOUT MADAGASCAR

The work on the MADAGASCAR package (previously named RSF for *Regularly Sampled Format*) started in 2003. The package got publicly released during an EAGE Workshop on open-source E&P Software in Vienna in June 2006. Since then, more than 25 people have contributed to its development, writing cummulatively almost 300,000 lines of code.

Ohloh, a website that maps the landscape of open-source software development, evaluates MADAGASCAR as

- Mature, well-established codebase;

- Increasing year-over-year development activity;

- Large, active development team.

Several events helped promote MADAGASCAR in the geophysical and signal processing communities (Table 1.)

REFERENCES

Claerbout, J., 1992, Electronic documents give reproducible research a new meaning: 62nd Ann. Internat. Mtg, Soc. of Expl. Geophys., 601–604.

Fomel, S., and J. Claerbout, 2010, Guest editors' introduction: Reproducible research: Computing in Science & Engineering, **11**, 5–7.

Schwab, M., M. Karrenbach, and J. Claerbout, 2000, Making scientific computations reproducible: Computing in Science & Engineering, **2**, 61–67.

Stockwell, J. W., 1997, Free software in education: A case study of CWP/SU: Seismic Unix: The Leading Edge, **16**, 1045–1049.

LIST OF TABLES

year	location	name
2006	Vancouver, Canada	School and Workshop on Reproducible Research in Computational Geophysics
2007	Austin, TX	Short Course on Using and Extending MADAGASCAR
2008	Golden, CO	Implementation Workshop
2009	Delft, the Netherlands	School on Reproducible Computational Geophysics
2010	Houston, TX	School on Reproducible Computational Geophysics and Hands-On Workshop

Table 1: MADAGASCAR annual events