

Functional Requirements for Developing ERAS – A Portal to Support Collaborative Geomechanical Simulations*

Maria Julia Lima, Melissa Lemos, Fernanda Pereira, Rodnei Couto, Deane Roehl

Tecgraf/ PUC-Rio Institute
PO Box 38097 – Rio de Janeiro – RJ – Brazil

{mjulia,melissa,nandalgp,rodnei,deane}@tecgraf.puc-rio.br

Abstract. *One of the most important tasks in geomechanical research is executing analytical and numerical simulations to understand geomechanical phenomena. In order to attain this objective, researchers have to prepare data to perform the simulations, build the models that define the appropriate physical representation and the mathematical modeling of the problem, run a computer system capable of simulating the phenomenon, and visualize and interpret the results. This paper presents the main functional requirements to support the development of solutions that encompass the simulation of geomechanical problems, taking into account a collaborative environment with access to an efficient computer infrastructure. The paper also describes ERAS, a portal developed according to these requirements, highlighting the advantages it brings to researchers in this area.*

Resumo. *Uma das tarefas mais importantes na pesquisa de geomecânica é a execução de simulação numérica e analítica para compreensão dos fenômenos geomecânicos. Com este objetivo, os pesquisadores preparam os dados para realizar as simulações do problema em questão; constroem o modelo que define a representação física apropriada e a modelagem matemática do problema; executam a computação capaz de simular o fenômeno estudado; e visualizam e interpretam os resultados. Este artigo apresenta os principais requisitos funcionais para apoio ao desenvolvimento de soluções que envolvam simulação de problemas geomecânicos, considerando um ambiente colaborativo com acesso a uma infraestrutura computacional eficiente. O artigo também apresenta o ERAS, um portal que está sendo desenvolvido de acordo com estes requisitos, destacando as vantagens que ele traz aos pesquisadores da área.*

1. Introduction

The application of analytical and numerical simulations to geomechanical problems is strategic to the Oil and Gas Industry. Research in this area provides a fundamental comprehension of subsoil geomechanical phenomena, which can optimize the industry's expenditure, while reducing geomechanical risks associated to drilling, completion, and development plan. Reservoir simulators, for example, have a fundamental role in the planning and optimization of new production field developments. Operating a deep-water well entails high costs, and thus minimizing the perforation risks of a dry well or

* This work received funds from Shell Brazil and ANP "Commitment to Investments in Research and Development", in the scope of the "BG- 47 Coupled Geomechanical Modeling" research project.

decreasing the number of wells necessary for the development of a new field can result in substantial savings.

Overall, during a project, the researchers' work in understanding subsoil geomechanical problems entails: (a) collecting, processing and interpreting data; (b) modeling, which defines the appropriate mathematical description of the problem; (c) processing, which executes the computer system capable of simulating the model; (d) analysis and visualization (post-processing), in which results are verified and studies are conducted; and finally (e) report generation and result dissemination. We note that this workflow is not always performed sequentially. In some cases, during analysis and visualization, the researchers may decide to go back to previous stages to prepare new data, revise the model or execute the simulation with new parameters, for example.

The teams involved in these projects are usually multidisciplinary, with professionals working in collaboration in different knowledge areas. In many cases, the simulations demand high computational processing power, producing a huge volume of data. Without the support of tools and an adequate computational infrastructure, researchers may spend a lot of energy and time collecting, organizing, processing and visualizing this data. Furthermore, the experience and the results obtained in those projects are a valuable knowledge source for future project enquiries and decision-making. Thus, the need for organizing, sharing, and reusing data, models, simulators, results, processes and knowledge is of great importance.

The importance of collaboration among researchers has led the development of social environments that support the whole lifecycle of experiments in scientific scenarios. E-ScienceNet [Classe et al. 2017] aims to provide support for geographically distributed researchers in the processes of creating, implementing and sharing experiments. ^{my}Experiment [Goble and De Roure, 2007] is an online environment that creates a social network of scientists who collaborate and share Taverna workflows [Wolstencroft et al. 2013]. The number of registered users and workflows available in ^{my}Experiment shows the potential of this kind of environment for scientific communities. An alternative is CrowdLabs [Mates et al. 2011], which adopts the model of social web sites to provide another rich collaborative environment for scientists.

This paper presents the main functional requirements to support research projects that involve simulation of geomechanical problems, taking into account a collaborative environment with access to an efficient computational infrastructure. These requirements stem from many years of experience of researchers at Institute Tecgraf/PUC-Rio working in this area. Furthermore, the paper describes ERAS, a portal developed by Tecgraf/PUC-Rio Institute, based on these requirements, as part of a research program in Geomechanics, in which Cambridge University and Berkeley University are also partners, as well as an O&G company.

The paper is organized as follows. Section 2 presents the functional requirements to support the development of solutions that involve simulations of geomechanical problems. Section 3 describes the main functionalities in ERAS Portal. Section 4 shows the services available in ERAS Portal to develop applications that use geomechanical simulations. We present an application for slope stability problems as a use case of the Portal. Finally, Section 5 presents comparisons with available related systems, final comments and future work.

2. Functional Requirements

2.1. Collaborative Project

Research that applies simulation in geomechanical problems is usually conducted by multidisciplinary teams, with professionals working collaboratively in different knowledge areas such as physics, engineering, mathematics, computer science, geology and geophysics. Therefore, it is crucial to have a work environment that promotes the exchange of knowledge and experience among researchers during an on-going project, as well in future projects.

In a geomechanical collaborative environment, researchers not only share data and simulators but also all the stages of the work in progress. In this context, data provenance is essential as it registers input data, models, simulators and results, providing a historical record of the data and its origins. This scenario allows for the reproducibility of results, which can be used to validate simulation results and improve future researchers.

2.2. Collaborative Simulation

The design and development of simulators is shifting towards collaborative paradigm. Therefore, simulation environments for such a paradigm need to take into account the cooperation between design teams in a distributed environment. Supporting the creation of a simulation can happen at different levels, all of which can benefit from working in collaboration. In the basic level, the researcher will use a pre-existing simulation that was carried out by another professional. Furthermore, he will need to know the application domain to create a model with the correct data (such as meshes, boundary conditions, etc.). In the advanced level, the researcher will develop new simulators, based on the available knowledge of the appropriate physical representation and the mathematical model of the problem.

Therefore, it is necessary to have an infrastructure in which expert researchers (from the advanced level) share simulators, models and existing data in such a way that they can be used by the other researchers from the basic level. Likewise, it is suitable to have a more advanced infrastructure, such as tools to build new simulators, to meet the demands of the advanced level.

2.3. The Infrastructure to Execute Simulations

Geomechanical simulations usually include complex calculations, which can be very time consuming in terms of execution. It is essential to have an infrastructure that allows for executing simulations applying distribution and parallelization techniques, taking advantage of heterogeneous resources such as clusters with GPU, fast I/O and CPUs with high speed cores and interconnectivity. Current resources such as cloud services, internet protocols and virtual machines offer accessible and appealing infrastructure that reduces cost with IT operations, without losing performance, availability and stability.

Furthermore, in a large scale geomechanical project, simulations often generate data at high speed and usually process a large volume of data of different types, which basically determines the characteristics of a big data scenario. Considering the complexity of the calculations and possibly re-executions, generating even more data, it is fundamental to have data intensive computing infrastructure that ensures run-time data

access efficiency, taking into account strategies to minimize data transfer and network latency.

2.4. Pre- and Post-Processing Applications

Data preparation to perform a simulation and the visualizations of the results are stages of research projects that use numerical tools. Geomechanics projects share some common characteristics such as complex geometry, mesh generation, definition of boundary conditions and result visualization. Computer graphics technologies are normally applied in the development of tools that support researchers in these stages. In geomechanical projects, the integration of a computational infrastructure necessary for simulation processing with tools that perform the pre-processing (preparation of input data) and post-processing (visualization of results) still represents a challenge. In many cases, researchers execute these stages independently and perform the integration manually. One example is data transfer and format conversion between an HPC cluster remote environment, which executes simulations, and desktops where users prepare and visualize the results.

Lately, high speed network services and advances in WebGL technologies for 2D and 3D drawing on browsers have allowed using the internet to build pre- and post-processing applications. Therefore, these new applications require a set of services in public or private clouds, that aims at integrating functionalities typical of a collaborative environment and hiding the complexity underlying the development stack of those applications.

3. The ERAS Portal

The ERAS Portal is developed by Institute Tecgraf/PUC-Rio according to the requirements presented in the previous section. It is currently used by researchers in geomechanical projects. Through the portal, researchers have access to software artifacts, such as models, documents and computational resources to access, execute and monitor their simulators, while working in a collaborative environment.

The conceptual architecture of ERAS presents the following modules: (a) **Dashboard Panel**: a front-end that allows researchers to define, execute and monitor their simulation, and also visualize the results. This module also provides tools for publishing and sharing data among researchers who take part in the same project. (b) **Execution Environment**: provides access to a data sharing and simulator execution infrastructure in a heterogeneous and distributed environment. (c) **Simulation Environment**: offers a catalogue of existing simulators that can be used by researchers by inserting an input data set. In its advance mode, this module supports the building of new simulators. (d) **Knowledge Base**: stores data and information on the research progress shared in the Portal, and also data provenance from simulations that were executed. (e) **Access Control**: provides control regarding access restrictions according to user profile and the projects they take part in. Next, we present the main features of ERAS. A detail discussion on building new simulators can be found in [Mendes 2016].

3.1. Collaborative Projects

Users organize their work area in the Portal by creating projects. A project is a space for collaboration, storing and sharing of resources used in the simulations, such as input files

with simulation model data, output files with simulation results. The project owner is the user who created it, and he can choose to share it with specific users. The type of sharing can be read-only or read and write. Besides the file storage area used in simulations, a database maintains the project's meta-information, its recent activities, and simulation execution history. This track record is an important register of the parameters used in the simulation and the provenance of the results.

Every project has a forum for discussions among its participants. The forum allows the user to reference the project files and the simulations submitted to execution, facilitating the discussion about the progress of the on-going research. Statistical information, such as project space usage and simulation history, are available to its participants.

3.2. Execution of Simulations

The Portal is responsible for managing and monitoring the execution of simulations in a distributed and heterogeneous environment. Users that take part in the same project can monitor in real-time the processing results. The Portal offers a monitoring dashboard, shown in Figure 1(a), which presents the execution output logs in text and chart format, thus facilitating execution monitoring and the convergence analysis (or not) of the simulation results.

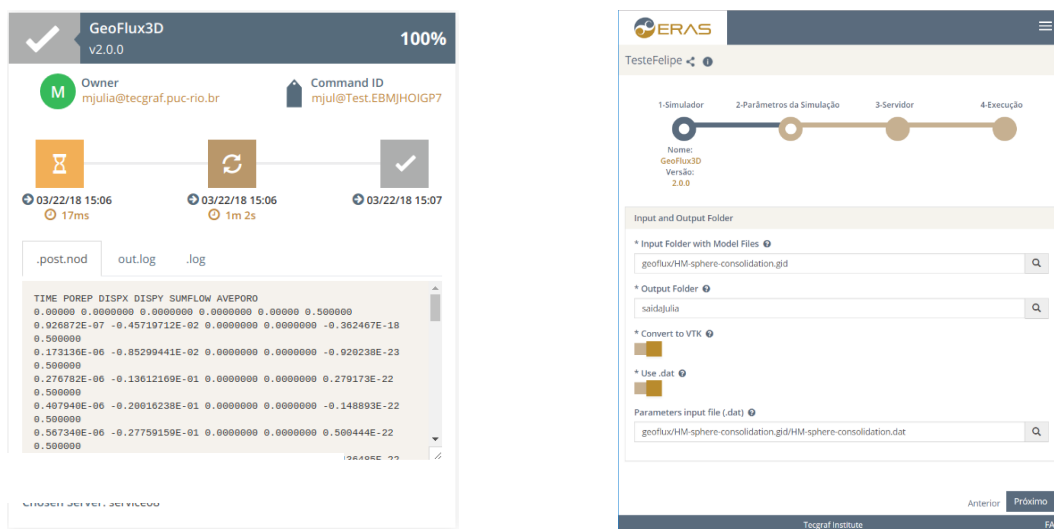


Figure 1. (a) Monitoring Dashboard (b) Form generated from a simulator configurator.

The execution of the simulations can be done in machines available in the Portal infrastructure. This infrastructure is comprised of physical and virtual machines with heterogenic resources and shared access to a data storage area. Monitor agents, which execute in the environment machines, provide information about computational resources, such as CPU, memory, disk and other specific properties set by the ERAS' administrator in each agent. In this way, the available infrastructure is deemed as a virtual cluster of computational resources. The scheduler adopted by Eras Portal's execution model submits the simulations to the cluster's machines that have the necessary requirements for the execution. Before submitting the simulation for execution on the selected machine, ERAS orchestrates the file staging needed. The Portal maintains an access control to its infrastructure that determines which user profiles can execute

simulations in designated machines in the environment. The access control can also limit which simulators and machines can be used by a group of users, establishing resource reserves for certain projects.

ERAS uses the middleware CSGrid [Lima, 2005] as backend. It mediates access to the computational resources of the ERAS environment. The CSGrid has been used in different projects. One of them is the mc2 tool, which supports the prototyping of scientific portals at SINAPAD [Gomes et al. 2015]. While already allocated resources (such as HPC and VM) are the main target hosts used by CSGrid, it can be integrated to other middlewares that manages the components of multiple virtualized infrastructures, including private and public clouds.

3.3. Sharing of Simulators

The Portal has a simulator repository that lists available simulators to which authenticated users are granted access. Each simulator has a configuration file that defines the input and output parameters which ERAS uses to create the simulation execution command in a machine of the environment. The dashboard client uses this same configurator to dynamically generate a web form to enter the parameters during submission. The simulator configurator specification language provides different types of parameters such as file, integer, float, list and table. This configuration feature allows new simulators installed in the portal to have a ready-made interface for parametrization and execution submission. Figure 1(b) shows a form generated from a simulator configurator that uses different types of input parameters for execution.

Simulators can be combined to build workflows, allowing users to create new orchestration models that can be saved in the project area. Therefore, existing simulators are also handled as re-usable and modular components that can be used across many different workflows.

4. Use Case: An application for Slope Stability Analysis

Besides the simulators, the Portal also presents a list of available applications in its dashboard. The applications offer a richer interface for the input data preparation stage (pre-processing) and the analysis and visualization of results (post-processing) stage. These applications employ ERAS Portal's API REST services to access the functionalities described in the previous section. An application for slope stability analysis is an example of a tool developed entirely with ERAS's API REST services.

Slope stability analysis is a procedure commonly used for verifying the safety of natural and artificial slopes (barriers, embankments, etc.). This assessment is done through the Safety Factor (SF), which provides a quantitative indicator of how far the slope is from rupture. Different approaches and methodologies are applied to assess the risks associated to engineering projects. Hybrid methods that combine the Shear Strength Reduction Method (SSRM) with probabilistic analysis and artificial neural networks are some of the techniques found in the literature [Shu and Gong, 2016]. The development of a probabilistic accumulated distribution curve as well as a neural network response surface need several simulation results to achieve a higher accuracy response. Even the SSRM needs a series of results to provide the SF. Therefore, the Slope Web application developed over the ERAS Portal services aims at providing support to this type of project.

In the Slope Web application, the slope geometry is built in a fast and simple manner. When filling in the input data, such as lithology and slope angle, a pre-visualization of the slope is generated, providing the user with a visual notion of the model being created. The user provides the specific weight, elastic properties (Young's module and Poisson's coefficient), and resistance parameters (cohesion and friction angle) of the layers. Finally, the user provides the position of the phreatic line through coordinates of points on the line, or through the "Pen" tool that allows drawing a line. Figure 2 illustrates the Slope Web application, which uses the ERAS Portal services.

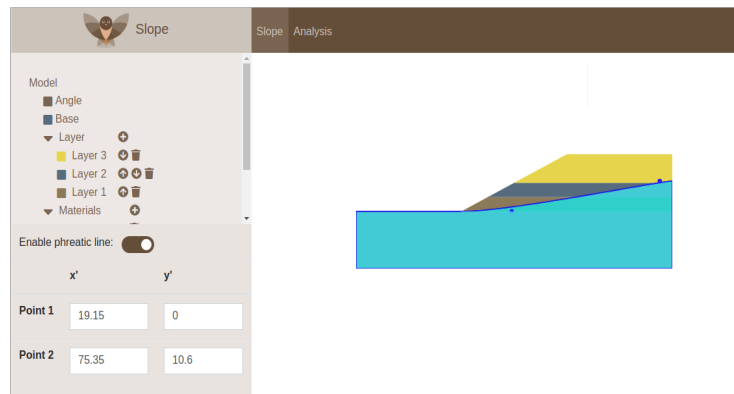


Figure 2. ERAS Portal Slope Web application

To obtain the SF, the user can define a search interval and precision of interest. Under user command, the application submits the simulation in ERAS Portal. The simulator uses the SSRM to calculate the SF, adjusting constantly the resistance parameters and assessing the slope instability through a non-linear finite element analysis. By the end of the execution, the application shows the 2D model generated by the simulation on the ParaViewWeb [Jourdain, 2010] application, integrated to the ERAS Portal for visualization of post-processing results. All the slope models used in the analysis are saved in the user project area, which can be shared with other researchers.

5. Conclusions

In this paper, we presented the main requirements to support the development of solutions that encompass the simulation of geomechanical problems, taking into account a collaborative environment with access to an efficient computer infrastructure. We also introduced the ERAS Portal, developed based on these set of requirements. The Slope Web is an example of an application oriented towards pre- and post-processing and developed over ERAS services.

ERAS is somewhat similar to many other related works in this area, with different approaches though. Different from E-ScienceNet peer-to-peer network architecture, ERAS is based on a client/server model. Currently, ERAS is being evolved to adopt a keyword-based query processing engine [Garcia et al. 2017]. Similar to E-ScienceNet assistance based on domain ontologies, this new capability will help users to search resources. The objectives of ^{my}Experiment and CrowdLabs are very related with the collaboration requirements (sections 2.1 and 2.2) that have driven our development in ERAS for Geomechanical users. However, ERAS project does not focus on Taverna and VisTrails workflows as the main shareable resources.

At present, the first version of ERAS Portal is operational with approximately 50 users, 10 simulators, and 10 machines for simulation execution. Besides the SlopeWeb,

new applications, such as well modeling, are being developed using the ERAS library. Future work includes integrating the framework GeMA [Mendes et al. 2016] to ERAS, which will provide support to the building of new simulators in the Portal. GeMA implements some important concepts of extensibility using of plugins and abstract interfaces, configurable orchestration and fast prototyping. Other future improvements also include the advanced searching mechanism that is being added to ERAS.

References

- Classe, T., Braga, R., David, J.M., Campos, F. and Arbex, W. (2017) “A Distributed Infrastructure to Support Scientific Experiments”, *J. Grid Comput.* 15, 4.
- García, G.M., Izquierdo, Y.T., Menendez, E., Dartayre, F. and Casanova, M.A. (2017) “RDF Keyword-based Query Technology Meets a Real-World Dataset”, *Proc. 20th International Conference on Extending Database Technology (EDBT)*, March 21-24.
- Goble, C.A. and De Roure, D. (2007) “^{my}Experiment: social networking for workflow-using e-scientists”, *WORKS '07: Proceedings of the 2nd workshop on Workflows in support of large-scale science*.
- Gomes, A. T. A., Bastos, B. F., Medeiros, V. and Moreira, V. M. (2015) “Experiences of the Brazilian national high-performance computing network on the rapid prototyping of science gateways”, *Concurrency and Computation: Practice and Experience*, v. 27 (2), p. 271–289.
- Griffiths, D. V. and Lane, P. A. (1999) “Slope stability analysis by finite elements”, *Géotechnique*, v. 49, n. 3, p. 387-403.
- Jourdain, S., Ayachit, U. and Geveci, B. (2011) “ParaViewWeb: A Web Framework for 3D Visualization and Data Processing”. *International Journal of Computer Information Systems and Industrial Management Applications*, v. 3, p. 870–877.
- Lima, M. J., Melcop, T., Cerqueira, R., Cassino, C., Silvestre, B., Nery, M. and Ururahy, C. (2005) “CSGrid: um Sistema para Integração de Aplicações em Grades Computacionais”, In: *Salão de Ferramentas do 23o. Simpósio Brasileiro de Redes de Computadores*, Fortaleza, v. 2. p. 1207-1214.
- Mates, P., Santos, E., Freire, J. and Silva, C.T. (2011) “CrowdLabs: Social Analysis and Visualization for the Sciences”, *International Conference on Scientific and Statistical Database Management SSDBM 2011*, p. 555-564.
- Mendes, C.A.T., Gattass, M. and Roehl, D. (2016) “The Gema Framework: An Innovative Framework for the Development of Multiphysics and Multiscale Simulations”, *Proceedings of the ECCOMAS Congress 2016*.
- Shu, S. X. and Gong, W. H. (2016) “An artificial neural network-based response surface method for reliability analyses of c-slopes with spatially variable soil”, *China Ocean Engineering*, v. 30, n. 1, p. 113-122.
- Wolstencroft, K., Haines, R., Fellows, D., Williams, A., Withers, D., Owen, S., Soiland-Reyes, S., Dunlop, I., Nenadic, A., Fisher, P., Bhagat, J., Belhajjame, K., Bacall, F., Hardisty, A., Nieva de la Hidalga, A., Balcazar Vargas, M., Sufi, S., Goble, C. (2013): The Taverna workflow suite: designing and Acids Research, 41(W1): W557–W561. <https://doi.org/10.1093/nar/gkt328>