LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

LLNL-TR-860182

SMART Deliverable 6.1.2a: Application of the ORION tool to the IBDP Carbon Storage Site

K. A. Kroll, C. S. Sheramn, G. M. Geffers, C. Wang, C. Layland-Bachmann

February 6, 2024

Disclaimer

This work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

Technical Report: SMART Deliverable 6.1.2a: Application of the ORION tool to the IBDP Carbon Storage Site

Kayla A. Kroll[1], Christopher S. Sherman[1], Gina-Maria Geffers[1], Chaoyi Wang[1], and Corinne Layland-Bachmann[2]

[1]Lawrence Livermore National Laboratory
[2]Lawrence Berkeley National Laboratory

January 30, 2024

1 Introduction

Forecasting and managing potential induced seismic activity is one of the challenges facing commercial-scale geologic carbon sequestration (GCS), as well as other geologic energy extraction and byproduct disposal technologies. Historically, the process to develop robust, science-based forecasts of induced seismicity has required an integrated effort from experts in seismology, geomechanics, and reservoir engineering to manage data, develop and evaluate models of subsurface processes, and to calibrate and interpret the results from a range of models to understand site behavior relative to prescribed standards and in the context of uncertainty in geologic characterization data, forecasting models, and operational scenario uncertainty. The Operational Forecasting of Induced Seismicity (ORION) toolkit is an open-source, observation-based forecasting toolkit that is being co-developed by two U.S. DOE-funded initiatives: the National Risk Assessment Partnership (NRAP) and the Science-informed Machine Learning for Accelerating Real Time Decisions in Subsurface Applications (SMART) Initiative. ORION is designed to provide functionality to support decision making about seismic hazard analysis and risk management for GCS stakeholders ranging from the public to site operators to expert seismologists. The tool, which is written as open-source code in the Python programming language, is composed of a desktop graphical user interface (GUI) and an underlying forecasting engine. The forecasting engine uses available reservoir properties, well and fluid injection scenario details, and observed seismic catalog data as inputs to produce a set of temporal and spatio-temporal seismic forecasts.

The in-situ fluid pressurization rate ($\dot{p}$) is one of the key factors that drives the induced seismicity model forecasts. Rather than evaluating a full-physics model of the reservoir, ORION has options that allow the user to estimate ($\dot{p}$) using efficient closed-form solutions (e.g., the Theis model), machine learning models, other reduced order reservoir models, or to interpolate pre-computed results from external tools. The forecasting engine implements multiple credible forecasting methods that are based on statistical principles and/or physical models. ORION uses a decision-tree model to mediate between this set of credible forecasts and produces an ensemble model that accounts for model uncertainty and supports robust decision making. In this report, we will review the technical basis for ORION, briefly describe the toolkit functionality, and demonstrate the utility of the ORION tool through a retrospective application to data collected during injection of CO2 at the Illinois Basin Decatur Project.

2 Background

2.1 Design Decisions

The design of ORION was guided by four primary pillars: 1) flexibility, 2) adaptability, 3) impartiality, and 4) automaticity. Firstly, design decisions were adopted to improve the *flexibility* of the tool with respect to input data requirements. For example, there are options to upload datasets (from the

simplest to the most robust) in various formats that can be read by built-in parsers or options to collect data from publicly available websites if a user doesn't have access to higher level data sources. Next, ORION was designed to be easily *adaptable* to the skills, data, and needs of the end-user, which can range from the general public, through regulators and operators, and up to Ph.D. level expertise (i.e. Super-users). Each end-user will have different goals and access to data when working with ORION, so a start-up wizard was created to identify the type of end-user and then reduce or expand number of interactive features on the dashboard. The features that get presented are those that are thought to be most interest to the identified end- user. As above, if end-users to do upload datasets, data is gathered from publicly available sources to produce forecasts.

Additionally, the authors recognize that there are many methods for forecasting earthquakes, therefore, ORION was designed to be *impartial* to the choice of forecasting method. Forecasting methods are historically rooted in the statistical properties of earthquake catalogs or the physics of earthquake occurrence. More recently, machine learning models have been derived for earthquake forecasting. All types of forecasting methods are employed within ORION. The pillar of impartiality also applies to the computation of relevant model parameters and reservoir pressure. When more than one method exists, all methods are used. Forecasts related to these methods are then combined into an ensemble spatial and temporal forecast. Finally, when possible ORION take measures to set reasonable default forecasting model parameters. If multiple methods exist to compute a parameter or if the parameter is thought to vary wildly depending on site-characteristics, ORION employs one or more *impartial* methods to *automatically* compute the relevant parameters from the input datasets. If more than one method exists to estimate a given parameter, a forecast is generated for all parameter estimates and passed to the ensemble forecast module.

ORION is organized as a pure-python pip installable package, and is hosted on the NRAP gitlab (https://gitlab.com/NRAP/orion). ORION contains a core seismic forecasting engine, a legacy tkinter-based GUI, and a modern STRIVE-based GUI. It is written using a modern, object-oriented python style, and follows the SMART Software Quality Assurance (SQA) guidelines.

2.2 Forecasting Engine

The forecasting engine contains all of the necessary components for producing an ensemble seismic forecast for a given site, and can be run independently from any GUI. Key inputs to the forecasting engine include a seismic catalog, a pressure model, and forecasting hyperparameters. The seismic catalog can be loaded from a local source (.hdf5, .csv, etc.), or can be acquired from an online database such as ComCat. Pressure models can be loaded from a local table file (.hdf5 or .csv), be calculated using a closed-form solution applied to reservoir and fluid injection data, through a pre-trained ML model, or by interfacing with the SMART USM. After loading data, the forecasting engine processes the seismic catalog and pressure model to estimate key forecasting parameters such as the Gutenberg-Richter b-value. The engine then evaluates a selection of statistical and physics-based forecasting models, compares them to the observed data, and then estimates the ensemble forecast. The forecasting engine is designed to be configured using a json-format file, which can be written by hand or produced by one of the following GUI frameworks.

2.3 Visualization

The original version of ORION v0.5 was built to run on an individual desktop or laptop through a graphical user interface (GUI) using Python and the "tkinter" package for visualization and was released in early 2023 [Kroll and Sherman, 2023] (hereinafter referred to as "legacy-ORION"). The legacy-ORION version uses simple API entries placed throughout the datastructure, which are parsed by the vizualization engine and used to create a desktop visualization on-the-fly. The resulting figure and configuration pages for the legacy GUI are organized using a tab-based interface, and are controlled by a set of buttons at the bottom of the screen and a set of drop-down menus at the top. This interface also includes a quickstart wizard, which is designed to identify different types of users and simplify the ORION configuration process.

Since this release, the ORION developers have been moving towards a web browser interface based on the Python package "plotly dash", within the SMART Tool Rapid Visualization Environment (STRIVE). The STRIVE package is derived from the legacy-ORION visualization engine, and relies on the same API entries. This version of ORION could be hosted on the cloud or on a user's local machine,

in which case STRIVE would point the user to a local IP address (e.g.: http://127.0.0.1:8050/). This version of ORION has not yet been released (hereinafter referred to as the ORION-STRIVE), however, we will present both visualizations from both versions in this document.

3 Data and Methods

The two fundamental input datasets required to produce an induced seismicity forecast are the pressurization rate ($\dot{p}$) and a catalog of seismic event locations containing the origin time, latitude, longitude, depth, and and magnitude of each event. These datasets can be uploaded by a user or derivatives can be pulled from public sources [i.e. state regulatory agencies and the United States Geological Survey (USGS)]. In Sections 3.1 and 3.3, we discuss how these datasets are derived and how they are used in the forecast.

3.1 Reservoir Models

The primary driving force of all induced seismicity forecasts within ORION is the Coulomb stressing rate ($\dot{S}$), given by

$$\dot{S} = \dot{\tau} + \mu(\dot{\sigma} + \dot{p}) \tag{1}$$

where $\dot{\tau}$ and $\dot{\sigma}$ are the tectonic shear and normal stressing rates in MPa/yr, respectively, projected onto optimally oriented faults, μ is the nominal coefficient of friction (generally taken to be 0.6), and $\dot{p}$ is the pressurization rate in each grid cell related to the injection and/or production of fluids. Note that the normal stresses are reckoned positive in tension.

Estimates of $\dot{p}$ can come from simple analytic solutions of pressure or from full-physics, high-fidelity reservoir simulations. If ORION only has access to injection rate information, $\dot{p}$ is taken to be the time-derivative of the pressure computed with the built-in Theis equation [Ferris et al., 1962]. ORION can also accept model results of all types of academic and commercial reservoir simulators (e.g. GEOS [Johnson et al., 2013], Fast Marching Method (FMM) [Yang et al., 2017], TOUGH-FLAC [Rutqvist and Moridis, 2008], CMG, Eclipse, etc.,). Irrespective of the source and unless otherwise supplied by the user, the time-derivative of the pressure is computed, $\dot{p}$, and passed to Equation 1 to calculate $\dot{S}$ on a user-defined grid and spatially interpolated for visualization purposes. Examples of the pressure distribution computed for the Illinois Basin Decatur Project with a TOUGH-FLAC model [Luu et al., 2022] and FMM and then visualized within legacy-ORION and ORION-STRIVE are presented in Figures 1 and 7, respectively. Note, $\dot{S}$ and $\dot{p}$ are also visualized if the user is interested.

3.2 Seismicity Forecasts

At present, ORION has the ability to compute seismicity forecasts based on three methodologies, 1) the seismogenic index model (SI; [Langenbruch et al., 2018]), 2) the coupled Coulomb rate-state model (CRS; [Kroll and Brudzinski, 2023]), and the 3) Rate-state Ordinary Differential Equation model (RSODE; [Luu et al., 2022]).

The SI model within ORION has been modified such that

$$SI_j = \log_{10} \mathrm{N}_j - \log_{10} \left\{ \sum_j [\dot{S}_{(j,t)})^2] \right\} + \mathrm{bM_c}, \tag{2}$$

$$\dot{R}_{(j,t)} = [\dot{S}_{(j,t)})^2] 10^{SI_j - \mathrm{bM_c}} \tag{3}$$

Where the SI is computed for every grid cell, j, at each time step where t is less than the current time, based on the number of events, N_j, and $\dot{S}_{(j,t)}$ with that grid cell. The Gutenberg-Richter b-value for all events larger than the magnitude of completeness, M_c, are computed from all events that occurred prior to time t and are applied to every grid cell. The event rate is then computed for each grid and time step based on Equation 3. All parameters necessary to compute an SI forecast are estimated automatically within ORION. To estimate M_c, the current default method is the b-value stability method [Cao and Gao, 2002, Wiemer and Wyss, 2000]. To estimate the b-value, we use two methods: the maximum likelihood estimate (MLE; [Aki, 1965]) which requires M_c and is the current

default method in ORION, as well as the *b-positive* method [van der Elst, 2021], which does not require an estimate of M_c.

The CRS model is defined based on the following equations, modified from [Dieterich, 1994] and [Kroll and Brudzinski, 2023]:

$$R = \exp\left(\frac{\Delta CFS}{a\sigma}\right) \tag{4}$$

$$\dot{R} = \frac{1}{\frac{1}{\dot{S}f_r} + \left(\frac{1}{R_0} - \frac{1}{\dot{S}f_r}\right)\exp\left(\frac{-\dot{S}t}{a\sigma}\right)}, \tag{5}$$

where ΔCFS is the static Coulomb stressing rate related to slip during a stress step (e.g. mainshock), a is a constitutive parameter, σ is the effective normal stress, R_0 is the seismicity rate at (t-1), and the scaling term f_r is the ratio of the background seismicity rate r and the background shear stressing rate $\dot{S}_r$. These terms can be difficult to estimate in regions where historic seismicity catalogs are incomplete and where tectonic stressing rates are low. In the absence of empirical data, this parameter can be approximated by fitting the observed number of events in the injection period since this terms acts like a scaling-factor in equation 5. In this formulation, the stressing rate between time steps is assumed to be constant. At large t, the seismicity relaxes to the steady-state rate where $R_{ss} = \dot{S}/\dot{S}_r$, which is the seismicity rate that would result from some arbitrary stressing rate, $\dot{S}$. R_{ss} is equal to r when $\dot{S} = \dot{S}_r$. The unknown parameters in Equations 4 and 5 are a and f_r and are automatically fit to the training data within ORION. This set of equations applies when short-term clustering of events in time and space are present in the observed seismicity catalog and can be used to forecast short-term transient rate changes as well.

Finally, the RSODE model is based on the following equations from [Segall and Lu, 2015]

$$\dot{R} = \frac{R}{t_a}\left(\frac{\dot{S}}{\dot{S}_0} - R\right) \tag{6}$$

where R is the seismicity rate relative to the background rate, $\dot{S}_0$ is the background Coulomb stressing rate and $t_a = \frac{a\dot{\sigma}}{\dot{S}_0}$. Here, $\dot{\sigma}$ is the background normal stressing rate. The ordinary differential equation is solved using a fifth order adaptive time step Runge-Kutta-Fehlberg algorithm [Fehlberg, 1970] and can only be fit to "declustered" seismicity catalogs, or catalogs with the short-term space-time clusters removed.

3.3 Seismicity Data

Due to the data-driven forecasting methodology, ORION requires a catalog of observed seismic events including the origin time, location, and magnitude. Ideally, these catalogs will have a low magnitude of completeness (e.g. $M_c = 0$ or lower) and location errors on the order of 10's of meters in the horizontal and vertical directions. The quality and completeness of the catalog has direct implications for the uncertainty in the seismic forecasts, as many forecast parameters are directly fit to the seismicity rates and frequency-magnitude distributions. For pre-injection, forecasts used for permitting or other applications, is especially important in the pre-injection periods and in regions where background seismicity is infrequent.

3.4 Parameter Estimation

Of the equations presented in Equation 1 and Section 3.2, there are 12 parameters that must be estimated (e.g. b, M_c, N, ΔCFS, $a\sigma$, $\dot{S}_0$, f_r, $\dot{\tau}$, and $\dot{\sigma}$). Of these, at least two have competing methods for estimation (e.g. b, M_c). Additionally, the RSODE method only applies to "declustered" catalogs, or catalogs, where short-term spatial and temporal clustering of events (i.e. aftershocks) have been excluded. Within ORION, we employ three methods to decluster catalogs [Gardner and Knopoff, 1974, Reasenberg and Jones, 1989, Zaliapin and Ben-Zion, 2016]. When more than one method exists, all methods are employed and a forecast generated for each method where those parameters are relevant. Automatic parameter estimation schemes have been employed within ORION, however, super-users have the ability to overwrite these calculations, if they believe they have better data to draw on or a preferred method for estimation.

Time Period	# Observed Events	# $M \geq M_c$	b-value
2011 - 2018	5420	2244	0.92
2011 - 2014	4889	2132	0.91
2014 - 2018	531	112	1.01

Table 1: Relevant statistical parameters computed from the NORSAR catalog for IBDP.

4 Application of ORION to IBDP

The Illinois Basin Decatur Project (IBDP) was a carbon storage demonstration project where ~1 million metric tonnes of CO_2 were injected into the Mt. Simon reservoir between 2011 and 2014. Seismic activity at the IBDP site was recorded by a network of two 4-component borehole geophones (installed at 1493m and 1660m in the CCS1 injection well). A string of 29 3-component geophones was installed between 418 and 844 m in the GM1 monitoring well, located 56 m northwest of CCS1. Two additional geophones were installed at 97 and 164 m above sea level within GM1. These sensors were operational for ~1.5 years prior to injection and continued to monitor through 2018. NORSAR processed this data and produced a relocated catalog of 5,420 earthquakes from the ~19,000 detections made by the network between December 2011 and July 2018. [Dando et al., 2021]. Table 1 lists the catalog statistics for the entire catalog (2011 to 2018), the injection period (2011 to 2014), and the post-injection period (2014 to 2018). All events with $M \geq M_c$ between 2011 and 2018 are used for forecast development and application within ORION. This seismic data is visualized within legacy-ORION (Figures 2 and 3) and in ORION-STRIVE (Figure 8).

There are publicly available models of the reservoir pressure due to injection at the IBDP site, one computed with the TOUGH-FLAC model [Luu et al., 2022] and the other with the Fast Marching Method (A. Datta-Gupta, *personal communication, 2023*). Both are incorporated into ORION along with the seismic catalog described above. Visualizations of the pressure data in legacy-ORION (with TOUGH-FLAC) and ORION-STRIVE (with the FMM pressure model) are presented in Figures 1 and 7, respectively. All spatiotemporal pressure data are converted to pressurization rate by taking the time-derivative for each grid cell and used to calculate a spatially varying Coulomb stressing rate via Equation 1. The Coulomb stressing rates are then used to produce spatial and temporal forecasts. Visualizations of the spatial forecasts within legacy-ORION and ORION-STRIVE are presented in Figures 4 and 8, respectively. Temporal forecasts are generated by collapsing all spatial data onto a single point and using all available data for training and forecasting. Visualization of the temporal forecasts within legacy-ORION and ORION-STRIVE are presented in Figures 5 and 9, respectively.

5 Application Challenges

Since its inception, the development team has made significant strides with ORION. The software is capable of forecasting seismicity at the site and basin-scale, is suitable to a range of end-user experience levels and needs, runs on high-quality, user-provided data or publicly available data, makes data-driven, automated estimates of necessary forecast parameters, and applies before, during, and after injection.

There are a number of challenges, nuisances, and/or open questions that are necessary to address in order to facilitate accurate forecasts. These challenges are itemized and explained below, however, please note, this is not meant to be an exhaustive list and summarizes many challenges we have encountered while applying ORION to the IBDP dataset.

1. There are cases where the pressure distribution does not lead to an accurate forecast because the spatiotemporal variation seismic event locations themselves do not propagate away from the well(s) with time, but respond the interaction of fluid pressures with sources of heterogeneity in the subsurface that vary in space. This leads to situations where there are non-zero values of $\dot{S}$ in grid cells, but there are no seismic events. The opposite is also true, sometimes there are seismic events in grid cells where $\dot{S}$ is close to or identically zero. Since the forecasts are mostly dependent on $\dot{S}$ and only fit to seismic data for parameter estimation, there will always be a discrepancy in the spatial forecasts and the actual event locations (in time and space). At the highest level, this will remain true unless some sort of spatial heterogeneity (e.g. permeability or geology) is included in the reservoir model. One method for overcoming this is to simply

compute temporal forecasts, where all data are lumped together and the spatial component is not considered. A second method to address this is to weight the spatial forecast by the event occurrence. Temporal forecasts are already provided within the ORION framework and we are currently exploring methods to address the latter. At present, the goal is to weight the computed spatial forecasts by the seismogenic index (Equation 2). Testing of this methodology is underway.

2. The accuracy of the forecasts is also dependent upon the spatial region and temporal window length used to train and test the models. For reasons mentioned above, there are often many grid cells or time windows with no data, which are difficult to treat in the models. Therefore, care must be taken when selecting default values of the grid size and window length and should be optimized to produce the smallest misfit between the observations and the forecasts. Additionally, the appropriate values will likely be site- and size-specific. At present, we are developing automated methods to test the forecast on a variety of grid-sizes and training window lengths. The grid-size and window length that produce the lowest misfit in the training period are retained for the forecasting period.

3. In future applications, knowledge of fault locations *a priori* may provide useful information for the forecasts which is not currently included. At present, most fault locations identified in active-source seismic have not become activated by injection and faults that are activated are not visible in active seismic data. This may be a consequence of the fact that most induced sequences to date have occurred on strike-slip faults which are difficult to identify in active-source seismic surveys. Of additional concern is that most other methods to identify faults rely on the occurrence of seismic activity, therefore, for this purpose it will be advantageous for operators to install a pre-injection monitoring network to capture microseismic events that might illuminate faults that may potentially host induced events.

4. Care should be taken to characterize the pre-injection event rates. It is useful to note that the background seismicity rates are primarily related to the background stressing rate. In the pre-injection period, the stressing rate is driven by tectonic loading which is generally constant over long-time periods. Once injection begins and stressing rates are modulated by fluid pressure, the character of the resulting seismic activity may differ considerably. Estimates of seismicity rates are plagued by catalog incompleteness issues arising from dynamic spatial and/or temporal seismic network coverage and the inability to detect individual events during periods of high seismic activity. Estimates of the seismicity rate are most robust when considering small magnitude events which improve the sample size. However, seismic catalogs are generally incomplete at the smallest magnitudes [van der Elst and Page, 2023]. Here again, installation of a seismic monitoring network prior to beginning injection may help address this issue of constraining the background rate. Such a network would need to ensure completeness down to a magnitude small enough that it captures events spanning 3-4 magnitude units and at least 1000 events to ensure robust statistical measurements [Marzocchi et al., 2020, Geffers et al., 2023]. Methods to estimate how long the seismic monitoring array must be installed for and the target completeness magnitude are in development.

5. Additional consideration should be given to the ambient absolute stress field. The amplitude of seismicity rate changes due to secondary (i.e. aftershock) triggering (see Equation 4) is related to the amplitude of the static Coulomb stress change, which, in turn is related to ambient stress in the crust. In other words, the seismic response to a single event, whether that event was triggered by fluid injection or by tectonic sources, is dependent on the stress state. If the first event was triggered by fluid injection, that event and entire resulting sequence should still be considered induced since that sequence would not have occurred at that time without the fluid injection.

6. Current pressure models provided to ORION from the IBDP site only compute the pressure in the reservoir, yet seismicity forecasts are being developed for events that occur within the basement rock, several kilometers below the reservoir. There are currently no methods that account for fluid migration from the reservoir to the basement (along faults/fractures or through the matrix), which may dramatically impact the pressurization and stressing rates passed to the forecasting methods.

7. There are a number of parameters that are required by the declustering algorithms (i.e. space window, time window, value of the nearest-neighbor distance that separates space-time clustered events from background events). These parameters are also likely to be site- and size-specific and should be calibrated accordingly. At present, we have tested the default values in a number of different tectonic regions for stability, however, more work should be done to automate the selection or test against different areas.

8. Finally, another onerous task is data curation and quality control. Seismicity catalogs uploaded to ORION should be provided in standardized units (e.g. locations should be reported in latitude and longitude and not state-plane coordinates, depths below ground elevation should be negative, time should be reported in UTC, SI units are preferred). In the case of the IBDP catalog used in this analysis, the event locations were reported in a local Illinois state-plane coordinate system and some depths were large and positive, resulting in what appeared to be "airquakes". In all cases, metadata should be provided that lists all coordinate systems, sign conventions, and units.

6 Applicability to Future Carbon Storage Sites

Ultimately, the ORION toolkit will be able to apply to all carbon storage sites. It should be useful across spatial scales from the site- to basin- scale, apply before, during, and after injection, and handle interacting operations. In order to achieve this goal, there are a number of items to consider. Again, this is not meant to be an exhaustive list and no preference or weight is to be inferred by the order these items are listed or the fact that some items are unintentionally not included below.

1. We have made the explicit design decision to not discriminate between models or methods, but to apply the forecasts to all methods. One question arises about the appropriateness of this methodology, especially when it comes to propagating pressure data (or any other data) through the forecasting workflow when the models completely disagree.

2. Often times, there is considerable disagreement between pressure models based on different simulation methods. It would be advantageous to collect pressure monitoring data that could be used to calibrate the model results. This would require a network of pressure monitoring well, perhaps acquired by states or through industrial partnerships to leverage existing abandoned and/or orphaned wells.

3. Forecasts made with ORION are only as good as the data provided to it. We currently have workflows established for users to provide seismicity catalogs, pressure data and/or models, well locations, and injection data. However, methods to gather these data from public sources are imperfect and require routine maintenance because they depend on architecture and structure of the websites and servers from which the data are collected. A well maintained and up-to-date central clearing house (e.g. EDX4CCS) where this data could be collected and queried within ORION would be ideal.

4. Advancements made within ORION to incorporate fault information would allow the exploration of the impact of poroelastic effects in addition to the direct pore-fluid effects currently being addressed. Resolving the poroelastic stresses typically requires fully coupled, high-fidelity numerical simulations of flow and mechanics that are run on super computers. Development of fast (possibly ML) methods to estimate the poroelastic stress changes on faults would be necessary to improve the speed of these calculations and extend their use to real-time forecasts of induced seismicity. Additionally, the fundamental assumptions made in ORION that govern earthquake nucleation need to be modified. Currently, ORION assumes that the volume is filled with a population of "nucleation point sources" and discrete fault surfaces will be required when considering poroelastic effects.

5. Extending to basin-scale applications will require modifications of the coordinate systems used in ORION, especially when sites span UTM zones.

6. Lastly, issues surrounding estimation of background seismicity rates, fault identification, and selection of grid size and time window, that were mentioned in Section 5 will continue to be a

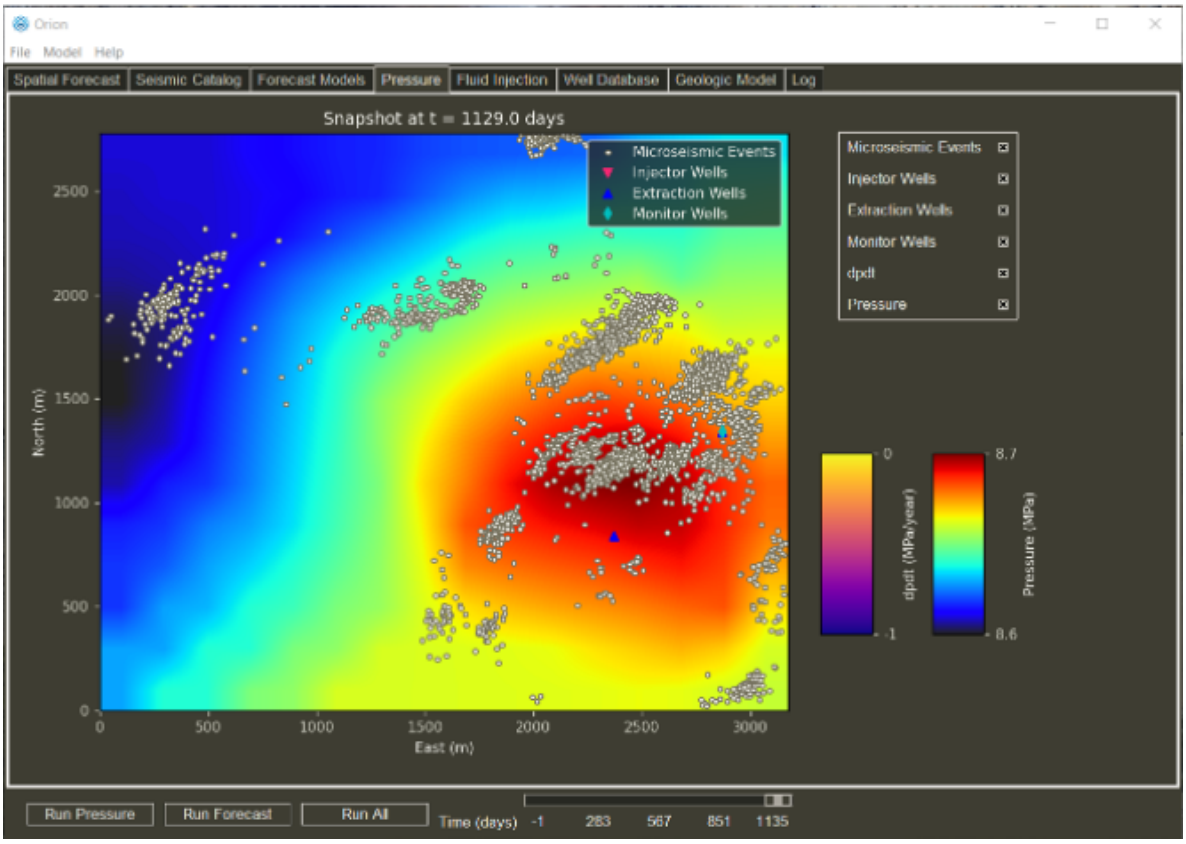

Figure 1: Pressure distribution computed with TOUGH-FLAC [Luu et al., 2022] related to injection at the IBDP site. Seismic event locations (computed by Norsar [Goertz-Allmann et al., 2017]]) are shown by the dots. Visualization of the pressure advances in time by dragging the slider bar on the bottom. Layers including the presence of the earthquake epicenters, well locations, forecasted event probability, and forecast event rate can be toggled on or off with the buttons on the right. This image is a screen capture from the legacy version of ORION which is built on the Python package "tkinter" and operates on a desktop machine.

concern when moving to larger spatial areas and regions with interacting operations. The two former issues could be somewhat addressed with high-precision and low Mc seismic catalogs for the pre-injection period and during on-going operations. Receiving this information in real-time would improve seismic forecasts during the operational period and support decision-making and agility to respond to adverse events. There are currently no public tools for real- time catalog generation in the FECM portfolio.

7 Current Advancements and Future Developments

Figure 10 provides a list of future improvements to be made to ORION. The sponsor for each attribute is listed in Columns 2 through 4. The NRAP and SMART programs provide avenues to support the development of these topics, however, in two cases, outside funding for R&D efforts is being sought. Lastly, the target completion date for each attribute is listed.

8 Acknowledgements

Portions of this work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344.

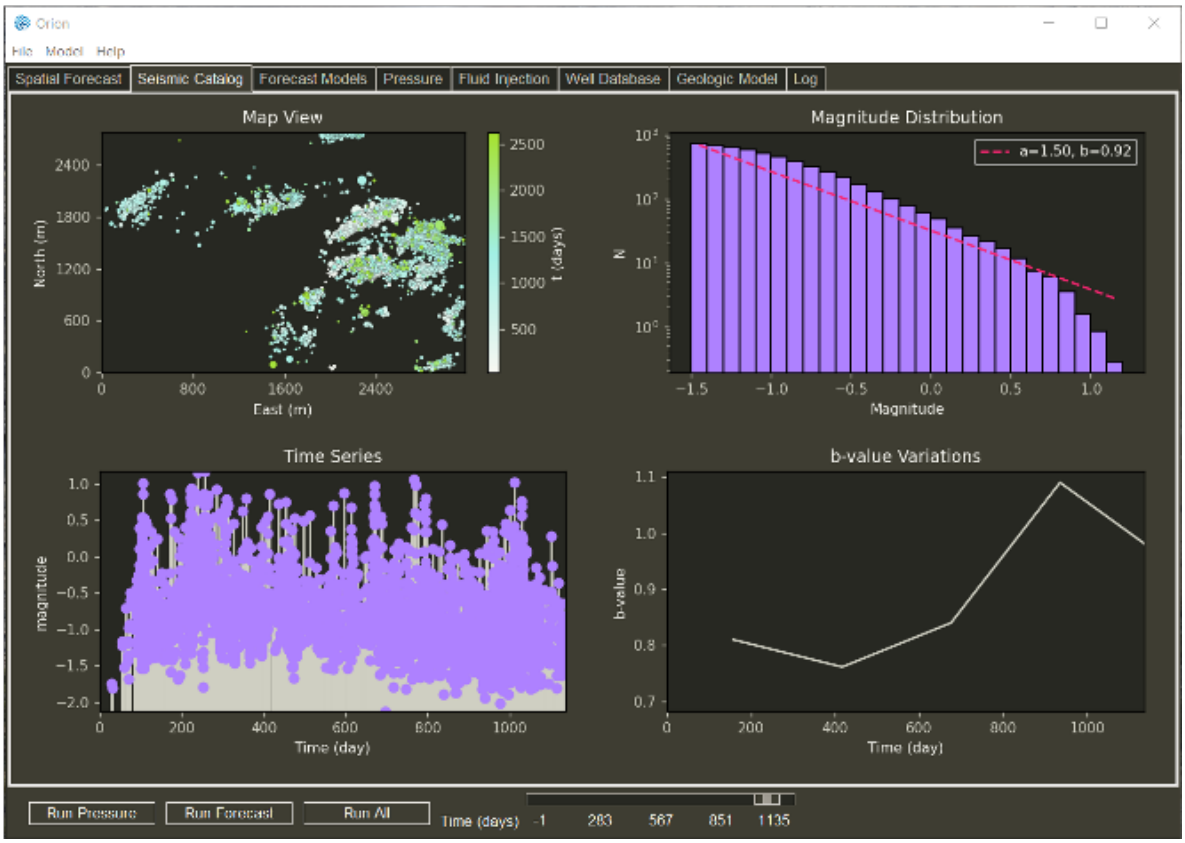

Figure 2: Attributes of the seismicity catalog. Top left: spatial distribution of event epicenters (latitude/northing, longitude/easting), colored by origin time and scaled by event magnitude. Top right: Frequency- magnitude distribution of the events in the IBDP seismic catalog. Red line indicates the best-fitting power law that relates frequency and magnitude, referred to at the "b-value". Bottom left: Event magnitudes as a function of time. Bottom right: Time-varying b-value derived from the b-value computation during moving windows. Image from the legacy-ORION version.

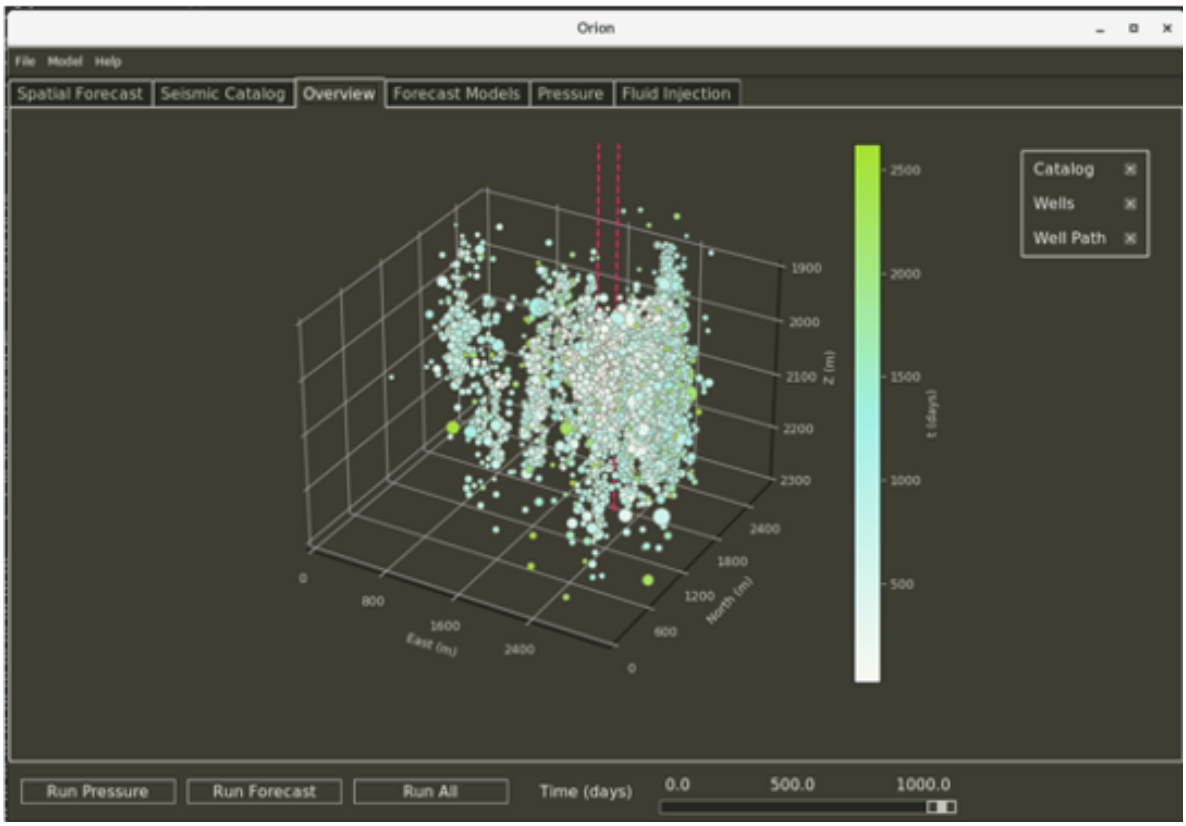

Figure 3: 3D view of the spatial distribution of event hypocenters (latitude/northing, longitude/easting, depth, colored by origin time and scaled by event magnitude. Image from the legacy-ORION version.

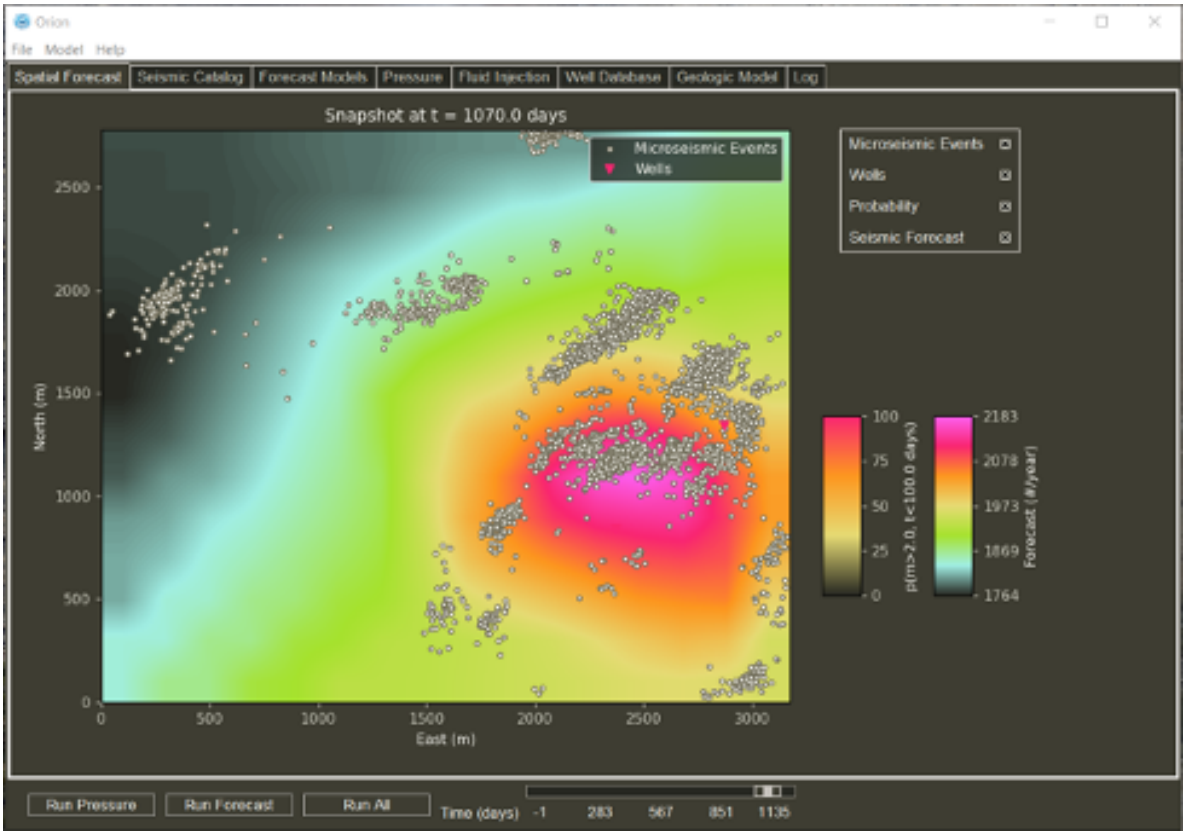

Figure 4: Spatial distribution of the forecasted yearly event rate per square kilometer. Forecast is computed for a set of grid cells that correspond to the grid on which the pressure was calculated in Figure 1. Dots are the earthquake epicenters. Visualization of the spatial forecast advances in time by dragging the slider bar on the bottom. This image is from the legacy-ORION version.

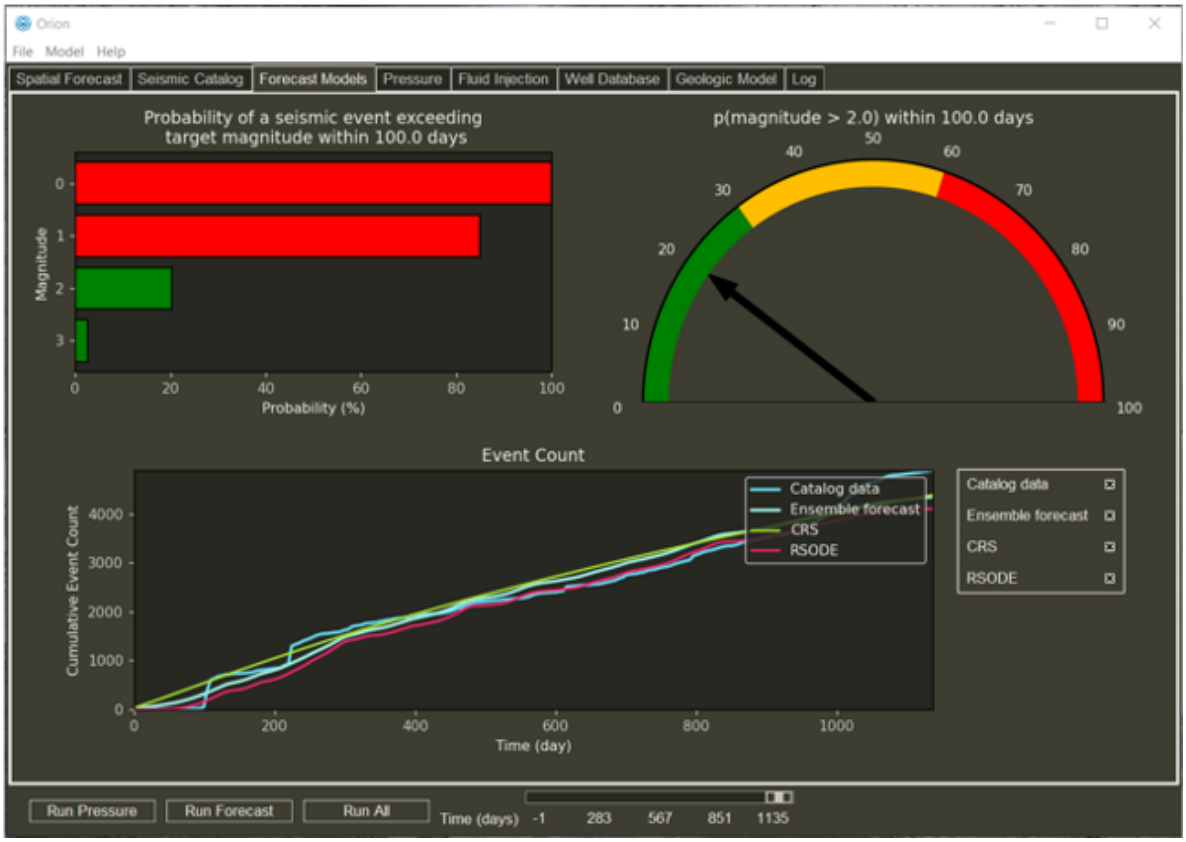

Figure 5: Attributes of the advanced traffic light system that captures the probability of exceeding a target magnitude earthquake in a given time period. Top left: Probability of exceedance for binned magnitudes listed on the y-axis over the next 100 days. The color is defined based on user-defined values of actionable probabilities. The red color indicates a remedial action is mandatory when the probability (P) exceeds 60%, in this example. Amber colors indicate a warning, where action is recommended, but not mandatory and is defined when $30\% \geq P \geq 60\%$. A green color indicates that no remedial action is suggested and is defined for $P \leq 30\%$. Top Right: Dial plot for a specified example that is defined for a target M=2 within the upcoming 100 days. Bottom: The cumulative number of events that were observed during the IBDP project (cyan) along with seismicity forecast generated via the CRS model (green), the RSODE model (red), the ensemble forecast (light blue). This image is from the legacy-ORION version.

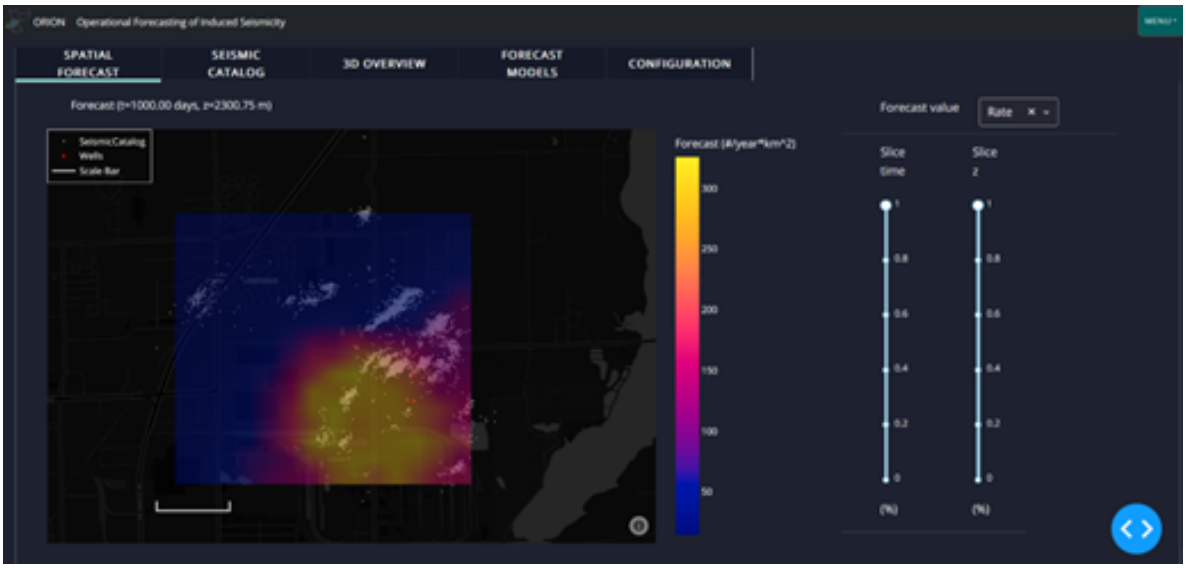

Figure 6: An example of the user interface created by ORION-STRIVE, which shows the spatial seismic forecast.

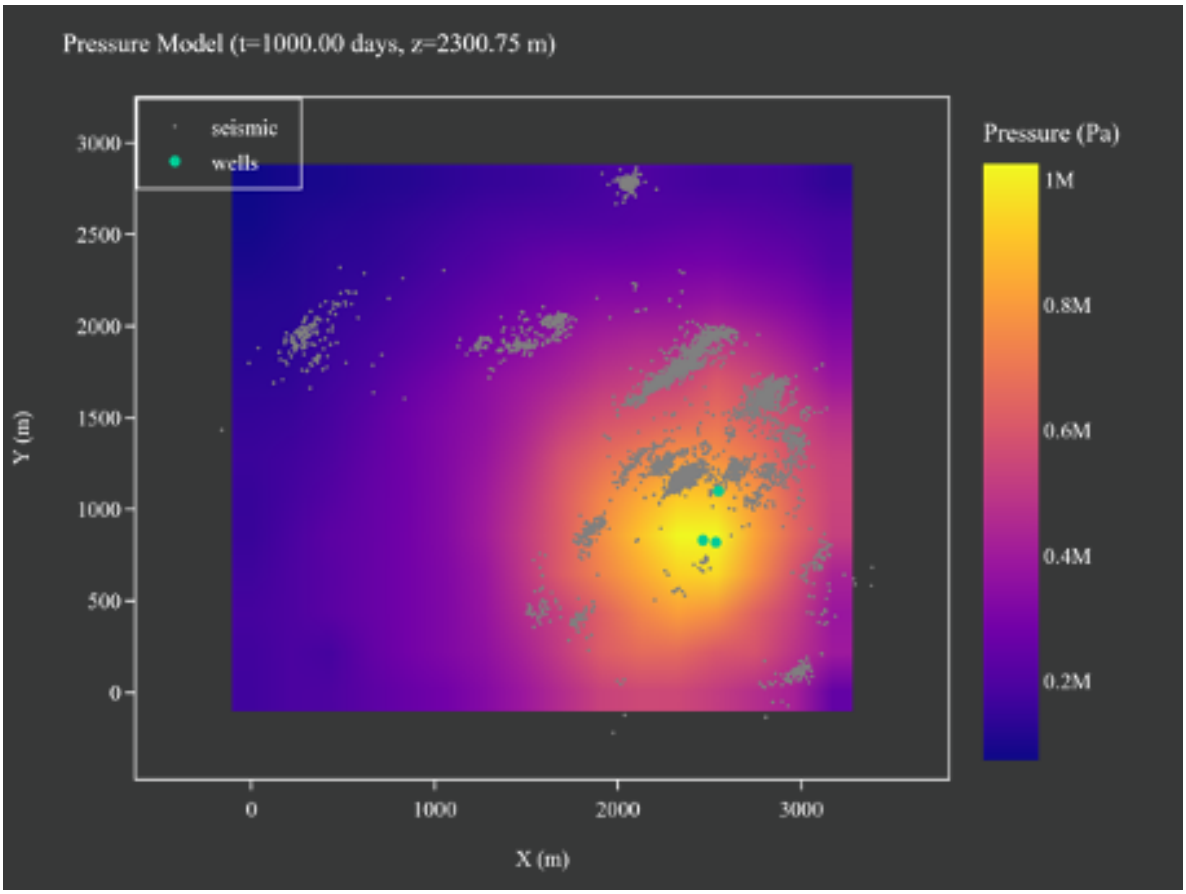

Figure 7: Pressure distribution computed with Fast Marching Method provided by the Datta-Gupta group at TAMU related to injection at the IBDP site. Seismic event locations (computed by Norsar) are shown by the dots. Layers including the presence of the earthquake epicenters and well locations, can be toggled on or off with the buttons in the inset legend. This image is a screen capture from the (currently unreleased) STRIVE version of ORION which is built on the Python package "plotly dash" and operates on a web browser.

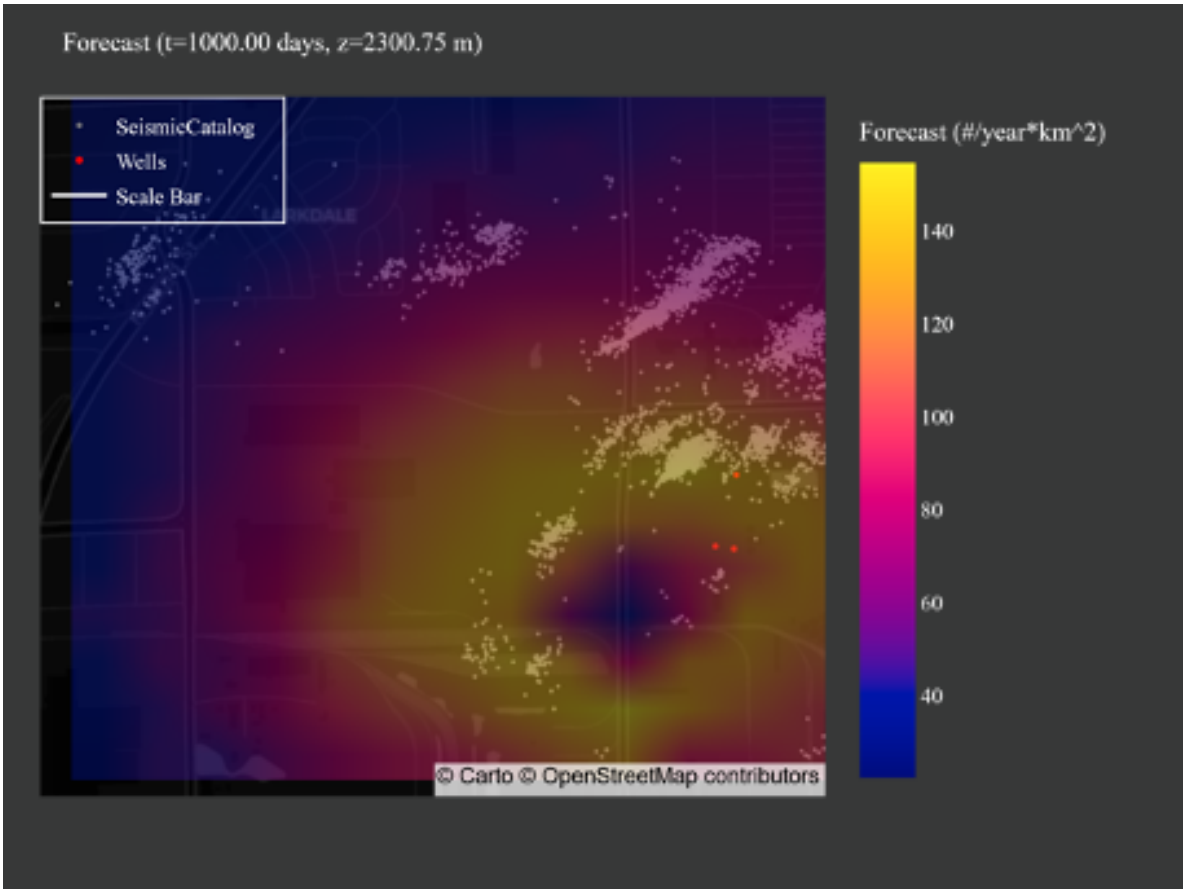

Figure 8: Spatial distribution of the forecasted yearly event rate per square kilometer. Forecast is computed for a set of grid cells that correspond to the grid on which the pressure was calculated in Figure 7. Dots are the earthquake epicenters. This image is from the STRIVE version of ORION.

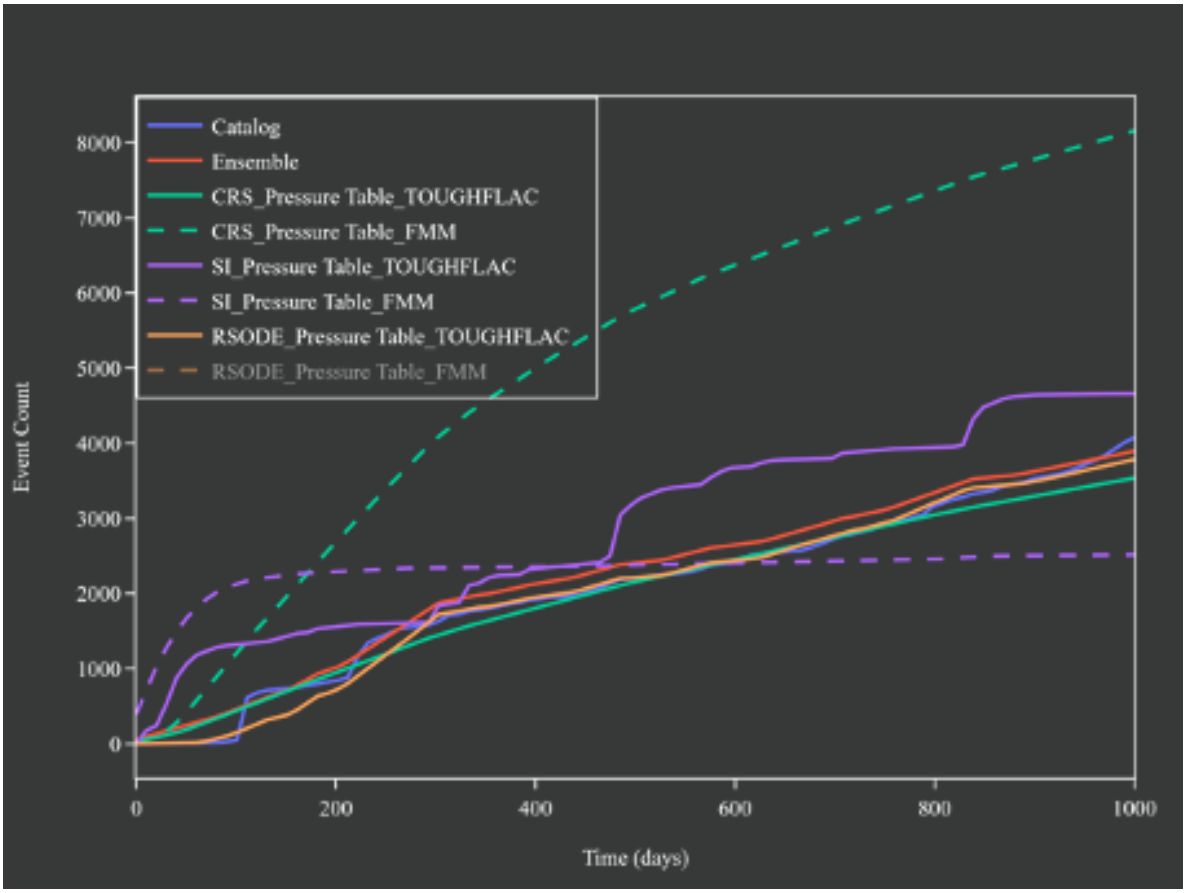

Figure 9: The cumulative number of events that were observed during the IBDP project (dark blue) along with seismicity forecasts generated via the CRS model (green), the RSODE model (orange), seismogenic index model (purple) the ensemble forecast (red). Each forecast method is applied to pressure distributions from the TOUGH-FLAC model shown in Figure 1 (solid lines) and the FMM model in Figure 7 (dashed lines). Note that forecasts applied to the FMM pressure model have not been calibrated and are built using default or automatically calculated parameters. This image is from the STRIVE version of ORION.

Future Developments in ORION	NRAP	SMART	Other funding sources	Target Development Date
Operational management strategies (OMS)	X			End EY24
Advanced traffic light system	X			End EY25
Demonstrate initial basin-scale applications	X			End EY24
Ground motion predicitons	X			End EY26
Forecast Evaluation	X			End EY24
Improved declustering methods	X			End EY24
Implement "Simple-ETAS" forecasting model	X			End EY24
Develop method to invert seismicity rates for pressurization rates	X			End EY26
Visualization of OMS		X		End EY24
Integration of SMART tools (USM) for rapid forecast evaluation / OMS		X		End EY24
ML-based decision tree for ensemble forecasting		X		End EY24
Evaluate/Build/Integrate pure-ML-based forecasting model		X		End EY26
Integration with SVDSP		X		End EY24
ML-based Mmax Forecast		X		End EY25
ML-based GR *a* and *b* value prediction with time		X		End EY25
Real-time foreacsts		X	X*	End EY26
Data-driven spatial forecast heterogeneity		X		End EY26
Incorporate fault/fracture prior information from SMART tools		X		End EY25
Incorporate focal mechanisms (into visualization tools, new data streams from smart tools?)		X		End EY26
Incorporate Poroelastic Stresses		X	X**	End EY26

*based on real-time catalog generation

**Requires rapid 1) fault identification and 2) focal mechanism detemination tools and incoporating FM results into ORION to project poroelastic stresses onto

Figure 10: The ORION toolkit for induced seismicity forecasting has been supported by both SMART and the National Risk Assessment Partnership (NRAP). NRAP supports the development of the forecasting "engine" or the back-end of the code-base, including all operational management strategies and traffic light systems. SMART supports the visualization components or the front-end of the code-base in addition to all the machine learning components. This table lists the future develop directions for ORION along with the supporting program and target development data. Funding for some of the functionality and/or input data products is being sought through funding sources external to SMART or NRAP.

References

[Aki, 1965] Aki, K. (1965). Maximum likelihood estimate of b in the formula logn=a-bm and its confidence limits. *Bulletin of the Earthquake Research Institute*, 43:237–239.

[Cao and Gao, 2002] Cao, A. and Gao, S. S. (2002). Temporal variation of seismic b-values beneath northeastern japan island arc. *Geophysical research letters*, 29(9):48–1.

[Dando et al., 2021] Dando, B., Goertz-Allmann, B., Kühn, D., Langet, N., Dichiarante, A., and Oye, V. (2021). Relocating microseismicity from downhole monitoring of the decatur ccs site using a modified double-difference algorithm. *Geophysical Journal International*, 227(2):1094–1122.

[Dieterich, 1994] Dieterich, J. (1994). A constitutive law for rate of earthquake production and its application to earthquake clustering. *Journal of Geophysical Research: Solid Earth*, 99(B2):2601–2618.

[Fehlberg, 1970] Fehlberg, E. (1970). Classical fourth-and lower order runge-kutta formulas with step-size control and their application to heat transfer problems. *Computing*, 6:61–71.

[Ferris et al., 1962] Ferris, J. G., Knowles, D. B., Brown, R. H., and Stallman, R. W. (1962). *Theory of aquifer tests*. US Government Printing Office Washington.

[Gardner and Knopoff, 1974] Gardner, J. and Knopoff, L. (1974). Is the sequence of earthquakes in southern california, with aftershocks removed, poissonian? *Bulletin of the seismological society of America*, 64(5):1363–1367.

[Geffers et al., 2023] Geffers, G.-M., Main, I. G., and Naylor, M. (2023). Accuracy and precision of frequency-size distribution scaling parameters as a function of dynamic range of observations: example of the gutenberg-richter law b-value for earthquakes. *Geophysical Journal International*, 232:2080–2086.

[Goertz-Allmann et al., 2017] Goertz-Allmann, B., Gibbons, S., Oye, V., Bauer, R., and Will, R. (2017). Characterization of induced seismicity patterns derived from internal structure in event clusters. *Journal of Geophysical Research: Solid Earth*, 122(5):3875–3894.

[Johnson et al., 2013] Johnson, S., Settgast, R., Fu, P., Antoun, T., and Ryerson, F. (2013). Geos code development road map-may, 2013. Technical report, Lawrence Livermore National Lab.(LLNL), Livermore, CA (United States).

[Kroll and Brudzinski, 2023] Kroll, K. A. and Brudzinski, M. R. (2023). Evaluating the Aftershock Duration of Induced Earthquakes. *Bulletin of the Seismological Society of America*.

[Kroll and Sherman, 2023] Kroll, K. A. and Sherman, C. S. (2023). Orion: Operational forecasting of induced seismicity. https://edx.netl.doe.gov/dataset/orion-operational-forecasting-of-induced-seismicity.

[Langenbruch et al., 2018] Langenbruch, C., Weingarten, M., and Zoback, M. D. (2018). Physics-based forecasting of man-made earthquake hazards in oklahoma and kansas. *Nature communications*, 9(1):3946.

[Luu et al., 2022] Luu, K., Schoenball, M., Oldenburg, C. M., and Rutqvist, J. (2022). Coupled hydromechanical modeling of induced seismicity from co2 injection in the illinois basin. *Journal of Geophysical Research: Solid Earth*, 127(5):e2021JB023496.

[Marzocchi et al., 2020] Marzocchi, W., Spassiani, I., Stallone, A., and Taroni, M. (2020). How to be fooled searching for significant variation of the b-value. *Geophysical Journal International*, 220:1845–1856.

[Reasenberg and Jones, 1989] Reasenberg, P. A. and Jones, L. M. (1989). Earthquake hazard after a mainshock in california. *Science*, 243(4895):1173–1176.

[Rutqvist and Moridis, 2008] Rutqvist, J. and Moridis, G. (2008). Development of a numerical simulator for analyzing the geomechanical performance of hydrate-bearing sediments. In *ARMA US Rock Mechanics/Geomechanics Symposium*, pages ARMA–08. ARMA.

[Segall and Lu, 2015] Segall, P. and Lu, S. (2015). Injection-induced seismicity: Poroelastic and earthquake nucleation effects. *Journal of Geophysical Research: Solid Earth*, 120(7):5082–5103.

[van der Elst, 2021] van der Elst, N. J. (2021). B-positive: A robust estimator of aftershock magnitude distribution in transiently incomplete catalogs. *Journal of Geophysical Research: Solid Earth*, 126(2):e2020JB021027.

[van der Elst and Page, 2023] van der Elst, N. J. and Page, M. T. (2023). a-positive: A robust estimator of the earthquake rate in incomplete or saturated catalogs. *Journal of Geophysical Research: Solid Earth*, 128(10):e2023JB027089.

[Wiemer and Wyss, 2000] Wiemer, S. and Wyss, M. (2000). Minimum magnitude of completeness in earthquake catalogs: Examples from alaska, the western united states, and japan. *Bulletin of the Seismological Society of America*, 90(4):859–869.

[Yang et al., 2017] Yang, C., Vyas, A., Datta-Gupta, A., Ley, S. B., and Biswas, P. (2017). Rapid multistage hydraulic fracture design and optimization in unconventional reservoirs using a novel fast marching method. *Journal of Petroleum Science and Engineering*, 156:91–101.

[Zaliapin and Ben-Zion, 2016] Zaliapin, I. and Ben-Zion, Y. (2016). A global classification and characterization of earthquake clusters. *Geophysical Journal International*, 207(1):608–634.