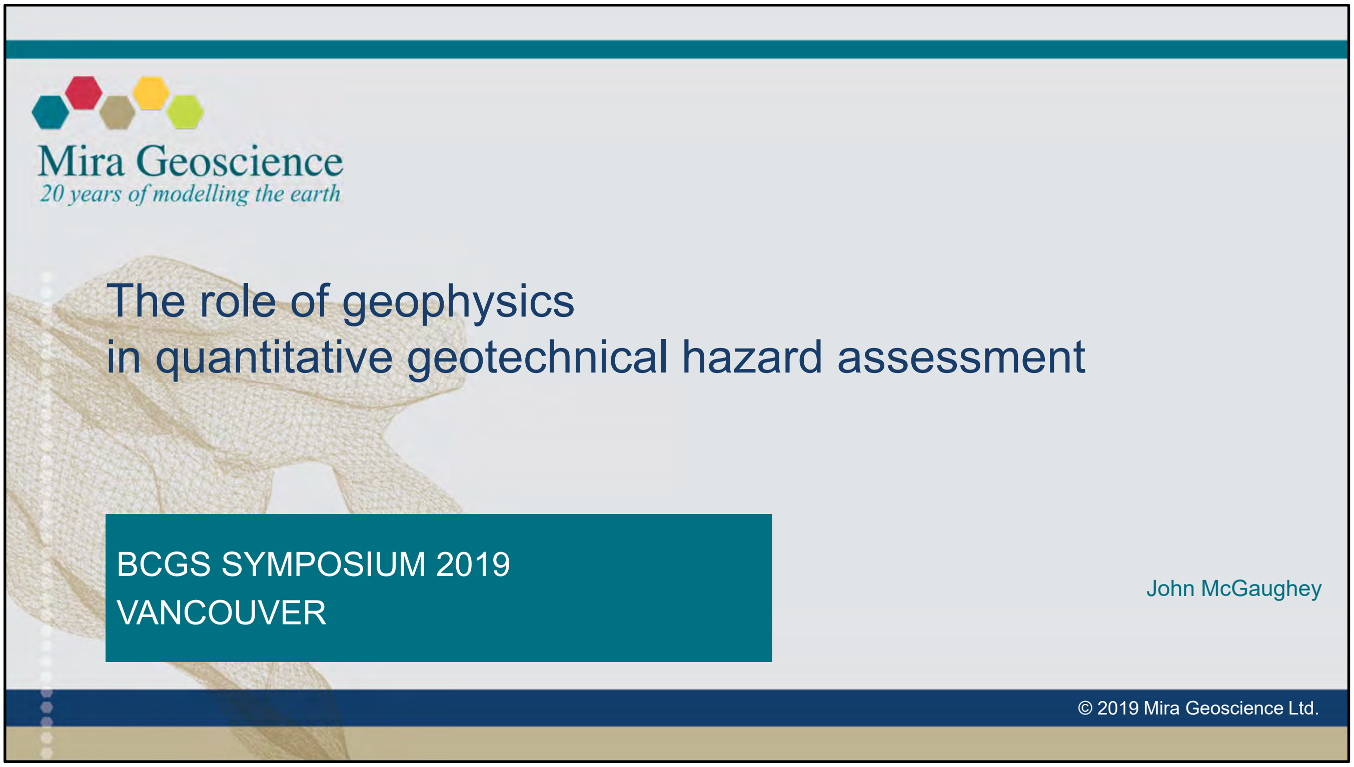
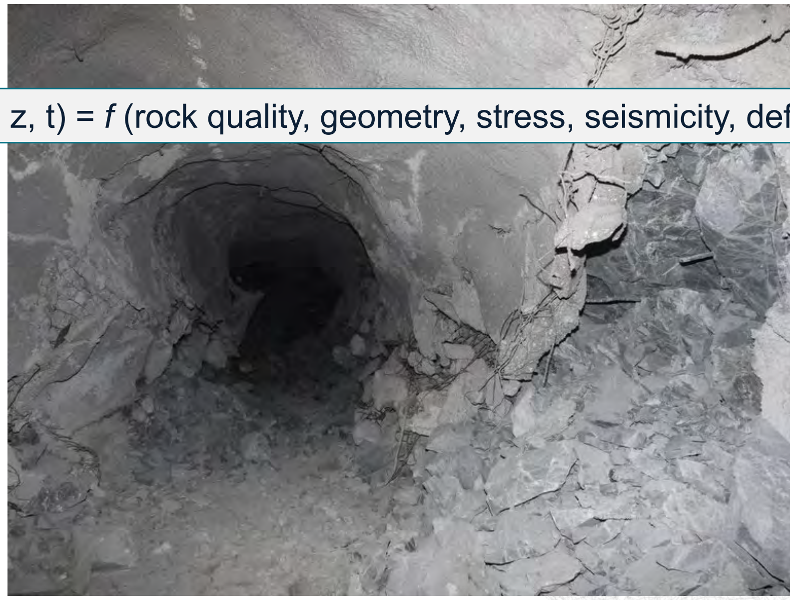


Mira Geoscience offers software and consulting services for the mining industry. Our focus is on 3D and 4D data management, earth modelling and analysis solutions for mineral exploration and geotechnical hazard assessment.

Geohazard

hazard (x, y, z, t) = *f* (rock quality, geometry, stress, seismicity, deformation, …)



Mira Geoscience
20 years of modelling the earth

Our premise is a simple one, and it's proven effective over many years of consulting projects in a wide variety of mining types (open pit, underground, hard rock, soft rock) and hazard types (e.g. rockburst, strain burst, roof fall, slope failure, flood).

We model hazard as a 4D function of time and space on the rock interface where the hazards are experienced. In a case like that shown in the photograph, we would model hazards along the drift, with the probabilistic hazard assessment computed as a function of numerous, quantitative hazard criteria in various classes – geological, rock mass quality, stress, seismicity, and others. We typically work with a couple of dozen candidate hazard criteria, which are modelled throughout the mine. Some evolve with time; some do not.

We use "predictive analytics" methods to examine the history of the state of the mine over time, compare it to the history of geohazards such as rockbursts, and to rigorously explore the relationships in order to define the data-driven hazard equation.

Geophysical contributions

Site characterization

- 2D/3D geophysical surveys
- wireline logging

Monitoring

- LiDAR / laser scanning
- laser prisms
- scanning radar
- InSAR
- time-lapse geophysical surveys
- passive seismic monitoring

Mira Geoscience
20 years of modelling the earth

Geophysics offers several data types that contribute to hazard assessment. They come in two major categories: geophysics for static (non-time dependent) site characterization and geophysics for time-dependent monitoring. Geophysics for site characterization is similar to the array of techniques used in mineral exploration. Geophysics for site monitoring, however, is different. Monitoring usually means deformation monitoring, time-lapse seismic or other geophysical surveys to image changes in structure or physical properties, or passive seismic monitoring with microseismic arrays.

The approach

- capture the 4D historical state of the mine
- capture the historical record of geohazards
- analyze and understand the 4D relationship between data and geohazard
- use the analysis to forecast the likelihood of geohazard events for any future mine state
- establish a system to automatically update the state of the mine and the geohazard forecast

Mira Geoscience
20 years of modelling the earth

The approach we have refined over the years is to understand the history of mining geohazard occurrence in terms of the state of the mine as expressed through numerous observed or modelled variables. We analyze the correlations between hazard event occurrence and the state of numerous variables using statistics or machine learning. This enables us to forecast hazard by recognizing patterns in data that have been previously associated with hazards. We encapsulate those relationships in a set of “rules” that can automatically be applied to future states as the mine evolves.

The workflow

- problem definition
- feature engineering
- data fusion
- analysis
- deployment

Mira Geoscience
20 years of modelling the earth

We have defined a workflow that enables to follow the same approach for any mine type and any hazard type.

The workflow

- problem definition →
- feature engineering
- data fusion
- analysis
- deployment

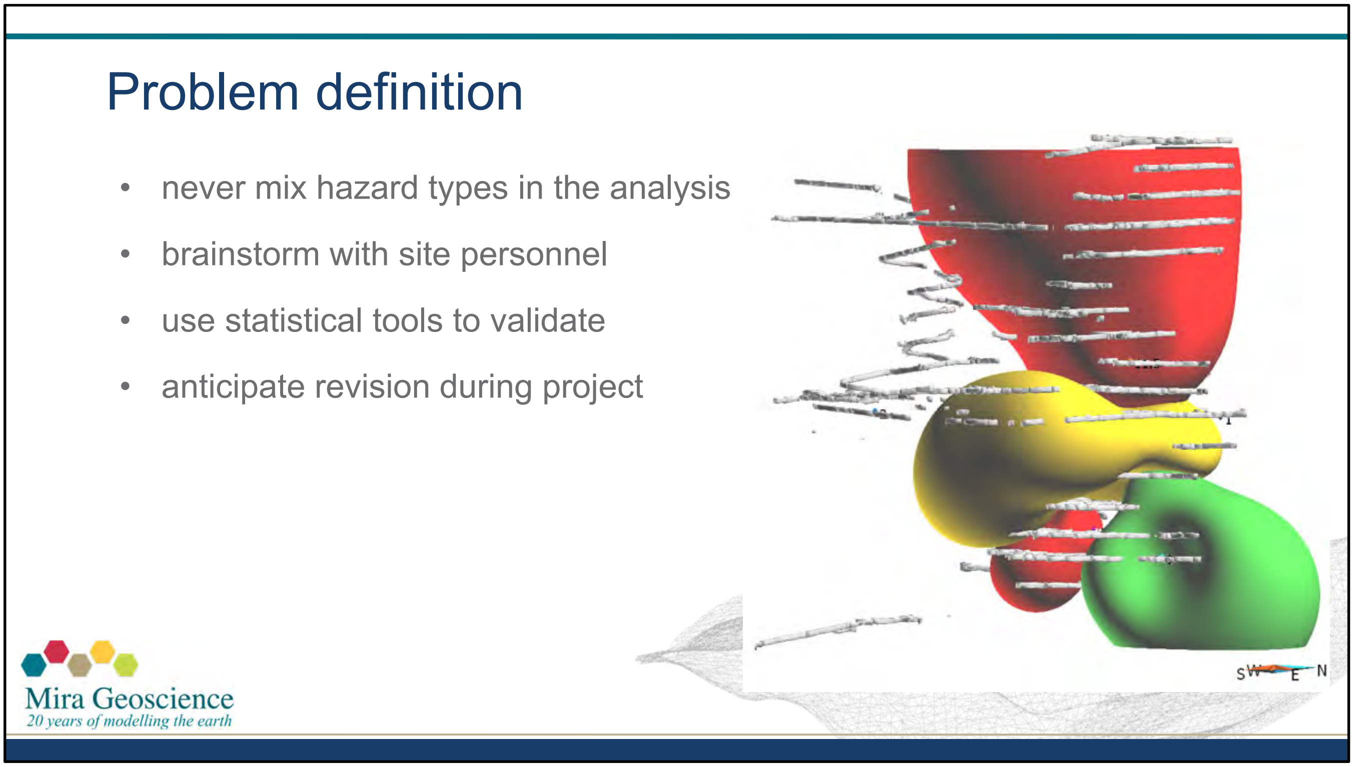
- hazard type (rockburst, slope failure, etc.)
- general hazard criteria and data sources

↓

structure, rock type, rock quality, stiffness, stress, blasting, mine geometry, production rate, sequencing, seismicity, deformation, slope angle, face angle, berm width, water

Mira Geoscience
20 years of modelling the earth

The first step is to understand the specific hazard(s) that are to be modelled, the sources of data available, and the general classes of variables that are believed to have potential correlation to the hazard. I have highlighted geophysical contributions to input data types in red.



Because we are correlating specific hazards with data, it is critical that each hazard type be treated as a separate problem. Beliefs by site personnel are generally refined over the course of a project as relationships between hazard occurrence and data are revealed through analysis.

The workflow

- problem definition
- feature engineering →
- data fusion
- analysis
- deployment

- mapping data sources to hazard criteria

↓

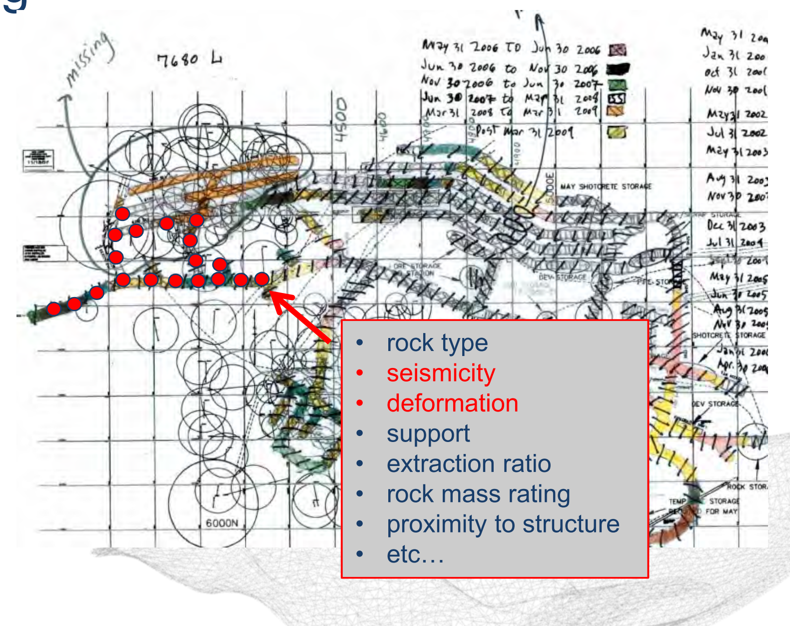
geological and geotechnical block models, geological wireframes, drillhole databases, mine development wireframes, seismicity, ground deformation data, blasting records, production records, structural data

Mira Geoscience
20 years of modelling the earth

"Feature engineering" is a term borrowed from machine learning. It refers to the assignment or creation of the individual variables derived from available data sets for correlation to the target variable. In our case the target variable is hazard occurrence.

Feature engineering

- hazard criteria are independently modelled
- digitization interval is a compromise between resolution and problem size
- each criteria class may yield many features



Mira Geoscience
20 years of modelling the earth

An important step is bring all the required hazard criteria onto the same support, with the values co-located in space and time. The co-location process is a complex one for many of the important criteria (although trivially simple for some others), and has been a significant part of our previous R&D, which now routinely and efficiently employ in a great variety of mining situations. It corresponds to what the analytics community call “feature extraction” or “data fusion”.

In practice, it means estimating a value for each of the criteria along the mine workings where the hazards are being assessed, typically spaced every few metres. When deployed in automated mode with Geoscience INTEGRATOR, the hazard criteria values are automatically updated as new data becomes available.

We digitize the mine model at locations where we want to model hazard. Each of the variables, or “features” to use the machine learning jargon, to be tested for correlation to the target variable is assigned to the appropriate point. The red dots in this image represent digitized points on a mine model with mine model points established along the drift centrelines at a spacing of a few metres. Some variable are time-dependent. This is handled by capturing the state of the mine model as a series of dates on which a hazard event was experienced, effectively capturing the 4D nature of the problem. A couple of variable types arising from geophysical data are highlighted in red.

Feature engineering

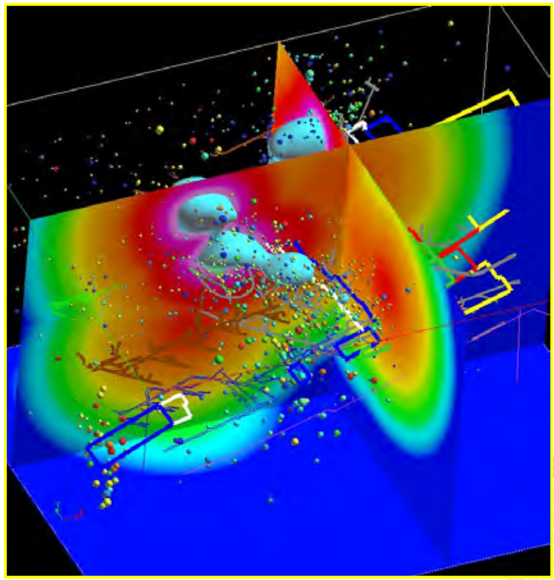
At the end of the day, some machine learning projects succeed and some fail. What makes the difference? Easily the most important factor is the features used. If you have many independent features that each correlate well with the class, learning is easy. On the other hand, if the class is a very complex function of the features, you may not be able to learn it.

Domingos (2015), A few useful things to know about machine learning, https://homes.cs.washington.edu/~pedrod/papers/cacm12.pdf

Mira Geoscience
20 years of modelling the earth

The field of predictive analytics, or machine learning, has seen a tremendous amount of progress in the past several years, where it is now ubiquitous in marketing, finance, and some scientific fields such as genetics. It is underutilized in mining but, as you can see from the definition, it provides a general description that exactly fits the geohazard assessment problem.

Feature engineering



microseismic event density

Mira Geoscience
20 years of modelling the earth

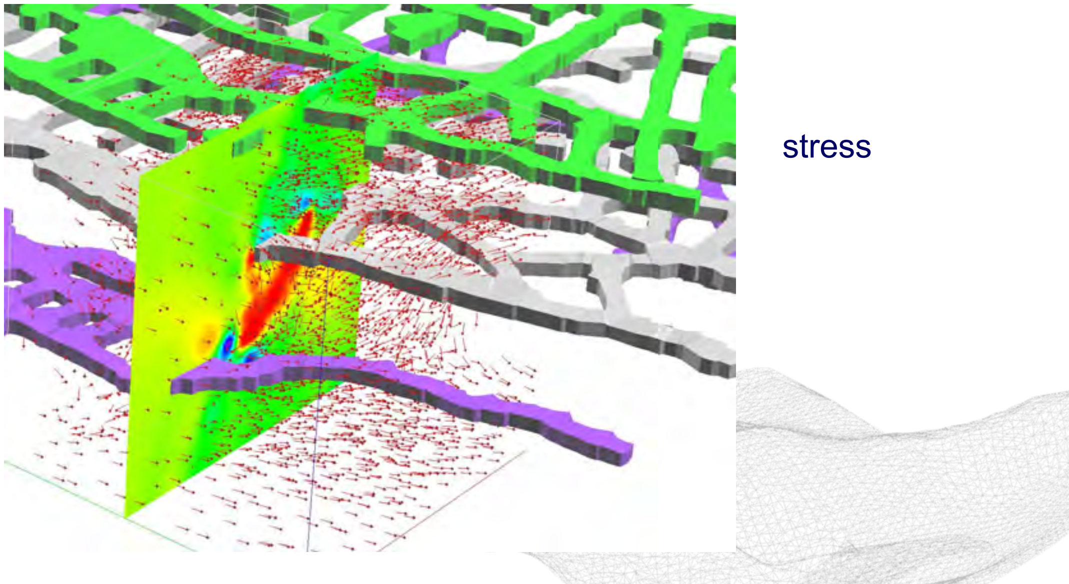
Some examples of the modelling that must be done to transform raw data to variables represented on the digitized mine model. This modelling process, the feature engineering, is the most time-consuming part of any project.

Here is an example of hazard criteria estimation, in this case from a South African gold mine. The hazard criteria is proximity to the edge of a volume containing an unusually high microseismic event density over a given time window.

You can see the distance to the edge of the blue volume has been computed everywhere in 3D space, from which it is projected onto the mine workings.

Such hazard criteria can be defined on a per-site basis with whatever level of complexity is required, and automatically computed and updated as new data arrive.

Feature engineering



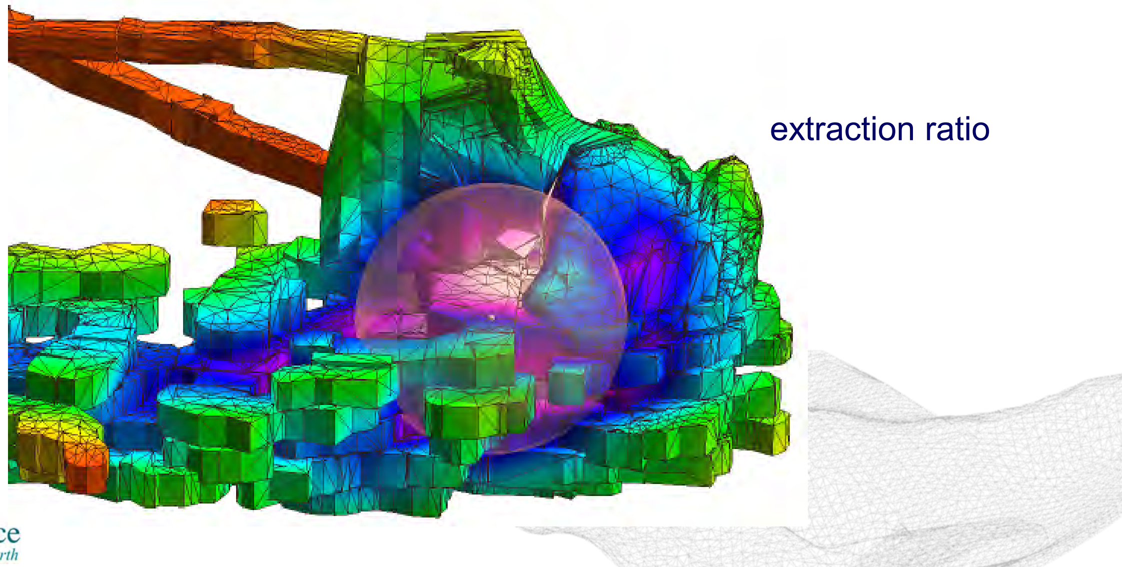
stress

Mira Geoscience
20 years of modelling the earth

Another example is deviatoric stress, which is the property shown on the 2D section. It is computed everywhere in 3D and projected onto the mine workings.

(The vectors are the principal stress direction, extracted from a 3DEC model.)

Feature engineering



Mira Geoscience
20 years of modelling the earth

Another example is excavation ratio, a variable that is typically an important hazard criteria, in part because it is often a local proxy for stress.

We compute excavation ratio as the volume of excavated versus intact material within a sphere of a given radius. In this picture we are showing excavation ration estimated on the each vertex of a wireframe. The computation is done automatically.

Feature engineering

Hazard criteria category	Example candidate hazard criteria
Mine development	Age of development, development rate, ground support category, age of support, span, orientation, proximity to intersections, depth.
Rock mass	Joint orientation, joint spacing, uniaxial compressive strength, fracture frequency, rock mass rating, rock quality designation.
Geology and structure	Rock type, proximity to contacts, proximity to waste gaps between ore zones, proximity to major structures, proximity to structural intersections, orientation of major structures, fault category, proximity to dykes.
Stress	Maximum principal stress, deviatoric stress, excavation ratio (as a local stress proxy), fault slip tendency.
Seismicity	Seismic event density, proximity to seismic cluster, local magnitude, E_s/E_p ratio, static stress drop, seismic moment, energy index.
Monitoring	Deformation from extensometers, convergence stations.

Mira Geoscience
20 years of modelling the earth

Some typical variables used in study of rockburst hazard in underground mines.

The workflow

- problem definition
- feature engineering
- data fusion →
- analysis
- deployment

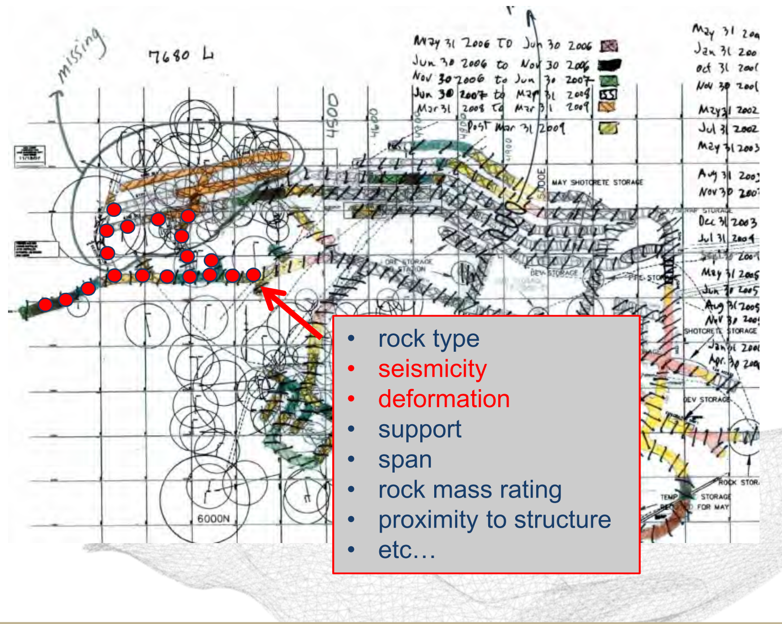
- establish list of historical hazards
- create mine model snapshots
- create data fusion table combining hazard criteria and hazard occurrences

Mira Geoscience
20 years of modelling the earth

The time-stepped mine model must be converted to a form amenable for statistical or machine learning analysis.

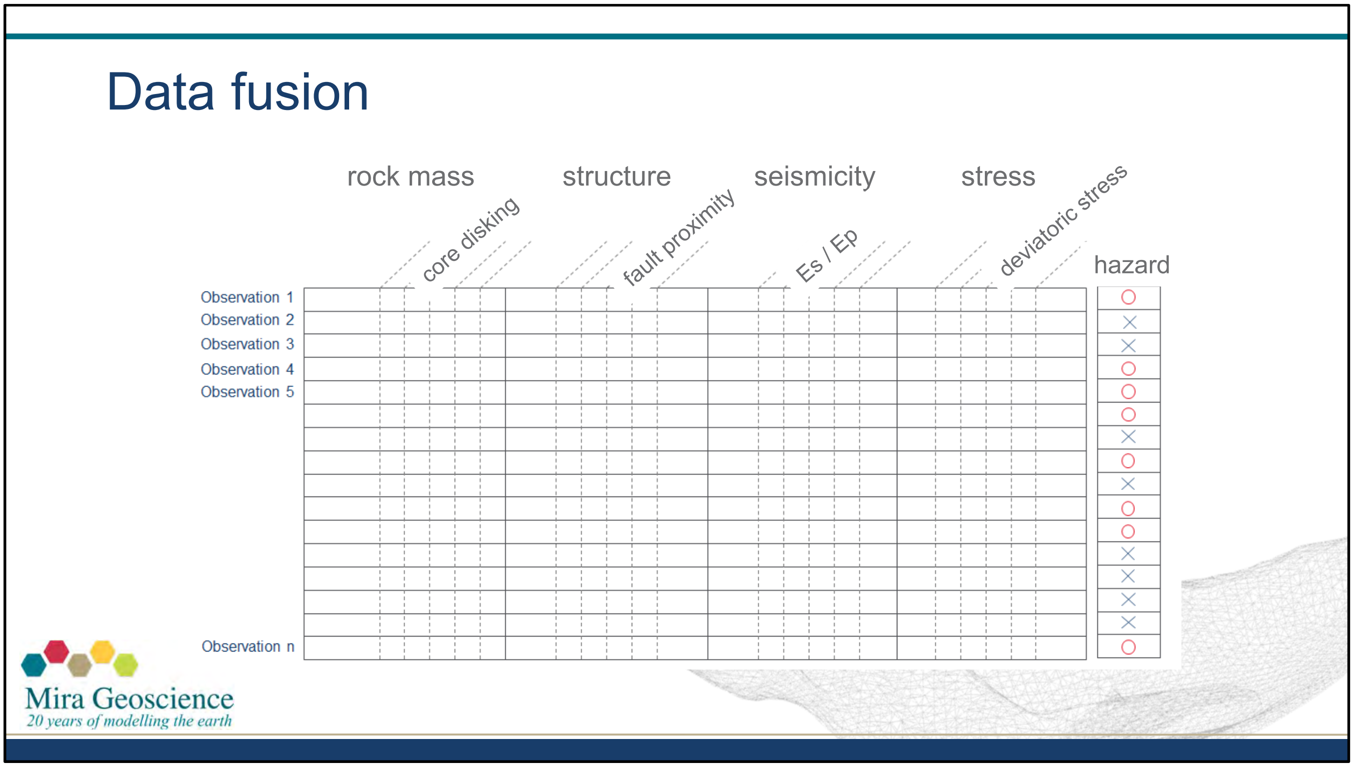
Data fusion

- hazard criteria are independently modelled
- digitization interval is a compromise between resolution and problem size
- each criteria class may yield many features



Mira Geoscience
20 years of modelling the earth

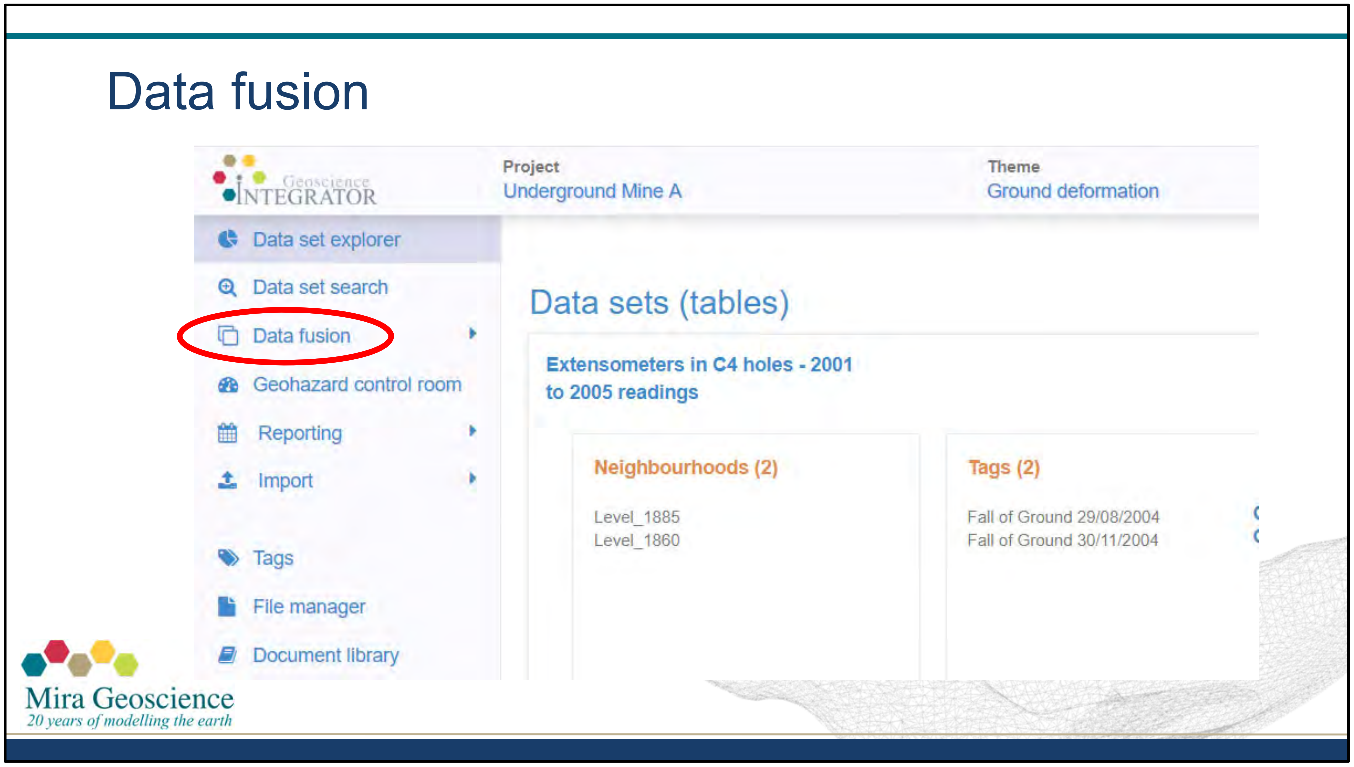
The analytics is carried out on the mine model points. Each point—the red dots in the image—correspond to a specific (x, y, z, date-time) in the mine, with a list of hazard criteria values corresponding to that point in space and time. If a hazard, such as a rockburst, occurred at or near that point in space or time, it is flagged in the mine model data structure.



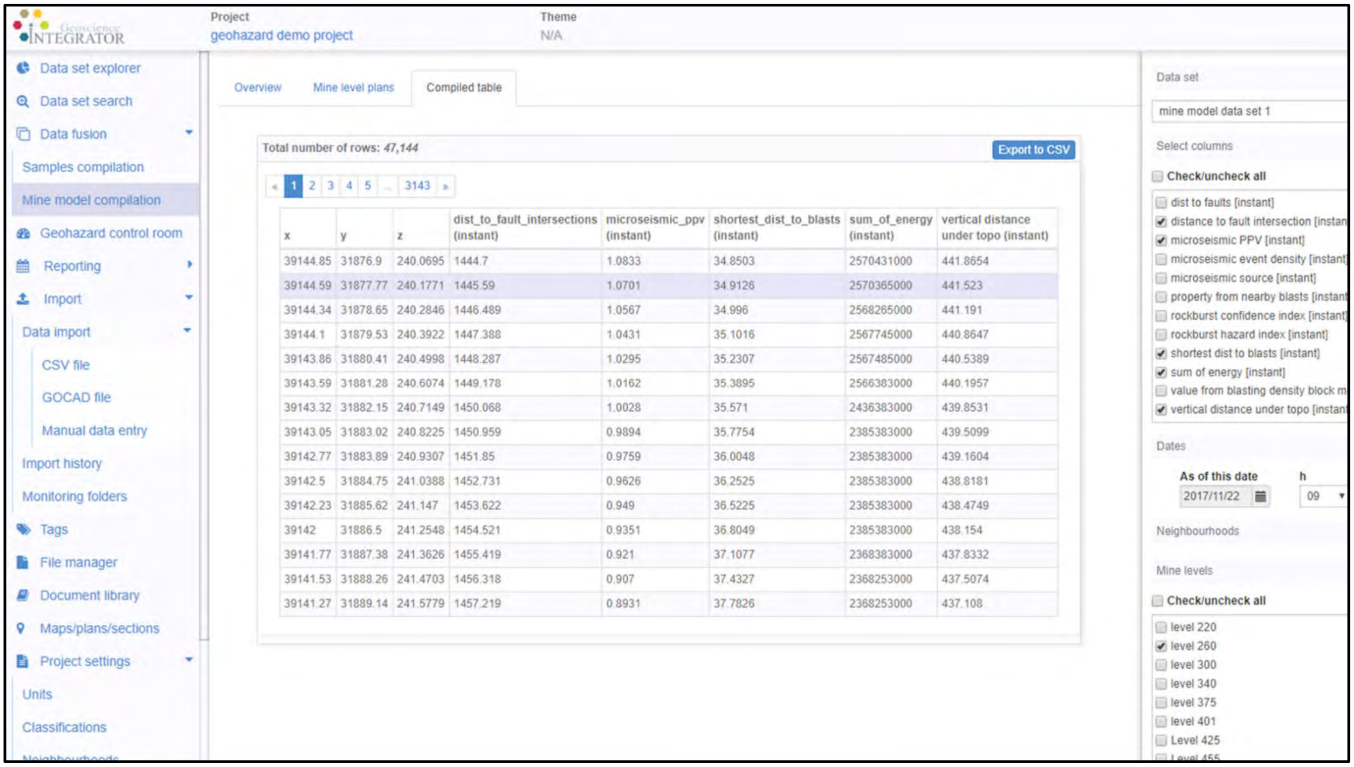
The first task in applying data-driven analysis is to construct a table of observations, in this case each observation corresponding to a position in the mine at a certain time. The rows of the table correspond to individual (x, y, z, t) observations—the red points shown in the previous image—while the columns correspond to the hazard criteria estimated at those positions and times. In the mining geohazard case, this table usually contains a few dozen columns and several million rows, with new rows being added at every time step.

We add an additional column called "hazard", which is a label applied to that row depending on whether or not the hazard being assessed occurred at that time and place. In the image, we use an "X" to indicate the hazard (e.g. a rockburst) occurred, and an "O" to indicate it did not occur. So the last column is a binary variable.

The goal of any predictive analytics or machine learning algorithm is to understand realtionships amongst the column variables and the hazard variable, with the objective of being able to forecast the probability of any new possible row being associated with the hazard or not.



Our Geoscience INTEGRATOR software system automatically carries out data fusion operations as new data are acquired.



This screen shot shows part of a data fusion table as seen in Geoscience INTEGRATOR's web interface.

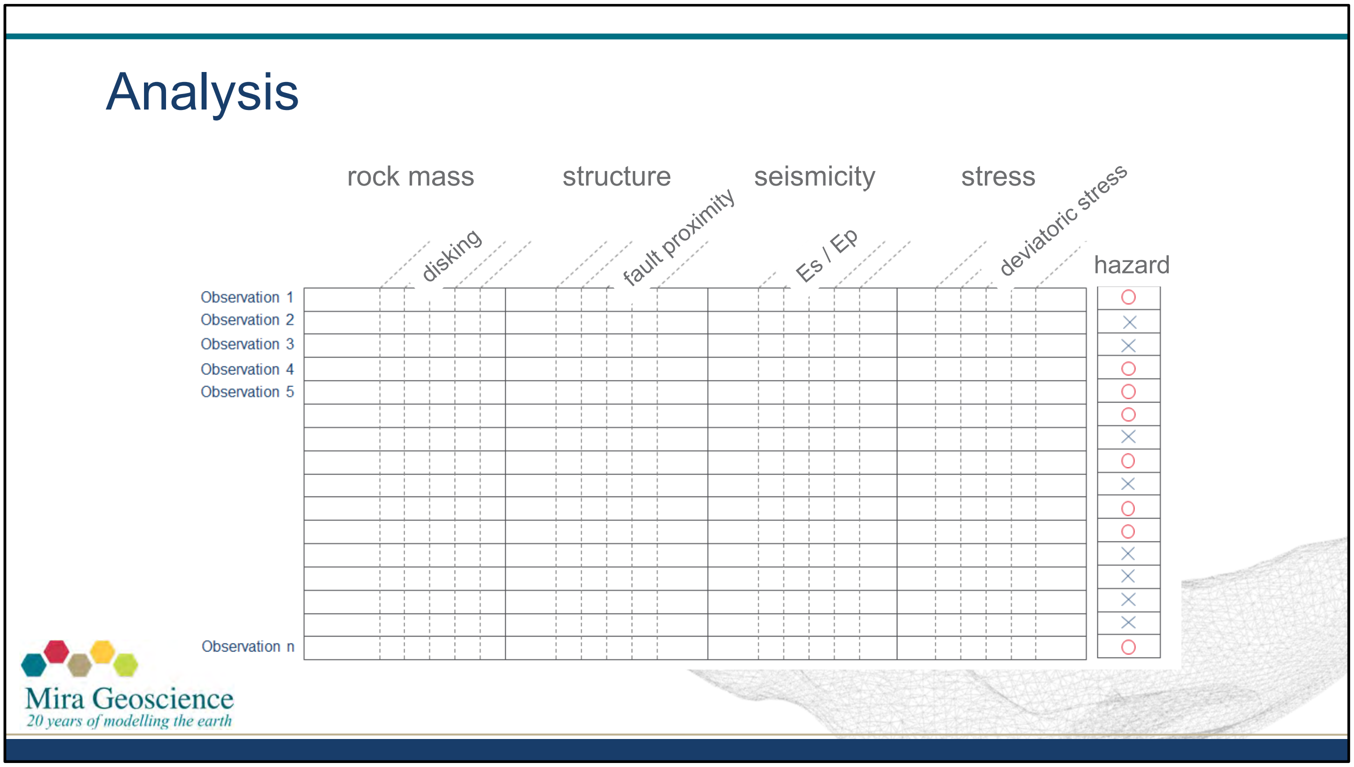
The workflow

- problem definition
- feature engineering
- data fusion
- analysis →
- deployment

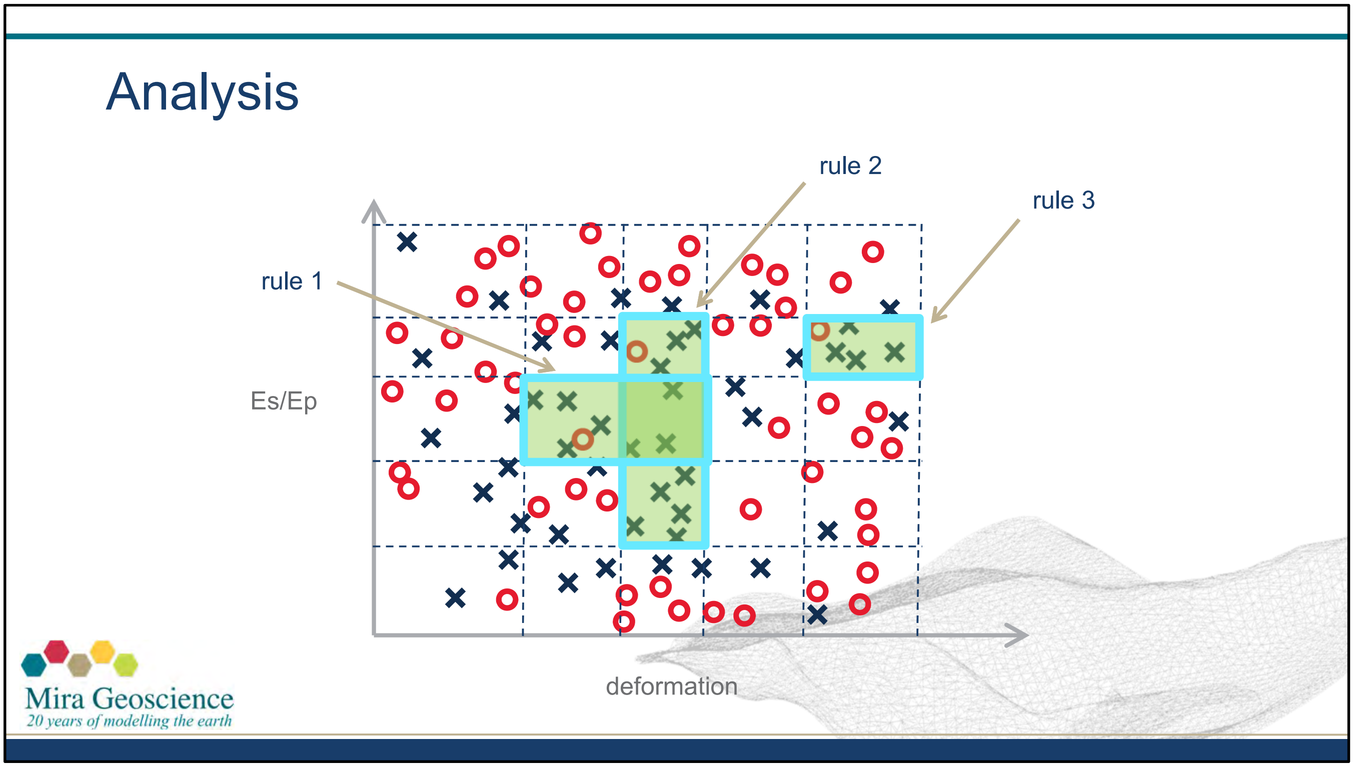
- statistical analysis
- machine learning

Mira Geoscience
20 years of modelling the earth

Once the data fusion table is created, we search for associations between the modelled variables and hazard using statistical analysis or machine learning.



The data fusion table in practice is typically represented as a large (a few 10s of columns and hundreds of thousands to a few million rows) csv file. Column files such as this are standard inputs to a wide variety of general statistical and machine learning tools.



This plot illustrates the general approach of the machine learning algorithms we employ. The X's and O's from the data fusion table are investigated in the multi-variate hazard criteria space (two of many hazard criteria dimensions being shown in the figure for simplicity) for areas of "over-concentration" of X's. Or to put it another way, what combinations of hazard criteria values, such as local fault-slip tendency and proximity to structure, have historically been associated with the occurrence of hazards such as rockbursts? Searching for such tell-tale associations in high-dimensional spaces (each hazard criteria being a dimension) cannot be readily identified with traditional statistical, visual, or observational approaches. However, machine learning algorithms can effectively comb through the multi-dimensional space using brute-force searches to discover relationships otherwise easily missed.

The associations of variables where hazards tend to be occur are captured by the system as "rules". The rules are the criteria-value bounds defining the multi-dimensional boxes shown by example in this two-dimensional plot. Output rules are directly interpretable by domain experts.

Analysis

Rule#	7		
	Criteria	Bounds	
3	blastDensity	0.004665	0.706523
13	distance_to_CuOre	1.672783	80.19519
15	distance_to_drift_intersections	5.160544	13.90811
6	dip_joints_dip_dir_175_185	43.21336	73.29842
10	distance_to_all_fault_intersections	58.96496	311.2795
17	distance_to_high_fault_slip	11.67736	144.8638
FOG#	6 4 7 9 15 111		

Rule#	44		
	Criteria	Bounds	
20	distance_to_NiOre	143.2019	186.1528
25	LocalMag	-2.75641	-2.6935
24	EsEp_log10	0.692114	0.784783
14	distance_to_dikes	7.413339	77.79344
10	distance_to_all_fault_intersections	14.61875	108.2473
FOG#	13 17 5		

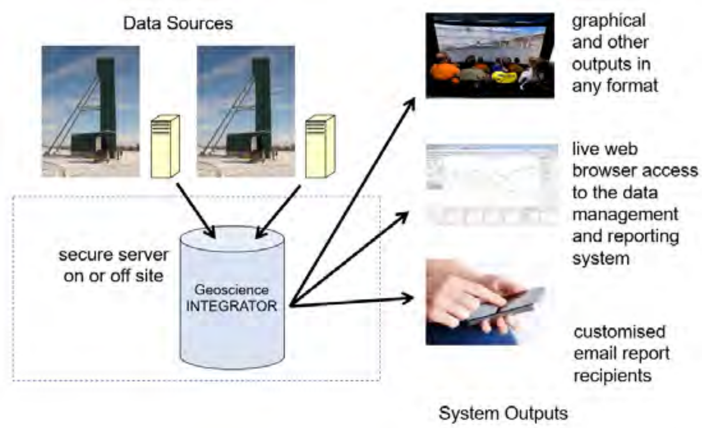
Mira Geoscience
20 years of modelling the earth

Rules output from machine learning can be inspected, interpreted, and validated. This process is called “rule mining”.

The workflow

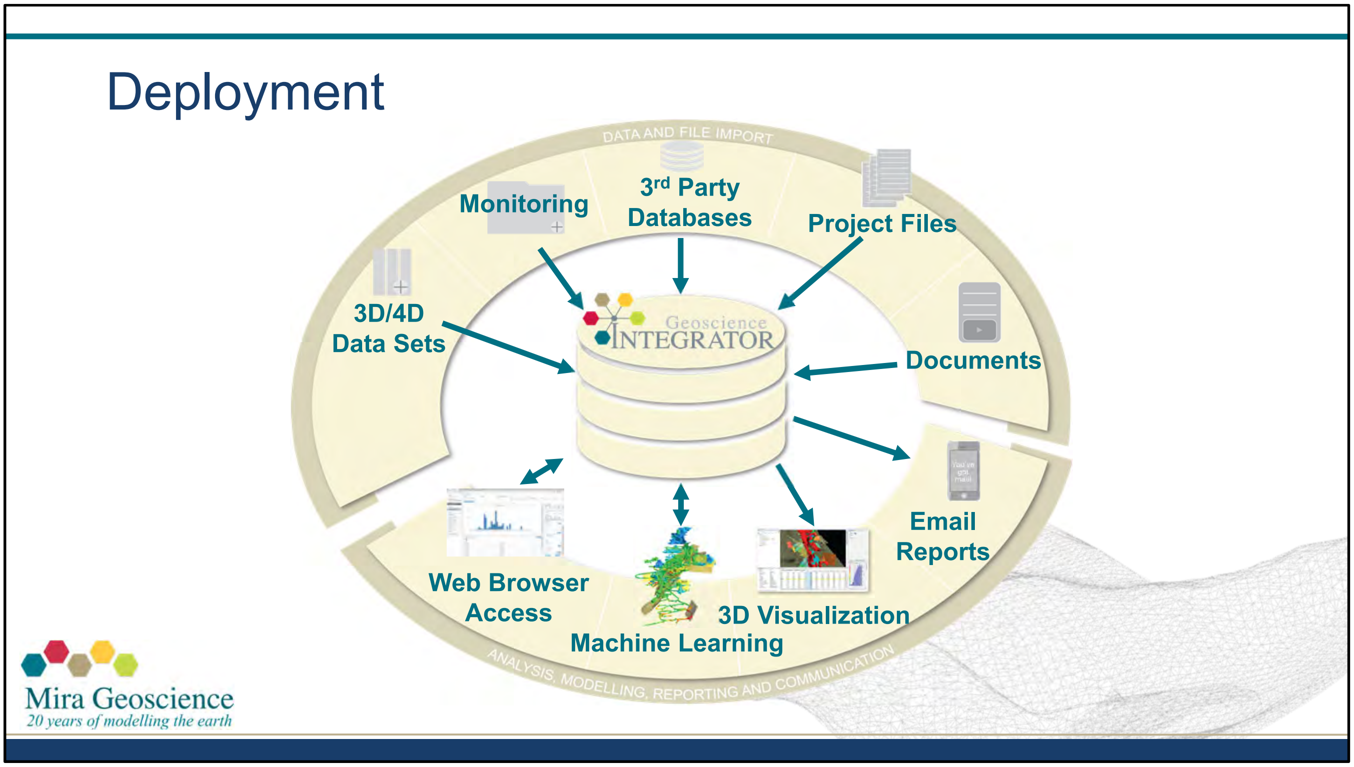
- problem definition
- feature engineering
- data fusion
- analysis
- deployment

- bringing it all together on the minesite



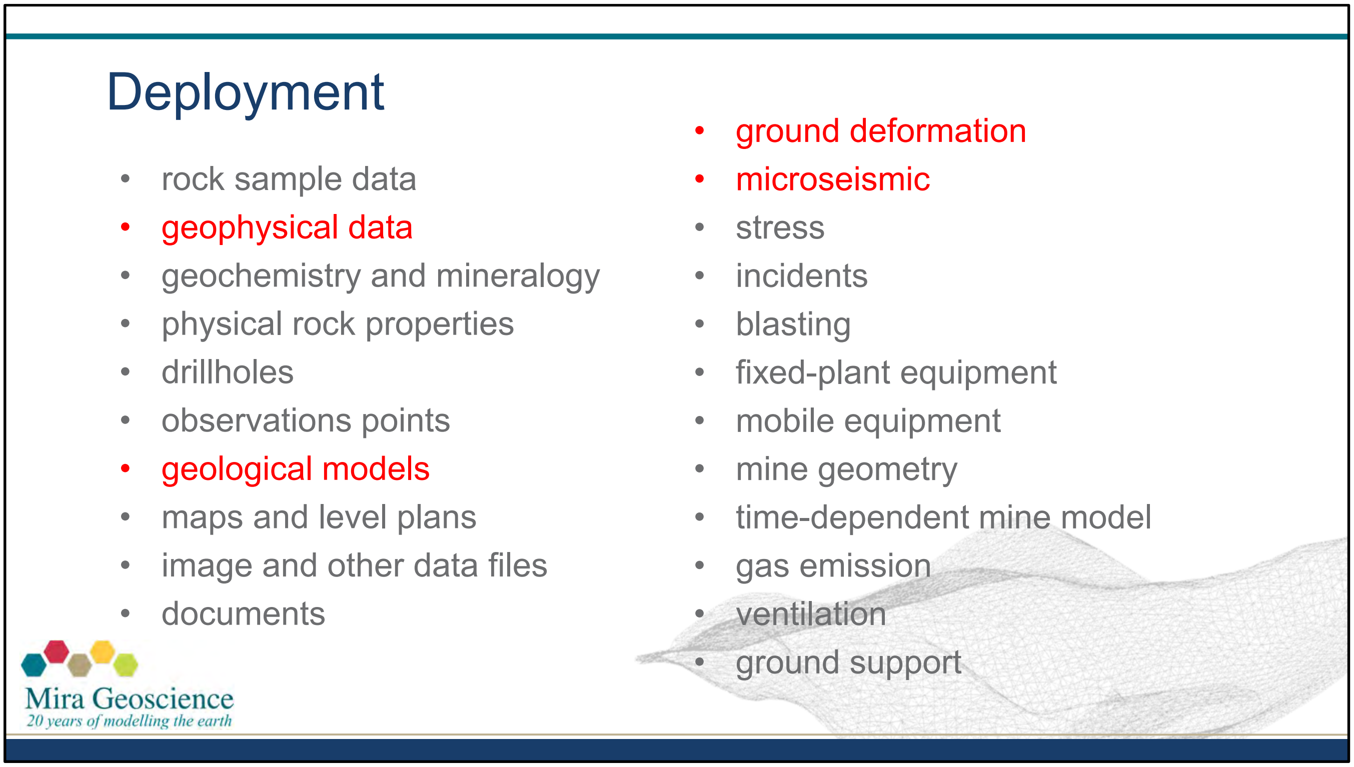
Mira Geoscience
20 years of modelling the earth

This is what it a deployed system looks like in its general design configuration – data input from multiple projects or mines into a structured system, making data and results available to all stakeholders, wherever they may be. Input of data from continuous monitoring systems is automated through monitored folders on the network file system or customized connections to site databases.

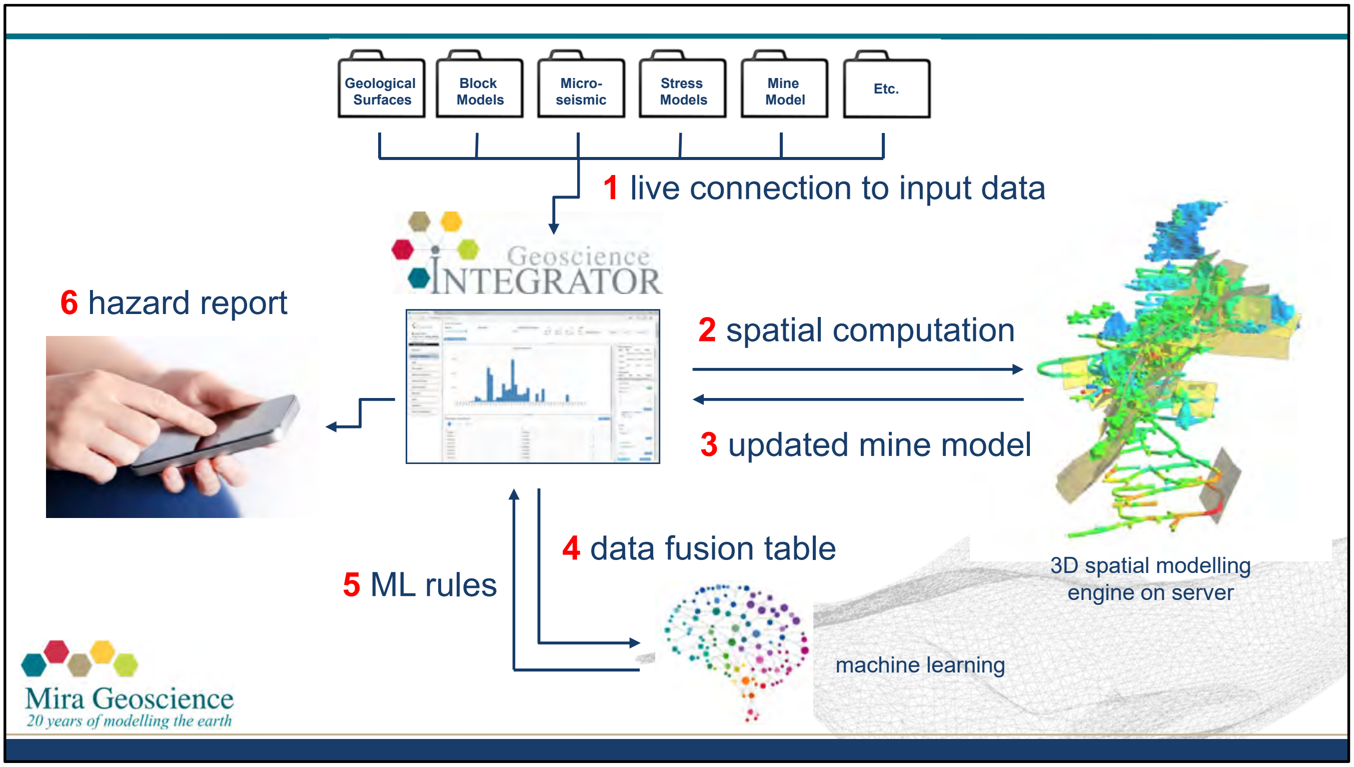


We have created a secure system which serves three main objectives:

1) Organization of multiple streams of data in a structured system, making that data available for all stakeholders, and structuring it for the geohazard assessment and other business problems requiring integrated interpretation.
2) Providing customized, ad-hoc access, reporting, and visualization of multiple data types from a secure, multi-user, easy-to-use interface.
3) Performing geohazard assessment in quasi-real time.

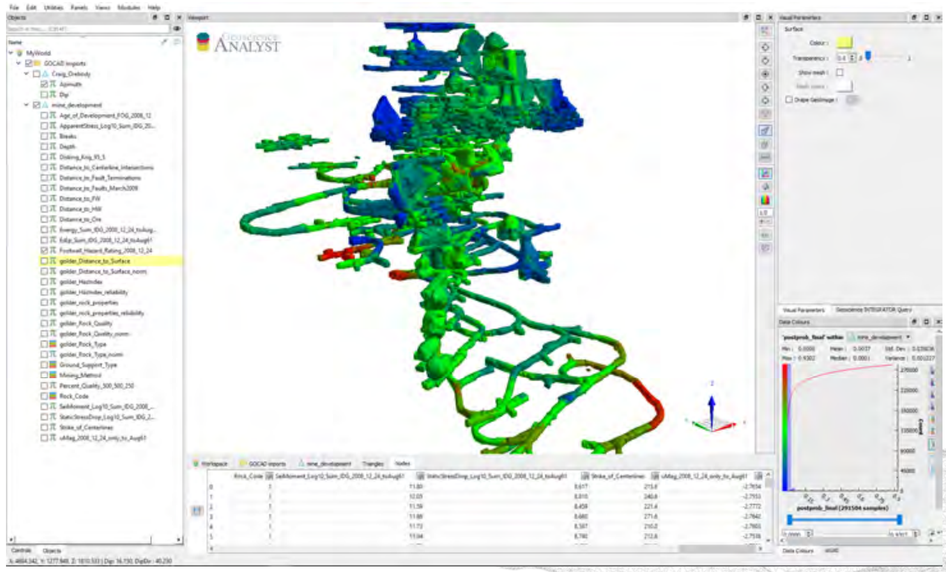


Data formats are free-form, using user-established templates containing any number of fields. Data may be integer, float, binary, classification, alphanumeric, or date/time. Time variance is supported on all data types for which it makes sense. QA/QC is automated on input according to user-defined rules.



This slide shows the data flow on the server. It is completely automated, with GOCAD running as run-time spatial computation engine on the server (without the GOCAD user interface). There is no need for users to learn complex modelling software.

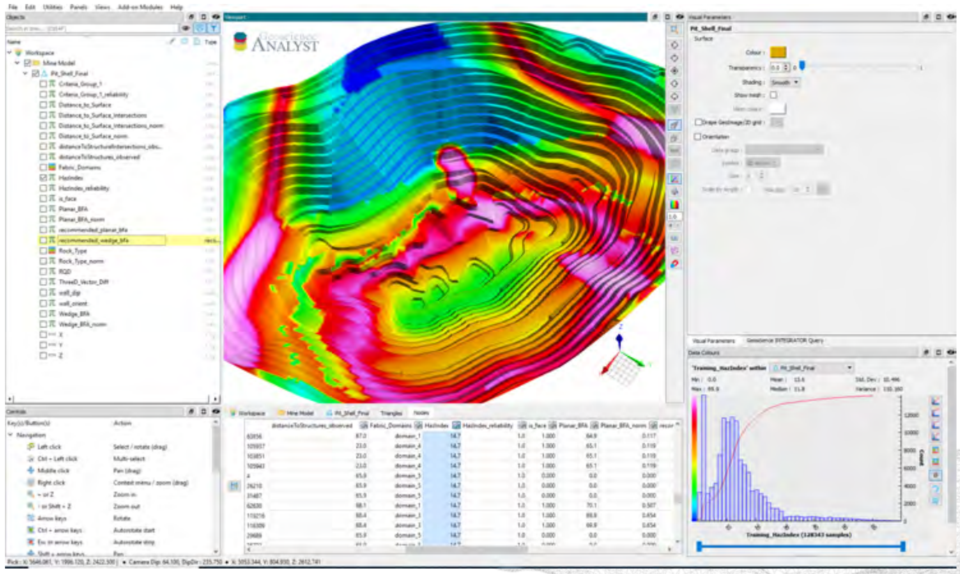
3D visual query interface



Mira Geoscience
20 years of modelling the earth

A powerful 3D visualizer is provided with an ability to query the server, and download only those files, documents, data set summaries, or data requested. Data are shown in linked graphical, tabular, and histogram displays.

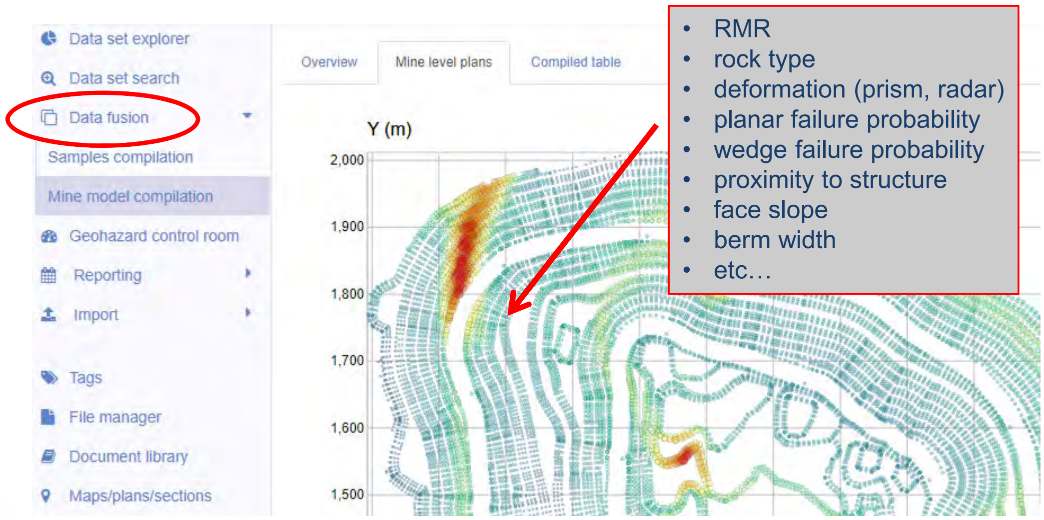
3D visual query interface



Mira Geoscience
20 years of modelling the earth

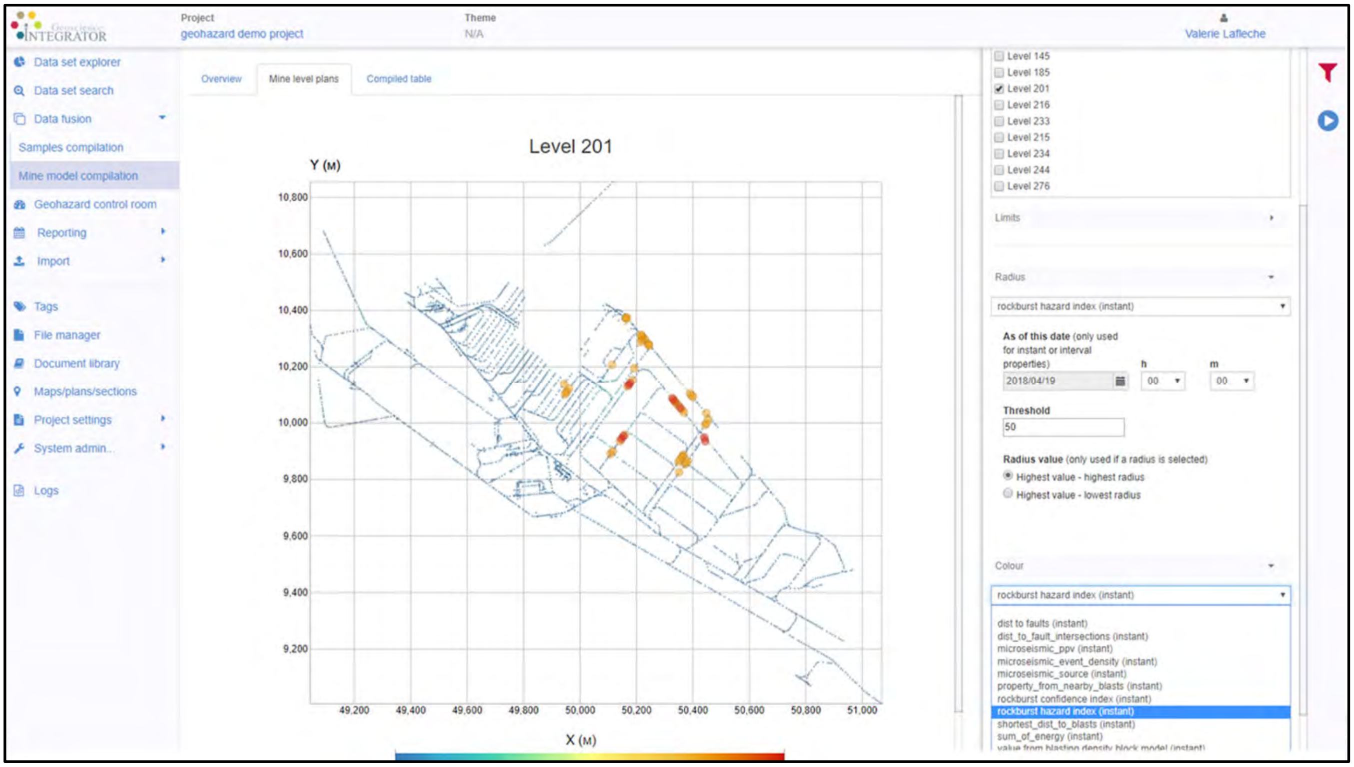
A powerful 3D visualizer is provided with an ability to query the server, and download only those files, documents, data set summaries, or data requested. Data are shown in linked graphical, tabular, and histogram displays.

Data fusion

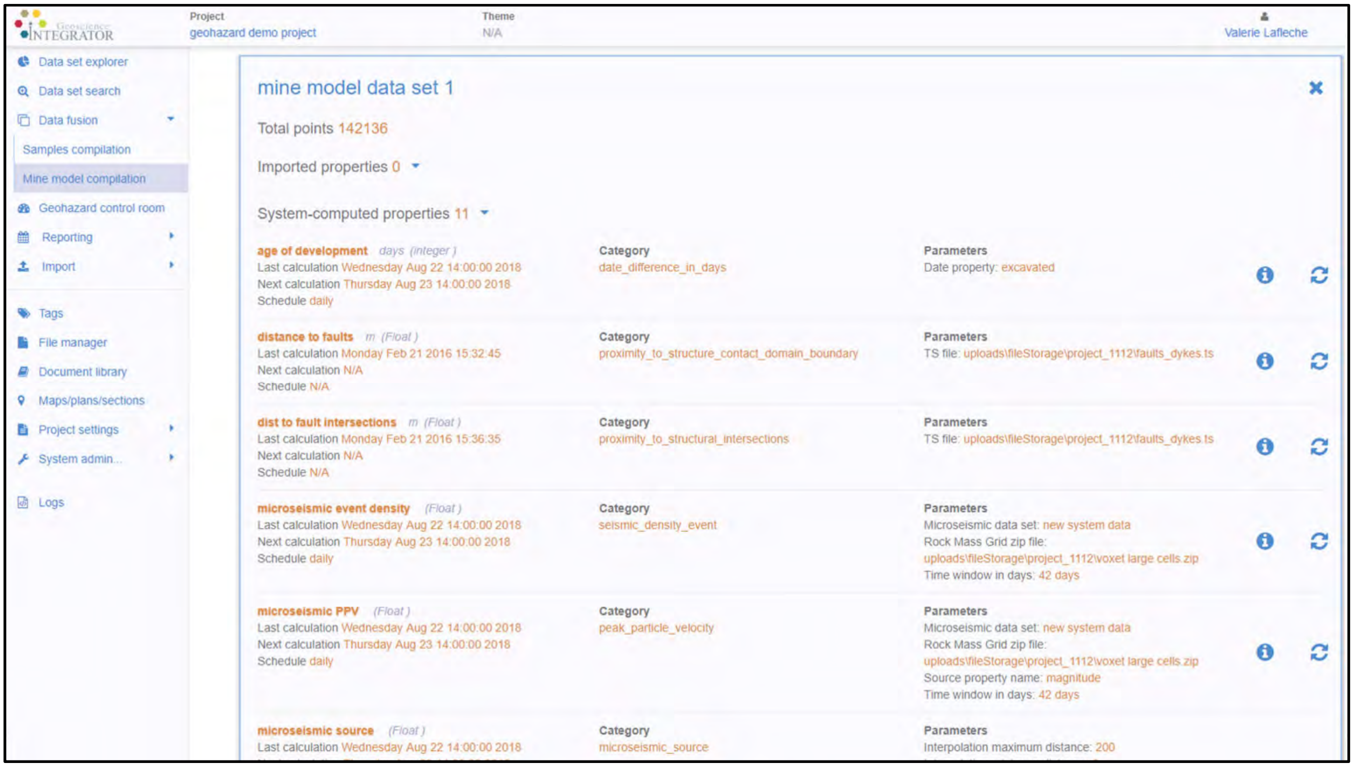


Mira Geoscience
20 years of modelling the earth

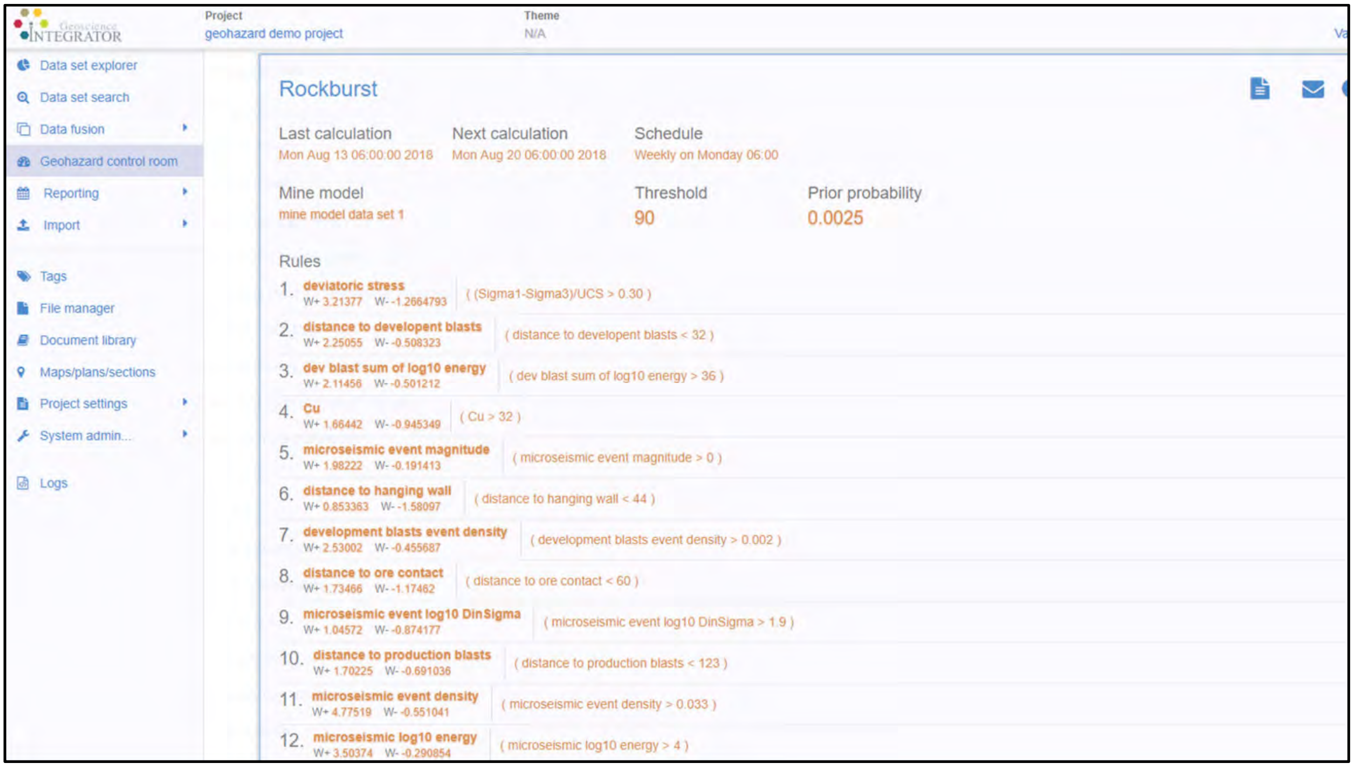
The mine model can also be conveniently viewed through the web browser, in this case showing a digitized pit shell.



Individual hazard criteria values, or the hazard probability forecast itself, can be displayed on selected mine levels in the web browser as well as in the 3D visualizer.



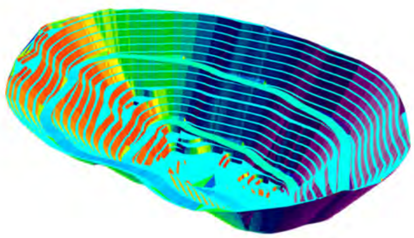
This slide shows the part of the UI where the automated hazard criteria computation is controlled. “System-computed properties” are updated on a defined schedule or on demand. The “Category” column corresponds to the class of function automatically run on the on-board GOCAD run-time engine within Geoscience INTEGRATOR.

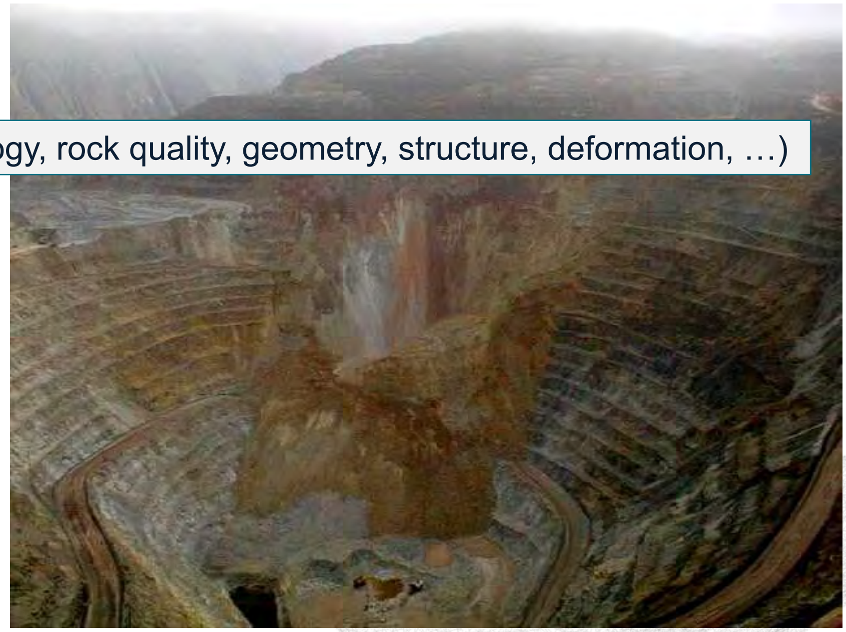


The definition of the hazard formulation "rules", threshold trigger alert levels, and calculation schedule is controlled in this interface.

Geohazard

hazard (x, y, z, t) = *f* (geology, rock quality, geometry, structure, deformation, …)





Mira Geoscience
20 years of modelling the earth

So, in summary, we have developed a workflow through which the geotechnical hazard assessment equation can be shifted from qualitative and knowledge-driven to quantitative and data-driven. Geophysical data is often a critical component in the solution. We believe it is a significant achievement.

Acknowledgements

- Centre for Excellence in Mining Innovation (CEMI)
- Ultra Deep Mining Network (UDMN)
- Canada Mining Innovation Council (CMIC)
- National Research Council (NRC)

Mira Geoscience
20 years of modelling the earth

Our work in developing Geoscience INTEGRATOR over the past 5 years has been supported by these research organizations, and in most cases funded by mining industry sponsors acting through these organizations.