Paper 36

Software Workflows for 3D Quantitative Data Integration

Perron, G. [1]

_____

1. Mira Geoscience Ltd.

ABSTRACT

During the last decade, breakthroughs in 3D geological and rock property modelling, geophysical inversion, and advanced visualization, coupled with the evolution of computer hardware, has made it possible to implement quantitative, multi-disciplinary, data integration projects in mineral exploration. Data integration is usually done through a series of pre-established "processes" that capture each discipline's expert knowledge. Standard exploration software applications are not designed as quantitative data integration platforms. They are usually built as expert toolkits focused on specific tasks and are lacking in interconnectivity. A new type of software application that bridges the gap between the different disciplines and delivers expert processes to general practitioners is presented. Workflows are software applications designed to execute a sequence of modelling tasks via a dynamic user interface presented in the appropriate earth science terminology. The expert's knowledge is captured in the form of a "process tree" where each branch is defined by a series of dynamically linked decision panels. By navigating through the different panels, the user makes decisions that are based on context-relevant choices, triggering appropriate subroutines to transform the input data into the expected output. Workflows are easily customizable, allowing the embedding of specifications proper to an exploration company's best practices. The example of a workflow designed for 3D quantitative integration of geological and geophysical data through geologically constrained potential field inversion is used in this paper to illustrate how such software applications are designed and implemented. Several benefits are associated with the use of workflows. Capturing the expert thought process leads to technically sound, reliable, and repeatable results. Workflows also have a direct positive impact on employees' learning curve and productivity. Productivity increase is achieved through streamlined work processes based on proven methodologies that focus on the science at hand instead of software mechanics. The workflow concept may in itself contribute to the emergence of a new type of explorationist, able to cross boundaries between different disciplines at an advanced level. The case study using the potential fields inversion workflow demonstrates these benefits.

INTRODUCTION

With the availability of new exploration technologies, such as geologically-constrained geophysical inversion, enabling practitioners to quantitatively integrate data from different earth science disciplines in 3D, the "common earth modelling" concept emerges as a new exploration paradigm for the mining industry. A common earth model, a term borrowed from the oil and gas world, is defined as a single 3D representation of the subsurface where models derived from the different exploration disciplines (geology, geophysics and geochemistry) are quantitatively consistent with each other. This new concept implies that explorationists work ever more closely as a team as their discipline-specific approaches no longer work in isolation from each other. New work processes based on new exploration technologies need to be defined or adapted to enable true quantitative data integration between disciplines. This also creates a challenge from a software point of view. A true 3D quantitative data integration software platform should provide the basic vector and raster representations for all types of exploration data as well as the required functionalities to construct and display 3D models interpreted from the initial raw observations. Gocad was originally designed and developed with the common earth modelling approach as its background philosophy. It encompasses all the requirements of a common earth modelling platform plus an advanced application programming interface (API) for developing new functionalities or links with external applications. While standard API's were proven very useful to develop discipline-specific software plug-ins, their structure does not adequately provide the ability to capture and represent the often non-linear, domain-expert thought process needed to perform complex data integration tasks. To answer this need a new programming environment within Gocad was developed. It is based on the concept of "decision graphs", the Model-View-Controller architectural pattern (Goldberg and Robson, 1983) and the XPDL programming language (van der Aalst, 2004), and provides a complete state-of-the-art environment to develop workflow applications needed for 3D data integration and modelling tasks. Workflows designed with these concepts have been in commercial implementation for several years for structural and resource modelling of oil and gas reservoirs.

Implementation of domain experts' thought processes as workflows allows practitioners with less expertise to execute complex 3D data integration tasks yielding robust, consistent, and repeatable results. Because workflows focus the user on the sequence of tasks at hand and not on the software mechanics, employee productivity is increased. The ability of workflows to report flow parameter choices while integrating user's notes and key images leaves behind a solid audit trail for standard reporting, quality control, and employee training purposes.

The main objectives of this paper are to provide a rationale for the need for workflows, a brief overview of why they are different from scripts, macros, and wizards, how they are designed and built, and the inherent benefits associated with this new type of software application for our industry.

WORKFLOW FUNDAMENTALS

Mineral exploration data integration is not a linear process with a unique path from beginning to end. The example of geologically-constrained 3D potential fields inversion reflects this complexity. Users need to be guided through a myriad of input parameter choices in order to produce results that are consistent with a priori geological knowledge. It is often difficult to learn and remember the list of operations required to do geophysical inversions, and the order in which these operations should be performed. This is why geophysical inversions are usually assigned to experts who are able to understand all the dependencies between the operations involved and execute them in the right order. For non-geophysicists on the exploration team, software workflows make complex processes like geophysical inversion more accessible. For the domain expert, they provide a mechanism for capturing, documenting, and re-using proven process parameters, and communicating them to colleagues. Errors are reduced and the time spent to complete complex processes is decreased by an order of magnitude.

Wizards shortcomings

Why do we need to develop workflows when we already have access to wizards? Software wizards are widely recognized in the world of computer science. There are successfully used to perform countless common tasks such has hardware/software installation, document printing, and connecting computers to office networks. Although they make seemingly complex tasks simpler, the tasks they automate are linear and can be seen as a concatenation of individual operations (nodes) in one specific order. Wizards, although useful, have several limitations that prevent them from capturing complex data integration processes:

1. Operations may be activated in any order even if there are dependencies between them.
2. If the task is too complex, it has to be divided into several independent wizards, generating the same problem of task sequencing faced with standard pull-down menus.
3. It is impossible to produce a report encapsulating the user's parameter selection and personal notes.

4. Wizards are based on the concept of "decision trees" that cannot really account for repeating a single operation (loop) or a series of operations (cycle). It is also difficult to track decisions and allow the user to invalidate a previous decision by going backwards.
5. In wizards, choices are not decoupled from operations, which can be confusing to the user.

Workflows are designed to overcome these deficiencies and to widen the scope of problems that can be addressed by software-guided processes.

Principal workflow components

Workflows are different from wizards in terms of their internal organization and the level of complexity for which they can account. Unlike wizards that are based on a "decision tree" concept, workflows are implemented as "decision graphs" (Figure 1). Decision graphs can have multiple entry and exit points allowing the user to momentarily quit the workflow environment to execute complementary actions, such as importing additional data to be used later in the workflow process. It is also possible to define loops and cycles within workflows to repeat a single, or a set of, key operations respectively. With decision trees, once the user makes a choice to go down a certain branch it is impossible to jump to another. Within decision graphs such behaviour is allowed. In decision graphs, junctions which prompt the user to make a choice are usually decoupled from operations. Junctions have an indirect impact on the process while operations directly influence the workflow results. Another important notion associated with decision graphs is the possibility to nest them. Each node of a graph may be a new decision graph. This allows dividing a complex data integration process into a series of individual sub-processes. This level of granularity helps to break down a complex sequence of events otherwise difficult to describe.
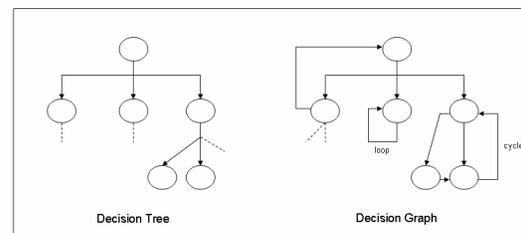
Figure 1: Examples of a decision tree and decision graph.

Sometimes when stepping through a process, it is mandatory to perform a certain task before another, thereby defining a compulsory rule. Sometimes a certain task should not be made available if another was performed prior to it. This constitutes an exclusive rule. Both compulsory and exclusive rules can be defined using a decision graph. These rules are required to map the complexity of approaches like geologically-constrained 3D potential field data inversion.

Navigation paths through a series of decision graphs and associated parameter selections can also be automatically recorded within a workflow session. At any time during the execution of a workflow, the session records its position within a

decision graph, the current status of an operation or junction. This allows easy tracking of information when a user is moving forward or backward in the workflow. The session is also at the basis of the reporting component of the workflow.

The graphical user interface of a Gocad workflow is based on a Model-View-Controller (MVC) architectural approach. Often, users would like to dissociate data/task (model) and user interface (view) concerns, so that changes to the user interface do not affect the data handling, and that the data can be re-organized without changing the user interface. The Model-View-Controller solves this problem by decoupling data access and algorithm execution from data visualization and user interaction. It does so by introducing the controller as an intermediate component. The controller translates inputs for model operations or for view display requests (Buschmann et al., 1996). The MVC architecture is required to implement the user interface style used by Gocad.

The Gocad workflow user interface layout

The workflow interface is split into two main areas (Figure 2). A first area shows the decision graph as a tree. Each tree branch corresponds to a junction in the workflow while leaves are associated with operations. It is possible to display the workflow decision graph as a tree because only the steps going forwards to the solution are registered in the tree.

The second area shows the decision/action panel corresponding to the selected branch/leaf of the tree. The tree and panel representations are dynamically linked. The elements viewed within a panel are dependent on the choices made in the previous panel, confronting the user only with the next logical choices. This way, it is impossible to drift away from the workflow path, preventing users from making inappropriate decisions. The decision panel area also includes two alternative views available by selecting their respective tabs. The info tab reveals contextual help specific to the corresponding decision panel while the report tab gives the user access to an editing window to take notes and snapshots of the 3D model.

Overall, the workflow architecture answers the shortcomings of its wizard predecessors, and uses state-of-the-art software concepts that provide the necessary tools to capture, implement, and deliver complex expert processes to the expert and non-expert alike.

THE POTENTIAL FIELDS WORKFLOW

The example of geologically constrained 3D geophysical inversion is chosen here to illustrate the use of workflows for 3D data integration in the context of mineral exploration. The Potential Fields Workflow (PFW) is presented. In order to quantitatively reconcile geological data with potential fields data, one needs to compute a 3D distribution of densities or magnetic susceptibilities that respects pre-established value ranges for each lithology present in a given geological model and reproduces the measured field within a tolerance margin. This expert process involves a complex series of operations,

some of which need to be done in a specific order, to achieve the desired output.

Planning of the workflow layout from a software-design perspective requires the contribution of geophysical inversion specialists as well as the software architects. Every possible path between input data and the final result needs to be investigated. Key steps have to be identified and grouped as workflow decision graphs.
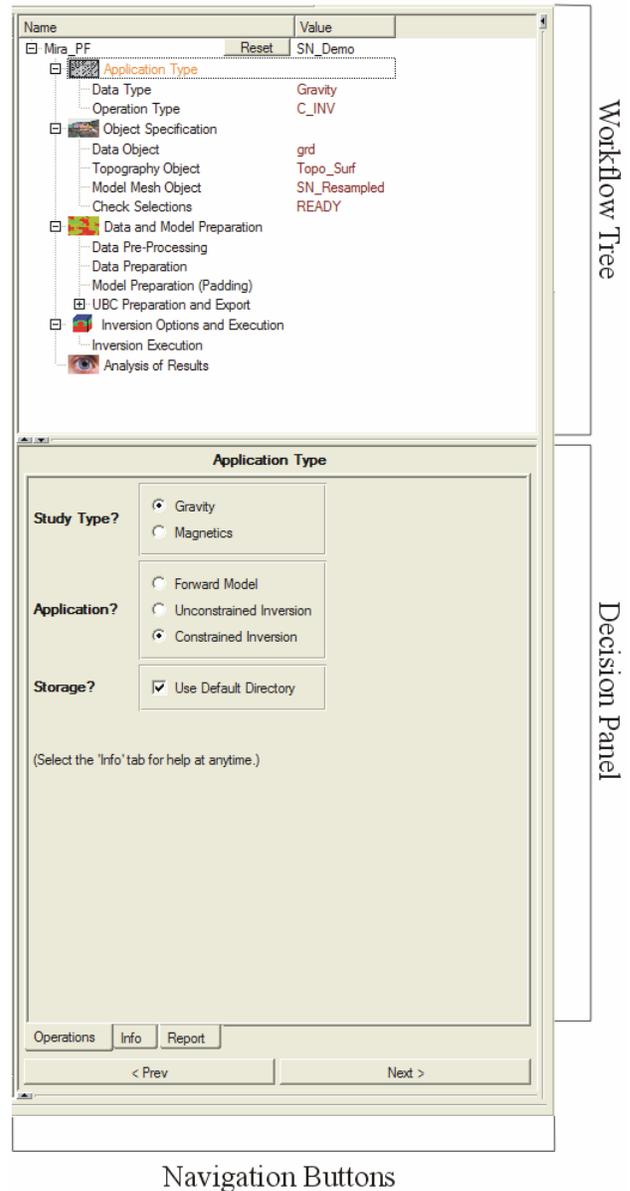

Figure 2: Workflow user interface.

Workflow specifications for constrained potential field inversion

PFW links the Mag3D (Li and Oldenburg, 1996) and Grav3D (Li and Oldenburg, 1998) inversion software from the

University of British Columbia's Geophysical Inversion Facility (UBC-GIF) with the Gocad 3D geological modelling and 3D-GIS environment. Gocad provides geologists tools to construct detailed 3D models attributed with rock types from surface and borehole observations. A 3D density or magnetic susceptibility "reference model", constructed from geological and rock property data, represents the a priori geological and petrophysical knowledge of the sub-surface. The constrained inversion process will perturb the reference model (as minimally as possible) by refining the density or magnetic susceptibility distribution so that the modelled response matches data within a prescribed uncertainty. The result is a 3D model quantitatively consistent with both geological and geophysical observation.

The general procedure to properly implement geologically constrained 3D potential field inversion involves the following key steps:

1. Selecting parameters that define the nature of the workflow study in terms of application and data types.
2. Selecting the required input data, topographic information, and geological constraints as Gocad objects.
3. Preparation of data and model, including data pre-processing and model definition.
4. Selecting inversion options and running the inversion.
5. Analysis of results.

Each of these steps includes workflow branches, loops and cycles.

The reference model and the measured potential fields data represent the minimum input required for a constrained gravity inversion. Systematically, the workflow presents the user with decision panels to select the input data and topography information and perform the first quality control on the input parameters. Checks are performed to ensure that the station locations are above the topographic surface, that the reference model grid is compliant with the UBC-GIF mesh file format, and that the topographic surface covers the extents of the grid.

Pre-processing of the data prior to the inversion is a key step. Four topics are addressed in this decision panel: 1) regional response; 2) DC shifts; 3) data resampling (gridding); and 4) data error as a convergence criterion. In the first prototype of PFW, regional response is accounted for by polynomial trend removal (first, second, and third-order). More sophisticated methods such as incising higher-resolution local models into coarser regional inversions will be implemented in future versions. DC adjustments to the gravity data, handled as a simple workflow step, are typically required in constrained inversion approaches so that anomaly amplitudes match computed responses from the reference model. Gridding of the data at a regular spatial interval is often desired to improve convergence. In constrained inversion, data are often gridded at the same interval as the reference model grid. PFW currently presents two gridding methods: Gocad's patented "discrete smooth interpolation" and conventional inverse distance. If data errors have not been defined, the workflow offers two options: constant or constant plus data amplitude percentage. Errors can also be assigned easily on a per-datum basis.

In the workflow sequence, model preparation immediately follows data pre-processing. Model preparation steps are highly dependent on the objective of the individual study, differing for gravity and magnetic data, forward modelling and inversion, the types of constraints imposed by the geological model, and model-specific characteristics such as the severity of topographic relief.

Padding the model with extra cells both laterally and at depth may be required to avoid grid-edge effects. These are often significant when considering realistic earth models having background geological variation throughout the model, as opposed to investigating anomalies situated within homogeneous half-spaces. Three different types of padding schemes are currently available within PFW: 1) constant cell size; 2) increasing cell size; and 3) a combination of the two previous cases. There is also a possibility of tapering the density and magnetic susceptibility values to a user-specified value as a function of distance from reference model edges.

Model padding sometimes extends beyond the aerial coverage of the available topographic surface. The inversion algorithms require that a topographic elevation measurement be assigned everywhere over the inversion grid. The workflow will automatically detect any lack of topography coverage and propose to extend the triangulated topographic surface to fit the extents of the newly padded model.

At this stage of the workflow, all model and data requirements for the inversion algorithm have been fulfilled, and the inversion engine itself is run using input passed from the workflow. Modifications will be made iteratively to the reference model in order for its computed potential fields response to converge towards the measured data. The final output model is consistent with both the geophysical data and the geological constraints, and thus fulfills the objective of quantitative 3D data integration.

The workflow is also used in post-processing, providing tools to understand the inversion results. A simple example is displaying the perturbation of the reference model as demanded by the data. This is a 3D spatial model indicating where the 3D geological model was most inconsistent with the geophysical data, often providing vital evidence for drillhole targeting.

SAN NICOLAS CASE STUDY

The San Nicolas deposit, Mexico, is a typical VMS system. It has been the focus of several geophysical studies showing a clear anomaly with both potential fields and electrical methods (Johnson et al., 1999). A simple constrained gravity inversion, using a litho-petrophysical reference model, illustrates a typical output of the potential fields workflow. A 3D geological model of the faults, formational contact surfaces, and interpreted massive suphide was built rapidly in Gocad from a set of interpreted cross-sections provided by TeckCominco. The model is automatically rasterized and represented as a 3D regular grid where each voxel is attributed with a lithology type. Each lithology is then characterized in terms of its density using basic statistical values (average, minimum and maximum). This constitutes the litho-petrophysical reference model (Figure 3a) that will be tested against the gravity data.

Figure 3 illustrates some results from a simple application of PFW. The reference density model of Figure 3a encapsulates the both the current understanding of density variation and is also used to communicate many constraints, such as property

bounds, weighting, and smoothness parameters to the inversion engine through the workflow. The 3D grid shown thus contains many numerical properties in addition the density property illustrated.
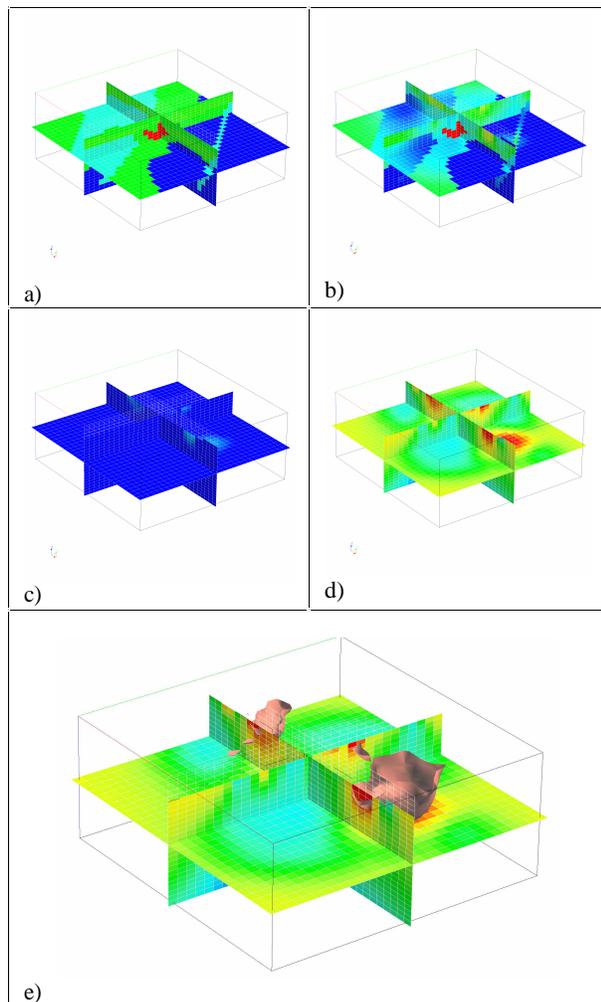
Figure 3: a) reference 3D density model shown as two cross-sections and one horizontal slice; b) constrained inversion results; c) model perturbation required by the gravity data in the same colour scale as a) and b); d) model pertubation of c) with colour scale adjusted to better show spatial model perturbation variation; e) automatic isosurface computation around volumes of excess mass not accounted for in the original geological model.

These properties are constructed as the user proceeds through the workflow, depending on many data-specific and model-specific issues as discussed. Figure 3b shows the post-inversion density model, in which the basic character of the initial reference model is respected (as are all the geological and petrophysical constraints), but heterogeneity has been added to force consistency with the gravity data. Figures 3c and 3d, which are the same except for a change in the colour scale, show the spatial variation of the density perturbations required of the geological reference model in order to explain the gravity data. The anomalies of Figure 3d can be thought of the residual

anomalies remaining after taking into consideration all a priori knowledge, including all geological knowledge. This is a vast step ahead of the traditional geophysical data-processing practice of "correcting" for terrain, elevation, latitude, etc., while not quantitatively accounting for prior knowledge of either 3D geological variation or physical rock properties. The automatic isosurface computation of Figure 3e provides a summary output showing the envelope of excess mass (at a certain user-specified threshold) not accounted for in the original geological-petrophysical reference model.

CONCLUSIONS

The simple example shown in the case study illustrates an important, valuable exploration result that can directly affect targeting or other business decisions on an exploration property. The process, and the sequence of steps required to implement it, is conceptually simple. Yet, in practice, organizing the sequence of steps to execute it is complex, demanding in its necessity of attention to detail, and replete with decisions requiring subtle understanding of potential field data and its optimum treatment. The software workflow paradigm successfully addresses these obstacles to proper deployment of modern potential field inversion technology by anticipating the details, decisions, and objectives of such work.

Productivity increase, improved quality of results, internal quality checks, repeatability, reporting, audit trail, bridging the gap between disciplines, and the need for organization-specific customization are all effectively addressed by the workflow, the next generation of exploration software interface.

REFERENCES

Buschmann, F., Meunier, R., Rohnert, H., Sommerland, P., and Stal, M., 1996, Pattern-oriented Software Architecture: John Wiley & Sons.

Goldberg, A. and Robson, D., 1983, Smalltalk-80: the Language and its Implementation: Addison-Wesley Longman Publishing.

Johnson, B., Montante, A., Kearvell, G., Janzen, J., and Scammell, R., 1999, Geology and exploration of the San Nicolas polymetallic (Zn-Cu-Au-Ag) volcanogenic massive sulphide deposit, in Jambor, J.J. ed., VMS and Carbonate-Hosted Polymetallic Deposits of Central Mexico, B.C. and Yukon: Chamber of Mines Cordilleran Roundup, 45-54.

Li, Y. and Oldenburg, D. W., 1996, 3–D inversion of magnetic data, Geophysics, 61, 394–408.

Li, Y. and Oldenburg, D. W., 1998, 3D inversion of gravity data, Geophysics, 63, 109–119

van der Aalst, W.M.P., 2004, Business Process Management Demystified: A Tutorial on Models, Systems and Standards for Workflow Management: Springer Lecture Notes in Computer Science, V3098.