3D EM Modelling and Inversion with Open Source Resources

Douglas W. Oldenburg
University of British Columbia, Geophysical Inversion Facility
Vancouver, BC, Canada
doug@eos.ubc.ca

Lindsey J. Heagy
University of British Columbia, Geophysical Inversion Facility
Vancouver, BC, Canada
lheagy@eos.ubc.ca

Seogi Kang
University of British Columbia, Geophysical Inversion Facility
Calgary, AB, Canada
skang@eos.ubc.ca

Rowan Cockett
University of British Columbia, Geophysical Inversion Facility
Vancouver, BC, Canada
rowanc1@gmail.com

SUMMARY

Great advancements in using airborne electromagnetic (EM) geophysical data have been made over the last few years but there is still much to be done. The next generation of problems will require multiple airborne and ground EM surveys to provide information about the three (generally frequency dependent) physical properties (σ,μ, and ε). The end goal will not be delineating the spatial distribution of these properties, but rather integrating that information to help solve multi-disciplinary problems associated with earth's resources and our environment. These problems are complex and require collaboration. This is especially true at the base scientific level where the underlying physical equations need to be solved and data, associated with physical experiments, need to be inverted. In this paper, we present arguments for adopting an open source model for geophysics and provide some background about open source software for electromagnetics. Immediate benefits are the reduced time required to carry out research, being able to disseminate results quickly, have results be reproducible, and the ability to collaborate. To illustrate the use of open source resources in electromagnetics, we present two challenges. The first is to simulate data from a time domain airborne system over a conductive plate buried in a more resistive earth. The second is to jointly invert airborne TDEM and FDEM data and ground TDEM. For the open source software we use SimPEG (Simulation and Parameter Estimation in Geophysics, http://simpeg.xyz). The figures in this paper can be reproduced by downloading the Jupyter Notebooks. Access to the codes allows the user to explore the simulations and inversions by changing model and inversion parameters, plot fields and fluxes to gain further insight about the EM phenomena, and solve a new research problem by using this software as a base. By providing results in a manner that allows others to reproduce, further explore, and even extend them, we hope to demonstrate that an open source paradigm has the potential to enable more rapid progress of our community as a whole.

Key words: frequency domain, time domain, inversion, open source software, finite volume

INTRODUCTION

The abstracts provided for the 7th AEM workshop provide a solid overview of the progress that has been made in solving airborne electromagnetic (EM) problems and in applying laboratory, field, and numerical research results to increase the use and usefulness of electromagnetics. With modern systems, particularly airborne (helicopter or fixed-wing) systems, it is now common to acquire hundreds-of-thousands of time and frequency domain soundings in a single survey. Inverting those data to recover a 3D electrical conductivity model of the earth in a reasonable amount of time has been one of the main focus for the EM community. Different approaches (including finite element, finite volume, and integral equations) can be used, but they all generally require designing a mesh, discretizing Maxwell's equations, and solving a linear system of equations. Today, the milestone of inverting the large data sets has been largely achieved, primarily through the combination of advances including: (a) development of better numerical solvers for the discretized Maxwell's equations involving a single source (e.g. direct solvers); (b) use of (semi-) unstructured meshes to reduce the number of variables; (c) separating the forward modelling and inversion meshes, (d) computing and storing sensitivities or applying the operations of a sensitivity times a vector; and (e) access to multi-cores so that subproblems can be solved in parallel. There are currently numerous companies and universities that can carry out this scale of 3D inversion. There will of course always be further developments to generate faster, more flexible algorithms and incorporate system parameters etc., but the goal of recovering 3D conductivity from airborne time domain EM (TDEM) and frequency domain EM (FDEM) systems has largely been achieved.

Tackling large-scale EM inversions is a milestone, however, it is not the only research question to be addressed in EM geophysics. Fundamentally, Maxwell's equations depend upon three physical properties, conductivity (σ), magnetic permeability (μ), and electric permittivity (ε), and depending upon the frequency band used in the transmitter, each of these properties, or some combination of them, might provide the needed information. For instance at low frequencies, permittivity can be ignored, but at higher frequencies it becomes the dominant property and that promotes the use of radar systems. The physical properties can also be frequency dependent. For conductivity, the characterization of this is the IP induced polarization (IP) signal; for magnetic permeability it is viscous remanent magnetization (VRM). A problem of current interest in the airborne EM (AEM) community is extracting and interpreting the IP signal in the data; there are multiple routes for a solution, such as how to best parameterize the complex conductivity or whether to determine it iteratively or all-at-once. The VRM problem, most evident in near surface surveys such as those carried out for man-portable units searching for UXOs, also provide scope for this research. Finally, there are surveys in which combinations of potentially dispersive physical properties need to dealt with. There is considerable research needed to solve such problems. The challenge is exacerbated when one wants to combine multiple surveys (e.g. airborne TDEM, FDEM along with their ground counterparts).

The above problems have focussed on the goal of extracting better physical property information from EM systems. That can be important but the real value of EM is in integrating this information in the context of the application to help solve diverse earth science problems. For example, there are many issues connected with groundwater. Fundamentally these problems are in the realm of government or district water managers who want to know where the groundwater exists, how much is available, how are aquifers are being recharged (or not) etc. This requires knowing the geologic structure and hydraulic conductivity of the earth. EM geophysics certainly has a role to play, but integration with stakeholders or researchers from other disciplines, many of whom are not experts in geophysics nor even understand our lexicon, is another level of challenge. Some of these challenges might be addressed by combining surveys from the different disciplines into process-based inversion (e.g. EM and fluid flow) and some addressed by understanding the information (sometimes not geophysical) that is important and incorporating it into the forward modelling and inversion algorithms.

There are still many EM problems to be solved and more will arise as we show the positive impact that EM has on existing applications. These challenges will further increase as geophysical information gets integrated with information from other disciplines, to tackle important, multidisciplinary problems in resource exploration, environmental, and geotechnical areas.

THE RESEARCHERS AND THEIR CHALLENGES

Who will solve these problems and what are the challenges? Industry has potential to make progress here but commercial research is generally confined to specific problems that provide financial benefits, usually in a relatively short time-frame. Conversely, academic research, which tends to be more exploratory, could be greatly accelerated if they had access to existing software that would allow them to explore solutions and to build further improvements and functionality into the software. Some of the most computationally efficient codes for handling large scale data and carrying out forward simulation and inversion are in the hands of industry. However, using those codes to explore research questions is often not practical. Firstly, companies are often unwilling to share their codes because they may contain important technical advances and give the owners a competitive edge in a commercial market. Even if the codes are provided, they are “fit for purpose”; they allow a particular type of data as input, have rigid control on inversion parameters, and provide a specific output. This makes sense for a company as it can make their workflows more efficient, but research is often more open-ended, and as such, it requires that the researcher have access to all aspects of the code, the ability to plug in new pieces, and the flexibility to experiment with methodology. Furthermore, many of the production-level codes are impenetrable (especially older Fortran codes) and it is often not possible to make desired alterations with confidence; it is more efficient for a researcher to begin coding from scratch. As a result, graduate students and research groups around the world duplicate efforts and re-invent technology and methodologies. This is impacting the current rate of research at universities and it will become a greater impediment as we attempt to solve complex multi-disciplinary problems. The plight of the graduate student is illustrated in Figure 1.

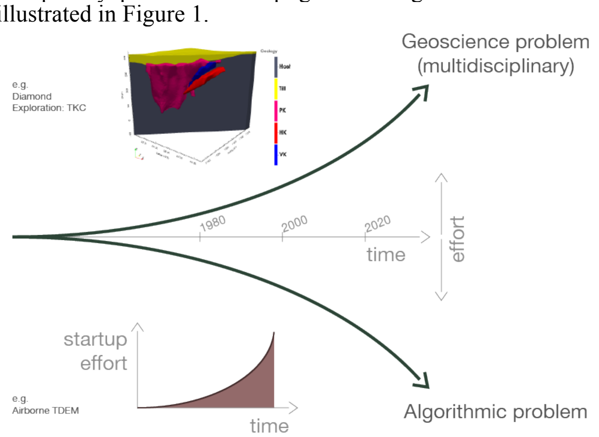

Figure 1. The plight of graduate student in geoscience.

The horizontal axis reflects time in calendar years. The lower solid line represents the effort required to solve an algorithmic problem; this has increased as the years have progressed and the problems have become harder. In 1980 students could obtain a PhD for solving a 1D forward problem in electromagnetics. To obtain a PhD in 2018 the representative state-of-the art forward modelling might be a 3D time domain EM problem with dispersive physical properties. There are far more complexities in solving the 3D problem than the 1D, and despite advances in computing power, it will take much longer for a student in 2018 to get his/her computation algorithms developed than an equivalent student in 1980. The effort in solving the geoscience problem (and hence the need for the algorithmic software) is also increasing. Data sets are higher quality (and therefore have more information to be extracted) and also the application they are being used to address require multi-physics and is generally multi-disciplinary. The vertical difference between the top and bottom curves is the total effort needed to solve a problem. This total effort continues to diverge and it is becoming impractical or impossible for one person to cross the gap. For those who attempt this, it takes longer to finish, and/or their research is marginalized because time and funding are limited. This is detrimental to the researcher and the development of geophysics as a whole. It is a paradigm that needs to be changed and working in a collaborative, open-source framework is a key ingredient for achieving this.

OPEN SOURCE DEVELOPMENT

As the curves continue to diverge, it becomes clear that solving the next generation of integrated geoscience problems requires effective collaborations between researchers with different backgrounds and skill sets. Other research communities, for example Astropy in astronomy (Astropy Collaboration, 2001), Scikit-learn in machine learning (Pederegosa et al., 2011), and SciPy in numerical computing (Jones et al., 2001), have embraced the open-source approach for collaboration and cooperation on the development software resources for research. Beyond simply making code publically available, scaling a collaborative effort beyond a handful of close colleagues requires the adoption of best practices such as version control, automated testing, comprehensive user and developer documentation, peer review of code, and issue tracking. Modern tools such as GitHub, Read the Docs, and Travis CI make this possible, accessible and scalable beyond close collaborators. One of the key factors in fostering large communities of contributors is the choice of a permissive license with clear terms, such as the MIT license (https://opensource.org/osd). Such licenses allows users to adapt the code for academic or industry-oriented purposes, and enables contributors from companies and universities to collaborate on a common set of tools.

The adoption of open source practices for software development has the potential to greatly accelerate research in geophysics. Often new avenues of questions can be explored with a few tweaks or combinations of existing pieces, and when the building-blocks are open-source, they can be re-used or adapted, rather than having to be created from scratch. For students and researchers exploring a problem unfamiliar to them, having access to source-code, which captures the details necessary for a successful implementation, can be an invaluable learning tool. In particular, non-linear inverse problems, such as those that we encounter in EM, require tuning parameters to be set and heuristics, such as a beta-cooling schedule, to be defined. These are critical aspects to the success of the algorithm, but are generally obfuscated in black-box codes. Furthermore, the Jupyter notebook is one of the tools that is enabling workflows and analyses to be widely communicated to other scientists and to the general public, thus making it easier to distribute results in a manner that is readily reproducible (Perez and Granger, 2015). These features speak to an individual getting up-and-running with software to address a research question.

Open source practices facilitate peer-review and collaboration at the level of the implementation. The re-use and adaptation of code requires that software be human-readable, and developed in a modular, flexible manner. Modern computational languages such as Python, Julia, and R have simplified code-readability, and when development is coupled with a peer-review of the code, ambiguities or unnecessary complexity in the implementation can be reduced.

In addition to well established software packages such as ModEM and the Aarhus Workbench and existing open source packages such as MARE2DEM, there is a growing ecosystem of open-source tools available for solving problems in electromagnetic geophysics (Auken et al., 2015). Our efforts have been focused on the development of SimPEG; the larger ecosystem also contains several other notable packages including empymod, fatiando, jInv, and pyGIMLi (Werthmüller, 2017; Uieda et al., 2013; Ruthotto, 2017; Rücker, 2017). Each of these packages differ in objectives, capabilities, structure, interactivity, license, and coding language (commonly Python and Julia). Our work at UBC, and the remaining focus in our paper, is the development and use of SimPEG for forward modelling and inversion of EM data.

SimPEG is an open source framework and set of tools for simulation and gradient-based parameter estimation in geophysics. It includes finite volume simulations and inversion routines for a variety of geophysical applications including Potential Fields, Vadose Zone flow, DC Resistivity, and Electromagnetics. Simulations may be performed on several different mesh types, including cylindrically symmetric meshes, 3D tensor meshes and OcTree meshes. A staggered-grid, finite volume solution approach is used to solve the quasi-static Maxwell's equations in both the time and frequency. The fields and fluxes computed everywhere in the simulation domain are readily accessible so that they can be easily visualized and explored. Such simulations and visualizations have proved valuable in the context of geoscience education (Oldenburg et al., 2017) and can be a useful tool for understanding the physical processes that contribute to the data we observe.

EM SIMULATIONS

Many of the research questions we encounter start with an exploration of the governing physics. To demonstrate the value of this we choose a canonical model of a conductive plate in a resistive earth and perform an airborne TDEM simulation. The setup is shown in Figure 2. The source loop has 1m radius and is located 30m above the surface; the source waveform is a step-off and the data are db_z/dt measured at a receiver that is coaxial with the transmitter. The simulation is performed in 3D on a tensor mesh. A 100m thick conductive plate (10 Ωm) is embedded 50 m below the surface and the background conductivity is 1000 Ωm as shown in Figure 2.

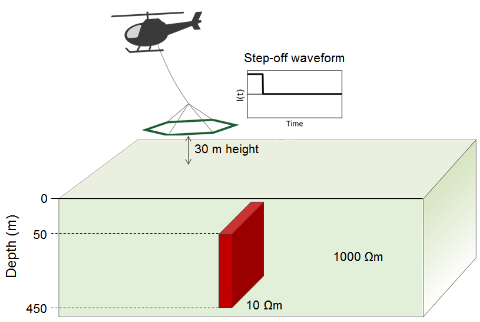

Figure 2. Airborne time domain EM survey over a vertical conductive plate.

In addition to the data values, a question that always arises is "What is the relative contributions to the data from the inductive and galvanic currents?" Insight can be obtained about this, and about the EM induction process in general, by viewing the currents and fields as a function of time. For example, in Figure 3 we show snapshots of the currents at two times. At early time (0.01 ms) the currents are primarily those that are channelled into the conductor (these are often referred to as galvanic currents). At later time (0.13 ms) induced vortex currents dominate. This transition between the two dominant current modes and their interactions can be further investigated as one toggles through the slider widgets (https://github.com/simpeg-research/heagy_2018_aem).

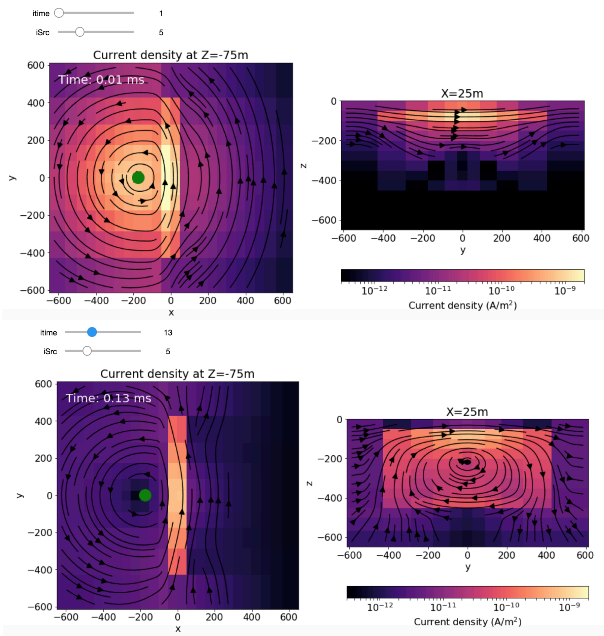

Figure 3. Snap-shots of Jupyter notebook widget which shows the current density at 0.01 ms and 0.13ms after shut-off. Depth slices (left) and cross-sections (right). The green dot indicates the source location.

EM INVERSIONS

Inversions are more complex than forward modelling; a framework and workflow are required to extract meaningful information from EM data. We use a deterministic Tikhonov style approach, where an objective function consisting of a data misfit and a regularization functional is minimized. Within this framework there are many specific formulations that a researcher might explore. They pertain to definition of: the data (linear, log, some function of observed fields); the misfit function (least squares, Ekblom); the model parameters (1D, 3D, linear, log, parametric); and the regularization function (what norm, what reference model, incorporation of a prior knowledge). In addition, there are many ways to solve the optimization problem, even with the spectrum of gradient-based methods, and there is always the issue of selecting an appropriate Tikhonov tradeoff parameter. In general the researcher will want to carry out multiple inversions using different definitions of important elements; the open source software should be tailored to allow this to happen as seamlessly as possible. In our example we carry out a joint inversion to recover electrical conductivity using data from different types of EM surveys. Independent software packages to invert each data type have been previously written, however, effective joint inversion approaches often prove problematic or complicated to implement, as each algorithm employs different assumptions and standards. In a modular framework however, each datum has its own forward modelling and sensitivity calculation so working with multiple surveys is straightforward.

For example, we consider a layered structure that includes a shallow resistor (100 Ωm) and deep conductor (1 Ωm) embedded in half-sace (10 ohm-m) as shown in Figure 4. To resolve this conductivity structure, three different EM surveys are selected. Each uses a loop transmitter and a receiver that measures db_z/dt. A Resolve (airborne FDEM), Geotem (airborne TDEM), and NanoTEM (ground TDEM), are illustrated in Figure 6. Each system has different sensitivity to the subsurface conductivity structure based upon its frequency-band and height above the surface. The frequency band of the Resolve system is 400Hz-130kHz, whereas the base frequency of the Geotem system is usually 30 (or 25 Hz) and hence even with very early time measurements (e.g. couple micro-s), the Geotem system will have less sensitivity to the near surface compared to Resolve. However, for a deep conductor Geotem will have greater sensitivity due to its lower frequency band compared to Resolve. Conversely, ground NanoTEM loops, which are much closer to the targets compared to two other airborne systems, will show greater sensitivity at the near surface. Jointly inverting all three systems together should therefore be beneficial in resolving the layered conductivity structure.

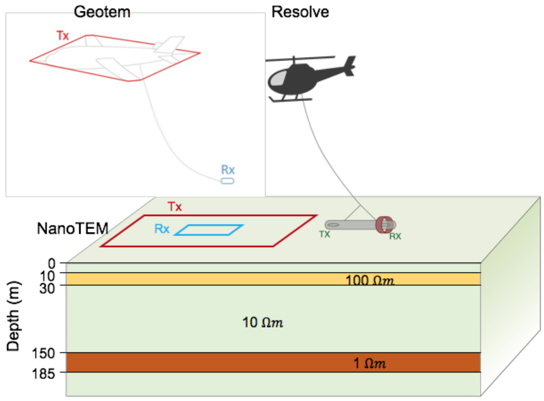

Figure 4. Three different EM systems: Resolve (airborne FDEM), Geotem (airborne TDEM), NanoTEM (ground TDEM) over a layered-earth including shallow resistor and deep conductor.

Synthetic data sets are generated with the model shown in Figure 4 using typical system specifications for Resolve, Geotem, and NanoTEM. Each of the three data sets is first inverted individually. The initial model is set to 10 ohm-m half-space and 5% percent standard deviation is assigned for the uncertainty. The recovered resistivity model from the three separate inversions have the following features (Figure 5). Each recognizes a shallow resistor but the value of the resistivity is underestimated compared to the true value. This is characteristic of all inductive source surveys. Of the three, the ground NanoTEM resistivity (red) shows the best match with the shallow resistor; this is due to its larger sensitivity at the near surface because the system is on the ground. Nano TEM does not see the deep conductor and neither does Resolve. However, it is well recovered in the Geotem resistivity (green). Joint inversion of all three EM data sets recovers both shallow resistor and deep conductor (purple). All four inversions used the same uncertainty and reached the root-mean-square target misfit of unity. Further exploration of the inversions, as well as reproducing the results shown here, can be achieved by downloading the Jupyter notebook (https://github.com/simpeg-research/oldenburg_2018_aem).

CONCLUSIONS

Unquestionably, the next generation of geoscience problems to be solved will be complex and multidisciplinary and, for electromagnetics, will likely involve multiple types of EM surveys coupled with process modelling and machine learning. These advances will be unlikely to be completed effectively by a single researcher or even a single research group. Collaboration, reproducibility and access to previous research achievements will be key. In this paper we presented arguments for adopting an open source model for geophysics and provided an introduction to some of the initiatives being pursued. With the use of Jupyter Notebooks, we illustrate how simulation and inversion research results can be shared and reproduced. We are not alone in our enthusiasm for this paradigm change in geophysical research. Even within the realm of electromagnetic geophysics, there are many groups worldwide that have the same aspirations. Our challenge therefore, is not only to enable continued development within each group, but to promote collaboration between groups so that the geophysical open source ecosystem grows coherently and sustainably.

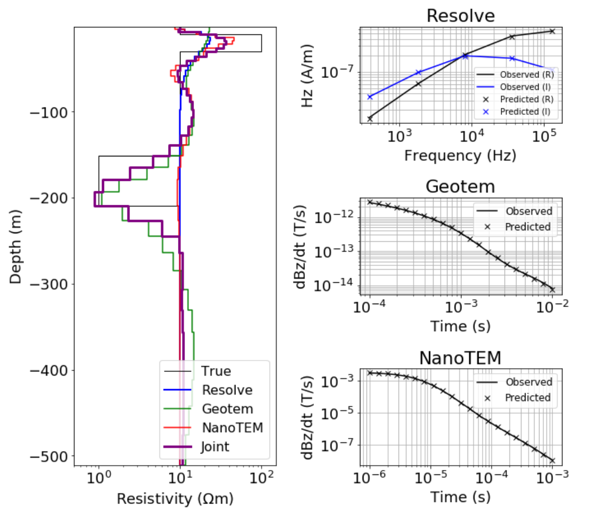

Figure 5. Joint inversion of three different EM data sets: (1) Resolve (airborne FDEM), (2) Geotem (airborne TDEM), and (3) NanoTEM (ground TDEM) in 1D space. Left: true (black) and recovered 1D conductivity models from each of separate EM inversions: Resolve (blue), Geotem (green), NanoTEM (red) and the joint inversion (purple). Right: observed and predicted data for the three EM data sets.

REFERENCES

Auken, E., Christiansen, A. V., Fiandaca, G., Schamper, C., Behroozmand, A. A., Binley, A., Nielsen, E., Effersø, F., Christensen, N. B., Sørensen, K. I., Foged, N., and Vignoli, G.,2015, An overview of a highly versatile forward and stable inverse algorithm for airborne, ground-based and borehole electromagnetic and electric data, Exploration Geophysics, 2015, 223-23

Cockett, R., Kang, S., Heagy, L., Pidlisecky, A., and Oldenburg, D., 2015, SimPEG: An open source framework for simulation and gradient based parameter estimation in geophysical applications: Computers & Geosciences, 85, 142-154. https://doi.org/10.1016/j.cageo.2015.09.015

Heagy, L., Cockett, R., Kang, S., Rosenkjaer, G., and Oldenburg, D., 2017, A framework for simulation and inversion in electromagnetics: Computers & Geosciences, 107, 1-19. https://doi.org/10.1016/j.cageo.2017.06.018

Oldenburg D., Heagy, L., and Kang, S., 2017, Geophysical Electromagnetics: Fundamentals and Applications: Society of Exploration Geophysics 2017 Distinguished Instructor Short Course.

Ruthotto L., Treister E., Haber E., 2017, jInv a flexible Julia Package for PDE Parameter Estimation, SIAM J. Sci. Comp., 39(5), S702–S722. https://doi.org/10.1137/16M1081063

Perez, F. and Granger, B., 2015, Project Jupyter: Computational Narratives as the Engine of Collaborative Data Science: Grant Proposal accepted by the Helmsley Trust, the Gordon and Betty Moore Foundation and the Alfred P. Sloan Foundation, http://archive.ipython.org/JupyterGrantNarrative-2015.pdf

Rücker, C., Günther, T., Wagner, F.M., 2017. pyGIMLi: An open-source library for modelling and inversion in geophysics, Computers and Geosciences, 109, 106-123, doi: 10.1016/j.cageo.2017.07.011

Uieda, L., V. C. Oliveira Jr, and V. C. F. Barbosa, 2013, Modeling the Earth with Fatiando a Terra, Proceedings of the 12th Python in Science Conference, pp. 91 - 98, http://conference.scipy.org/proceedings/scipy2013/pdfs/uieda.pdf

Werthmüller D., 2017. An open-source full 3D electromagnetic modeler for 1D VTI media in Python: empymod, GEOPHYSICS, 82(6), WB9-WB19. https://doi.org/10.1190/geo2016-0626.1