

# A framework for simulation and inversion in electromagnetics

Lindsey J. Heagy<sup>1,1</sup>, Rowan Cockett<sup>1</sup>, Seogi Kang<sup>1</sup>, Gudni K. Rosenkjaer<sup>1</sup>, Douglas W. Oldenburg<sup>1</sup>

<sup>a</sup>604-836-2715, lheagy@eos.ubc.ca

<sup>b</sup>Geophysical Inversion Facility, University of British Columbia

---

## Abstract

Simulations and inversions of electromagnetic geophysical data are paramount for discerning meaningful information about the subsurface from these data. Depending on the nature of the source electromagnetic experiments may be classified as time-domain or frequency-domain. Multiple heterogeneous and sometimes anisotropic physical properties, including electrical conductivity and magnetic permeability, may need be considered in a simulation. Depending on what one wants to accomplish in an inversion, the parameters which one inverts for may be a voxel-based description of the earth or some parametric representation that must be mapped onto a simulation mesh. Each of these permutations of the electromagnetic problem has implications in a numerical implementation of the forward simulation as well as in the computation of the sensitivities, which are required when considering gradient-based inversions. This paper proposes a framework for organizing and implementing electromagnetic simulations and gradient-based inversions in a modular, extensible fashion. We take an object-oriented approach for defining and organizing each of the necessary elements in an electromagnetic simulation, including: the physical properties, sources, formulation of the discrete problem to be solved, the resulting fields and fluxes, and receivers used to sample to the electromagnetic responses. A corresponding implementation is provided as part of the open source simulation and parameter estimation project SimPEG (<http://simpeg.xyz>). The application of the framework is demonstrated through two synthetic examples and one field example. The first example shows the application of the common framework for 1D time domain and frequency domain inversions. The second is a field example that demonstrates a 1D inversion of electromagnetic data collected over the Bookpurnong Irrigation District in Australia. The final example shows how the modular implementation is used to compute the sensitivity for a parametric model where a transmitter is positioned inside a steel cased well.

**Keywords:** Geophysics, Numerical Modelling, Finite Volume, Sensitivities, Object Oriented

---

## <sup>1</sup> 1. Introduction

The field of electromagnetic (EM) geophysics encompasses a diverse suite of problems with applications across mineral and resource exploration, environmental studies and geotechnical engineering. EM problems can be formulated in the time or frequency domain. Sources can be grounded electric sources or inductive loops driven by time-harmonic or transient currents, or natural, plane wave sources, as in the case of the magnetotelluric method. The

6 physical properties of relevance include electrical conductivity, magnetic permeability, and electric permittivity. These  
7 may be isotropic, anisotropic, and also frequency dependent. Working with electromagnetic data to discern information  
8 about subsurface physical properties requires that we have numerical tools for carrying out forward simulations  
9 and inversions that are capable of handling each of these permutations.

10 The goal of the forward simulation is to solve a specific set of Maxwell's equations and obtain a prediction the  
11 EM responses. Numerical simulations using a staggered grid discretization (?), have been extensively studied in their  
12 application for finite difference, finite volume and finite element approaches (c.f. ??), with many such implementations  
13 being optimized for efficient computations for the context in which they are being applied (????).

14 Finding a model of the earth that is consistent with the observed data and prior geologic knowledge is the ‘inverse  
15 problem’. It presupposes that we have a means of solving the forward problem. The inverse problem is generally  
16 solved by minimizing an objective function that consists of a data misfit and regularization, with a trade-off parameter  
17 controlling their relative contributions. (???). Deterministic, gradient-based approaches to the inverse problem are  
18 commonplace in EM inversions. Relevance of the recovered inversion model is increased by incorporating *a priori*  
19 geologic information and assumptions. This can be accomplished through, the regularization term (??) or param-  
20 eterizing the inversion model (???). Multiple data sets may be considered through cooperative or joint inversions  
21 (??).

22 Each of these advances relies on a workflow and associated software implementation. Unfortunately, each soft-  
23 ware implementation is typically developed as a stand-alone solution. As a result, these advances are not readily  
24 interoperable with regard to concepts, terminology, notations *and* software.

25 The advancement of EM geophysical techniques and the expansion of their application requires a flexible set of  
26 concepts and tools that are organized in a framework so that researchers can more readily experiment with, and ex-  
27 plore, new ideas. For example, if we consider research questions within the growing application of EM for reservoir  
28 characterization and monitoring in settings with steel cased wells (cf. ??????), the numerical tools employed must  
29 enable investigation into factors such as the impact of variable magnetic permeability (??) and casing integrity (?)  
30 on electromagnetic signals. Various modelling approaches in both time and frequency domain simulations are being  
31 explored, these include employing highly-refined meshes (?), using cylindrical symmetry (?) or approximating the  
32 casing on a coarse-scale (?), possibly 3D anisotropic approximations (?). Beyond forward simulations that predict  
33 EM responses, to enable the interpretation of field data with these tools requires that machinery to address the inverse  
34 problem and experiment with approaches for constrained and/or time lapse inversions be in place (??). Typically,  
35 addressing each of these complexities would require a custom implementations, particularly for the frequency domain  
36 and time domain simulations, although aspects, such as physical properties, are common to both. Inconsistencies be-  
37 tween implementations and the need to implement a custom solution for each type of EM method under consideration  
38 presents a significant barrier to a researcher’s ability to experiment with and extend ideas.

39 Building from the body of work on EM geophysical simulations and inversions, the aim of our efforts is to identify  
40 a common, modular framework suitable across the suite of electromagnetic problems. This conceptual organization

41 has been tested and developed through a numerical implementation. The implementation is modular in design with the  
42 expressed goal of affording researchers the ability to rapidly adjust, interchange, and extend elements. By developing  
43 the software in the open, we also aim to promote an open dialog on approaches for solving forward and inverse  
44 problems in EM geophysics.

45 The implementation we describe for EM forward and inverse problems extends a general framework for geo-  
46 physical simulation and gradient based inverse problems, called SIMPEG (?). The implementation of SIMPEG is  
47 open-source, written in Python and has dependencies on the standard numerical computing packages NumPy, SciPy,  
48 and Matplotlib (???). The contribution described in this paper is the implementation of the physics engine for prob-  
49 lems in electromagnetics, including the forward simulation and calculation of the sensitivities. Building within the  
50 SIMPEG ecosystem has expedited the development process and allowed developments to be made in tandem with  
51 other applications (<http://simpeg.xyz>). SIMPEGEM aspires to follow best practices in terms of documentation, testing,  
52 continuous integration using the publically available services Sphinx, Travis CI, and Coveralls (???). As of the writ-  
53 ing of this paper, when any line of code is changed in the open source repository, over 3 hours of testing is completed;  
54 documentation and examples are also tested and automatically updated (<http://docs.simpeg.xyz>). We hope these prac-  
55 tices encourage the growth of a community and collaborative, reproducible software development in the field of EM  
56 geophysics.

57 The paper is organized as follows. To provide context for the structure and implementation of SIMPEGEM, we  
58 begin with a brief overview of the SIMPEG inversion framework as well as the governing equations for electromag-  
59 netics in Section ???. In Section ??, we discuss the motivating factors for the EM framework, and in Section ??, we  
60 discuss the framework and implementation of the forward simulation and calculation of sensitivities in SIMPEGEM. We  
61 demonstrate the implementation with two synthetic examples and one field example in Section ???. The first example  
62 shows the similarities between the time and frequency implementations for a 1D inversion. In the second example,  
63 we invert field data from the Bookpurnong Irrigation district in Australia. The final example demonstrates how the  
64 modular implementation is used to compute the sensitivity for a parametric model of a block in a layered space where  
65 a transmitter is positioned inside a steel cased well.

## 66 **2. Background**

67 We are focused on geophysical inverse problems in electromagnetics (EM), that is, given EM data, we want to find  
68 a model of the earth that explains those data and satisfies prior assumptions about the geologic setting. We follow the  
69 SIMPEG framework, shown in Figure ??, which takes a gradient- based approach to the inverse problem (?). Inputs  
70 to the inversion are the data and associated uncertainties, a description of the governing equations, as well as prior  
71 knowledge and assumptions about the model. With these defined, the SIMPEG framework accomplishes two main  
72 objectives:

- 73        1. the ability to forward simulate data and compute sensitivities (Forward Simulation - outlined in green in Figure ??),  
 74  
 75        2. the ability to assess and update the model in an inversion (Inversion Elements and Inversion as Optimization -  
 76        outlined in red in Figure ??).

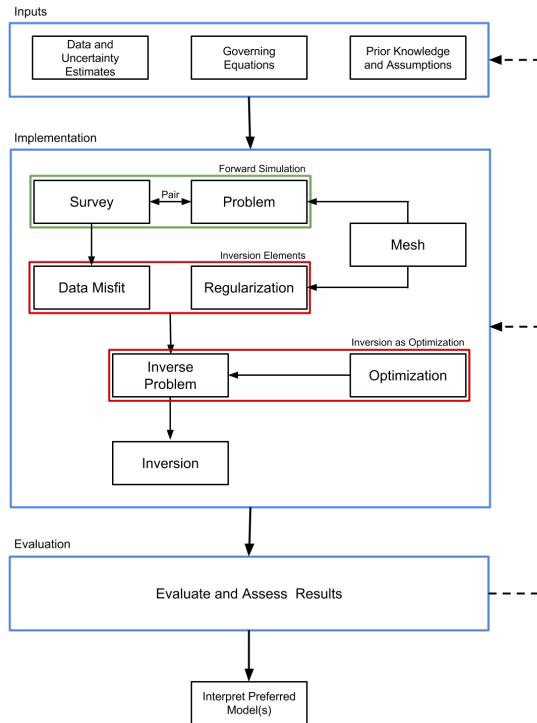


Figure 1: Inversion approach using the SimPEG framework. Adapted from ?

77        The implementation of the framework is organized into the self-contained modules shown in Figure ??; each mod-  
 78       ule is defined as a base- class within SimPEG. The **Mesh** provides the discretization and numerical operators. These  
 79       are leveraged by the **Problem**, which is the numerical physics engine; the **Problem** computes fields and fluxes when  
 80       provided a model and **Sources**. The **Sources** are specified in the **Survey**, as are the **Receivers**. The **Receivers**  
 81       take the **Fields** computed by the **Problem** and evaluate them at the receiver locations to create predicted data. Each  
 82       action taken to compute data, when provided a model, has an associated derivative with respect to the model; these  
 83       components are assembled to create the sensitivity. Having the ability to compute both predicted data and sensitivities  
 84       accomplishes the first objective.

85        To accomplish the second objective of assessing and updating the model in the context of the data and our as-  
 86       sumptions, we consider a gradient-based approach to the inversion. For this, we specify an objective function which  
 87       generally consists of a **DataMisfit** and **Regularization**. The **DataMisfit** is a metric that evaluates the agreement  
 88       between the observed and predicted data, while the **Regularization** is a metric constructed to assess the model's

89 agreement with assumptions and prior knowledge. These are combined with a trade-off parameter to form a math-  
 90 ematical statement of the `InvProblem`, an optimization problem. The machinery to update the model is provided  
 91 by the `Optimization`. An `Inversion` brings all of the elements together and dispatches `Directives` for solving  
 92 the `InvProblem`. These `Directives` are instructions that capture the heuristics for solving the inverse problem; for  
 93 example, specifying a target misfit that, once reached, terminates the inversion, or using a beta-cooling schedule that  
 94 updates the value of the trade-off parameter between the `DataMisfit` and `Regularization` (cf. ?? and references  
 95 within).

96 The output of this process is a model that must be assessed and evaluated prior to interpretation; the entire process  
 97 requires iteration by a human, where underlying assumptions and parameter choices are re-evaluated and challenged.  
 98 Be it in resource exploration, characterization or development; environmental remediation or monitoring; or geotech-  
 99 nical applications – the goal of this model is to aid and inform a complex decision.

100 Here we note that the inversion framework described above is agnostic to the type of forward simulation employed,  
 101 provided the machinery to solve the forward simulation and compute sensitivities is implemented. Specific to the  
 102 EM problem, we require this machinery for Maxwell's equations. As such, we focus our attention on the `Forward`  
 103 `Simulation` portion of the implementation for the EM problem and refer the reader to ? and ? for a more complete  
 104 discussion of inversions.

### 105 2.1. Governing Equations

106 Maxwell's equations are the governing equations of electromagnetic problems. They are a set of coupled par-  
 107 tial differential equations that connect electric and magnetic fields and fluxes. We consider the quasi-static regime,  
 108 ignoring the contribution of displacement current (???)<sup>1</sup>

We begin by considering the first order quasi-static EM problem in time,

$$\vec{\nabla} \times \vec{e} + \frac{\partial \vec{b}}{\partial t} = \vec{s}_m$$

$$\vec{\nabla} \times \vec{h} - \vec{j} = \vec{s}_e \quad (1)$$

109 where  $\vec{e}$ ,  $\vec{h}$  are the electric and magnetic fields,  $\vec{b}$  is the magnetic flux density,  $\vec{j}$  is the current density, and  $\vec{s}_m$ ,  $\vec{s}_e$   
 110 are the magnetic and electric source terms.  $\vec{s}_e$  is a physical, electric current density, while  $\vec{s}_m$  is “magnetic current  
 111 density”. Although  $\vec{s}_m$  is unphysical, as continuity of the magnetic current density would require magnetic monopoles,  
 112 the definition of a magnetic source term can be a useful construct, as we will later demonstrate in Section ?? (see also  
 113 ?).

---

<sup>1</sup>In most geophysical electromagnetic surveys, low frequencies or late-time measurements are employed. In these scenarios  $\sigma \gg \epsilon_0\omega$  (eg. conductivities are typically less than 1S/m,  $\epsilon_0 = 8.85 \times 10^{-12} F/m$  and frequencies considered are generally less than  $10^5$  Hz), so displacement current can safely be ignored.

By applying the Fourier Transform (using the  $e^{i\omega t}$  convention), we can write Maxwell's equations in the frequency domain:

$$\begin{aligned}\vec{\nabla} \times \vec{E} + i\omega \vec{B} &= \vec{S}_m \\ \vec{\nabla} \times \vec{H} - \vec{J} &= \vec{S}_e\end{aligned}\tag{2}$$

where we use capital letters to denote frequency domain variables. The fields and fluxes are related through the physical properties: electrical conductivity  $\sigma$ , and magnetic permeability  $\mu$ , as described by the constitutive relations

$$\begin{aligned}\vec{J} &= \sigma \vec{E} \\ \vec{B} &= \mu \vec{H}\end{aligned}\tag{3}$$

114 The physical properties,  $\sigma$  and  $\mu$  are generally distributed and heterogeneous. For isotropic materials,  $\sigma$  and  $\mu$  are  
115 scalars, while for anisotropic materials they are  $3 \times 3$  symmetric positive definite tensors. The same constitutive  
116 relations can be applied in the time domain provided that the physical properties,  $\sigma, \mu$  are not frequency-dependent.

117 In an EM geophysical survey, the sources provide the input energy to excite responses that depend on the physical  
118 property distribution in the earth. These responses, electric and magnetic fields and fluxes, are sampled by receivers to give the observed data. The simulation of Maxwell's equations may be conducted in either the time or frequency domain,  
119 depending on the nature of the source; harmonic waveforms are naturally represented in the frequency domain,  
120 while transient waveforms are better described in the time domain.

The aim of the inverse problem is to find a model,  $\mathbf{m}$  (which may be a voxel-based or a parametric representation) that is consistent with observed data and with prior knowledge and assumptions about the model. Addressing the inverse problem using a gradient-based approach requires two abilities of the forward simulation: (1) the ability to compute predicted data given a model

$$\mathbf{d}_{\text{pred}} = \mathcal{F}[\mathbf{m}]\tag{4}$$

and (2) the ability to compute or access the sensitivity, given by

$$\mathbf{J}[\mathbf{m}] = \frac{d\mathcal{F}[\mathbf{m}]}{d\mathbf{m}}.\tag{5}$$

122 To employ second order optimization techniques, we also require the adjoint of the sensitivity,  $\mathbf{J}^\top$ . These two elements,  
123 when combined into the SIMPEG framework, enable data to be simulated and gradient-based inversions to be run. As  
124 such, this work benefits from other peoples' contributions to the underlying inversion machinery, including: discrete  
125 operators on a variety of meshes, model parameterizations, regularizations, optimizations, and inversion directives  
126 (?).

### 127 3. Motivation

128 The motivation for the development of this framework is that it be a resource for researchers in the field of  
129 electromagnetic geophysics. To best serve this goal, we require a framework that is modular and extensible in order to

130 enable exploration of ideas. An associated numerical implementation is essential for this work to be tested and acted  
131 upon. As such, we provide a tested, documented, fully open-source software implementation of the framework (under  
132 the permissive MIT license).

133 Specific to the EM problem, we require the implementation of Maxwell's equations in both the time domain  
134 and frequency domain. The implementation must allow for variable electrical conductivity and magnetic perme-  
135 ability, anisotropic physical properties; various model parameterizations of the physical properties (e.g. voxel log-  
136 conductivity or parametric representations); a range of sources including wires, dipoles, natural sources; variable  
137 receiver types; variable formulations of Maxwell's equations; solution approaches such as using a primary-secondary  
138 formulation; and the flexibility to work with and move between a variety of meshes such as tensor, cylindrically sym-  
139 metric, curvilinear, and octree discretizations. Furthermore, the sensitivity computation must be flexible enough to be  
140 computed for any sensible combination of these approaches. In the following section, we will outline the framework  
141 we have used to organize and implement these ideas.

#### 142 **4. Simulation Framework**

143 The aim of the forward simulation is to compute predicted data,  $\mathbf{d}_{\text{pred}}$ , when provided with an inversion model<sup>2</sup>,  
144 **m** and Sources. SIMPEGEM contains implementations for both time domain (TDEM) and frequency domain (FDEM)  
145 simulations, allowing data from commonly used EM methods to be simulated.

146 The framework we follow to perform the forward simulation is shown in Figure ??; it consists of two overarching  
147 categories:

- 148 1. the Problem, which is the implementation of the governing equations,
- 149 2. the Survey, which provides the source(s) to excite the system as well as the receivers to samples the fields and  
150 produce predicted data at receiver locations.

151 Here, we provide a brief overview of each of the components, and discuss them in more detail in the sections that  
152 follow.

153 The ‘engine’ of the forward simulation is the physics; it contains the machinery to solve the system of equations  
154 for EM fields and fluxes in the simulation domain when provided with a description of the physical properties and  
155 sources. In general, the physics engine may be an analytic or numeric implementation of Maxwell's equations. Here,  
156 we focus our attention on the numerical implementation using a standard staggered-grid finite volume approach,  
157 requiring that the physical properties, fields, fluxes and sources be defined on a mesh (cf. ???). We discretize fields  
158 on edges, fluxes on faces and physical properties in cell centers, as shown in Figure ???. To construct the necessary  
159 differential and averaging operators, we leverage the Mesh class within SIMPEG (??).

---

<sup>2</sup>We use the term *inversion model* to describe a parameterized representation of the earth (e.g. voxel-based or parametric), even if the model is solely used for forward modelling, its form sets the context for the inverse problem and the parameter-space that is to be explored.

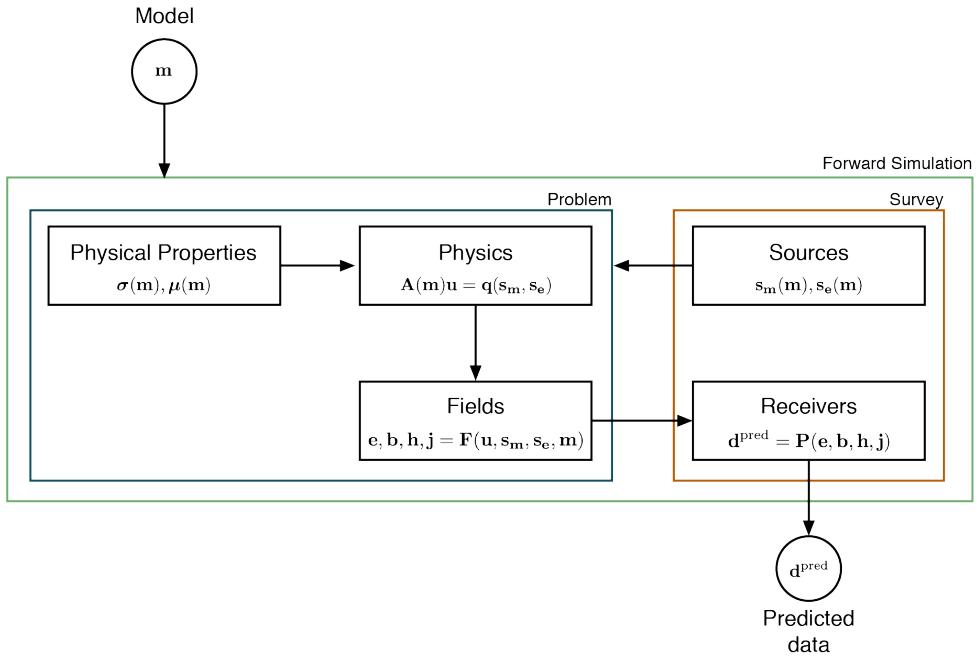


Figure 2: Forward simulation framework.

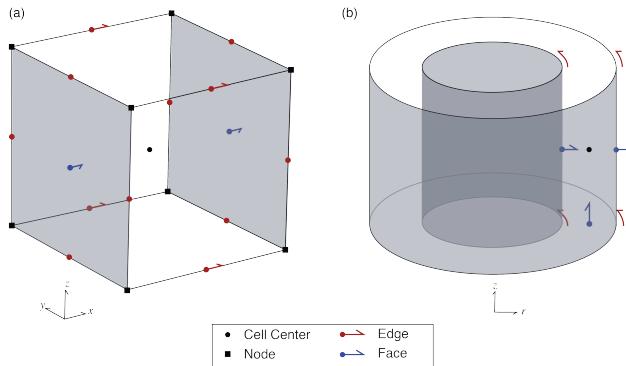


Figure 3: Location of variables in the finite volume implementation for both a unit cell in (a) cartesian and (b) cylindrical coordinates (after ?)

To compute electromagnetic responses, the forward simulation requires the definition of a physical property model describing the electrical conductivity ( $\sigma$ ) and magnetic permeability ( $\mu$ ) on the simulation mesh, as well as discrete representations of the sources used to excite EM responses ( $s_e, s_m$ ). Often in solving an inverse problem, the model which one inverts for (the vector  $\mathbf{m}$ ), is some discrete representation of the earth that is decoupled from the physical property model. This decoupling requires the definition of a Mapping capable of translating  $\mathbf{m}$  to physical properties on the simulation mesh. For instance, if the inversion model is chosen to be log-conductivity, an exponential mapping is required to obtain electrical conductivity (i.e.  $\sigma = \mathcal{M}(\mathbf{m})$ ). To support this abstraction, SimPEG provides a number

167 of extensible Mapping classes (??).

With both the physical property model and the source specified, we define and solve the physics, a Maxwell system of the form

$$\mathbf{A}(\mathbf{m})\mathbf{u} = \mathbf{q}(\mathbf{s}_m, \mathbf{s}_e), \quad (6)$$

for an electric or magnetic field or flux. Here,  $\mathbf{A}$  is the system matrix that may eliminate a field or flux to obtain a system in a single field or flux,  $\mathbf{u}$ , the solution vector. Correspondingly, the vector  $\mathbf{q}$  is the second order right-hand-side. Note, if there are necessary manipulations to make equation ?? easier to solve numerically (e.g. symmetry) we can add these here; doing so has no effect on the derivative. The remaining fields and fluxes can be computed from  $\mathbf{u}$  anywhere in the simulation domain, through an operation of the form

$$\mathbf{f} = \mathbf{F}(\mathbf{u}(\mathbf{m}), \mathbf{s}_e(\mathbf{m}), \mathbf{s}_m(\mathbf{m}), \mathbf{m}) \quad (7)$$

where  $\mathbf{f}$  is conceptually a vector of *all* of the fields and fluxes (i.e.  $\mathbf{e}$ ,  $\mathbf{b}$ ,  $\mathbf{h}$  and  $\mathbf{j}$ ). This vector is never stored in the implementation, instead the fields are computed on demand through the subset of stored solution vectors ( $\mathbf{u}$ ). From the computed fields ( $\mathbf{f}$ ), predicted data are created by the Receivers through an operation of the form

$$\mathbf{d}_{\text{pred}} = \mathbf{P}(\mathbf{f}) \quad (8)$$

168 In the simplest case, the action of  $\mathbf{P}$  selects the component of interest and interpolates the fields to the receiver  
169 locations, more involved cases could include the computation of ratios of fields, as is the case for impedance or tipper  
170 data. Obtaining predicted data from the framework concludes the forward simulation.

171 The same framework is employed for both time domain (TDEM) and frequency domain (FDEM) implementations  
172 within SIMPEGEM. In the case of the FDEM implementation, the matrix  $\mathbf{A}(\mathbf{m})$  and the solution vector  $\mathbf{u}$  represent all  
173 frequencies. As these frequencies are independent (i.e. a block diagonal matrix, ), each frequency can be solved  
174 independently. In the TDEM code, the matrix  $\mathbf{A}(\mathbf{m})$  and the solution vector  $\mathbf{u}$  represent all timesteps (??) and take the  
175 form of a lower triangular block matrix (bidiagonal in the case of Backward Euler, ) meaning the computation of  
176 each time-step depends on previous time-steps. The form of these matrices will be discussed further in the Physics  
177 section (Section ??)

178 To perform a gradient-based inversion, we require the sensitivity of the data with respect to the inversion model,  
179 thus, each action taken to calculate data from the model must have an associated derivative. The full sensitivity is  
180 a dense matrix and is expensive to form and store, but when the optimization problem is solved using an iterative  
181 optimization approach, it does not need to be explicitly formed; all that is required are products and adjoint-products  
182 with a vector. We treat this using a modular approach so that individual elements of the framework can be rapidly  
183 interchanged or extended. The process we follow to compute matrix-vector products with the sensitivity is shown with  
184 red arrows in Figure ?? (b). The sensitivity-vector product  $\mathbf{J}\mathbf{v}$  is built in stages by taking matrix vector products with  
185 the relevant derivatives in each module, starting with the derivative of the physical property with respect to the model.  
186 The product with the adjoint is similarly shown in Figure ?? (c) starting with the adjoint of the receiver operation.

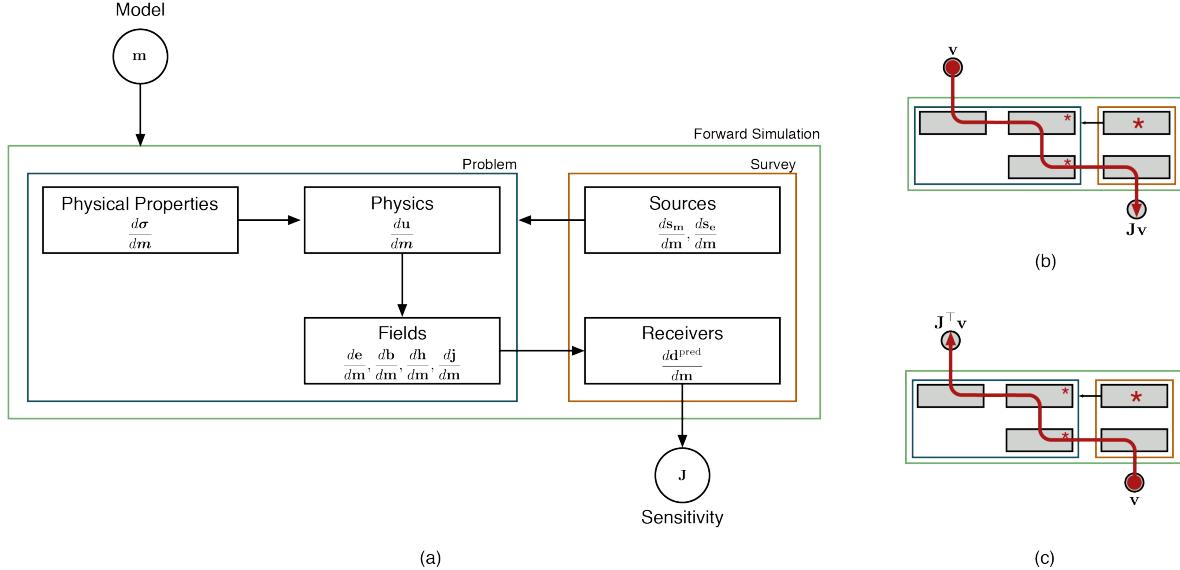


Figure 4: (a) Contributions of each module to the sensitivity. (b) process for computing  $Jv$  and (c)  $J^\top v$ ; stars indicate where the source derivatives are incorporated.

Using electrical conductivity,  $\sigma$ , as the only active property described by the inversion model  $\mathbf{m}$  for brevity, the sensitivity takes the form

$$\mathbf{J}[\mathbf{m}] = \frac{d\mathbf{P}(\mathbf{f})}{d\mathbf{f}} \frac{d\mathbf{f}}{d\sigma} \frac{d\sigma}{dm} = \underbrace{\frac{d\mathbf{P}(\mathbf{f})}{d\mathbf{f}}}_{\text{Receivers}} \left( \underbrace{\frac{\partial \mathbf{f}}{\partial \mathbf{u}} \overbrace{\frac{du}{d\sigma}} + \frac{\partial \mathbf{f}}{\partial s_m} \overbrace{\frac{ds_m}{d\sigma}}}_{\text{Fields}} + \underbrace{\frac{\partial \mathbf{f}}{\partial s_e} \overbrace{\frac{ds_e}{d\sigma}} + \frac{\partial \mathbf{f}}{\partial \sigma}}_{\text{Properties}} \right) \underbrace{\frac{d\sigma}{dm}}_{\text{Properties}} \quad (9)$$

187 The annotations denote which of the elements shown in Figure ?? are responsible for computing the respective con-  
 188 tribution to the sensitivity. If the model provided is in terms of  $\mu$  or a source/receiver location, this property replaces  
 189 the role of  $\sigma$ . The flexibility to invoke distinct properties of interest (e.g.  $\sigma$ ,  $\mu$ , source location, etc.) in the inversion  
 190 requires quite a bit of ‘wiring’ to keep track of which model parameters are associated with which properties; this is  
 191 achieved through a property mapping or PropMap (physical properties, location properties, etc.) within SimPEG.

192 Although typically the source terms do not have model dependence and thus their derivatives are zero, the deriva-  
 193 tives of  $s_e$  and  $s_m$  must be considered in a general implementation. For example, if one wishes to use a primary-  
 194 secondary approach, where source fields are constructed by solving a simplified problem, the source terms may have  
 195 dependence on the model meaning their derivatives have a non-zero contribution to the sensitivity (c.f. ???); this will  
 196 be demonstrated in the Casing Example in Section ??.

The derivative of the solution vector  $\mathbf{u}$  with respect to the model is found by implicitly taking the derivative of

equation ?? with respect to  $\mathbf{m}$ , giving

$$\frac{d\mathbf{u}}{d\mathbf{m}} = \mathbf{A}^{-1}(\mathbf{m}) \left( -\underbrace{\frac{\partial \mathbf{A}(\mathbf{m})\mathbf{u}^{\text{fix}}}{\partial \mathbf{m}}}_{\text{getADeriv}} + \underbrace{\frac{\partial \mathbf{q}}{\partial \mathbf{s}_m} \frac{d\mathbf{s}_m}{d\mathbf{m}} + \frac{\partial \mathbf{q}}{\partial \mathbf{s}_e} \frac{d\mathbf{s}_e}{d\mathbf{m}} + \frac{\partial \mathbf{q}}{\partial \mathbf{m}}}_{\text{getRHSDeriv}} \right) \quad (10)$$

197 The annotations below the equation indicate the methods of the `Problem` class that are responsible for calculating the  
 198 respective derivatives. Typically the model dependence of the system matrix is through the physical properties (i.e.  $\sigma$ ,  
 199  $\mu$ ). Thus, to compute derivatives with respect to  $\mathbf{m}$ , the derivatives are first taken with respect to  $\sigma$  and the dependence  
 200 of  $\sigma$  on  $\mathbf{m}$  is treated using chain rule. The chain rule dependence is computed and tested automatically in SimPEG  
 201 using the composable Mapping classes.

202 In the following sections, we discuss the implementation of elements shown in Figure ?? and highlight their  
 203 contribution to the forward simulation and calculation of the sensitivity. We begin by discussing the inversion model  
 204 and its relationship to the physical properties (Section ??), move on to the core of the forward simulation, the Physics  
 205 (Section ??), and to how Sources which excite the system are defined (Section ??). Following these, we then discuss  
 206 how Fields are calculated everywhere in the domain (Section ??) and how they are evaluated by the Receivers to  
 207 create predicted data (Section ??). We conclude this section with a Summary and discussion on testing (Section ??).

#### 208 4.1. Model and Physical Properties

209 For all EM problems, we require an inversion model that can be mapped to meaningful physical properties in  
 210 the discretized Maxwell system. Typically, we consider the model to be a description of the electrical conductivity  
 211 distribution in the earth. Often, the model is taken to be log-conductivity, in which case, an exponential mapping is  
 212 required (`ExpMap`) to convert the model to electrical conductivity. The inversion model may be defined on a subset  
 213 of a mesh and referred to as an ‘active cell’ model. For instance, air cells may be excluded and only the subsurface  
 214 considered; in this case an `InjectActiveCells` map is used to inject the active model into the full simulation  
 215 domain. In the case of a parametric inversion, the inversion model is defined on a domain that is independent of  
 216 the forward modelling mesh and the mapping takes the parametric representation and defines a physical property on  
 217 the forward modelling mesh (e.g. a gaussian ellipsoid defined geometrically) (????). Maps can be composed, for  
 218 instance, a layered, 1D log conductivity model defined only in the subsurface may be mapped to a 2D cylindrical  
 219 Mesh, as shown in Figure ??.

```
220 import numpy as np
221 from SimPEG import Mesh, Maps
222 mesh = Mesh.CylMesh([20, 20])      # SimPEG cylindrically symmetric mesh
223 m_air = np.log(1e-8)              # value of the model in the air cells
224 indAct = mesh.vectorCCz < 0.0    # define active cells to be subsurface only
225 mapping = ( Maps.ExpMap(mesh) *
226             Maps.SurjectVertical1D(mesh) *
227             Maps.InjectActiveCells(mesh, indAct, m_air, nC=mesh.nCz) )
```

228 In the code above, the ‘multiplication’ performs the composition of the mappings. For the contribution of this action  
 229 to the sensitivity, the derivative of the electrical conductivity with respect to the model is computed using the  
 230 chain rule for the composed maps (cf. ??). During an inversion, the electrical conductivity on the simulation mesh  
 231 associated with the current inversion model and its derivative are accessed through the `BaseEMProblem`, which is  
 232 inherited by both the TDEM and FDEM problems. In some cases, variable magnetic permeability must be considered;

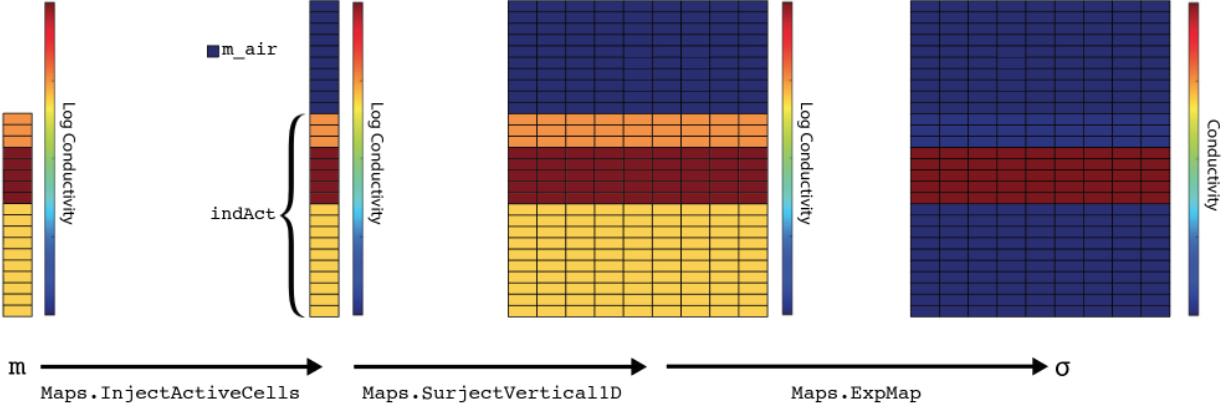


Figure 5: Mapping an inversion model, a 1D layered, log conductivity model defined below the surface, to electrical conductivity defined in the full simulation domain.

233 this is accomplished through a property mapping (PropMap). The PropMap handles the organization and independent  
 234 mappings of distinct physical properties (i.e.  $\sigma$ ,  $\mu$ ).

235 *4.2. Physics*

To formulate a system of equations from Maxwell's equations in time (equation ??) or frequency (equation ??) that can be solved numerically using a finite volume approach, we require a statement of the problem in terms of two equations with two unknowns, one of which is a field (discretized on edges), and the other a flux (discretized on faces). Thus, we can consider either the E-B formulation, or the H-J formulation. For the frequency-domain problem, we can discretize the electric field,  $\vec{e}$ , on edges, the magnetic flux,  $\vec{b}$ , on faces, physical properties  $\sigma$  and  $\mu^{-1}$  at cell centers, and the source terms  $\vec{s}_m$  and  $\vec{s}_e$  on faces and edges, respectively (see Figure ??). Doing so, we obtain the discrete system:

$$\begin{aligned} \mathbf{Ce} + i\omega \mathbf{b} &= \mathbf{s}_m \\ \mathbf{C}^\top \mathbf{M}_{\mu^{-1}}^f \mathbf{b} - \mathbf{M}_\sigma^e \mathbf{e} &= \mathbf{s}_e \end{aligned} \tag{11}$$

236 where  $\mathbf{C}$  is the discrete edge curl,  $\mathbf{M}_{\mu^{-1}}^f$  is the face inner-product matrix for  $\mu^{-1}$ ,  $\mathbf{M}_\sigma^e$  is the edge inner-product matrix  
 237 for  $\sigma$ ; these inner product matrices can be computed for isotropic, diagonally anisotropic or fully anisotropic physical  
 238 properties using operators within SimPEG's Mesh class (??).

Note that the source-term  $\mathbf{s}_e$  is an integrated quantity. Alternatively, the H-J formulation discretizes  $\vec{h}$  on edges,  $\vec{j}$  on faces,  $\rho$  and  $\mu$  at cell centers, and the source terms  $\vec{s}_m$ ,  $\vec{s}_e$  on edges and faces, respectively, giving

$$\begin{aligned} \mathbf{C}^\top \mathbf{M}_\rho^f \mathbf{j} + i\omega \mathbf{M}_\mu^e \mathbf{h} &= \mathbf{s}_m \\ \mathbf{Ch} - \mathbf{j} &= \mathbf{s}_e. \end{aligned} \tag{12}$$

239 Similarly,  $\mathbf{s}_m$  is an integrated quantity. In a full 3D simulation, the electric and magnetic contributions for the two  
 240 formulations are merely staggered from one another. However, if using an assumption of cylindrically symmetry, the

appropriate formulation must be used to simulate either rotational electric or magnetic contributions (?). For both the basic FDEM and TDEM implementations, natural boundary conditions ( $\mathbf{b} \times \hat{\mathbf{n}} = 0 \forall \vec{x} \in \partial\Omega$  in E-B formulation or  $\mathbf{j} \times \hat{\mathbf{n}} = 0 \forall \vec{x} \in \partial\Omega$  in H-J formulation), in which the fields are assumed to have decayed to a negligible value at the boundary, are employed to construct the differential operators, the framework and implementation are however, extensible to consider other boundary conditions (cf. ??).

In order to solve either equation ?? or equation ??, we eliminate one variable and solve the second order system. This elimination is performed by the FDEM problem classes. For instance, in `FDEM Problem_e`, we eliminate  $\mathbf{b}$  and obtain a second order system in  $\mathbf{e}$

$$\underbrace{\left( \mathbf{C}^\top \mathbf{M}_{\mu^{-1}}^f \mathbf{C} + i\omega \mathbf{M}_\sigma^e \right)}_{\text{getA}} \underbrace{\mathbf{e}}_{\text{u}} = \underbrace{\mathbf{C}^\top \mathbf{M}_{\mu^{-1}}^f \mathbf{s}_m - i\omega \mathbf{s}_e}_{\text{getRHS}} \quad (13)$$

`FDEM Problem_e` has methods `getA` and `getRHS` to construct the system

```
247     def getA(self, freq):
248         MfMui = self.MfMui
249         MeSigma = self.MeSigma
250         C = self.mesh.edgeCurl
251         return C.T*MfMui*C + 1j*omega(freq)*MeSigma
252
253     def getRHS(self, freq):
254         s_m, s_e = self.getSourceTerm(freq)
255         MfMui = self.MfMui
256         C = self.mesh.edgeCurl
257         return C.T * (MfMui * s_m) - 1j * omega(freq) * s_e
```

and associated methods `getADeriv` and `getRHSDeriv` to construct the derivatives of each with respect to the inversion model. These function definitions are methods of the `Problem` class, where the `self` variable refers to the instance of the class, and is standard Python (cf. Python documentation - <https://docs.python.org/3/tutorial/classes.html>). For FDEM `Problem_e`, `getRHSDeriv` is zero unless one or both of the source terms have model dependence. However, if we eliminate  $\mathbf{e}$  and solve for  $\mathbf{b}$  (`Problem_b`), the right hand side contains the matrix  $\mathbf{M}_\sigma^e$ , and therefore will, in general, have a non-zero derivative. To solve this linear system of equations, SimPEG interfaces to standard numerical solver packages (e.g. SciPy, Mumps (????)), using for example `pymatsolver` <https://github.com/rowanc1/pymatsolver>). The components used to perform the forward simulation are assembled in the `fields` method of the `BaseFDEMProblem` class; the `fields` method solves the forward simulation for the solution vector  $\mathbf{u}$  (from equation ??) at each frequency and source considered.

Similarly, for the time-domain problem, the semi-discretized E-B formulation is given by

$$\begin{aligned} \mathbf{Ce} + \frac{d\mathbf{b}}{dt} &= \mathbf{s}_m \\ \mathbf{C}^\top \mathbf{M}_{\mu^{-1}}^f \mathbf{b} - \mathbf{M}_\sigma^e \mathbf{e} &= \mathbf{s}_e \end{aligned} \quad (14)$$

and the semi-discretized H-J formulation is given by

$$\begin{aligned} \mathbf{C}^\top \mathbf{M}_{\rho^{-1}}^f \mathbf{j} + \frac{d\mathbf{M}_\mu^e \mathbf{h}}{dt} &= \mathbf{s}_m \\ \mathbf{Ch} - \mathbf{j} &= \mathbf{s}_e. \end{aligned} \quad (15)$$

For the time discretization, we use Backward Euler (cf. [?](#)). To form the TDEM Problem\_b, we eliminate  $\mathbf{e}$  from equation [??](#) and apply Backward Euler for the time discretization. A single timestep takes the form

$$\underbrace{\left( \mathbf{C} \mathbf{M}_\sigma^{e-1} \mathbf{C}^\top \mathbf{M}_{\mu^{-1}}^f + \frac{1}{\Delta t^k} \right)}_{\mathbf{A}_0^{k+1}(\mathbf{m})} \underbrace{\mathbf{b}^{k+1}}_{\mathbf{u}^{k+1}} + \underbrace{\frac{-1}{\Delta t^k} \mathbf{I}}_{\mathbf{A}_{-1}^{k+1}(\mathbf{m})} \underbrace{\mathbf{b}^k}_{\mathbf{u}^k} = \underbrace{\mathbf{C} \mathbf{M}_\sigma^{e-1} \mathbf{s}_e^{k+1} + \mathbf{s}_m^{k+1}}_{\mathbf{q}^{k+1}(\mathbf{s}_m, \mathbf{s}_e)} \quad (16)$$

where  $\Delta t^k = t^{k+1} - t^k$  is the timestep and the superscripts  $k, k+1$  indicate the time index. Each TDEM problem formulation (ie. `Problem_e`, `Problem_b`, `Problem_h`, `Problem_j`) has methods to create the matrices along the block-diagonals,  $\mathbf{A}_0^{k+1}(\mathbf{m})$  and  $\mathbf{A}_{-1}^{k+1}(\mathbf{m})$ , as well as a method to construct the right hand side,  $\mathbf{q}^{k+1}(\mathbf{s}_m, \mathbf{s}_e)$ , at each timestep. When inverting for a model in electrical conductivity using `Problem_b`, the subdiagonal matrices are independent of  $\mathbf{m}$ , however, in other formulations, such as `Problem_e`, the subdiagonal matrices do have dependence on electrical conductivity, thus in general, the model dependence must be considered. Depending on the solver chosen, it can be advantageous to make the system symmetric; this is accomplished by multiplying both sides by  $\mathbf{M}_{\mu^{-1}}^{f^\top}$ . To solve the full time-stepping problem, we assemble all timesteps in a lower block bidiagonal matrix, with on-diagonal matrices  $\mathbf{A}_0^k(\mathbf{m})$  and sub-diagonal matrices  $\mathbf{A}_{-1}^k(\mathbf{m})$ , giving

$$\underbrace{\begin{pmatrix} \mathbf{A}_0^0(\mathbf{m}) & & & \\ \mathbf{A}_{-1}^1(\mathbf{m}) & \mathbf{A}_0^1(\mathbf{m}) & & \\ & \mathbf{A}_{-1}^2(\mathbf{m}) & \mathbf{A}_0^2(\mathbf{m}) & \\ & & \ddots & \ddots \\ & & & \mathbf{A}_{-1}^{n-1}(\mathbf{m}) & \mathbf{A}_0^{n-1}(\mathbf{m}) \\ & & & & \mathbf{A}_{-1}^n(\mathbf{m}) & \mathbf{A}_0^n(\mathbf{m}) \end{pmatrix}}_{\mathbf{A}(\mathbf{m})} \underbrace{\begin{pmatrix} \mathbf{u}^0 \\ \mathbf{u}^1 \\ \mathbf{u}^2 \\ \vdots \\ \mathbf{u}^{n-1} \\ \mathbf{u}^n \end{pmatrix}}_u = \underbrace{\begin{pmatrix} \mathbf{q}^0 \\ \mathbf{q}^1 \\ \mathbf{q}^2 \\ \vdots \\ \mathbf{q}^{n-1} \\ \mathbf{q}^n \end{pmatrix}}_{\mathbf{q}(\mathbf{s}_m, \mathbf{s}_e)} \quad (17)$$

268 When solving the forward simulation, the full time-stepping matrix,  $\mathbf{A}(\mathbf{m})$ , is not formed, instead the block system  
269 is solved using forward substitution with each block-row being computed when necessary. The initial condition,  $\mathbf{u}^0$ ,  
270 depends on the source type and waveform; it is computed numerically or specified using an analytic solution. For  
271 example, if using a grounded source and a step-off waveform,  $\mathbf{u}^0$  is found by solving the direct current resistivity  
272 or the magnetometric resistivity problem, depending on which field we choose to solve for. When a general current  
273 waveform is considered, the initial condition will be  $\mathbf{u}^0 = \mathbf{0}$ , and either  $\mathbf{s}_m$  or  $\mathbf{s}_e$ , depending on type of the source used,  
274 will have non-zero values during the on-time.

275 Derivatives of the matrices along the block-diagonals of  $\mathbf{A}(\mathbf{m})$  along with derivatives of the right-hand-side are  
276 stitched together in a forward time stepping approach to compute the contribution of  $\frac{d\mathbf{u}}{d\mathbf{m}}$  to  $\mathbf{J}^\top \mathbf{v}$  and in a backwards  
277 time stepping approach for the contribution of  $\frac{d\mathbf{u}}{d\mathbf{m}}^\top$  to  $\mathbf{J}^\top \mathbf{v}$ .

278 **4.3. Sources**

279 Sources input EM energy into the system. They can include grounded wires, loops, dipoles and natural sources.  
280 Controlled sources are implemented in the FDEM and TDEM modules of SIMPEGEM, and natural sources are implemented

281 in the NSEM module. For simulations, we require that the sources be discretized onto the mesh so that a right-hand-side  
 282 for the Maxwell system can be constructed (i.e. `getRHS`). This is addressed by the `eval` method of the source which  
 283 returns both the magnetic and electric sources ( $\mathbf{s}_m, \mathbf{s}_e$ , shown in Figure ??) on the simulation mesh.

In some cases, a primary-secondary approach can be advantageous for addressing the forward problem (cf. ???). We split up the fields and fluxes into primary and secondary components ( $\mathbf{e} = \mathbf{e}^P + \mathbf{e}^S, \mathbf{b} = \mathbf{b}^P + \mathbf{b}^S$ ) and define a “Primary Problem”, a simple problem, often with an analytic solution, that is solved in order to construct a source term for a secondary problem. For instance, a point magnetic dipole source may be simulated by defining a zero-frequency primary which satisfies

$$\begin{aligned}\mathbf{e}^P &= 0 \\ \mathbf{C}^\top \mathbf{M}_{\mu^{-1}}^f \mathbf{b}^P &= \mathbf{s}_e^P.\end{aligned}\tag{18}$$

If we define  $\mu^{-1}^P$  to be a constant, equation ?? has an analytic solution for  $\mathbf{b}^P$  that may be expressed in terms of a curl of a vector potential (cf. ?). When using a mimetic discretization, by defining the vector potential and taking a discrete curl, we maintain that the magnetic flux density is divergence free as the divergence operator is in the null space of the edge curl operator ( $\nabla \cdot \nabla \times \vec{v} = 0$ ), so numerically we avoid creating magnetic monopoles (c.f. ?). The secondary problem is then

$$\begin{aligned}\mathbf{C}\mathbf{e}^S + i\omega\mathbf{b}^S &= -i\omega\mathbf{b}^P \\ \mathbf{C}^\top \mathbf{M}_{\mu^{-1}}^f \mathbf{b}^S - \mathbf{M}_\sigma^e \mathbf{e}^S &= -\mathbf{C}^\top \left( \mathbf{M}_{\mu^{-1}}^f - \left( \mathbf{M}_{\mu^{-1}}^f \right)^P \right) \mathbf{b}^P\end{aligned}\tag{19}$$

284 The source terms for the secondary problem are  $\mathbf{s}_m = -i\omega\mathbf{b}^P$ , and  $\mathbf{s}_e = -\mathbf{C}^\top (\mathbf{M}_{\mu^{-1}}^f - \mathbf{M}_{\mu^{-1}}^{f,P}) \mathbf{b}^P$ . In scenarios where  
 285 magnetic permeability is homogeneous, the electric source contribution is zero.

286 The left hand side is the same discrete Maxwell system as in equation ??; the distinction is that we are solving  
 287 for secondary fields, and a primary problem was solved (analytically or numerically) in order to construct the source  
 288 terms. To obtain the total fields, which we sample with the receivers, we must add the primary fields back to the  
 289 solution. To keep track of the primary fields, they are assigned as properties of the source class.

290 In most cases, source terms do not have a derivative with respect to the model. However, in a primary-secondary  
 291 problem in electrical conductivity the source term depends on the electrical conductivity and derivatives must be  
 292 considered (see Section ??). This is similar to inverting for magnetic permeability using a primary-secondary approach  
 293 described in equation ?? (??). It is also possible to consider your inversion model to be the location or waveform  
 294 of the source, in which case the derivative is also non-zero and source derivatives can be included in the optimization  
 295 procedure.

#### 296 4.4. Fields

297 By solving the second-order linear system, as in equation ??, we obtain a solution vector,  $\mathbf{u}$ , of one field or flux  
 298 everywhere in the domain. In the case of a primary-secondary problem, this solution is a *secondary* field. To examine

299 all of the fields, we require easy access to the total fields and total fluxes everywhere in the domain. This is achieved  
300 through the `Fields` object.

For efficient memory usage, only the solution vector is stored, all other fields and fluxes are calculated on demand through matrix vector multiplications. As such, each problem type (**e**, **b**, **h**, **j**) has an associated `Fields` object with methods to take the solution vector and translate it to the desired field or flux. For instance, `Fields_j` stores the solution vector from `Problem_j` and has methods to compute the total magnetic field in the simulation domain by first computing the secondary magnetic field from the solution vector (**u**; in this example, **u** = **j**) and adding back any contribution from the source

$$\mathbf{h} = \frac{1}{i\omega} \mathbf{M}_\mu^{e-1} (-\mathbf{C}^\top \mathbf{M}_\rho^f \mathbf{u} + \mathbf{s}_m) \quad (20)$$

For their contribution to the sensitivity (equation ??), the fields have methods to compute derivatives when provided the vectors **v** and  $\frac{d\mathbf{u}}{d\mathbf{m}}\mathbf{v}$  (from the `Physics`). For instance, for **h**

$$\frac{d\mathbf{h}}{d\mathbf{m}}\mathbf{v} = \frac{d\mathbf{h}}{d\mathbf{u}} \left( \frac{d\mathbf{u}}{d\mathbf{m}}\mathbf{v} \right) + \left( \frac{d\mathbf{h}}{ds_e} \frac{ds_e}{d\mathbf{m}} + \frac{d\mathbf{h}}{ds_m} \frac{ds_m}{d\mathbf{m}} + \frac{\partial \mathbf{h}}{\partial \mathbf{m}} \right) \mathbf{v} \quad (21)$$

301 The derivatives for **e**, **b**, and **j** take the same form. Conceptually, the product of the full derivative and a vector  
302  $\left( \frac{d\mathbf{f}}{d\mathbf{m}}\mathbf{v} \right)$  can be thought of as a stacked vector of all of the contributions from all of the fields and fluxes, however, this  
303 is never formed in practice.

#### 304 4.5. Receivers

305 The measured data consist of specific spatial components of the fields or fluxes sampled at the receiver locations at  
306 a certain time or frequency. Receivers have the method `eval` that interpolates the necessary components of the fields  
307 and fluxes to the receiver locations and evaluates the data required for the problem, such as the frequency domain fields  
308 or natural source impedance data. For the frequency domain problem, real and imaginary components are treated as  
309 separate data so that when inverting, we are always working with real values. The separation of the data evaluation  
310 from fields in receiver objects allows the derivative computation to be performed and tested in a modular fashion; this  
311 enables rapid development and implementation of new receiver types.

#### 312 4.6. Summary

313 Having defined the role of each of the elements in the forward simulation framework outlined in Figure ??, the  
314 necessary machinery to compute predicted data and sensitivities is at hand for both FDEM and TDEM problems. The  
315 modular nature of the framework allows us to make several abstractions which make the code more transparent and  
316 ensure consistency across implementations. For instance, the definition of the physical properties and associated  
317 inner product matrices is common to all formulations in both time and frequency domains. Thus, these are defined  
318 as properties of a `BaseEM` class which is inherited by both the TDEM and FDEM modules. Within each of the TDEM  
319 and FDEM modules, common methods for the calculation of the fields, sensitivities and adjoint are defined and shared  
320 across the approaches that solve for **e**, **b**, **h**, or **j** (see the documentation <http://docs.simpeg.xyz>).

321 Testing is conducted using comparisons with analytics, cross-comparisons between formulations, order tests on  
322 the sensitivity, adjoint tests, examples, tests on the finite volume operators, projections, interpolations, solvers, etc.  
323 Tests are run upon each update to the repository through the continuous integration service TravisCI (?). This ensures  
324 that we can trust the tools that we use and move faster in our research into new methods and implementations. This  
325 also supports new developers and researchers in contributing to the code base without fear of breaking assumptions  
326 and ideas laid out by previous development.

327 **5. Examples**

328 To demonstrate the application and structure of the framework, we explore three examples, one field example and  
329 two synthetic examples. The purpose of the first synthetic example is to show simple time and frequency domain  
330 electromagnetic inversions, and highlight the common framework. For this, we invert for a 1D layered Earth using a  
331 2D cylindrically symmetric mesh for the forward simulation. In the second example, we show 1D inversions of field  
332 data (RESOLVE and SkyTEM) collected over the Bookpurnong Irrigation district in Australia. The final example is a  
333 3D synthetic example that demonstrates a sensitivity analysis using a parametric model of a block in a layered space  
334 for a reservoir characterization problem where the transmitter is positioned down-hole in a steel-cased well. We use  
335 this example to demonstrate how mappings, multiple physical properties (both electrical conductivity and magnetic  
336 permeability), and multiple meshes, a cylindrically symmetric and a 3D tensor mesh, can be composed in a primary-  
337 secondary approach for performing the forward simulation and computing the sensitivities. The scripts used to run  
338 these examples are available on <http://docs.simpeg.xyz>.

339 *5.1. Cylindrically Symmetric Inversions*

340 The purpose of this example is to demonstrate the implementation of the electromagnetic inversion in both time  
341 and frequency domains. We have chosen this example as it is computationally light, can be run on any modern laptop  
342 without installing complex dependencies, and yet it uses most of the elements and functionality needed to solve a large  
343 3D EM problem. The script used to run this simulation is available at: <https://doi.org/10.6084/m9.figshare.5035175>.

344 We consider two 1D inversions for log-conductivity from an EM survey, one frequency domain experiment and  
345 one time domain experiment. Both surveys use a vertical magnetic dipole (VMD) source located on the surface. For  
346 simplicity, we consider a single receiver, measuring the vertical magnetic field, located 50m radially away from the  
347 source. The magnetic permeability is taken to be that of free space ( $\mu = \mu_0$ ), and electrical conductivity is assumed to  
348 be frequency-independent.

349 Figure ?? shows the setup used for: (a) the frequency domain simulation, (b) the time domain simulation, and  
350 (c) the common inversion implementation. In both, a cylindrical mesh is employed for the forward simulation and  
351 a 1D layered earth, described in terms of log-conductivity. To map the inversion model to electrical conductivity,  
352 a composite mapping is used to inject the 1D subsurface model into one including air cells (InjectActiveCells),

353 subject the 1D model onto the 2D simulation mesh (`SurjectVertical1D`) and take the exponential to obtain electrical  
 354 conductivity (`ExpMap`), as described in the Model and Physical Properties section (Section ??).

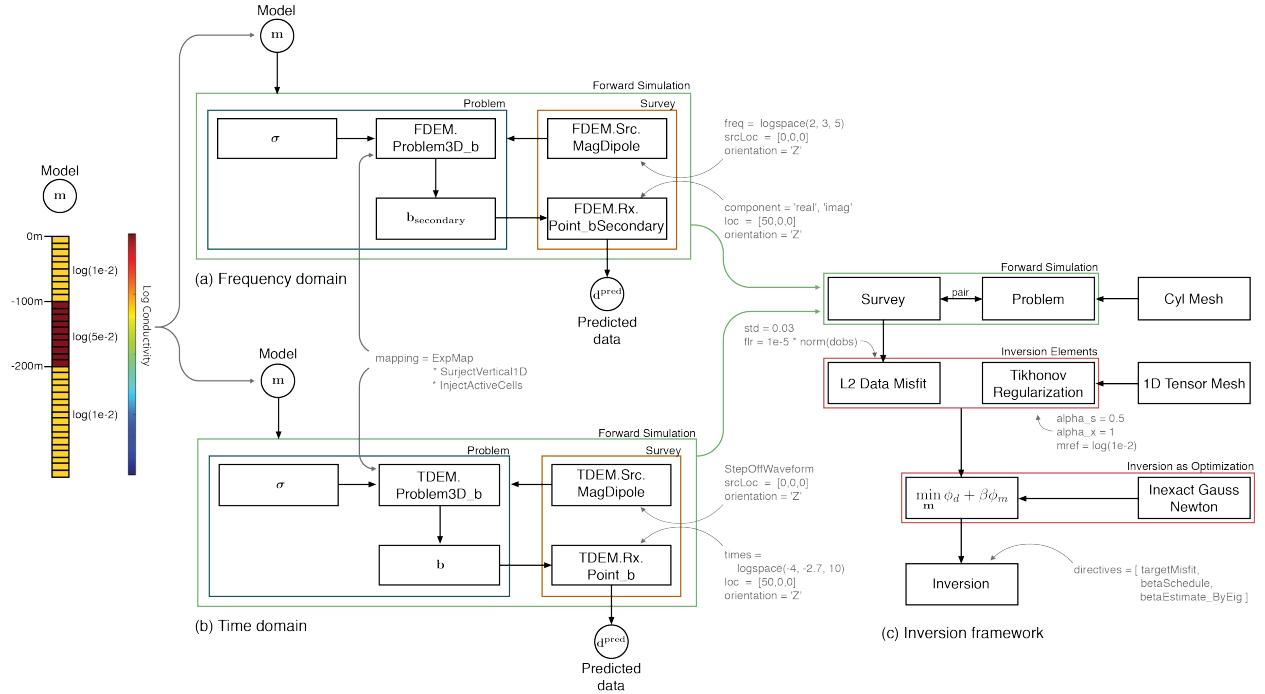


Figure 6: Diagram showing the entire setup and organization of (a) the frequency domain simulation; (b) the time domain simulation; and (c) the common inversion framework used for each example. The muted text shows the programmatic inputs to each class instance.

355 The distinction between the frequency and time domain inversions comes in the setup of the forward simulations.  
 356 Each employs the appropriate description of the physics (FDEM or TDEM) in the problem, and the definition of the  
 357 survey, consisting of both sources and receivers, must be tailored to the physics chosen. For the FDEM survey, a  
 358 vertical harmonic magnetic dipole located at the origin transmits at five frequencies logarithmically spaced between  
 359 100 Hz and 1000 Hz. The receiver is located at (50 m, 0 m, 0 m) and measures the secondary magnetic flux (with the  
 360 primary being the free-space response of a harmonic magnetic dipole). The observed response is complex-valued,  
 361 having both real and imaginary components. We consider these as separate data, giving a total of ten data points for  
 362 this example. For the time domain survey, we again use a vertical magnetic dipole at the origin, however, we now  
 363 use a step-off waveform. The observed responses are defined through time, and thus are all real-valued. For this  
 364 example, we sample 10 time channels, logarithmically spaced between  $10^{-4}$  s and  $2 \times 10^{-3}$  s. These time channels  
 365 were selected to be sensitive to depths similar to the FDEM simulation.

366 With the forward simulation parameters defined in both the time and frequency domain simulations, we can  
 367 generate synthetic data. The model used consists of a 100m thick conductive layer (0.05 S/m) whose top boundary  
 368 is 100 m-below from the surface, as shown in Figure ???. The conductivity of the half-space earth is 0.01 S/m. In  
 369 both cases, 3% gaussian noise is added to the simulated data, and these are treated as the observed data ( $\mathbf{d}^{\text{obs}}$ ) for the

370 inversion.

For the inversions, we specify the inversion elements: a data misfit and a regularization. We use an L2 data misfit of the form

$$\phi_d = \frac{1}{2} \|\mathbf{W}_d(\mathbf{d}^{\text{pred}} - \mathbf{d}^{\text{obs}})\|_2^2 \quad (22)$$

where  $\mathbf{W}_{d_{ii}} = 1/\epsilon_i$  and we define  $\epsilon_i = 3\%|d_i^{\text{obs}}| + \text{floor}$ . For both simulations the floor is set to  $10^{-5}\|\mathbf{d}^{\text{obs}}\|$ . The regularization is chosen to be a Tikhonov regularization on the 1D model

$$\phi_m = \frac{1}{2} (\alpha_s \|\mathbf{m} - \mathbf{m}_{\text{ref}}\|_2^2 + \alpha_x \|\mathbf{D}_x \mathbf{m}\|_2^2) \quad (23)$$

371 where  $\mathbf{m}_{\text{ref}}$  is the reference model which is set to be a half-space of  $\log(10^{-2})$ . The matrix  $\mathbf{D}_x$  is a 1D gradient operator.  
372 For both examples  $\alpha_s = 0.5$  and  $\alpha_x = 1$ . The data misfit and regularization are combined with a trade-off parameter,  
373  $\beta$ , in the statement of the inverse problem. To optimize, we use the second-order Inexact Gauss Newton scheme. In  
374 this inversion, we use a beta-cooling approach, where  $\beta$  is reduced by a factor of 4 every 3 Gauss Newton iterations.

375 The initial  $\beta$  is chosen to relatively weight the influence of the data misfit and regularization terms. We do this  
376 by estimating the largest eigenvalue of  $\mathbf{J}^\top \mathbf{J}$  and  $\mathbf{W}_m^\top \mathbf{W}_m$  using one iteration of the power method. We then take  
377 their ratio and multiply by a scalar to weight their relative contributions. For this example, we used a factor of 10.  
378 For a stopping criteria, we use the discrepancy principle, stopping the inversion when  $\phi_d \leq \chi \phi_d^*$ , with  $\chi = 1$  and  
379  $\phi_d^* = 0.5N_{\text{data}}$  (with  $\phi_d$  as defined in equation ??.)

380 The FDEM inversion reaches the target misfit after 9 iterations, and the TDEM inversion reaches the target misfit  
381 after 6 iterations. Figure ?? shows the recovered models (a), predicted and observed data for the FDEM inversion (b)  
382 and predicted and observed data for the TDEM inversion (c). In both the FDEM and TDEM inversions, the data are  
383 fit well. The recovered models are smooth, as is expected when employing an L2, Tikhonov regularization and both  
384 the location and amplitude of the conductive layer. The structure of both models are comparable, demonstrating that  
385 the information content in both the FDEM and TDEM data are similar. The recovered model can be improved by  
386 many additional techniques that are not explored here (e.g. using compact norms in the regularization). The SimPEG  
387 package provides a number of additional directives and regularization modules which can be useful for this purpose.

### 388 5.2. Bookpurnong Field Example

389 The purpose of this example is to demonstrate the use of the framework for inverting field data and provide an  
390 inversion that can be compared with other results in the literature. In particular, we invert frequency and time domain  
391 data collected over the Bookpurnong Irrigation District in Southern Australia. The Murray River and adjacent flood-  
392 plain in the Bookpurnong region have become extensively salinized, resulting in vegetation die-back (??). Multiple  
393 electrical and electromagnetic data sets have been collected with the aim of characterizing the near-surface hydrologic  
394 model of the area (?). For a more complete background on the geology and hydrogeology of the Bookpurnong region,  
395 we refer the reader to ?.

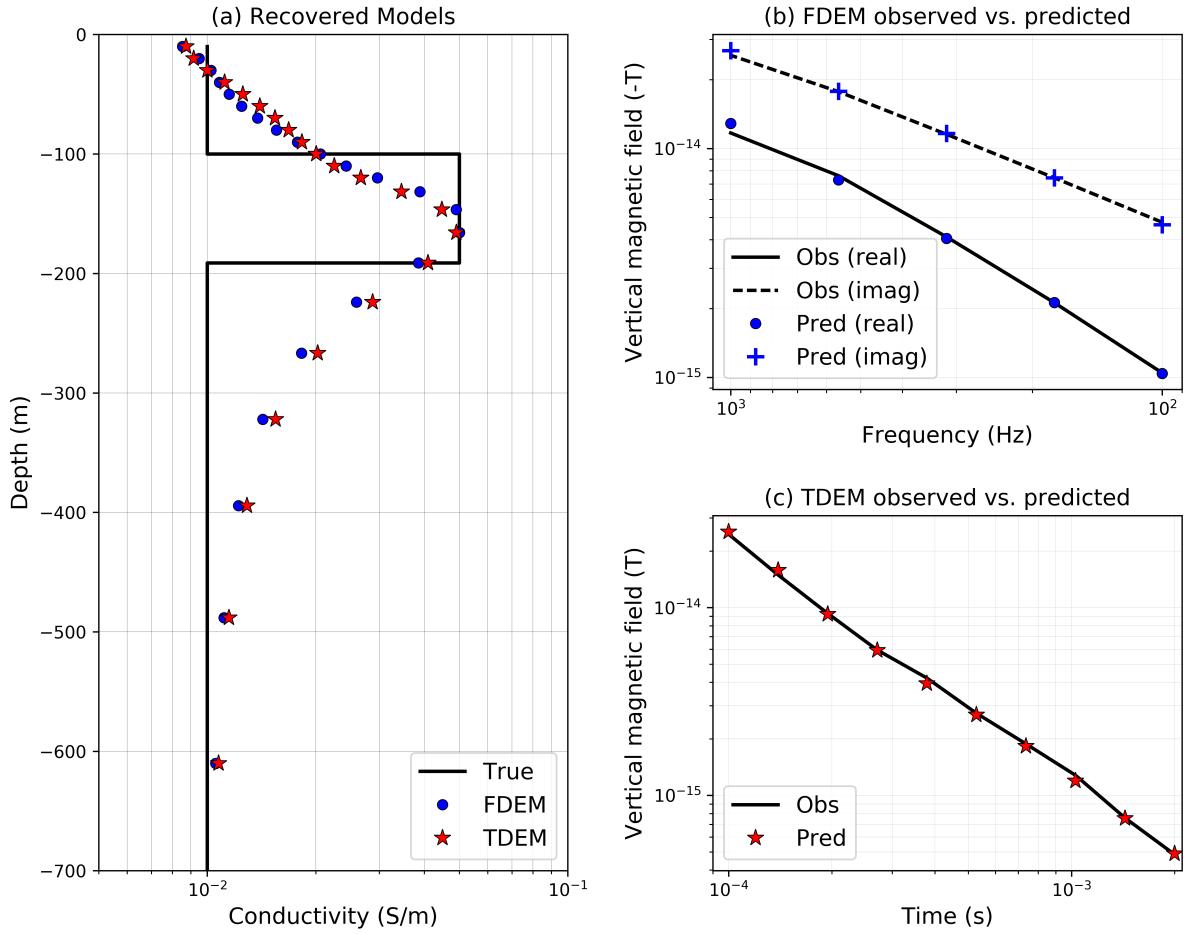


Figure 7: (a) True and recovered models for the FDEM and TDEM inversions; predicted and observed data for (b) the FDEM example, and (c) the TDEM example. In (b) the magnetic field data are in the negative z-direction.

396 Here, we will focus our attention to the RESOLVE frequency-domain data collected in 2008 and the SkyTEM  
 397 time-domain data collected in 2006. These data are shown in Figure ???. The RESOLVE system consists of 5 pairs of  
 398 horizontal coplanar coils, with nominal frequencies of 400 Hz, 1800 Hz, 8200 Hz, 40 000 Hz, and 130 000 Hz as well  
 399 as a vertical coaxial coil pair of coils which operates at 3200Hz. For the Bookpurnong survey, the bird was flown at  
 400 ~50m altitude (?). The SkyTEM time-domain system operates in two transmitter modes that can be run sequentially.  
 401 The high moment mode has high current and operates at a low base frequency (25 Hz and can be lowered to 12.5  
 402 Hz), and the low moment operates at a lower current and higher base frequency (222.5 Hz) (?). The Bookpurnong  
 403 SkyTEM survey was flown at an altitude of ~60m (?).

404 Multiple authors have inverted these data sets; 1D spatially constrained inversions of the SkyTEM and RE-  
 405 SOLVE data were performed by (??). ? independently inverted these data in 1D and provides a discussion at

406 [http://em.geosci.xyz/content/case\\_histories/bookpurnong/index.html](http://em.geosci.xyz/content/case_histories/bookpurnong/index.html). The SkyTEM data (high moment) were inverted  
 407 in 3D by (?). In the example that follows, we select a location where both the RESOLVE and SkyTEM datasets have  
 408 soundings and invert them in 1D, we then proceed to perform a stitched 1D inversion of the RESOLVE data. The  
 409 data have been made available with the permission of CSIRO and are accessible, along with the script used to run the  
 410 inversions at <https://doi.org/10.6084/m9.figshare.5107711>.

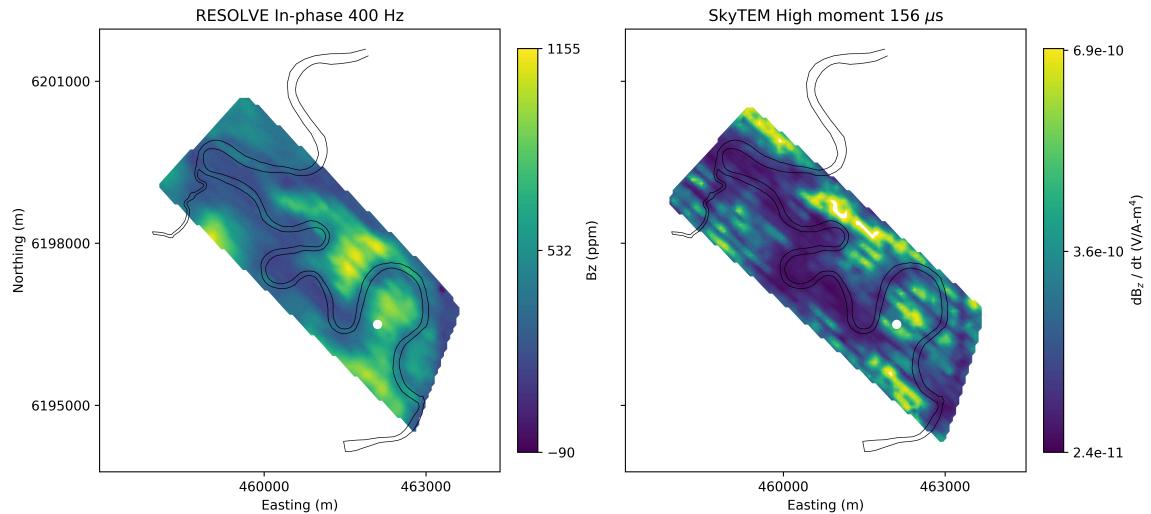


Figure 8: 400 Hz In-phase RESOLVE data at (left) and High Moment SkyTEM data at  $156 \mu\text{s}$ . The white dot at (462100m, 6196500m) on both images is the location of the stations chosen to demonstrate the 1D inversions in frequency and time.

#### 411 5.2.1. 1D Inversion of RESOLVE and SkyTEM soundings

412 We have selected a sounding location (462100m, 6196500m) at which to perform 1D inversions of the RESOLVE  
 413 and SkyTEM (High Moment) data. The observed data at this location are shown in Figure ?? (b) and (c). For the  
 414 RESOLVE inversion, we consider the horizontal co-planar data collected at 400 Hz, 1800 Hz, 8200 Hz, 40 000 Hz,  
 415 and 130 000 Hz. For the noise model, we assign 10% error for the three lowest frequencies and 15% error for the two  
 416 highest; a noise floor of 20 ppm is assigned to all data. The inversion mesh uses cells that expand logarithmically with  
 417 depth, starting at the surface with a finest cell size of 1m. The forward simulation is carried out on the cylindrically  
 418 symmetric mesh, similar to the previous example. In the inversion, we employ a Tikhonov regularization in which  
 419 length scales have been omitted in the regularization function. A fixed trade-off parameter of  $\beta = 2$  is used,  $\alpha_z$  is  
 420 set to be 1, and  $\alpha_s$  is  $10^{-3}$ . A half-space reference model with conductivity 0.1 S/m is used, this also served as the  
 421 starting model for the inversion. The inversion reached target misfit after 2 iterations. The resulting model and data  
 422 fits are shown in Figure ???. Very close to the surface, we recover a resistor, while below that, we recover a conductive  
 423 unit ( $\sim 2$  S/m). Examining the data (Figure ??b), we see that the real components are larger in magnitude than the  
 424 imaginary, and that with increasing frequency, the magnitude of the imaginary component decreases while the real

425 component increases; such behaviour is consistent with an inductive- limit response, and we thus expect to recover  
 426 conductive structures in the model.

427 For the time domain inversion, we consider the SkyTEM high moment data. We use the source waveform shown  
 428 in the inset plot in Figure ?? (c). For data, we use 21 time channels from  $47 \mu\text{s}$  to  $4.4 \text{ ms}$ ; the latest three time  
 429 channels ( $5.6\text{ms}$ ,  $7\text{ms}$  and  $8.8 \text{ ms}$ ) are not included. For data errors, we assign a 12% uncertainty and a floor of  
 430  $2.4 \times 10^{-14} \text{ V/Am}^4$ . We again use a Tikhonov regularization, here with  $\alpha_z = 1$  and  $\alpha_s = 10^{-1}$ . The trade-off parameter  
 431 is  $\beta = 20$ . A half-space starting model of  $0.1 \text{ S/m}$  is again employed. For the reference model, we use the model  
 432 recovered from the RESOLVE 1D inversion. As we are using the high-moment data, we do not expect the SkyTEM  
 433 data to be as sensitive to the near surface structures as the RESOLVE data. By using the model recovered in the  
 434 RESOLVE inversion as the starting model for the SkyTEM inversion, we can assess agreement between the two and  
 435 isolate structures that are introduced by the SkyTEM inversion. The inversion reached the target misfit after 3 iteration  
 436 and the results are shown in Figure ???. At this location, there is good agreement in the models recovered from the  
 437 RESOLVE and SkyTEM data, with both supporting a near-surface resistor and showing a deeper conductive structure.

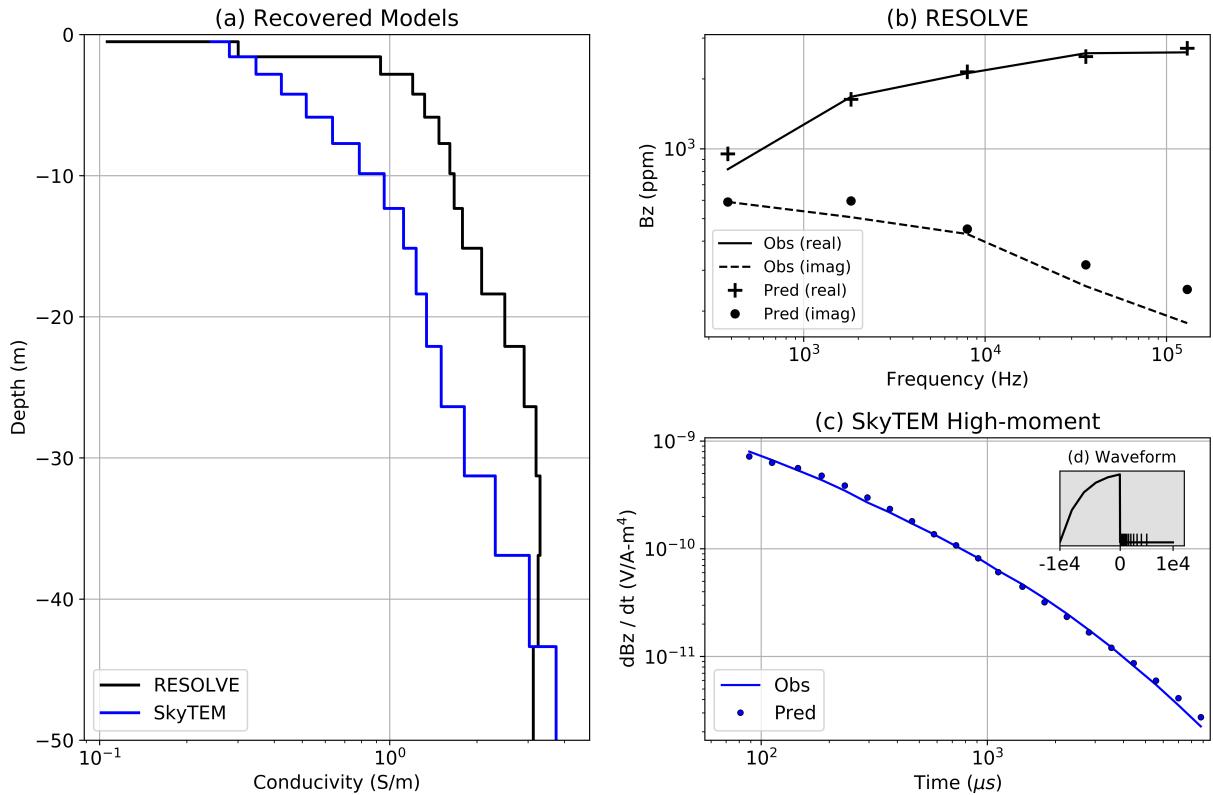


Figure 9: (a) Models recovered from the 1D inversion of RESOLVE (back) and SkyTEM (blue) data at the location (462100m, 6196500m). (b) Observed (lines) and predicted (points) frequency domain data. (c) Observed and predicted time domain data. (d) Source waveform used in for the SkyTEM inversion, the x-axis is time ( $\mu\text{s}$ ) on a linear scale.

438 5.2.2. *Stitched 1D inversion of RESOLVE data*

439 Next, we perform a stitched 1D inversion of the RESOLVE data set. With this example, we aim to demonstrate a  
 440 practical inversion workflow that will run on modest computational resources. As such, we have heavily downsampled  
 441 the data set, taking 1021 stations of the 40 825 collected. A 1D stitched inversion is a relatively straight-forward  
 442 approach for creating a conductivity model - each sounding is inverted independently and the inversion results are  
 443 then assembled to create a 3D model. This can be a valuable quality-control step prior to adopting more advanced  
 444 techniques such as including lateral or 3D regularization across soundings or even performing a 3D inversion. In cases  
 445 where the geology is relatively simple, a stitched 1D inversion may be sufficient. The inversion parameters are the  
 446 same as those used in the inversion of the RESOLVE sounding discussed in the previous section. A plan- view of the  
 447 recovered model 9.9m below the surface is shown in Figure ??a. A global  $\chi$ -factor of 0.74 was reached, and plots  
 448 comparing the real component of the observed and predicted data at 400Hz are shown in Figures ?? (b) & (c).

449 The recovered model (Figure ??a), bears similar features to the models found by ? (Figure 4 of ?) and by ?. In  
 450 general, the northwestern portion of the Murray river is more resistive, in particular near (459 000m, 6 200 000m)  
 451 and (460 000m, 6 198 000m) while the southeastern portion of the river is more conductive. Two mechanisms of  
 452 river salinization have been discussed in ??: the resistive regions are attributed to a “losing” groundwater system, in  
 453 which freshwater from the Murray River discharges to adjacent banks, while the conductive regions are attributed to  
 454 a “gaining” system, in which regional saline groundwater seeps into the river.

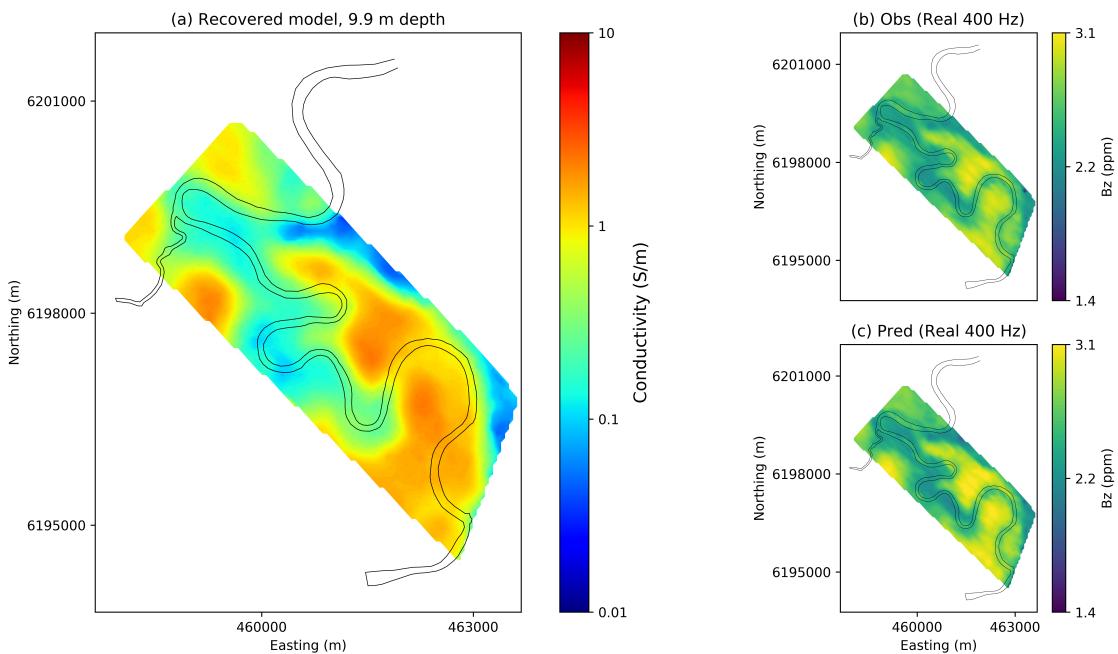


Figure 10: (a) Conductivity model 9.9m below the surface from a stitched 1D inversion of RESOLVE data. (b) Real component of the observed RESOLVE data at 400Hz. (c) Real component of the predicted data at 400Hz.

455    5.3. Steel-Cased Well: Sensitivity Analysis for a Parametric Model

456    The purpose of this example is to demonstrate the modular implementation of simpegEM and how it can be  
457    used to experiment with simulation and inversion approaches. Conducting electromagnetic surveys in settings where  
458    steel casing is present is growing in interest for applications such as monitoring hydraulic fracturing or enhanced  
459    oil recovery (?????????). Steel is highly conductive ( $\sim 5.5 \times 10^6 S/m$ ), has a significant magnetic permeability ( $\sim$   
460     $50\mu_0 - 100\mu_0$ ) (?). This is a large contrast to typical geologic settings, with conductivities typically less than 1  
461    S/m and permeabilities similar to that of free space,  $\mu_0$ . In addition to the large physical property contrast, the  
462    geometry of well casing also presents a significant computational challenge. Well casing is cylindrical in shape and  
463    only millimeters thick, while the geologic structures we aim to characterize are on the scale of hundreds of meters to  
464    kilometers. Inverting electromagnetic data from such settings requires that we have the ability to accurately simulate  
465    and compute sensitivities for models with casing and 3D geologic variations. One strategy that may be considered  
466    is using a primary- secondary approach, simulating the casing in a simple background and using these fields to  
467    construct a source for the secondary problem which considers the 3 dimensional structures of interest (?). Here, we  
468    demonstrate how the framework can be employed to implement this approach and compute the sensitivities. The  
469    parametric representation of the model allows us to investigate the expected data sensitivity to specific features of the  
470    model such as the location, spatial extent and physical properties of a geologic target. Such an analysis may be used to  
471    investigate how well we expect certain features of the model to be resolved in an inversion and it could be employed  
472    as a survey design tool. In what follows, we outline the general approach and then discuss a specific implementation.  
473    The script used to generate this example is available at: <https://doi.org/10.6084/m9.figshare.5036123>.

474    5.3.1. Approach

475    In this example we design a survey to resolve a conductive body in a reservoir layer in the presence of a vertical,  
476    steel-cased well as shown in Figure ???. To calculate the sensitivity of the data with respect to each model parameter  
477    requires that we be able to simulate and calculate derivatives of each component used to simulate data.

We use a primary-secondary approach, as described in ?. The physical properties, fields and fluxes are composed of two parts, a primary and a secondary part. For example in the E-B formulation,  $\sigma = \sigma^P + \sigma^S$ ,  $\mu = \mu^P + \mu^S$ ,  $\vec{E} = \vec{E}^P + \vec{E}^S$ ,  $\vec{B} = \vec{B}^P + \vec{B}^S$ . A primary problem, which includes the cylindrically symmetric part of the model (casing, source, and layered background) is defined

$$\begin{aligned}\vec{\nabla} \times \vec{E}^P + i\omega \vec{B}^P &= 0 \\ \vec{\nabla} \times \mu^{-1} \vec{B}^P - \sigma^P \vec{E}^P &= \vec{s}_e.\end{aligned}\tag{24}$$

This primary problem is solved on a cylindrically symmetric mesh with cells fine enough to capture the width of the casing and its solution yields the primary fields. The primary fields are then interpolated to a 3D tensor mesh, suitable for discretizing 3D reservoir-scale features. The primary fields are used to construct the source current density for the

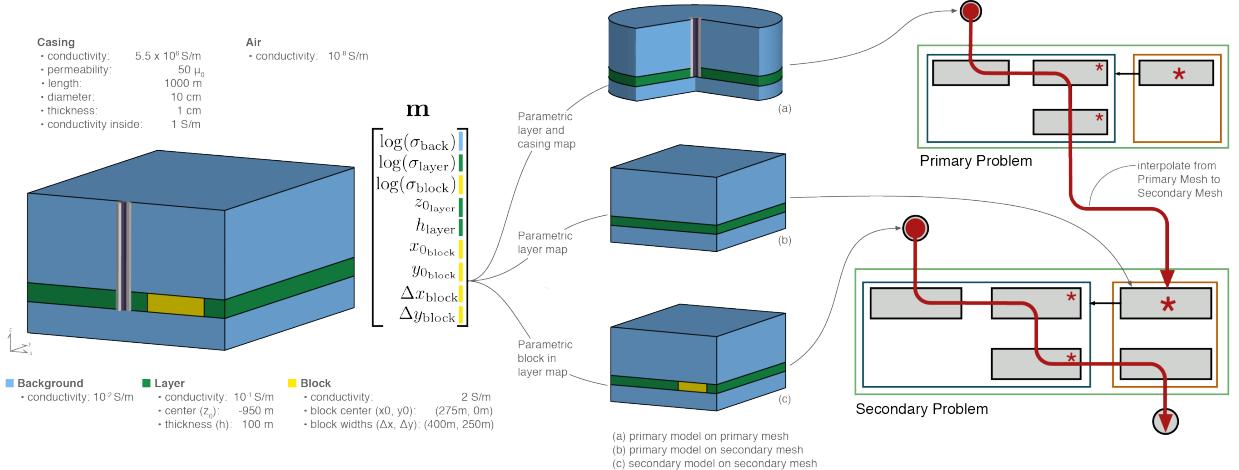


Figure 11: Setup of parametric models and calculation of the sensitivity for a primary secondary approach of simulating 3D geology and steel casing.

secondary problem, given by

$$\begin{aligned} \vec{\nabla} \times \vec{E}^S + i\omega \vec{B}^S &= 0 \\ \vec{\nabla} \times \mu^{-1} \vec{B}^S - \sigma \vec{E}^S &= \vec{q} \\ \vec{q} &= (\sigma - \sigma^P) \vec{E}^P. \end{aligned} \quad (25)$$

478 By solving the secondary problem, we then obtain secondary fields and fluxes. These are sampled by the receivers to  
 479 create predicted data.

480 In equation ??, we see that the source term,  $\vec{q}$  has model dependence through  $\sigma$ ,  $\sigma^P$  and  $\vec{E}^P$ . Typically primary-  
 481 secondary approaches are used when the background is assumed to be known, as it is captured in the primary. Here,  
 482 however, we do not wish to assume that the background is known; in practice it may be constrained, but it is not  
 483 generally well known. The primary solution is used instead to separate the contributions of the casing and the block  
 484 so that we can avoid a potentially crippling assumption. This approach allows an appropriately tailored mesh to be  
 485 constructed for each problem. Thus, we require derivatives not only on the 3D secondary mesh, but also derivatives of  
 486 the primary fields (in this case on a cylindrically symmetric mesh). To implement this type of primary-secondary prob-  
 487 lem, we construct a Primary-Secondary source which solves the primary problem to provide the primary fields. Since  
 488 all derivatives are implemented for the primary problem, when computing sensitivities for the secondary problem, the  
 489 derivatives due to the primary problem are accounted for in the contributions of the source term to the derivative. This  
 490 is conceptually shown in Figure ??.

491 For this example, we wish to investigate how sensitive the specified survey is to aspects of the model which we  
 492 might want to resolve in a field survey, such as the geometry and location of the anomalous body, as well as the  
 493 physical properties of the geologic units. A voxel-based description of the model does not promote investigation of

494 these questions, so we will instead apply a parametric description of the model. The model is parameterized into nine  
 495 parameters which we consider to be unknowns ( $\log(\sigma_{\text{background}})$ ,  $\log(\sigma_{\text{layer}})$ ,  $\log(\sigma_{\text{block}})$ ,  $z_{0,\text{layer}}$ ,  $h_{\text{layer}}$ ,  $x_{0,\text{block}}$ ,  $\Delta x_{\text{block}}$ ,  
 496  $y_{0,\text{block}}$ ,  $\Delta y_{\text{block}}$ ). In what follows, we examine the sensitivity of the data with respect to these model parameters.

497 *5.3.2. Implementation*

498 The model we use is shown in Figure ???. It consists of a 1km long vertical steel cased well (diameter: 10 cm,  
 499 thickness: 1cm) with conductivity  $\sigma = 5.5 \times 10^6$  S/m, and magnetic permeability  $\mu = 50\mu_0$ . The casing is assumed  
 500 to be filled with fluid having a conductivity of 1S/m. The background has a resistivity of  $100\Omega\text{m}$ , and the 100m thick  
 501 reservoir layer has a resistivity of  $10\Omega\text{m}$ . The target of this survey is the conductive block (2S/m) with dimensions  
 502  $400\text{m} \times 250\text{m} \times 100\text{m}$ . The source used consists of two grounded electrodes, a positive electrode coupled to the casing  
 503 at a depth of 950m, and a return electrode 10km from the wellhead on the surface. We consider a frequency-domain  
 504 experiment at a transmitting frequency of 0.5Hz and 1A current. For data, we consider two horizontal components ( $x$   
 505 and  $y$ ) of the real part of the electric field measured at the surface.

506 To accomplish this simulation and sensitivity calculation, we construct 3 mappings, shown conceptually in Figure ???, in order to obtain: (1)  $\sigma^P$  on the primary (cylindrical) mesh, (2)  $\sigma^P$  on the secondary mesh (as is needed in equation ???) and (3)  $\sigma$  on the secondary mesh. Differentiability of the electrical conductivity models with respect to each of the 9 parameters is achieved by constructing the model using arctangent functions (cf. ??). Each of these parameterizations can be independently tested for second-order convergence to check the validity of the computation of the derivatives (cf. ?).

512 The source term for the secondary fields requires that we simulate the primary fields. For this, we use the mapping  
 513 of  $\mathbf{m}$  to  $\sigma^P$  on the primary mesh and employ the H-J formulation of Maxwell's equations in the frequency domain in  
 514 order to describe a vertically and radially oriented current density and a rotational magnetic field. In this simulation,  
 515 we also consider the permeability of the casing. The source consists of a wire-path terminating downhole at -950m  
 516 where it is coupled to the casing. At the surface, the return electrode is 10km radially away from the well<sup>3</sup>. With  
 517 these parameters defined, we have sufficient information to solve the primary problem and thereby obtain the primary  
 518 electric field everywhere in the simulation domain. The real, primary current density for this example is shown in  
 519 Figure ???.

520 This primary field is described on the cylindrical mesh, so in order to use it to construct the source term for the  
 521 secondary problem, we interpolate it to the 3D tensor mesh. The remaining pieces necessary for the definition of the  
 522 secondary source on the 3D mesh are defining  $\sigma$  and  $\sigma^P$ ; this is achieved through the mappings defined above. The  
 523 primary problem and source, along with the mapping required to define  $\sigma^P$ , are used to define a primary-secondary  
 524 source, which solves a forward simulation to compute the secondary source-current,  $\mathbf{s}_e$ , shown in Figure ???. Note that

---

<sup>3</sup>Due to the symmetry employed, the return electrode is a disc. Numerical experiments over a half-space show that the real, radial electric field from the cylindrical simulation exhibits the same character as the 3D simulation but is slightly reduced in magnitude.

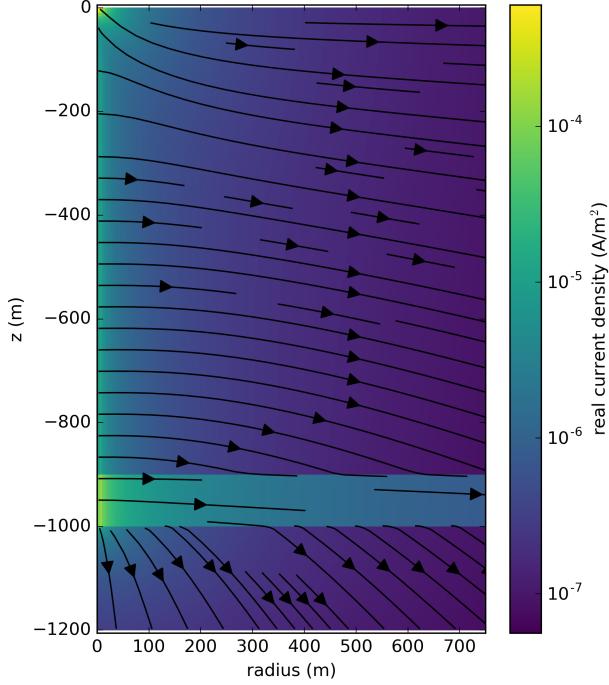


Figure 12: Cross sectional slice of primary (casing + background) real current density. The colorbar is logarithmically scaled and shows the amplitude of the real current density.

525 the source current density is only present where there are structures in the secondary model that were not captured in  
 526 the primary, in this case, where the conductive block is present.

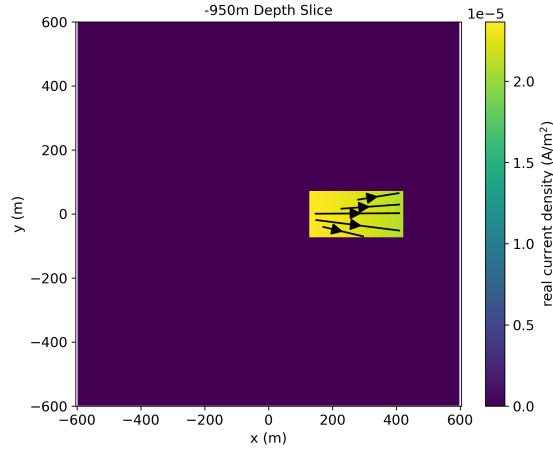


Figure 13: Depth slice at  $z = -950\text{m}$  showing the source current density for the secondary problem.

527 With the source term for the secondary problem defined, the secondary problem is then solved resulting in the  
 528 predicted data at the surface. Here, we focus our attention to the real  $x, y$  components of the electric field, as shown

529 in Figure ???. The top two panels show the total (casing and conductive target) x-component (a) and y-component (b)  
 530 of the electric field while the bottom two panels show the secondary (due to the conductive target, outlined in white)  
 531 x-component (c) and y-component (d) of the electric field. As expected, the total electric field is dominated by the  
 532 source that is located in the casing. As shown in Figure ?? the majority of the current is exiting into the layer at depth,  
 533 but current is still emanating along all depths of the casing. Measured electric fields at the surface are sensitive to the  
 534 currents that come from the top part of the casing and hence the observed fields are strongest closest to the pipe and  
 535 they fall off rapidly with distance. The behavior of the secondary electric field is, to first order, like that expected from  
 536 a dipole at depth oriented in the x-direction. It has a broad smooth signature at the surface.

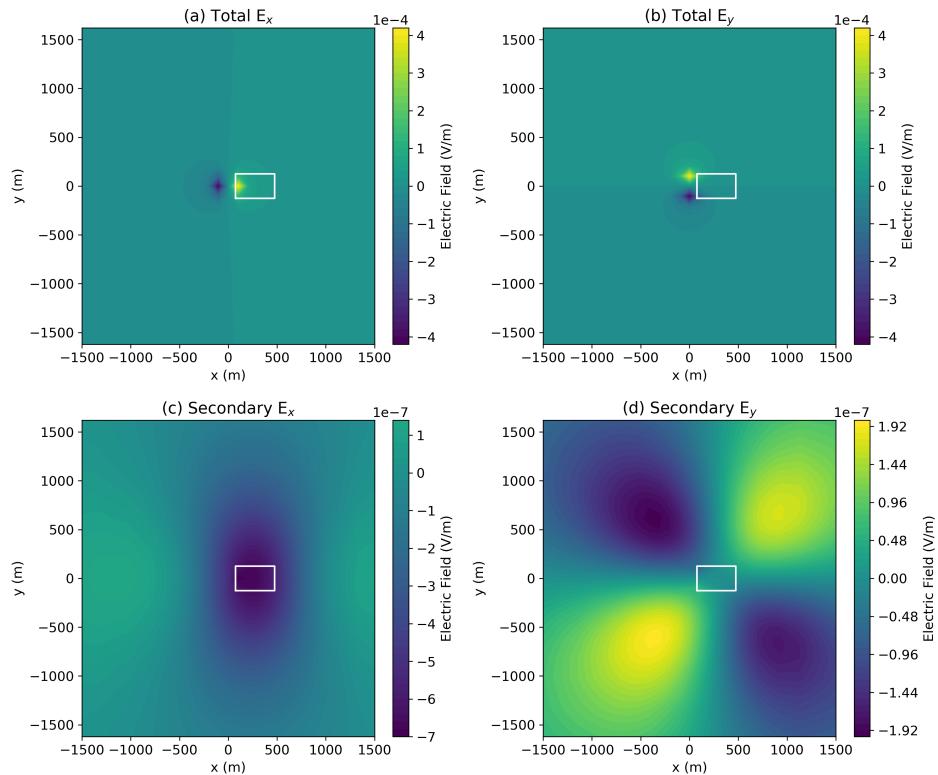


Figure 14: Simulated real electric field data as measured at the surface using a primary secondary approach for casing and a conductive target (outlined in white). The upper panels show the total  $E_x$  (a) and  $E_y$  (b); the lower panels show the secondary (due to the conductive block)  $E_x$  (c) and  $E_y$  (d). Note that the colorbars showing the secondary electric fields are not on the same scale. The limits of the colorbars have been set so that the zero-crossing is always shown in the same color.

537 Now that the pieces are in place to perform the forward simulation, we want to compute the sensitivity. Generally,  
 538 we do not form the full sensitivity when performing an inversion as it is a large, dense matrix. Here however, since the  
 539 inversion model is composed of only nine parameters, the final sensitivity matrix is small (nine by number of data).  
 540 The steps followed to stitch together and compute the sensitivity are shown in the diagram in Figure ???. To check the  
 541 simulation approach for this example, the sensitivity is tested for second-order convergence (cf. ?).

542 Figures ??, ?? and ?? shows the sensitivity of both the real  $E_x$ (left), and real  $E_y$  (right) data with respect to each  
 543 of the 9 model parameters. Note that the colorbars are not identical in each image and the units of the sensitivity are  
 544 dependent on the parameter under consideration. In each image, the white outline shows the horizontal location of the  
 545 block.

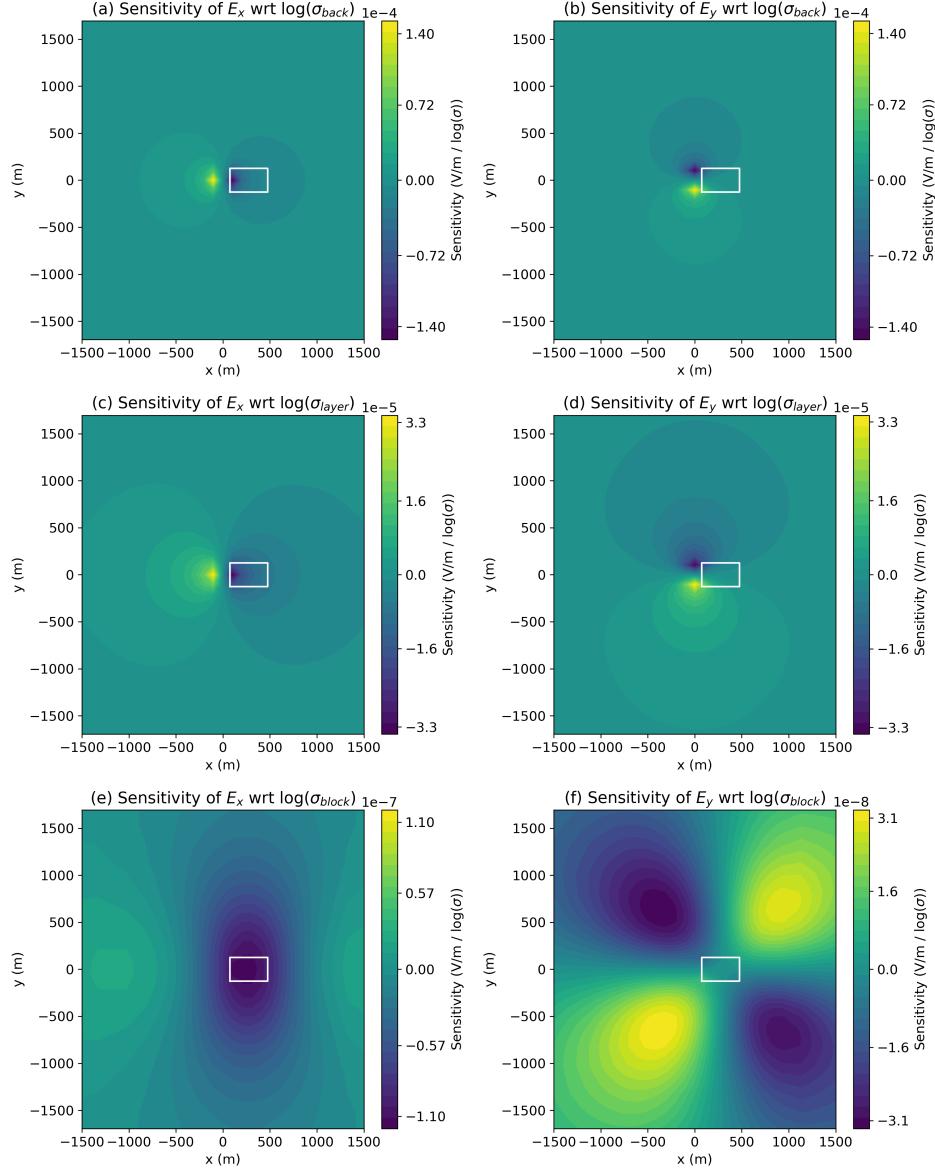


Figure 15: Sensitivity of surface real  $E_x$  (left) and  $E_y$  (right) data with respect to the physical properties,  $((V/m)/(\log(\sigma)))$

546 In Figure ??, we focus on the physical properties of the background layer and block, all parametrized in terms of  
 547  $\log(\sigma)$ . Clearly, the conductivity of the background has the largest influence on the data, in particular near the well  
 548 (at the origin), followed by the conductivity of the layer, where the injection electrode is situated. There are 4 orders

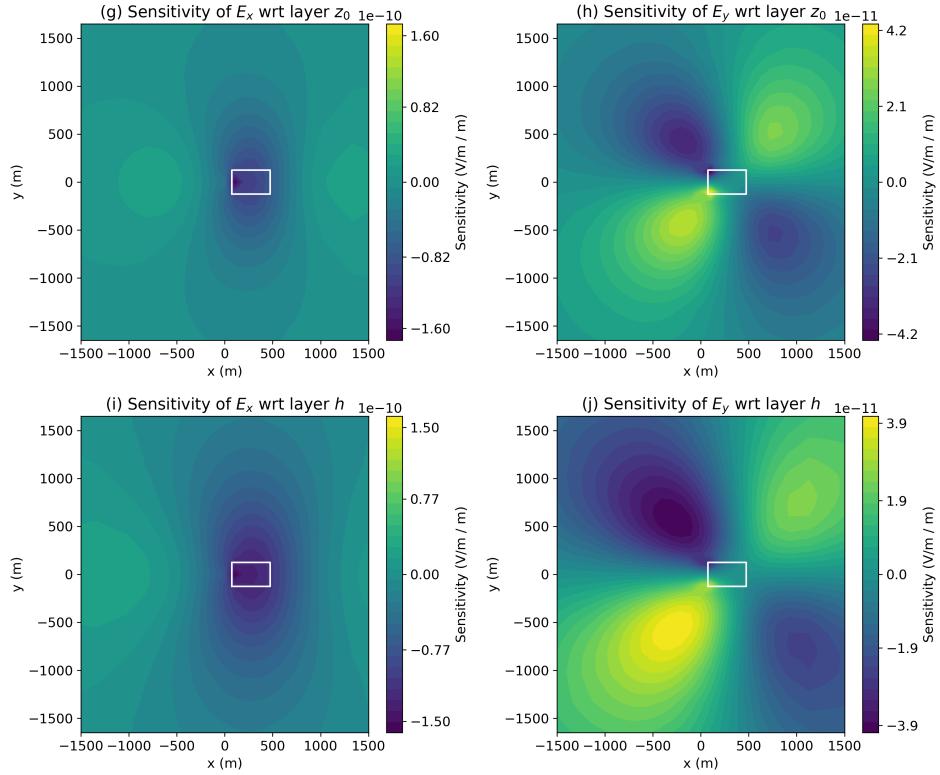


Figure 16: Sensitivity of surface real  $E_x$  (left) and  $E_y$  (right) data with respect to the layer geometry,  $((\text{V}/\text{m})/\text{m})$

of magnitude difference between the maximum sensitivity of the data with respect to the conductivity of the block and that of the background. This indicates that in order to resolve such an anomalous body, the background must be well-constrained. When looking at Figure ?? (f), we see that the areas of largest sensitivity of the  $E_y$  data with respect to the physical properties of the block are spatially distant from the body and the well. This indicates that if one is designing a survey, it may be advantageous to collect data in these regions as these are also regions where the influence of the properties of the background are less dominant.

In Figure ??, we focus on the depth and thickness of the layer. Note that the depth and thickness of the block are constrained to be the same as the layer, so the character of the sensitivity is influenced by the presence of the block. Here, the units of the sensitivity are  $(\text{V}/\text{m})/\text{m}$ . Similarly, Figure ?? shows the sensitivity with respect to the geometric properties of the block.

To compare between the physical properties and geometry of the model, the scales of interest must be taken into consideration. In Table ??, we show the maximum amplitude of the sensitivity with respect to each individual model parameter. From this, we approximate the sensitivity as linear about the true model and compute the perturbation required to cause a change of  $10^{-9} \text{ V}/\text{m}$  in the data ( $\Delta\mathbf{m}_i = 10^{-9} / \max |\mathbf{J}_i|$ ). For ease of comparison, the perturbations

in the log-conductivity of the background, layer, and block were converted to linear conductivity by

$$\Delta\sigma_{\text{unit}} = \frac{\exp[\log(\sigma)_{\text{unit}} + \Delta \log(\sigma)_{\text{unit}}] - \exp[\log(\sigma)_{\text{unit}} - \Delta \log(\sigma)_{\text{unit}}]}{2}. \quad (26)$$

559 In table ??, we see that to cause a perturbation in the  $E_x$  data by  $\sim 10^{-9} \text{ V/m}$ , requires a 0.007% change in  
 560 the conductivity of the background, while the conductivity of the block would need to change by 0.8% to have a  
 561 comparable impact in the  $E_x$  data. In comparing between physical properties and geometric features of the model, we  
 562 see that a change in the conductivity of the block by 0.8% has a similar impact in the  $E_x$  data as moving  $x_0$  of the  
 563 block by  $\sim 16 \text{ m}$ . For a change in  $y_0$  of the block to have a comparable impact in the  $E_x$  data would require that it  
 564 be perturbed by  $\sim 85 \text{ m}$ . However, the  $E_y$  data are more sensitive to  $y_0$ ; a perturbation of  $\sim 24 \text{ m}$ , about 1/3 of that  
 565 required in the  $E_x$  data, would result in a  $\sim 10^{-9} \text{ V/m}$  change in the measured responses.

parameter $\mathbf{m}_i$	Units of Sensitivity, $\mathbf{J}_i$	max $ \mathbf{J}_i $ wrt $E_x$	perturbation required to cause $\pm 10^{-9} \text{ V/m}$ in $E_x$	max $ \mathbf{J}_i $ wrt $E_y$	perturbation required to cause $\pm 10^{-9} \text{ V/m}$ in $E_y$
$\log(\sigma_{\text{back}})$	( $\text{V}/\text{m}$ ) / $\log(\sigma)$	1.5e-04	6.6e-08 $S/\text{m}$ (6.6e-04%)	1.5e-04	6.6e-08 $S/\text{m}$ (6.6e-04%)
$\log(\sigma_{\text{layer}})$	( $\text{V}/\text{m}$ ) / $\log(\sigma)$	3.5e-05	2.9e-06 $S/\text{m}$ (2.9e-03%)	3.4e-05	2.9e-06 $S/\text{m}$ (2.9e-03%)
$\log(\sigma_{\text{block}})$	( $\text{V}/\text{m}$ ) / $\log(\sigma)$	1.2e-07	1.7e-02 $S/\text{m}$ (8.4e-01%)	3.3e-08	6.1e-02 $S/\text{m}$ (3.1e+00%)
$z_0_{\text{layer}}$	( $\text{V}/\text{m}$ )/ $m$	1.7e-10	5.8e+00 $m$	4.4e-11	2.3e+01 $m$
$h_{\text{layer}}$	( $\text{V}/\text{m}$ )/ $m$	1.6e-10	6.2e+00 $m$	4.1e-11	2.4e+01 $m$
$x_0_{\text{block}}$	( $\text{V}/\text{m}$ )/ $m$	6.2e-11	1.6e+01 $m$	1.8e-11	5.6e+01 $m$
$y_0_{\text{block}}$	( $\text{V}/\text{m}$ )/ $m$	1.2e-11	8.5e+01 $m$	4.2e-11	2.4e+01 $m$
$\Delta x_{\text{block}}$	( $\text{V}/\text{m}$ )/ $m$	4.8e-11	2.1e+01 $m$	1.5e-11	6.6e+01 $m$
$\Delta y_{\text{block}}$	( $\text{V}/\text{m}$ )/ $m$	1.4e-11	7.3e+01 $m$	6.5e-12	1.5e+02 $m$

Table 1: Comparison of the maximum amplitude of the sensitivity with respect to each model parameter, and the approximate perturbation in that parameter required to produce a  $10^{-9} \text{ V/m}$  change in the measured data. The conversion from a perturbation in log-conductivity to conductivity is given by equation ???. The perturbation in conductivity is also provided in terms of a percentage of the true model conductivity.

566 Examining the nature of the sensitivity with respect to parameters describing the target of interest provides insight  
 567 both into how one might design a survey sensitive to the target, and how well we may be able to resolve various geo-  
 568 metric features or physical properties in the model. For the example shown here, we see that it may be advantageous  
 569 to collect data away from the well and hundreds of meters offset from the block. These are regions where both the  
 570  $E_x$  and  $E_y$  data have high sensitivity to features of the target and are distant from the steel-cased well, where we have  
 571 the highest sensitivity to the background. Thus, data collected in these regions may improve our ability to resolve the  
 572 target of interest. The parametric definition of the model provides a mechanism for examining how well we might  
 573 expect to resolve various aspects of the target, such as its spatial extent. There are clearly further questions that may be  
 574 investigated here, including exploring survey parameters such as the impact of varying the frequency on our ability to  
 575 resolve the block, or performing the same analysis for a time-domain survey. A modular framework, with accessible

576 derivatives, is an asset for exploring these types of questions.

577 **6. Conclusion**

578 The framework we have laid out has rigorously separated out various contributions to the electromagnetic equa-  
579 tions in both time and frequency domain. We have organized these ideas into an object oriented hierarchy that is  
580 consistent across formulations and attends to implementation details and derivatives in a modular way. The organi-  
581 zation of the EM framework and numerical implementation are designed to reflect the math. The goal is to create  
582 composable pieces such that electromagnetic geophysical inversions and forward simulations can be explored and  
583 experimented with by researchers in a combinatorial, testable manner.

584 We strive to follow best practices in terms of software development including version control, documentation  
585 unit testing, and continuous integration. This work and the SimPEG project are open-source and licensed under the  
586 permissive MIT license. We believe these practices promote transparency and reproducibility and we hope that these  
587 promote the utility of this work to the wider geophysics community.

588 **Acknowledgments**

589 The authors thank Dr. Eldad Haber for his invaluable guidance in the development and implementation of the  
590 framework. We also thank the growing community of SimPEG developers, in particular: Dr. Dave Marchant, Dr.  
591 Brendan Smithyman, Michael Mitchell, Thibaut Astic, Dominique Fournier, and Joseph Capriotti, have contributed to  
592 the SimPEG code-base; we are grateful for your contributions both to the code and the discussions in the development  
593 of SimPEG.

594 Thanks also to CSIRO for making the Bookpurnong data available and to Dikun Yang for conversations on the  
595 inversion of those data, and finally to the three anonymous reviewers whose comments improved the quality of the  
596 paper.

597 The funding for this work is provided through the Vanier Canada Graduate Scholarships Program.

598 **References**

599 ,

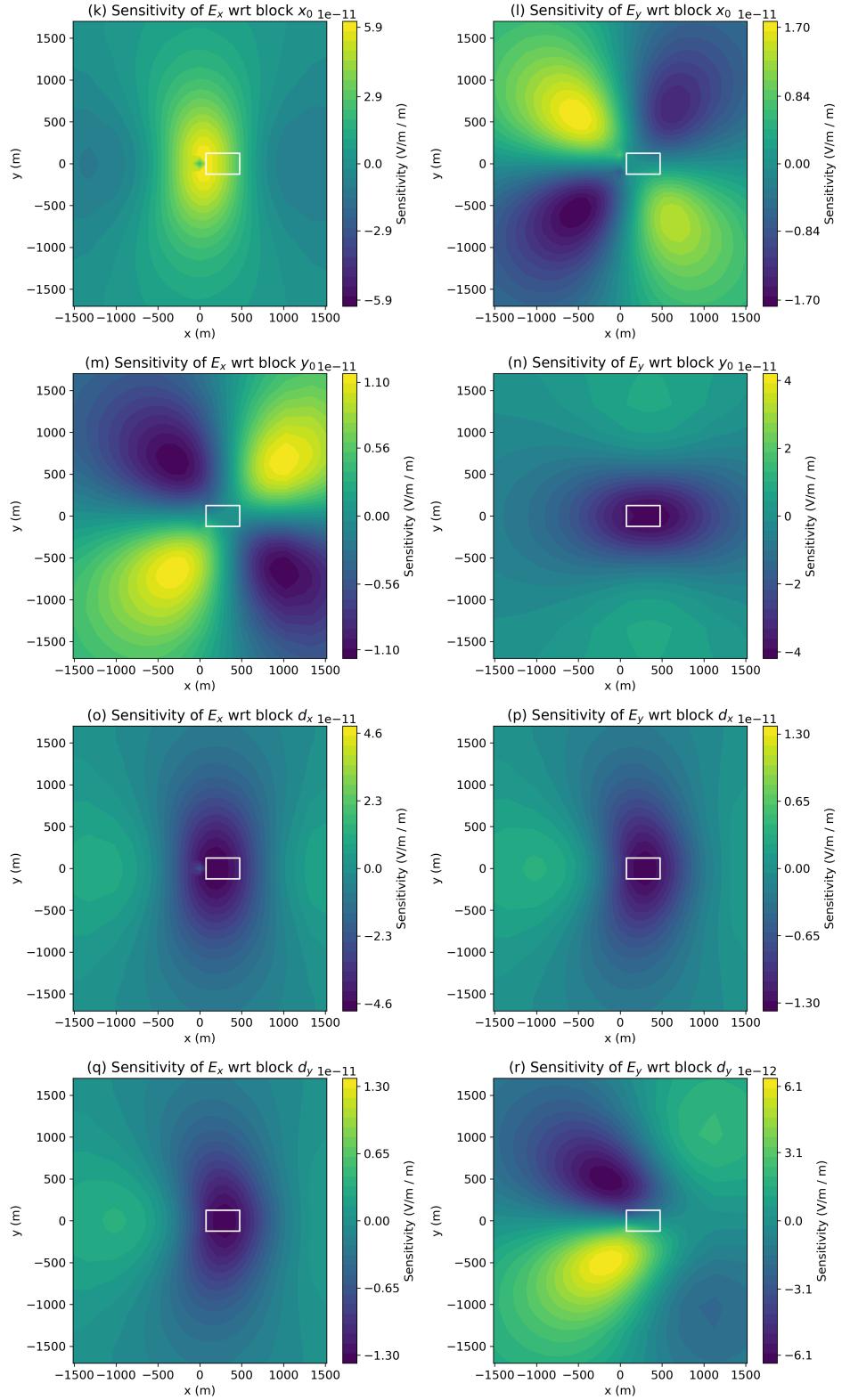


Figure 17: Sensitivity of surface real  $E_x$  (left) and  $E_y$  (right) data with respect to the block geometry,  $((V/m)/m)$