Thesis

Exploring novel algorithms for an improved ERT monitoring workflow

Student ID: 4257589

Registration: 01/10/2015

Expected Thesis Pending: 01/10/2018

Original submission date: 12/08/2021

Corrections submission date: 09/03/2023

Luke Richard Hawley-Sibbett

June 22, 2023

Abstract

Electrical Resistivity Tomography (ERT) is a technique for estimating the resistivity of the subsurface. Long term ERT monitoring is increasingly used to estimate changes in resistivity over time, but data processing workflows are not fully developed. In this thesis, new methods are explored to improve important elements of the workflow: the regularisation of time-lapse inversions, and the detection of data quality anomalies which can be missed by conventional methods.

One of the limitations of ERT inversion is that regularisation methods are often biased towards either sharp or smooth changes in resistivity in both the temporal or spatial domains. A novel alternating-minimisation inversion algorithm is developed, using Total Generalised Variation regularisation, a functional which is able to represent both smooth and sharp changes better than existing methods. Unlike minimum gradient support type approaches, it does not depend on a pre-defined threshold in the smoothing function which limits smooth behaviours to a specific resistivity or gradient range. Comprehensive parameter testing of the algorithm demonstrated that it converges reliably, and outperforms conventional TV and l_2 regularisation functions for a test model when properly configured. However, the solutions are significantly influenced by the initial model. The smooth behaviour is also limited near the model boundaries. Potential limitations of the finite difference approach used in this implementation are discussed, and alternatives are proposed for future improvements.

Existing data quality measures such as reciprocal errors and contact resistance measurements are insensitive to certain sources of noise. This work is motivated by a case study where a damaged cable connector led to short circuiting errors. Short-circuiting errors have the potential to occur at any monitoring installation, as they are a consequence of any physical damage to the cables. A Principal Component Analysis (PCA) control chart method is developed in order to detect the onset of these errors, which went unnoticed until months later. Previous data from the site was used to establish a PCA model. Deviation from the model is measured using a Q^2 statistic. The new method is able to successfully detect the onset of these shorting errors, using both resistivity and reciprocal error data. Therefore, this approach may be used

where reciprocal error data is unavailable. Suggestions are made for the further development of this method though testing with a wider array of data.

Acknowledgements

Thank you to my supervisors, Jon Chambers and Paul Wilkinson, for your help, support and encouragement through difficult times, both personally and professionally.

Thank you to Jon Garibaldi and Tony Pridmore for being generous with their time and guiding me through the end of my project. Thank you also to my original supervisor Bai Li, for her close interest in my work.

Particular thanks to my colleague and friend Jinming Duan, for introducing me to variational image processing, and for many fruitful discussions about my TGV implementation. Likewise, thank you to Wil Ward for taking me under your wing in the early days of my project. Thank you also to my friends and colleagues in the Geophysical Tomography group at the British Geological survey.

Most of all, I am grateful to my wife, Ruby, for supporting me through my illness and beyond. Without you, this work would never have been completed.

Contents

Chapter 1

Introduction

1.1 Background

Electrical Resistivity Tomography (ERT) is a geophysical imaging technique for generating models of the subsurface resistivity. These can be 2D, 3D, or 4D, including both spatial and temporal resistivity changes. Resistivity is related to several ground properties of interest to geoscientists, including geological material, porosity, moisture saturation, temperature and poor fluid compositions, which makes it relevant to a wide range of applications and disciplines including hydrology, civil engineering, contaminated land and mineral exploration [Loke et al., 2013]. Over the last 10-15 years, advances in hardware and computing power have led to increased capabilities and the use of long term ERT monitoring, which allows dynamic resistivity changes to be modelled. Monitoring applications include: hydrological processes e.g.(Johnson et al. [2012]; Kuras et al. [2016]), including those related to landslide stability [Perrone et al., 2012]; geothermal systems [Hermans et al., 2014]; and carbon capture and storage (e.g. Schmidt-Hattenberger et al. [2014]). ERT monitoring provides a different set of challenges to individual ERT surveys. Monitoring requires either a permanent installation or repeated visits to the same site, which is demanding in equipment or manpower. Furthermore, the volume of data which can be generated by a monitoring installation far exceeds that which can be produced by individual survey methods, which strains established data processing methodologies. New, temporal, information is also available to be used for diagnostic purposes. The inversion approach, used to produced resistivity models, should also be adjusted for time-series data. In this thesis, new methods to improve the ERT monitoring processing pipeline are explored.

A typical ERT setup consists of an array of electrodes, usually embedded along the ground surface or in boreholes beneath the surface. The geometry of the array determines the imaging

region. Currents are injected into the ground between pairs of electrodes, and the resulting potential differences are measured between other electrode pairs. These are used to calculate a transfer resistance or an apparent resistivity, using a normalising factor. An inverse problem is then solved to calculate a ground resistivity model from the data, a process often known as inversion. The complementary forward problem, of calculating the resistivity measurements from a known ground resistivity distribution is directly solvable. However, the ERT inverse problem is mixed determined (the model solution is under determined in some regions and over determined in others) [Menke, 1989], and is non-unique and ill-posed, so additional constraints are required in order to regularise the problem. These are usually conditions of spatial smoothness.

The purpose of ERT monitoring is to estimate changes in ground resistivity over time. The magnitude of these can be small, and errors in individual inverted models can result in artefacts of the order of the true resistivity changes. A range of inversion protocols aim to improve the accuracy of this process, by inverting two or more models together and applying constraints between them. The 4D simultaneous inversion approach applies structural regularisation to both the spatial and time domains, which substantially reduces inversion artefacts. These constraints favour either smooth or sharp changes across the entire space or time domain, which can be chosen based on the expected variation of the resistivity model. Where both smooth and sharp changes are observed in the same time series, inaccuracies can result.

A traditional ERT data processing workflow is manually performed by an expert, relying on experience and judgement to determine sensible data quality thresholds and appropriate constraints and parameters for the inversion. Some degree of trial and error may be required to determine the optimal parameters for a given data set. This is informed by data quality measurements and the available geological knowledge for the site. As monitoring data are acquired at daily intervals, or even more frequently, this approach demands prohibitive amounts of manpower to process data in real time, incentivising either a retrospective batch processing approach or, preferably, an automated processing pipeline. A representative monitoring system workflow is shown in figure 1.1 [Chambers, 2021]. Currently, the data analysis and delivery process is not yet fully automated and manual input is still required for data processing.

1.2 Aims and Objectives

The aim of this work was to develop new methods to improve the processing of ERT monitoring data, informed by techniques and technologies used in Computer Science.

The objectives are:

1. Identify new techniques which have the potential to improve the ERT monitoring processing pipeline: specifically the regularisation of time lapse inversions; and new data quality assurance methods utilising the time-dependence of the data.
2. Create a novel Total Generalised Variation algorithm for the inversion of ERT data, initially in the spatial domain.
3. Evaluate the performance of the new algorithm for suitability in ERT inversion; and configure the optimum parameters for different resistivity distributions.
4. Design and test a new data quality method for ERT, capable of the detecting short-circuiting errors in the motivating case study.

1.3 Structure

In Chapter 2, I review the ERT inversion and data quality methods from a monitoring perspective. In the field of image processing and computer vision, several variational approaches have been developed, many of which aim to allow some mixture of smooth and sharp changes. These are used across a range of applications, from image denoising to medical image reconstruction. I explore these methods, and the suitability of their use for improved ERT inversion in the context of the existing work on time-lapse inversion. Another important area of the monitoring pipeline, are the data quality processes. The direct data quality measures used in ERT have limited ability to detect certain type of errors, for example short-circuiting errors, which can lead to a significant loss and degradation of data in a monitoring series if they are not identified quickly. Short-circuiting errors are not unusual, as they result from damage to the monitoring installation cables. Potential anomaly detection data quality frameworks for the detection of these errors are reviewed, and a Statistical Process Control spectral anomaly detection approach is identified as a promising approach.

In Chapter 3, I derive a new algorithm for piecewise affine ERT regularisation, and present the results of preliminary testing to establish the basic behaviour of the algorithm. The initial algorithm is reviewed, and improvements to the model configuration are made.

In Chapter 4, an in-depth parameter exploration evaluates the performance of the new algorithm. The influence of the initial conditions and several parameters involved in the inversion are explored, and general guidelines for the configuration and use of the new regularisation algorithm are developed.

In Chapter 5, a new data quality method for ERT monitoring is developed, using data from

a historical case study as a basis. A Principal Component Analysis based control chart approach is used. This approach aims to identifies short-circuiting errors from their effects on the correlations between the data.

The work is summarised in Chapter 6, and future work is discussed.

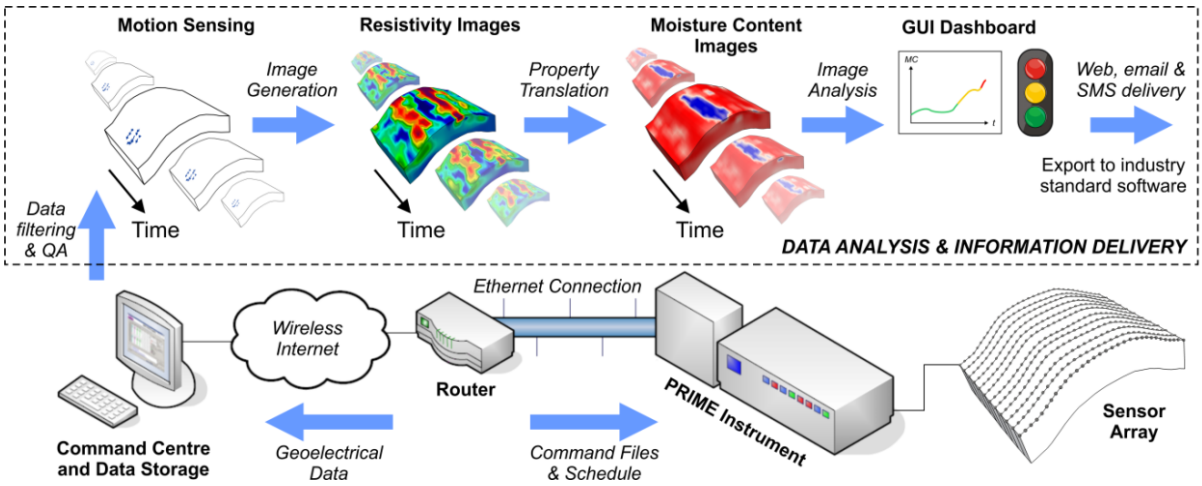


Figure 1.1: Schematic workflow for the PRIME ERT monitoring system, currently in the late stages development at the British Geological Survey. The dotted lines contain the data processing and delivery processes.

Chapter 2

Literature Review

In this chapter, I review the current literature on the regularisation of ERT time-lapse inversion and ERT data quality measurements. I review the Variational Image Processing literature for new regularisation functions, with the potential to deliver improved piecewise smooth regularisation for the inversion of monitoring data. Suitable methods for the detection of unusual data quality issues are explored. The data quality problem is framed as one of anomaly detection. A spectral anomaly detection method derived from the statistical process control literature is identified as a viable candidate for ERT.

2.1 ERT inversion

The process of generating a resistivity model from the ERT data is an inverse problem. The corresponding forward problem, to calculate the ERT data for a given sensor geometry and resistivity distribution, has a unique solution and can be solved numerically. For a point current source at (x_s, y_s, z_s), the potential is given by

$$\nabla . \left[\frac{1}{\mathbf{m}(x,y,z)\nabla\phi(x,y,z)} \right] = -\frac{\partial j}{\partial t}\delta(x_s, y_s, z_s), \tag{2.1.1}$$

where $\mathbf{m}$ is the resistivity, ϕ is the potential and j is the charge density.

There are several approaches to ERT inversion, the most common being the optimisation method, as detailed here, where an objective function is iteratively minimised over a parameter grid, to find a model that is consistent with the data subject to some regularisation constraint.

In this thesis I use an optimisation approach to solve the ERT inverse problem. The complementary forward problem, $f(\mathbf{m})$, can be solved numerically for resistivity distribution, $\mathbf{m}$ [Dey

and Morrison, 1979]. The general form of the inverse problem objective function is

$$\mathbf{\Phi}(\mathbf{m}) = \phi_d(f(\mathbf{m}) - \mathbf{d}) + \lambda\phi_R(\mathbf{m}). \tag{2.1.2}$$

The data misfit term, ϕ_d, is a measure of how well the model fits the data, $\mathbf{d}$; ϕ_R is a regularisation term; and λ is the scalar damping parameter, a Lagrange multiplier, which determines the step length in the inverse solution space.

The non-unique nature of the ERT inverse problem means that many solution models fit the data equally well. It is a mixed-determined problem, where some parts of the model are over-determined and others are under-determined. It is the structural constraints imposed by the regularisation term which stabilise the inversion.

The choice of regularisation function should reflect any available *a priori* information about the expected resistivity distribution. The resulting model should be verified against all available ground truth to ensure the choice of regularisation is reasonable.

Due to noise and modelling errors, parts of the ERT inverse problem are somewhat over-determined. To limit the potential for over-fitting, the inversion is terminated when the solution fits the data to within the level of noise. The root-mean square difference between the data and $f(\mathbf{m})$, or RMS data misfit, is used to assess if the inversion has produced a result which reasonably fits the data. If good measurement error estimates are available, they can be used to weight the data.

The objective function is minimised using an iterative Gauss-Newton approach: starting with an initial model, the objective is repeatedly linearised and solved about the present solution [Menke, 1989]. The value of λ at each iteration can be determined either via a line search targeting a given decrease in data misfit, or else by a 'cooling off' scheme where the parameter is gradually decreased (which can involve re-running the inversion to find a suitable schedule) (Farquharson and Oldenburg [1998]; Karaoulis et al. [2013]).

A structural regularisation approach applied in the widely used form of equation (2.1.2) is

$$\mathbf{\Phi}(\mathbf{m}) = \|\mathbf{W}_d(f(\mathbf{m}) - \mathbf{d})\|_q + \lambda\|\nabla\mathbf{m}\|_t, \tag{2.1.3}$$

where matrix $\mathbf{W}_d$ can be used to weight the data using an error model. $\|.\|_q$ and $\|.\|_t$ are any combination of l_1 and l_2 norms. The regularisation term is a constraint on the structure of the model: l_2 regularisation favours smooth models, while l_1 regularisation favours a piecewise constant model. The $l1$ structural term is known as a Total Variation, or TV, constraint. In

the data misfit term, the l_1 norm is more robust to outliers, and so is better suited to noisy data.

Unlike l_2 norms, the l_1 norm is not directly differentiable. The iteratively reweighted least squares (IRLS) approach approximates l_1 norms as weighted l_2 norms by linearising about the present solution; this can be efficiently carried out simultaneously with the Gauss-Newton iterations [Farquharson and Oldenburg, 1998].

2.1.1 Spatial Regularisation Methods

In optimisation schemes, the regularisation term usually consists of an additional constraint on the model, often a structural constraint minimising some function of the model gradient, such as in equation (2.1.3). The inverted model is ideally the model which minimises those constraints, within the set of possible models which all fit the ERT data equally well. However, this may not be precisely the case, as the IRLS Gauss-Newton method converges on a local solution, which is not necessarily the same as the global solution which fully minimises the objective function.

Here, I review the main approaches to spatial regularisation in the literature, which applies to the spatial component of monitoring data, alongside time constraints.

A conservative approach to regularisation is Occam's inversion [deGroot Hedlin and Constable, 1990], which favours the model with the minimum possible structure consistent with a given data misfit, in the sense of a maximally smooth model gradient. This is achieved by penalising the l_2 norm of the model gradient. These solutions do not necessarily reflect the physical reality best, but limiting the structure in the model solution prevents over-interpretation of features which aren't directly supported by the ERT data, and the smoothness of the solution provides an indication of model resolution. In practice, some degree of site information is generally available, and where the subsurface is known not to be generally smooth, a careful choice of regularisation scheme allows for solutions which are better representative of the physical ground truth.

A well established alternative, which favours piecewise-constant (blocky) models can be achieved by instead penalising the l_1 norm of the model [Farquharson and Oldenburg, 1998]. This is also known as the Total Variation, or TV penalty, which I use here. The choice of constraint is determined by *a priori* information about the resistivity structure [Loke et al., 2003]. However, where the subsurface displays a mix of sharp and gradated changes, neither produces a model which reflects the physical reality well.

Various approaches have been proposed to tackle this problem. Active constraint balancing [Yi et al., 2003] spatially varies an l_2 type smoothness constraint with the model resolution, so that

the smoothing regularisation is weaker where the solution is well constrained by the data. The laterally constrained inversion [Auken and Christiansen, 2004] [Auken et al., 2005] is a limited piecewise smooth inversion scheme, allowing for smooth horizontal variation within a pre-specified number of sharply delineated layers. Blaschek et al. [2008] use a minimum gradient support (MGS) functional, which heavily penalises gradients above a defined threshold, whilst allowing small smooth variations below it. If it is available, information from other imaging modalities can be used to similar effect, by locally alleviating smoothing constraints across known boundaries, e.g. (Doetsch et al. [2012a]; Zhou et al. [2014]; Uhlemann et al. [2017]). In the scheme introduced by Song et al. [2015a], a mix of TV and (smooth) Tikhonov penalties are applied across the entire model, weighted by the average absolute gradient in the image. In Song et al. [2015b], the same authors use a spatially adaptive Total Variation regulariser with a second order term which applies weak TV-like penalties at edge regions, and strong l_2 constraints on flat or smooth regions. Both approaches demonstrably reduce the number of artefacts present in blocky images, but unfortunately the authors do not extend their analyses to cases where gradated resistivity changes are also present.

2.1.2 Time Lapse Inversion

In a monitoring context, changes in resistivity are usually of more interest than absolute values. The simplest approach is to invert each measurement set independently, and take the difference of the model at each time with respect to a baseline, for example [Daily et al., 1992]. As the log resistivity is typically used due to the range of resistivity contrasts, this difference is equivalent to the ratio of the changed model to a baseline.

However, noise in the data can produce artefacts in the models of the order of the true resistivity changes. Noise levels vary with time, which cause variation in the model resolution, and affects the inverted images, resulting in artefacts [LaBrecque et al., 1996].

A range of approaches have been developed to minimise these artefacts whilst remaining sensitive to true resistivity changes. These schemes generally involve inverting two or more time steps together, or else inverting a data difference or ratio between them.

The difference inversion [LaBrecque and Yang, 2001] dampens the prevalence and severity of artefacts by directly inverting the data difference between each individual data set and a baseline data set, using the inverted baseline model as a reference. This has the effect of suppressing systematic errors, and is computationally efficient to apply, comparable to independent inversion. The ratio inversion [Daily et al., 1992] is a related method that instead uses the ratio of the data differences for the same purpose. Difference or ratio inversions are highly dependent on the quality of the original model, and can be difficult to systematically apply over long monitoring periods as the noise profile, and hence the model resolution, changes over time. A

solution to the changing noise profile has been proposed by Lesparre et al. [2017]. The authors find that the differences in both the normal and reciprocal measurements between the baseline and each individual data can be used to fit a resistance dependent error model which captures the changing noise profile.

The cascaded inversion [Miller et al., 2008], involves using an inverted model from a baseline set as a reference model to constrain future time steps. The authors suggest that this may be a useful alternative to the difference inversion where the noise profile changes significantly between the baseline and the time step under observation.

An extension of the minimum support functional in the context of time-lapse inversion is carried out in Fiandaca et al. [2015]. The authors use a difference inversion approach, inverting for the resistivity change against a baseline. A minimum support functional is applied to the spatial distribution of the changes. The minimum support is an s-shaped sigmoid-type functional, such that values below a transition region $\rightarrow 0$ and values above the transition region $\rightarrow 1$. Their model may be tuned using three interdependent parameters, which together determine the resistivity threshold; the width and asymmetry of the transitional region; and the relative magnitude of the data and regularisation terms. If the threshold is set to a small value, the regularisation favours models where the changes are confined or 'focused' to certain regions. The sharper the functional transition region, the sharper the resistivity change at the boundaries of these regions. This flexibility allows the regularisation function to be adjusted based on suspected ground conditions, but it also increases the user complexity, and increases the potential for operator bias in the inversion result. As a difference inversion approach, it is also dependent on the initial baseline reference model, and the optimum tuning parameters may vary if a mix of smooth or sharp edged changes occur. Related approaches utilising the minimum gradient support have also been used [Hermans et al., 2014] [Nguyen et al., 2016].

The 4D time lapse approach first proposed by Kim et al. [2009] inverts the full ERT time series simultaneously, applying regularisation constraints between models in adjacent timesteps. If the subsurface is changing more rapidly than the measurement rate, the the method allows for continuous changes in resistivity during the inversion, although this feature is not retained in some implementations of the 4D time lapse scheme. The authors used an l_2 time constraint in this case, but a range of closely related methods exist. As this type of approach involves inverting every model in the time series simultaneously, it is very computationally resource intensive.

Loke et al. [2014] uses a TV temporal time constraint for 4D inversion. Karaoulis et al. [2013] use an active constraint balancing type approach to vary the strength of l_2 constraints both spatially and temporally, called 4D-ACT inversion. The authors' proposed 4D-ATC allows the

strength of the regularisation term to vary both spatially and temporally, by spatially varying the lagrange multiplier λ, so that areas of active change are not over smoothed. Just as in the 2D active time constraint method [Yi et al., 2003], the resolution matrix needs to be calculated at each iteration of the inversion to identify the areas of focus. For 4D models this is more computationally intensive. Furthermore, the spread of λ values are parameters which need to be independently set prior to the inversion. Kim et al. [2013] also use active constraint balancing, for a mix of l_1 and l_2 constraints, with an automatically determined method of selecting the λ parameters at each stage of the inversion, although this does introduce some additional parameters to initialise the inversion.

In Karaoulis et al. [2013] a range of time lapse regularisation approaches were compared using synthetic models containing sharp temporal resistivity changes. Independent inversion captures true changes well, at the cost of a high rate of false change artefacts. Difference inversion reduces these artefacts, whilst still capturing the true changes. The authors found that l_2 4D inversion almost eliminates false change artefacts, but leads to oversmoothing and reduced sensitivity to sharp changes. However, it would be expected that an l_1 time constraint would be better suited to their data. Loke et al. [2014] finds that l_1 4D inversion contains fewer artefacts and captures sharp resistivity changes in synthetic models better than l_2 4D inversion. Kim et al. [2013] found that different spatio-temporal combinations of TV and l_2 constraints were superior depending on the resistivity structure of the model, and therefore that there was a need to choose the optimal regularisation regime on a model by model basis.

If the resistivity changes under observation occur on a timescale faster than the time taken to acquire a full ERT dataset, then additional uncertainty is introduced to the inversion results, for example in a tracer experiment [Kuras et al., 2009]. Generally, this is less significant in long term monitoring set ups, which tend to monitor changes on a diurnal or even seasonal time scale, although rapid changes may still occur during the time series. As well as the approach of Kim et al. [2009], spline interpolations can be used to interpolate data onto common time steps [Coscia et al., 2012] [Doetsch et al., 2012b].

Throughout the literature, there is a general consensus that choosing a regularisation method which captures the type of resistivity change present can result in a more physically representative model. Difference inversions have good sensitivity to true changes, but are more prone to artefacts than the full 4D time lapse simultaneous inversion approach, as well as being strongly dependent on the baseline data set.

In the 4D approach, a choice between l_2 and TV constraints still needs to be made for each time series, based on the type of change present. However, when both smooth and sharp changes are both present in either the spatial or time dimension, a trade off is involved in the choice.

A 4D approach which is able to capture both smooth and sharp changes in different model regions is highly desirable, and this is one of the key aims of this thesis.

2.2 Variational Image Processing

Variational image processing methods formulate image processing problems as the minimisation of an energy function by the solution of a partial differential equation [Chan et al., 2003]. They are traditionally applied to problems including image denoising, inpainting (interpolating missing pixels), deblurring, texture extraction, and segmentation. Similar methods are also used in a range of imaging techniques, for example in the regularisation of MRI images (e.g. Knoll et al. [2011]). This family of approaches bears a close resemblance to the regularised ERT inverse problem; the primary difference being the form of the data misfit term.

In ERT the data misfit term is highly nonlinear, and is iteratively linearised through a Gauss-Newton solution to the inverse problem. In image processing, the data misfit term is typically a linear difference measure between the solution and the original image, and the regularisation function is responsible for the transformation. These regularisation functions are often much more sophisticated than those used in ERT, however their solution relies on algorithms which are unsuitable where the data misfit term is nonlinear.

The use of Total Variation regularisation for image denoising originated with Rudin et al. [1992]. However, it has some limitations. TV image denoising can lead to a loss of image contrast, as well as a tendency to form stepped contours in the solutions, dubbed 'the staircase effect'. Examples of these can be found in Getreuer [2012].

A variety of higher order variational methods have been developed over the years to address these limitations of TV regularisation. A number of these are reviewed in Lu et al. [2016]. Higher order derivatives such as the Laplacian, $\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}$, and the Bounded Hessian,

$$\nabla^2 u = \begin{bmatrix} \frac{\partial^2 u}{\partial x^2} & \frac{\partial^2 u}{\partial y \partial x} \\ \frac{\partial^2 u}{\partial x \partial y} & \frac{\partial^2 u}{\partial y^2}, \end{bmatrix}$$

feature in many of these. The minimisation of these functionals requires the solution of fourth order nonlinear partial differential equations. Primal-Dual or Split Bregman algorithms are often used to transform the energy minimisation problem into simple subproblems [Lu et al., 2016]. However, these approaches are not well-suited to ERT data, due to the highly non-linear forward model.

The Total Generalised Variation (TGV) functional of Bredies et al. [2010] is a convex higher order generalisation of TV, which favours piecewise affine models based on a local evaluation of the first and second order model gradients. In the TGV functional, higher order derivatives are not calculated directly, but appear as the first derivatives of a new parameter, $\mathbf{p}$, which is related to the model gradient. This is advantageous for use in ERT inversion. In addition, unlike the MGS regularisation, there is no need to choose a pre-determined threshold at which sharp changes occur ub the model and it does not require *a priori* knowledge of layers or boundaries, which are often unavailable. Originally introduced in the context of image denoising (e.g. [Duan et al., 2016]), it has been used extensively in image restoration applications, including the reconstruction of MRI and CT images from under-sampled data (Knoll et al. [2011]; Niu et al. [2014]). The TGV functional is described in detail in Chapter 3, and an algorithm for the second order Total Generalised Variation (TGV) is derived and implemented.

2.2.1 Finite difference methods

The finite difference method is a means of approximating the derivative of a function in a discretised model. In the case of the ERT inverse model, finite differences are used to calculate the local gradient of the resistivity, $\nabla \mathbf{m}$.

The Taylor expansion of a function $f(x)$ in a regular discretised 1D domain, $\mathbf{x}$, about point x_0 is

$$f(x_0 \pm l) = f_n(x_0) \pm l f'(x_0) + f''(x_0)\frac{l^2}{2!} \pm f^{(3)}(x_0)\frac{l^3}{3!} + ..., \tag{2.2.1}$$

where l is the small distance between adjacent cells and the 1st, 2nd and qth derivatives of $f(x)$ are $f'(x)$, $f''(x)$ and $f^{(q)}(x)$ respectively.

The forward difference approximation can be directly obtained by solving equation (2.2.1) for $f'(x_0)$, using the points at x_0 and $x_0 + l$,

$$f'_{(+)}(x_0) = \frac{f(x_0 + l) - f(x_0)}{l} - \Big[\frac{l}{2}f''(x_0) - ...\Big]. \tag{2.2.2}$$

The square brackets separate the higher order terms, which are discarded in the forward difference approximation. From this, it can be seen that the error is approximately $(l/2)f''$. The analogous backward difference,

$$f'_{(-)}(x_0) = \frac{f(x_0) - f(x_0 - l)}{l} - \Big[\frac{l}{2}f''(x_0) - ...\Big]. \tag{2.2.3}$$

can also be used.

The central difference approximation is a closer approximation to f'_n, and has the advantage of being symmetrical. It arises from the difference between $f(x_0 + l)$ and $f(x_0 - l)$,

$$f'_{(cen)}(x_0) = \frac{f(x_0 + l) - f(x_0 - l)}{2l} - \Big[\frac{l^2}{3!}f^{(3)}(x_0) - ...\Big]. \tag{2.2.4}$$

From the truncated terms in square brackets, it can be seen that the error is approximately $\frac{l^2}{6}$. This is an order of magnitude smaller than the forward difference, making it a better approximation to the true derivative. Furthermore, the symmetry of the central difference makes it less dependent on the orientation of the mesh.

Despite the apparent advantages of the central difference, it is poorly suited for use in TV and related methods, as well as in l_2 ERT regularisation, as it is insensitive to thin structures, such as edges.

A second order difference can be calculated by summing equation (2.2.1)$\pm$.

$$f''(x_0) = \frac{f(x_0 + l) - 2f(x_0) + f(x_0 - l)}{l^2} - [\frac{l^2}{12}f^{(4)}(x_0) - ...]. \tag{2.2.5}$$

This is both symmetrical and accurate to $O(l^2)$. The same approximation for the second difference can also be arrived at by a composition of forward and backward differences: $f''(x_0) = (f'_{(-)} \circ f'_{(+)})(x_0)$. As the minimisation of a TV regularisation function involves second order gradient terms, 'forward-backward' difference schemes are typically used in TV regularised problems.

2.2.2 Finite differences in ERT

Several existing ERT inversions utilising unstructured meshes use an approximate measure of the first order difference (eg. Rücker [2010]). The gradient is defined as the difference in resistivity along all cell edges (or faces in 3D) between adjacent parameter cells. This is very similar to the forward differences, but with no adjustment for parameter size or shape. This has the advantage of flexibility, allowing the mesh to be refined near the sensors, model topography or other features; and efficiency due to the use of larger cells in low sensitivity regions. Gradual changes in cell size reduce the size of the errors involved. While the derivatives in this formulation of the smoothness are not strictly equal to the gradients, in practice it proves effective.

2.3 Data Quality and Errors

Time-Lapse electrical Resistivity Tomography (ERT) is a valuable tool for monitoring dynamic subsurface processes [Loke et al., 2013], including hydrology [Brunet et al., 2010], slope stability [Perrone et al., 2012], contaminated land and geothermal systems [Hermans et al., 2014]. Permanent installations are often used for monitoring at important sites (e.g. Schmidt-Hattenberger et al. [2014], Wilkinson et al. [2010]), allowing for regular acquisition of data over a long period of time with a fixed array geometry, and remote data transmission allows processes to be monitored in near real time. Such installations typically consist of an array of electrodes and cables, down boreholes or buried just beneath the surface to avoid damage, connected to a multichannel resistivity meter with a data log and transmission system. Each individual resistivity measurement utilises multiple electrodes; a pair of electrodes are used to transmit current, whilst the potential is measured between other receiver pairs. A single survey will consist of 100-100000 measurements, and these will be repeated at regular intervals throughout the monitoring period.

In such long running experiments, ongoing data quality control is extremely important, as prompt detection of data deterioration allows for remedial works to be carried out and can prevent the loss of months of measurement data. There are many potential sources of noise and system error in ERT, many of which can to develop over time. Polarisation errors arise when an electrode is used to measure potential soon after being used to transmit current [Dahlin, 2000]. Cable and electrode connections are weak spots; while outright disconnections are easy to detect, poor connectivity can cause many problems. Current leakages can occur at damaged cable insulation or connectors, causing additional electrodes to be involved in the measurements in the resulting electrical shorting (e.g. Mitchell and Oldenburg [2016]). Individual electrodes become noisy due to poor galvanic contact with the ground. This can vary over time, and is sensitive to changes in ground conditions such as water content, ionic concentration and temperature, as well as ongoing electrode corrosion [LaBrecque and Daily, 2008]. Instrumental malfunction can lead to low or noisy signal due to problems with current injection, or the contamination of measurements with spurious potentials. Other sources of noise include induced polarisation effects and external currents (both natural telluric currents and power cables).

Data quality control in ERT tends to rely on measurement derived uncertainty estimates and direct contact resistance measurements, using empirically derived thresholds to exclude poor data in conjunction with manual inspection of error distributions and measurement parameter cross plots. However, not all potential error sources are captured by these measures. These methods are discussed in more detail in Section 5.2. However, in the case of monitoring data, the time series of resistivity and error data contains additional information which can be ex-

ploited for data quality purposes. In section 2.4, I review the anomaly detection approaches for detecting errors in time series data.

2.4 Anomaly detection methods

Anomaly detection is the process of detecting rare, unexpected patterns within data, which deviate from expected behaviour [Chandola et al., 2009]. Typically, this approach involves a 'normal' system state, which is known or can be derived from either theory or existing data. The characteristics of any potential anomaly are unknown, and do not need to be defined before they are observed. Instead, the anomaly is identified by a deviation from normal behaviour.

Variation in ambient noise levels is expected over the course of any ERT monitoring period. However, there are many ways that data quality can deteriorate unpredictably, whether due to hardware damage, software errors, or changes to the local environment which may contaminate the data. These are all anomalies in the anomaly detection paradigm. An anomaly detection paradigm for ERT aims to detect these changes by establishing a measure of a 'normal' system state, and detecting deviations from it.

The nature of ERT data means that anomaly detection techniques are best applied to a single times series, as ERT data from different sites are hard to generalise. ERT surveys differ significantly from one another in experimental geometry, noise profile, ground resistivity, and resistivity change patterns. Therefore, the amount of data required for a representative training set is prohibitive, and an anomaly detection method for ERT must instead use data from an ongoing time series.

Anomaly detection methods can be *supervised* or *unsupervised*. *Supervised* methods require labelled examples of both normal and anomalous data from the system in question. However, anomalies are, by their very nature, rare, and representative anomalies cannot be expected to appear in the training data, so fully *supervised* methods are not appropriate. This suggests a focus on *unsupervised* methods. On the other hand, a *semi-supervised* approach, which assumes that all the data in the training set is normal, is also valid, particularly when conventional data quality measures can be used to filter the normal data.

Unsupervised and *semi-supervised* anomaly detection techniques can be broken down into broad categories [Chandola et al., 2009]. Information theoretic techniques detect anomalies through changes in the information content in data, using measures such as entropy. Clustering approaches define normal data as lying in a cluster defined by a clustering algorithm, and anomalies as data lying outside of, or on the edge of, these clusters. Statistical anomaly detection approaches fit a model to the normal data, using a statistical inference test to identify anomalous

data which lie outside of the model. Spectral techniques project data into a lower dimensional subspace, and relies on the assumption that the anomalous data appears significantly different in that subspace.

ERT monitoring data typically consist of $10^2 - 10^5$ resistivity measurements. These are generally highly correlated, due to overlaps in the sensitivity regions of different quadropoles. As such, multivariate data processing and dimensionality reduction techniques will be required before many anomaly detection techniques can be applied. Therefore, spectral techniques seem particularly promising for ERT anomaly detection.

2.4.1 Statistical Process Control

The field of Statistical Process Control (SPC), sometimes known, more generally, as Statistical Process Monitoring, is concerned with monitoring time series data in many contexts, including industrial, economic and IT processes. In this field, control charts are developed in order to identify when a system is deviating from typical behaviours, which is a form of anomaly detection. This framing is attractive for a data quality approach, as control charts allow a small number of parameters to be tracked over time.

Multivariate SPC methods have been developed to handle high-dimensional, time dependent-data. Principal Component Analysis (PCA) based methods project this data on a lower dimensional subspace, where monitoring approaches are applied [De Ketelaere et al., 2015]. In the anomaly detection framework, these are spectral anomaly detection techniques. When working with non-stationary data, moving windows or incremental down-weighting of older data are often used in the PCA process. However, ERT monitoring time series are relatively short and sparsely sampled, and often experience repeating or cyclical weather or seasonal based changes, which suggests that there is value in retaining all previous data which captures normal conditions. However, the non-stationarity of ERT data means that careful attention must be paid to the control chart thresholds and updating the model for any PCA-based control chart method.

As ERT data are highly correlated, spectral anomaly detection techniques similar to those used in SPC are a good candidate for use in ERT data quality purposes, provided that the data quality anomalies are separable from normal behaviour in the PCA subspace.

2.5 Summary

ERT monitoring is a valuable tool in a wide range of applications. However, inversion and data quality methods are not fully adapted to the new challenges that monitoring installations present.

A major limitation of time-lapse ERT inversion schemes is that common regularisation methods favour either smooth or blocky constraints uniformly across all time-steps and spatial dimensions, when in practice a mix of gradual and fast onset changes occur. Existing attempts tend to favour sharp or smooth changes. Some schemes, such as the 4D-ATC, aim to dampen smoothing effects where fast changes are occurring, and MGS approaches allow for smoothing behaviours to occur around a specified resistivity threshold, in an otherwise sharp model. After reviewing a range of regularisation schemes used in variational imaging, a Total Generalised Variation scheme is identified as most promising for application to ERT inversion. As a convex higher order functional containing only first order derivatives of $\mathbf{m}$ and a 'pseudo' gradient parameter, $\mathbf{p}$, the TGV functional is well suited to the iterative linearisation schemes which are necessitated by the nonlinear ERT data term. The TGV trade off parameter between sharp and gradated changes is not dependent on a specific rate of resistivity change, unlike minimum gradient norm type solutions, and allows for sharp or smooth changes of various magnitudes to coexist in the same model.

Regardless of the choices of regularisation schemes, the ERT problem remains ill-posed and highly sensitive to anomalous data. Even the current best practice ERT data quality methods are insensitive to certain types of error. Using the additional information contained within time series data, an anomaly detection method may be suitable for detecting changes in the data which are indicative either of data errors, or else of significant resistivity changes. Due to the large number of variables in a typical ERT setup, a data-reduction approach is valuable. A control chart PCA method, based on those used in Statistical Process Control literature, seems promising for this application, due to the utilisation of the correlations between resistivity measurements, and for the interpretability of the results.

Chapter 3

Total Generalised Variation Inversion

The Total Generalised Variation (TGV) functional was identified by the literature review as a promising method for capturing both slow and rapid changes in the time series inversion of ERT monitoring data. It also has the potential to better capture both smooth and sharp spatial variations in ERT models.

In this chapter, the second order TGV functional is introduced in Section 3.1, and an inversion scheme is derived and implemented in Section 3.2, based on the existing inversion code IP4DI [Karaoulis et al., 2013] . An initial evaluation of the implementation is performed using synthetic models, and the results are compared to the widely used l_2 and TV regularisations. The aim of these initial tests is to evaluate the viability of using TGV regularisation for ERT inversion. While the eventual goal is to apply this regularisation to time lapse inversion, the initial implementation is limited to 2D (with a 2.5D forward model) in order to reduce the computational run time and simplify the analysis. After the initial tests were performed, the findings were used to inform a review of the methodology in 3.4. Following updates to the algorithm and testing methodology, an in-depth analysis, including parameter testing, was carried out in Chapter 4.

3.1 The Total Generalised Variation approach

A Total Generalised Variation (TGV) functional approach to ERT regularisation has the potential to capture both sharp and gradual transitions in resistivity, both spatially and in 4D (time-lapse) inversions. Specifically, the piecewise affine TGV functional is given by

$$\mathrm{TGV}_{\mu}(\mathbf{m}) = \min_{\mathbf{p}} \|\nabla\mathbf{m} - \mathbf{p}\|_1 + \mu\|\frac{1}{2}(\nabla\mathbf{p} + \nabla\mathbf{p}^T)\|_1, \tag{3.1.1}$$

where μ is a fixed scalar parameter. The parameter $\mathbf{p}$ controls whether the regularisation locally favours TV-like behaviour or seeks a constant gradient, and is determined by the underlying resistivity distribution.

Intuitively, the behaviour of the TGV functional can be understood by considering the behaviour of the functional at edges and linearly smooth regions. In smooth regions, the second order resistivity gradient, $\nabla^2\mathbf{m}$, is constant. Therefore, the TGV functional is minimised locally when $\mathbf{p} = \nabla\mathbf{m}$, as both terms in the functional vanish. At sharp edges, $\nabla^2\mathbf{m} > \nabla\mathbf{m}$, and the TGV functional is minimised when $\mathbf{p} \to 0$, reducing to the familiar TV penalty.

As in the MGS (and other higher order regularisation approaches), an extra parameter, μ needs to be specified. μ controls the relative weighting of the first and second order terms, and so can bias the TGV function towards smooth or sharp edges. An empirical value of $\mu = 2$ is typically used for image processing and reconstruction (e.g. [Knoll et al., 2011]), where edges are of primary interest, whilst outperforming TV in smooth regions. By contrast, both smooth and sharp features are of interest in ERT models. Unlike the absolute gradient threshold in MGS, μ is a relative measure, so TGV performs well over a range of contrasts.

The form of the TGV functional has several advantages for ERT regularisation. The symmetrised derivative in the second term ensures the TGV regularised problem is convex. It is rotationally and translationally invariant, and is therefore independent of mesh orientation. The formation of the TGV functional as first order in two variables is highly beneficial, as it limits both the linearisation of the forward problem and the expensive forward modelling calculations to a first order subproblem. On the other hand, the variable resolution of ERT data may complicate matters; the tradeoffs involved in setting μ may vary across the model. Implementing TGV regularisation may also prove less straightforward on unstructured meshes. The vector nature of $\mathbf{p}$ requires the inclusion of directional information beyond the simple adjacency conditions that are generally sufficient for l_2 or TV regularisation [Rücker, 2010].

3.2 TGV regularisation algorithm

3.2.1 Algorithm strategy

An alternating-minimisation algorithm in $\mathbf{m}$ and $\mathbf{p}$ is used to solve the TGV regularised inverse problem. The full objective function,

$$\mathbf{\Phi}(\mathbf{m}) = \|\mathbf{W}_d(f(\mathbf{m}) - \mathbf{d})\|_q + \lambda\|\nabla\mathbf{m} - \mathbf{p}\|_1 + \lambda\mu\|\frac{1}{2}(\nabla\mathbf{p} + \nabla\mathbf{p}^T)\|_1, \tag{3.2.1}$$

is minimised with regard to $\mathbf{m}$ and $\mathbf{p}$ in turn, with the other held constant. First, the objective to be minimised with regards to $\mathbf{m}$ is

$$\mathbf{\Phi}_{TGV^m}(\mathbf{m}) = \|\mathbf{W}_d(f(\mathbf{m}) - \mathbf{d})\|_1 + \lambda\|\nabla\mathbf{m} - \mathbf{p}\|_1. \tag{3.2.2}$$

This subproblem resembles the TV regularised objective function (equation (2.1.3)), where a target value of the gradient is sought, instead of $\nabla\mathbf{m} = 0$. The l_1 data misfit term is used in this derivation deliberately; an l_2 data misfit term can be recovered by setting the IRLS terms to the identity in 3.2.6. Consequently, equation (3.2.2) can be minimised using an IRLS Gauss-Newton approach to obtain an updated model, $\mathbf{m}$. Next, $\mathbf{m}$ is held constant, and the objective to be minimised with regards to $\mathbf{p}$ is

$$\mathbf{\Phi}_{TGV^p}(\mathbf{p}) = \|\nabla\mathbf{m} - \mathbf{p}\|_1 + \mu\|\frac{1}{2}(\nabla\mathbf{p} + \nabla\mathbf{p}^T)\|_1. \tag{3.2.3}$$

3.2.2 TGV algorithm derivation

Here, the alternating-minimisation inversion algorithm is derived in more detail. The outer loop consists of single IRLS Gauss-Newton iterations solving equation (3.2.2) about the present solution. Each iteration is followed by a full calculation of $\mathbf{p}$ for the present model, through IRLS minimisation of equation (3.2.3) until convergence.

The 2D discretised form of equation (3.2.2), linearised about the present model, $\mathbf{m}^{(n)}$ is

$$\begin{aligned}\mathbf{\Phi}_{TGV^m}(\mathbf{m}^{(n)}) = \|\mathbf{W}_d(\mathbf{d}^{(n)} + \mathbf{J}^{(n)}(\mathbf{m}^{(n+1)} - \mathbf{m}^{(n)}) - \mathbf{d})\|_1 \\ + \lambda^{(n)}\|\sqrt{(\mathbf{C}_x\mathbf{m}^{(n)} - \mathbf{p}_x)^2 + (\mathbf{C}_z\mathbf{m}^{(n)} - \mathbf{p}_z)^2}\|_1,\end{aligned} \tag{3.2.4}$$

where $\mathbf{d}_n$ and $\mathbf{J}_{(ij)}^{(n)} = \partial d_i^{(n)}/\partial m_j^{(n)}$ are the forward model solution and Jacobian matrix of sensitivities for the present model, $\mathbf{m}^{(n)}$; $\mathbf{C}_x$ and $\mathbf{C}_z$ are the first order finite difference matrices in the horizontal and vertical directions [deGroot Hedlin and Constable, 1990], and $\mathbf{p}_x$ and $\mathbf{p}_z$ are the vertical and horizontal components of $\mathbf{p}$. Note that I use the isotropic gradient, $|\nabla\mathbf{m}_x + \nabla\mathbf{m}_z|$, which better handles gradients aligned in an arbitrary direction.

Using the approach of Rodriguez and Wohlberg [2006], IRLS can be used to represent l_1 norms of vector and tensor fields as reweighted l_2 norms. Equation (3.2.4) becomes

$$\Phi_{TGV^m}(\mathbf{m}^{(n)}) = \frac{1}{2}\|(\mathbf{R}_d^{\frac{1}{2}})^{(n)}\mathbf{W}_d(\mathbf{d}^{(n)} + \mathbf{J}^{(n)}(\mathbf{m}^{(n+1)} - \mathbf{m}^{(n)}) - \mathbf{d})\|_2 + \frac{\lambda^{(n)}}{2}\|(\mathbf{R}_m^{\frac{1}{2}})^{(n)}((\mathbf{C}_x\mathbf{m}^{(n)} - \mathbf{p}_x))\|_2 + \frac{\lambda^{(n)}}{2}\|(\mathbf{R}_m^{\frac{1}{2}})^{(n)}(\mathbf{C}_z\mathbf{m}^{(n)} - \mathbf{p}_z)\|_2, \quad (3.2.5)$$

where $(\mathbf{R}_d)_{i,k}^{(n)} = 1/\sqrt{(\mathbf{W}_d(\mathbf{d}^{(n)} - \mathbf{d}))^2 + \gamma_d^2}$ and $(\mathbf{R}_m)_{i,k}^{(n)} = 1/\sqrt{(\mathbf{C}_x\mathbf{m}^{(n)} - \mathbf{p}_x)^2 + (\mathbf{C}_z\mathbf{m}^{(n)} - \mathbf{p}_z)^2 + \gamma_m^2}$ are the diagonal reweighting matrices, which are functions of $\mathbf{m}^{(n)}$ and so change each iteration. $\mathbf{R}_d = \mathbf{I}$ for the l_2 data norm. The γ terms are small constants to avoid singularities in the solution. I follow the approach of [Farquharson and Oldenburg, 1998] in solving the iterative Gauss Newton and IRLS solutions simultaneously.

Differentiating and minimising for the updated model, $\mathbf{m}^{(n+1)}$, gives

$$\mathbf{m}_{n+1} = \mathbf{m}^{(n)} + \left[\mathbf{J}^{(n)T}\mathbf{R}_d^{(n)}\mathbf{J}^{(n)} + \lambda^{(n)}(\mathbf{C}_x^T\mathbf{R}_m\mathbf{C}_x\mathbf{m}^{(n)} + \mathbf{C}_z^T\mathbf{R}_m^{(n)}\mathbf{C}_z\mathbf{m}^{(n)})\right]^{-1}\left[\mathbf{J}^{(n)T}\mathbf{R}_d\mathbf{W}_d(\mathbf{d}^{(n)} - \mathbf{d}) - \lambda^{(n)}\mathbf{Y}\right], \quad (3.2.6)$$

where $\mathbf{Y} = \mathbf{C}_x^T\mathbf{R}_m^{(n)}(\mathbf{C}_x\mathbf{m}^{(n)} - \mathbf{p}_x) + \mathbf{C}_z^T\mathbf{R}_m^{(n)}(\mathbf{C}_z\mathbf{m}^{(n)} - \mathbf{p}_z)$. This reduces to the TV solution when $\mathbf{p} = 0$, and the l_2 solution if $\mathbf{R}_m = \mathbf{I}$ also.

The discretised objective for equation (3.2.3) is

$$\Phi_{TGV^p}(\mathbf{p}) = \|\sqrt{(\mathbf{C}_x\mathbf{m} - \mathbf{p}_x)^2 + (\mathbf{C}_z\mathbf{m} - \mathbf{p_z})^2}\|_1 + \mu\|\sqrt{(\mathbf{C}_x^T\mathbf{p}_x)^2 + (\mathbf{C}_z^T\mathbf{p}_z)^2 + 0.5(\mathbf{C}_x^T\mathbf{p}_z + \mathbf{C}_z^T\mathbf{p}_x)^2}\|_1. \quad (3.2.7)$$

Transposing the second order difference matrices ensures alternating forward-backward differences for accuracy. Minimising using IRLS,

$$\Phi_{TGV^p}(\mathbf{p}^{(k)}) = \frac{1}{2}\|(\mathbf{R}_m^{\frac{1}{2}})^{(k)}\mathbf{C}_x\mathbf{m}_x - \mathbf{p}_x^{(k)}\|_2 + \frac{1}{2}\|(\mathbf{R}_m^{\frac{1}{2}})^{(k)}\mathbf{C}_z\mathbf{m}_z - \mathbf{p}_z^{(k)}\|_2 + \frac{\mu}{2}\|(\mathbf{R}_p^{\frac{1}{2}})^{(k)}\mathbf{C}_x^T\mathbf{p}_x^{(k)}\|_2 + \frac{\mu}{2}\|(\mathbf{R}_p^{\frac{1}{2}})^{(k)}\mathbf{C}_z^{\frac{1}{2}}\mathbf{p}_z^{(k)}\|_2 + \frac{\mu}{4}\|(\mathbf{R}_p^{\frac{1}{2}})^{(k)}(\mathbf{C}_x^T\mathbf{p}_z^{(k)} + \mathbf{C}_z^T\mathbf{p}_x^{(k)})\|_2. \quad (3.2.8)$$

where $\mathbf{R}_p = 1/\sqrt{(\mathbf{C}_x^T\mathbf{p}_x^{(k)})^2 + (\mathbf{C}_z^T\mathbf{p}_z^{(k)})^2 + 0.5(\mathbf{C}_x^T\mathbf{p}_z^{(k)} + \mathbf{C}_z^T\mathbf{p}_x^{(k)})^2 + \gamma_p^2}$. Differentiating with regard to $\mathbf{p}_x^{(k)}$ and $\mathbf{p}_z^{(k)}$ obtains the minimiser as the solution to the coupled equations,

$$\begin{bmatrix} \mathbf{R}_m + \mu \mathbf{C}_x \mathbf{R}_p^{(k)} \mathbf{C}_x^T + \frac{\mu}{2} \mathbf{C}_z \mathbf{R}_p^{(k)} \mathbf{C}_z^T & \frac{\mu}{2} \mathbf{C}_z \mathbf{R}_p^{(k)} \mathbf{C}_x^T \\ \frac{\mu}{2} \mathbf{C}_x \mathbf{R}_p^{(k)} \mathbf{C}_z^T & \mathbf{R}_m + \mu \mathbf{C}_z \mathbf{R}_p^{(k)} \mathbf{C}_z^T + \frac{\mu}{2} \mathbf{C}_x \mathbf{R}_p^{(k)} \mathbf{C}_x^T \end{bmatrix} \begin{bmatrix} \mathbf{p}_x^{(k+1)} \\ \mathbf{p}_z^{(k+1)} \end{bmatrix} = \begin{bmatrix} \mathbf{R}_m^{(k)} \mathbf{C}_x \mathbf{m} \\ \mathbf{R}_m^{(k)} \mathbf{C}_z \mathbf{m} \end{bmatrix} \tag{3.2.9}$$

This is a local approximation for $\mathbf{p}$ ($\mathbf{R}_m^{(k)}$ and $\mathbf{R}_p^{(k)}$ are functions of $\mathbf{p}^{(k)}$). I calculate $\mathbf{p}$ to convergence for every intermediate model, ensuring the accuracy of the outer loop model updates. As $f(\mathbf{m})$ does not need to be recomputed, the inner loop iterations are relatively cheap.

The full algorithm is summarised in algorithm 1, where x_d are the electrode locations and equation (3.2.9) is expressed as $\mathbf{Qp} = \mathbf{t}$.

Algorithm 1 ERT inversion with TGV regularisation

Input $\mathbf{d}, \mathbf{W}_d, x_d$
Initialise $\mathbf{m}, \mathbf{p}_x, \mathbf{p}_y \leftarrow \mathbf{0}; \lambda^{(1)}$
repeat
  $\mathbf{d}?^{(m)}, \mathbf{J} \leftarrow \text{forwardSolver}(\mathbf{m}, x_d)$
  $\mathbf{m} \leftarrow \mathbf{m} + \left[\mathbf{J}^T \mathbf{R}_d \mathbf{J} + \lambda \mathbf{Y}\right]^{-1} \left(\mathbf{J}^{(n)T} \mathbf{R}_d \mathbf{W}_d (\mathbf{d}^{(n)} - \mathbf{d}) - \lambda \mathbf{Y}\right)$
  $\mathbf{p}_x, \mathbf{p}_y \leftarrow \mathbf{0}$
  repeat
    $\mathbf{p}_x, \mathbf{p}_y \leftarrow \mathbf{Q}^{-1}\mathbf{t}$
  until $\mathbf{p}$ converged
  update λ
until $\mathbf{m}$ converged, or iteration limit reached

3.2.3 Implementation

The implementation of TGV regularisation for ERT is based on the open source IP4DI inverse code for MATLAB [Karaoulis et al., 2013]. The code for my implementation is available online [Hawley-Sibbett, 2023].

IP4DI uses separate meshes for the forward and inverse problems: a coarse rectangular grid for the resistivity parameters, $\mathbf{m}$, used for the finite difference inverse problem; and a finer, triangular finite element mesh constructed between the nodes of the rectangular grid is used for the forward problem. The finite element boundary conditions at the air-surface boundary are homogeneous Neumann (zero current) and homogeneous Dirichlet (zero potential) at the internal boundaries, which which extend far from the sensors, outside the rectangular parameter mesh, to model the potential dropping to zero at infinity. The finite difference boundary conditions for $\mathbf{m}$ were also homogeneous Neumann, enforcing $\nabla \mathbf{m} = 0$ across the boundary normals.

The forward modeling elements of IP4DI were unmodified, except for some optimisation changes; these were thoroughly tested before being used in the inversion. The Jacobian is calculated analytically (as in Tsourlos [1995]) as is the case in the IP4DI base code [Karaoulis et al., 2013], using the secondary potential approach. In this initial implementation, homogeneous Neumann boundary conditions were applied to the finite difference equation in $\mathbf{p}$ also.

3.3 Initial testing of the TGV algorithm

In this section I present the initial testing and benchmarking of the TGV algorithm. These early results provided useful insights into the formulation of the inverse algorithm and model configuration. These changes are described in Section 3.4, and implemented for the more extensive testing in Chapter 4.

3.3.1 Algorithm configuration and parameterisation for initial testing

The inverse problem was calculated in the log resistivity domain. Otherwise, as the typical transfer resistances in ERT span several orders of magnitude, the low transfer resistance data is effectively down-weighted in the data fitting process. The same principal applies to the model resistivity. The model, $\mathbf{m}$ was initialised to a uniform zero. An l_1 data misfit term was used. The IRLS stabilising parameters were $\gamma_m = 10^{-3}$, $\gamma_p = 10^{-3}$ and $\gamma_d = 5 \times 10^{-3}$. In the inversion, I used a cooling off scheme for the Lagrnage multiplier, λ, terminating the algorithm when the change in root mean square (RMS) error was 2% or less. The inversion was repeated several times for various initialisation of $\lambda^{(0)}$ and cooling off schemes. The data weighting matrix, $\mathbf{W_d}$ was set to the identity matrix; data noise was not accounted for in the inversion stage, which risks over-fitting the data to the noise. To compensate for any over-fitting in the synthetic examples, the solution that is closest to the synthetic model across all iterations of all inversion runs with a good fit to the data were were chosen for comparison.

The rectangular finite difference parameter mesh was generated around the electrode positions. The top layer of cells have width equal to the inter-electrode spacing, s_{elec}, and depth $0.5s_{elec}$, and the depth of each cell layer increases by 10% relative to the last. This type of scheme with small model cells in close proximity to the electrodes and larger ones at depth mirrors the ERT resolution, allowing for increased computational efficiency (see for example [Rücker, 2010], [Tsourlos, 1995]). The varying cell widths were accounted for in the finite difference matrices $\mathbf{C}_x$ and $\mathbf{C}_y$.

The convergence criterion for $\mathbf{p}$ was a root mean square change in $\mathbf{p}$ <1% relative to the last iteration. $\mathbf{p}$ was initialised to $[0; 0]$ at all points.

3.3.2 Benchmarking the TGV functional solver

The alternating-minimisation algorithm for TGV inversion solves the TGV regularisation function, equation (3.2.3), for each intermediate resistivity model in the inversion. This serves a dual purpose: verifying that the TGV implementation is functioning as expected, and providing some insight into the values of μ which can be expected to perform well in the inversion. This part of the algorithm can be benchmarked independently of the inversion by solving equation (3.2.9) for a given resistivity model until convergence, which was was defined as a mean change in $\mathbf{p}$ of less than 1% across the entire model.

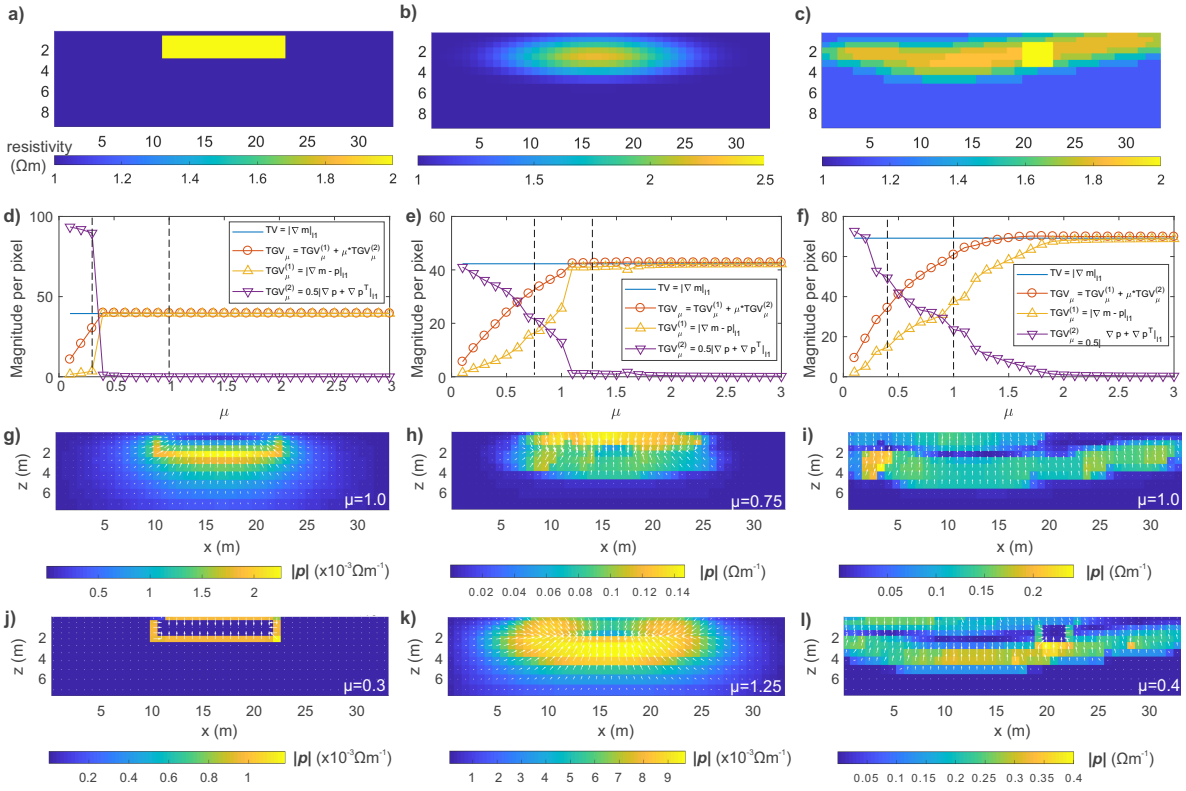


Figure 3.1: Effect of μ on the TGV functional, d-e, for fixed synthetic models, a-c, during the initial testing of the TGV algorithm. From top to bottom, each column contains: (i) The synthetic model used to test the TGV functional; (ii) The variation of the TGV objective function with μ for the same model, normalised by the number of cells in the model. The vertical dotted lines denote the values of μ presented in g-f; (iii) The magnitude and direction of $\mathbf{p}$ for a value of μ chosen for its piecewise smooth behaviour; (iv) An example of $\mathbf{p}$ for a value of μ which poorly captures the desired piecewise smooth behaviour. Note that the magnitudes of $\mathbf{p}$ vary significantly across g-f.

Figure 3.1 benchmarks the TGV function for three synthetic models of different roughnessess: a piecewise constant rectangular block (figure 3.1a); a smooth Gaussian like feature (figure 3.1b); and a piecewise smooth model containing a constant resistivity block in a smoothly varying

background (figure 3.1c). The variation in each term of the TGV objective function for each model are given in figure 3.1(d-f).

The limiting behaviour of these curves can be used to test whether the TGV functional behaves as expected. Theoretically, as μ becomes large, $\nabla \mathbf{p}$ is heavily penalised, and the TGV functional reduces to $|\nabla \mathbf{m}|_1$, the Total Variation. In all three models' μ dependency curves, TGV_μ monotonically approaches TV as μ increases, with little observable differences for $\mu > 2$. Similarly, in the case of an entirely piecewise constant model, such as figure 3.1a, the TGV solution should resemble TV. The results show that TGV_μ is almost indistinguishable to the TV result for much of its range, showing small differences when $\mu < 0.5$ and breaking down entirely for $\mu <= 0.3$. Therefore, $\mu = 0.5$ is the lower bound at which sharp edges can be reasonably represented by the TGV functional, although in practice the relatively low resolution of the ERT inversion is likely to require the TGV functional to be weighted further in favour of sharp edges.

Figure 3.1(e,f), show that both terms of the TGV functional contribute significantly to the objective in the approximate range $0.5 < \mu < 1.5$, while larger values of μ differ from TV function by a smaller amount.

The magnitude and direction of $\mathbf{p}$ for each model are shown in Figures 3.1(g-l). If the TGV functional is capturing piecewise smooth behaviour, $\mathbf{p}$ should be approximately equal to the model gradient in smooth regions, small at edges and in constant resistivity regions.

For the blocky model, $\mathbf{p}$ was of the order of 10^{-3} (Figure 3.1g). This is much smaller than the gradient of the block, but is not negligible. At $\mu = 0.3$, this breaks down, and $\mathbf{p} = \nabla \mathbf{m}$ at the sharp edges of the resistive block 3.1j. This effect also occurs at the edges of the block in figure 3.1l.

Figures 3.1(h,i) were chosen from the range of μ where both terms of the TGV functional contribute significantly to the objective. In both cases, $\mathbf{p}$ broadly follow the gradients of the smooth features, and $\mathbf{p}$ is small at the sharp edges of the resistive block in 3.1i. This agrees well with the expected behaviour of $\mathbf{p}$ over piecewise smooth models. The magnitude of $\mathbf{p}$ in figure 3.1g approaches the lower limit of γ_p, reflecting the piecewise constant model. However, the asymmetry in $\mathbf{p}$ within the symmetrical smooth feature in figure 3.1h is unexpected, and may be indicative of errors in the model.

The lower limit of $\mathbf{p} \approx \gamma_p = 10^{-3}$ is approached in 3.1k, where the objective function is close to TV. In this case, it is expected that TV-like behaviour would occur at this μ.

Overall, this initial analysis of the TGV inner loop implementation yielded promising results. The TGV functional behaved generally as expected in the limits of large and small μ and in the case of an entirely piecewise constant model. A range of $0.75 < \mu < 1.5$ appears effective

in capturing piecewise smooth features for these static models. These initial results suggest that the default $\mu = 2$ value used in image processing may favour sharp boundaries more than smooth gradients, and so smaller μ may be preferred for ERT inversion. The minimum $\mathbf{p}$ observed was of the order of 10^{-3}, which is not negligible in relation to the model gradient. Some asymmetries in $\mathbf{p}$ were also observed, where the underlying model is symmetrical. Both of these factors are reviewed in section 3.4.

3.3.3 Synthetic testing of initial TGV implementation

Next, the full inversion algorithm was tested, under the same conditions used in Section 3.1. These preliminary results provide some useful insights into the algorithm configuration, and so are presented here, despite the issues which were identified in Section 3.1. ERT data were generated from synthetic resistivity models, and used to test the relative performance of TGV, TV and l_2 constraints.

As modelling bias can be introduced by using the same mesh for the forward and inverse problems, these data were generated using the forward modelling software Res2DMod [Loke and Barker, 1996] from models formed on a regular rectangular grid (0.15m $\times$0.18m cells). As both the Res2DMod and IP4DI grids define their meshes around vertices on the electrode positions, some edges were shared, however. The models themselves were higher resolution realisations of the models in 3.1(a-c). In these tests no noise was added to the results.

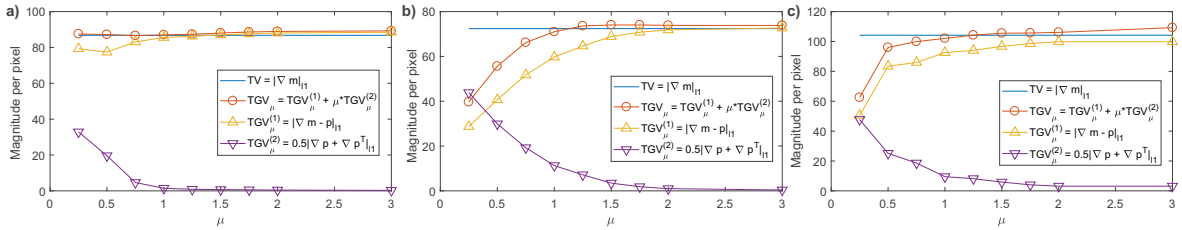


Figure 3.2: Variation of the TGV function with μ for the final iteration of the inverted image, during the initial testing of the TGV algorithm. (a-c) respectively correspond to the inversion of data derived from the synthetic models shown in figure 3.2(a-c). The corresponding TV result is shown for reference. It should be remembered that exact comparisons between the data should be treated with caution, as they each are calculated on a different model grid.

Figure 3.2 shows the μ dependence of the TGV regularisation. Unlike the static example in figure 3.1, the models output by the inversion have been shaped by the regularisation process. The inversion seeks the model that minimises TGV_μ at each iteration, whilst also being consistent with the resistivity data. All three models tend towards a constant value close to the TV solution as $\mu \to 3$, at a slower rate than for a static model, and hence a wider range of μ

appears viable than for the static models.

An analysis of the inverted models, and the behaviour of $\mathbf{p}$, provides insight into the performance of this initial implementation. I show examples of the inversions using each regularisation method. μ were chosen to highlight significant trends from Figure 3.2. As the TV implementation is also new, I analyse it's performance in parallel.

The three regularisation methods perform broadly as expected on the piecewise constant model (figure 3.3). Both the TV and $TGV_{0.75}$ solutions are sharp and well defined, in contrast to the characteristically smooth l_2 solution. The $TGV_{0.75}$ regularisation unexpectedly defined the lower edge of the model more sharply than the TV solution. As μ increases, the TGV solution approaches the TV result (not shown), and the solution for $\mu > 2$ is almost identical to the the TV solution.

Figure 3.4 shows that $\mathbf{p}$ supports a weak gradient within the block in the $TGV_0.75$ solution. Such a gradient is present in the inverted model (figure 3.3c), although it is too weak to be clear in the image. It is plausible that this adherence to a local gradient acts as a stabilising feature on the dimensions of the block. As μ increases, this gradient is diminished, and the result tends towards the TV solution.

These effects occur at the lower edge of the block, which is further from the electrodes, and therefore lower resolution than the other block boundaries. The regularisation method has a greater influence on this part of the model.

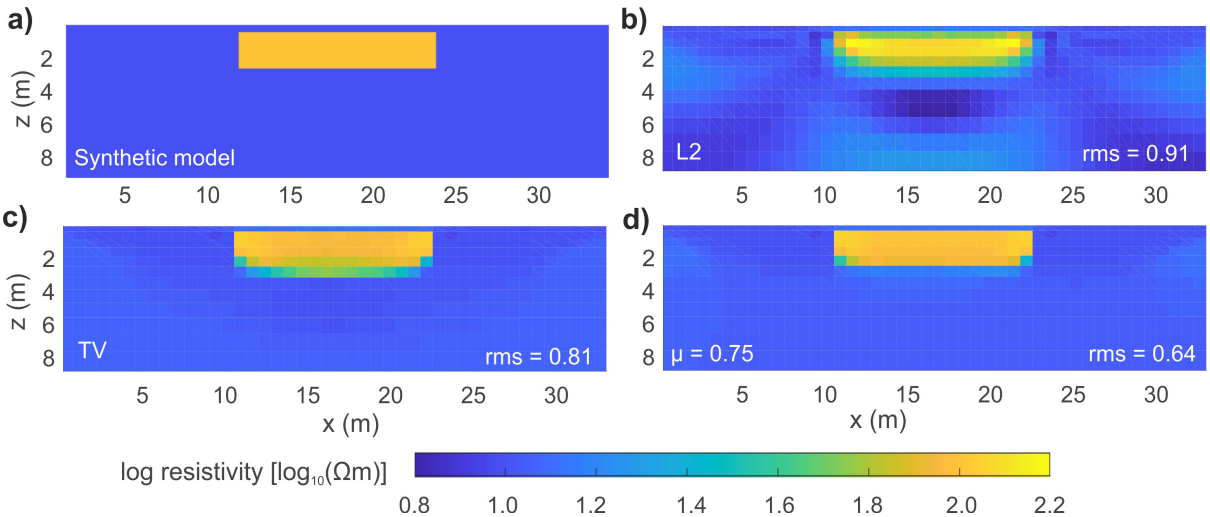


Figure 3.3: Results of initial TGV inversion of data from a synthetic piecewise constant model. The synthetic model is shown in (a). (b-d) show inversion results with l_2, TV, and $TGV_{0.75}$ regularisation respectively. The root-mean-square (rms) model misfits with regard to (a) are included.

In the smooth model, figure 3.5, the TV and TGV solutions share a sharp edge at the top of

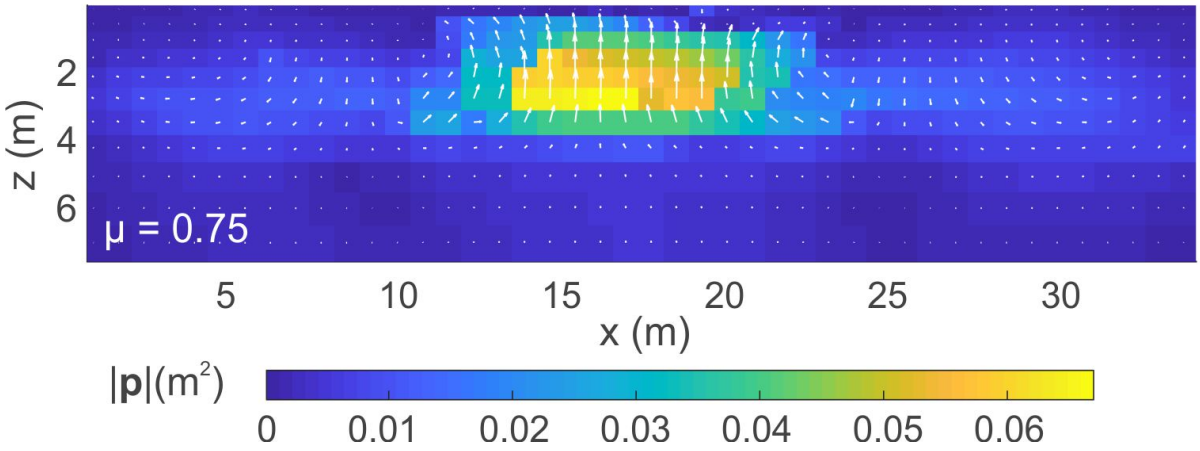


Figure 3.4: Distribution of $\mathbf{p}$ (white vectors) for $TGV_{0.75}$ corresponding to the inversion result in 3.3. The magnitude, $|\mathbf{p}|$, is also shown as a colour map. The small magnitude of $\mathbf{p}$ doe not give rise to a meaningful gradient within the resolving power of the ERT data, but may play a role in stabilising the sharp block boundary in the inverted model.

the Gaussian feature ($z = 0.9m$), despite varying significantly in smoothness across the rest of the model. This is in contrast to the l_2 solution, in which the resistivity accurately drops away smoothly in all directions around the peak, as is expected, as l_2 regularisation favours Gaussian smooth models.

This local breakdown in the smooth behaviour of the TGV inversions, regardless of μ, is concerning, especially in a relatively high resolution region near the surface. Figure 3.6, provides some insight. $\mathbf{p}$ becomes small in this region, when I would expect it to match the resistivity gradient of the original model. As $\mathbf{p} \to 0$, the TGV functional reduces to TV, resulting in the observed sharp edges in the TGV solutions. This behaviour near the surface was observed in the $\mathbf{p}$ distributions regardless of μ. When $\mu \geq 2$, $\mathbf{p}$ was at least an order of magnitude smaller than the gradient over the entire model.

The different regularisation functions behaved much as expected throughout the rest of the model. The peak of the Gaussian feature is replaced by a flat, constant resistivity region in the TV model, although the smooth edges of the Gaussian are captured remarkably well; the use of the isotropic gradient approximation may be at least partially responsible, as it is advantageous in capturing gradients which are not aligned with the mesh axis. The TGV models decrease in smoothness as μ increases, approaching the TV solution at $\mu = 2$. The $TGV_{0.75}$ solutions captures the smoothness of the Gaussian feature better than the TV solution, although the lower bound of the feature is more accurately defined in the TV solution; the $TGV_{0.75}$ model linearly extends several metres below the original feature boundary.

I again see similarities between the TGV and TV solutions in the upper regions of the piecewise

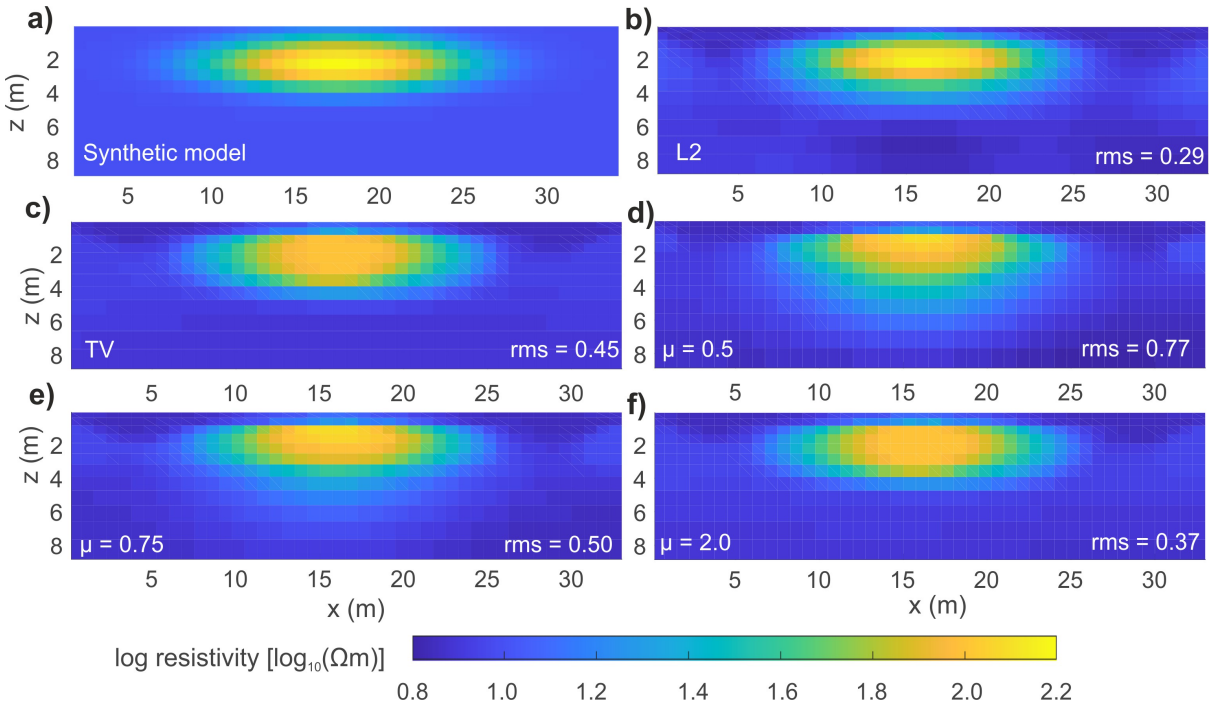


Figure 3.5: Results of initial TGV inversion of data from a synthetic smooth Gaussian model. The synthetic models is shown in (a). (b-f) show inversion results with l_2, TV, $TGV_{0.5}$, $TGV_{0.75}$, and $TGV_{2.0}$ regularisation respectively. The root-mean-square (rms) model misfits with regard to (a) are included.

smooth model, particularly from $0 < x < 20m$. In this area $\mathbf{p}$ shrinks to zero (figure 3.8), and the $TGV_{1.0}$ solution is sharp edged (figure 3.7d).

Both the $TGV_{1.0}$ and the TV models capture the sharp edges and resistivity of the constant-resistivity block much better than the l_2 solution, which is smooth and significantly overestimates the peak resistivity.

The lower boundary of the smooth region is not well realised in any of the models, likely due to the limitations of the resolution in this region, especially near the corners of the image. The l_2 model smoothly extends into the background region; the flat peak of the TV solution is wider than the original model, and projects beyond it smoothly; and the $TGV_1.0$ solution captures the width of the feature well, but significantly overestimates the background resistivity. In all three cases the resistivity of the smooth feature is overestimated and the background resistivty underestimated near the surface. Once again, the TV solution is smoother than one would expect in the deeper regions of the model, where the regularisation function tends to have a larger influence on the solution relative to the data. This suggests that an additional source of smoothing is affecting these solutions.

Elsewhere the $TGV_{1.0}$ solution is smooth. The results for different values of μ show the same broad behaviour as in figure 3.6 (not shown): decreasing μ leads to smooth extensions into

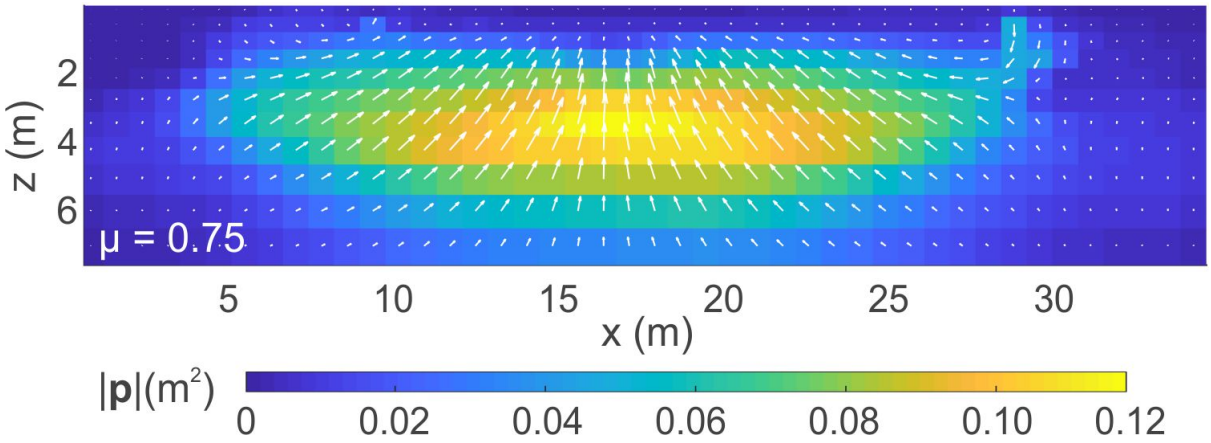


Figure 3.6: Distribution of $\mathbf{p}$ (white vectors) for $TGV_{0.75}$ corresponding to the inversion result in 3.5. The magnitude, $|\mathbf{p}|$, is also shown as a colour map. Near the surface the direction and magnitude of $\mathbf{p}$ shrinks to small values, while elsewhere it reflects the model gradient.

the background regions, and increasing μ causes the model to incrementally resemble the TV solution more closely.

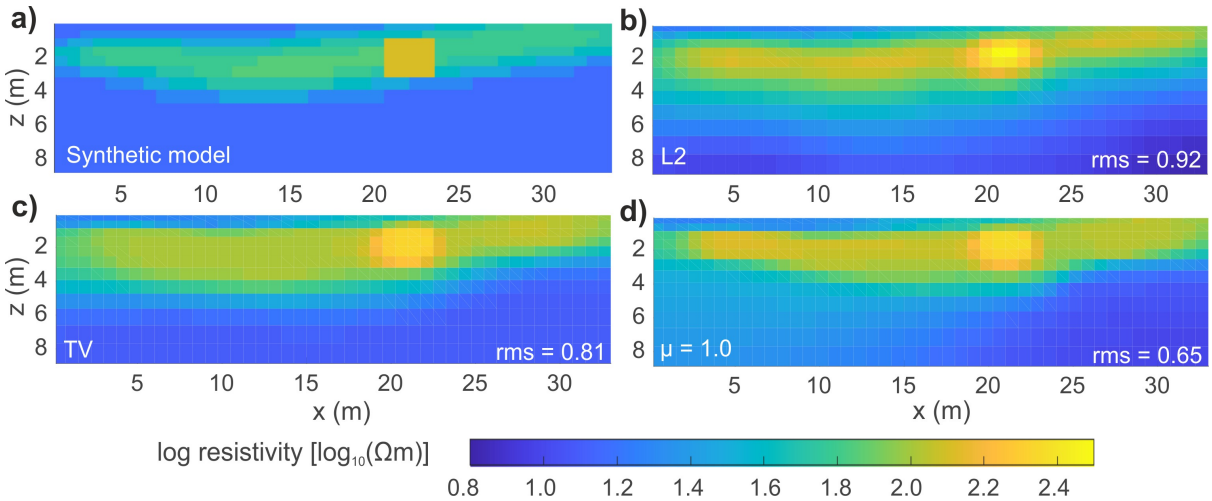


Figure 3.7: Results of initial TGV inversion of data from a synthetic piecewise smooth model. The synthetic model is shown in (a). (b-d) show inversion results with l_2, TV and $TGV_{1.0}$ regularisation respectively. The root-mean-square (rms) model misfits with regard to (a) are included.

3.3.4 Initial testing results

These initial tests demonstrated the viability of TGV regularisation for ERT, while highlighting some potential problems with the current implementation. The TGV solutions converged reliably, and displayed the expected piecewise smooth behaviour, suggesting that the alternating-minimisation IRLS approach is valid for ERT inversions. As μ increased, the TGV solution approached the TV solution, futher benchmarking the method. However, there were several

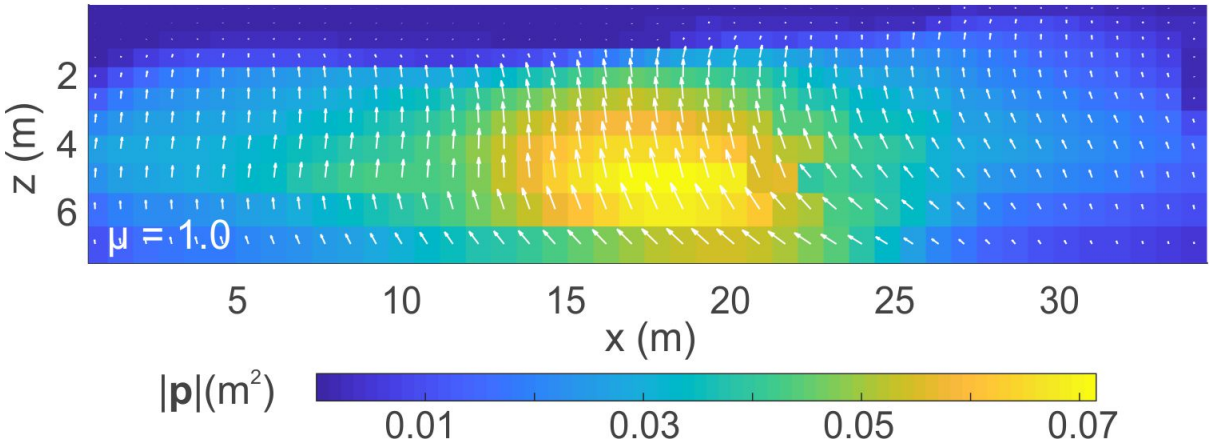


Figure 3.8: Distribution of $\mathbf{p}$ (white vectors) for $\text{TGV}_{1.0}$ corresponding to the inversion result in 3.7. The magnitude, $|\mathbf{p}|$, is also shown as a colour map. In this instance also $\mathbf{p}$ vanishes instead of changing direction near the surface, while following the gradient elsewhere.

inconsistencies with expected behaviours.

Smoothing effects occurred in the TV solutions, and the minimum observed $\mathbf{p}$ was of the order of 10^{-3}, which is not negligible compared to the model gradient. These may be related to the size of the IRLS stabilising parameters. $\gamma_m = 10^{-3}$ is not small compared to the resistivity gradients. Similarly, $\gamma_p = 10^{-3}$ places a lower bound on $\mathbf{p}$ at the same magnitude, which may result in unwanted smoothing effect at $O(10^{-3})$ in the TGV solutions too. These constants exist to prevent singularities in the solutions, and should be sufficiently small that they are negligible, ideally close to machine precision (10^{-15}).

Another anomaly in the TGV solutions is the suppression of $\mathbf{p}$ near the upper model boundary, instead of following the changes in direction of the model gradient. This did not affect the stability of the inversion, as the regularisation behaves like TV in these areas. However, it does reduce the effectiveness of the TGV regularisation when smooth gradients are found in these areas.

In addition, there were notable horizontal asymmetries in $\mathbf{p}$, for the symmetrical fixed model in Figure 3.1h. This effect is seen independently of the inversion, and so is related to the solution of equation (3.2.9). One possibility is that the solver was terminated early, before $\mathbf{p}$ had fully converged. Alternatively, it may be due to the finite difference method used in this analysis, or due to numerical errors in the code.

3.4 Investigating sources of error in the initial TGV algorithm

There were several inconsistencies in the initial solutions to the TGV algorithm, which may be due to challenges encountered when working with inverse problems. In this section, some of those challenges are discussed. A review of the methodology used in the initial TGV testing is performed, and potential sources of error are identified, in preparation for further testing in Chapter 4.

3.4.1 Challenges of debugging an inverse problem solver

The term 'bugs' is generally used to refer to programming errors in an algorithm or piece of software. However, the performance of an ERT inversion is affected by the configuration of many interlocking system components that do not necessarily constitute errors. In this section, I use the term 'debugging' to refer to the joint process of identifying any errors in the algorithm and adjusting any poorly configured elements.

A given solution to an inverse problem is not unique, instead it is drawn from the unknown distribution of all valid solutions [Boyle, 2016]. The data misfit is a measure of validity, assuming functioning forward modelling software, and models with a large data misfit can be discarded. This non-uniqueness presents a challenge in the evaluation of an inversion algorithm: even if a solution is valid and fits the data, it does not mean that the algorithm is functioning as desired. As such, the related processes of 'debugging' and configuring an inversion algorithm require strategies which combine evidence from several tests and observed behaviours, which overall demonstrate that the algorithm is performing as expected.

Various parameters influence the results of an inversion. The initial model and the chosen schedule for λ affect the descent path through solution space, and ultimately which local solution is reached. The optimum choice of these parameters may be different for different regularisation functions, and different regularisation methods may be more sensitive to changes in these conditions.

Other factors which affect the outcomes of the inversion are modelling errors, arising from the finite difference approximation used, the mesh geometry, and the boundary conditions. The IRLS stabilising parameters, which are introduced solely to avoid numerical singularities, may also unduly influence the model if not sufficiently small. The convergence criterion, or the selection of the best iteration of an inversion will affect the chosen reconstruction. Noise, and the error weighting matrix $\mathbf{W}$, may also introduce artefacts into the solutions, and complicates the convergence criterion as over-fitting noise results in a poorer reconstruction at a lower data misfit.

[Boyle, 2016] writes that "When debugging any problem, a clear idea of the expected outcome is important. When a result is outside of the expected range or somehow unusual a deeper investigation can be initiated". These expectations provide an additional layer of validation, and increase our confidence in an algorithm.

For TGV regularisation, these expected behaviours centre around two main factors: the effect of μ and the relationship to TV, and the general smoothing behaviours. As μ increases, the TGV functional converges on TV. This is a useful property for benchmarking, as it can be quantified through the objective function of the model, and by comparison to the TV solution. Less concretely, it is expected that models should progressively become smoother as μ decreases from this point. The relationship between $\mathbf{p}$ and the gradient is also useful. It is expected that $\mathbf{p}$ will align itself with the significant smooth gradients in the model, and vanish near sharp edges. All else being equal, linear gradients will be preferred.

In this section, a range of validation and correctness tests are performed on appropriate constituent parts of the algorithm. First, consistency checks are carried out to ensure that errors have not been introduced to the underlying code base. Next, investigations are carried out into three potential sources of error in the inversion which may be responsible for some of the inconsistencies which were identified in the initial tests. Finally, some changes to the configuration of the test inversions were performed in order to reduce the potential sources of uncertainty further. Once this process is complete, the performance of the improved TGV regularised algorithm will be evaluated through a parameter testing process in Chapter 4.

3.4.2 Consistency checks

Consistency checks were performed on established parts of the code base, to ensure that these were functioning correctly, and that the inconsistencies in the initial inversion results were not due to coding errors.

As the TGV inversion algorithm utilised and partially modified the existing forward modelling and meshing components of the inversion code, I was able to use the base IP4DI code as a benchmark for consistency checks. The purpose of this is to ensure that the information is flowing correctly through my algorithm and that no errors or bugs have been introduced during the integration of the TGV algorithm and existing code sections. While it can be generally assumed that the original code, which has been independently tested, is accurate [Karaoulis et al., 2013], additional validation of the forward model was performed through comparison with other established forward modelling codes, such as Res2DMod Loke and Barker [1996].

Some modifications to the IP4DI forward modelling code were made in order to improve its

computational efficiency. These were previously consistency checked with the original upon implementation, but I repeated the process to ensure that forward modelling errors were not responsible for the artefacts noticed in the initial testing.

When the modified forward model was benchmarked against the original IP4DI forward model, using the synthetic model 3.1b, the results were numerically identical. The results were also the same as those of the Res2DMod forward modeling software, to a tolerance of $< 0.1\%$, which can be attributed to the different meshes and modelling approaches used. As these forward modelling elements have both been independently tested, it did not seem necessary to perfrom further tests using analytical models.

As IP4DI does not have capacity for l_1 inversions, an l_2 regularisation was chosen for a full inversion test. The same data were inverted using l_2 regularisation in both IP4DI and using the TGV algorithm, using the same cooling off scheme for λ. The inversion results were numerically identical, smooth reconstructions of the original model. This limits any bugs or errors to the parts of the code responsible for the TGV and l_1 inversions only.

As part of the consistency checks, visual examination of the finite difference matrices were performed, with particular attention paid to the boundary conditions. These were implemented correctly.

As part of the consistency checks the minimisation of $\mathbf{p}$, equation (3.2.9), was also independently benchmarked before inversions were performed. This was performed after the 'debugging' process was complete, and can be found in Sections 4.1.

3.4.3 IRLS stabilising parameters

The TV solutions in the initial tests were more smooth than expected, and the minimum of $\mathbf{p}$ was 10^{-3}, which is not small relative to the magnitude of the model gradient. Furthermore, the TGV results for the blocky model were sharper at small μ, and decreased in smoothness as μ increased, contrary to the expected behaviour.

It is suspected that this is due to the fact that the IRLS stability factors $\gamma_p = \gamma_m = 10^{-3}$ are not strictly small compared to the model gradient, particularly as this is the same magnitude as the minimum observed $\mathbf{p}$.

This should be easily rectified by using a suitably small value for the stability factors. However, setting $\gamma_p = \gamma_m = 10^{-12}$, for example, caused both the TV and TGV inversions to fail on the first iteration, due to a singularity error in the solution of Equation (3.2.6).

To identify the source of this failure, the individual terms of Equation (3.2.6) were examined. It was found that the initialisation as a homogeneous half space was responsible for the failure

to solve the inverse equation. Previously, the large γ parameters had acted to stabilise the solution.

Replacing the initial model with the result of a single iteration of an l_2 regularised inversion, as in Loke et al. [2003], resulted in the TV and TGV inversions functioning as expected with $\gamma_m = \gamma_p = 10^{-12}$. These results are explored in Chapter 4, alongside the other changes to the inversion regime.

3.4.4 Asymmetric p

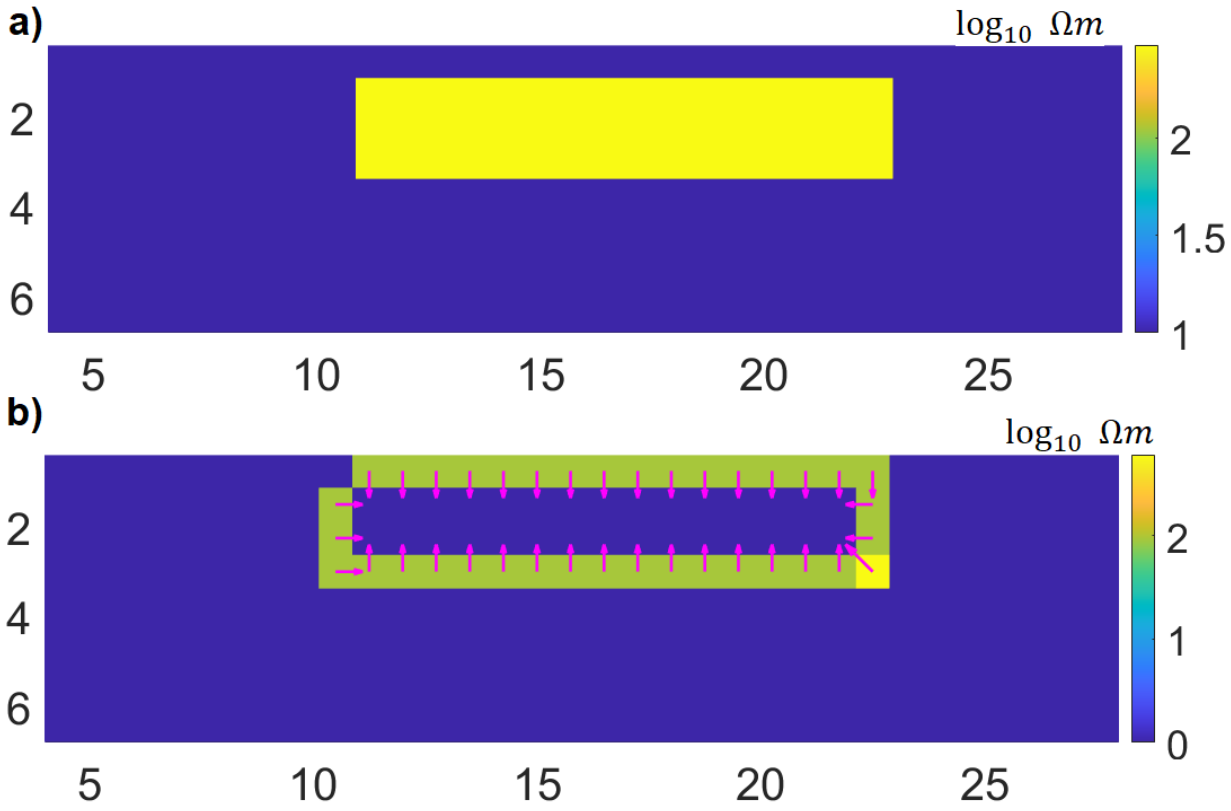


Figure 3.9: Demonstration of the asymmetry of the forward gradient of the piecewise constant model (a). (b) is model gradient calculated using forward differences. The colour map shows the magnitude, while the vectors include both magnitude and direction.

Another potential source of error in the TGV inversion is the forward-backward finite difference model used in gradient calculations. The one sided nature of the forward-backwards difference model introduces asymmetries into the TGV inversions. In some preliminary inversions of field data (not shown), artefacts tended to appear at the corners of the domain, which may be related to this.

The forward difference operator, $\mathbf{C}_x$, at point i approximates the gradient to be $\mathbf{m}_{i+1,k} - \mathbf{m}_{i,k}$

(see Section 2.2.1), where i and k are the cell indices in x and z. In matrix form, with homogeneous Neumann boundary conditions, this is

$$\mathbf{C}_x = \begin{pmatrix} -1 & 1 & 0 & \dots & 0 \\ 0 & -1 & 1 & \dots & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & \dots & 0 & -1 & 1 \\ 0 & \dots & 0 & 0 & 0 \end{pmatrix}, \quad \mathbf{C}_x^T = \begin{pmatrix} -1 & 0 & 0 & \dots & 0 \\ 1 & -1 & 0 & \dots & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & \dots & 1 & -1 & 0 \\ 0 & \dots & 0 & 1 & 0 \end{pmatrix}, \tag{3.4.1}$$

This single sided gradient is asymmetrical, as can be seen in figure 3.9. The position at which a gradient is recorded differs between forward and backward gradients. Here, the gradient at the block edge is assigned to the first cell outside the block at $x \approx 10$, and the first cell within the block at $x \approx 22$. At the block corners this results in errors. In this example, only the lower right hand corner is sensitive to both gradient directions.

However, the combined forward-backward operator $\mathbf{C}^T\mathbf{C}$ is symmetrical, and is equivalent to the second order central difference (2.2.5),

$$\mathbf{C}_x^T\mathbf{C}_x = \begin{pmatrix} -1 & 1 & 0 & \dots & 0 \\ 1 & -2 & 1 & \dots & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & \dots & 1 & -2 & 1 \\ 0 & \dots & 0 & 1 & -1 \end{pmatrix} \tag{3.4.2}$$

For homogeneous Neumann boundary, this reduces to a single sided difference at the boundary. This affects the various regularisations in various ways (see equation (3.2.6)). The l_2 inversion only contains second derivatives, and so is unaffected by the the asymmetry. The IRLS reweighting terms introduce some asymmetry into the isometric TV solution, as only the forward differences of the gradient components orthogonal to the gradient direction are included. i.e. In $\mathbf{C}^T\mathbf{R}_c\mathbf{C}$, only forward differences appear in the y components of the gradient in the $\mathbf{R}$ terms.

The TGV solution, equation (3.2.9), is more affected by this asymmetry. Single sided derivatives appear in the cross terms of the solution matrix, the $\mathbf{R}_m\mathbf{Cm}$ terms, and in the cross terms of the IRLS term $\mathbf{R}_p$. $\mathbf{C}_x^T$ and $\mathbf{C}_y^T$ appear independently of their adjoints. The impact of these asymmetries appears most strongly at the corners and edges of features, and at the domain boundaries. It is possible that this may contribute to the suppression of $\mathbf{p}$ near the boundaries, and to other artefacts at the model boundaries or around the edges of features.

It may be possible to alleviate this affect through better handling of the cross term boundaries. The backward operator's boundary conditions in particular are not physically meaningful when applied separately to the forward operator. An alternative boundary scheme is described in Section 3.4.5. A more accurate alternative would be to use a symmetrical approximation to the gradient. This idea is explored further in Section A, in the context of future work on TGV.

3.4.5 Boundary conditions

The initial implementation uses homogeneous first order Neumann boundary conditions in both $\mathbf{m}$ and $\mathbf{p}$; these enforce zero gradient across the model boundaries. The influence of boundary effects on the $\mathbf{p}$ distribution is apparent in Figures 3.6 and 3.8, where $\mathbf{p}$ shrinks to zero rather than change direction to follow the model gradient near to the boundary. There may also be errors in the application of the Neumann boundary conditions for the single sided backward operation (Section 3.4.4).

It is perhaps not surprising that homogeneous boundary conditions in both $\mathbf{m}$ and $\mathbf{p}$ suppress $\mathbf{p}$ locally. The Neumann boundary conditions enforce a constant $\mathbf{p}$ across the boundary. In smooth regions, $\nabla\mathbf{p} = 0$ corresponds to a constant resistivity gradient, $\nabla\mathbf{m} = constant$. However, as $\nabla\mathbf{m} = 0$ is also imposed at the boundary, the boundary conditions are only satisfied when $\mathbf{p} = 0$ at the boundary, which will tend to suppress $\mathbf{p}$ locally.

The imposition of $\nabla\mathbf{p}|_{boundary} = 0$ by the homogeneous Neumann boundary conditions may to be responsible for the suppression of $\mathbf{p}$ near the boundaries. The periodic boundary condition is a simple alternative. Opposite boundaries are considered to be adjacent, such that $\nabla_x\mathbf{m}|_{boundary} = \mathbf{m}_{x=n} - \mathbf{m}_{x=1}$, where there are n model cells in the x dimension. Periodic boundary conditions are commonly used in image processing contexts, including for TGV regularised problems (for example Bredies et al. [2010]; Duan et al. [2016]). One advantage of this is that the domain becomes a torus, with no distinct boundary edges which require special treatment, and so the effects of asymmetry around boundary handling are dampened. Any sharp discontinuities in resistivity across the periodic boundaries are naturally treated by the TGV functional as an edge.

For the piecewise smooth model, Figure 3.10, the change to periodic boundary conditions leads to an improvement in the ability of $\mathbf{p}$ to capture the model gradient near the boundary for smaller values of μ, Figure 3.11 and 3.12. Away from the boundary, $\mathbf{p}$ is unaffected.

This difference is most significant for $\mu \leq 1$ for this model, and therefore does not appear to fully solve the suppression of $\mathbf{p}$ near the boundary. Nevertheless, the improvement is significant enough to adopt periodic boundary conditions for all further evaluation of the algo-

rithm.

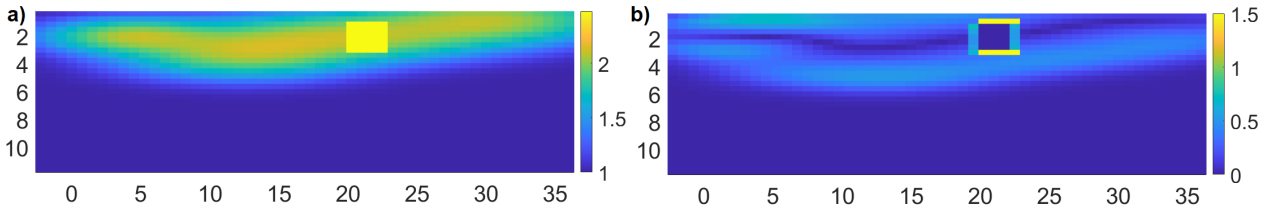


Figure 3.10: Piecewise smooth synthetic model (a) with the corresponding model gradient (b). The units are $log_{10}(\Omega m)$. The model is repeated from figure 3.1c for convenience.

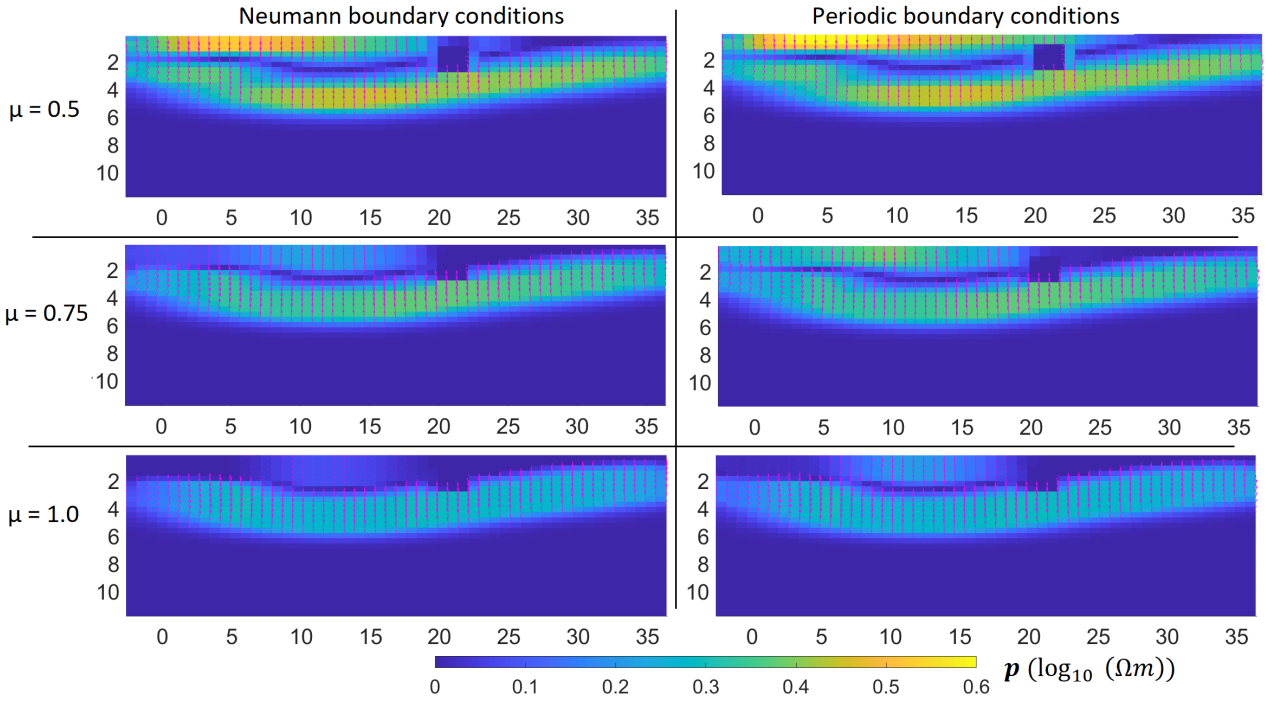


Figure 3.11: **p** distribution (pink vectors) for the piecewise smooth model 3.10a, for different boundary conditions and tradeoff parameters. The magnitude, $|\mathbf{p}|$, is also shown as a colormap. Unlike the color scaling, the arrow length scale varies between images.

3.5 Removing sources of error in the inversion scheme

As part of the 'debugging' process, the configuration of the inversion was reviewed, and a number of changes were made in order to reduce the potential sources of error in the results.

3.5.1 Line search

A significant source of uncertainty in inversion results comes from the search length parameter, λ. In practice, for l_2 and TV inversions, pre-scheduled cooling off schemes for λ are often

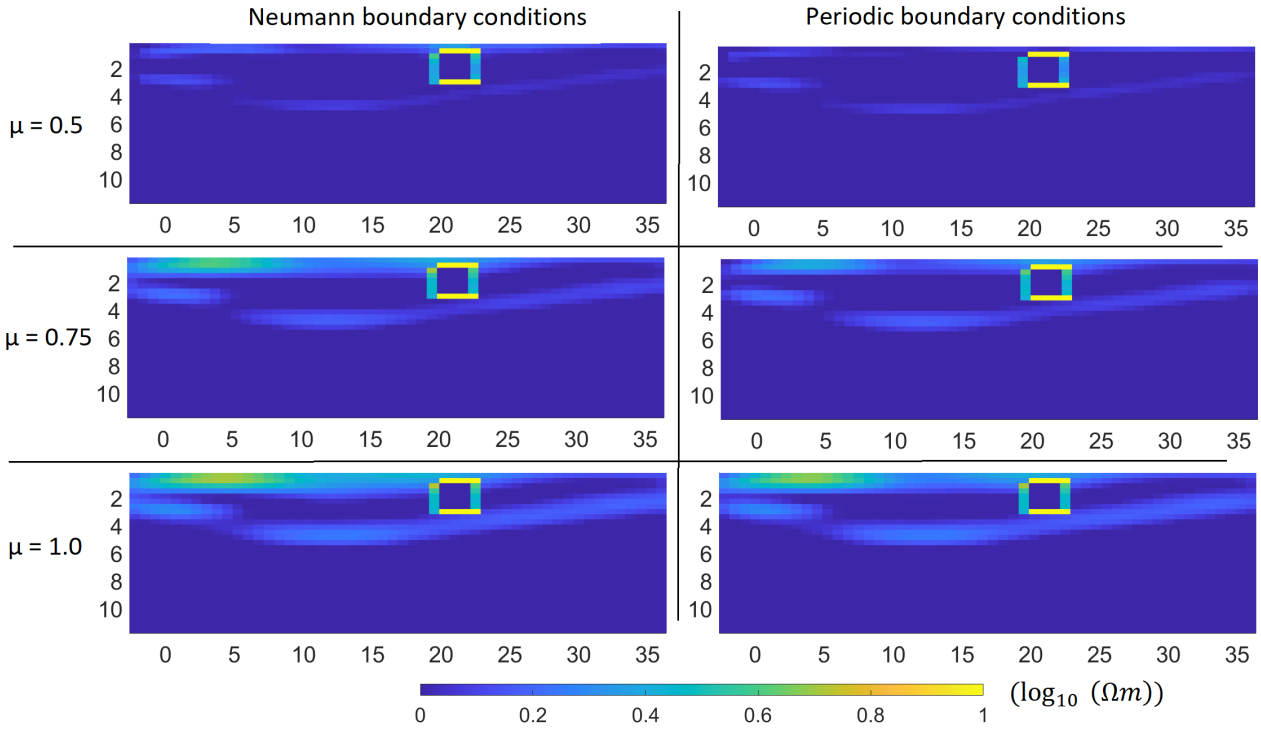


Figure 3.12: Difference between the model gradient and **p** for the piecewise smooth synthetic model (figure 3.10).

used [Farquharson and Oldenburg, 1998]. If the inversion result does not converge well, the schedule is adjusted and the inversion is repeated. This is computationally more efficient, however it also introduces a subjective element to the selection of representative models.

An alternative is to use a line search approach, as used by deGroot Hedlin and Constable [1990]. For each iteration in the Gauss-Newton solution for **m**, equation (3.2.6) is solved for a range of λ. The RMS data misfit for each is compared to the result of the previous iteration, and the model which is closest to a target percentage decrease in data misfit is chosen. This is usually the model with the smallest misfit, however a slower descent is favoured by some so as to avoid local minima in the search space.

As the search space at each iteration is convex, the golden section search method was chosen, due to its improved efficiency Press et al. [2007]. This method evaluates the function across a series of three point brackets to identify the minimum. An additional point is evaluated in the interval containing the largest range in function value, with its position defined by the golden ratio. The function minimum will be found between three of the points: the fourth is discarded. The process is repeated. By using the golden ratio to define the positions of the search points, the search space is covered more efficiently. As the optimal λ varies by orders of magnitude across the inversion, the line search is applied to $\log_{10}(\lambda)$. A total of 11 values of λ were tested on each iteration of the inversion.

3.5.2 Uniform mesh

The mesh used in Section 3.3.2 mesh used model cells increasing by 10% with each layer, as model resolution drops deeper in the model. This creates some inaccuracies in the finite difference model (but not in the forward model which uses a separate mesh). The original IP4DI makes no adjustment for this, however my earlier implementation weighted the finite differences by the inverse distance between cell centres. Accurate gradients are much more important in TGV than in TV, as it is a higher order method.

In future tests, I revert to using a uniform grid for the parameter testing. This avoids any secondary effects in the solutions introduced by errors due to the variable cell sizes.

3.5.3 Extending edge cells

Another potential source of error occurs at the edges of the model. The sensitivity of the measurements can extend laterally outside of the region bounded by the electrodes, and so it is a good precaution to extend the width of the model a few cells either side of the electrode array to prevent this causing errors.

Corner artefacts were observed in the TGV inversions of some field data (not shown). One possible source of this is the sensitivity coverage, or interactions with the boundary conditions at the model boundary. To test this, the model boundary was extended to twice the width of the electrode grid. This did not resolve the errors in the specific model, and therefore these errors were likely due to other sources.

3.5.4 Noise inclusion and error weighting

Noise is absent for the initial test examples, which is unrealistic. In order to ensure that the TGV algorithm is tested under realistic conditions, 1% Gaussian random noise was added to the apparent resistivity data in all future synthetic examples. A 1% error weighting was applied through $\mathbf{W}$. This noise estimate is more accurate than can usually be applied to field data, but this data provides a reasonable upper bound on the performance of the inversion with all three regularisation methods.

Although this contradicts the general approach of simplification, testing the TGV algorithm without noise is not a fair evaluation of its behaviour in realistic use.

3.5.5 Data misfit term

The initial tests used an l_1 norm for the data misfit term, which is more robust to outliers than an l_2 data norm. This was newly implemented into the IP4DI code for this algorithm. The

stability factor, $\gamma_d = 5 \times 10^{-3}$, was not strictly small in these tests, which may have introduced some errors into the solution in a similar way to γ_m and γ_p.

As the synthetic data is only affected by low magnitude Gaussian noise, the need for a robust data norm is reduced. As such, the l_2 norm is used for the data terms in all subsequent tests, in order to simplify the results. Not only does this remove the need to tune another parameter, γ_d, it also removes the possibility of any errors arising from an incorrect implementation of the $l1$ norm.

3.5.6 Convergence of the p subproblem

The convergence criterion for $\mathbf{p}$ used previously was a $< 1\%$ relative change in $|\mathbf{p}|_1$ between subsequent iterations of equation (3.2.9). One disadvantage of this metric is that the denominator changes between subsequent iterations.

In future inversions, the convergence of $\mathbf{p}$ is measured as the change relative to the model gradient norm, $|\nabla \mathbf{m}|_1$. Not only is this constant for any given model, but $\mathbf{p} = \nabla \mathbf{m}$ can be considered to be an approximate 'upper bound' on the expected magnitude of $\mathbf{p}$. I chose the cutoff to be $10^{-5}|\nabla \mathbf{m}|_1$, as being sufficiently small relative to the model gradient for the effect on the solution to be negligible.

3.6 Summary

In this chapter, an alternating minimisation algorithm for the regularisation of the ERT inversion using the Total Generalised Variation was derived and implemented. The results of the initial tests included some promising results, however there were some inconsistencies in the results. Most notably, a breakdown in the capture of gradients near the model boundary by $\mathbf{p}$, unexpected smoothing in the TV solution, some asymmetries in $\mathbf{p}$, and the expected relationship between μ and model smoothness was not universally seen.

In response, a methodical evaluation of the algorithm was performed. The integrity of the forward model was verified. The IRLS stabilising parameters where identified as the source of oversmoothing, and a non-homogenous initial model was introduced to allow γ_p and γ_m to be properly small. Periodic boundary conditions were implemented to both reduce the suppression of $\mathbf{p}$ near the boundary, and to remove any errors caused by the use of one sided finite differences at the boundaries. In addition, several further adjustments were made to the algorithm configuration in order to reduce potential sources of error.

In Chapter 4, a parameter test of the TGV algorithm is carried out, in order to understand how the performance of the algorithm is affected by the important parameters μ, λ, and the

initial model. These results are benchmarked against the TV and l_2 inversion results, and are compared to the expected behaviour of TGV regularisation.

Chapter 4

Parameter exploration for Total Generalised Variation Inversion

In this chapter the improved TGV algorithm is tested, and some best practice guidelines for its use are developed. The limitations of the current implementation are explored, and potential future avenues of work are identified. Working with inverse problems can be challenging, and there are many input settings and variables to tune, therefore it is important to provide some general guidelines for any future users.

As TGV regularisation is an extension of the TV model, the behaviour of the TV model under different parameterisations can be used to establish a baseline from which the TGV specific effects can be distinguished.

The parameters explored in this chapter are: the initialisation of $\mathbf{m}$ and $\mathbf{p}$; the effect of the descent schedule λ; and the TGV trade off parameter, μ. The configuration space of all possible combinations of these parameters is too large to fully explore. Instead, combinations of initialisation and descent schedule are evaluated for a test model, in order to establish the effects of varying each. The synthetic test model chosen for the parameter testing contains distinct smooth and sharp resistivity changes, in order to showcase the piecewise-smooth performance of the TGV inversion.

Evaluation of the inversions is achieved through comparison to the original synthetic model. The model misfit is measured using two indices. The rms model misfit is the root mean squared sum of the differences between the model cells in the inverted and original models. The Structural Similarity Index, or SSIM, instead compares measures of luminance, structure and contrast over groups of pixels (or cells in this case) [Zhou Wang et al., 2004]. This thesis uses the standard form of SSIM, with the exponents all equal to 1. By capturing structural

information, it is possible that the SSIM may be a more suitable tool for measuring the fidelity of structural features in the model than a simple point to point comparison. The two methods are compared at the end of the analysis.

The lower corners of the rectangular model domain used in these inversions are poorly captured by the ERT data, and therefore the resolution there is low. While these cells are present at the inversion stage, I filter out all model cells with a cumulative sensitivity of $< 0.2\%$ prior to the analysis in order to avoid biasing these results with the relatively meaningless fluctuations in these background cells. The inverted models are cropped accordingly. I do display $\mathbf{p}$ in these regions, as wider patterns in $\mathbf{p}$ impact the smoothing effects seen in the higher resolution areas of the model.

4.1 Initialisation of p

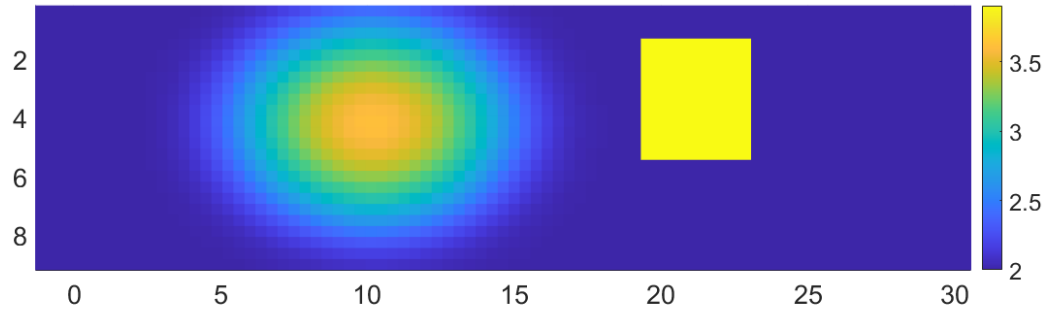


Figure 4.1: Synthetic resistivity model with both smooth and blocky features used to benchmark the inner loop of the TGV algoritm.

The calculation of $\mathbf{p}$ in the inner loop of the alternating-minimisation algorithm (Algorithm 1) depends only on the current resistivity model, and therefore can be tested independently of the inversion.

In the initial testing in Section 3.3, $\mathbf{p}$ was initialised as $\mathbf{p}_x = \mathbf{p}_y = \mathbf{0}$, uniformly throughout the model. With this initialisation, $\nabla\mathbf{p} = \mathbf{0}$ throughout the model, causing Equation (3.2.9) to become singular, unless γ_p and γ_m are sufficiently large compared to the gradients of $\mathbf{p}$ and $\mathbf{m}$. As such, $\gamma_p = \gamma_m = 10^{-3}$ were used in the initial tests, which resulted in undesirable smoothing, as discussed ins Section 3.4.3.

Three alternative initialisations of $\mathbf{p}$ were tested with $\gamma_m = 10^{-12}$ and $\gamma_p = 10^{-12}$. The first initialisation applies a single iteration of the TGV subproblem, replacing the usual l_1 norms

with l_2, to a uniform $\mathbf{p} = \mathbf{0}$ model. During this iteration, $\mathbf{R}_p$ and $\mathbf{R}_m$ in equation (3.2.9) reduce to the identity matrix. The second initialisation sets $\mathbf{p} = \nabla\mathbf{m}$ across the entire model; this solution is expected to be close to the final solution where only linearly-smooth gradients are present. The third initialisation applies random noise, generated from the standard Normal Distribution, to a $\mathbf{p} = 0$ model. This initialisation is far from the true model and was expected to be the most likely to experience difficulties with convergence before the tests were carried out.

The three initialisation methods were tested using the fixed synthetic resistivity model in figure 4.1, which contains a mix of smooth and sharp gradients, for four different values of the trade off parameter, μ. The solution was terminated when the rms change in $\mathbf{p}$ from the last iteration was $< 10^{-5}$ of the rms model gradient.

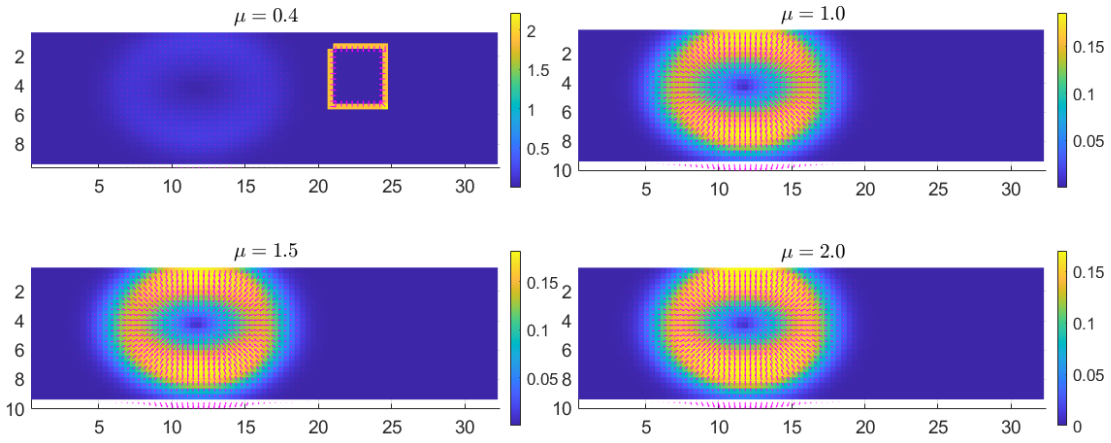


Figure 4.2: Solutions for $\mathbf{p}$ (pink vectors) from the $\mathbf{p}$ initialisation tests. The magnitude, $|\mathbf{p}|$, is also shown as a colour map. Only when μ is small, is $\mathbf{p}$ equal to the gradient on the block edges.

	RMS distance			
Pair of Initial Models Being Compared	$\mu = 1.0$	$\mu = 1.5$	$\mu = 1.5$	$\mu = 2.5$
l_2 vs $\nabla\mathbf{m}$	6.5×10^{-7}	9.4×10^{-7}	2.1×10^{-5}	7.8×10^{-6}
l_2 vs Gaussian	3.4×10^{-6}	6.8×10^{-6}	3.0×10^{-6}	7.6×10^{-6}
$\nabla\mathbf{m}$ vs Gaussian	3.4×10^{-6}	6.7×10^{-6}	1.9×10^{-5}	1.5×10^{-5}

Table 4.1: Root mean squared distances in $|\mathbf{p}|$ between different initialisation methods.

All three initialisations converged on the solution in Figure 4.2, to within $< 2 \times 10^{-5}$ rms distance of one another (see table 4.1). This is small compared to the true model gradient and the ERT modelling errors. The convergence measure (Figure 4.3) decreased monotonically

in each case. There is a small difference in the number of iterations before termination in favour of the Gaussian random noise initialisation, however it is not a significant factor in the overall computational run time of the inversion, as each iteration takes >0.01s in this example, compared to around 10s for the forward modelling process.

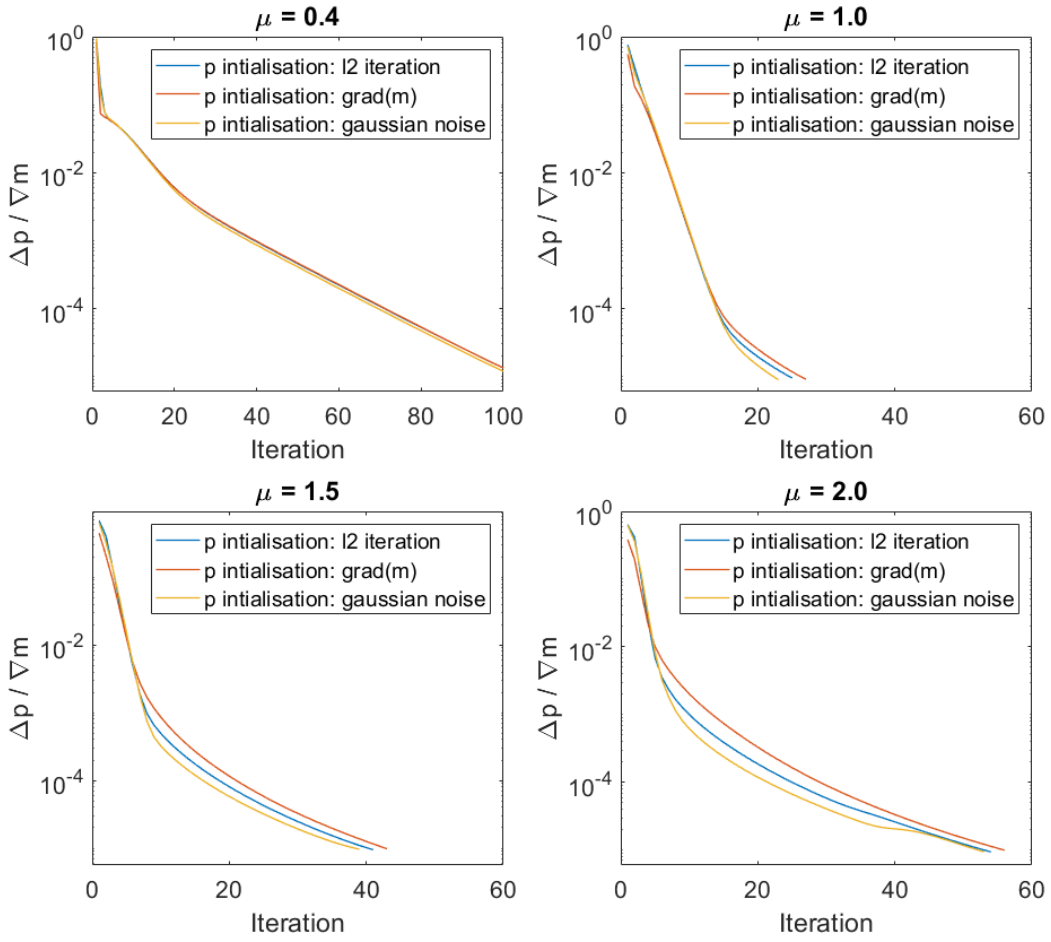


Figure 4.3: The root-mean-square change in **p** with respect to the root-mean-square model gradient, for each iteration of the solution to equation (3.2.9). Show for a range of initialisation of **p**, and different values of μ.

In general, the initialisation of **p**, and the solution of equation (3.2.9) are not significant sources of error in the TGV inversion. All the initialisations are equally effective, with a small, but not significant, computational advantage for the Gaussian noise initialisation. However, it is possible that the random noise initialisation is not well suited to all models, so in general I would recommend using either the l_2 or $\nabla\mathbf{m}$ initialisations, as they will be reliably nearer to the true solution in parameter space, given that the computational cost is small.

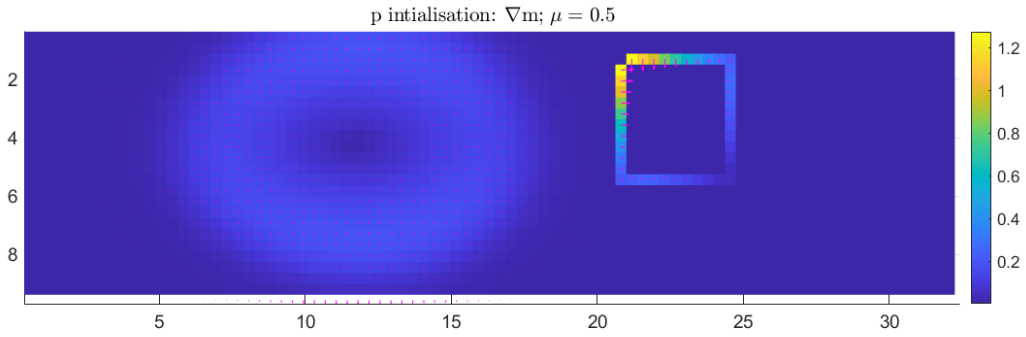


Figure 4.4: $\mathbf{p}$ (pink vectors), for $\mu = 0.5$ for the synthetic model in figure 4.1. The magnitude, $|\mathbf{p}|$, is also shown as a colour map. The solution is affected by asymmetry in $\mathbf{p}$ around the edges of the blocky feature.

4.1.1 Stability of p solution

When $\mu = 0.5$, the symmetry in the solution for $\mathbf{p}$ breaks down around the edge of the block (Figure 4.4). $\mathbf{p}$ gradually decreases around the block edges, from a maximum near the upper left corner to a minimum at the opposite corner. This suggests that the asymmetric gradient approximation, which affects the corner gradient values (as described in 3.4.4), is responsible. $\mu = 0.5$ is an inflection point in the solution: for $\mu > 0.5$, $\mathbf{p} \approx 0$ at the block edge, while $\mathbf{p} \approx \nabla\mathbf{m}$ for $\mu < 0.5$.

This is accompanied by an instability in the convergence rate of $\mathbf{p}$, which begins to oscillate an order of magnitude above the usual convergence criterion of $\nabla\mathbf{p}/\nabla\mathbf{p} = 10^{-5}$, as shown in figure 4.5a. The magnitude of the oscillations are small: there is no substantial change in $\mathbf{p}$ during the oscillatory period.

To better understand the general stability of the convergence rate, the $\mathbf{p}$ solver was iterated past the normal convergence criterion for a range of μ, over the same synthetic model, figure 4.1. Representative examples are shown in figure 4.5b-d. These all decreased monotonically until the convergence rate was at least $< 10^{-10}$, and none experienced any asymmetry in the solutions. Note that the convergence rate decreased at significantly different rates for small changes in μ without any corresponding change in the solution: the difference in the the $\mu = 1.00$ and the $\mu = 1.01$ solution was negligible.

In practice, the specific inflection point at $\mu = 0.5$ is unlikely to be relevant, as these low values of μ do not provide a good balance between smooth and blocky features, which is the aim of TGV regularisation in ERT. However, deviation from a monotonic convergence of $\mathbf{p}$ may

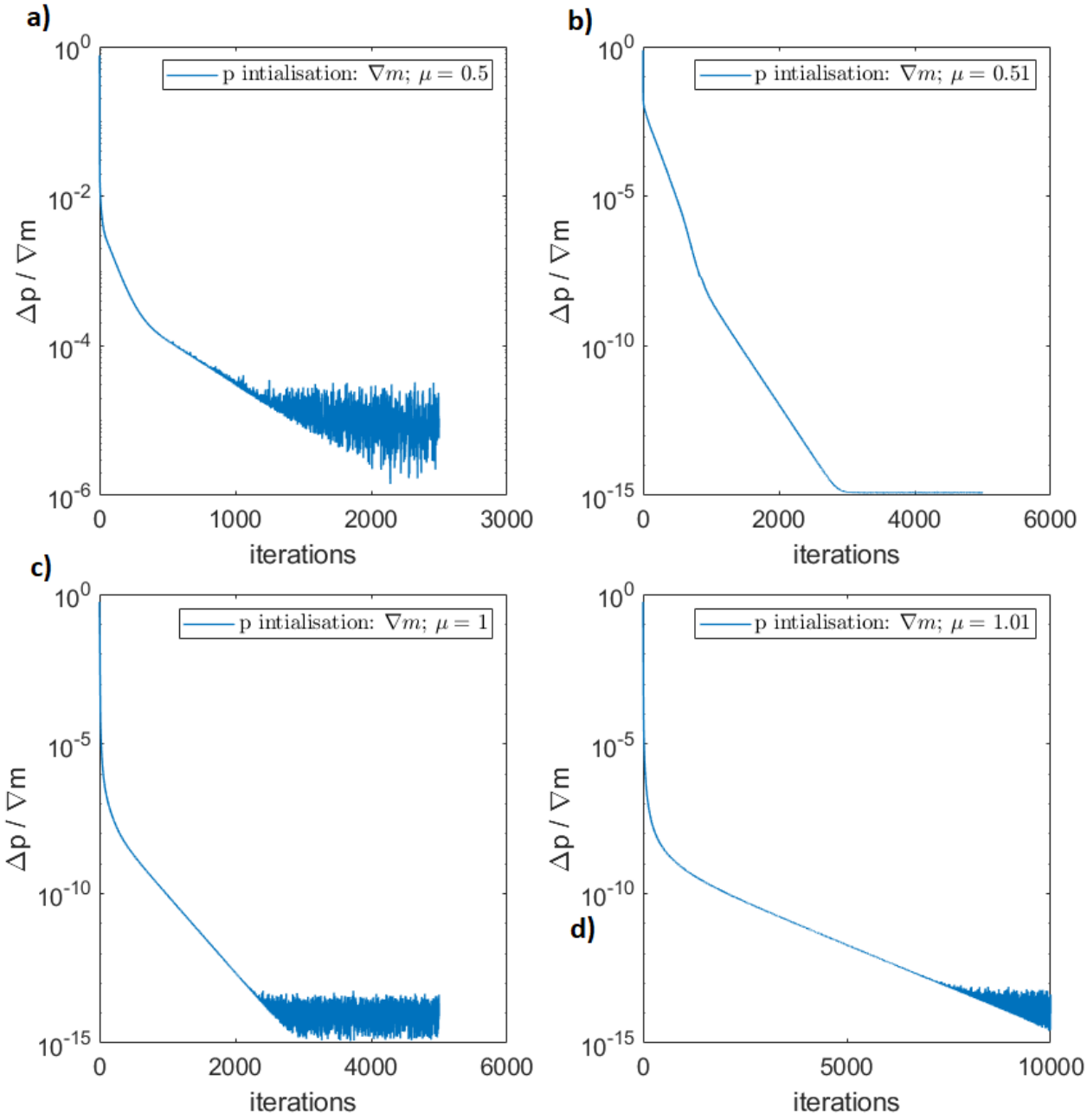


Figure 4.5: Stability of the IRLS algorithm for $\mathbf{p}$ at large number of iterations. These tests used an initialisation of $\mathbf{p} = \nabla \mathbf{m}$ for the synthetic model in figure 4.1. Only the $\mu = 0.5$ model shows any instability at magnitudes $> 10^{-}10$.

be indicative of other errors, and the convergence of $\mathbf{p}$ should therefore be monitored during TGV inversions.

4.2 Baseline for parameter exploration

The alternating-minimisation scheme is a novel addition to the IRLS Gauss-Newton inversion for ERT, and therefore should be thoroughly tested. The solutions to the ERT inversion are local, and so may be affected by the model initialisation, or by the descent path controlled by λ. The TV inversion provides a good baseline with which to evaluate the performance of the TGV algorithm, as the second order TGV functional is an extension of TV.

Synthetic ERT data from the model in Figure 4.6 were used in these tests. These data were generated using the forward modelling software res2dmod [Loke and Barker, 1996], on a rectangular grid with square cells of width 0.1875m. 50 electrodes separated by 0.75m were used. The data consisted of 952 measurements, using a dipole-dipole configurations with parameters $n = 1 - 4, a = 1 - 8$. See Loke et al. [2013] for details of the dipole-dipole array. 1% Gaussian random noise was added to the data.

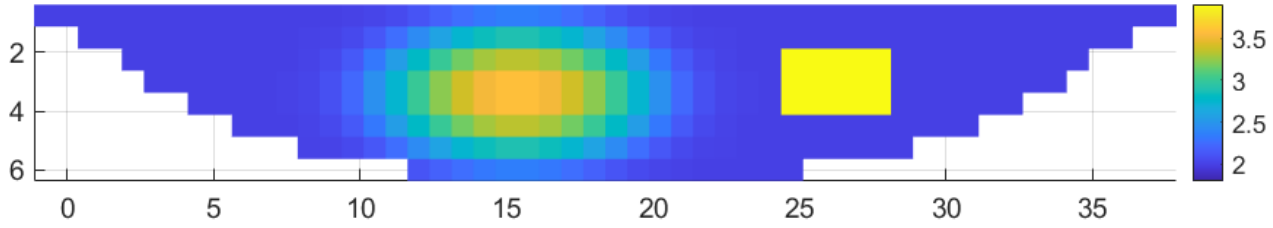


Figure 4.6: Synthetic resistivity model used for parameter testing, containing both smooth and sharp features. The data was generated from a version of this model using the independent forward modelling code Res2DMod, with a rectangular model domain extending to 8m depth.

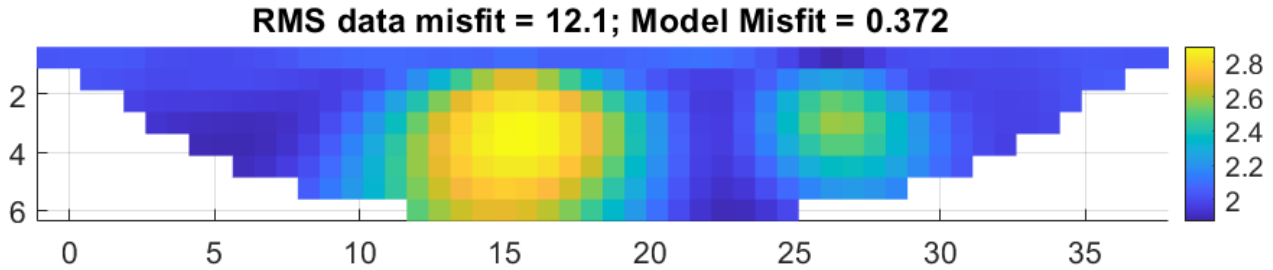


Figure 4.7: Result of the first iteration of an l_2 line search inversion of data from the model in Figure 4.6. This model was used as to initialise TV and TV inversions. The color map is not scaled to match other results in order to provide contrast to the smoothing.

4.2.1 TV inversion: l_2 initialisation

As TGV regularisation is an extension of TV, the behaviour of the TV inversion is a good benchmark from which to understand the performance of TGV regularisation. First, different descent schemes for the Lagrange multiplier, λ, were explored for the same initial model. These inversions were initialised by the first iteration of an l_2 inversion acting on a homogeneous half space, with λ determined by a line search, as used by e.g. [Loke et al., 2003] (Figure 4.7).

The test data were inverted using three different cooling off schemes for λ: a golden section line search; the 'step cool' scheme where λ remains constant for several iterations between sharp decreases; and the 'halving' cooling off scheme, where λ halves each iteration. Both these cooling off schemes begin with larger values of λ which then gradually decrease. This is deliberate, as undesirable structures may appear in the model as a result of overfitting data noise if, in the early iterations, λ is close to the optimum value at convergence [Walker, 1999]. Additionally, larger λ corresponds to larger step sizes through the model solution space, which can enable a faster convergence.

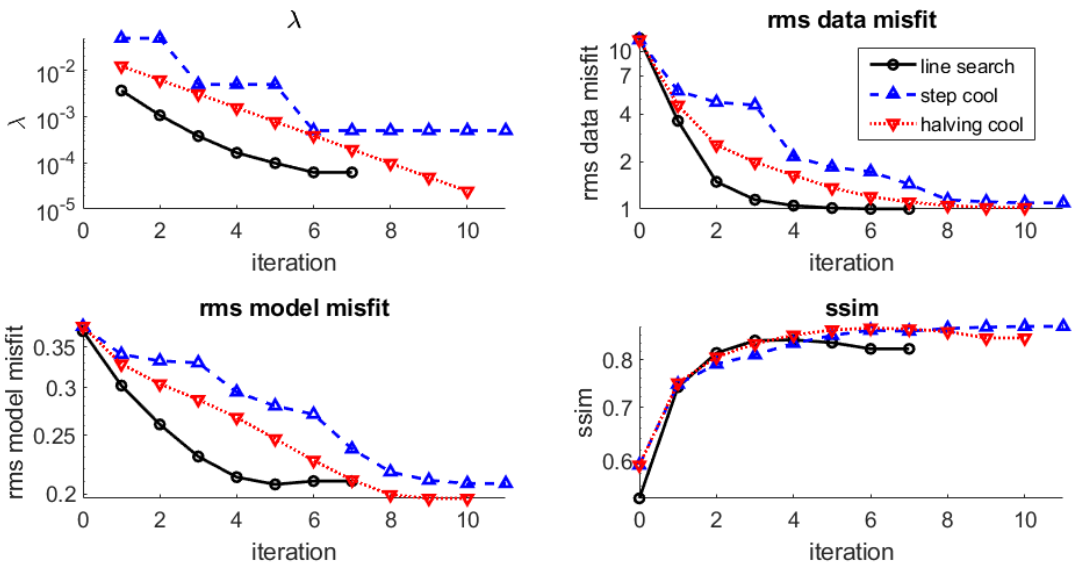


Figure 4.8: Lagrange multiplier, and measures of model and data misfit for line search TGV inversions of the piecewise smooth synthetic model in Figure 4.6. The inversion was initialised by the result of a single iteration of the l2 inversion with $\lambda = 0.447$, determined by a golden-section line search. 'rms' is the root mean square, and the SSIM is the Structural Similarity Index of [Zhou Wang et al., 2004]. 1% Gaussian random noise was present in the data.

These cooling off schemes are shown in figure 4.8, alongside graphs showing the convergence behaviour of the model and data misfit in these inversions. It is expected that a successful in-

version would converge with a data misfit approximately equal to the noise level of 1%.

In all three cases, the rms data misfit decreases monotonically, before converging on a constant value close to 1%. Further tests (not shown) indicate that this convergent behaviour is generally the case when λ is equal to or smaller than the previous iteration, while divergence occurs when λ increases relative to the last iteration, as the inversion overshoots the solution.

The rms model misfit also decreases monotonically, except for during the penultimate iteration of the line search inversion. There, the model misfit increases from 0.207 to 0.210 while the data misfit continues to fall. This therefore is likely due to the overfitting of noise in the data.

The convergence rate of the line search inversion is fastest in terms of both model and data misfit, and converges in the fewest iterations. It is possible that any overfitting is a consequence of this faster convergence behaviour, which enforces less regularisation to the intermediate solutions through smaller λ.

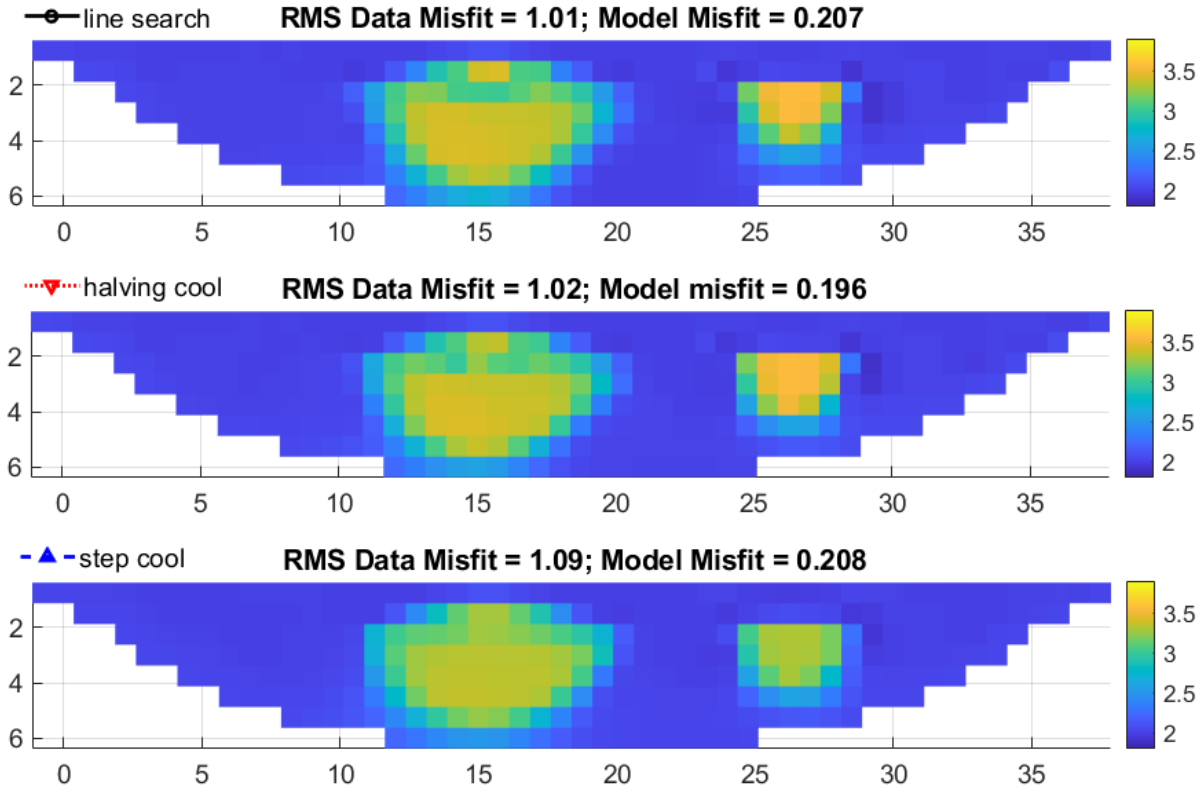


Figure 4.9: Best solutions from the l_2-initialised TV benchmarking inversions in figure 4.8, using data from the model in figure 4.6. These solutions were chosen to have the lowest RMS model misfit. Note that the color scale is truncated relative to 4.6, to better highlight the differences between these models.

The SSIM measure of model misfit favours earlier iterations compared to the rms misfit model, for both the line search and 'halving cool' inversions. These solutions also have larger λ and data misfits relative to the solutions favoured by the rms misfit. The best SSIM solution across

all models is the final iteration of the 'step cool' inversion, with a higher λ and data misfit than the best solution model misfit sense. This suggests that the SSIM measure is favouring models which have a more regular structure then the rms misfit measure, as a larger λ imposes more regularisation on the model solution. The larger data misfit may be due to an intolerance of overfitting in the SSIM, but it could be an indication that the rms data misfit is a more reliable measure of model fidelity.

The best solutions of each inversion, in terms of rms model misfit, are shown in Figure 4.9. All three are structurally similar. The smooth feature on the left contains a large constant-resistivity peak, obscuring the smooth changes in the original model, Figure 4.6, while the resistive block to the right is better defined. The peak resistance of each feature is underestimated in each case.

For comparison, the equivalent solution for an l_2 line search inversion is Figure 4.10. Both features contain smooth gradients: appropriate for the smooth feature, but a poor representation of the constant-resistivity block. The constant-resistivity background region is also less well reconstructed, blending into the edges of the block, resulting in an overall poorer model misfit than for the TV solutions.

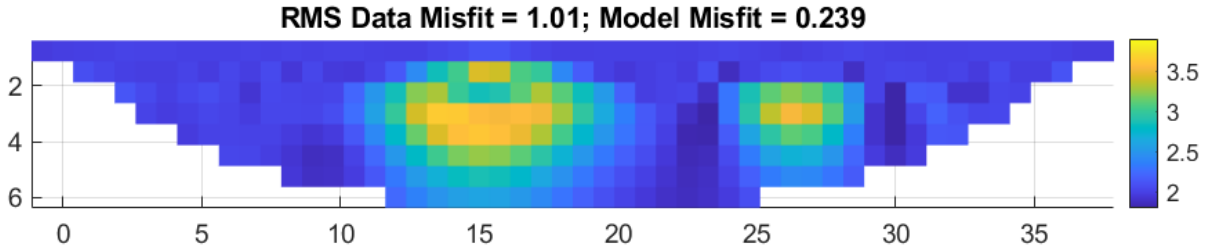


Figure 4.10: Best solution, in a model misfit sense, from an l2 line search inversion, using data from the model in Figure 4.6.

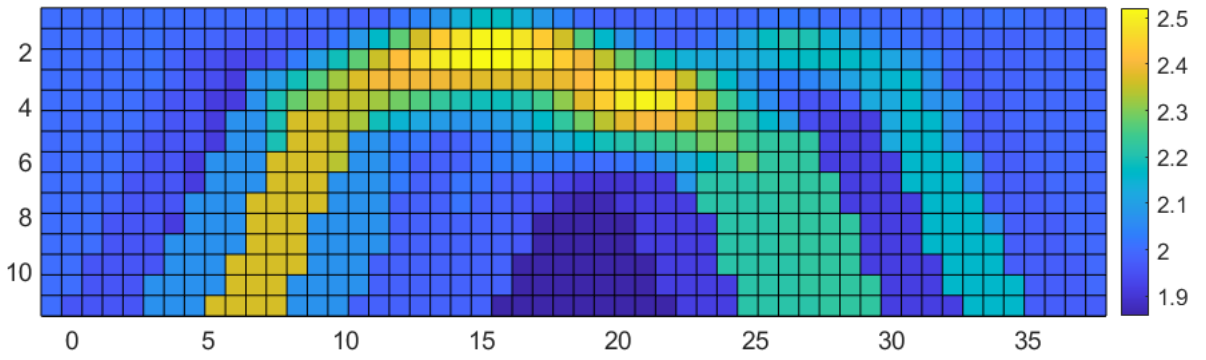


Figure 4.11: Pseudosection of data from 4.6, with 1% noise added. Used for initialisation of TV and TGV inversions

Overall, the differences between the $l1$ initialised TV solutions are small for any practical purpose. The 'step cool' solution, with a larger regularisation term, produces a flatter, more reg-

ularised model, and underestimates the resistivity of the block by a larger margin. The more slowly converging 'halving cool' method outperforms the line search method in this case, producing a slightly smaller model misfit, possibly through the avoidance of overfitting.

4.2.2 TV inversion: psuedosection initialisation

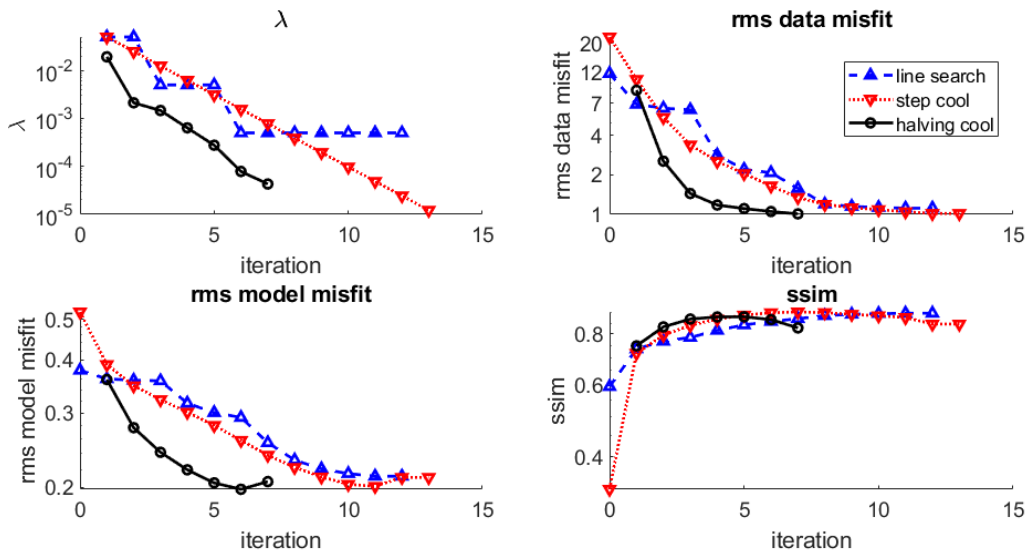


Figure 4.12: Lagrange multiplier, and measures of model and data misfit for line search TGV inversions of the piecewise smooth synthetic model in Figure 4.6, initialised by the model pseudosection, Figure 4.11. 'rms' is the root mean square, and the SSIM is the Structural Similarity Index of [Zhou Wang et al., 2004]. 1% Gaussian random noise was present in the data.

To test the sensitivity of the TV inversion to the initial model, the TV inversions were repeated using an alternative initialisation: a representation of the pseudosection of the ERT data. This is a spatial projection of the raw ERT apparent resistivity data. Each datum is horizontally positioned at the midpoint of the four electrodes that make up the quadropole. The depth is approximated by the 'median depth of investigation'. Half the integrated sensitivity of the measurement lies above the median depth, and half below it [Barker, 1992]. An interpolation was used to project these onto the inversion mesh: a linear interpolation was used between pseudosection points, while the cells outside of the pseudosection area were determined using a nearest neighbour extrapolation. This pseudosection is shown in Figure 4.11

The convergence plots, Figure 4.12, show similar interrelationships between the cooling schedules to those found in the l_2 initialisation inversions. The most notable differences are that the line search solution performs best in terms of model misfit, and there is evidence of overfitting in the rms model misfit of the 'halving cool' scheme in addition to the line search. The line search and 'halving cool' methods converged at a higher data misfit 1.05% compared to

1.01% – 1.02% for the l_2 initialisations.

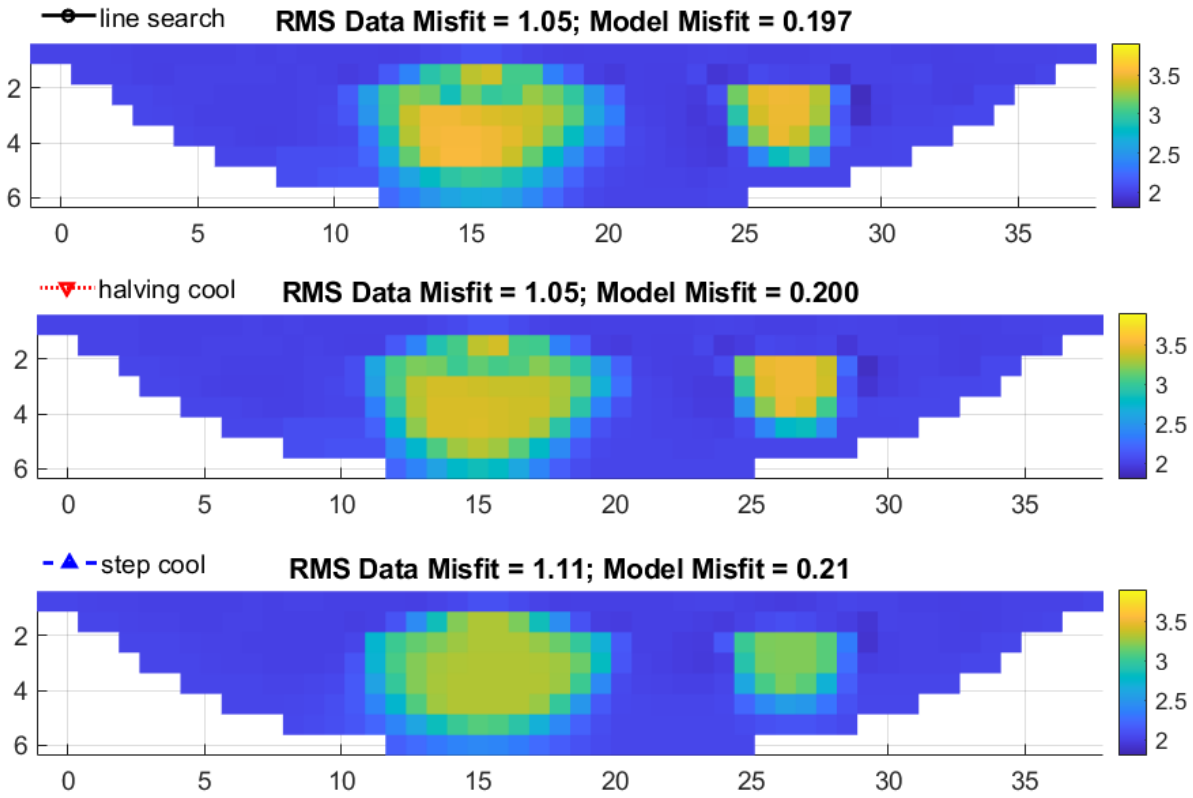


Figure 4.13: Best solutions from the pseudosection-initialised TV benchmarking inversions in Figure 4.12, using data from the model in Figure 4.6. These solutions were chosen to have the lowest RMS model misfit. Note that the color scale is truncated relative to 4.6, to better highlight the differences between these models.

The best solution for each pseudosection-initialised TV inversion are shown in Figure 4.13. The smooth feature in the line search solution contains smoother gradients than in the other solutions. It is also noticeably asymmetrical in the horizontal plane. It is likely some of the asymmetry present in the pseudosection has been preserved during the faster convergence of the line search solution (Figure 4.11). The higher data misfit compared to the l_2 solutions supports the idea that this smoothing is due to fortuitously inherited structure from the initialisation, not through a better fit to the data.

The other solutions are more regular, and structurally resemble the l_2 initialisation results (Figure 4.9). The larger influence of the regularisation term in the early iterations, compared to the line search, may be responsible for smoothing these details. The 'step cool' solution converges at a larger data and model misfit than the 'halving cool' scheme for this initialisation too.

Overall, the change in initialisation results in small changes in the model solutions. In terms of model misfit, the differences between the different initialisation methods are of the same order as the differences between the Lagrange multiplier schemes. The initialisation only had a sig-

nificant impact on the solution in one of the models, which was the pseudosection-initialised line search. Indeed, the rapid descent of the line search scheme may not necessarily result in a superior model, due to overfitting and the preservation of structures from the initial model, despite its advantages in objectivity, reliability, and convergence rate.

4.3 Parameter testing: TGV analysis

The same model was used to carry out parameter testing on the TGV inversion, investigating the effects of λ, the initial model, and μ on the inversion. The TGV trade-off parameter, μ, introduces an additional degree of freedom into the parameter tests. For this reason, these tests will each vary a subset of parameters, in order to isolate the influence of each one.

The constraint $\lambda > 10^{-4}$ was used, as smaller values occasionally caused instabilities in the inversion. These instabilities often led to artefacts, including single model cells with much higher or lower resistivity than the surrounding ones. In some cases the matrix inversion in Equation (3.2.6) became singular. This is likely due to insufficient regularisation.

4.3.1 TGV Parameter Testing: Initialisation and Line Search

First, a baseline performance of the TGV inversion was established, using the two initialisation methods above: an l_2 iteration and the pseudosection. A line search scheme was chosen for λ, as a reliable method which is not dependent on operator trial and error. Three values of μ were tested, in order to capture the broad impacts of varying μ for each initialisation, as the impact of the initialisation may differ with μ. μ is assessed more thoroughly in Section 4.3.5. These were chosen as $\mu = 1.0, \mu = 1.5, \mu = 2.0$, to bracket the range in which μ contributes significantly to the inversion in the initial tests (Section 3.3.3).

Figures 4.14 and 4.15 respectively show the convergence metrics from the l_2 and pseudosection initialised inversions.

The pseudosection initialised inversions are well behaved, and converge at a data misfit of $\approx 1\%$ although the $\mu = 1.0$ solution does not fit the model as well as the TV solutions. The $\mu = 2.0$ and $\mu = 1.5$ l_2-initialised inversions converge similarly well, although the $\mu = 1.5$ solution does not fit the model as well. On the other hand, the l_2-initialised inversions for $\mu = 1.0$ behaves differently. λ does not monotonically decrease. The data misfit of 1.3% at convergence is large relative to the noise of 1%, and the model misfits are relatively poor.

The SSIM measure of misfit largely agrees with the rms model misfit in terms of model preference, although it has a smaller preference for the earlier iterations of the pseudosection initialised inversion than the rms model misfit. While this might be indicative of overfitting, they

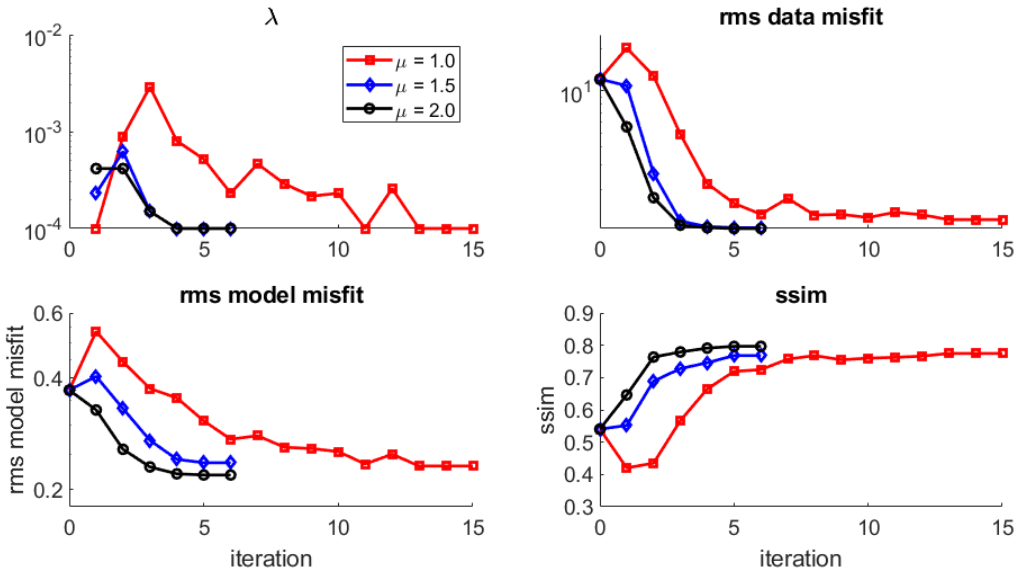


Figure 4.14: Lagrange multiplier, and measures of model and data misfit for line search TGV inversions of the piecewise smooth synthetic model in Figure 4.6, initialised by the first iteration of an l_2 inversion. 'rms' is the root mean square, and the SSIM is the Structural Similarity Index of [Zhou Wang et al., 2004]. 1% Gaussian random noise was present in the data.

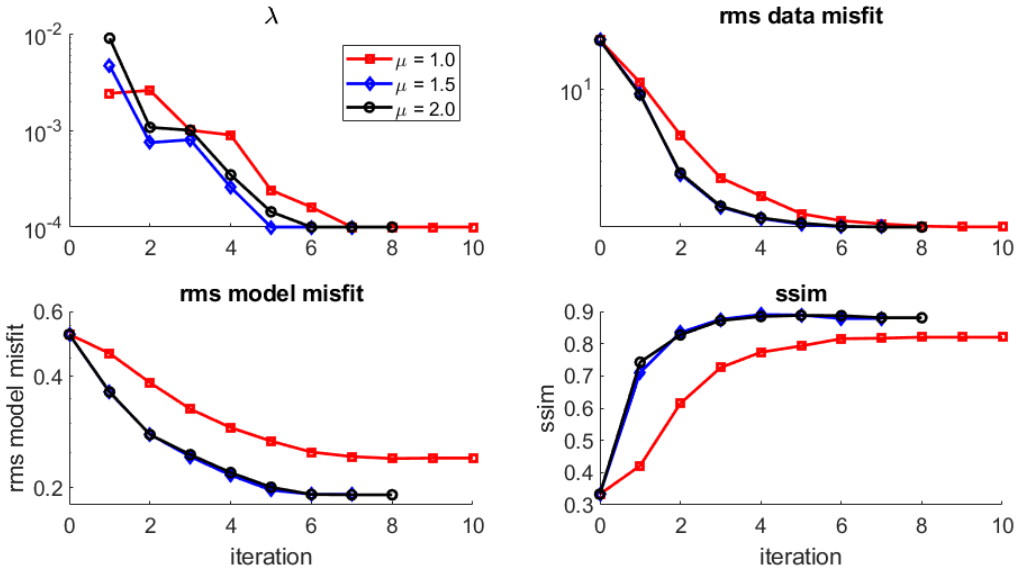


Figure 4.15: Lagrange multiplier, and measures of model and data misfit for line search TGV inversions of the piecewise smooth synthetic model in Figure 4.6, initialised by the model pseudosection, Figure 4.11. 'rms' is the root mean square, and the SSIM is the Structural Similarity Index of [Zhou Wang et al., 2004]. 1% Gaussian random noise was present in the data.

have higher data misfit than the rms misfit preferred solutions, which suggests that the rms model misfit is a more useful measure.

The results of these inversions, and the corresponding $\mathbf{p}$, are shown in Figure 4.16. There is a noticeable difference between the two initialisations for all μ. The l_2 initialised models tend to contain stronger linear gradients while the pseudosection initialised models contain larger constant resistivity regions. For both initialisations, the models become less smooth as μ increases, alongside a corresponding decrease in $\mathbf{p}$.

Just as in the initial tests, $\mathbf{p}$ tends to be suppressed near the upper boundary of the model, resulting in TV-like regularisation in these areas, despite the use of periodic boundary conditions. Only for the l_2 initialisation at $\mu = 1.0$ does $\mathbf{p}$ capture a gradient in the upper regions of the smooth feature. However, unlike the symmetrical original model, this gradient is weaker than that of the lower half of the feature. This solution also contains a spurious vertical gradient in the constant resistivity block, and fits the model poorly. The pseudosection-initialised solution for $\mu = 1.0$, also contains a weak gradient inside the block. This, alongside the relatively high model misfits of these solutions, suggests that $\mu = 1.0$, which weights linear gradients equally to sharp edges, is too weak a constraint to capture piecewise constant regions. It appears to be the case that TGV inversions may revert to TV-like behaviour close to the electrodes, especially if a sharp change in the orientation of $\mathbf{p}$ is required.

The models produced by the l_2 initialisation contain strong linear gradients in the lower half of the smooth feature. These gradients follow $\mathbf{p}$ closely in these areas, to a greater extent than in the pseudosection-initialised models. These gradients are particularly strong in the vertical direction; the horizontal components of $\mathbf{p}$ are much weaker, and as a result the models are somewhat blocky laterally. At the peak of these gradients, the resistivity is overestimated by over 10%, more than in any of the other solutions. These features appear to be responsible for the high model and data misfits of these solutions.

The pseudosection-initialised $\mu = 1.0$ solution is also noticeably smooth, with $\mathbf{p}$ driving a strong linear gradient in the lower half of the smooth feature, alongside a weaker gradient in the blocky feature. The high model misfit reflects this. This suggests that $\text{TGV}_{1.0}$, which weights linear gradients equally to sharp edges, is too weak a constraint for piecewise smooth ERT inversion.

For $\mu = 1.5$ and $\mu = 2$, the pseudosection-initialised models are similar to the TV solution, with a constant resistivity peak at the centre of the smooth feature, close to the true resistivity. These were the most successful inversions in terms of rms model misfit. The smooth feature in these solutions are slightly asymmetrical in the horizontal plane, similar to the equivalent TV line search solution (Figure 4.13). This is accentuated in the $\text{TGV}_{1.5}$ solution, where $\mathbf{p}$ drives an

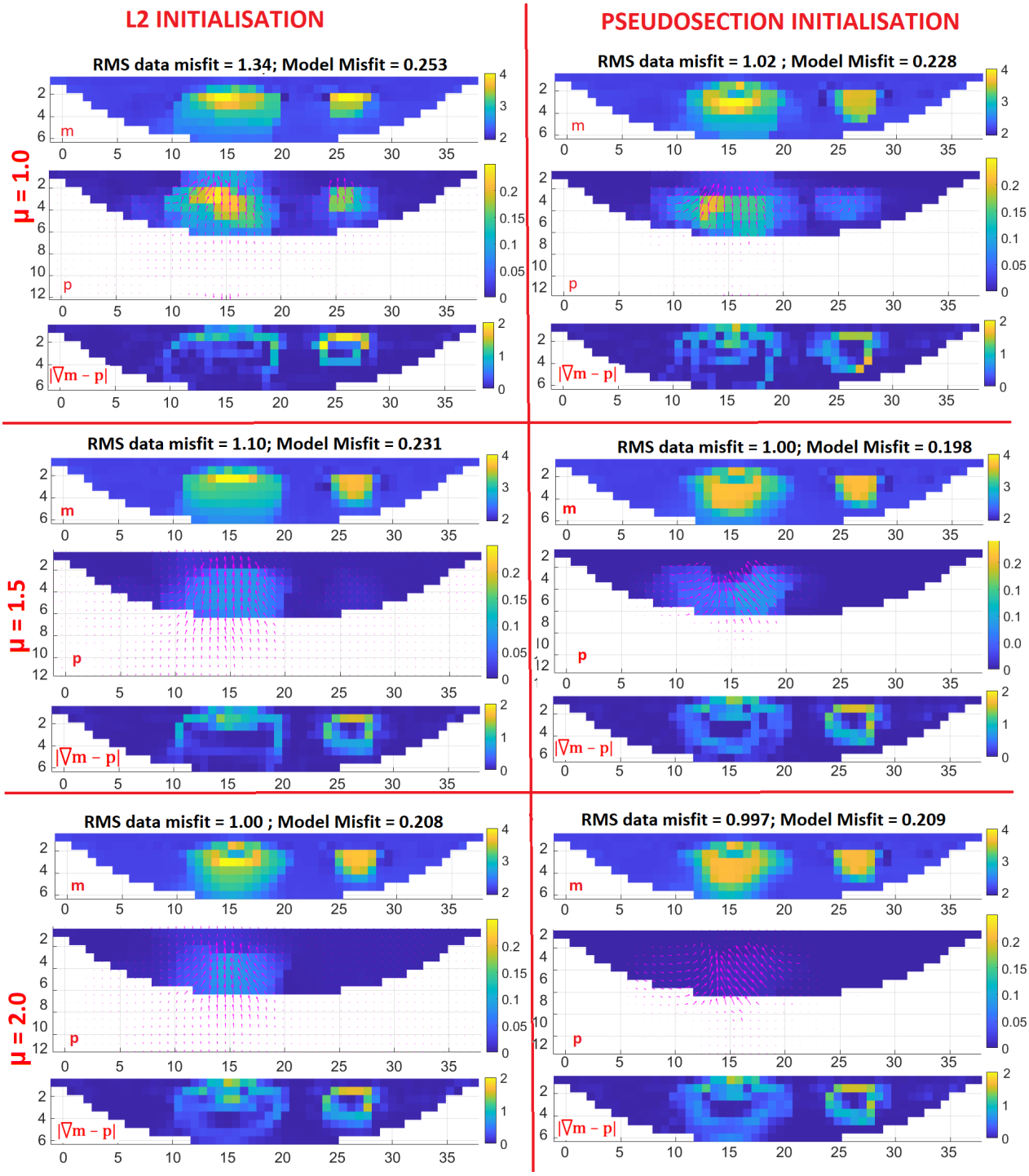


Figure 4.16: TGV inversion results for two different initial models, using a golden section line search for λ, for trade off parameters $\mu = 1.0, 1.5$, and 2.0. The initialisation was a single iteration of an l_2 inversion, and the model pseudosection (Figure 4.11). The synthetic data was generated from the model in Figure 4.6, with 1% Gaussian random noise added. These are the best iterations in terms of data misfit. $\mathbf{p}$ (pink vectors) is accompanied by a colour map of the magnitude, $|\mathbf{p}|$.

extended gradient in the lower right hand region of the feature. This is the only region where a non-zero $\nabla \mathbf{m} = \mathbf{p}$ in the model, in other areas, $\mathbf{p}$, contributes to the gradient, but it does not define it in the manner of the l_2-initialised models. As a similar asymmetry appears in the TV solution, it is likely that this asymettric smoothing is a consequence of the initial model and not of the TGV inversion itself.

The choice of initial model has a much greater effect on the TGV solution than the TV inversion. The new alternating minimisation algorithm introduces an additional parameter, $\mathbf{p}$, which depends on the model gradient of the previous solution. It seems likely that this tends to limit the changes in $\nabla \mathbf{m}$ between one iteration and the next, resulting in local solutions with gradients similar to the initial model.

In the TV inversion, the line search for λ resulted in a stronger influence of the initial model on the solution than alternative cooling schemes. Next, in Section 4.3.2, the extent to which this applies to TGV inversions is tested.

4.3.2 TGV Parameter Testing: 'Halving Cool' Schedule in λ

The TGV inversions for each initialisation were repeated using the 'halving cool' schedule for λ, where the Lagrange multiplier halves every iteration until a minimum of $\lambda = 10^{-4}$.

The l_2-initialised 'halving cool' inversion, Figure 4.17 converged more slowly than the line search equivalent (Figure 4.14), and resulted in significantly higher model misfits across all solutions.

The pseudosection initialisation, Figure 4.18, converged at a comparable rate to the line search (Figure 4.15). The $\text{TGV}_{1.5}$ and $\text{TGV}_{2.0}$ solutions are very similar to the line search equivalents, but perform slightly better in terms of misfit measures. These solutions, Figure 4.19, resemble the line search equivalents (Figure 4.16) except for an increase in the magnitude of $\mathbf{p}$ on the lower right hand side of the smooth feature, leading to smoother gradients which closer match the original model. This is the area in which asymmetries already arise in the TV and line search TGV models. The TGV inversion here accentuates a gradient which appears in the initial model.

The smoothing parameter, $\mathbf{p}$, has an influence over a much greater extent of the l_2-initialised 'halving cool' solutions than in the line search solutions, where $\mathbf{p}$ is negligible in the 100 Ωm background areas of the model. This is reminiscent of the broad smoothing in the initial l_2 model, Figure 4.7.

As a result, resistivity gradients extend into the model background. The shape of the piecewise smooth feature is also adversely affected at $\mu = 1.5$.

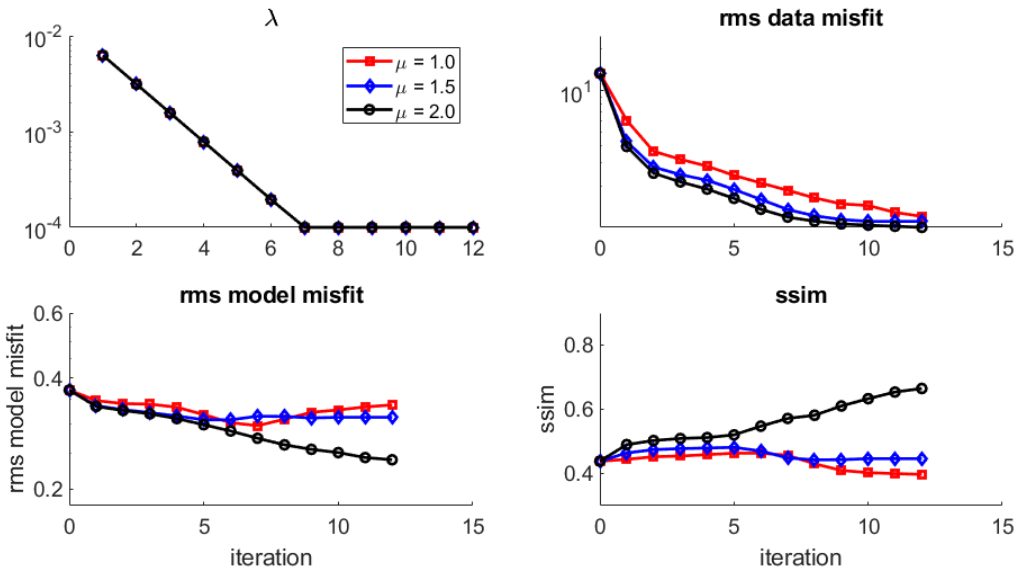


Figure 4.17: Lagrange multiplier, and measures of model and data misfit for TGV inversions of the piecewise smooth synthetic model in Figure 4.6, using the 'halving cool' schedule for λ, initialised by the first iteration of an l_2 inversion. 'rms' is the root mean square, and the SSIM is the Structural Similarity Index of [Zhou Wang et al., 2004]. 1% Gaussian random noise was present in the data.

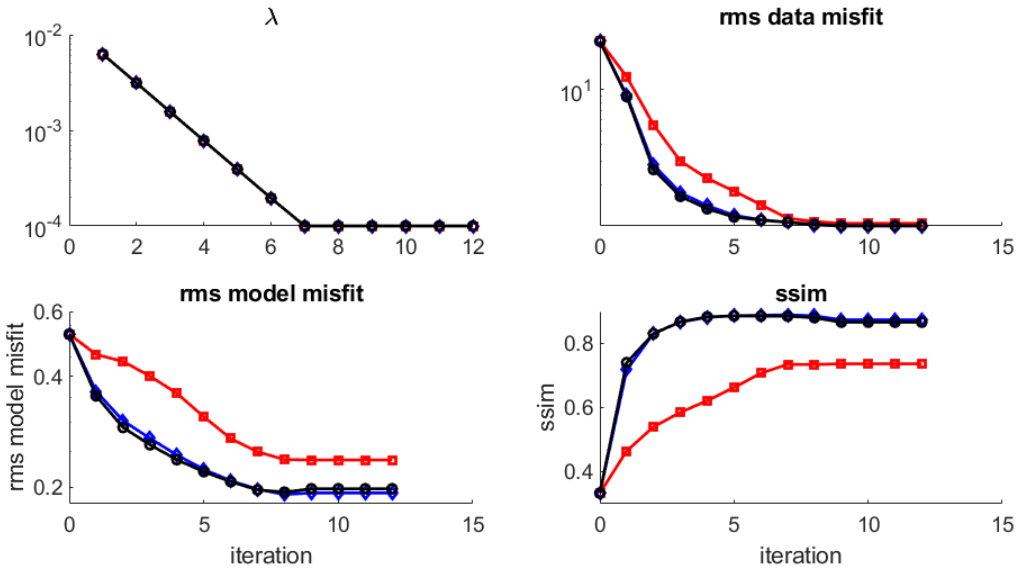


Figure 4.18: Lagrange multiplier, and measures of model and data misfit for TGV inversions of the piecewise smooth synthetic model in Figure 4.6, using the 'halving cool' schedule for λ, initialised by the model psuedosection. 'rms' is the root mean square, and the SSIM is the Structural Similarity Index of [Zhou Wang et al., 2004]. 1% Gaussian random noise was present in the data.

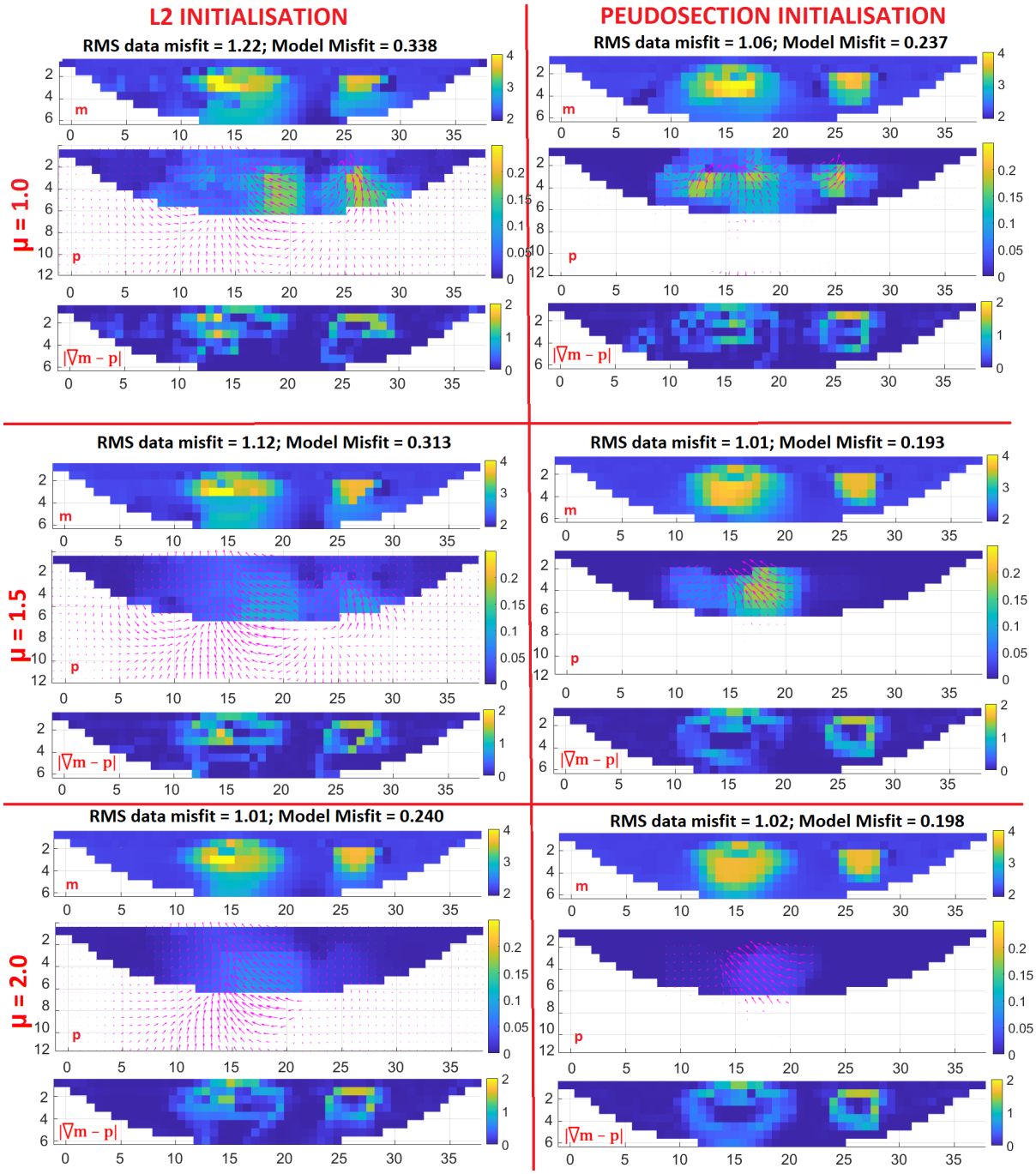


Figure 4.19: TGV inversion results for two different initial models, using the 'halving cool' scheme for λ, for trade off parameters $\mu = 1.0, 1.5$, and 2.0. The initialisation used were a single iteration of an l_2 inversion, and the model pseudosection (Figure 4.11). The synthetic data was generated from the model in Figure 4.6, with 1% Gaussian random noise added. These are the best iterations in terms of data misfit. $\mathbf{p}$ (pink vectors) is accompanied by a colour map of the magnitude, $|\mathbf{p}|$.

For both initial models, the use of the 'halving' schedule for λ resulted in increases in $\mathbf{p}$, and hence additional smoothing, in the solutions. The areas of additional smoothing related to the original models, and is further evidence of a strong hysteresis, or initialisation dependence, on the descent path of the TGV inversion.

In the case of the l_2 initialisation, this smoothing affected the constant resistivity background and blocky feature, resulting in a poorer model misfit. For the pseudosection initialisation, the additional smoothing was localised to one quadrant of the smooth feature where the original model was also smooth, leading to an improved model misfit for $\mu = 1.5$ and $\mu = 2$.

4.3.3 TGV Search Space

The IRLS Gauss Newton method involves linearising the inverse problem around the current solution at each iteration, in order to form a local approximation of the objective function. The TGV regularisation functional is evaluated at each of these intermediate solutions, and applies and additional constraint on the solution based on the gradient of the previous solution. It is possible that the influence of the model gradient from previous iterations biases the inversion strongly towards models with similar gradient structures, and that larger step sizes are required in the inversion to allow significant changes in $\mathbf{p}$. Alternatively, it may be that a greater number of incremental evaluations of the TGV functional are required to allow cumulative changes in $\mathbf{p}$ to take effect.

A 'slow' cooling schedule, in which λ decreases by 20% each iteration, was chosen to test whether a more gradual convergence process, with more incremental evaluation of the TGV functional might allow the TGV functional to be less affected by the initial model. However, the l_2-initialised inversion, Figure 4.20, continues to converge at a much higher model misfit than the pseudosection-initialised inversion, Figure 4.21. The solutions themselves (not shown) resemble the 'halving cool' results in Figure 4.19 closely, although it should be noted that the $\text{TGV}_{1.5}$ result achieves a lower model misfit than both the line search and 'halving' cool methods, 0.186. This is largely due to small differences in the gradient in the lower right quadrant of the smooth feature, where $\mathbf{p}$ is somewhat smaller than in the 'halving cool' solution, 0.14, compared to 0.21.

The more gradual convergence process failed to result in a different gradient structure of the solutions. Therefore, it is expected that the bias towards local solutions is due to too small step sizes, rather than too large ones. This regime is much more difficult to navigate with a cooling schedule than a gradual convergence; large step sizes (large λ) have the potential to overshoot local minima, and so the schedule of the decreases in λ can be very important.

The impact of the initial model persists in this 'slow' cooling schedule too. In Section 4.3.4, the

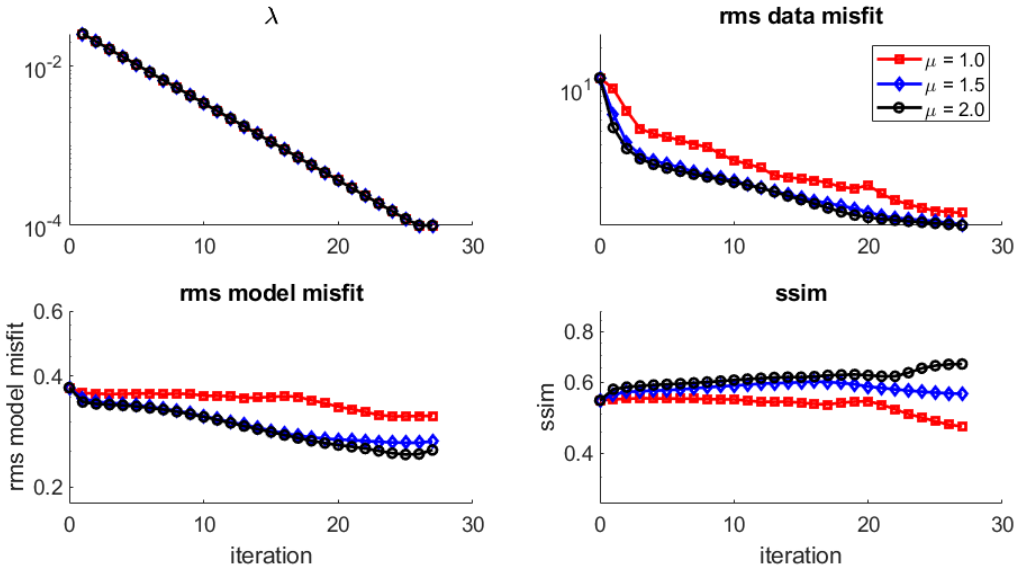


Figure 4.20: Lagrange multiplier, and measures of model and data misfit for TGV inversions of the piecewise smooth synthetic model in Figure 4.6, using the 'slow cool' schedule for λ, initialised by the model psuedosection. 'rms' is the root mean square, and the SSIM is the Structural Similarity Index of [Zhou Wang et al., 2004]. 1% Gaussian random noise was present in the data.

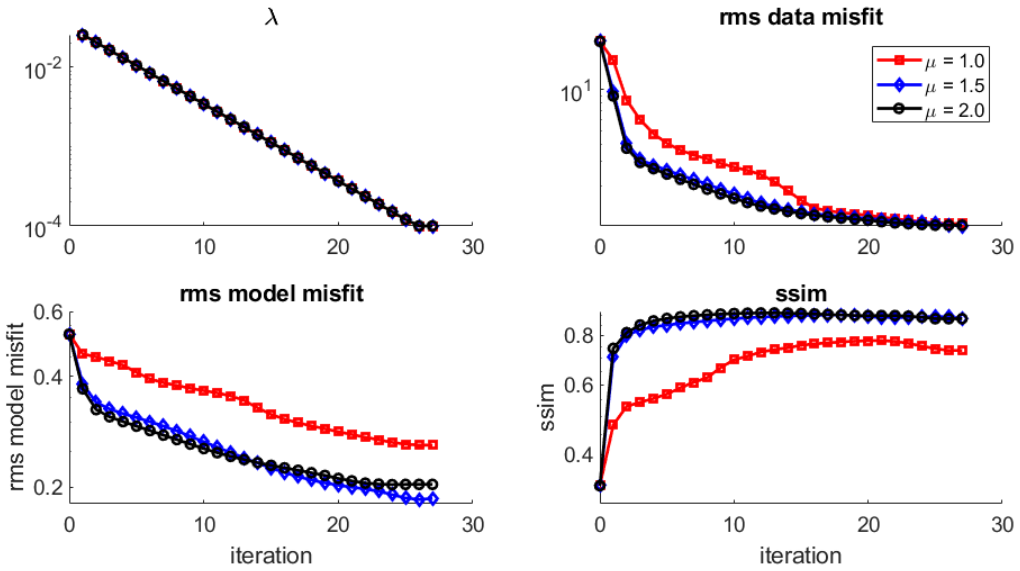


Figure 4.21: Lagrange multiplier, and measures of model and data misfit for TGV inversions of the piecewise smooth synthetic model in Figure 4.6, using the 'slow cool' schedule for λ, initialised by the model psuedosection. 'rms' is the root mean square, and the SSIM is the Structural Similarity Index of [Zhou Wang et al., 2004]. 1% Gaussian random noise was present in the data.

solution to a previous inversion is used as the initial model, in order to reduce the influence of initial models which are far from the true solution.

4.3.4 Parameter Testing: 'Double' inversion

The TGV solutions have been shown to be strongly influenced by the initial model used in the inversion. For the TV inversion, the rapidly converging line search descent method resulted in solutions which were more heavily influenced by the initial model than other, more gradual, λ schedules (see Section 4.2.2). It would be desirable to find a similar means of reducing the influence of the initial model on the solution of the TGV inversion.

Ideally, one would widen the search space in an attempt to find a global minimum. However, such an approach is completely unfeasible due to the computational demands. Instead, I approach the problem by choosing an initial model which is closer to the final solution. The gradients in the model will be closer to their final values, and the changes in the solution between iterations will be smaller and more incremental. These changes in the model will be more focused on the 'fine tuning' of the model structure, and less on defining the broad structural features. This may allow the TGV functional to update more smoothly.

The natural initial models 'close to the final solution', are the solutions of previous inversions. I used the solutions to the TGV line search inversions from Figure 4.16 to initialise a second set of inversions, with the same μ as the original solutions. I did this to avoid biasing the inversions with the effects of different regularisation functions. For simplicity, I call these 'double' inversions. The 'pseudosection-initialised double inversion' is an inversion which is initialised by the solution to a previous pseudosection-initialised inversion, for example.

A 'halving cool' schedule is used, as the initial model is already a local minimum, so a line search in λ will not result in further improvements to the solution. Figure 4.20, shows that the l_2-initialised double inversion results in an improved rms model misfit for all three values of μ, as well as an improvement in data misfit for the $TGV_{1.5}$ and $TGV_{1.0}$ solutions. All three converge smoothly on a solution, although $TGV_{1.0}$ performs notably poorer. The pseudosection-initialised double inversions, Figure 4.21, all converge at a similar rate of data misfit to one another, with the $TGV_{1.5}$ solution performing best in terms of model misfit. A small increase in model misfit prior to convergence may be indicative of overfitting.

The SSIM order of preference of the different solutions is similar to the RMS model misfit, and does not provide any significant insights.

The solutions to all the double inversions, Figure 4.24, are much more consistent and symmetrical with one another, with smaller model misfits than the line search and 'halving' cool solutions. There remain traces of the broad differences between the original initialisations,

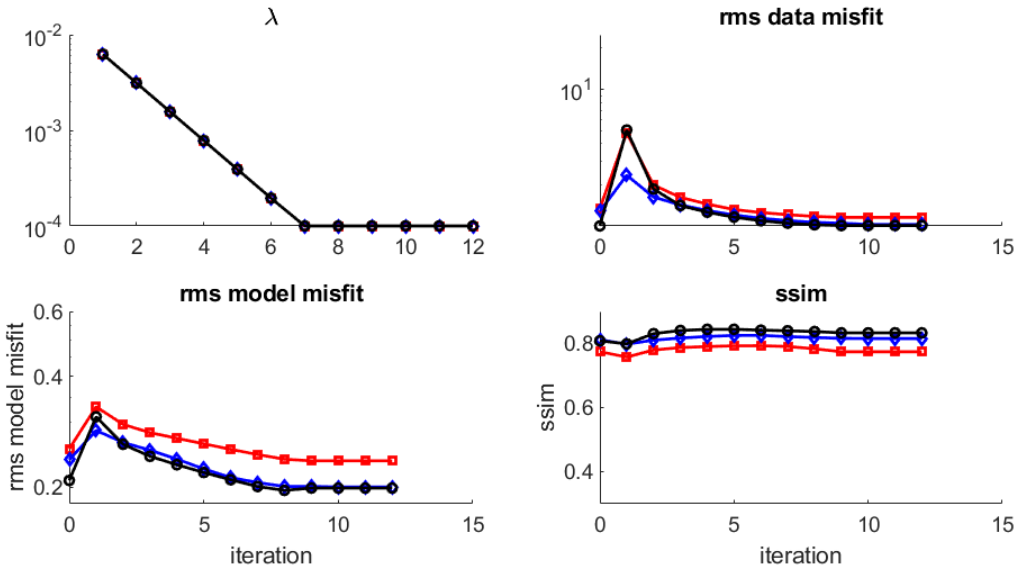


Figure 4.22: TGV 'double inversions' of the synthetic data from Figure 4.6, with 1% Gaussian noise, for different values of μ. An l_2-initialised line search inversion was used for the initial model. The 'halving cool' scheme for λ was used. (a) shows the iteration with the smallest rms model misfit; (b) shows the line search damping parameter, λ; (c) is the root-mean-square data misfit; (d) is the root-mean-square misfit to the original model; and (e) is the structural similarity index.

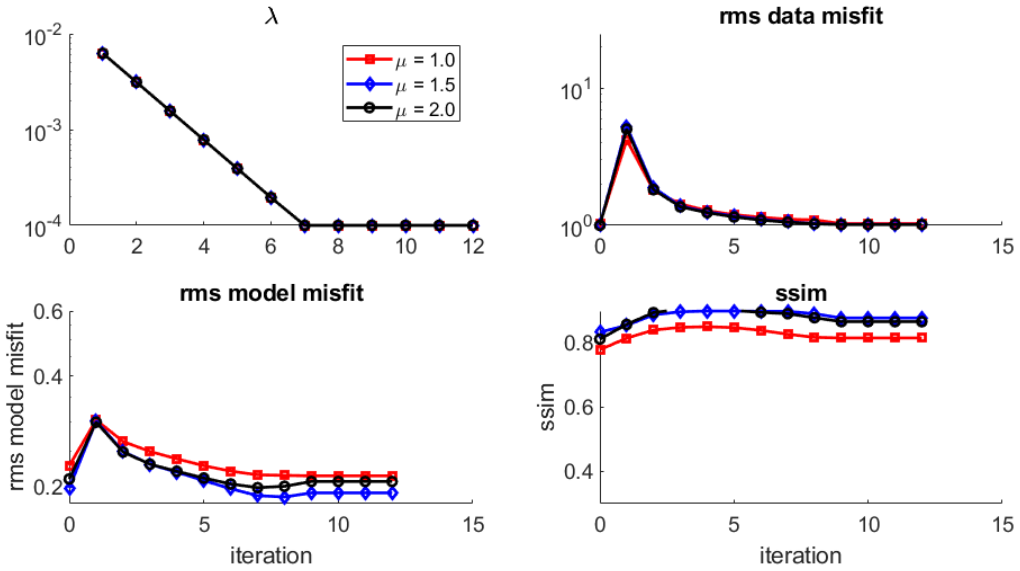


Figure 4.23: TGV 'double inversions' of the synthetic data from Figure 4.6, with 1% Gaussian noise, for different values of μ. The 'halving cool' scheme for λ was used. A pseudosection-initialised line search inversion was used for the initial model. (a) is the iteration with the smallest rms model misfit; (b) is the line search damping parameter, λ; (c) is the root-mean-square data misfit; (d) is the root-mean-square misfit to the original model; and (e) is the structural similarity index.

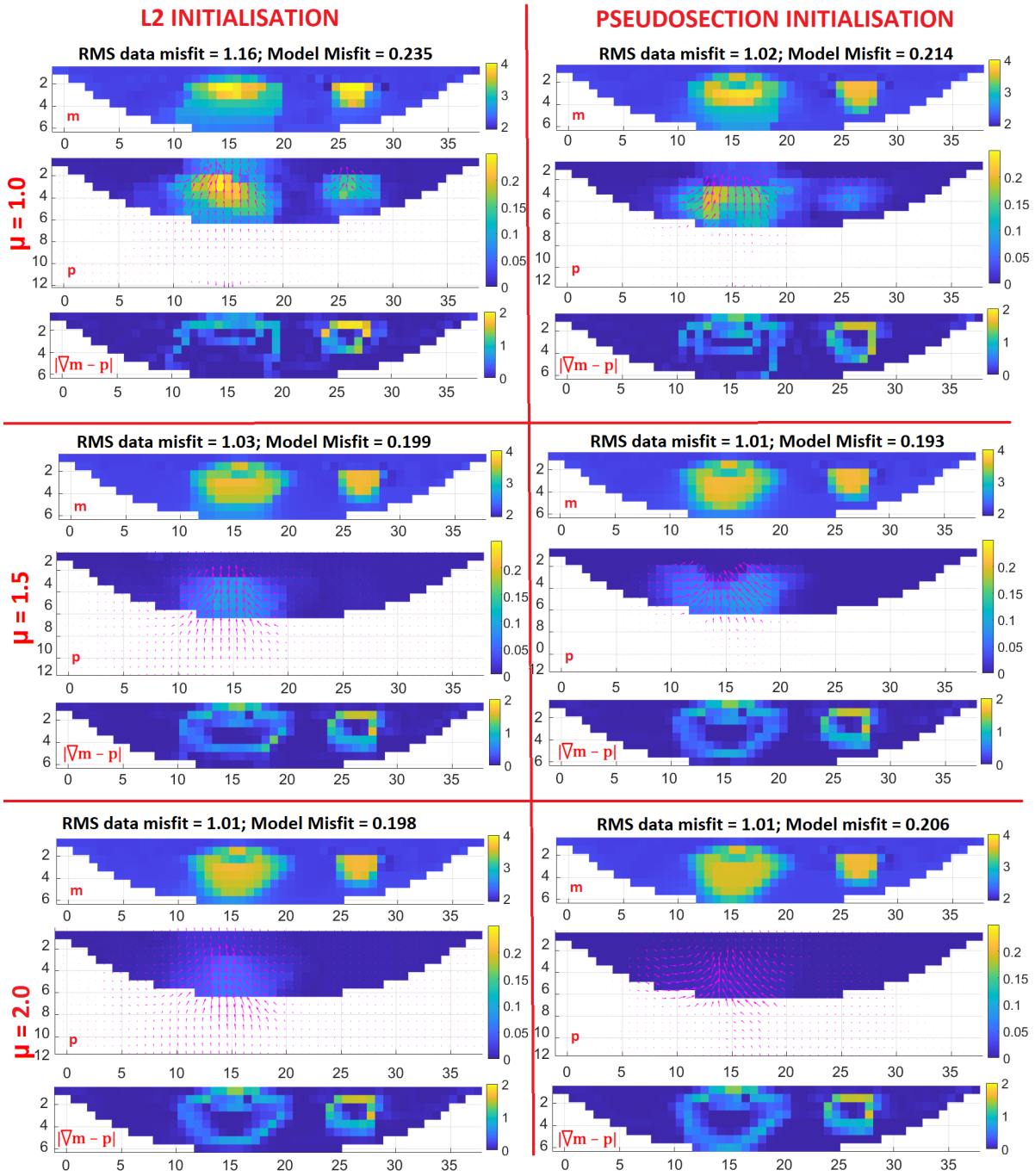


Figure 4.24: TGV 'double inversion' results, initialised using the solutions to the previous l_2- and pseudosection-initialised line search inversions from Figure 4.6. The 'halving' cooling schedule was used for λ. The synthetic data was generated from the model in Figure 4.6, with 1% Gaussian random noise added. These are the best iterations in terms of data misfit. $\mathbf{p}$ (pink vectors) is accompanied by a colour map of the magnitude, $|\mathbf{p}|$. For the pseudosection and $\mu = 2.0$, $\mathbf{p}$ has a maximum of 10^{-2}.

with **p** being more vertically orientated in the l_2-initialised double inversions.

The doubly inverted l_2-initialised models significantly outperformed the l_2-initialised inversions, and the high resistivity peaks have disappeared. Only the $TGV_{1.0}$ solution did not converge with a data misfit close to 1. This may be at least partly due to the original line search solution for this inversion being particularly far from the final solution. There is still a strong vertical component in **p** for the $TGV_{1.5}$ solution, but **p** is more spatially compact, and the horizontal components are larger. This results in a more compact solution which better approximates the shape and gradients of the original model.

By contrast, the changes in the pseudosection-initialised double inversions are smaller and more incremental. Most notably, the asymmetries in the smooth feature are diminished, and the resulting model misfits are incremental improvements on the line search solutions. As the original model is symmetrical, one would expect the solutions to be symmetrical, and so this symmetry is arguably as an improvement in the solution. An inversion which better captures the model gradient is likely to result in more consistent performance across many data. However, this does not necessarily result in the best model misfits, as both the 'halving cool' $TGV_{2.0}$ and the 'slow cool' $TGV_{1.5}$ pseudosection-initialised solutions had lower model misfits than the double inversion, due to asymmetrical smoothing effects.

Indeed, the magnitude of the $TGV_{1.5}$ asymmetrical **p** in Figure 4.19 is much closer to the local model gradient in magnitude, than **p** is in the double inversion solution.

The $TGV_{1.5}$ solutions both capture some of the smoothness of the gradient around the smooth feature. However, **p** skews more vertically in the l_2 model, while in the pseudosection model it is more closely aligned with the gradient of the original model. As one might expect, the solution which more closely approximates the original gradient has a lower model misfit.

In further tests, not shown, the results of the double inversion were used to initialise another, third inversion. The same cooling scheme was used. This was in order to investigate whether this might result in a progressive refinement of the TGV solutions towards a common solution. However, this did not result in any further changes to the solution for either the l_2 or pseudosection initialisations.

The double inversion approach has been partially successful in making the results of the TGV inversion more consistent, and closer to the desired piecewise smooth behaviour. The solutions converge more reliably, are more symmetrical than the single inversions, and **p** follows the smooth gradients of the model more closely. Though the differences are smaller, the solutions still differ meaningfully between different initialisations, which makes interpreting the inversion results more difficult.

4.3.5 TGV parameter testing: TGV Trade-Off Parameter

The effects of varying μ, which controls the model smoothness, were properly assessed. The double inversion was chosen for these tests, due to its increased consistency and improved piecewise smooth behaviour compared to the other cooling schedules. Figures 4.25 and 4.26 show how the objective function and model misfits of the solution to the inversion vary with μ.

The pseudosection initialisation displays some clear trends. The objective function tends towards TV as the $\mathbf{p}$ terms drop, becoming near identical to TV for $\mu > 2.0$. The data misfit is consistently in the range $1.01\% - 1.03\%$ for $\mu \geq 1$, approaching the 1% noise added to the model. The model misfit also varies smoothly, with the minimum between $1.3 < \mu < 1.5$. There is a small bias towards sharp edges, which is presumably needed to prevent oversmoothing where the model resolution is low, but is significantly smaller than the typical $\mu = 2.0$ used in image processing applications.

By contrast, the l_2-initialised double inversion experiences a more gradual decrease in the $\mathbf{p}$-influenced objective terms with μ, reflecting the larger influence of $\mathbf{p}$ in these solutions. The data and model misfits are significantly larger for $\mu < 1.5$, where it appears the additional smoothing in the l_2-initialised solutions is not beneficial. The best solution occurs at $\mu = 2.25$, closer to the values used in image processing. It is likely that larger μ is required to limit the over-smoothing effects outside of the smooth features in the model.

It should be noted that $|\nabla \mathbf{m} - \mathbf{p}|_1$ decreases as μ increases. This might be contrary to expectations, as $|\mathbf{p}|_1$ is expected to approximate the model gradient, and $|\mathbf{p}|_1$ decreases as μ increases. However, in these solutions $\mathbf{p}$ is observed to extend into the constant resistivity regions to a greater or lesser extent, and the positive contribution of these areas of the objective function term outweighs the areas where the gradient equals $\mathbf{p}$.

The snapshots at $\mu = 1.0, 1.5$ and 2.0 in Figure 4.24 are representative of the μ dependence a whole. There is a progressive decrease in smoothness as μ increases, and the magnitude and spread of $\mathbf{p}$ around the smooth feature gradually drops.

Overall, the control of μ over the smoothness of the solutions behaves as expected for a given initialisation, which increases confidence that the TGV inversion is performing correctly. There is a significant difference in optimal μ between the different initialisations. The larger μ preferred for the l_2-initialised model appears to be calibrated to counteract the excessive smoothing arising from the initialisation. The weaker constraint, and weaker initial smoothing, of the pseudosection-initialised model seems preferable.

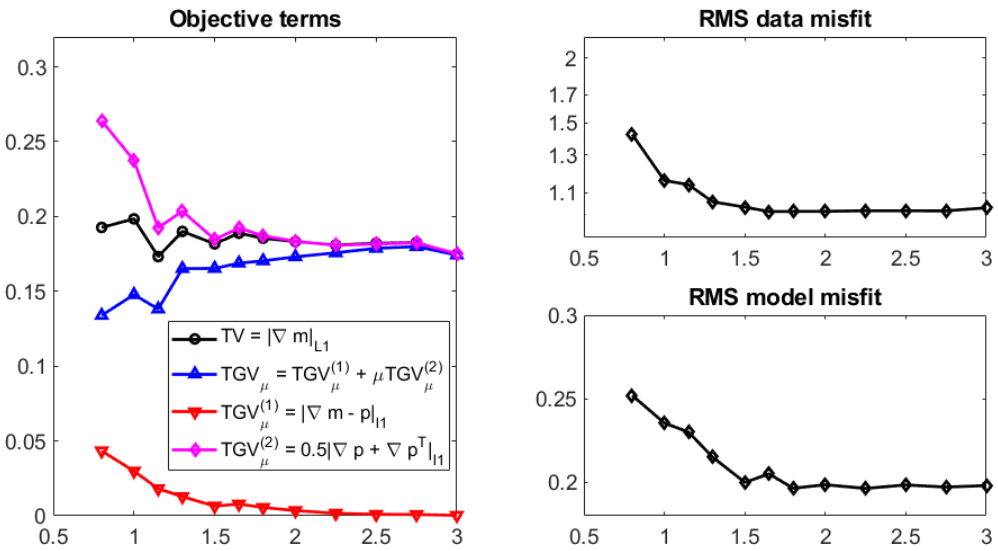


Figure 4.25: Effect of μ on the TGV solution, initialised by the solution to the l_2-initialised inversion. The terms of the objective function for the model solution per model cell, the root mean square data misfit, and the model misfit are shown.

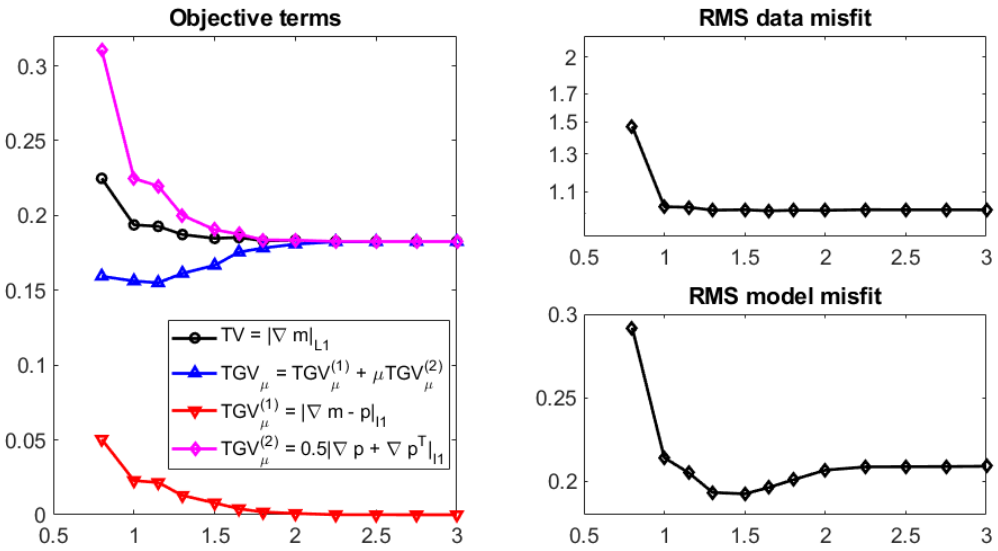


Figure 4.26: Effect of μ on the TGV solution, initialised by the solution to the pseudosection-initialised inversion. The terms of the objective function for the model solution per model cell, the root mean square data misfit, and the model misfit are shown.

4.4 Discussion

Overall, these parameter tests have shown that the TGV inversion is capable of improving upon the TV solution when inverting a piecewise smooth model, by better capturing smooth gradients. However, they have also demonstrated that the TGV solution is much more dependent on the initial conditions and descent path than the TV inversion.

These tests have also led to the identification of the pseudosection-initialised double inversion, as being the most likely regime to produce consistent piecewise smooth results without being overly influenced by the initialisation. This is where the result of a pseudosection-initialised TGV inversion is used as the initial model in a second TGV inversion.

Following the changes made to the inversion algorithm configuration made in Chapter 3, the TGV inversion is behaving much more as expected. The inner loop calculation of $\mathbf{p}$ now converges smoothly and reliably, and produces a close approximation to the model gradient. It is generally stable, although instabilities may exist at specific values of μ, which may be identified by high magnitude oscillations in the convergence rate. The inversion decrease in smoothness as μ increases, and tend towards the TV solution at large μ, at a rate which depends on the initialisation. The smoothness is controlled by $\mathbf{p}$, which tends to align itself with the model gradients. $\mathbf{p}$ tends towards linear gradients. The tendency for $\mathbf{p}$ to vanish near the upper boundary of the model remains in the inverted models, although not in the $\mathbf{p}$ solution in isolation. The use of periodic boundary conditions has reduced the incidence of artefacts at the domain edges and corners.

These parameter tests have established that the initial model and the λ descent scheme both have a significant influence on the TGV inversion, to a greater extent than for the TV inversion. This is perhaps not surprising, as the IRLS Gauss-Newton updates find local solutions based on the current model. In the TGV solutions, these depend on $\mathbf{p}$ for the previous iteration, which tends to preserve the form of gradient structures in the solution. This is compounded by the low resolution of ERT imaging, as there are many gradients structures which are valid, or approximately valid, solutions to the inverse problem.

Two different initialisations were tried. The l_2 initialisation generally performed poorly, with poor model and data misfits. The gradients were not well localised, driven by a $\mathbf{p}$ which extended into the background of the model, while overestimating the vertical component. The l_2 initial model is itself very smooth, with gradients which overestimate the extent of the model features, and these features are a legacy of the smooth initial model. The pseudosection-initialisation is less smooth, and the gradients in the resulting solutions were far more localised. These solutions were flatter, and largely resembled the TV model, with some additional smoothing around the edge of the smooth feature, which generally led to improvements

in model misfit for some μ. These solutions also retained features from the initial model, in the form of an asymmetrical gradient on one side of the smooth feature. This asymmetry was accentuated when the step size of λ was decreased, which supports the hypothesis that the algorithm tends to get trapped in the vicinity of the previous solutions. Increasing the step size λ too much violates the Gauss-Newton assumption of local linearity in the solutions, and so is not a straightforward remedy.

The alternative descent schemes had significantly less impact on the final model than the initialisation. For both TV and TGV inversion, the line search inversion lead to convergence in the fewest iterations. For the TV inversion, the line search resulted in an increase in retained structure from the initialisation. There was also evidence of residual structures in the TGV line search solution. The use of the 'halving' and 'slow' cool schemes for λ only increased the influence of the initial model on the solution.

Further refinements to the consistency of the solution were achieved through a 'double inversion': initialising the TGV inversion using the results from a previous inversion. This reduced the artefacts in the solutions due to the original model. The pseudosection-initialised solutions became more symmetrical, and $\mathbf{p}$ became more compact in the l_2-initialised solutions, and the gradient no longer overestimated the peak resistivity of the smooth feature.

For this reason, I would recommend the pseudosection double inversion as the preferred configuration for TGV inversion. Pseudosections are not generally as smooth as l_2 models, and so it is likely that this will avoid over-smoothing in the solutions. The double inversion acts to refine the solution, and reduce any remaining artefacts or asymmetries remaining after the initial model. For the test model, $\mu \approx 1.4$ produced the best solution. This weights edges approximately 40% higher than smooth regions, and was sufficient to capture the gradients of the smooth feature while also recognising the sharp edges of the block.

The model smoothness has been shown to consistently decrease with the TGV trade off parameter, μ. However, the rate at which it does so depends upon the smoothness of the initial model. For the double inversion of the test model, the optimum solution is in the range $1.3 < \mu < 1.5$ for the pseudosection-initialisation, while for the l_2-initialisation it was $1.8 < \mu < 2.5$. The relatively high μ of the l_2 double inversion, compensates for the overly smooth initial model through an increased anti-smoothness penalty in the regularisation functional. Where the initial model is blockier, the optimal anti-smoothing penalty is much weaker. Conceptually, this is favourable, as the inversion should be free to behave more smoothly or more sharply as the data dictates. It also results in better solutions, and $\mathbf{p}$ is a closer representation of the true gradient than in the vertically-skewed l_2 results.

Therefore the recommended modality for TGV inversion is for a pseudosection-initialised dou-

ble inversion for $\mu \approx 1.4$.

In principle, it is straightforward to extend the approach to the time-lapse ERT inversion of monitoring data. However, the suppression of $\mathbf{p}$ near the model boundaries is still occurring, despite the change to periodic boundary conditions in this section. Potential modifications to the boundary conditions, and the finite difference method, which might help alleviate this, are discussed in Chapter 6. It is possible that some of these changes may also improve the convergence of the TGV inversion, due to increased accuracy and reduced influence of the boundaries.

4.5 Summary

In this chapter, comprehensive parameter testing of the new alternating-minimisation algorithm for TGV inversion was performed. The results were mixed. The TGV inversion outperformed both l_2 and TV solutions, and captured the smooth gradients in the test model away from the surface, under optimum settings. However, the smoothing behaviour continues to be suppressed near the model boundary, and the TGV inversion was found to be highly dependent on the initial model. The 'double inversion' approach, which initialises a second TGV inversion with the result of the first, reduces the initialisation dependence and produces more consistent smoothing effects. On the other hand, the TGV trade off parameter, μ, controlled the smoothness of the solution in a grated manner, and the solutions were robust to small changes in μ. The reconfigured $\mathbf{p}$ subproblem is generally stable and converges reliably, independently of the initial model. Overall, the pseudosection-initialised double inversion, with $\mu \approx= 1.4$ is the recommended inversion configuration from these tests. It is possible that the performance and reliability of the TGV algorithm could be improved further with changes to the finite difference scheme; these are discussed in Section 6.2.

Chapter 5

Data Quality in ERT monitoring

5.1 Chapter overview

One of the challenges which arises from the increasing use of ERT remote monitoring systems is data quality management. Monitoring systems remain in place for months or even years, and some degree of environmental damage or degradation are almost inevitable.

Commonly used data quality measurements, contact resistance and reciprocal errors, have limits to the types of errors they can detect. The contact resistance is a measure of the electrical contact between the electrode and the ground, and the reciprocal error is derived from the difference between a measurement and its reciprocal, where the current and potential electrodes are exchanged. Reciprocal errors are able to detect some systematic errors where reciprocity is violated, such as polarisation errors (see 5.2 for further details). But in general, systematic errors are particularly difficult to identify if they do not violate reciprocity. In a monitoring setting, additional information is available in the form of the time series data for both the resistivity data and the error measurements: to some degree we can expect future data to resemble the observations made so far.

The aim is to capture errors which are not well detected by conventional means. As these errors can take many forms, from physical damage to components, to short circuiting of current pathways or the introduction of external current sources, then it is difficult to predict the form of any potential error. As such, the problem can be framed as one of anomaly detection [Chandola et al., 2009]. This is the process of detecting data which do not fit an established statistical pattern.

The motivating case study for this chapter is an example of short circuiting errors which occurred during a 14 month long monitoring experiment at the Sellafield site in the UK. Damage to an electrical detector led to short circuiting errors which were undetected at the time. The

aim of the chapter is to develop a new data quality assessment which is able to detect the onset of the short circuiting errors. The method will be tested on both the reciprocal error data, and the apparent resistivity data directly.

ERT data are highly correlated and contain a large amount of redundant information due to the spatial overlap between the regions of sensitivity of different quadropoles. Measurements which are sensitive to the same geological units, even if spatially separated, will also be strongly related. It is also reasonable to expect significant correlations between the reciprocal or repeatability error estimates. Measurements with common electrodes or cables will both be exposed to errors related to those components. Electrodes embedded in the same geological units may also experience similar trends in error levels due to experiencing the same wetting and drying cycles. Where external noise sources are involved, measurements with a physical or temporal proximity to the noise source, and to one another, will also tend to be correlated.

Principal Component Analysis (PCA), is a versatile data analysis technique which utilises the correlation or covariances of a data set to generate a reduced dimensional model of the data. In this chapter, I develop a PCA control chart anomaly detection method, and evaluate its performance at detecting the short circuiting errors from the Sellafield data set.

In section 5.2, existing data quality methods are described and reviewed. Section 5.3 contains an overview of PCA theory and its use in control charts. A methodology behind applying this method to ERT monitoring data is described in section 5.4. The motivating case study is described in detail in section 5.5. The control chart method for detecting short circuiting errors is applied to reciprocal error data evaluated in section 5.6.1 and resistivity data in section 5.6.2. These results are discussed in 5.7.

5.2 Data quality in ERT

One of the challenges of using ERT monitoring systems is ensuring the quality of the data. It is typical for environmental or infrastructure monitoring sites to be located far from the operating base of the monitoring team. Access can be limited to remote or sensitive sites, and manpower and travel costs may prohibit regular site visits.

There are several potential sources of noise or error that may develop over time. Animal or human activity can damage cables or power sources. Wetting and drying cycles can affect the electrode contact resistance with the ground, as can corrosion and other effects at the electrode-ground boundary.

Excessive moisture can also lead to short circuiting where components are not sufficiently wa-

terproofed. External current sources or metal structures (e.g. new fencing) could be introduced nearby. Ground motion cause changes in electrode geometry, or even cable damage. Malfunctions of the data acquisition system switch box or power supply may also occur.

In this section, we limit our analysis to those sources of error which affect either or both of the measured resistivity and reciprocal error data. Other sources can be identified separately. For example, power supply problems can usually be identified through a voltage drop.

There are two main types of direct data quality measurement used in ERT systems: contact resistance and three types of reliability measures. Contact resistance measurements are carried out between adjacent pairs of electrodes independently of the resistivity data acquisition, and allow individual electrodes with poor galvanic contact to be identified. These electrodes can only transmit small currents into the ground, which results in a low signal to noise ratio for measurements they are involved in. Contact resistance measurements are a powerful, but specific, tool, and are less useful for detecting other types of noise.

The commonly used forms of reliability measure are each associated with an individual multi-electrode transfer resistance measurement. Stacking errors are a measure of the variability of the transfer resistance over a number of current cycles during current injection. These do not take any additional time to acquire, but have a tendency to underestimate the error such that they are not useful for weighting the inversion [Tso et al., 2017].

Repeat errors are obtained by repeating each measurement two or more times, often through the repetition of the full measurement sequence. The data are taken as the mean of the repeat measurements, and the repeat error is the standard error of the mean or a similar measure of variability.

Reciprocal errors rely on the principal of reciprocity: the same result should be obtained when the current and potential electrodes are swapped, provided no polarisation effects are present. Reciprocal errors function similarly to repeat errors, with additional sensitivity to chargeability effects due to non-ohmic behaviour [Parasnis, 1988] and interference between cables and external current sources.

When an electrode is used to transmit current a build up of charge can occur, and the electrode becomes temporarily polarised [Dahlin, 2000]. A carefully ordered measurement schedule avoids polarisation errors, by allowing time for any such charge to dissipate before the same electrode is used to measure potential [Dahlin, 2000] [Wilkinson et al., 2012].

All three of these reliability measurements are limited in that they are measures of precision rather than accuracy. Of the three, there are strong reasons to prefer reciprocal error measurements, due to their additional sensitivity to polarisation errors and other non-ohmic effects

such as external currents. They are also more independent than repeat and stacking errors, as the current tends to follow different signal pathways. Due to this sensitivity to several physical effects, the reciprocal error is closer to a measure of accuracy than repeat or stacking errors. However, these are not a complete set, and the reciprocal error is not sensitive to other sources of error, including electrode positioning errors or short circuiting for example.

Typically, only a single repeat or reciprocal measurement is available to calculate these error estimates, and as such the calculated error values themselves carry a large degree of uncertainty. This is unsurprising, given that each repeat essentially doubles the acquisition time, and there is a trade off between volume of data collected and the number of repeats that are made.

A statistical error model is often used to generate the data weights for the inverse problem, so that random noise in the individual error measurements does not bias the inversion. In Binley et al. [1995], the authors fit a linear error model, under the assumption that the errors are normally distributed and proportional to the transfer resistance. A related approach, described in Koestel et al. [2008], makes the same assumption that measurements with similar resistance will have similar errors. The normal and reciprocal measurements are first binned according to transfer resistance, and the standard deviation of each bin is used to fit a linear error model. More recently, Tso et al. [2017] extended this to a linear mixed effects model, which also accounts for noise variation between electrodes, by grouping the data by the electrodes used in each measurement. This approach can allow stacking error derived models to perform comparably to repeat or reciprocal errors at assigning error weights, at least in some cases.

Deceuster et al. [2013] developed a Bayesian methodology for identifying individual faulty electrodes in a time lapse context from the the stacking error quality factor. Specifically, this methodology was designed to detect changes in electrical contact properties of electrodes due to changing conditions over time. A prior distribution of the stacking errors is determined from a baseline data set, with a reciprocal error threshold used to ensure the quality of this initial data set. During monitoring, likelihood ratio contrasts are used to identify electrodes as faulty when the measurements involving those electrodes deviate from the prior distribution by more than would be expected due to chance, as estimated by Monte Carlo simulations. By establishing a per-system baseline using reciprocal errors, better results are able to be obtained using stacking errors alone. However, unless the measurement time constraints are especially sensitive (perhaps due to a rapidly changing ground resistivity), it is usually feasible to measure reciprocal errors directly during monitoring as the system is already installed.

In Mitchell and Oldenburg [2016], the authors develop a detailed methodology for data qual-

ity control for highly noise-contaminated data in the absence of reciprocal error data. Instead, this process utilises the data misfit from inverted models as a proxy for data quality. An initial manual investigation of the data was carried out, before an automated clustering method was used to identify and remove bad data. As part of this approach, Singular Value Decompositon (SVD) was performed on binary indicator matrices for both low and high misfit data. The singular vectors showed that the high misfit data was generally associated with one of the two cables used for their dataset. This was followed up by boxplot analysis of various data parameters which suggested that current leakages were occurring in some measurements. Next *k*-means data clustering approach was used to identify data subgroups based on the input and output voltages, current and data misfit of the data. Each subgroup was inverted independently, and based on the data misfit, the subgroup was either rejected, retained or subjected to further rounds of clustering analysis. The clusters were then recombined and inverted to ensure that they were consistent with the wider data set.

5.3 Principal Component Analysis

Principal Component Analysis (PCA) is a technique used for dimensionality reduction and exploratory data analysis, among other applications [Jolliffe, 2002]. PCA applies a linear transformation to the data into a new set of orthogonal variables, called the principal components (PCs), such that the variance of the first few principal components captures most of the variance present in the data. This allows a large data set containing many correlated variables to be represented by a few principal components, while retaining most of the important information. The coefficients of this transformation can also be used to offer insights into the significance and relationships between the variables.

A data set of m observations of n variables can be represented as the $n \times m$ data matrix, $\mathbf{X}$. Columns of $\mathbf{X}$, $\mathbf{x}_i$ correspond to a single observation across all variables, while each row, $\mathbf{x}_j$, contains all the observations of a single variable. To simplify the analysis, each variable in $\mathbf{X}$ is first centred by subtracting its mean.

Provided $m > n$, this can be represented exactly through a PCA transformation as

$$\mathbf{A} = \mathbf{XP}, \tag{5.3.1}$$

where $\mathbf{A}$ are the $n \times m$ PC scores, which are the magnitudes of the data projected onto the PC axes. $\mathbf{P}$ is the the $m \times m$ matrix of PC coefficients, also known as loadings, which determine the transformation. The columns of $\mathbf{P}$ are the eigenvectors of the covariance matrix, $\mathbf{\Sigma}_{cov}$, of $\mathbf{X}$, which define a new orthogonal basis for the data through a linear combination of the original

variables. **A** are proportional to the eigenvalues. The principal component scores are naturally ordered by the proportion of variance of **X** that they contain. The first k scores will contain most of the variation in the dataset. This allows a good approximation to the data matrix to be made using only the first k components:

$$\mathbf{X} = \mathbf{A}\mathbf{P}^{\mathbf{T}} + \mathbf{E}, \tag{5.3.2}$$

where **A** is the $n \times k$ score matrix, **P** is the $m \times k$ coefficient matrix and **E** is the $n \times m$ residual matrix.

5.3.1 Data prepossessing

Each PC is defined so that it maximises the proportion of the variance of **X** that it explains. As the variance of a variable depends on the units it is measured in, the PCA solution is dependent on the unit scaling of each variable. When the variables are all measurements of the same quantity, then covariance PCA, as described in section 5.3, may be appropriate. In covariance PCA, the magnitudes of the variances within each variable are captured.

Alternatively, the data matrix **X** can be standardised by dividing each mean-centred variable by its standard deviation. The PC coefficients of the standardised data are the eigenvalues of the correlation matrix $\mathbf{\Sigma}_{cor}$. This is known as correlation matrix PCA, or simply correlation PCA.

5.3.2 Control chart statistics

There is great potential to use control charts to identify anomalous behaviour in monitoring systems, in order to identify the onset of measurement errors which are not captured by traditional measures. An initial measurement period is used to generate a PCA model which captures the typical behaviour of the system. If at any point the system under observation deviates substantially from the PCA model, a change of state can be identified. This may either be 1) a significant change in subsurface resistivity or 2) the introduction of a new source of error.

Hotelling's T^2 statistic is the multivariate analogue to student's t-test, which is used to compare the differences between different groups. For mean-centred **X**, Hotelling's T^2 is

$$T^2 = \mathbf{X}\mathbf{\Sigma}_{cor}\mathbf{X} = \mathbf{A}^T\mathbf{A} \tag{5.3.3}$$

If the underlying samples, **X**, are *i.i.d.* (independent and identically distributed), then control limits for T^2 can be derived based on the F-statistic [Jackson, 1959],

$$T^2_{\alpha(n,m)} = \frac{n(m-1)}{m-n} F_{\alpha(n,m-n)}, \tag{5.3.4}$$

where $F_{\alpha(n,m-n)}$ is the F distribution with n and $m-n$ degrees of freedom. If T^2 falls outside these limits for a given observation, it is identified as anomalous with respect to the training distribution. Unfortunately, the ERT time series' under investigation are not i.i.d and so these control statistics cannot be directly applied. However, T^2 may still act as an indicator for any egregious data quality issues, or sudden interesting changes in the underlying resistivity.

The Q statistic measures the square sum of the data residual which is not captured by the k principal components used in the model [Jackson and Mudholkar, 1979]. For a given observation, $\mathbf{x}$,

$$Q = e^T e = x^T(\mathbf{I} - \mathbf{P}\mathbf{P}^T)x \tag{5.3.5}$$

where e is the row of the residual matrix, $\mathbf{E}$, corresponding to observation $\mathbf{x}$.

Q captures the proportion of the data which lie outside of the subspace defined by the PCA model, and hence it can be used as an indicator of a change of state in the system. Assuming *i.i.d.* variables, a control limit for Q can be derived (see Jackson and Mudholkar [1979] for details).

5.4 A PCA approach for Data Quality

The ERT resistivity measurements are highly correlated, containing a great deal of redundant information. This is evident simply from the numbers of measurements required; the number of measurements used to generate an ERT model can be an order of magnitude larger than the number of cells in the model it constrains. There is a high degree of spatial overlap in sensitivity between adjacent measurements, as well as between those with the same geometric factors. This is related to the mixed-determined nature of the ERT inverse problem, as parts of the model tend to be over-determined, while others are under-determined. Similarly, correlations also arise in the reciprocal error data, as measurements with common electrodes or in regions undergoing simultaneous wetting and drying patterns are exposed to the same sources of error.

A PCA model generated using a baseline period of data will reflect the correlations present within that data. Control charts based on this PCA model will be able to detect potential data quality issues due to changes in the correlation structure of the data, such as occurs during

short circuiting errors, providing that the model properly captures the behaviour of the system.

5.4.1 Method

In this analysis the viability of a control chart approach for ERT monitoring data quality control is assessed. The goal is to use an anomaly detection approach to identify and analyse the deterioration of data quality in the Sellafield case study (see section 5.5). This consists of short circuiting errors caused by damage to a borehole connector. The analysis was split into two halves:

1. Using reciprocal error data, use PCA control charts to identify the data quality deterioration due to short circuiting.
2. Using apparent resistivity data, use PCA control charts to identify the data quality deterioration due to short circuiting.

As short circuiting errors only affect the reciprocal errors indirectly, the post-damage error increases are of relatively low magnitude. Objective (1) aims to identify these changes as anomalous despite their relative low magnitude. Objective (2) detects data quality problems without direct error estimates. This is useful when limited error data is available, or with short circuiting type errors which are poorly captured by reciprocal error.

The ERT monitoring problem differs from a standard statistical process control system in a number of ways. Most notably, it consists of time series data where the process under observation is continuously changing. The rate of data acquisition is slow, and the total number of raw observations, n, is much smaller than the number of variables, m. The shorting errors are not a point anomaly, but a sudden onset change in the system, which continues to develop over time. The aim of our analysis is to detect the onset of the connector damage, using a model derived from the preceding time series, before identifying the source of the errors if possible.

PCA is well suited to exploratory analysis of data, as the principal components reflect the correlation or covariance structures within the data, and the contributions of each measurement to the PC subspace can be detected in the residual. After an anomaly is detected, data residuals are examined for all measurements between each pair of boreholes, called inter-borehole panels in order to identify the source of the anomaly.

For these analyses, the first n_{train} timesteps of $\mathbf{X}$, were used to generate a PCA model. The full timeseries of data was then projected onto k components, selected such that 99% of the variance of the training data was retained in the model.

T^2 and Q control charts will be tested in order to identify anomalous events in the ERT time series. An increase in Q corresponds to an increase in variance outside of the PC subspace. A large increase in T^2 indicates that the magnitude of the signal inside the PC subspace is abnormally large with respect to the training data. This would imply that some previously observed behaviour of the system is now occurring to an unprecedented degree. A hypothetical example of this in ERT might be at a site which experiences regular variation in the surface resistivity varies over time, through wetting and drying cycles, which are captured by components of the PC model. An unprecedented drought might be captured by a large increase in those components, which would appear in the control chart as a large increase in T^2.

One consequence of the non-stationary nature of the ERT measurements system is that the PCA model will need to be incrementally updated during measurements. As the resistivity measurements are not *i.i.d* random variables, then analytically derived thresholds in T^2 and Q cannot be used to identify anomalies, so a more heuristic approach is required.

After the training period, some drift in Q is expected over time. For the model to be useful in detecting anomalous system behaviour, we assume that the rate of this background drift will be small and gradual, in order that the model remains approximately valid; and predictable, such that anomalous changes can easily be distinguished. The rate at which the confidence limit for Q is exceeded may be useful to benchmark the drift rate.

By periodically expanding the training data sample used to generate the PCA model, the drift can be reset. A true anomaly should appear in the control charts over a range of training lengths. If it does not, it is an indication that the training model is fragile and poorly generalisable.

In contrast to the analysis of Deceuster et al. [2013], this approach does not rely on the stacking errors for a positive signal. Unlike the method of [Mitchell and Oldenburg, 2016], PCA control charts are available in real time as the data is acquired, and do not require computationally intensive inversions to solve. One challenge for assessing data quality methodologies, especially those aimed at detecting anomalies, is that it is very difficult to simulate anomalous events accurately, especially alongside realistic noise profiles.

5.5 Case study: Damaged borehole connector

Our motivating case study is a 14 month long geoelectrical monitoring experiment which experienced significant degradation of data following damage to a borehole connector. This resulted in the ingress of moisture to the connector, causing short-circuiting between electrode connector pins. This went undetected for some time, as the data quality processes used at the time were relatively insensitive to shorting effects. The inclusion of this data resulted in signif-

icant artefacts in contemporaneous inverted models, suggesting erroneous hydrological connectivity between regions thought to be physically separated by a low permeability aquitard. The onset of errors was only later discovered by a manual inspection of the data, after further deterioration of the data prompted an investigation.

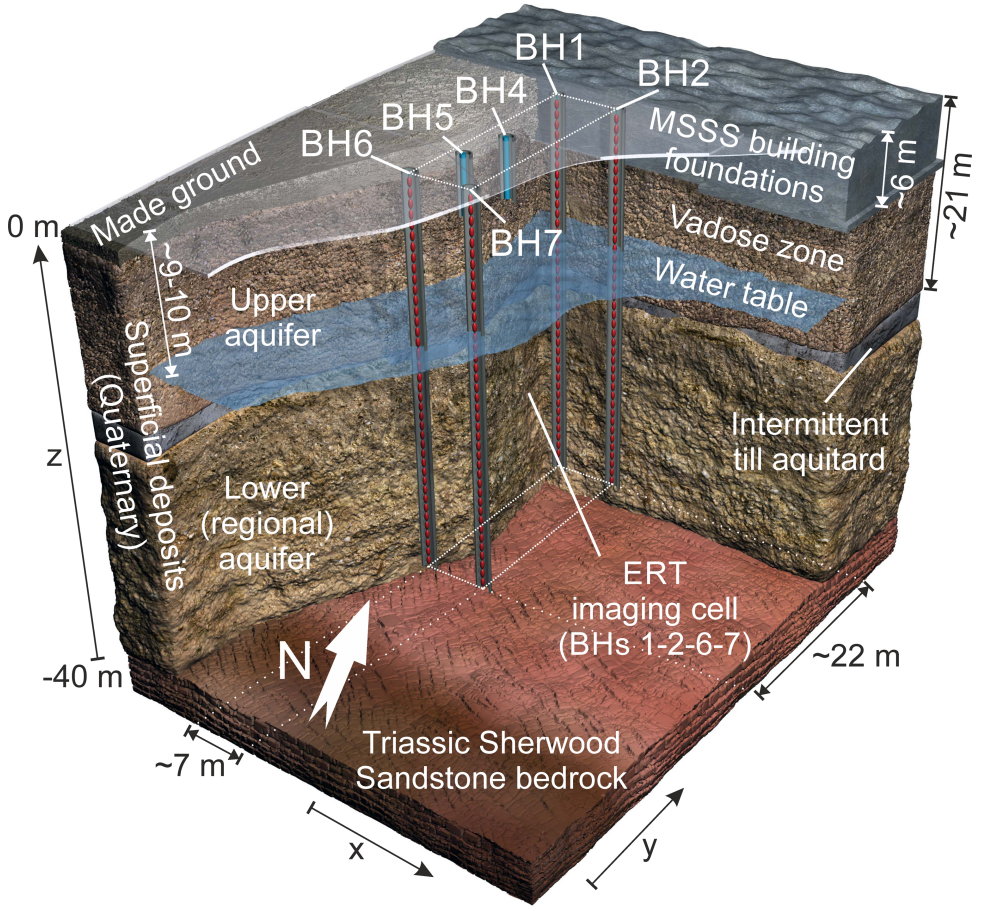


Figure 5.1: Geological model and borehole sensor array schematic for the Sellafield monitoring experiment. This data quality analysis uses the resistivity measurements from the primary ERT imaging cell, consisting of boreholes BH1, BH2, BH6, and BH7.

This case study was a proof of technology trial which involved detecting simulated contaminant leaks at the Sellafield site in Cumbria, UK [Kuras et al., 2016]. The imaging cell was defined by four vertical boreholes, each instrumented with 40 electrodes at 1m spacing. Each measurement consisted of 11000 forward and reciprocal pairs, following a crosshole bipole-bipole scheme, alongside contact resistance measurements. Polarisation errors were minimised by ensuring that there was sufficient delay between an electrode transmitting current and being used to measure potential. Full measurements were acquired on alternate days from January 2013 until April 2014, with occasional gaps, each one taking 22.5h to complete. For further experimental details, see Kuras et al. [2016]. The objective of the monitoring was to detect plumes of low-contrast simulated contaminants in the model imaging cell. As the

system undergoing observation is only experiencing subtle changes, the 'signal to noise' ratio of an error detection system will larger relative to more dynamic systems.

The contaminants were introduced in three phases, 06/02/2013 to 16/03/2013; 28/06/2013 to 02/09/2013; and 24/09/2013 to 27/09/2013. These were all detectable in inverted ERT difference models.

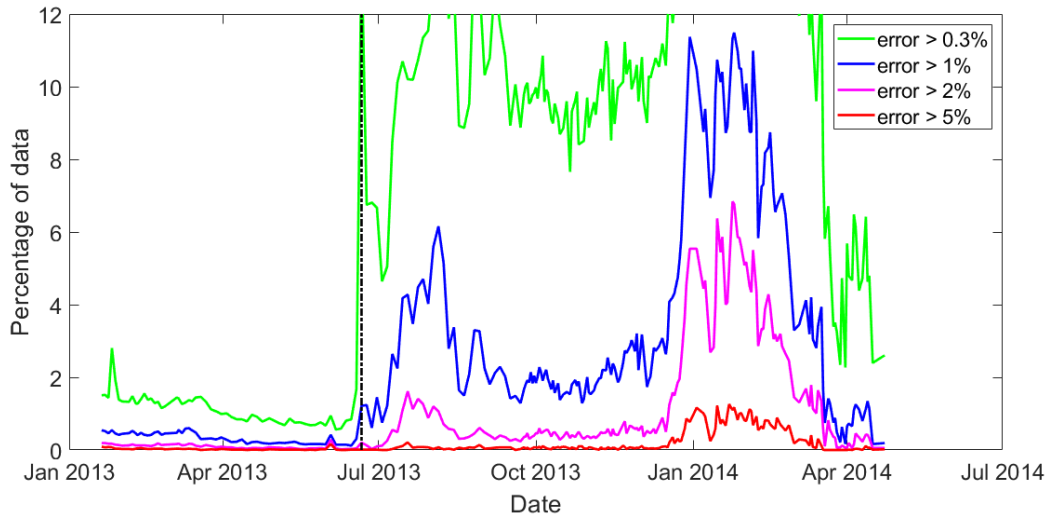


Figure 5.2: Percentage of data points exceeding a range of error thresholds over time for the Sellafield data set. The dashed line marks the time at which connector damage is believed to have occurred

Two main methods were originally used to monitor data quality: the contact resistances and reciprocal error estimates. The contact resistance measurements identified three electrodes with poor electrical contact ($> 2k\Omega$) which were excluded from the measurements. Data with a reciprocal error $> 5\%$ were excluded from inverse modelling. This threshold was determined by expert judgement after viewing the initial error distribution, based on experiences at similar sites. The number of data excluded by each metric provided a broad overview of the health of the system.

In the Sellafield monitoring set, the main source of data deterioration was short circuiting errors, which are challenging to detect using these conventional methods. After decommissioning, an examination of the borehole 1 connector discovered significant mechanical damage, as well as signs of moisture ingress and associated corrosion of the electrical connections. This moisture caused some conductive contact between adjacent connector pins corresponding to individual electrodes in the borehole. Therefore, whenever any of these electrodes were involved in data acquisition, these other electrodes were also passing current, which affecting the measured potential difference for each data point.

As the electrical contact between the electrodes and the earth remains unaffected, these short-

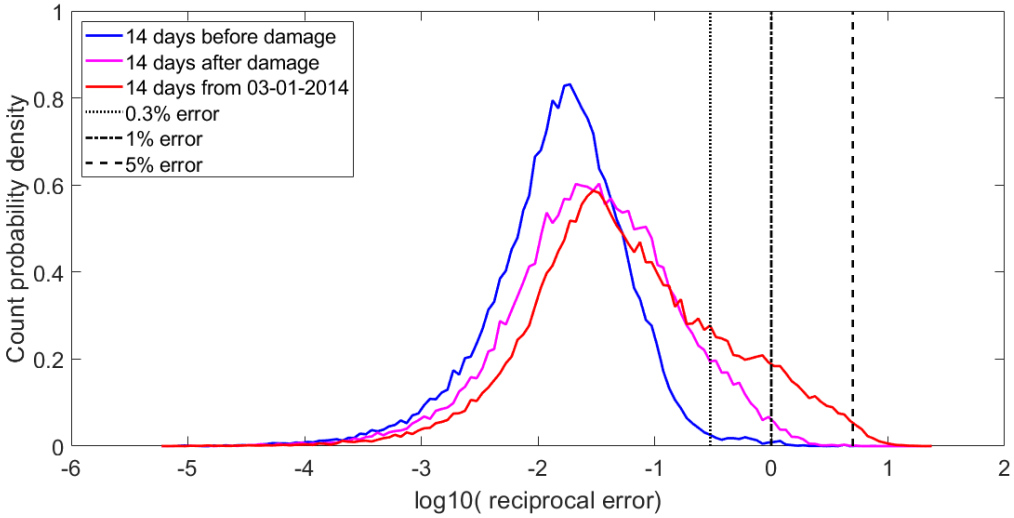


Figure 5.3: Reciprocal error distributions for fortnightly periods of the Sellafield monitoring data. After the connector damage there is a noted increase in error levels. However, most of this change is not well captured by the fixed error thresholds at 0.3%, 1% and 5%.

ing effects were not detectable through high contact resistances. Similarly, one would not necessarily expect short circuit errors to be detectable by reciprocal error measurements, as the principle of reciprocity still holds for a shorted circuit. However, in this case an increase in reciprocal errors was observed, as shown in figures 5.2 and 5.3. This increase in error is relatively subtle, especially at the onset of the connector damage.

While reciprocal error measurements are not directly sensitive to shorting errors, they are sensitive to polarisation errors: these occur when an electrode is used to measure potential whilst still being polarised after recently being used to transmit current. The measurement scheme was ordered so as to avoid any such polarisation effects [Wilkinson et al., 2012]. However, short circuiting effects led to electrodes other than those specified for each measurement transmitting current. It is likely that the observable increase in reciprocal errors was due to the secondary polarisation effects induced by these errant currents in recently used electrodes.

The onset of the connector damage and the resultant short circuiting errors are believed to have occurred between the 20th and 23rd of June 2013, when the increases in the reciprocal error distribution first occur in reciprocal errors $<=$ 1% (figures 5.2, 5.3). This date corresponds with the emergence of artefacts in the inverted models.

A rise in errors $>$ 2% followed in early July. As there was negligible change in the errors $>$ 5%, this change to the error distribution was not detected, and hence contaminated data from borehole 1 persisted into the initial inversions of the data, resulting in artefacts in the inverted models. Six months later, in January 2014, the error counts rose across all thresholds,

including above $> 5\%$. The short circuiting may have increased due to the wet winter weather. This triggered an investigation which uncovered the previous errors.

The impact of the connector damage on the reciprocal errors is more clearly visible in the error distribution, figure 5.3. The majority of the initial error increase occurred in the 0.1% to 1% region, and so even a 1% filter would not have detected it.

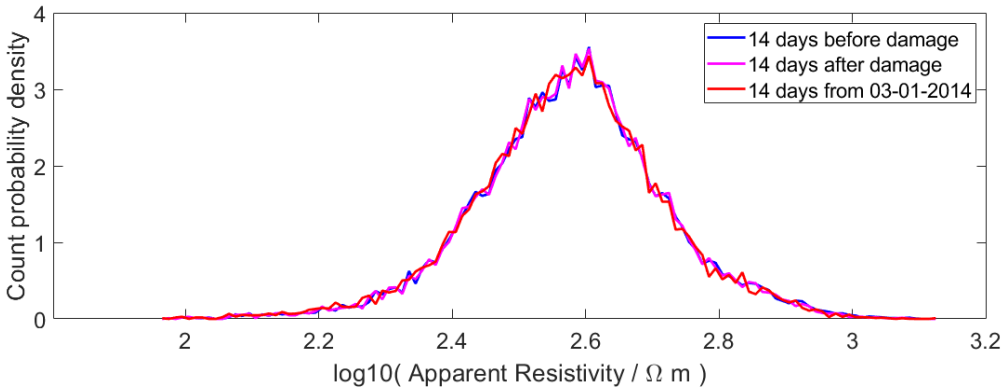


Figure 5.4: Apparent resistivity distributions for fortnightly periods of the Sellafield monitoring data. The onset of connector damage is not apparent.

The apparent resistivity distribution does not change significantly at the onset of error (figure 5.4).

This case study suggests that threshold based measures have significant weaknesses as a data quality metric, and disregard information from both the broader error distribution. The aim of the PCA control chart methodology is to utilises the time series information in both the error and the resistivity distributions to detect the data deterioration, through changes in the variances and covariances of the data.

5.5.1 Weaknesses of a reciprocal error threshold approach

In the Sellafield case study a reciprocal error threshold of 5% was used for data quality control. Figure 5.2 suggests that this threshold was too high to detect a significant change in the reciprocal error distribution immediately after the connector damage occurred, and that a lower threshold would be better suited to detecting the onset of data deterioration.

When choosing an error threshold, there is a trade off to be made between the sensitivity of the detector to true errors, while minimising the amount of useful data that is discarded. Assuming that the reciprocal error measurements are an accurate measure of the data uncertainty, a 5% threshold is a reasonable choice which aims to discard the smallest amount of data. An alternative approach to determine an error threshold is to base the choice on the initial error distribution. A 1% threshold excludes 0.2% of the pre-damage data, whilst a 0.3% threshold

excludes 0.9%. Both of these could be determined acceptable rates of loss. However, true data errors in the range of $0.3\% < e < 1\%$ are commonly seen in ERT, and the modelling errors in ERT inversion are typically around this magnitude, and often somewhat greater. Therefore, data with errors in this range can effectively be considered "accurate" for the purposes of ERT inversion, even more so when the data are properly weighted. Such a strict threshold risks false alarms and an excessive exclusion of usable data given a small change in monitoring conditions.

This is a more fundamental weakness in any reciprocal error based method. Implicitly, these methods assume that the reciprocal error measurements are a reasonably accurate estimate of the true data error. This is not the case for short circuiting errors, which are only observed through secondary polarisation effects.

The net change in error levels between sample periods can be estimated from the area between the curves in Figure 5.3. 21% of measurements show a net increase in error in the fortnight after the damage occurs. Any reasonable threshold method fails to detect the majority of the affected measurements. Indeed, a 1% threshold only detects 4% of the affected measurements. The strict 0.3% threshold captures 30% of the affected measurements, at the cost of over sensitivity to small increases in noise levels from other sources.

Indeed, the significant overlap in magnitude between the contaminated and uncontaminated data errors ensures that no threshold can be calibrated to capture the full error increases without also including a significant proportion of data where the errors remain 'normal'.

An alternative to a reciprocal error threshold approach would be to monitor the error distribution directly. The change to the Sellafield error distribution is captured in the gaussian moments of the distribution: the first three being the mean, variance, and skewness. This would increase detection of sudden, large scale, low-amplitude error increases which may be related to polarisation errors from short circuiting.

There are limitations in the ability of error distribution monitoring to detect short circuiting errors. There is no guarantee that secondary polarisation effects will occur, so a means of detecting short circuiting errors directly from the resistivity data is preferable. Without independent information, the secondary effects of short-circuiting errors cannot be distinguished from low-amplitude noise increases from other sources.

More generally, distribution monitoring is effective at identifying error sources which affect a significant proportion of the measurement array. Those which affect a small number of electrodes or measurements may be undetected.

By contrast, the proposed PCA control chart anomaly detection approach can detect changes

affecting individual electrodes (see section 5.6.3), as well as systemic errors and may be applied separately to resistivity and reciprocal error data.

5.5.2 An indirect approach to error analysis

The proposed method aims to provide a means of detecting unusual error types, such as the short circuiting errors seen in this case study, in ERT monitoring systems, based on changes to the correlation structure of the data at the onset of errors.

Such an approach has several advantages over direct error measurements. It can be applied to both the raw resistivity data and to any error measurements. Sources of error which do not violate the principles of reciprocity can be potentially identified from the resistivities alone. It may also be used where reciprocal error measurements are unavailable, such as when observing rapid resistivity changes. Unlike approaches based on error thresholds, it can detect low-magnitude changes to the error data similar to those seen in this case study, which may not be captured by a reasonable error threshold.

This approach requires an initial training period before it can be applied, during which time conventional data quality metrics must be used.

5.5.3 Data preprocessing

The data was processed before performing PCA. For this analysis, only data consisting of averaged reciprocal pairs were used, due to the availability of reciprocal error estimates. Two electrodes in the vertical cell were in poor electrical contact, and all measurement configurations involving these were removed from the dataset [Kuras et al., 2016].

A median of 13170 forward-reciprocal pairs per time step were contained in the data set before filtering. These measurements consisted of parallel current and potential bipoles between pairs of boreholes in the vertical cell. In 28 of 243 time steps instrument error led to missing measurements. Of these, 15 were missing <6 measurements. The remaining 13 were missing between 255 and 17893, with a mean of 7718 missing points.

The data were then filtered to remove any reciprocal pairs that contained failed or unreliable measurements, according to the measured voltage readings. Note that there is overlap between the subsets of data identified by each of the filters.

Filters were also applied to remove data with voltages $< 10^{-10}$mV, which are failed measurements, as well as a voltage waveform ratio outside the range $0.95 < w_{ratio} < 1.05$. The ALERT measurements system applied a symmetrical square DC waveform during measurements. The waveform ratio is a measure of the asymmetry between the positive to negative cycles. If the

waveform is not symmetrical, it is an indication that the measurement is occurring in a non-ohmic regime, and polarisation effects are occurring. On a challenging industrial site such as this, some of this is likely due to metallic objects buried in the area. After early March, two months into the sequence, the measurement order was adjusted in order to reduce these polarisation effects, resulting in fewer removed measurements and a reduction in reciprocal errors after this point (see figure 5.2).

0.65% of all measurements in the first two months were failed measurements (voltages $< 10^{-10}$mV), dropping to .00007% during the rest of the series. 4.3% of data across the series were removed due to having a voltage waveform ratio outside the range $0.95 < w_{ratio} < 1.05$. The data were also filtered by instrument error codes, which corresponded closely with the low voltage measurements.

After filtering, the data consisted of 12872 electrode quadropoles measured at 285 time steps. 4.9% of these were missing.

Two approaches for back-filling missing values were considered. A commonly used strategy is to replace each missing value with the mean of that variable across all observations. This works best when it can be assumed that each variable is stationary, which is not the case for time series data. Instead, a linear interpolation in time was used, which more closely reflects the current state of the system. As large numbers of missing values in close proximity cause errors in the interpolation, configurations and dates with large numbers of missing values were removed from the data set before interpolation was applied.

1860 electrode configurations with less than 200 time entries were removed from the analysis. Next, 58 dates with more than 3000 missing values were also removed. These were all after April 2014, long after the connector damage had occurred and so unimportant for the analysis. The processed data matrix contained 11012 configurations for 227 dates, of which 1.15% were back filled by interpolation. The median number of missing data per date was 58, or 0.5%. The six remaining dates with the most missing data (in the range $500 < N_{missing} < 3000$) all neighboured time steps with a more typical data coverage, and sequential interpolated values were rare.

A more rigorous approach to estimating missing values would be to use an expectation minimisation or fit based on the PCA model, but this is much more computationally intensive, especially considering the size of $\mathbf{X}$, which is not well suited for repeat calculations as the model is repeatedly updated.

5.6 Case study analysis

In my analysis, I describe the dataset as consisting of a 'baseline' period, consisting of 41 usable measurements (after filtering) between 14th January 2013 and 20th June 2013, during which the monitoring system was working as expected. The 'damaged period' is the from 21st June onwards after the borehole 1 connector damage had occurred. 172 measurements (post filtering) were recorded in this time. Both error and apparent resistivity data are available in the Sellafield dataset. First, I apply the PCA control chart methodology to the reciprocal error data in section 5.6.1, followed by the resistivity in section 5.6.2.

The objective of the PCA analysis is to establish a model which captures the behaviour of the training data, which can be used as a baseline to detect future changes. The control chart analysis relies on identifying significant changes to the covariance/correlation structure of the measured data, both within and outside the model subspace. As such, a model which captures the majority of the variance in the training set is desired. For that reason, I choose to retain k principal components in the training model, such that 99% of the model variance is retained in the k components.

As the data are time dependent, and not *i.i.d.* random variables, there is likely to be a natural drift in the measurements away from the model space as the system evolves over time. For an anomaly detector to function efficiently, this drift should be minimal, and significantly smaller than the change caused by an anomaly. As this is the case, models which can account for the majority of the variance with a small number of principal components are be preferred. The broad trends and correlations in the data are of interest. If $k \approx m$, the model subspace is almost the entire subspace occupied by the original data. As such it is likely that the model is overfitting the noise and small variations within the sample data training set, and hence will not generalise well to new data. If, on the other hand, k is small ($k < 10$) and the PCA model only spans a small subspace of the sample data, then it is likely that the model captures the broad trends in the data and will be applicable to future measurements. As the rank of $\mathbf{X}$ is $\leq m$, the number of principal components is limited to m.

One means of testing this is to update the model as new data is received, so that the new data remains close to the model. If a sharp change in the control chart appears across several training sets, it is more likely to be a true anomaly and not due to incremental drift. The greater the time lag between training set and a suspected anomaly, the less reliable the model and the greater the uncertainties involved.

In this analysis, I typically present a training length of $n = 35$ in my results. This is long enough to train a representative model that remains valid at the onset of the connector damage, while also allowing the background drift of the model to be seen and characterised. This also ensures

that the separate June 3rd anomaly lies outside the training set, avoiding the need for special treatment (see section 5.6.3). The anomaly detection capabilities of the different models are consistent when similar or longer training lengths are used (not shown).

The first step in a control chart analysis is to determine the model. This is complicated by the time-dependent nature of the monitoring data. The distribution and correlation structure of the data is not stationary, therefore I expect to have to update the model periodically as more data is acquired. If the normal rate of change in the data distribution is sufficiently slow, these models will be sensitive enough to detect gross anomalous changes. Further investigation can then be applied to determine whether this is a data quality issue or due to a change in ground conditions.

Several choices need to be made in the determination of a PCA model. A choice of data scaling needs to be made, and the number of principal components retained in the model needs to be chosen. As each of our analyses contain data of a single type, it is a valid option to use covariance matrix PCA. Correlation matrix PCA, however, is less influenced by outliers in the data set, and so may be more robust. It is responsive to trends in the data, regardless of magnitude.

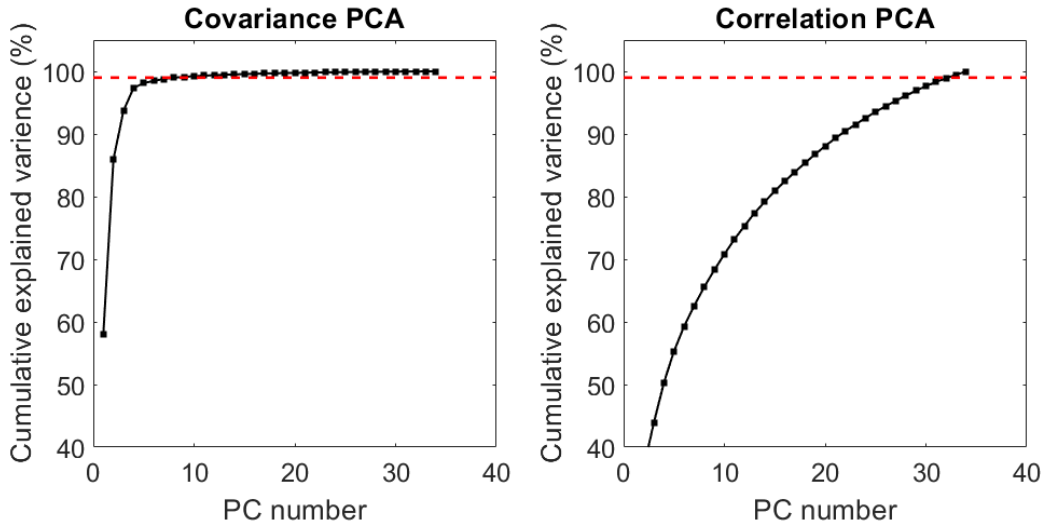


Figure 5.5: Cumulative variance explained across all Principal Components in the Sellafield reciprocal error training data. Here, the training data consisted of the first $n = 35$ measurements. Similar results were found across a range of n in the baseline period. The dashed red line indicates 99% variance.

The PCA model in an anomaly detector aims to capture the normal behaviour of the system over the training period. As such, it is hoped that it captures a high proportion of the variance present in the training data. Our training set is low rank, as the number of observations m is two orders of magnitude smaller than the the number of variables, n. As such, there is a

danger of overfitting the training set if the number of retained components approaches the number of observations, $k \approx m$. An overfit model is likely to have captured spurious features and noise in the training data, and hence will be poorly generalisable to future data. Therefore, the ideal PCA model explains the majority of variance in the training data with a small number of principal components.

5.6.1 Anomaly detection results using the reciprocal errors

First, I apply the methodology to the reciprocal error data. The initial step is to determine the appropriate data scaling for the PCA model.

The cumulative variance plots in figure 5.5 suggest that covariance PCA is better able to model the broad data structure using a small number of principal components, with $k = 9$ sufficient to explain 99% of the variance in the training set, compared to $k = 33$ for correlation PCA. This indicates that a correlation PCA model will either omit a significant proportion of the variance in the training data, or risk overfitting the training model.

The failure of the correlation PCA model to capture the model variance in a small number of principal components is likely due to the over-interpretation of tiny error fluctuations in low noise electrodes; correlation PCA weights trends in both high and low variance configurations equally. By contrast, in a covariance PCA model the measurements with the largest error variance dominate. These will tend to be the measurements with larger absolute errors which are most exposed to external error sources, and thus these carry more useful information about the system's reciprocal error regime.

Representative control charts for covariance and correlation PCA of the reciprocal error data are shown in figures 5.6 and 5.7. A spike appears in several control charts on 3rd June; this appears unconnected with the connector failure, and is discussed further in section 5.6.3.

Qualitatively, the Q profiles look similar. They both show a sustained increase from baseline after the damage occurs, followed by a further escalation in January 2014. However, quantitatively, they behave very differently.

In the correlation PCA model, Q exceeds the 99% control limit by an order of magnitude (figure 5.6c) immediately after the training period ends. This indicates that the correlation model is not generalisable outside the training period. This is likely due to overfitting, in the $k = 33$ model, as predicted from figure 5.5. A similar pattern is observed in T^2, although the confidence limit only detects anomalies after January 2014, where the magnitude of the errors increases significantly.

As the covariance PCA Q control chart remains within the control limits after the training

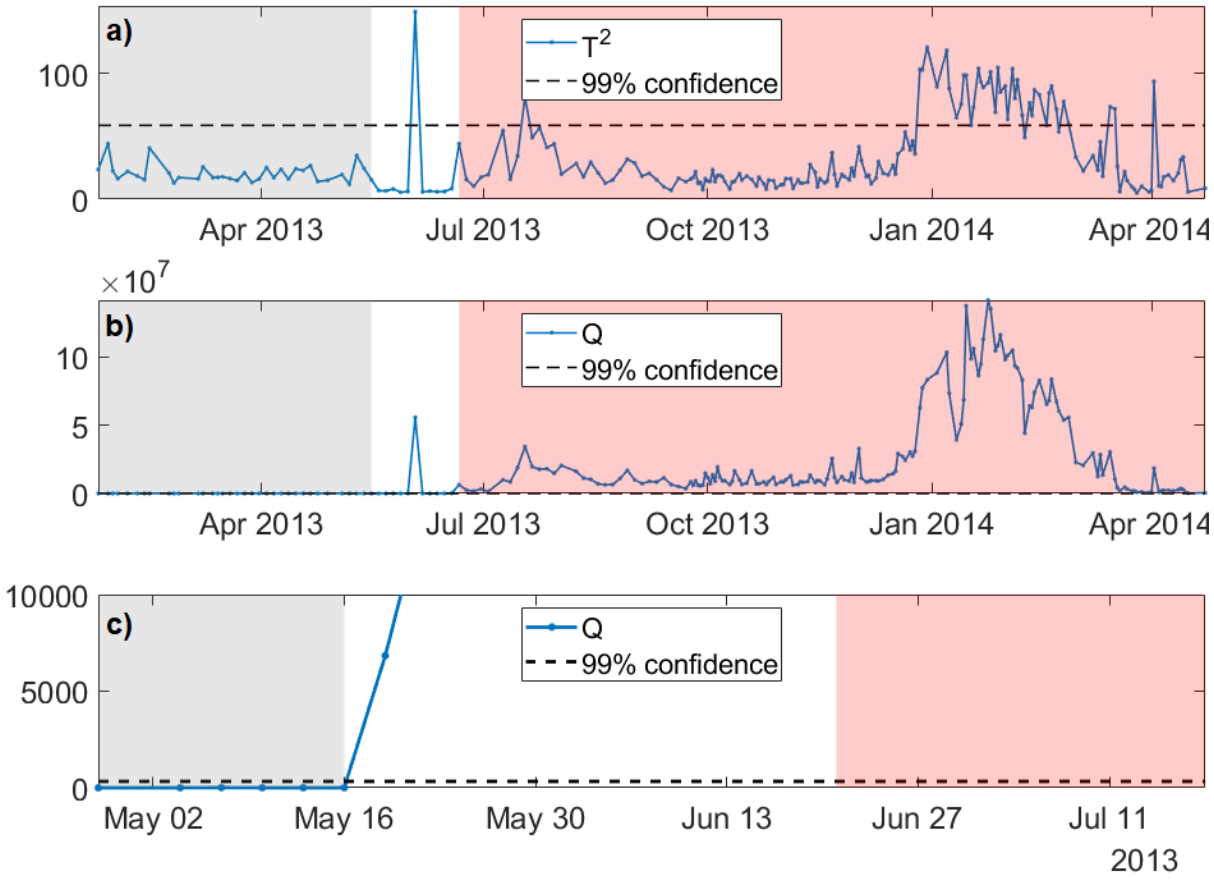


Figure 5.6: Correlation PCA control charts for the reciprocal error data from Sellafield, showing T^2 (a) and Q (b,c). (c) focuses on the period around the connector damage. The training set consisted of the first $n = 35$ measurements, highlighted by the grey region. The period after the borehole 1 connector was damaged is highlighted in red. The 99% confidence intervals are based on the assumption that the data consist of *i.i.d* random variables.

period ends, and the rate in increase in Q is small, the covariance model can be relied upon outside of the training period. The background rate of drift away from the training model is small relative to the control limit, which suggests that the *i.i.d.* assumption is useful during the 22 day period between the end of the training period and the onset of short-circuiting.

Two anomalies exceeding the control period are apparent in Q: the 3rd of June spike, and an ongoing effect beginning around the 20th or 23rd of June, which corresponds to the connector damage (figure 5.7c). The gradients here are significantly steeper than the drift rate. Both features remain present if the end of the training period is varied (not shown).

The T^2 control limit is exceeded on several occasions during the training period. This is probably due to the time dependence of the data within the training set, which violates the assumptions used to estimate the control limits. Note that much, but not all, of this rise tends to occur

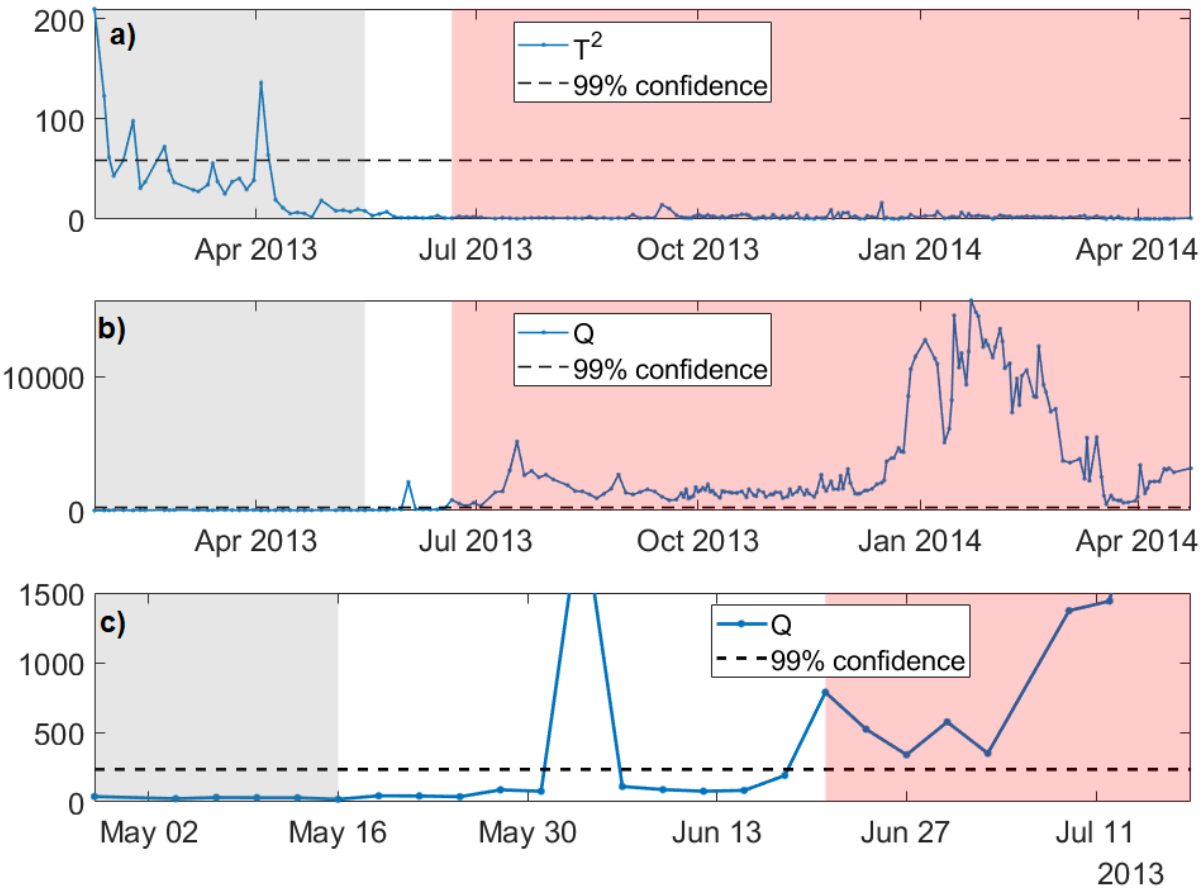


Figure 5.7: Covariance PCA control charts for the reciprocal error data from Sellafield, showing T^2 (a) and Q (b,c). (c) focuses on the region around the connector damage. (c) focuses on the period around the connector damage. The training set consisted of the first $n = 35$ measurements, highlighted by the grey region. The period after the borehole 1 connector was damaged is highlighted in red. The 99% confidence intervals are based on the assumption that the data consist of *i.i.d* random variables.

early in the monitoring sequence, before the acquisition schedule was reordered to minimise polarisation effects. I prefer to retain this data in the training period in order to maximise the information on the past state of the system. The lack of any anomalies in T^2 after the training period ends shows that the changes due to the connector damage are distinct from the original state of the system as captured by the training model.

Having identified an anomaly, we now assess the potential of PCA to identify potential sources of the data error. As Q is a square sum of the residuals, subsets of Q can easily be calculated for subsets of the data. Normalisation by the number of measurements in each subset ensures the data are comparable. Unlike the true Q statistic, the approximation of the control chart thresholds is not valid, as only part of the distribution is sampled. As such, this breakdown should only be used to investigate anomalies after they have been identified by the control

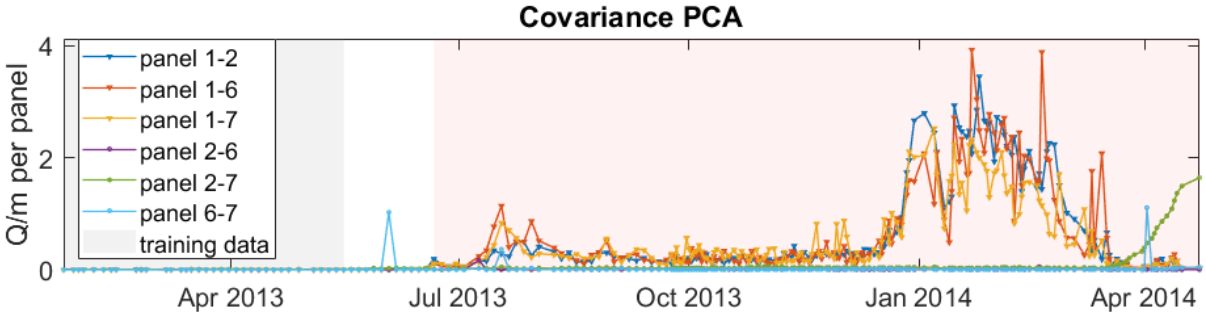


Figure 5.8: Q statistic per measurement of each inter-borehole panel for covariance PCA of the reciprocal error data. $n = 35$. The grey shading is the training period and the red shading highlights the time after the borehole connector damage occurred.

charts.

Figure 5.8 shows the subsets of Q for a series of inter-borehole 'panels' for covariance matrix PCA. Each panel is defined by a pair of boreholes, and consists of the set of all measurements between electrodes located in those boreholes. Following the connector damage, there is a clear rise in Q for only the panels containing borehole 1, a clear indication that the errors are originate from borehole 1.

The spike in Q on the 3rd June can wholly be attributed to the imaging panel between boreholes 6 and 7, and therefore it appears to be independent from the shorting effects involving borehole 1. A detailed analysis of the spike can be found in section 5.6.3.

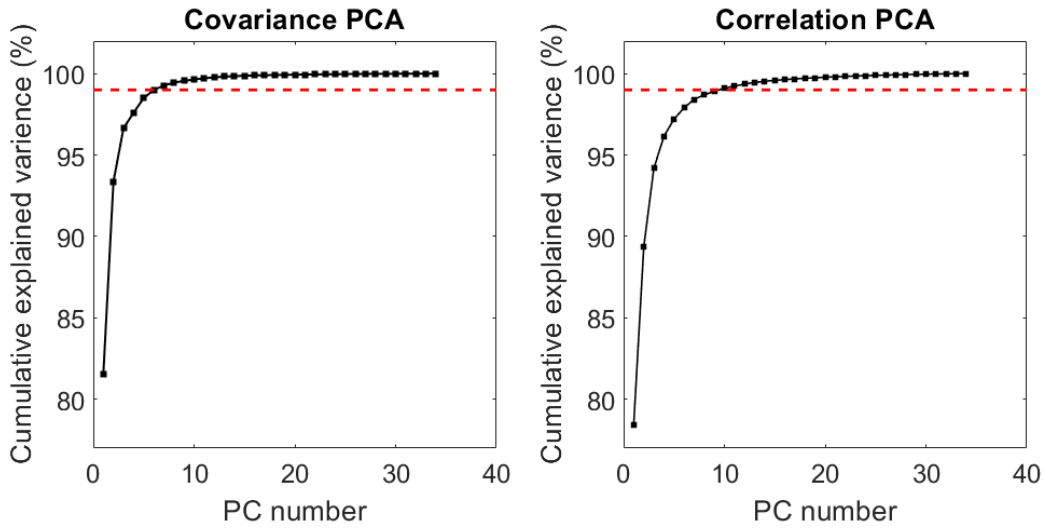


Figure 5.9: Cumulative variance explained across all Principal Components in the Seallafield apparent resistivity training data. Here, the training data consisted of the first $n = 35$ measurements. Similar results were found across a range of n in the baseline period. The dashed red line indicates 99% variance.

5.6.2 Anomaly detection using the apparent resistivity

Next, the potential for identifying the onset of shorting errors directly from the ERT measurements was investigated. In this analysis, I show the results using the apparent resistivity, where the resistances are normalised by a geometric factor, as these are more closely related to the true ground resistivity. The results are qualitatively similar if resistance is used instead.

Unlike in the case of the error data, there is no significant change in the resistivity distribution at the onset of connector damage (figure 5.4). The anomaly detection therefore relies upon detecting changes to the correlation/covariance structures in the resistivity data.

Both covariance and correlation matrix PCA capture 99% of the variance with $k < 10$ principal components (figure 5.9), therefore both are viable candidates for a control chart model.

The control charts for covariance matrix PCA of the resistivity data for a range of training periods are shown in figure 5.10. Again, T^2 exceeds the control limit early in the training period, reflecting the non-stationarity of the data, and potentially also the measurement scheduling changes in early March. T^2 remains below he confidence limit at later times, suggesting no anomalous behaviour inside the PCA model subspace. For the shorter training periods, Q begins to rise steadily immediately after the training period ends, suggesting that these models are poorly generalisable to future time periods. There is no clear increase in Q indicative of a sharp departure from the training model (figure 5.12a,b).

The Q control chart for $n = 35$, figures 5.10f and 5.15, displays a lower rate of drift initially, remaining within the control limits until the connector damage occurs. Due to the background drift, the gradient of the control chart is the primary signifier of an anomaly, not the point at which the control chart is exceeded, which is affected by the interval since the training period. A five-fold increase in the gradient of Q between the 18th-23rd June marks the connector damage (figure 5.12c). This effect is weaker, relative to both the control limit and the background drift rate, than in the reciprocal error control chart (figure 5.10). Without independent knowledge of the connector damage, a small increase in gradient could easily be missed. Likewise. the modest increase in Q on June 3rd is too weak to be identified as an anomaly.

The rate of background drift away from the training models for $n = 10$ and $n = 20$ renders them unsuitable for use as anomaly detectors, as it obscures any increase in Q arising from unusual system changes (figure 5.12). However, the $n = 35$ model appears more stable. This example highlights the need to ensure that a stable training model has been achieved before using the control chart for anomaly detection.

A similar rate of background drift in Q contributions from all panels makes interpretation of the inter-borehole panel breakdown more challenging than in the reciprocal error analysis.

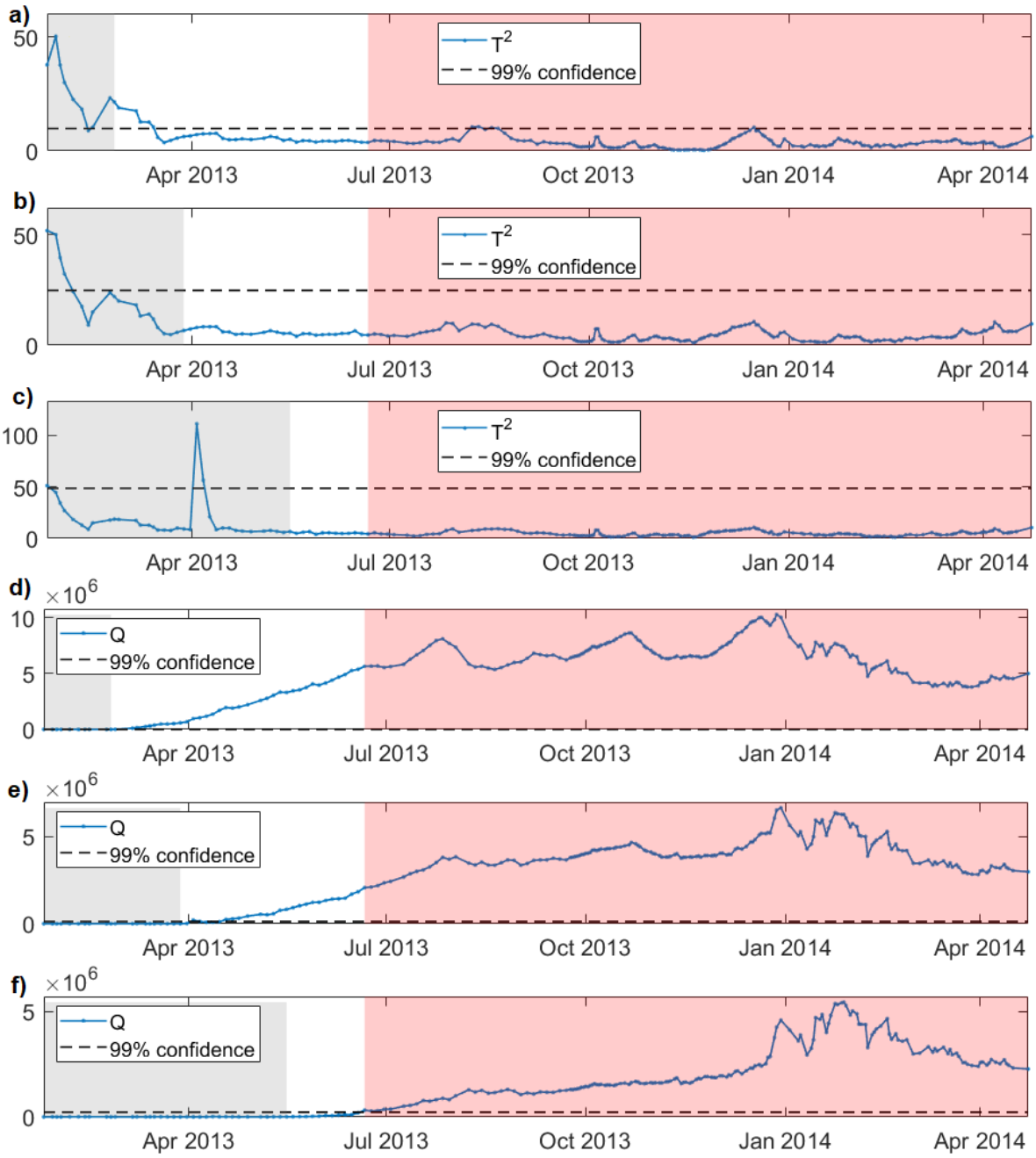


Figure 5.10: Covariance matrix PCA control charts T^2 for the Sellafield resistivity data. The grey shading highlights the training period, which consisted of $n = 10$ time steps for (a) and (d); $n = 20$ for (b) and (e); and $n = 35$ for (c) and (f). $k = 5, 6, 7$ respectively. The red shading highlights the period after which the connector damage is thought to have occurred. The 99% confidence intervals are based on the assumption that the data consist of *i.i.d* random variables.

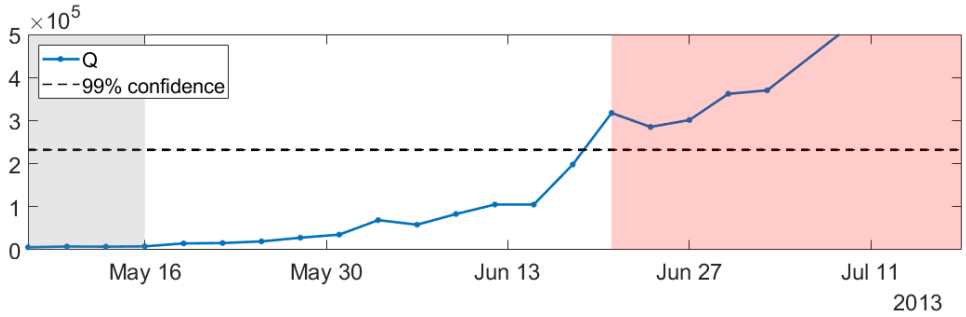


Figure 5.11: Detail from figure 5.10f, showing the drift in the Q statistic after the end of the ($n = 35$) training period. The grey shaded region marks the training period, whilst the red shaded region marks the period after the connector-damage is believed to have occurred.

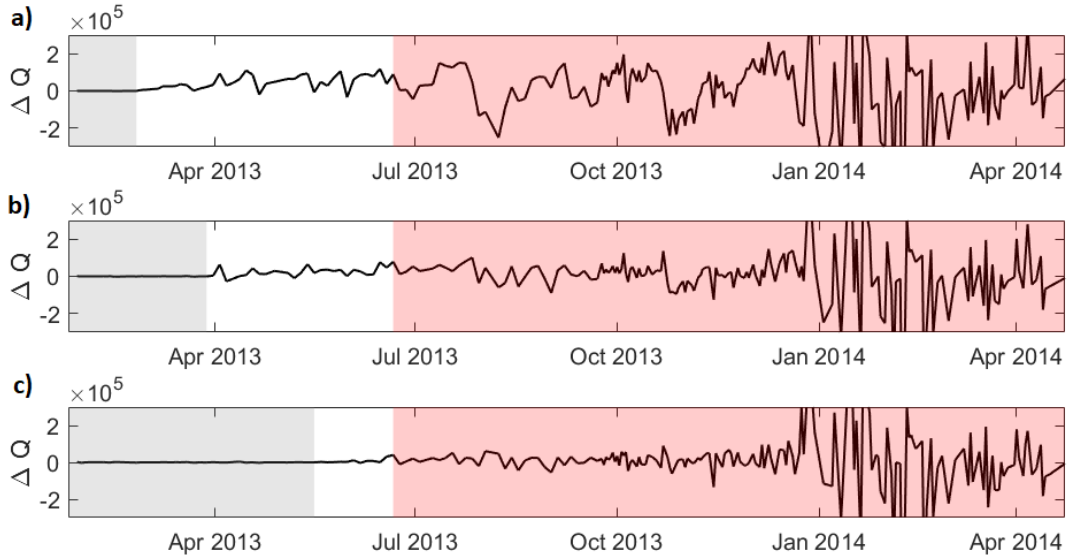


Figure 5.12: Data differences, $\Delta Q = (Q_n - Q_{n-1})./(t_n - t_{n-1})$, for covariance matrix PCA. (a), (b) and (c) are for training periods of length $n = 10, 20$, and 35 respectively, marked by the grey shaded regions. The red regions denote the period post connector-damage.

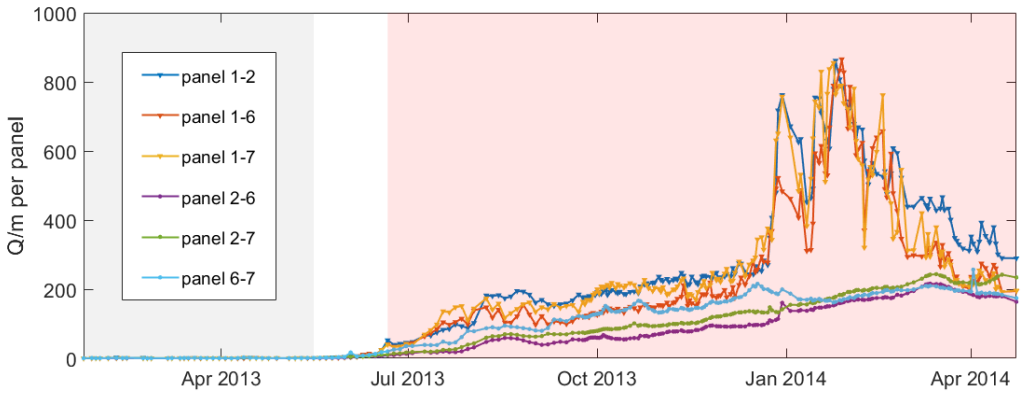


Figure 5.13: Q statistic of each inter-borehole panel for covariance matrix PCA of the resistivity data. The training period was $n = 35$. The grey shaded region marks the training period, whilst the red shaded region marks the period after the connector-damage.

These are shown in figure 5.13 for $n = 35$. While the Q contributions for panels 1-2 and 1-7 are largest, the Q contributions for panels 1-6 and 6-7 overlap significantly before December 2014. The contributions from each panel appear to exist in a continuum, rather than in clearly separable groups. The equivalent plots for the shorter training periods display no clear separation between panels.

Overall, the performance of the covariance PCA resistivity control charts are unsatisfactory, due to low sensitivity to the anomaly and indistinct separation of the Q contributions for the inter-borehole panels.

The correlation PCA control charts performed better for the resistivity. For all three training lengths shown, the connector damage is identifiable by an increase in the gradient of Q (figures 5.14, 5.15, 5.16). These were larger than the background drift by factor of 4 for $n = 10$ and over an order of magnitude for $n = 20$ and $n = 35$; the anomaly is unambiguous. For $n = 35$ the distinction is most clear, emphasising the need to update the PCA model with more recent results (providing no unaddressed anomalies are included in the training period). Confidence in the identification of an anomaly is increased if it appears in multiple training lengths. The detail in figure 5.15 shows that the drift rate in Q is larger relative to the control limit than seen in previous examples; the control limit is exceeded from the third of June onwards. The anomalies on June 3rd and 18th July are an order of magnitude larger than both the gradient of the drift and the control limit, and thus are easily identified.

The T^2 control charts in figure 5.14a-c show few signs of anomalous behaviour outside of the training period. A spike in late December/early January should not be over-interpreted, as it occurs months after the training period has ended.

Borehole 1 is clearly identifiable as the source of the anomaly in the per panel breakdown of Q for $n = 35$ (figure 5.17c). The drift in the other inter-borehole panels is consistently small, which allows the changes to the borehole 1 panels to be immediately distinguished. For $n = 15$ and $n = 10$, borehole 1 is not identifiable as the main source of the anomaly until, respectively, several weeks or months later, due to a higher rate of drift (figure 5.17a,b). This demonstrates the necessity of using an up-to date PCA model.

Correlation matrix PCA is preferred for the analysis of the Sellafield resistivity data due to its strong detection of the anomaly and the isolation of the anomaly to borehole 1.

5.6.3 Peak on the 3rd of June

A large spike in the Q statistic was observed on the 3rd of June in both control charts. As this spike far exceeds the control limits, it is classed as a point anomaly. Investigations into the source of the spike were made in order to understand the potential cause.

The panel by panel breakdown of Q in figure 5.8, indicates that this anomaly may occur in the imaging panel between borehole six and seven, and therefore it appears unrelated to the short circuiting errors which occur later.

One possibility is that errors in the interpolation process may cause a data anomaly, particularly if many data were missing during that time step. On the 3rd of June there were 36 missing data values (0.3%), compared to a median of 48. Therefore it is unlikely that missing data are responsible for the observed behaviour.

A more likely source of the anomaly is the small increase in errors on the 3rd of June, which can be seen in figure 5.2. There were 31 errors $> 2\%$ on the 3rd of June, compared to an average of 6 for dates prior to that time. Of these $31 < 2\%$, 27 involve electrode 104, which is located in borehole six. Therefore, the peak on the 3rd of June can be attributed to a temporary fault or disturbance to electrode 104, even if the source of that noise is unidentified. As there is no indication of short-circuiting effects here, and the effect was not ongoing, these effects can be accounted for by assigning appropriate error weighting to the data from electrode 104 in the inversion, or else removing them from the analysis.

Using the covariance PCA reciprocal error control chart, the Q statistic can be used to identify electrode 104 as the source of the 3rd June peak directly, from the sum of the contributions to Q for each electrode. On the 3rd of June, the contribution to Q from electrode 104 was a factor of 5000 larger than those electrodes with the largest contribution to Q in the neighbouring time periods. The contributions from the other electrodes that were involved in measurements with electrode 104 were also elevated, suggesting that an additional step is required to fully isolate individual noisy electrodes.

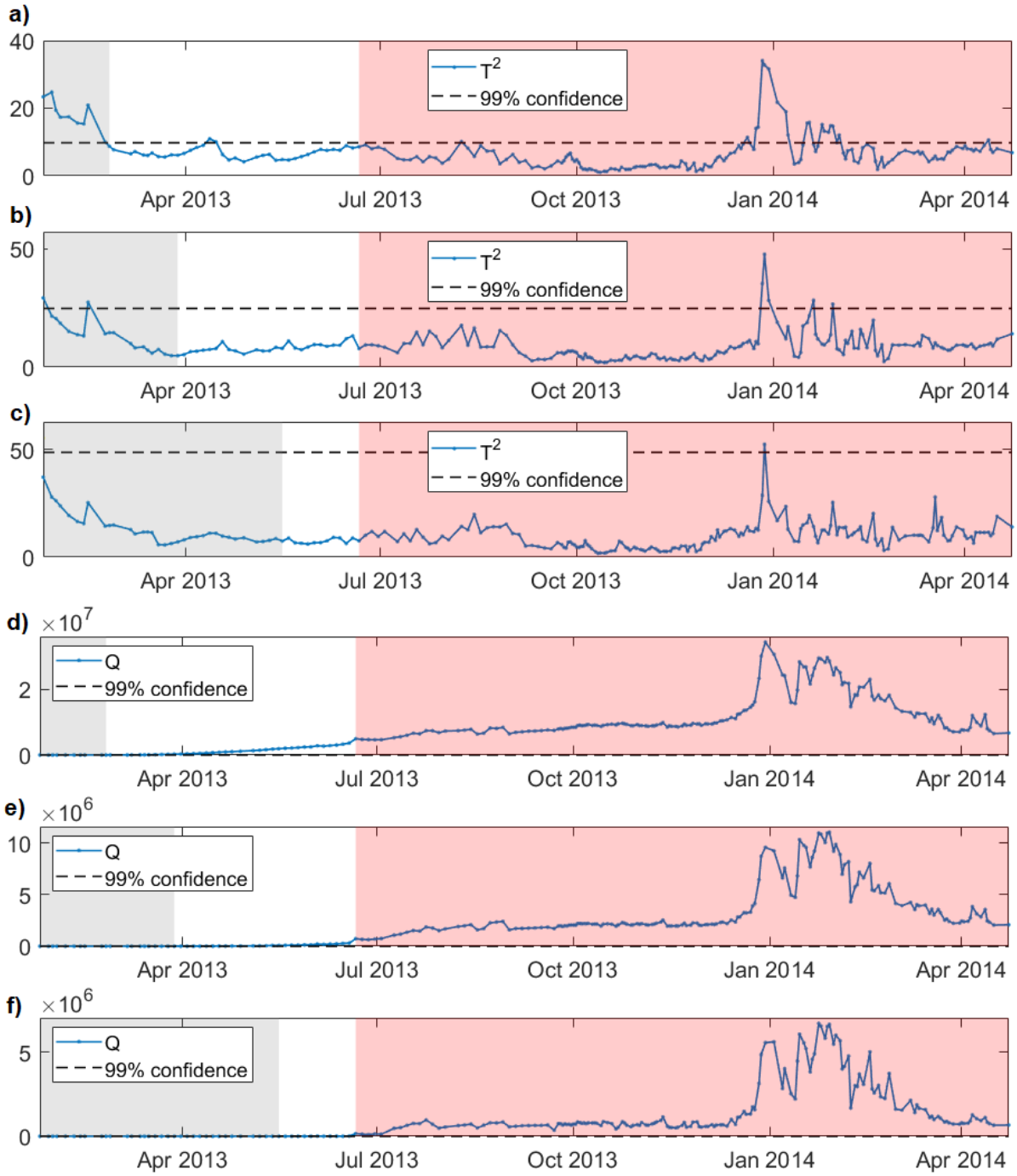


Figure 5.14: Correlation matrix PCA control charts T^2 for the Sellafield resistivity data. The grey shading highlights the training period, which consisted of $n = 10$ time steps for (a) and (d); $n = 20$ for (b) and (e); and $n = 35$ for (c) and (f). $k = 7, 9, 10$ respectively. The red shading highlights the period after which the connector damage is thought to have occurred. The 99% confidence intervals are based on the assumption that the data consist of *i.i.d* random variables.

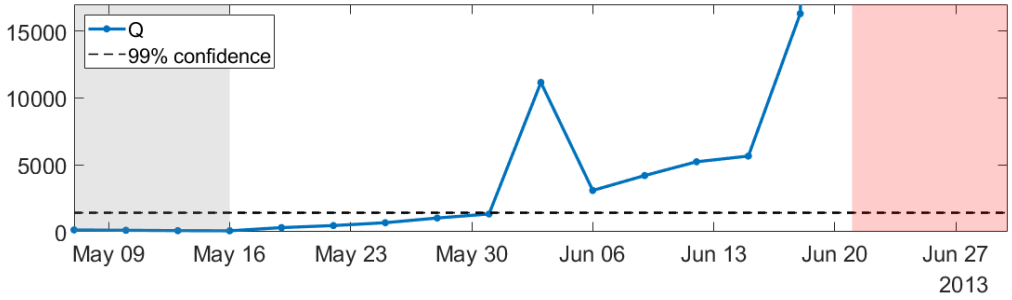


Figure 5.15: Detail from figure 5.14f, showing the drift in the Q statistic after the end of the ($n = 35$) training period. The grey shaded region marks the training period, whilst the red shaded region marks the period after the connector-damage is believed to have occurred.

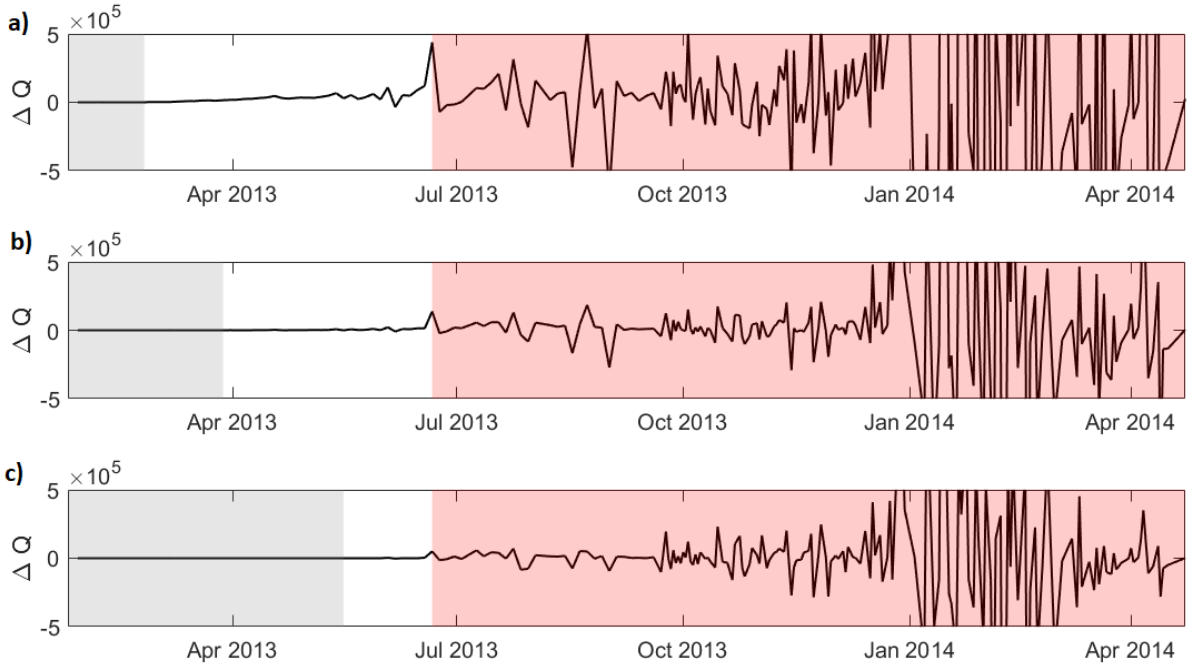


Figure 5.16: Data differences, $\Delta Q = Q_t - Q_{t-1}/(t_n - t_{n-1})$, for correlation matrix PCA. (a), (b) and (c) are for training periods of length $n = 10, 20$ and 35 respectively, marked by the grey shaded regions. The red regions denote the period post connector-damage.

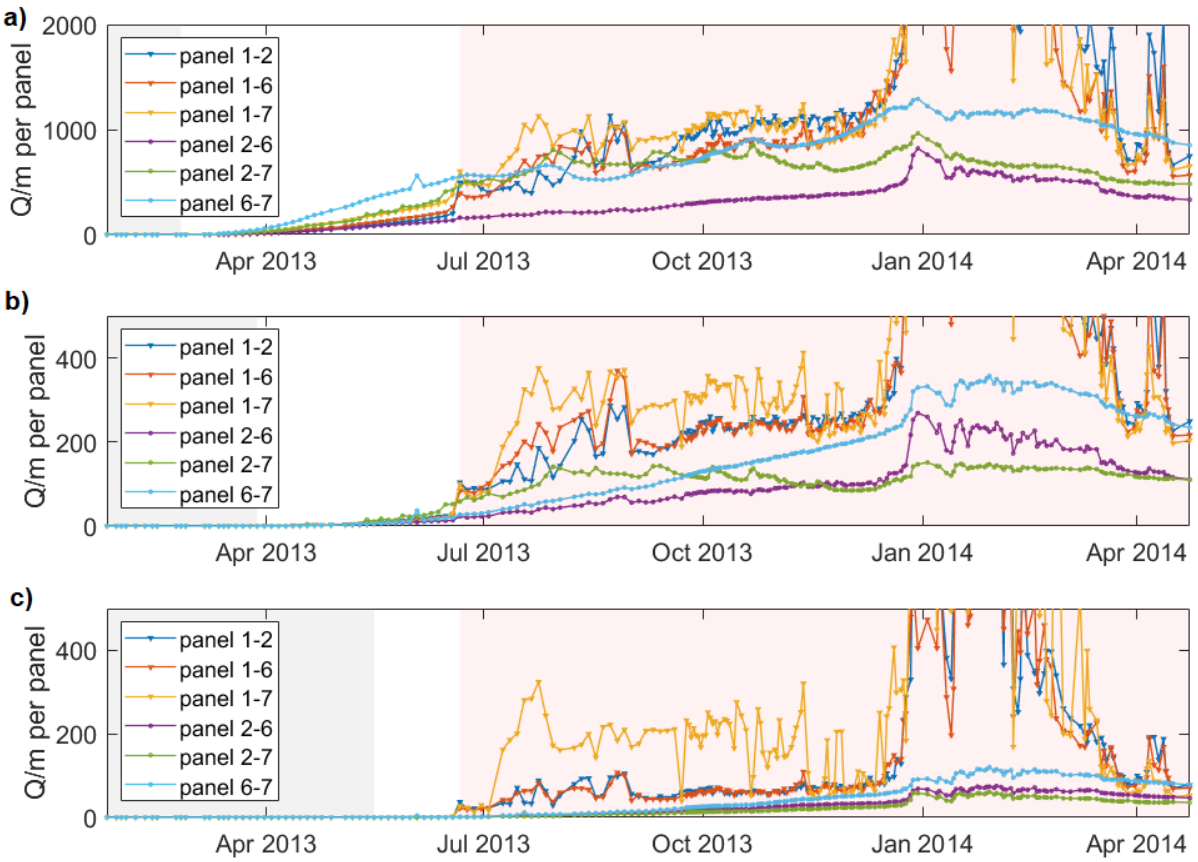


Figure 5.17: Q statistic of each inter-borehole panel for correlation matrix PCA of the resistivity data. (a), (b) and (c) are for training periods of length $n = 10, 20$ and 35 respectively, marked by the grey shaded regions. The red regions denote the period after the connector-damage.

If all errors $> 2\%$ are filtered out of the reciprocal error analysis, the 3rd of June spike vanishes entirely and there is no appreciable change in any of the control chart responses to the short circuiting errors.

This sensitivity to a new error pattern in measurements involving a single electrode has implications for the interpretation and analysis of the PCA control charts. On the one hand, the reciprocal error covariance PCA control chart successfully detects unusual behaviour in measurements involving electrode 104, which can be identified using a stem plot of the Q contributions from each electrode. This would allow remedial action to be taken.

On the other hand, this anomaly in a small subset of measurements linked to a single electrode produced a stronger, if transient, response in Q in the error control chart than the initial connector damage. This runs the risk of misleading operators as to the severity of the anomaly. The magnitude of an anomaly in Q is not necessarily proportional to its impact, and further investigations are needed to identify the source of any anomaly. For best practice, an interpretation work flow should be developed in order to triage anomalies effectively.

5.6.4 An alternative visualisation

As a sum-of-squares residual metric, the Q statistic can be misleading. The magnitude of the Q statistic depends on the sum-of-squares magnitude of the underlying data. In a traditional Statistical Process Control setting this distinction is unimportant, as the underlying data is assumed to be drawn from a stationary distribution. Any increase in Q over the control limit is positively identified as an anomaly. As the ERT monitoring data is not stationary, Q naturally increases over time after the training period ends as the system state drifts away from the model. From the rate of this drift, and the Q control limit, it has been possible to identify the anomalous changes which correspond to the connector damage in the case study. However, care needs to be taken not to over-interpret changes in Q which may be due to variation in the sum-of-squares magnitude of the data itself. For example, the correlation PCA control chart in figure 5.6 generalises poorly to new data, so the observed changes in Q are largely due to the increase in reciprocal errors seen at that time.

In order to control for these changes in the underlying data, the ratio of the root-square-sum data residual to the root-square-sum of the data can be used: the 'Q ratio'. The Q ratio can be intuitively interpreted as the proportion of the data which lie outside of the PCA model space. Moreover, using this normalisation, it is possible to compare the rate of drift and the magnitude of anomalies between control charts to some extent.

This presentation somewhat obscures the relationship between the Q ratio and the control limits. The transformed control limit is now a function of the root-square-sum data magnitude,

and the difference between the control limit and the Q ratio at different time steps are not easily visually comparable. Only the proportional difference between Q ratio and the control ratio is comparable between time steps. Fluctuations in the Q ratio below the control limits are not meaningful in the context of anomaly identification, and are not directly comparable along the time axis to changes which later exceed the limits.

Figure 5.18 contains the Q ratio equivalents of the control charts in sections 5.6.1 and 5.6.2. The same PCA modalities are preferred: covariance PCA of the reciprocal error and the correlation PCA of the resistivity show increases of 40% and 60% in the proportion of data outside of the PCA model between the 15th and 21st of June. Clear anomalies on the 3rd of June also occur in both.

The uptick in Q ratio begins on the 18th of June in both the resistivity and reciprocal error results, suggesting that the connector damage may have occurred some days sooner than was originally identified from the reciprocal error data. This is one data point prior to the red shaded region on the control charts.

The Q ratio jumps to 80% immediately after the training period ends for the correlation PCA model of the reciprocal errors, again demonstrating that this model overfits the training set and is poorly generalisable outside of it.

The onset of connector damage appears as a modest 12% increase in the Q ratio between the 15th and 21st June for covariance PCA of the resistivity. This is largely masked by the steady drift rate away from the model. Without prior knowledge of the connector damage, it would be difficult to draw conclusions from this signal.

5.7 Discussion

A new anomaly detection data quality methodology for ERT monitoring has been developed and successfully used to detect the onset of short-circuiting errors in the Sellafield case study, and some basic heuristics have been developed for establishing a valid PCA model. Reciprocal error and resistivity control charts were both used to independently identify the onset of short circuiting errors, and both correctly localised their source. Short-circuiting errors are poorly captured by conventional ERT data quality measures, such as the reciprocal error, so the ability to infer their presence through statistical changes in the resistivity data is a particularly promising development.

The Q statistic control chart, which measures the data residual outside the PCA model, proved effective in identifying the short-circuiting errors. The interpretability of the control charts can be somewhat improved through the Q ratio re-scaling, which estimates the magnitude of the

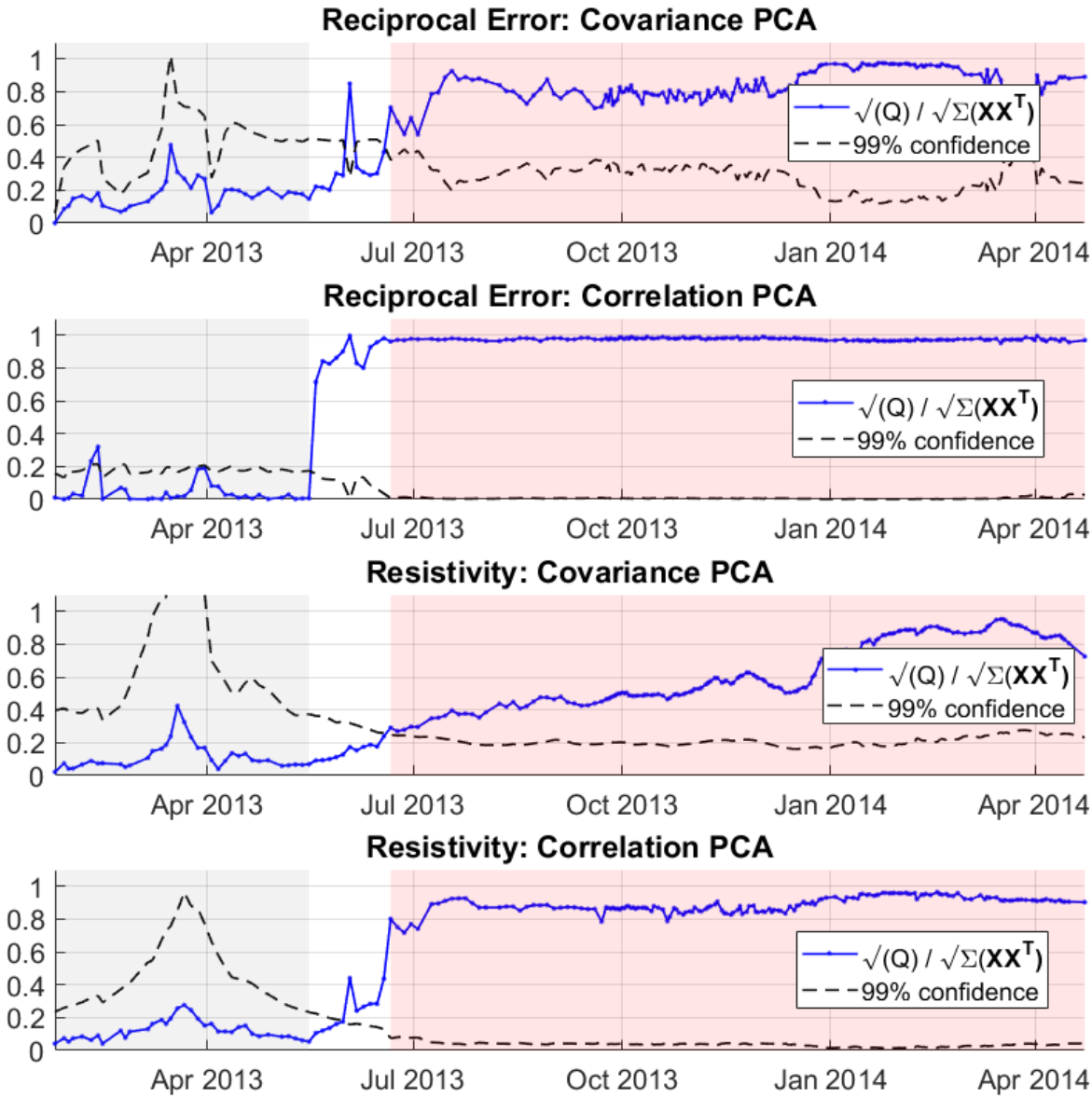


Figure 5.18: Q control charts for both covariance and correlation PCA of the reciprocal error and resistivity data, presented as the ratio of the root-square-sum of the data residual to the root-square-sum of the data. The Q control limits have been similarly scaled. The grey and red shaded areas represent the training and post-damage periods respectively. The training period was $n = 35$ and the k was chosen to retain 99% of the variance.

data residual relative to the data. A breakdown of the inter-borehole panel contributions to the Q statistic of both data localised the source of the errors to borehole 1. It is possible to break down the Q contributions further, to the level of individual electrodes. As each measurement involves four such electrodes, this process involves significant cross contamination whenever an electrode is included in a quadropole involving noisy electrodes, and therefore further research is required to reliably identify individual faulty electrodes from Q contributions alone.

The ability of the PCA control chart method to detect short-circuiting errors directly from statistical changes to the observed resistivity data is perhaps the most promising result, as these errors are typically challenging to identify. An anomaly corresponding to faulty electrode 104 was also detected by the resistivity control chart; there is at least some sensitivity to conventional error types. The control chart method may be useful for general error detection when reciprocal error data is unavailable, although more research and testing is required. The reciprocal error is generally more reliable, easier to interpret, and requires no training period. The main use-case for the control chart method is to complement the reciprocal error measurements, and detect anomalies corresponding to error types which the reciprocal error is not sensitive to.

When applied to the reciprocal error data, the PCA control chart detected anomalies for both the June 3rd single electrode fault as well as for the connector damage. This is in contrast to threshold-based methods, which cannot be calibrated satisfactorily to detect low magnitude changes. This anomaly detection paradigm, alongside the ability to use Q contributions to identify potential error sources, could form a useful data quality paradigm for ERT monitoring. However, this depends heavily on the sensitivity of the method, the rate of false positives, and the ability to quickly and easily identify the sources of noise. More work is needed. An alternative approach could utilise a combination of thresholds and monitoring the moments of the error distribution.

In this case study, short circuiting errors appear in the error estimates through secondary polarisation effects, due to short circuited electrodes passing current out of order. It also allows for an independent method to test for short circuiting errors, or other potential anomalies, alongside other data quality measures, or where reciprocal error measurements are not available.

An important part of the method is ensuring that the PCA model generated during the training period can be usefully generalised to future time periods. For each data modality, a choice needs to be made between Covariance or Correlation matrix PCA. An initial training period needs to be determined, as does a suitable update rate for the PCA model. An initial screening

of the cumulative variance plots per principal component was able to identify PCA models which cannot capture 99% of the variance of the data for $k < 10$ components, and therefore are overfitting the training data. In the next step, the gradient, or drift rate of the Q statistic away from the PCA model relative to the 99% confidence thresholds for stationary data are compared. Where the initial drift rate is smaller than the control limit, confidence in the model is increased. The Q ratio provides a direct estimate of the rate of departure from the PCA model, which is valuable for identifying models which generalise poorly and for separating anomalies from the background drift rate. An anomaly can be confidently identified as a large increase in Q. Confidence in an anomaly is increased if it appears in control charts of varying training length.

In the case study, covariance matrix PCA was preferred for the reciprocal error data, whilst correlation matrix PCA performed best for the resistivity data. This is perhaps unsurprising. Covariance matrix PCA depends on the relative magnitudes of the data variances. In the reciprocal error data, these carry significant information about the severity of the detected errors. A change in reciprocal error magnitude is potentially indicative of an anomaly. Therefore, I can tentatively recommend using covariance PCA for reciprocal error control charts in future, although it remains best practice to trial both types.

By contrast, correlation matrix PCA scales the data variances so that they are independent of the magnitude of the data. It is the correlations between the ERT measurements which are important in forming the model. We expect to find these correlations between measurements with common electrodes and overlapping sensitivity regions. This correlation pattern is disturbed by the short circuiting errors, and hence appears as a clear anomaly in Q.

Another important variable is the length of the extrapolation period between the end of the training window and the detection of an anomaly. Ideally, this period is short, so that all the previous behaviour of the model is included. However, too short a period obscures useful information about the background drift rate away from the PCA model, which increases confidence in a true anomaly.

5.7.1 Experimental limitations

This analysis used a single case study for a proof-of-concept test of the methodology. Wider and more rigorous testing is required to determine the behaviour of the control charts in the wide variety of circumstances which may occur during ERT monitoring. This is needed, as this is primarily an empirical method.

The Sellafield case study involves a relatively stable system, which is likely to be advantageous for establishing a training model which generalises well to future data. The Sellafield

data measured a subsurface cell using electrodes embedded in boreholes, which dampens the effect of short term temperature and rainfall changes relative to a surface array. The resistivity changes under observation were gradual, and low contrast, with cumulative changes in resistivity due to the leak reaching a maximum of $\pm 10\%$ per model in the most effected areas. It would be interesting to test the method on more dynamic resistivity systems. If rapid resistivity changes are frequently misidentified as anomalies, the usefulness of the method in those systems will be limited.

A transient anomaly was detected on June 3rd due to a noise increase in a single electrode (see section 5.6.3). The ability to detect both small and large scale anomalies is potentially both a strength and a weakness of the method. However, a robust workflow is needed in order to quickly assess the type of anomaly when one does appear, and to filter out those anomalies, like the June 3rd peak, from the expanded training set as more measurements are received. Based on the case study, I would expect an analysis of the Q contributions per electrode and per cable to be an important part of this. An analysis of the individual electrode contributions to Q may need additional statistical filtering, such as the linear mixed effect model used by Tso et al. [2017], in order to avoid the inadvertent flagging of a functional electrode as anomalous if it is involved in quadropoles alongside faulty electrodes. If the Q contributions can be analysed efficiently, this may prove a valuable means of identifying the sources of potential errors.

As the PCA control chart method needs a training model to be established, it requires a lead in period before it can be useful. This method is therefore complementary to existing data quality techniques, which must be relied upon during the initial set up and training period, by identifying those anomalies that are poorly detected by traditional methods.

While these results are promising, further testing on historical field data is required to estimate the robustness, or consistency, of the method. Ideally a range of different known errors would be tested (faulty individual electrodes, external noise sources, large changes in contact resistance, moving electrodes), as well as control time series when no significant error sources were present. In addition, testing data sets where significant rapid resistivity changes would be useful to identify the sensitivity to non-error related anomalies.

5.8 Summary

A PCA control chart anomaly detection data quality method for ERT monitoring has been developed and poof-of-concept tested on a case study data set. The onset of short-circuiting errors in the Sellafield case study data was successfully detected using both reciprocal error and resistivity data. The Q statistic, which measures the deviation of new data from the trained PCA model, proved the most valuable tool. Reciprocal error and contact resistance measures

are insensitive to short-circuiting errors, so the ability to detect short circuiting errors directly from the resistivity distribution is valuable. Further work is now required to understand the behaviour and sensitivity of the method to other error types, as well as its effectiveness under different survey and ground conditions.

Chapter 6

Conclusions

6.1 Summary of Contributions

In this thesis, new methods and techniques for the processing of time-lapse ERT monitoring data have been developed and tested. The two main areas of interest were the development of an improved regularisation method for piecewise smooth inversion, which is of particular interest for time-lapse data, and the utilisation of time-domain information to detect data quality anomalies.

In Chapter 2 I reviewed the literature on the the ERT data processing pipeline. I identified promising techniques for the improvement of two key areas: regularisation of the inversion and data quality measures.

In Chapter 3 I derived and implemented a novel algorithm for the the inversion of ERT data, using the piecewise-smooth TGV regularisation functional. Piecewise-smooth regularisation is beneficial for time-lapse data, where a mix of smooth and sharp changes are expected. Initial tests of the algorithm in 2.5D produced generally promising results, which nevertheless contained inconsistencies. A systematic evaluation of the algorithm was performed, leading to adjustments to the algorithm. The initial model and the IRLS stabilising parameters were modified in order to reduce over-smoothing in the TV and TGV solutions, and periodic boundary conditions were introduced.

In Chapter 4 a comprehensive parameter exploration of the 2.5D TGV algorithm was performed, its performance was assessed, and best practice guidelines for its use were developed. The results of the TGV inversion were found to be sensitive to the initial model, and smooth behaviours were suppressed near the model boundary. A double inversion approach, where the result of a TGV line search inversion are used to initialise a second TGV inversion, was found to reduce the influence of the initial model on the solution. Despite this, there were still

significant differences in the solutions for different initial models. The preferred initialisation was the model pseudosection, which resulted in solutions which were much closer to the expected TGV smoothing behaviour. The TGV trade-off parameter, μ, regulates the smoothness of the solution as expected. The optimum choice of μ depends on the initialisation, but for the pseudosection-initialised double inversion, $\mu \approx 1.4$ performed well in the tests. The effects of varying μ are gradual, and the solutions are not overly sensitive to small changes. This is desirable, as a more sensitive μ dependency would rely too much on the human operator's judgement and bias. While the TGV inversion produced small improvements in the model misfit compared to the TV inversion, there is scope for further refinement. Challenges for future development include additional mitigation of the sensitivity to the initial conditions, and the reversion to TV-like behaviour near the boundary. A discussion of potential future work on the TGV functional is found in Section 6.2.

In Chapter 5, a PCA based control chart data was developed in order to detect errors in monitoring data. A training period is used to develop a PCA model, which needs to be updated as the system changes over time. Deviations from that model are measured using a Q^2 statistic control chart. The method proved successful in detecting the onset of short circuiting errors in the Sellafield case study. This is a promising proof of concept, which justifies further testing and development of the method on more data sets. Further discussion of this can be found in Section 6.3.

The aims of this thesis were to develop new methods to improve the processing of ERT monitoring data, in both the data quality and inversion stage. The new piecewise smooth TGV regularisation method was successfully developed and tested, and parameter testing identified a recommended optimum configuration, in which the method outperforms both TV and l_2 inversions. However, the remaining dependence on the initial model and performance limitations near the model boundary may require further work to be overcome. The PCA control chart method was developed and successfully tested on the Sellafield case study as a proof of concept. It is now ready for further development and, ultimately, trialling on ERT monitoring data.

6.2 TGV Future Work

The aim of the development of a TGV regularisation algorithm for ERT inversion is to be able to better represent smooth and sharp changes in the same model, particularly where the smooth regions vary in magnitude and gradient and so cannot be well represented by MGS-type techniques. This would be particularly valuable for time-lapse ERT inversions, where both smooth and sharp transitions occur in the time domain.

In this thesis, a viable TGV algorithm for ERT inversion has been developed, and demonstrably performs better than l_2 and TV at capturing piecewise smooth behaviour for a 2D test model when configured correctly. However, it has also been shown to be significantly more sensitive to the initial model than the other regularisation functions, and often converges at local minima where the piecewise smooth behaviour is poorly represented. Best practice guidelines have been developed to increase the reliability of the TGV inversion, but the dependence on the initial conditions remains, and the best initial model may vary depending on the smoothness of the ground resistivity. In addition, the best performance of the TGV regularisation was limited by a reversion to TV-like behaviour near the model boundary, where smooth gradients were poorly captured.

In this section, I discuss a range of directions for future work on the TGV regularisation of ERT data, which have the potential to deliver further improvements. Several of these relate to the finite difference method and the modelling of the boundary conditions.

6.2.1 Testing on further data

The parameter space of the TGV inversion is large, due to the range of parameters which must be captured. This led to compromises in the testing regime in this thesis. A single synthetic model, containing a mix of smooth and sharp features, was used for a large set of tests in order to establish the baseline properties of the TGV algorithm. However, it would be desirable to test the algorithm with a wider range of data, for a variety of model smoothness structures, including field data, and realistic noise conditions for example.

6.2.2 Extension to time-lapse data

The motivation for the development of the TGV regularisation algorithm was for use in the inversion of time-lapse monitoring data. For ease of interpretation, the implementation and analysis in this thesis used 2.5D models, with TGV regularisation in the spatial domain . The application of TGV regularisation to the time domain, using the 4D inversion approach of Kim et al. [2009], is one of implementation, rather than research. In the current implementation, there is likely to be a small advantage in the representation of piecewise smooth changes for an appropriate initial models. In 4D ERT, the combination of spatial and time domain regularisation means that there are stronger constraints on each model cell, which may reduce the dependence of the TGV solution on the initial model, and therefore lead to better performance. If that is the case, then a hybrid regularisation scheme, where TGV regularisation in the time domain is combined with a different spatial regularisation function, may prove effective. However, as improvements to the spatial performance of the TGV inversion algorithm will also benefit its performance on time-lapse data, it may be more effective to focus initial

efforts on the performance of the algorithm before extending the regularisation to the time domain.

6.2.3 A spatially varying TGV trade-off parameter

Only when $\mu <= 1.0$ was the TGV inversion able to capture the smooth gradients in the upper reaches of the model. Otherwise $\mathbf{p}$ was suppressed in these areas, and the local regularisation reduced to TV.

However, these small values of μ resulted in over-smoothing in other areas of the model, particularly in the vicinity of the blocky feature. It is likely that the optimum μ will vary spatially within models, due to factors like model resolution, proximity to the model boundary and the local rate of change in the gradient.

The Active Constraint Balancing (ACB) [Yi et al., 2003], and the time-lapse equivalent, the Active Time Constraint (ATC) [Karaoulis et al., 2013] vary the regularisation across the model via a spatially varying Lagrange multiplier, λ. For ACB inversion, the parameter spread functions of the model resolution matrix are used to vary the Lagrange multiplier inversely to the resolution at each model cell. For the 4D-ATC inversion, the Lagrange multipliers are varied based on local estimates of the rate of resistivity change, which are derived from the first iteration of an independent inversion of each time step.

A similar approach could be used to vary μ spatially within models, which might allow the TGV inversion to better capture gradients near the model boundaries. For a time-lapse TGV inversion, an approach similar to the 4D-ATC inversion would be appropriate; in this case the local smoothing would be regulated by tuning μ, instead of by increasing or decreasing the strength of an l_2 smoothing term.

It is possible that an approach which utilises the model resolution in a similar manner to the ACB might be effective at controlling μ. Where the model resolution is higher, the data is better able to determine the smoothness of the model, so a μ approaching 1 allows for smooth or sharp gradients to form based on the gradient alone. Where the resolution is lower, a larger μ allows sharp edges to form.

6.2.4 Second order Neumann boundary conditions

An alternative approach to solve the suppression of $\mathbf{p}$ near the model boundary might be to change the boundary conditions themselves. The adoption of periodic boundary conditions did not lead to improvements in the ability of $\mathbf{p}$ to capture gradients near the boundary without promoting oversmoothing in the rest of the model. Periodic boundary conditions are not physically realistic, but were chosen for their internal consistency between $\mathbf{m}$ and $\mathbf{p}$, and in

order to reduce errors from the asymmetrical finite differences by avoiding any special treatment at the boundary. It is likely that a better choice of boundary condition might improve the performance of the TGV algorithm for smooth features near the model boundaries.

A proper treatment of the boundary conditions should reflect the relationship between $\mathbf{p}$ and $\mathbf{m}$. This can be achieved using homogeneous Neumann boundary conditions of different orders: first order in $\mathbf{p}$ and second order in $\mathbf{m}$.

In smooth regions, $\mathbf{p}$ is an approximation of the local model gradient. Therefore, where the resistivity varies smoothly, first order homogeneous Neumann boundary conditions in $\mathbf{p}$, $\mathrm{d}\mathbf{p}/\mathrm{d}\mathbf{x} = 0$, are approximately equivalent to a constant gradient in $\mathbf{m}$, $\mathrm{d}^2\mathbf{m}/\mathrm{d}\mathbf{x}^2 = 0$. This is the second order homogeneous Neumann boundary condition. Therefore, these boundary conditions encourage the continuation of gradients up to the model boundaries. This is in contrast to periodic boundary conditions which could result in step changes in resistivity wherever different geological units occur at opposite model boundaries.

The second order boundary conditions in $\mathbf{p}$ can be implemented by using so-called 'ghost points' outside the edge of the finite difference grid to extend the grid outside of the model during the finite difference calculation.

6.2.5 A half-point finite difference scheme for TGV

The asymmetry of the forward-backward difference scheme for TGV is suspected of contributing to errors in the inversion, particularly at the edges and corners of features in the model. It may also be possible that these errors are affecting the convergence of the algorithm, and contribute to the undue influence of the initial conditions on the results of the model.

To avoid these errors, a new finite difference scheme can be used. The central difference scheme, equation (2.2.4), is symmetrical, but its insensitivity to thin structures such as edges makes it poorly suited to TGV type regularisation. Instead, to retain the edge sensitivity, I recommend the use of a half-point central difference scheme, where the distance is set to half the cell spacing. I describe and derive a half point central difference scheme for TGV inversion in Appendix A. The half-point central difference scheme is more complex, and more challenging to implement, as terms cannot be represented as the product of forward-backward difference matrices. Instead, each term in the inverse solution is constructed individually.

6.2.6 Higher order TGV

The second order TGV functional, which favours piecewise affine (linear) models has been implemented in this work. The TGV framework is generalisable to higher orders, which may deliver improved performance in areas of variable gradient. This does present several challenges

to the inversion process, most notably introducing additional trade off parameters into the inversion to weight the contributions from the terms at each order. Tuning these parameters is a source of potential bias in the inversion, and the increase in complexity is multiplicative with the number of parameters. Furthermore, errors in the finite difference model may compound at higher order, and it is likely that the shortcomings of the second order TGV algorithm will remain at higher order TGV.

6.3 PCA Future Work

This proof of concept case study demonstrated that the PCA control chart method is able to detect anomalies corresponding to short-circuiting errors in both the reciprocal error and the resistivity data. However, further work is required to understand the method's general effectiveness.

The priorities of this future work should be to determine how the method performs over different resistivity distributions, resistivity change rates, noise profiles, and training lengths. Methods of identifying types of anomaly and filtering these from the time series are also desirable.

There are two main approaches to this testing. A synthetic study allows for control over the experimental parameters, and allows for the response of the system to a range of simulated errors to be understood. On the other hand, using historical ERT monitoring data is more realistic, and would be particularly useful in establishing the sensitivity of the method to changes in ground resistivity and real noise fluctuations.

A carefully designed synthetic study allows for control over the type, timing and duration of anomalies, whilst controlling for a known background noise profile. For example, short-circuiting anomalies can be simulated by replacing the effected electrodes in the forward model by a weighted grouping of the effected electrodes, whilst stochastic random noise can be added to any electrode or electrode grouping.

However, changes in the resistivity distribution also influence the control charts. This can be accounted for in any synthetic experiments. Clean synthetic data can be generated from the inverted models produced by real monitoring data, and a known noise profile can be added if desired. This allows the effects of realistic ground changes on the control charts to be isolated from the unknown noise profile of real data. Both 'normal' background noise and anomalous background errors can be added to the data in order to study their effects on the control charts.

Aside from the new PCA control chart method, a key finding of Chapter 5 is that threshold-

based error detection methods have significant limitations in the detection of low magnitude changes to the reciprocal error distribution, particularly low-magnitude errors resulting from short circuiting. More attention should be paid to changes in the reciprocal error distribution, which would enable similar large scale, low magnitude error changes to be detected and evaluated, without the need for a pre-determined threshold. Monitoring the moments (mean, variance and skewness) of the reciprocal error distribution should prove more effective at detecting low magnitude changes to the resistivity distribution, such as occurred in this case study, than a simple threshold, and are easier to use than a PCA control chart. However, unlike the PCA control charts, these methods are not sensitive to changes in the correlation structure of the measurements, and so would not have been able to detect the short circuiting directly from the resistivity in the case study.

References

Esben Auken and Anders Vest Christiansen. Layered and laterally constrained 2d inversion of resistivity data. *Geophysics*, 69(3):752–761, 2004.

Esben Auken, Anders V Christiansen, Bo H Jacobsen, Nikolaj Foged, and Kurt I Sørensen. Piecewise 1d laterally constrained inversion of resistivity data. *Geophysical Prospecting*, 53 (4):497–506, 2005.

R Barker. A simple algorithm for electrical imaging of the subsurface. *First break*, 10(2), 1992.

Andrew Binley, Abelardo Ramirez, and William Daily. Regularised image reconstruction of noisy electrical resistance tomography data. In *Proceedings of the 4th Workshop of the European Concerted Action on Process Tomography, Bergen, Norway*, pages 6–8, 1995.

R. Blaschek, A. Hordt, and A. Kamna. A new sensitivity controlled focusing regularization scheme for the inversion of induced polarisation data based on the minimum gradient support. *Geophysic*, 73(2):F45–F54, 2008.

Alistair James Scott Boyle. *Geophysical applications of electrical impedance tomography*. PhD thesis, Carleton University, 2016.

K. Bredies, K. Kunsich, and T. Pock. Total Generalised Variation. *SIAM Journal on Imaging Sciences*, 3(3):492–526, 2010.

Pascal Brunet, Rémi Clément, and Christophe Bouvier. Monitoring soil water content and deficit using electrical resistivity tomography (ert)–a case study in the cevennes area, france. *Journal of Hydrology*, 380(1-2):146–153, 2010.

J. Chambers. Geophysical monitoring of natural and engineered slopes: Towards improved early warning of landslides. 2021(1):1–7, 2021. ISSN 2214-4609. doi: https://doi.org/10.3997/2214-4609.202152175. URL `https://www.earthdoc.org/content/papers/10.3997/2214-4609.202152175`.

Tony F Chan, Jianhong Shen, and Luminita Vese. Variational pde models in image processing. *Notices AMS*, 50(1):14–26, 2003.

Varun Chandola, Arindam Banerjee, and Vipin Kumar. Anomaly detection: A survey. *ACM computing surveys (CSUR)*, 41(3):1–58, 2009.

Ilaria Coscia, Niklas Linde, Stewart Greenhalgh, Thomas Günther, and Alan Green. A filtering method to correct time-lapse 3d ert data and improve imaging of natural aquifer dynamics. *Journal of Applied Geophysics*, 80:12–24, 2012.

Torleif Dahlin. Short note on electrode charge-up effects in dc resistivity data acquisition using multi-electrode arrays. *Geophysical Prospecting*, 48(1):181–187, 2000.

William Daily, Abelardo Ramirez, Douglas LaBrecque, and John Nitao. Electrical resistivity tomography of vadose water movement. *Water Resources Research*, 28(5):1429–1442, 1992.

Bart De Ketelaere, Mia Hubert, and Eric Schmitt. Overview of pca-based statistical process-monitoring methods for time-dependent, high-dimensional data. *Journal of Quality Technology*, 47(4):318–335, 2015.

John Deceuster, Olivier Kaufmann, and Michel Van Camp. Automated identification of changes in electrode contact properties for long-term permanent ert monitoring experiments. *Geophysics*, 78(2):E79–E94, 2013.

C. deGroot Hedlin and S. Constable. Occam's inversion to generate smooth, two-dimensional models from magnetotelluric data. *Geophysics*, 55(12):1613–1624, 1990.

A Dey and H Frank Morrison. Resistivity modelling for arbitrarily shaped three-dimensional structures. *Geophysics*, 44(4):753–780, 1979.

Joseph Doetsch, Niklas Linde, Mirco Pessognelli, Alan G Green, and Thomas Günther. Constraining 3-d electrical resistance tomography with gpr reflection data for improved aquifer characterization. *Journal of Applied Geophysics*, 78:68–76, 2012a.

Joseph Doetsch, Niklas Linde, Tobias Vogt, Andrew Binley, and Alan G Green. Imaging and quantifying salt-tracer transport in a riparian groundwater system by means of 3d ert monitoring. *Geophysics*, 77(5):B207–B218, 2012b.

Jinming Duan, Wenqi Lu, Christopher Tench, Irene Gottlob, Frank Proudlock, Niraj Nilesh Samani, and Li Bai. Denoising optical coherence tomography using second order total generalized variation decomposition. *Biomedical Signal Processing and Control*, 24:120–127, 2016.

C.G. Farquharson and D.W. Oldenburg. Non-linear inversion using general measures of data misfit and model structure. *Geophysical Journal International*, 134:213–227, 1998.

G. Fiandaca, J. Doetsch, G. Vignoli, and E. Auken. Generalized focusing of time-lapse changes

with applications to direct current and time-domain induced polarization inversions. *Geophysical Journal International*, 203:1101–1112, 2015.

Pascal Getreuer. Rudin-osher-fatemi total variation denoising using split bregman. *Image Processing On Line*, 2:74–95, 2012.

L. Hawley-Sibbett. Tgvinv: an adaption of ip4di for total generalised variation inversion. `https://github.com/Hawley-SibbettL/TGVinv`, 2023.

Thomas Hermans, Frédéric Nguyen, Tanguy Robert, and Andre Revil. Geophysical methods for monitoring temperature changes in shallow low enthalpy geothermal systems. *Energies*, 7(8):5083–5118, 2014.

J. Edward Jackson. Quality control methods for several related variables. *Technometrics*, 1(4): 359–377, 1959. doi: 10.1080/00401706.1959.10489868.

J Edward Jackson and Govind S Mudholkar. Control procedures for residuals associated with principal component analysis. *Technometrics*, 21(3):341–349, 1979.

Timothy C Johnson, Lee D Slater, Dimitrios Ntarlagiannis, Frederick D Day-Lewis, and Mehrez Elwaseif. Monitoring groundwater-surface water interaction using time-series and time-frequency analysis of transient three-dimensional electrical resistivity changes. *Water Resources Research*, 48(7), 2012.

Ian Jolliffe. *Principal Component Analysis*, pages 1094–1096. Springer Berlin Heidelberg, Berlin, Heidelberg, 2 edition, 2002. ISBN 978-3-642-04898-2. doi: 10.1007/978-3-642-04898-2_455. URL `https://doi.org/10.1007/978-3-642-04898-2_455`.

M Karaoulis, A Revil, P Tsourlos, DD Werkema, and BJ Minsley. Ip4di: A software for time-lapse 2d/3d dc-resistivity and induced polarization tomography. *Computers & Geosciences*, 54:164–170, 2013.

J-H Kim, M-J Yi, Park S-G, and Kim J.G. 4-D inversion of DC resistivity monitoring data aquired over a dynamically changing earth model. *Journal of Applied Geophysics*, 68:522–532, 2009.

J-H Kim, R. Supper, P. Tsourlos, and M-J Yi. Four-dimensional inversion of resistivivty monitoring data through L*p* norm minimizations. *Geophysical journal international*, 195, 2013.

F. Knoll, K. Bredies, T. Pock, and R. Stollberger. Second order Total Generalised Variation (TGV) for MRI. *Magnetic Resonance in Medicine*, 65(2):480–491, 2011.

Johannes Koestel, Andreas Kemna, Mathieu Javaux, Andrew Binley, and Harry Vereecken.

Quantitative imaging of solute transport in an unsaturated and undisturbed soil monolith with 3-d ert and tdr. *Water Resources Research*, 44(12), 2008.

Oliver Kuras, Jonathan D. Pritchard, Philip I. Meldrum, Jonathan E. Chambers, Paul B. Wilkinson, Richard D. Ogilvy, and Gary P. Wealthall. Monitoring hydraulic processes with automated time-lapse electrical resistivity tomography (alert). *Comptes Rendus Geoscience*, 341 (10):868–885, 2009.

Oliver Kuras, Paul B Wilkinson, Philip I Meldrum, Lucy S Oxby, Sebastian Uhlemann, Jonathan E Chambers, Andrew Binley, James Graham, Nicholas T Smith, and Nick Atherton. Geoelectrical monitoring of simulated subsurface leakage to support high-hazard nuclear decommissioning at the sellafield site, uk. *Science of the Total Environment*, 566:350–359, 2016.

Douglas LaBrecque and William Daily. Assessment of measurement errors for galvanic-resistivity electrodes of different composition. *Geophysics*, 73(2):F55–F64, 2008.

Douglas J LaBrecque and Xianjin Yang. Difference inversion of ert data: A fast inversion method for 3-d in situ monitoring. *Journal of Environmental & Engineering Geophysics*, 6(2): 83–89, 2001.

Douglas J LaBrecque, Michela Miletto, William Daily, Aberlardo Ramirez, and Earle Owen. The effects of noise on occam's inversion of resistivity tomography data. *Geophysics*, 61(2): 538–548, 1996.

Nolwenn Lesparre, Frederic Nguyen, Andreas Kemna, Tanguy Robert, Thomas Hermans, Moubarak Daoudi, and Adrian Flores-Orozco. A new approach for time-lapse data weighting in electrical resistivity tomography. *Geophysics*, 82(6):E325–E333, 2017.

Meng Heng Loke and RD Barker. Rapid least-squares inversion of apparent resistivity pseudosections by a quasi-newton method. *Geophysical prospecting*, 44(1):131–152, 1996.

Meng Heng Loke, Ian Acworth, and Torleif Dahlin. A comparison of smooth and blocky inversion methods in 2d electrical imaging surveys. *Exploration geophysics*, 34(3):182–187, 2003.

M.H. Loke, J.E Chambers, D.F. Rucker, O. Kuras, and P.B. Wilkinson. Recent developments in the direct-current geoelectrical imaging method. *Journal of applied Geophysics*, 95:135–156, 2013.

MH Loke, Torleif Dahlin, and DF Rucker. Smoothness-constrained time-lapse inversion of data from 3d resistivity surveys. *Near Surface Geophysics*, 12(1):5–24, 2014.

Wenqi Lu, Jinming Duan, Zhaowen Qiu, Zhenkuan Pan, Ryan Wen Liu, and Li Bai. Imple-

mentation of high-order variational models made easy for image processing. *Mathematical Methods in the Applied Sciences*, 39(14):4208–4233, 2016.

LW. Menke. *Geophysical Data Analysis: Discrete Inverse Theory*, page 289. Academic Press Inc., 2 edition, 1989.

Carlyle R Miller, Partha S Routh, Troy R Brosten, and James P McNamara. Application of time-lapse ert imaging to watershed characterization. *Geophysics*, 73(3):G7–G17, 2008.

Michael A Mitchell and Douglas W Oldenburg. Data quality control methodologies for large, non-conventional dc resistivity datasets. *Journal of Applied Geophysics*, 135:163–182, 2016.

Frederic Nguyen, Andreas Kemna, Tanguy Robert, and Thomas Hermans. Data-driven selection of the minimum-gradient support parameter in time-lapse focused electric imaging. *Geophysics*, 81(1):A1–A5, 2016.

Shanzhou Niu, Yang Gao, Zhaoying Bian, Jing Huang, Wufan Chen, Gaohang Yu, Zhengrong Liang, and Jianhua Ma. Sparse-view x-ray ct reconstruction via total generalized variation regularization. *Physics in Medicine & Biology*, 59(12):2997, 2014.

Dattatray Parasnis. Reciprocity theorems in geoelectric and geoelectromagnetic work. *Geoexploration*, 25(3):177–198, 1988.

Angela Perrone, Sabatino Piscitelli, and Vincenzo Lapenna. Electrical resistivity tomographies for landslide monitoring: a review. *Geoelectric Monitoring*, page 129, 2012.

William H Press, Saul A Teukolsky, William T Vetterling, and Brian P Flannery. *Numerical recipes 3rd edition: The art of scientific computing*, chapter 10.2. Cambridge university press, 3 edition, 2007.

Paul Rodriguez and Brendt Wohlberg. An iteratively reweighted norm algorithm for total variation regularization. In *Signals, Systems and Computers, 2006. ACSSC'06. Fortieth Asilomar Conference on*, pages 892–896. IEEE, 2006.

Carsten Rücker. *Advanced electrical resistivity modelling and inversion using unstructured discretization*. PhD thesis, Universität Leipzig, 2010.

Leonid I Rudin, Stanley Osher, and Emad Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60(1-4):259–268, 1992.

Cornelia Schmidt-Hattenberger, P Bergmann, Tim Labitzke, and F Wagner. Co2 migration monitoring by means of electrical resistivity tomography (ert)–review on five years of operation of a permanent ert system at the ketzin pilot site. *Energy Procedia*, 63:4366–4373, 2014.

Xizi Song, Yanbin Xu, and Feng Dong. A hybrid regularization method combining tikhonov

with total variation for electrical resistance tomography. *Flow Measurement and Instrumentation*, 46:268–275, 2015a.

Xizi Song, Yanbin Xu, and Feng Dong. A spatially adaptive total variation regularization method for electrical resistance tomography. *Measurement Science and Technology*, 26(12): 125401, 2015b.

Chak-Hau Michael Tso, Oliver Kuras, Paul B Wilkinson, Sebastian Uhlemann, Jonathan E Chambers, Philip I Meldrum, James Graham, Emma F Sherlock, and Andrew Binley. Improved characterisation and modelling of measurement errors in electrical resistivity tomography (ert) surveys. *Journal of Applied Geophysics*, 146:103–119, 2017.

Panagiotis Tsourlos. *Modelling, interpretation and inversion of multielectrode resistivity survey data.* PhD thesis, University of York, 1995.

Sebastian Uhlemann, Jonathan Chambers, Paul Wilkinson, Hansruedi Maurer, Andrew Merritt, Philip Meldrum, Oliver Kuras, David Gunn, Alister Smith, and Tom Dijkstra. Four-dimensional imaging of moisture dynamics during landslide reactivation. *Journal of Geophysical Research: Earth Surface*, 122(1):398–418, 2017.

Sean Eugene Walker. *Inversion of em data to recover 1-d conductivity and geometric survey parameter*. PhD thesis, University of British Columbia, 1999.

Paul Wilkinson, Meng Loke, Phil Meldrum, Jonathan Chambers, Oliver Kuras, D.A. Gunn, and Richard Ogilvy. Practical aspects of applied optimized survey design for electrical resistivity tomography. *Geophysical Journal International*, 189:428–440, 04 2012. doi: 10.1111/j.1365-246X.2012.05372.x.

Paul B Wilkinson, Jonathan E Chambers, Philip I Meldrum, David A Gunn, Richard D Ogilvy, and Oliver Kuras. Predicting the movements of permanently installed electrodes on an active landslide using time-lapse geoelectrical resistivity data only. *Geophysical Journal International*, 183(2):543–556, 2010.

Myeong-Jong Yi, Jung-Ho Kim, and Seung-Hwan Chung. Enhancing the resolving power of least-squares inversion with active constraint balancing. *Geophysics*, 68(3):931–941, 2003.

Jieyi Zhou, Andre Revil, Marios Karaoulis, Dave Hale, J Doetsch, and S Cuttler. Image-guided inversion of electrical resistivity data. *Geophysical Journal International*, 197(1):292–309, 2014.

Zhou Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, 2004.

Appendices

Appendix A

Derivation of a Half Point Difference Scheme for TGV

In this section, I derive an alternative, symmetrical, finite difference scheme, which has the potential to avoid these artefacts: the half point central difference. This derivation may form the basis of future work on the TGV regularisation of ERT data. The half point central difference scheme is a variation of the central difference scheme (equation 2.2.4) where the distance is halved, so that the function, f, is evaluated at the cell edges and corners. In x, this is

$$f'(x_0) = \frac{f(x_0 + \frac{l}{2}) - f(x_0 - \frac{l}{2})}{l}. \tag{A.0.1}$$

This scheme combines the advantages of symmetry (in both $\mathbf{m}$ and $\mathbf{p}$) and accuracy of the central difference scheme with an increased sensitivity to thin features. The cost is increased complexity; interpolation is needed to evaluate the function at edge points, and the solution can no longer be represented by simple difference matrices with a diagonal IRLS weighting matrix. A half point scheme was chosen over a higher order approximation for f', as these require stencils involving $x_0 \pm 2l$ and beyond to approximate the first derivative, and this creates undesirable complexities around boundary handling and higher order derivatives (including of $\mathbf{p}$).

Here, I derive the half point difference scheme for the TGV regularised ERT inversion. The solution reduces to the TV solution if $\mathbf{p} = 0$, and to the l_2 solution if the IRLS denominators are removed. An isotropic half point gradient operator is used, $\nabla^h = \nabla^h_x + \nabla^h_z$, where ∇^h_x and ∇^h_z are the half point central difference operators along x and z. In order to simplify the algebra without loss of generality, $l = 1$ hereafter.

A.0.1 Half-difference component blocks

In a cell-centred finite difference scheme, the derivatives and their constituent terms all map to the cell centres. In the half difference scheme, the model values at edges and corners are also needed. In a simple interpolation, edge values are approximated as the mean of the two adjacent model cells, while corner values are approximated by the mean of the four surrounding cells.

Examples of central half differences evaluated at a cell centre and edge are shown in figure A.1.

Symmetry can be used to find the derivatives at the other edges and at different orientations, which are used in the following derivation.

A.0.2 Subproblem in m

I approach the TGV regularised inversion as an alternating-minimisation problem in $\mathbf{m}$ and $\mathbf{p}$, as in Section 3.2. First the subproblem in $\mathbf{m}$ is derived. The half-point difference TV result can be obtained by setting $\mathbf{p} = \mathbf{0}$ in the following solution. The l_2 solution can be obtained by further setting all $\mathbf{r}^{(m)}$ terms to 1.

The data misfit term is unchanged, so I only consider the regularisation terms here. The continuous representation of the TGV regularisation term (see equation((2.1.3)) is,

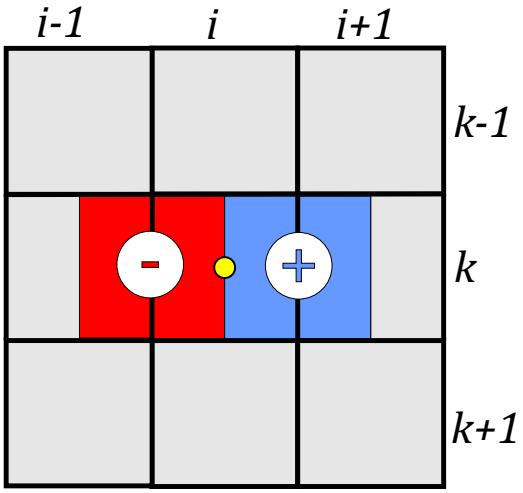
$$\mathbf{\Phi}^h_{\mathrm{TGV}}(\mathbf{m}, \mathbf{p}) = \int_x \int_z \frac{(\nabla^h \mathbf{m} - \mathbf{p})^2}{|\nabla^h \mathbf{m} - \mathbf{p}|} \mathrm{d}x\mathrm{d}z + \mu \int_x \int_z \frac{(\nabla^h \mathbf{p} + \nabla^h \mathbf{p}^T)^2}{|\nabla^h \mathbf{p} + \nabla^h \mathbf{p}^T|} \mathrm{d}x\mathrm{d}z \tag{A.0.2}$$

where $|\cdot|$ denotes the absolute value, used in the IRLS approximation of the l_1 norm as a square interval. The regularisation function is minimised with regard to $\mathbf{m}$,

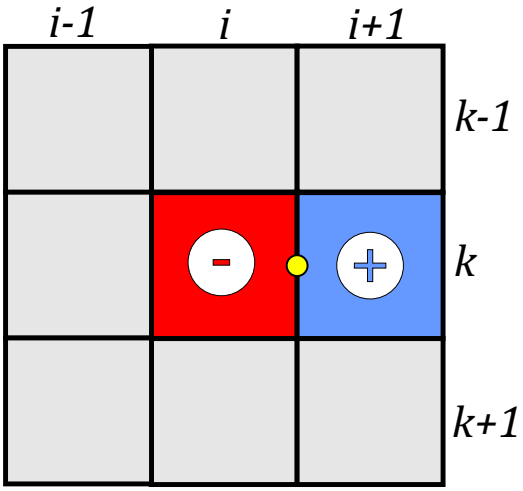
$$0 = \int_x \int_z \nabla^h_x \cdot \left(\frac{(\nabla^h_x \mathbf{m} - p_x)}{|\nabla^h \mathbf{m} - \mathbf{p}|} \right) \mathrm{d}x\mathrm{d}z + \int_x \int_z \nabla^h_x \cdot \left(\frac{(\nabla^h_z \mathbf{m} - p_z)}{|\nabla^h \mathbf{m} - \mathbf{p}|} \right) \mathrm{d}x\mathrm{d}z, \tag{A.0.3}$$

Discretising and solving for the x term for the model cell at index (i, k),

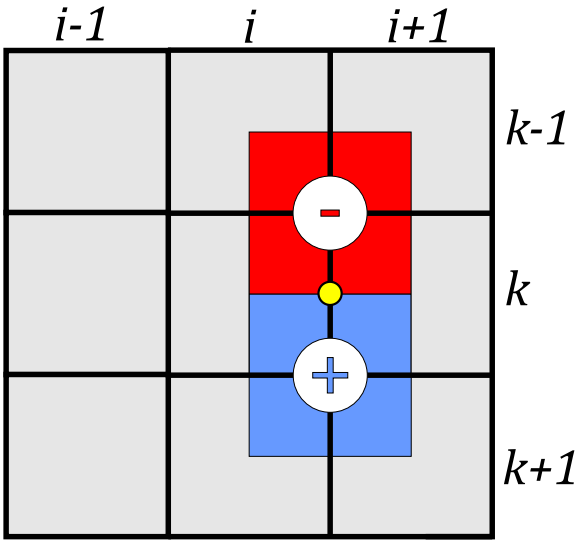
$$\begin{aligned} \nabla_x \cdot \left(\frac{(\nabla^h_x \mathbf{m} - p_x)}{|\nabla^h \mathbf{m} - \mathbf{p}|} \right)_{i,k} &= \left(\frac{\nabla^h_x \mathbf{m} - p_x}{|\nabla^h \mathbf{m} - \mathbf{p}|} \right)_{i+\frac{1}{2},k} - \left(\frac{\nabla^h_x \mathbf{m} - p_x}{|\nabla^h \mathbf{m} - \mathbf{p}|} \right)_{i-\frac{1}{2},k} \\ &= \frac{\nabla^+_x m_{i,k} - 0.5\big((p_x)_{i,k} + (p_x)_{i+1,k}\big)}{r^{(m)}_{i+\frac{1}{2},k}} - \frac{\nabla^-_x m_{i,k} - 0.5\big((p_x)_{i,k} + (p_x)_{i-1,k}\big)}{r^{(m)}_{i-\frac{1}{2},k}} \end{aligned} \tag{A.0.4}$$



(a) $(\nabla_x^h \mathbf{m})_{i,j} = (m_{i+\frac{1}{2},k} - m_{i-\frac{1}{2},k})/2$



(b) $(\nabla_x^h \mathbf{m})_{i+\frac{1}{2},j} = \nabla_x^+ m_{i,k}$
$= m_{i+1,k} - m_{i,k}$



(c) $(\nabla_y^h \mathbf{m})_{i+\frac{1}{2},j} =$
$(m_{i+1,k+1} + m_{i,k+1} - m_{i+1,k-1} - m_{i,k-1})/4$

Figure A.1: Examples of half-point finite differences at the cell centre (a), across an edge (b) and perpendicular to the edge (c). The yellow dot marks the position of the derivative. The positive and negative half points are marked by the blue (+) and red (-) symbols respectively. The coloured regions around the half points are a geometric representation of the interpolation used. Each model cell contributes to the derivative in exact proportion to its fractional shaded area. The colour denotes the sign.

where ∇_x^+ and ∇_x^- are respectively the forward and backward difference operators in x. The IRLS terms, $\mathbf{r}^m$, are evaluated on the midpoints of the vertical edges at $i \pm \frac{1}{2}$,

$$r^{(m)}_{i+\frac{1}{2},k} = \sqrt{\left(\nabla_x^+ m_{i,k} - (p_x)_{i,k}\right)^2 + \left(\nabla_z^{||} m_{i,k} - \frac{(p_z)_{i,k} + (p_z)_{i+1,k}}{2}\right)^2 + \gamma_m^2} \tag{A.0.5}$$

$$r^{(m)}_{i-\frac{1}{2},k} = \sqrt{\left(\nabla_x^- m_{i,k} - (p_x)_{i,k}\right)^2 + \left(\nabla_z^{|} m_{i,k} - \frac{(p_z)_{i,k} + (p_z)_{i-1,k}}{2}\right)^2 + \gamma_m^2} \tag{A.0.6}$$

where γ_m is a small constant to avoid singularities. $\nabla_x^{||} m_{i,k} = (\nabla_x^h \mathbf{m})_{i,k+\frac{1}{2}}$ and $\nabla_x^{|} m_{i,k} = (\nabla_x^h \mathbf{m})_{i,k-\frac{1}{2}}$ are the half point differences in x evaluated on the vertical edges. The corresponding $\nabla_z - p_z$ terms can be found from symmetry. A full, detailed, set of terms and definitions can be found in appendix B. Substituting these results, into the full ERT objective function, the IRLS Gauss-Newton update for $\mathbf{m}$ is found to be

$$m^{(n+1)} = \mathbf{m}^{(n)} + \left[\mathbf{J}^{(n)T}\mathbf{R}_d\mathbf{J}^{(n)} + \lambda^{(n)}\mathbf{b}(\mathbf{m}^{(n)})\right]^{-1}\left[\mathbf{J}^{(n)T}\mathbf{R}_d\mathbf{W}_d(\mathbf{d}^{(n)} - \mathbf{d}) - \lambda^{(n)}(\mathbf{b}(\mathbf{m}^{(n)}) + \mathbf{c}(\mathbf{p}))\right], \tag{A.0.7}$$

where the (second order) model gradient terms are

$$\mathbf{b}(m_{i,k}) = \frac{\nabla_x^+ m_{i,k}}{r^{(m)}_{i+\frac{1}{2},k}} + \frac{\nabla_y^+ m_{i,k}}{r^{(m)}_{i,k+\frac{1}{2}}} - \frac{\nabla_x^- m_{i,k}}{r^{(m)}_{i-\frac{1}{2},k}} - \frac{\nabla_y^- m_{i,k}}{r^{(m)}_{i,k-\frac{1}{2}}} \tag{A.0.8}$$

and the half point central difference terms in $\mathbf{p}$ are

$$\mathbf{c}(p_{i,k}) = -\frac{(p_x)_{i,k} + (p_x)_{i+1,k}}{2r^{(m)}_{i+\frac{1}{2},k}} - \frac{(p_z)_{i,k} + (p_z)_{i,k+1}}{2r^{(m)}_{i,k+\frac{1}{2}}} + \frac{(p_x)_{i,k} + (p_x)_{i-1,k}}{2r^{(m)}_{i-\frac{1}{2},k}} + \frac{(p_z)_{i,k} + (p_z)_{i,k-1}}{2r^{(m)}_{i,k-\frac{1}{2}}} \tag{A.0.9}$$

In this formulation, the IRLS and gradient terms are symmetrical, equally weighting contributions from all the surrounding cells, unlike $(\mathbf{R}_m)_{i,k}$ in the forward-backward scheme, which utilises the forward difference only.

A.0.3 Subproblem in p

As in Section 3.2, an IRLS method is used to solve for $\mathbf{p}$ over the current model $\mathbf{m}^{(n)}$ by iteratively linearising the objective function about $\mathbf{p}$. In continuous form, the minimiser of the TGV functional (A.0.2) about $\mathbf{p_x}$ is

$$0 = -\int_{x,y}\left(\frac{(\nabla_x^h\mathbf{m}-\mathbf{p_x})}{|\nabla^h\mathbf{m}-\mathbf{p}|}\right)dxdz+\mu\int_{x,y}\nabla_x^h\cdot\left(\frac{\nabla_x^h\mathbf{p}_x}{|\nabla^h\mathbf{p}+\nabla^h\mathbf{p}^T|}\right)dxdz+\frac{\mu}{2}\int_{x,y}\nabla_y^h\cdot\left(\frac{\nabla_x^h\mathbf{p}_y+\nabla_y^h\mathbf{p}_x}{|\nabla^h\mathbf{p}+\nabla^h\mathbf{p}^T|}\right)dxdz$$

Each term is discretised for the model cell at (i,k):

$$\left(\frac{(\nabla_x^h\mathbf{m}-p_x)}{|\nabla^h\mathbf{m}-\mathbf{p}|}\right)_{i,k}=\frac{\nabla_x^h m_{i,k}-(p_x)_{i,k}}{r_{i,k}^{(m)}} \tag{A.0.10}$$

$$\int_{x,y}\nabla_x^h\cdot\left(\frac{\nabla_x^h\mathbf{p}_x}{|\nabla^h\mathbf{p}+\nabla^h\mathbf{p}^T|}\right)_{i,k}dxdz=\frac{\nabla_x^+(p_x)_{i,k}}{r_{i+\frac{1}{2},k}^{(p)}}-\frac{\nabla_x^-(p_x)_{i,k}}{r_{i-\frac{1}{2},k}^{(p)}} \tag{A.0.11}$$

$$\int_{x,y}\nabla_y^h\cdot\left(\frac{\nabla_y^h\mathbf{p}_x+\nabla_x^h\mathbf{p}_z}{|\nabla^h\mathbf{p}+\nabla^h\mathbf{p}^T|}\right)_{i,k}dxdz=\frac{\nabla_z^+(p_x)_{i,k}+\nabla_x^{||}(p_z)_{i,k}}{r_{i,k+\frac{1}{2}}^{(p)}}-\frac{\nabla_z^-(p_x)_{i,k}+\nabla_x^{|}(p_z)_{i,k}}{r_{i,k-\frac{1}{2}}^{(p)}} \tag{A.0.12}$$

The definitions of the IRLS terms $\mathbf{r}^{(m)}$ and $\mathbf{r}^{(p)}$, as well as the equivalent results for $\mathbf{p}_z$ can be found in Appendix B. Substituting these results back into the minimiser of $\mathbf{p}$ sets up the following system of equations, to be iteratively solved for $\mathbf{p}$,

$$\begin{aligned}
\frac{\nabla_x^h m_{i,k}}{r_{i,k}^{(m)}} &= \left(\frac{1}{r_{i,k}^{(m)}}+\frac{\mu\nabla_x^+}{r_{i+\frac{1}{2},k}^{(p)}}-\frac{\mu\nabla_x^-}{r_{i-\frac{1}{2},k}^{(p)}}+\frac{\nabla_z^+}{r_{i,k+\frac{1}{2}}^{(p)}}-\frac{\nabla_z^-}{r_{i,k-\frac{1}{2}}^{(p)}}\right)(p_x)_{i,k}+\frac{\mu}{2}\left(\frac{\nabla_x^{||}}{r_{i,k+\frac{1}{2}}^{(p)}}-\frac{\nabla_x^{|}}{r_{i,k-\frac{1}{2}}^{(p)}}\right)(p_z)_{i,k}\\
\frac{\nabla_z^h m_{i,k}}{r_{i,k}^{(m)}} &= \frac{\mu}{2}\left(\frac{\nabla_z^{||}}{r_{i+\frac{1}{2}}^{(p)},k}-\frac{\nabla_z^{|}}{r_{i-\frac{1}{2}}^{(p)},k}\right)(p_x)_{i,k}+\left(\frac{1}{r_{i,k}^{(m)}}+\frac{\mu\nabla_z^+}{r_{i,k+\frac{1}{2}}^{(p)}}-\frac{\mu\nabla_z^-}{r_{i,k-\frac{1}{2}}^{(p)}}+\frac{\nabla_x^+}{r_{i+\frac{1}{2},k}^{(p)}}-\frac{\nabla_x^-}{r_{i-\frac{1}{2},k}^{(p)}}\right)(p_z)_{i,k}.
\end{aligned} \tag{A.0.13}$$

In equation (3.2.9), the analogous forward-backward solution to (A.0.13), there is significant asymmetry, not only in the IRLS $\mathbf{R}_m$ and $\mathbf{R}_p$ terms, but also in the cross terms $\mathbf{C}_z\mathbf{R}_p\mathbf{C}_x^T\mathbf{p}_x$ and $\mathbf{C}_x\mathbf{R}_p\mathbf{C}_z^T\mathbf{p}_z$, which each feature only single sided derivatives in x and z. By contrast, equations (A.0.13) is fully symmetrical, thereby eliminating asymmetry as a possible source of error at the domain boundary or corner features.

A.0.4 Boundary conditions in the half point difference scheme

Unlike the forward-backward difference model, where the terms of the inverse solution are formed from the products of the forward-backward matrices, and the boundary conditions are enforced implicitly, the half point solutions are assembled directly from the explicit terms of the solution. Boundary conditions can be applied directly, by substituting the appropriate condition whenever a cell outside the model appears in the difference equations.

Appendix B

Half-point central differences: Full definitions and terms

The definitions for the terms used in the half-point central difference TGV inversion algorithm in Appendix A are:

$$\nabla_x^+ m_{i,k} = (\nabla_x^h \mathbf{m})_{i+\frac{1}{2},k} = m_{i+1,k} - m_{i,k} \tag{B.0.1}$$

$$\nabla_x^- m_{i,k} = (\nabla_x^h \mathbf{m})_{i-\frac{1}{2}c,k} = m_{i,k} - m_{i-1,k} \tag{B.0.2}$$

$$\nabla_z^+ m_{i,k} = (\nabla_z^h \mathbf{m})_{i,k+\frac{1}{2}} = m_{i,k+1} - m_{i,k} \tag{B.0.3}$$

$$\nabla_z^- m_{i,k} = (\nabla_z^h \mathbf{m})_{i,k-\frac{1}{2}} = m_{i,k} - m_{i,k-1}. \tag{B.0.4}$$

$$\nabla_x^h m_{i,k} = \frac{m_{i+1,k} - m_{i-1,k}}{2} \tag{B.0.5}$$

$$\nabla_z^h m_{i,k} = \frac{m_{i,k+1} - m_{i,k-1}}{2} \tag{B.0.6}$$

$$\nabla_x^{||} m_{i,k} = (\nabla_x^h \mathbf{m})_{i,k+\frac{1}{2}} = \frac{m_{i+1,k+1} + m_{i+1,k} - m_{i-1,k+1} - m_{i-1,k}}{4} \tag{B.0.7}$$

$$\nabla_x^{|} m_{i,k} = (\nabla_x^h \mathbf{m})_{i,k-\frac{1}{2}} = \frac{m_{i+1,k-1} + m_{i+1,k} - m_{i-1,k-1} - m_{i-1,k}}{4} \tag{B.0.8}$$

$$\nabla_z^{||} m_{i+\frac{1}{2},k} = (\nabla_z^h \mathbf{m})_{i+\frac{1}{2},k} = \frac{m_{i+1,k+1} + m_{i,k-1} - m_{i+1,k-1} - m_{i,k-1}}{4} \tag{B.0.9}$$

$$\nabla_z^{|} m_{i+\frac{1}{2},k} = (\nabla_z^h \mathbf{m})_{i-\frac{1}{2},k} = \frac{m_{i-1,k+1} + m_{i,k-1} - m_{i-1,k-1} - m_{i,k-1}}{4} \tag{B.0.10}$$

$$r_{i,k}^{(m)} = \sqrt{\left(\nabla_x^h m_{i,k} - (p_x)_{i,k}\right)^2 + \left(\nabla_z^h m_{i,k} - (p_z)_{i,k}\right)^2 + \gamma_m^2} \tag{B.0.11}$$

$$r_{i+\frac{1}{2},k}^{(m)} = \sqrt{\left(\nabla_x^+ m_{i,k} - (p_x)_{i,k}\right)^2 + \left(\nabla_z^{||} m_{i,k} - \frac{(p_z)_{i,k} + (p_z)_{i+1,k}}{2}\right)^2 + \gamma_m^2} \tag{B.0.12}$$

$$r_{i-\frac{1}{2},k}^{(m)} = \sqrt{\left(\nabla_x^- m_{i,k} - (p_x)_{i,k}\right)^2 + \left(\nabla_z^{|} m_{i,k} - \frac{(p_z)_{i,k} + (p_z)_{i-1,k}}{2}\right)^2 + \gamma_m^2} \tag{B.0.13}$$

$$r_{i,k+\frac{1}{2}}^{(m)} = \sqrt{\left(\nabla_x^{||} m_{i,k} - \frac{(p_x)_{i,k} + (p_x)_{i,k+1}}{2}\right)^2 + \left(\nabla_z^+ m_{i,k} - (p_z)_{i,k}\right)^2 + \gamma_m^2} \tag{B.0.14}$$

$$r_{i,k-\frac{1}{2}}^{(m)} = \sqrt{\left(\nabla_x^{|} m_{i,k} - \frac{(p_x)_{i,k} + (p_x)_{i,k-1}}{2}\right)^2 + \left(\nabla_z^- m_{i,k} - (p_z)_{i,k}\right)^2 + \gamma_m^2} \tag{B.0.15}$$

$$r_{i+\frac{1}{2},k}^{(p)} = \sqrt{\left(\nabla_x^+ (p_x)_{i,k}\right)^2 + \left(\nabla_z^{||} (p_z)_{i,k}\right)^2 + \left(\frac{\nabla_x^+ (p_z)_{i,k} + \nabla_z^{||} (p_x)_{i,k}}{2}\right)^2 + \gamma_p^2} \tag{B.0.16}$$

$$r_{i-\frac{1}{2},k}^{(p)} = \sqrt{\left(\nabla_x^- (p_x)_{i,k}\right)^2 + \left(\nabla_z^{|} (p_z)_{i,k}\right)^2 + \left(\frac{\nabla_x^- (p_z)_{i,k} + \nabla_z^{|} (p_x)_{i,k}}{2}\right)^2 + \gamma_p^2} \tag{B.0.17}$$

$$r_{i,k+\frac{1}{2}}^{(p)} = \sqrt{\left(\nabla_x^{||} (p_x)_{i,k}\right)^2 + \left(\nabla_z^+ (p_z)_{i,k}\right)^2 + \left(\frac{\nabla_x^{||} (p_z)_{i,k} + \nabla_z^+ (p_x)_{i,k}}{2}\right)^2 + \gamma_p^2} \tag{B.0.18}$$

$$r_{i,k-\frac{1}{2}}^{(p)} = \sqrt{\left(\nabla_x^{|} (p_x)_{i,k}\right)^2 + \left(\nabla_z^- (p_z)_{i,k}\right)^2 + \left(\frac{\nabla_x^{|} (p_z)_{i,k} + \nabla_z^- (p_x)_{i,k}}{2}\right)^2 + \gamma_p^2} \tag{B.0.19}$$