



Adaptive Regularization of the Reference Model in an Inverse Problem

MEIJIAN AN¹

Abstract—The solution to an inverse problem is often resolved by inverting the perturbation to a reference model of physical parameters and using regularizations. However, the most commonly used higher-order Tikhonov regularizations, which are unrelated to the reference model, are generally unable to correct false variations in the reference model, since these regularizations tend to minimize the model variations in the inverted perturbations. A viable approach to overcome this shortcoming is to adapt the regularization for the reference model such that a sharp variation around a given position in the reference model (regardless of whether the variation is true or false) receives a smaller weighting in the regularizations. Linear and nonlinear inversion tests show that this new adaptive regularization can improve the inversion results at or around positions with either badly constructed or true variations in the reference model.

Keywords: Tikhonov regularization, linear or linearized inversion, tomography, adaptive regularization.

1. Introduction

Inverse problems can be resolved via two different inversion schemes: Scheme 1, with inversions to determine the best-fit model of physical parameters, and Scheme 2, with inversions investigating perturbations relative to a reference model. Scheme 2 is termed the *creeping* strategy (Shaw and Orcutt 1985; Parker 1994) because the inverted solution is the model perturbation, and the model solution is the summation of the reference model and inverted perturbation. This scheme is effectively driven by the fact that it is easier to *discover truth by building on previous discoveries* (Keith et al. 2016) or by *standing on the shoulders of giants* (I. Newton, Letter to R.

Hooke, 5 Feb. 1676). All previous discoveries can be used to construct the reference model, such that Scheme 2 is widely applied in practical studies. A reference model allows many nonlinear inverse problems to be resolved via linearization using the first-order Taylor series approximation (e.g., Backus and Gilbert 1967, 1968; Aster et al. 2005). Many modern physical and geophysical tomographic studies, such as electrical resistance (e.g., Xu et al. 2016), impedance (e.g., Vauhkonen et al. 1998; Babaeizadeh and Brooks 2007), optical emission (e.g., Darne et al. 2013), and seismic body-wave (Nolet 2008) studies, define the reference model as a crucial and necessary component of the inversion process.

Both inversion schemes should yield the same model solution if the inverse problem is linear and well-posed. However, most practical inverse problems (e.g., Vauhkonen et al. 1998; Aster et al. 2005; Xu et al. 2016) are often ill-posed, and ad hoc artificial regularization (e.g., Tikhonov 1963; Engl et al. 2000) is required to stabilize the inversion and reduce bias in the solution. The above two schemes can yield different model solutions when regularization is applied. The model solution obtained via Scheme 2 (inversion for perturbations relative to a reference model) may contain strange or false anomalies; however, the appearance of these anomalies due to the inversion process has not been comprehensively investigated to date.

Here, the two inversion schemes using the most commonly applied Tikhonov regularizations are reviewed, the effects of Tikhonov regularizations on the model solutions for the two inversion schemes are described, and an adaptation of the Tikhonov regularization to the reference model is proposed to suppress false anomalies in the model solutions. Examples of one-dimensional (1D) linear and nonlinear (linearized) inversions are provided to illustrate

Electronic supplementary material The online version of this article (doi:<https://doi.org/10.1007/s00024-020-02530-z>) contains supplementary material, which is available to authorized users.

¹ Chinese Academy of Geological Sciences, Beijing 100037, China. E-mail: meijianan@live.com; meijianan@cags.ac.cn

how adaptive regularization can improve inversion results. This new adaptive regularization is also applicable to two-dimensional (2D) and three-dimensional (3D) inverse problems.

2. Linear or Linearized Inversion

A series of observations \mathbf{d} can be expressed simply in nonlinear form as a function g of a physical model \mathbf{m} :

$$\mathbf{d} = g(\mathbf{m}) \quad (1)$$

or in linear form as:

$$\mathbf{d} = \mathbf{G}\mathbf{m} \quad (2)$$

where the operator g is replaced by a matrix \mathbf{G} .

The least-squares or minimum-norm solution \mathbf{m} of Eq. (2) can be determined by minimizing:

$$\|\mathbf{d} - \mathbf{G}\mathbf{m}\|^2 \quad (3)$$

However, regularization (e.g., Levenberg 1944; Tikhonov 1963; Engl et al. 2000) is necessary to stabilize the inversion and reduce bias in the solution when the inversion is ill-posed. The model solution \mathbf{m} is then obtained by minimizing (e.g., Golub et al. 1999; Scales et al. 2001):

$$\|\mathbf{d} - \mathbf{G}\mathbf{m}\|^2 + \lambda^2 \|\mathbf{L}\mathbf{m}\|^2 \quad (4)$$

where \mathbf{L} is an additional transformation to \mathbf{m} and λ is the regularization parameter. Equation (4) represents the inversion equation for the first scheme, which does not incorporate a reference model.

The inclusion of a reference model \mathbf{m}_{ref} in the inversion for \mathbf{m} involves modifying Eqs. (1) and (2) to invert the data residuals $\Delta\mathbf{d}$ ($= \mathbf{d} - \mathbf{G}\mathbf{m}_{\text{ref}}$), to determine model perturbation $\Delta\mathbf{m}$ ($= \mathbf{m} - \mathbf{m}_{\text{ref}}$):

$$\Delta\mathbf{d} = \mathbf{G}\Delta\mathbf{m} \quad (5)$$

where \mathbf{G} is either a Jacobian matrix (the partial-derivative matrix of \mathbf{d} with respect to \mathbf{m}_{ref}) for the nonlinear problem in Eq. (1) or the same \mathbf{G} in Eq. (2) for the linear problem. Ill-posed problems require regularizations, such that a similar equation to Eq. (4) is employed to determine the solution:

$$\|\Delta\mathbf{d} - \mathbf{G}\Delta\mathbf{m}\|^2 + \lambda^2 \|\mathbf{L}\Delta\mathbf{m}\|^2 \quad (6)$$

Equation (6) represents the inversion equation for the second scheme, which requires a reference model. The perturbation solution $\Delta\mathbf{m}$ is obtained by minimizing Eq. (6), and the model solution \mathbf{m} is then determined as $\mathbf{m} = \mathbf{m}_{\text{ref}} + \Delta\mathbf{m}$.

Equation (6) can be rewritten as:

$$\|(\mathbf{d} - \mathbf{G}\mathbf{m}_{\text{ref}}) - \mathbf{G}(\mathbf{m} - \mathbf{m}_{\text{ref}})\|^2 + \lambda^2 \|\mathbf{L}\Delta\mathbf{m}\|^2 \quad (7)$$

and simplified to:

$$\|\mathbf{d} - \mathbf{G}\mathbf{m}\|^2 + \lambda^2 \|\mathbf{L}\Delta\mathbf{m}\|^2 \quad (8)$$

A comparison of Eqs. (4) and (8) highlights that both equations contain two terms: the data-fitting term ($\|\mathbf{d} - \mathbf{G}\mathbf{m}\|^2$) and the regularization term ($\|\mathbf{L}\mathbf{m}\|^2$ or $\|\mathbf{L}\Delta\mathbf{m}\|^2$). The data-fitting term is the same in both equations, but the regularization term is not. Therefore, the regularization terms in the two inversion schemes can potentially yield different model solutions. The two inversion schemes will produce the same model solution for the extreme case where the reference model is $\mathbf{0}$ ($\mathbf{m}_{\text{ref}} = \mathbf{0}$, such that $\Delta\mathbf{m} = \mathbf{m}$ and $\|\mathbf{L}\mathbf{m}\|^2 = \|\mathbf{L}\Delta\mathbf{m}\|^2$). However, a nonzero reference model raises the likelihood of $\|\mathbf{L}\mathbf{m}\|^2 \neq \|\mathbf{L}\Delta\mathbf{m}\|^2$, with potentially different model solutions obtained via the two inversion schemes.

3. Tikhonov Regularization

There are many types of regularization (e.g., Engl et al. 2000; Silva et al. 2001; Benning and Burger 2018) all of which are, by definition, artificial. The most widely used regularizations in physical and geophysical studies are the Tikhonov regularizations, especially zeroth- and first-order Tikhonov regularizations, which are also known as damping and flatness, respectively. The effects of these regularizations on the inversion results are illustrated by resolving the synthetic linear problem shown in Fig. 1. I present a simple synthetic 1D ray-tracing problem (Fig. 1a) to illustrate the effects of regularization on the inverted model parameters, where model \mathbf{m} contains the slowness (reciprocal of speed) in each unit segment (1 km) to be resolved (m_i , $i = 1, \dots, 100$), each arrow indicates the raypath coverage,

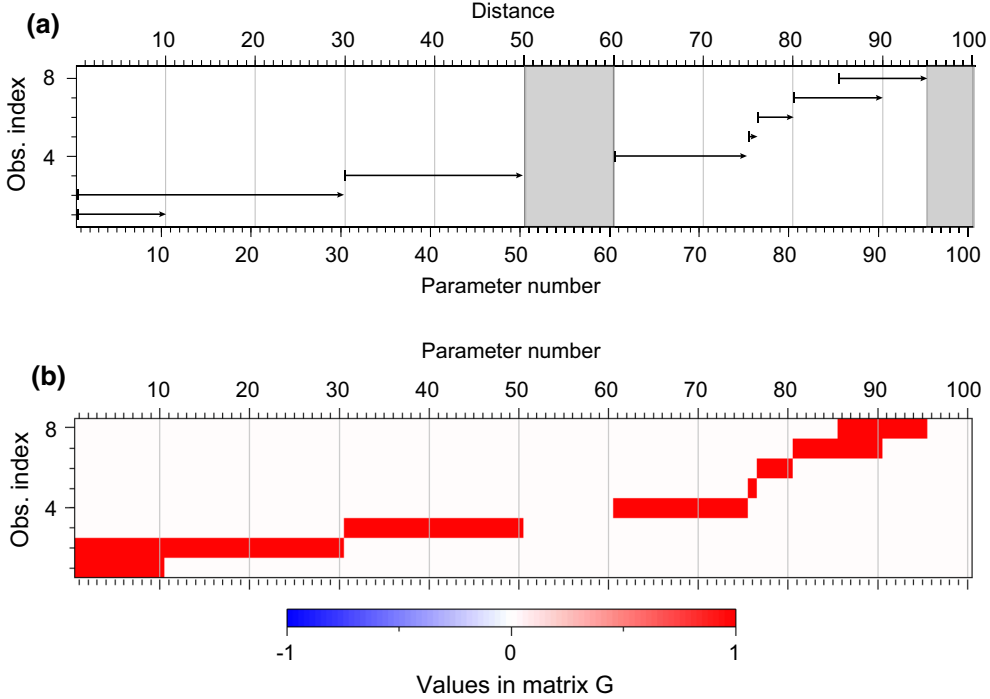


Figure 1

Illustration of a simple 1D ray-tracing linear problem. The model vector \mathbf{m} includes the slowness (reciprocal of speed) in each unit segment to be resolved (m_i ; $i = 1, \dots, 100$). The matrix \mathbf{G} , which is shown in **b**, contains the distance segments traveled by the eight rays (arrow lines in **a**). The observation vector \mathbf{d} contains the travel times for the eight rays (d_j , $j = 1, \dots, 8$). \mathbf{d} is linearly related to \mathbf{m} via Eq. (2). This example is an underdetermined inverse problem, as there are 8 observations and 100 unknowns. The gray-shaded parameters, m_{51-60} and m_{96-100} , are unconstrained, as no rays pass through these model parameters

matrix \mathbf{G} (Fig. 1b) contains the lengths of the eight raypaths, and observation vector \mathbf{d} contains the travel times for the eight rays (d_j , $j = 1, \dots, 8$). Equation (2) highlights the linear relationship between the data \mathbf{d} and model \mathbf{m} , and is applied here. This example is an underdetermined inverse problem, as there are 8 observations and 100 unknown model parameters. The gray-shaded model parameters (m_{51-60} and m_{96-100}) are intentionally set as being unconstrained by the observations, because no rays pass through these model parameters.

3.1. Zeroth-Order Tikhonov Regularization

The zeroth-order Tikhonov regularization \mathbf{L}^0 , which is also termed L_2 -norm regularization or damping (Levenberg 1944), is defined as an identity matrix that regularizes individual parameter values (e.g., Aster et al. 2005; Gregor and Fessler 2015).

Therefore, Eqs. (4) and (8) for the two inversion schemes become:

$$\|\mathbf{d} - \mathbf{G}\mathbf{m}\|^2 + \lambda^2 \|\mathbf{m}\|^2 \quad (9)$$

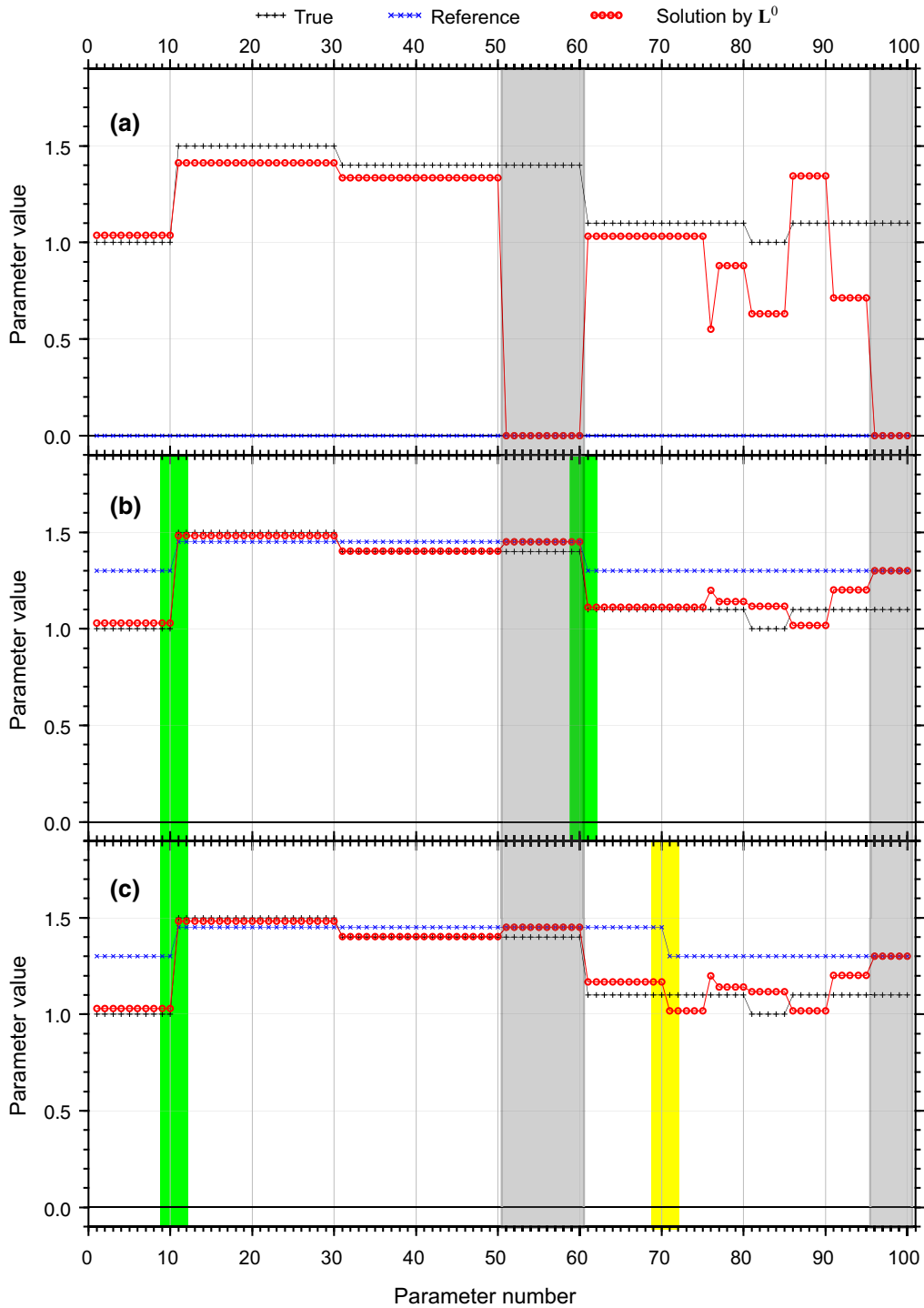
and

$$\|\mathbf{d} - \mathbf{G}\mathbf{m}\|^2 + \lambda^2 \|\Delta\mathbf{m}\|^2 \quad (10)$$

respectively.

The damping regularizations in Eqs. (9) and (10) try to minimize the norms of \mathbf{m} ($\|\mathbf{m}\|^2$) and $\Delta\mathbf{m}$ ($\|\Delta\mathbf{m}\|^2$), respectively. The model solution of the first inversion scheme (Eq. 9) tends to range from $\mathbf{0}$ to the true model (\mathbf{m}_{true}), whereas the model solution of the second scheme (Eq. 10) tends to range from the reference model (\mathbf{m}_{ref} , when $\Delta\mathbf{m} = \mathbf{0}$) to the true model (\mathbf{m}_{true} , when $\Delta\mathbf{m} \neq \mathbf{0}$).

The inversion results via damping (\mathbf{L}^0) regularization of the synthetic problem in Fig. 1 are shown in Fig. 2. The inversion solutions in Fig. 2c were



determined using different regularization parameters λ . A comparison of the inversion results, which highlights the tradeoff between the data-fitting and regularization terms (Eq. 10) for each inversion,

indicates that $\lambda = 1$ is the optimal parameter via Morozov's discrepancy principle (Morozov 1984). The same λ value ($\lambda = 1$) is used for all of the inversions in Fig. 2a-c for comparison. The

◀Figure 2

Inversions of the problem in Fig. 1 via \mathbf{L}^0 regularization (damping) using three different reference models (blue lines). **a** All model parameters in the reference model are zero, such that there is effectively no reference model in the inversion. **b** The reference model possesses two discontinuities (green-shaded areas) at the 10th and 60th parameters, as shown in the synthetic (true) model \mathbf{m}_{true} (line with black crosses). **c** The reference model is similar to that in **b**, with the exception of moving the discontinuity at the 60th parameter to the 70th parameter (yellow-shaded area)

inversions are performed via singular value decomposition with the same observation \mathbf{d} and \mathbf{L}^0 regularization in each case. The results from Scheme 1 via Eq. (9) are shown in Fig. 2a, and the results from Scheme 2 via Eq. (10) are shown in Fig. 2b, c using reference models with different values for the 61st–71st model parameters. The 95th and 96th parameters are neighbors that possess the same values in the synthetic model \mathbf{m}_{true} (line with black crosses, $m_{95} = m_{96}$ in Fig. 2). However, the 95th and 96th inverted parameters from the damped inversion via Eq. (9) (red circles in Fig. 2a) are quite different to each other ($m_{95} \neq m_{96}$), and exhibit greater deviations from their true values than those from the inversions via Eq. (10) (Fig. 2b, c). Such a large deviation also occurs for the 50th and 51st parameters from the inversion via Eq. (9) (Fig. 2a); however, this deviation is not present in the results via Eq. (10) (Fig. 2b, c). The synthetic results are in line with expectations, since Scheme 1 permits a larger solution range (from $\mathbf{0}$ to \mathbf{m}_{true}) than that of Scheme 2 (from \mathbf{m}_{ref} to \mathbf{m}_{true}). The other model parameters that are well-constrained by the observations are reasonably retrieved by both inversion schemes.

The test results imply that damping is helpful in resolving ill-posed inverse problems, but it cannot resolve poorly constrained parameters without a reference model. An inversion that uses a reference model [Scheme 2; either Eq. (6) or Eq. (10)] is therefore preferred over Scheme 1 [either Eq. (4) or Eq. (9)] for ill-posed practical studies using damping. The results from different reference models (blue lines in Fig. 2b, c) show that the true model can be better retrieved when heterogeneities in the reference model are consistent with the true model (green-

shaded region), whereas the true model cannot be retrieved when the heterogeneities in the reference model are different from those in the true model (yellow-shaded region in Fig. 2c). This highlights that the damping effect on inversion results relies on the confidence of the reference model.

3.2. Higher-Order Tikhonov Regularization

A higher-order Tikhonov regularization \mathbf{L}^n (Tikhonov 1963) is defined as either the higher-order derivative or partial derivative of a model or model perturbation that further minimizes the variations among neighboring inverted parameters. Therefore, the $\mathbf{L}^n \mathbf{m}$ and $\mathbf{L}^n \Delta \mathbf{m}$ regularizations in Eqs. (4) and (6), respectively, can be expressed as the n th-order derivatives $d^n \mathbf{m}$ and $d^n \Delta \mathbf{m}$, respectively.

Here, only the first-order ($n = 1$) Tikhonov regularization (\mathbf{L}^1), which uses the first-order derivative ($d^1 \mathbf{m}$ or $d^1 \Delta \mathbf{m}$), is tested. The \mathbf{L}^1 regularization, also called flatness, has been widely applied in various practical inversions (e.g., Vauhkonen et al. 1998; Aster et al. 2005; Gregor and Fessler 2015). The \mathbf{L}^1 matrix for the synthetic problem in Fig. 1 is given as follows (Menke 1989):

$$\mathbf{L}^1 = \begin{vmatrix} 1 & -1 & & & \\ & \ddots & \ddots & & \\ & & & 1 & -1 \end{vmatrix}_{99 \times 100} \quad (11)$$

All of the entries on the i th row (\mathbf{L}^1_i) are zero, with the exception of the i th and $(i + 1)$ th entries. The regularization part ($\mathbf{L}^1 \mathbf{m}$) around the i th parameter m_i in Eq. (4), $\mathbf{L}^1_i \mathbf{m}$, can therefore be written as:

$$d^1 m_i = m_i - m_{i+1} \quad (12)$$

where $d^1 m_i$ represents the difference between two neighboring parameters (m_i and m_{i+1}) in \mathbf{m} . Therefore, the inversion via Eq. (4) tries not only to fit the observation \mathbf{d} , but also to minimize the difference between two neighboring parameters in an attempt to make neighboring parameters have the same values. However, the inversion via Eq. (6), where the regularization part ($\mathbf{L}^1_i \Delta \mathbf{m}$) around the i th parameter Δm_i is:

$$d^1 \Delta m_i = \Delta m_i - \Delta m_{i+1} \quad (13)$$

tries to minimize the difference between the model perturbations of two neighboring parameters in an attempt to make neighboring parameters have the same perturbations.

Figure 3 shows the inversion results for the synthetic linear problem in Fig. 1 with flatness (\mathbf{L}^1) regularization. The inversions in Fig. 3c were run using different λ values, with a value of one selected on the basis of Morozov's discrepancy principle (Morozov 1984). This λ value ($\lambda = 1$) is used for all of the inversions in Fig. 3a, b. The results from Scheme 1 [via Eq. (4)] are shown in Fig. 3a, and those from Scheme 2 [via Eq. (6)] in Fig. 3b, c using different reference models. A comparison of the results in Fig. 3a and Fig. 2a shows that the model obtained via flatness regularization overcomes the false abrupt variation around the 95th and 96th parameters that appears on the model obtained via damping regularization. The inversion with flatness regularization is also less sensitive to sharp variations than the inversion with damping regularization, and retrieves a smoother model (line with red circles) than the true model (line with black crosses). This result is reasonable because \mathbf{L}^1 regularization tries to make the neighboring parameters equal.

The results in Fig. 3 also demonstrate the importance of the reference model. A good reference model whose heterogeneity is consistent with that in the true model is helpful in improving the inversion results (green-shaded regions), whereas a poor reference model with inconsistent heterogeneity may worsen the inversion results (yellow-shaded region), which is similar to the damping regularization in Fig. 2. This problem also exists in inversions using other higher-order Tikhonov regularizations (\mathbf{L}^n , $n > 1$; not shown here).

3.3. Problems and Possible Solutions

All of the above-mentioned inversions based on a reference model are more capable of retrieving the true model than the inversions without a reference model for ill-posed inverse problems. However, the model solutions in the above-mentioned regularized inversions (e.g., \mathbf{L}^0 and \mathbf{L}^1) obviously rely on the confidence of the reference model, as a badly

constructed reference model may worsen the model solution. Both the damping (\mathbf{L}^0) and flatness (\mathbf{L}^1) regularizations failed to correct the badly constructed heterogeneities in the reference model (yellow-shaded regions in Figs. 2c, 3c), and damping produced fake heterogeneities (the 75th–100th parameters in Fig. 2b, c).

There are three possible methods for improving the model solution in the second inversion scheme: by (1) constructing a reference model that is as reliable as possible; (2) using different regularization parameters (λ) or applying multiple regularizations together; or (3) constructing a better regularization to minimize the effects of a badly constructed reference model. Method (1) is not feasible in most cases because the confidence of the reference model is generally unknown. Method (2) can influence the solution when the regularization parameter λ is changed (e.g., Liu 2013; Pazos and Bhaya 2015). However, tests (not shown here) using different λ values in Eq. (6) cannot improve the model solution when it is contaminated by a badly constructed reference model. Simultaneous flatness and damping regularizations (e.g., Chang et al. 2010) are better than using damping alone, but this approach is still unable to avoid the problem caused by a badly constructed reference model. Method (3) appears to be the most feasible way to improve the model solution, whereby a better regularization can weaken the effects of a badly constructed reference model. Here, the focus is on customizing a higher-order Tikhonov regularization (Sect. 4), as damping is seldom used alone during the inversion process.

4. Adapting A Higher-Order Tikhonov Regularization For \mathbf{m}_{ref}

A reference model is normally constructed based on a priori information. The reference model (\mathbf{m}_{ref}) is not homogenous but rather heterogeneous, with variation ($|d^n m_{ref_i}| \geq 0$) around any model parameter (e.g., the i th parameter m_{ref_i}). It is expected that a false variation in the reference model should be corrected by a variation in the opposite direction. However, the \mathbf{L}^1 regularization used in the above tests failed to correct badly constructed

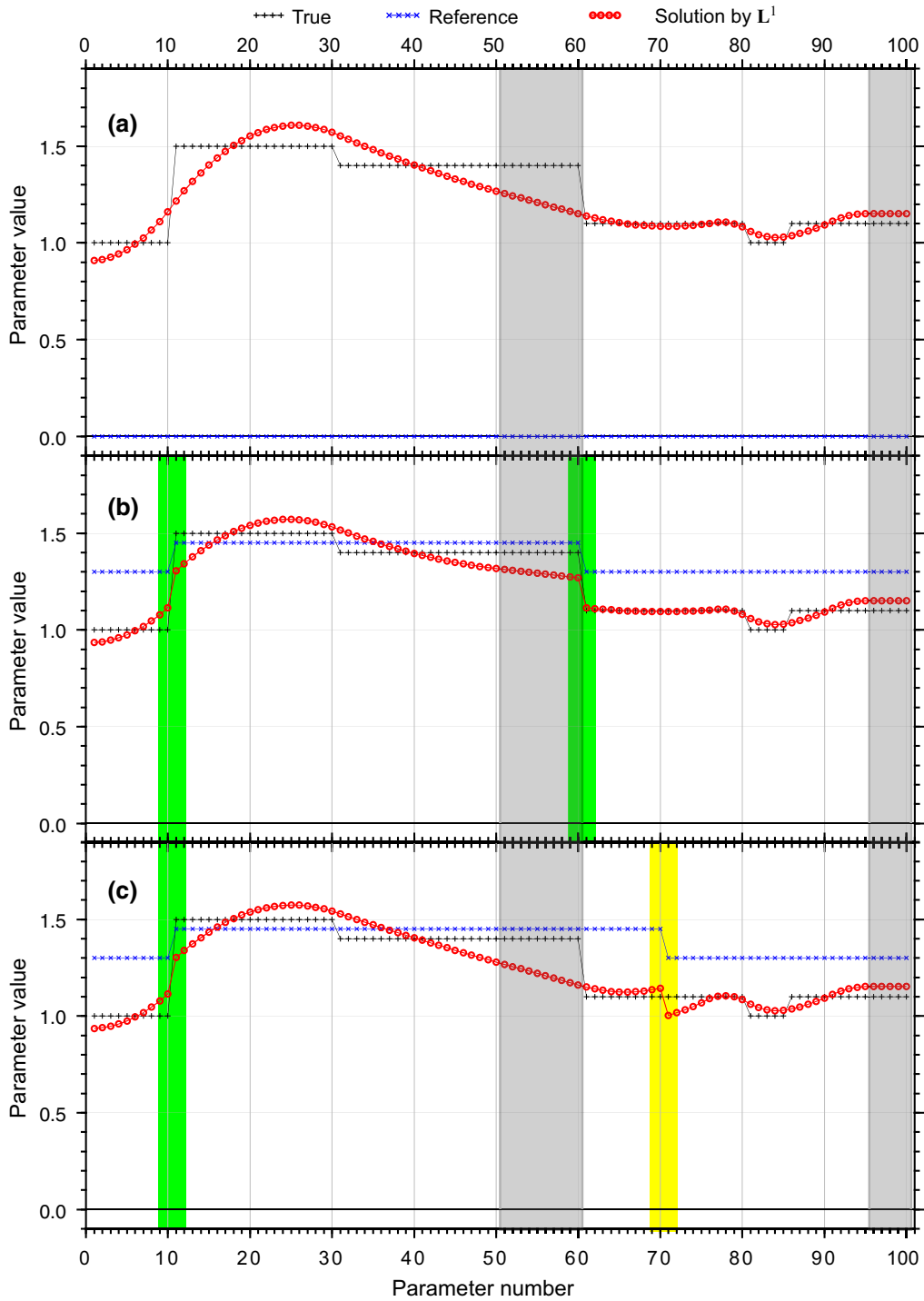


Figure 3

Inversions of the problem in Fig. 1 via L^1 regularization (flatness) using three different reference models (blue lines). The reference models and symbols are the same as in Fig. 2

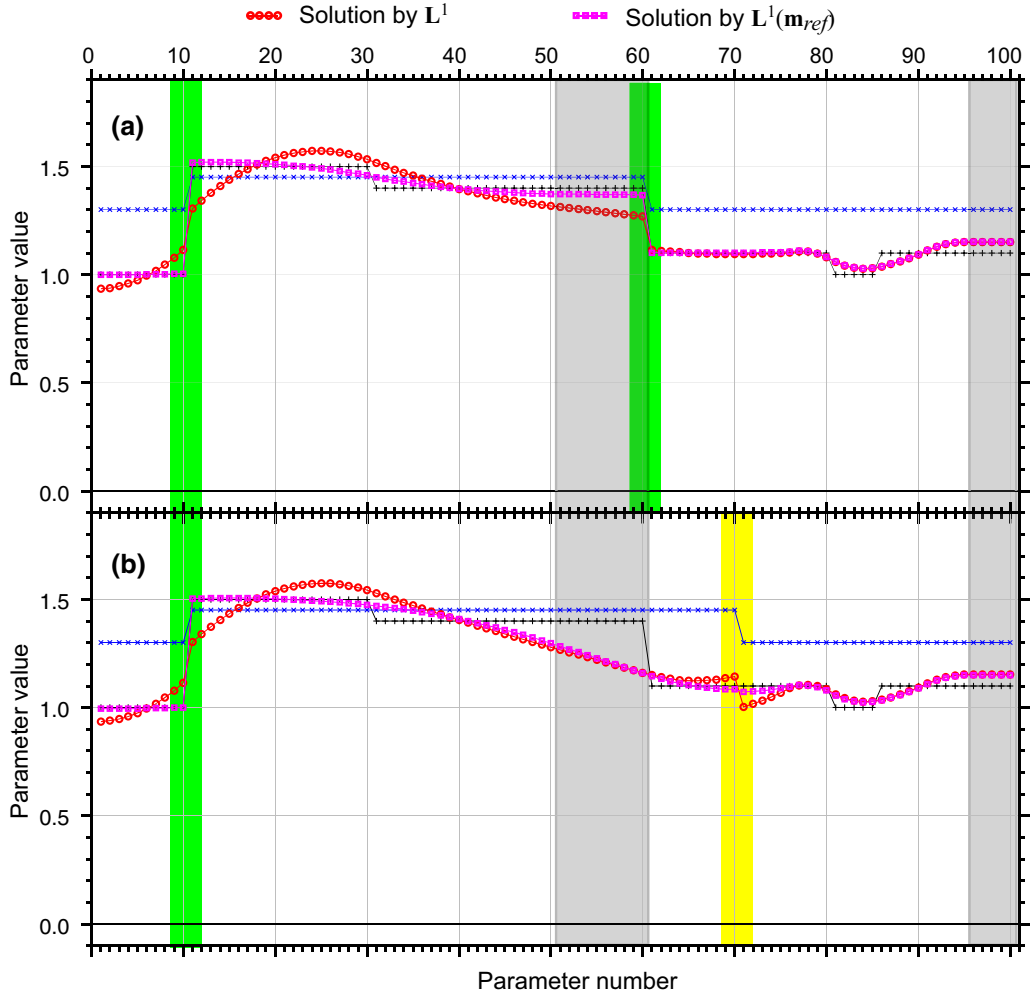


Figure 4

Inversions of the problem in Fig. 1 via $L^1(\mathbf{m}_{ref})$ regularization (adaptive flatness). The lines with magenta squares are the $L^1(\mathbf{m}_{ref})$ -regularized model solutions. The reference models and other symbols in **a**, **b** are the same as in Fig. 3b, c, respectively

is a better regularization approach, even if the reference model is well constructed. Note that another large gradient around the 60th parameter (right green-shaded region) in Fig. 4b is not improved due to the lack of observations (gray-shaded region).

The test case in Fig. 4 illustrates the effectiveness of the proposed adaptive regularization when the number of model parameters (m) is much larger than the number of observations (n_{obs}) ($n_{obs} \ll m$). It is ideal to perform further tests for the cases with $n_{obs} \approx m$ and $n_{obs} > m$ observations, as more observations generally indicate a larger contribution of the observations during the inversion, which weakens the

effect of regularization ($\|\mathbf{L}\mathbf{m}\|^2$) on the solution [Eq. (4)]. I conducted an additional test with $n_{obs} \approx m$ observations, which is shown in Figure S1 in the supporting information. The test includes 10 model parameters ($m = 10$) and 8 observations ($n_{obs} = 8$). A smaller λ value (0.1) is needed for this quasi-balanced problem, and the contribution of regularization to the solution is much weaker than in the problem shown in Fig. 4, as expected. However, the solutions via the proposed adaptive flatness are still much better than those via general flatness, thereby confirming the effectiveness of the proposed adaptive flatness regularization approach.

Data sampling of the model parameters in a practical study is often uneven when $n_{\text{obs}} \geq m$, such that the model may contain repeatedly and/or poorly controlled model parameters. The numbers of independent observations and well-controlled parameters are therefore less than or equal to m , such that the rank of the observation matrix \mathbf{G} is smaller than or equal to m . Therefore, an inversion with $n_{\text{obs}} \geq m$ observations is similar to an inversion with $n_{\text{obs}} \leq m$ observations, which means that the test cases with $n_{\text{obs}} \ll m$ (Fig. 4) and $n_{\text{obs}} \approx m$ (Figure S1) observations can also reflect the effect of regularization for the case with $n_{\text{obs}} \geq m$ observations. This implies that the proposed adaptive flatness regularization is superior to the general regularization for a given case with any number of observations ($n_{\text{obs}} > m$, $n_{\text{obs}} \approx m$, and $n_{\text{obs}} < m$).

5. Application of $\mathbf{L}^1(\mathbf{m}_{\text{ref}})$ Regularization to a Linearized Inverse Problem

Surface-wave dispersion inversion for the S -wave velocity structure is a typical seismological nonlinear inverse problem that is widely applied to study 1D or 3D sedimentary, crustal, or lithospheric earth structures (e.g., Snoke and James 1997; Xia et al. 1999; Feng and An 2010). This new adaptive $\mathbf{L}^1(\mathbf{m}_{\text{ref}})$ regularization was applied here to synthetic surface-wave dispersion inversions (Fig. 5) and practical studies (Fig. 6) to demonstrate the robustness of this new regularization approach.

The model parameters \mathbf{m} in the synthetic inversions are the S -wave velocities at each depth (Fig. 5b), and the observations \mathbf{d} are the Rayleigh-wave group velocities at each period (circles in Fig. 5a). The nonlinear relationship between the Rayleigh-wave group velocity and S -wave velocity indicates that this problem is best resolved by employing the second inversion scheme via Eq. (6), which uses a reference model, where the matrix \mathbf{G} contains partial derivatives of the observations to the reference model $\partial \mathbf{d} / \partial \mathbf{m}$. The synthetic observation and partial-derivative matrix \mathbf{G} were calculated using the program *surf96* (Herrmann and Ammon 2002). The regularization parameter was set to $\lambda = 1$ after a series of tests based on the trade-off between

the misfit and model flatness. Two inversions, each with a different a value (0.2 and 0.02 km s⁻¹), were performed to demonstrate the effects of adaptive regularization. The smallest w_i^n weights for the two a values are 0.4 and 0.067, respectively.

The inversion was run for only one linearized inversion iteration to compare how well general flatness regularization (\mathbf{L}^1 in Eq. 11) and adaptive flatness regularization ($\mathbf{L}^1(\mathbf{m}_{\text{ref}})$ in Eq. 14) fit the observations to the true model. The S -wave velocity profiles are shown in Fig. 5b. The true model (black line) contains two layers with a sharp velocity increase at 35 km depth. The reference model (blue line), which is quite different from the true model, contains a sharp velocity increase at 3 km depth and a gradual velocity increase from 35 to 50 km depth. The $\mathbf{L}^1(\mathbf{m}_{\text{ref}})$ -regularized models (green and magenta lines) exhibit a marked improvement over the \mathbf{L}^1 -regularized models (red line), and are therefore more similar to the true model (black line), especially in the yellow-shaded area where the reference model contains a large false gradient. The $\min(w_i^n)$ value, 0.067, in the inversion when using a smaller constant a (0.02 km s⁻¹) is much smaller than 0.4 when $a = 0.2$ km s⁻¹, which indicates a stronger adaptive regularization with $a = 0.02$ km s⁻¹, and that the inverted result (magenta line in Fig. 5b) is closer to the true model than the inverted model with $a = 0.2$ km s⁻¹ (green line). The structures at other depths are not worsened (magenta line almost overlaps with red line) where the reference model possesses a reliable structure. Therefore, adaptive flatness is also better than general flatness in a nonlinear (linearized) inverse problem, similar to the presented linear inversion tests.

The lithospheric structure beneath the Paraná Basin, SE Brazil, has been carefully studied using surface waves recorded at seismic station RIFB via linearized inversion (Snoke and James 1997) and global optimization methods (Snoke and Sambridge 2002; An and Assumpção 2006). Here, a linearized inversion was performed for the lithospheric structure (Fig. 6) using the observations (phase and group velocities of the Rayleigh and Love waves) in Fig. 5 of Snoke and James (1997). The inversions were run for only one iteration by applying \mathbf{L}^1 and $\mathbf{L}^1(\mathbf{m}_{\text{ref}})$. The reference model (Fig. 6b) is the same as in

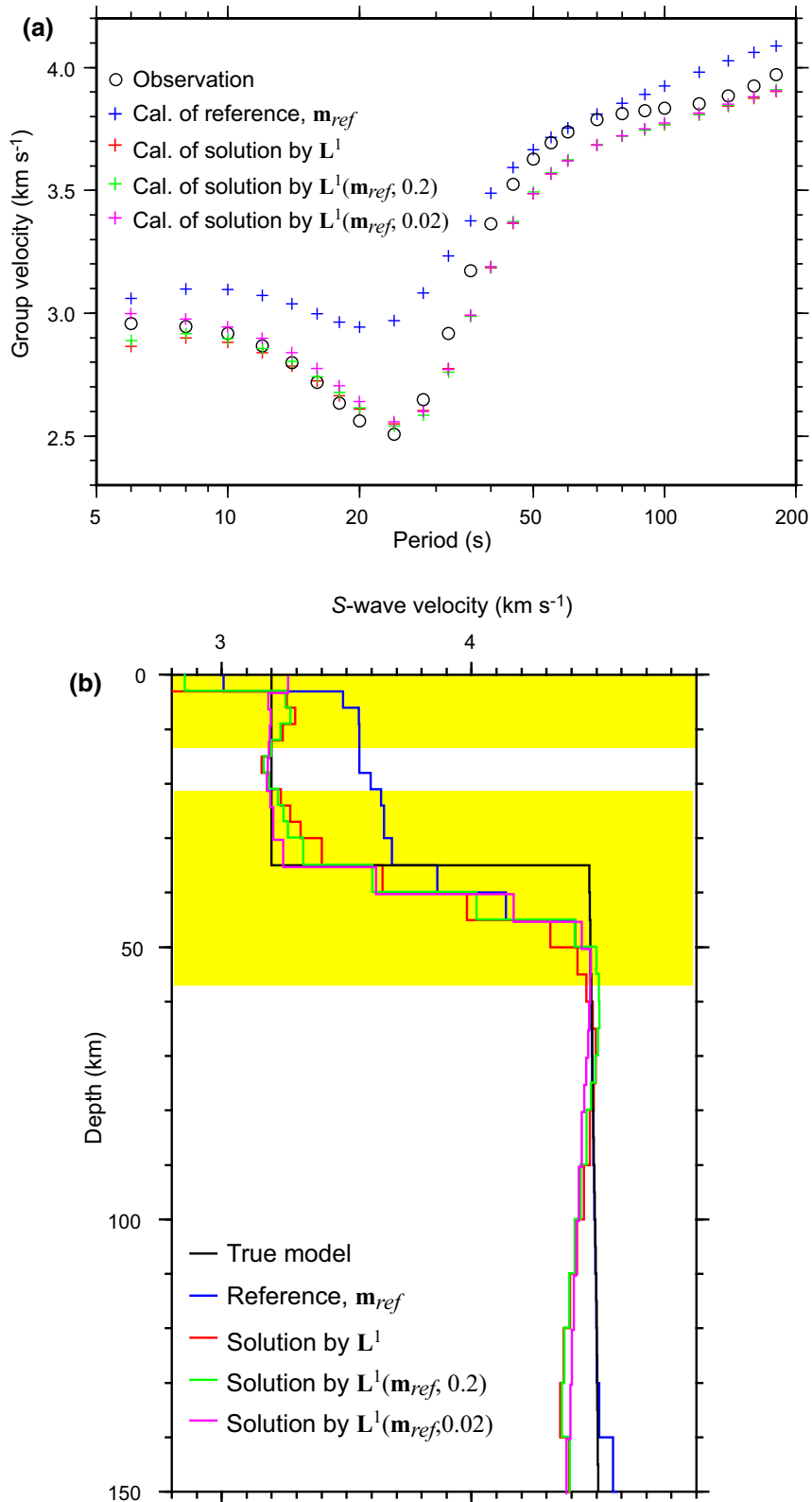


Figure 5

Synthetic linearized surface-wave inversion using the $L^1(\mathbf{m}_{ref})$ adaptive flatness regularization. $L^1(\mathbf{m}_{ref}, 0.2)$ and $L^1(\mathbf{m}_{ref}, 0.02)$ represent the inverted results using $L^1(\mathbf{m}_{ref})$, and $a = 0.2$ and 0.02 km s^{-1} , respectively. **a** Synthesized observations of the fundamental-mode Rayleigh-wave group velocities (black circles) and predictions from the 1D S -wave velocity models in **b**. **b** The inverted 1D S -wave velocity models via L^1 (red), $L^1(\mathbf{m}_{ref}, 0.2)$ (green), and $L^1(\mathbf{m}_{ref}, 0.02)$ (magenta) regularization, together with the true (black) and reference (blue) models

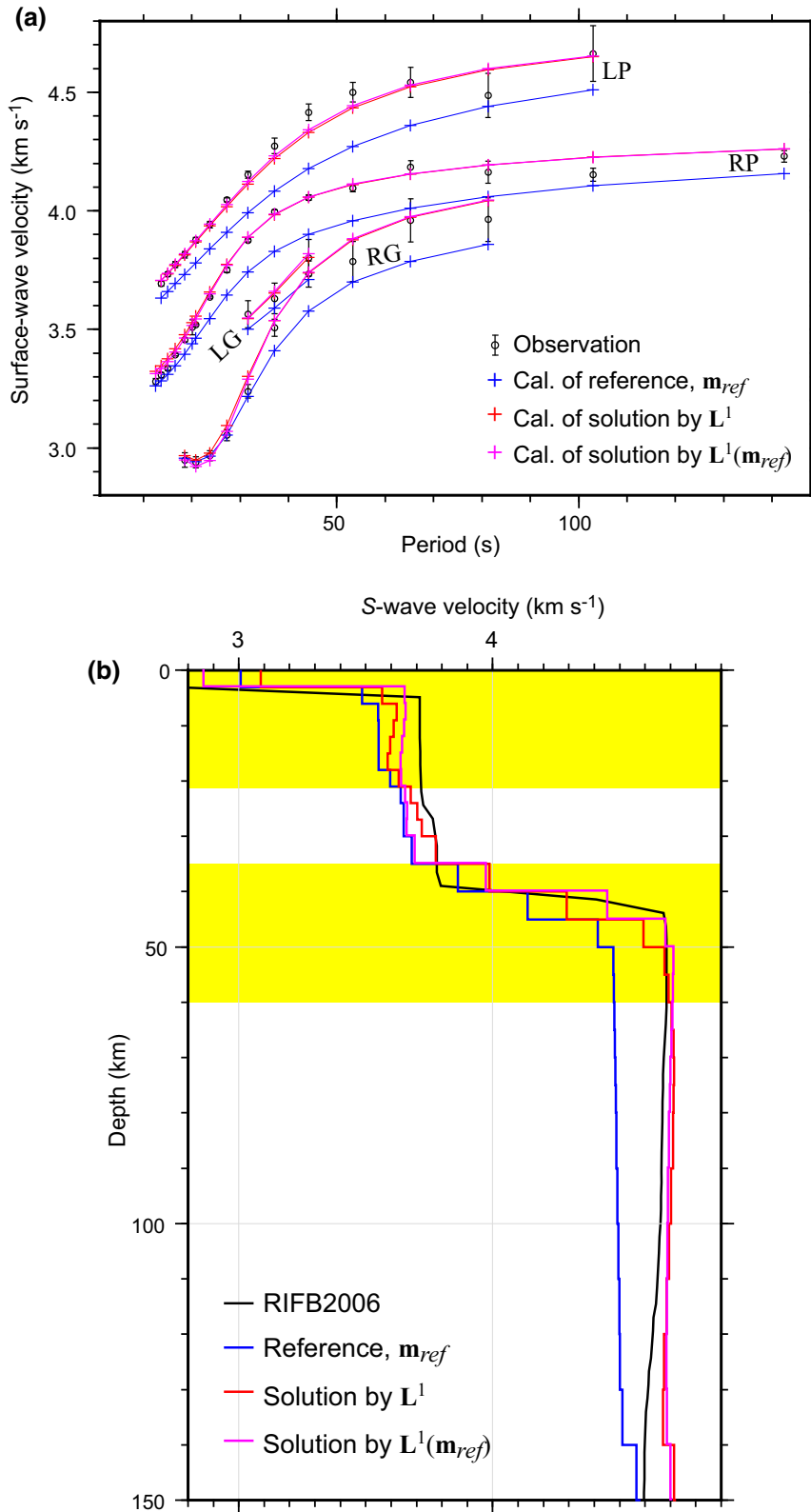


Figure 6

Linearized fundamental-mode surface-wave inversion for the Paraná Basin using the adaptive flatness regularization. The black circles with error bars in **a** represent the observations from Snoke and James (1997). *RP* Rayleigh-wave phase velocity, *RG* Rayleigh-wave group velocity, *LP* Love-wave phase velocity, *LG* Love-wave group velocity. Model RIFB2006 in **b** is the inverted model via global optimization in An and Assumpção (2006). The other symbols are as in Fig. 5

Fig. 5b, with $a = 0.02 \text{ km s}^{-1}$ and $\lambda = 1$. The inverted results are shown in Fig. 6. For comparison, the model via global optimization (An and Assumpção 2006) is shown in Fig. 6 (labeled RIFB2006). The inverted model in the first iteration via $L^1(\mathbf{m}_{\text{ref}})$ regularization is much closer to RIFB2006 than that via L^1 regularization. This practical example confirms that adaptive flatness regularization is better than general flatness regularization.

6. Conclusion

Regularized inversions of a reference model involve regularizing the perturbation between inverted and reference models. The inversions demonstrate that a well-constructed reference model can help the inversion, as expected. However, an ill-constructed variation around a given position in the reference model is difficult to be corrected when a non-adaptive regularization is used, as the regularization tends to weaken variations in the inverted perturbation. The adaptive regularization of the reference model that is suggested here can improve the inversion, particularly if the reference model possesses ill-constructed heterogeneities.

Figure S1. Inversions of a model with 10 model parameters regularized by adaptive flatness $L^1(\mathbf{m}_{\text{ref}})$. (a) is the observation matrix G . The symbols in (a) are the same as those in Fig. 1b. The symbols in (b) and (c) are the same as in Fig. 4a and b, respectively.

Acknowledgements

This work was financed by the National Natural Science Foundation of China (Grants 41574049 and 41974051), Basic Research Foundation of Chinese Academy of Geological Sciences (YWF201904), and the China Geological Survey Program (DD20190010). I thank two anonymous reviewers for their constructive comments and suggestions. All figures were generated using Generic Mapping Tools (Wessel and Smith, 1991).

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

REFERENCES

- An, M., & Assumpção, M. S. (2006). Crustal and upper mantle structure in intracratonic Paraná Basin, SE Brazil, from surface wave dispersion using genetic algorithm. *Journal of South American Earth Sciences*, 21(3), 173–184. <https://doi.org/10.1016/j.jsames.2006.03.001>.
- Aster, R. C., Borchers, B., & Thurber, C. H. (2005). *Parameter estimation and inverse problems*. Burlington: Academic Press.
- Babaeizadeh, S., & Brooks, D. H. (2007). Electrical impedance tomography for piecewise constant domains using boundary element shape-based inverse solutions. *IEEE Transactions on Medical Imaging*, 26(5), 637–647. <https://doi.org/10.1109/TMI.2006.887367>.
- Backus, G., & Gilbert, J. F. (1967). Numerical applications of a formalism for geophysical inverse problems. *Geophysical Journal of the Royal Astronomical Society*, 13, 247–276.
- Backus, G., & Gilbert, J. F. (1968). The resolving power of gross earth data. *Geophysical Journal of the Royal Astronomical Society*, 16, 169–205.
- Benning, M., & Burger, M. (2018). Modern regularization methods for inverse problems. *Acta Numerica*, 27, 1–111. <https://doi.org/10.1017/S0962492918000016>.
- Chang, S.-J., van der Lee, S., Flanagan, M. P., Bedle, H., Marone, F., Matzel, E. M., et al. (2010). Joint inversion for three-dimensional S velocity mantle structure along the Tethyan margin. *Journal of Geophysical Research*, 115, B8. <https://doi.org/10.1029/2009jb007204>.
- Darne, C., Lu, Y., & Sevic-Muraca, E. M. (2013). Small animal fluorescence and bioluminescence tomography: A review of approaches, algorithms and technology update. *Physics in Medicine and Biology*, 59(1), R1–R64. <https://doi.org/10.1088/0031-9155/59/1/r1>.
- Engl, H. W., Hanke, M., & Neubauer, A. (2000). *Regularization of inverse problems. Mathematics and its applications*, 375 (p. 322). Berlin: Kluwer Academic Publishers.
- Feng, M., & An, M. (2010). Lithospheric structure of the Chinese mainland determined from joint inversion of regional and teleseismic Rayleigh-wave group velocities. *Journal of Geophysical Research*, 115, B06317. <https://doi.org/10.1029/2008JB005787>.
- Golub, G. H., Hansen, P. C., & O'Leary, D. P. (1999). Tikhonov regularization and total least squares. *SIAM Journal on Matrix Analysis and Applications*, 21(1), 185–194. <https://doi.org/10.1137/s0895479897326432>.
- Gregor, J., & Fessler, J. A. (2015). Comparison of SIRT and SQS for regularized weighted least squares image reconstruction. *IEEE Transactions on Computational Imaging*, 1(1), 44–55. <https://doi.org/10.1109/TCI.2015.2442511>.
- Herrmann, R. B., & Ammon, C. J. (2002). Computer programs in seismology—surface waves, receiver functions and crustal structure. St. Louis University, St. Louis, MO. <https://www.eas.slu.edu/People/RBHerrmann/ComputerPrograms.html>.

- Keith, B., Vitasek, K., Manrodt, K., & Kling, J. (2016). *Strategic sourcing in the new economy: Harnessing the potential of sourcing business models for modern procurement* (p. 448). US: Palgrave Macmillan.
- Levenberg, K. (1944). A method for the solution of certain non-linear problems in least squares. *Quarterly of Applied Mathematics*, 2, 164–168.
- Liu, C.-S. (2013). A dynamical tikhonov regularization for solving ill-posed linear algebraic systems. *Acta Applicandae Mathematicae*, 123(1), 285–307. <https://doi.org/10.1007/s10440-012-9766-3>.
- Menke, W. (1989). *Geophysical data analysis: Discrete inverse theory (revised edition)* (p. 289). San Diego: Academic Press.
- Morozov, V. A. (1984). *Methods for solving incorrectly posed problems* (p. 257). New York: Springer.
- Nolet, G. (2008). *A breviary of seismic tomography: Imaging the interior of the earth and sun*. Cambridge: Cambridge University Press.
- Parker, R. L. (1994). *Geophysical inverse theory* (p. 386). Princeton: Princeton University Press.
- Pazos, F., & Bhaya, A. (2015). Adaptive choice of the Tikhonov regularization parameter to solve ill-posed linear algebraic equations via Liapunov Optimizing Control. *Journal of Computational and Applied Mathematics*, 279, 123–132. <https://doi.org/10.1016/j.cam.2014.10.022>.
- Scales, J. A., Smith, M. L., & Treitel, S. (2001). *Introductory geophysical inverse theory* (p. 193). Toronto: Samizdat Press.
- Shaw, P. R., & Orcutt, J. A. (1985). Waveform inversion of seismic refraction data and applications to young Pacific crust. *Geophysical Journal International*, 82(3), 375–414. <https://doi.org/10.1111/j.1365-246X.1985.tb05143.x>.
- Silva, J. B. C., Medeirosz, W. E., & Barbosa, V. C. F. (2001). Potential-field inversion: Choosing the appropriate technique to solve a geologic problem. *Geophysics*, 66(2), 511–520. <https://doi.org/10.1190/1.1444941>.
- Snoko, J. A., & James, D. E. (1997). Lithospheric structure of the Chaco and Paraná Basins of South America from surface-wave inversion. *Journal of Geophysical Research*, 102, 2939–2951.
- Snoko, J. A., & Sambridge, M. (2002). Constraints on the S wave velocity structure in a continental shield from surface wave data: Comparing linearized least squares inversion and the direct search Neighbourhood Algorithm. *Journal of Geophysical Research*, 107(B5), 2094. <https://doi.org/10.1029/2001JB000498>.
- Tikhonov, A. N. (1963). On the solution of ill-posed problems and the method of regularization. *Doklady Akademii Nauk*, 151(3), 501–504.
- Vauhkonen, M., Vadasz, D., Karjalainen, P. A., Somersalo, E., & Kaipio, J. P. (1998). Tikhonov regularization and prior information in electrical impedance tomography. *IEEE Transactions on Medical Imaging*, 17(2), 285–293. <https://doi.org/10.1109/42.700740>.
- Wessel, P., & Smith, W. H. F. (1991). Free software helps map and display data. *Eos, Transactions American Geophysical Union*, 72, 441. <https://doi.org/10.1029/90EO00319>.
- Xia, J., Miller, R. D., & Park, C. B. (1999). Estimation of near-surface shear-wave velocity by inversion of Rayleigh waves. *Geophysics*, 64(3), 691–700.
- Xu, Y., Pei, Y., & Dong, F. (2016). An adaptive Tikhonov regularization parameter choice method for electrical resistance tomography. *Flow Measurement and Instrumentation*, 50, 1–12. <https://doi.org/10.1016/j.flowmeasinst.2016.05.004>.

(Received November 6, 2019, revised June 2, 2020, accepted June 5, 2020)