Pure and Applied Geophysics



On Resolution Matrices

Meijian An¹

Abstract-Solution appraisal, which has been realized on the basis of projections from the true medium to the solution, is an essential procedure in practical studies, especially in computer tomography. The projection operator in a linear problem or its linear approximation in a nonlinear problem is the resolution matrix for the solution (or model). Practical applications of a resolution matrix can be used to quantitatively retrieve the resolvability of the medium, the constrainability of the solution parameters, and the relationship between the solution and the factors in the study system. A given row vector of the matrix for a solution parameter can be used to quantify the resolvability, deviation from expectation, and difference between that solution parameter and its neighbor from the main-diagonal element, rowvector sum, and difference between neighboring elements in the row vector, respectively. The resolution length of a solution parameter should be estimated from the row vector, although it may be unreliable when the vector is unstable (e.g., due to errors). Comparatively, the resolution lengths that are estimated from the column vectors of the observation-constrained parameters are reliable in this instance. Previous studies have generally employed either the direct resolution matrix or the hybrid resolution matrix as the model resolution matrix. The direct resolution matrix and hybrid resolution matrix in an inversion with damping (or general Tikhonov regularization) are Gramian (e.g., symmetric). The hybrid resolution matrix in an inversion using zero-row-sum regularization matrices (e.g., higher-order Tikhonov regularizations) is one-row-sum but is not a stochastic matrix. When the two resolution matrices appear in iterative nonlinear inversions, they are not a projection of the solution, but rather the gradient of the projection or a projection of the solution improvement immediately after a given iteration. Regardless, their resultant resolution lengths in iterative nonlinear inversions of surface-wave dispersion remain similar to those from the projection of the solution. The solution is influenced by various factors in the study, but the direct resolution matrix is derived only from the observation matrix, whereas the hybrid resolution matrix is derived from the observation and regularization matrices. The limitations imply that the appropriateness using the two resolution matrices may be questionable in practical applications. Here we propose a new complete resolution matrix to overcome the limitations, in which all of the factors (e.g., errors) in linear or nonlinear (inverse or non-inverse) studies can be

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00024-022-03211-9.

incorporated. Insights on all of the above are essential for ensuring a reliable and appropriate application of the resolution matrix to appraise the model/solution and understand the relationship between the solution and all of the factors in the study system, which is also important for improving the system.

Keywords: Linear projection, linear regression, regularization, observational constraints, gramian matrix.

1. Introduction

Geophysics is the investigation of the Earth based on the principles of physics, whereby the physical properties of the Earth medium (i.e., the unknown model) are inverted from either surface- or spacebased observations. Geophysical inverse problems are often underdetermined (e.g., Menke 2015; Tarantola and Valette 1982; Wiggins 1972) owing to the various limitations of these observations. However, advances in generalized or regularized inverse theory over the last century now make it possible to solve underdetermined inverse problems (e.g., Hansen 1992; Lawson and Hanson 1995; Levenberg 1944; Moore 1920; Morozov 1984; Penrose 1955; Tikhonov 1963). Although the quality and reliability of such an under-constrained solution is essential in a data-poor environment, the verification of a geophysical solution is difficult or even impossible, as the investigated medium is generally inaccessible. As such, geophysicists have to make great efforts in the methodology of solution appraisal. However, understanding what is reliable in some geophysical solutions remains perhaps the most exciting challenge to date (Foulger et al. 2015).

An inverted solution (\underline{x}) that represents investigated medium (\underline{x}) can be described as:

¹ Chinese Academy of Geological Sciences, Beijing, China. E-mail: meijianan@live.com

where the operator *r*, which can also be denoted by $r: x \to \underline{x}$, represents the relationship between \underline{x} and x. This problem can be linearized as:

$$\underline{x} = Rx, \qquad (2)$$

where **R**, which can also be denoted by $\mathbf{R:x} \rightarrow \underline{x}$, is the resolution matrix (Backus and Gilbert 1968, 1970), a linear projection approximation of $r:\mathbf{x} \rightarrow \underline{x}$ (Eq. (1)). **R** has been widely used in solution appraisals in geophysics (e.g., Aki et al. 1977; An 2012; Aster et al. 2005; Menke 2015; Tarantola and Valette 1982; Wiggins 1972; Yao et al. 1999) and other research areas (e.g., Lütkenhöner and Grade de Peralta Menendez 1997; Katamreddy and Yalavarthy 2012).

However, the properties of R remain poorly understood in practical problems, and considerable uncertainties and/or inconsistencies regarding R still exist. For example, early studies mostly analyzed the diagonal entries to evaluate the resolvability of the target medium (e.g., Aki et al. 1977; Day-Lewis et al. 2005; Wiggins 1972), even though the solution is actually related to all of the entries in R (Eq. (2)). \boldsymbol{R} has been mostly applied to estimate the resolution length (or resolution width). This length should be retrieved from a given row in R (e.g., An 2012; Barmin et al. 2001; Crosson 1976), although the result appears to be similar to that from the corresponding column (e.g., Alumbaugh and Newman 2000; Miller and Routh 2007; Pilkington 2016). Why and when are they similar? An inverse problem is often solved using reference model, whereby the inverted solution is not a model of the medium but rather a perturbation (Δx) of the reference model (x^{1}) . R can be provided in the inversions (e.g., Jackson 1972; Ren and Kalscheuer 2020), but it represents the $\Delta x \rightarrow \Delta x$ projection, not $x \rightarrow x$. What is the relationship between the matrix and $r: x \to x$? Can the matrix be used to estimate the resolution length of the solution $x (= \Delta x + \underline{x}^i)$? All the above questions will be addressed in this paper.

R is particularly useful in model appraisal, but it comes with various limitations. For example, some of the general factors (e.g., observational and data processing errors) cannot be reflected by the matrix. The

resolution estimated from the matrix with limitations may be unrealistically high (Pilkington 2016). A matrix $\mathbf{R}: \mathbf{x} \to \mathbf{x}$ with all of the factors in a complete process can overcome these limitations. However, no such matrix exists to date.

In total, **R** is a unique quantitative indicator of the reliability of a given solution, and it is also important in understanding the relationships between the solution and the observations, regularization, and other factors. However, various uncertainties and limitations regarding R remain. This paper reviews previous resolution matrices and clarifies both the significance and properties of the matrices, which often appear in practical inversions, to explain how to appropriately employ such a matrix in a given study. Furthermore, this paper clarifies the resolution matrices in nonlinear inversions and suggests a new resolution matrix, which can include all of the factors in a linear/nonlinear problem. This study can therefore assist in the appropriate selection and implementation of a resolution matrix and provide the reader with a better understanding of both the quality and reliability of the solution and the relationships between the solution and all of the factors in the study system, which are important for further improving the system.

2. Resolution Matrices from Observations and Regularization Matrices

Resolution matrices that are derived from observations and regularization matrices have been widely applied in various research studies. This section provides a review of resolution matrices and their significance. Furthermore, the matrix properties, which are important in matrix applications, are first clarified.

2.1. Variables Used

 x, \underline{x} : Real medium (or true model) vector, solution (or inverted model) vector.

 x_i , \underline{x}_i : *i*th real-medium parameter, *i*th solution parameter.

 $r_{i,j}$: Entry at the *i*th row and *j*th column of matrix **R**.

 $r_{i,*}$ (or $r_{*,j}$): *i*th row (or *j*th column) vector of matrix **R**.

 $\Sigma \mathbf{r}_{i,*}$ or $\Sigma_i \mathbf{R}$: Sum of all of the entries in the *i*th row of matrix \mathbf{R} .

2.2. Solutions of the Inverse Problem

The goal of a geophysical investigation is to directly retrieve the solution of the medium (\underline{x}_D) from observations $(\underline{d} = d + \delta d)$ that are contaminated with errors (δd) via the inverse equation:

$$\underline{x}_{\mathrm{D}} = g^{-\mathrm{g}}\underline{d},\tag{3}$$

which is based on the physical relationship between the medium parameters $(m \times 1 \text{ vector } \boldsymbol{x}; [x_1, x_2,..., x_m]^T)$ and observational data $(n \times 1 \text{ vector } \boldsymbol{d}; [d_1, d_2,..., d_n]^T)$, with the latter defined as:

$$\boldsymbol{d} = g(\boldsymbol{x}),\tag{4}$$

the operator g in Eq. (4) is often not invertible, such that a generalized inverse of g, g^{-g} , is used, as in Eq. (3).

Equation (3) can be expressed in linear form as:

$$\underline{x}_{\mathrm{D}} = \boldsymbol{G}^{-\mathrm{g}} \underline{\boldsymbol{d}},\tag{5}$$

and rewritten as:

$$\boldsymbol{d} = \boldsymbol{G}\boldsymbol{x},\tag{6}$$

the $n \times m$ matrix G, which is normally termed the observation matrix, is composed of the sensitivities of d with respect to x. The generalized inverse of G, G^{-g} (e.g., Lawson and Hanson 1995; Moore 1920; Penrose 1955; Tan 2017; Tarantola and Valette 1982), can be either a right inverse:

$$\boldsymbol{G}^{-\mathrm{g}} = \boldsymbol{G}^{\mathrm{T}} \left(\boldsymbol{G} \boldsymbol{G}^{\mathrm{T}} \right)^{-1}, \quad n < m, \tag{7}$$

or a left inverse:

$$\boldsymbol{G}^{-\mathrm{g}} = \left(\boldsymbol{G}^{\mathrm{T}}\boldsymbol{G}\right)^{-1}\boldsymbol{G}^{\mathrm{T}}, \quad n \ge m.$$
(8)

 G^{-g} can also be obtained via singular value decomposition (SVD) (Golub and Kahan 1965; Golub and Reinsch 1970; Varah 1973) or some form of matrix factorization (e.g., Gentle 2007; Golub 1965). The calculated G^{-g} from those methods is tightly related or exactly equivalent.

Figure 1a shows an example (Example 1) of a one-dimensional (1-D) underdetermined inverse

problem that mimics the relationship between distance, slowness (x), and travel time (d). The problem is described by 10 linear equations (Eq. (S1) in the supporting information), with 100 (m = 100)unknown parameters in x and 10 (n = 10) observations (d). The coefficient of the *i*th parameter (x_i) in the *j*th equation for d_i is the product of the segment length x_i that is traveled by the *j*th observation (Fig. 1a) and the sensitivity of d_i with respect to x_i . All of the coefficients are stored in G (Fig. 1b). Several coefficients in Eq. (S1) (or the entries in (G) for the second and ninth observations are greater than one, which means that higher sensitivities are set for the parameters because the segment lengths of all the parameters are equal to one (Fig. 1a). Figure 1c shows a synthetic x and its pseudo-inverse solution $x_{\rm D}$. The synthetic error-free observation $d (\delta d = \{0\})$ and prediction of \underline{x}_{D} (d_{D}) are shown in Fig. 1d.

Both the data coverage (observation distribution) and sensitivities, which are stored in G, influence the solution in the synthetic example (Fig. 1). For example, the solution parameters from x_1 to x_{10} are constrained by only one observation (the first travel path; the top left arrow in Fig. 1a), such that the nonzero entries for all ten parameters in the first row of G (Fig. 1b) and the coefficients in the first row of Eq. (S1) are the same and equal to one. Similarly, the parameters from x_{11} to x_{20} are also constrained by one observation (the second path), but the nonzero entries for the ten parameters in the second row of G (Fig. 1b) and the coefficients in the second row of Eq. (S1) are different. The resultant $x_1 - x_{10}$ values in \underline{x}_{D} (Fig. 1c) are the same and equal to the average of the synthetic model parameters $x_1 - x_{10}$ (circles in Fig. 1c), whereas the resultant $x_{11}-x_{20}$ values are quite different, both from each other and their respective synthetic values $(x_{11}-x_{20})$. The differences in $x_{11}-x_{20}$ are caused by different sensitivities because they all have the same path segments. Parameters $\underline{x}_{21} - \underline{x}_{40}$ and $\underline{x}_{41} - \underline{x}_{60}$ are constrained by two paths, but their path-overlapping patterns (Fig. 1a and b) differ. The differences in $x_{21}-x_{60}$ in $x_{\rm D}$ (Fig. 1c) are related to the path coverage. Parameters $x_{61} - x_{90}$ are constrained by several observations. $x_{91}-x_{100}$ are not constrained by any observations, such that they are all equal to zero in $\underline{x}_{\mathrm{D}}$.



Figure 1

Inverse problem Example 1. This is a simple 1-D ray-propagation (linear) problem that is defined by either d = Gx or Eq. (S1). Arrow lines in (a) illustrate the ray paths. The observation vector d contains the travel times for all ten rays $(d_j; j = 1, ..., 10)$. The vector x includes the slowness (reciprocal of speed) in each unit segment to be resolved $(x_i; i = 1, ..., 100)$. The matrix G in (b) contains the combinations of the distance segments travelled by the ten rays and their synthetic sensitivities. Several parameters are set to a high sensitivity in the second and ninth observations, which simulates a study with parameters that possess different sensitivities. c Synthetic model x (circles) and solutions \underline{x}_D , $\underline{x}(I)$, and $\underline{x}(L^1)$ (lines) obtained via generalized inversions, with zeroth- (I) and first-order (L^1) Tikhonov regularizations included in the latter two inversions. d The synthetic observation data \underline{d} (circles) from x, and the predictions d_D , d(I), and $d(L^1)$ (lines) from the solutions \underline{x}_D , $\underline{x}(I)$ and $\underline{x}(L^1)$, respectively

However, most geophysical inverse problems are ill-posed (i.e., there are not sufficient observations to obtain a unique and stable solution), and the generalized solution $x_{\rm D}$, such as that in Example 1 (Fig. 1c), is not physically rational. Furthermore, large discrepancies may exist between the real model x and solution $x_{\rm D}$, especially for parameters with poor or no observation coverage (Fig. 1a-c), even though there is a good fit between the observation data and model predictions (Fig. 1d). Additional artificial constraints (often called regularization) (e.g., Aster et al. 2005; Benning and Burger 2018; Engl et al. 2000; Levenberg 1944; Menke 1989; Tikhonov 1963) on the model parameters must therefore be included during the inversion to obtain a physically rational solution. The regularization-based forward equation then becomes:

$$\boldsymbol{b} = \boldsymbol{A}\boldsymbol{x},\tag{9}$$

where the $n_{\rm b} \times m$ matrix **A** and $n_{\rm b} \times 1$ vector **b** are:

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{G} \\ \boldsymbol{C} \end{bmatrix} \quad \text{and} \quad \boldsymbol{b} = \begin{bmatrix} \boldsymbol{d} \\ \boldsymbol{c} \end{bmatrix}, \quad (10)$$

respectively, which contain an $n_c \times m$ ($n_c = n_b - n$) matrix **C** and an $n_c \times 1$ vector **c**, respectively, both of which are related to the regularization.

Tikhonov regularization (Levenberg 1944; Tikhonov 1963) is used largely in geophysical inversions (e.g., Aster et al. 2005; Constable et al. 1987; Menke 1989). *C* is denoted by λL^n in n^{th} -order Tikhonov regularizations, and *c* (Eq. (10)) is a zero vector ({0}). The factor λ is the regularization parameter, which balances the contributions of either the observations and regularization in the inversion, or the fit to the observational data *d* and regularization vector c. Tests that employ ad hoc methods (e.g., Craven and Wahba 1978; Hansen 1992; Morozov 1984) are often used to determine λ . The matrix L^0 for zerothorder Tikhonov regularization (damping regularization), which minimizes the model (Levenberg 1944), is the identity matrix I. L^n (n > 0; e.g., L^1 in Eq. (S2) in the supporting information) is a zero-row-sum band matrix. L^1 regularization (flatness regularization) flattens the model via minimizing the first-order gradient of the model. L^n Tikhonov regularization exerts uniform a priori constraints to all of the solution parameters. Therefore, regularization ($C = \lambda L^n$) with an optimal weighting factor λ produces the optimal average regularizing effect for all of the parameters. However, if C contains a diagonal matrix W (= diag($w_1, w_2,...$)) (i.e., $C = \lambda WL^n$), then the regularization is heterogeneous and spatially variant for the model (e.g., An 2020; Katamreddy and Yalavarthy 2012; Pogue et al. 1999; Sanny et al. 2018).

Regularization at least makes A a full-column rank matrix, such that A^{-1} (or the left inverse A^{-g}), which is an $m \times n_b$ matrix, can be uniquely obtained. The solution (\underline{x}) of an inversion using regularization is:

$$\underline{x} = A^{-g}\underline{b}.$$
 (11)

where **b** and **d** are replaced by \underline{b} and \underline{d} , respectively, in Eq. (10). \underline{x} can also be obtained via least-squares or minimum-norm inversions using the objective function:

$$\min \left\| \begin{bmatrix} \boldsymbol{G} \\ \boldsymbol{C} \end{bmatrix} \boldsymbol{x} - \begin{bmatrix} \boldsymbol{d} \\ \boldsymbol{c} \end{bmatrix} \right\|^2.$$
(12)

When Tikhonov regularization ($C = \lambda L$, $c = \{0\}$) is used, a truncated form of A^{-g} ($m \times n$ matrix A^{-t}) can be obtained via (e.g., Aster et al. 2005; Barmin et al. 2001; Crosson 1976):

$$\boldsymbol{A}^{-t} = (\boldsymbol{G}^{\mathrm{T}}\boldsymbol{G} + \lambda^{2}\boldsymbol{L}^{\mathrm{T}}\boldsymbol{L})^{-1}\boldsymbol{G}^{\mathrm{T}}.$$
 (13)

x is then given by:

$$\underline{x} = A^{-t}\underline{d},\tag{14}$$

which is the same as that in Eq. (11). \underline{x} in Example 1 (Fig. 1a), which uses first-order Tikhonov regularization, is shown in Fig. 1c. An optimal factor λ of one is selected by Morozov's discrepancy principle (Morozov 1984) for the inversion. The resultant \underline{x} via Eq. (11), which is obtained using regularization, is generally more rational than \underline{x}_{D} . For example, \underline{x} (denoted as $\underline{x}(L^{1})$) in Example 1 (Fig. 1c) is closer to the synthetic model than \underline{x}_{D} , such that \underline{x} is preferred over \underline{x}_{D} (Eq. (5)) in practical

inversions (i.e., the solution provided in a practical ill-posed inversion is \underline{x} rather than \underline{x}_D). If the errors (δd) in the measured observation data

 \underline{d} (= $d + \delta d$) are known, then the solution \underline{x} in Eq. (11) becomes:

$$\underline{\boldsymbol{x}} = \boldsymbol{A}^{-g} \boldsymbol{b} + \delta \boldsymbol{x}_{\mathrm{d}}, \qquad (15)$$

where:

$$\delta \mathbf{x}_{d} = \mathbf{A}^{-g} \begin{bmatrix} \delta \mathbf{d} \\ \mathbf{0} \end{bmatrix} = \mathbf{A}^{-t} \delta \mathbf{d}.$$
(16)

It is noted that δx_d is only the contribution of the observational errors to the solution and is therefore not the solution error. The solution error or residual, $\delta x \ (= \underline{x} - \underline{x})$, is related to both δd and the other factors in the $x \rightarrow \underline{x}$ process.

2.3. Resolution Matrices from the Observations and Regularization Matrices

The process to obtain the solution (\underline{x}) of a true medium (x) is a projection $(r:x \rightarrow \underline{x})$ from x to \underline{x} , which can be described using Eq. (1) (Fig. 2a). If the projection is linear, then \underline{x} can be written as a linear regression equation (Fig. 2a):

$$\underline{\boldsymbol{x}} = \boldsymbol{R}\boldsymbol{x} + \delta\boldsymbol{x}_{\rm of},\tag{17}$$

where **R** is the slope of the regression, $\delta \mathbf{x}_{of}$ is a constant offset. If $\delta \mathbf{x}_{of}$ is ignored, then the linear regression becomes the linear projection of Eq. (2) (Fig. 2a), where **R** is the so-called resolution (Backus and Gilbert 1968, 1970) or projection matrix. For a nonlinear problem, either **R** or $\mathbf{R}: \mathbf{x} \to \mathbf{x}$ can be considered a linear approximation of $r: \mathbf{x} \to \mathbf{x}$ (Fig. 2a).

However, this matrix has been obtained via matrix operations, with several different resolution matrices potentially being obtained (An 2012). If the observational errors (δd) in data \underline{d} in Eq. (10) are ignored, then $\underline{d} = d$. Replacing \underline{d} in Eq. (5) with d in Eq. (6) causes the projection from x to \underline{x}_{D} ($x \rightarrow \underline{x}_{D}$) to become:



Figure 2

Illustration of resolution matrices **a** in a general study and **b** in a study with offset error. The model $x = [x_1]$ in the illustration constains one parameter (x_1) . A nonlinear relation $r:x \to \underline{x}$ between solution (\underline{x}) and true model (x) is approximated by linear projection $(\underline{x} = Rx, \text{Eq. }(2))$ and linear regression $(x = Rx + \delta x_{\text{off}}, \text{Eq. }(17))$. *R* is the slope of the linear relations

$$\underline{\boldsymbol{x}}_{\mathrm{D}} = \boldsymbol{R}_{\mathrm{D}}\boldsymbol{x},\tag{18}$$

where the resolution matrix R_D (Table 1) is of the following form (e.g., Jackson 1972; Menke 1989; Wiggins 1972):

$$\boldsymbol{R}_{\mathrm{D}} = \boldsymbol{G}^{-\mathrm{g}}\boldsymbol{G}.$$
 (19)

The transformation from x to \underline{x} ($x \rightarrow \underline{x}$) is obtained by inserting Eq. (9) into Eq. (11):

$$\underline{\boldsymbol{x}} = \boldsymbol{R}_{\mathrm{I}}\boldsymbol{x},\tag{20}$$

where resolution matrix R_{I} (Table 1) is of the form (An 2012):

$$\boldsymbol{R}_{\mathrm{I}} = \boldsymbol{A}^{-g} \boldsymbol{A}. \tag{21}$$

When the vector c is a zero vector (e.g., in Tikhonov regularization) and d in Eq. (11) is replaced by that in Eq. (6), the transformation from x to \underline{x} $(x \rightarrow \underline{x})$ is:

$$\underline{\boldsymbol{x}} = \boldsymbol{R}_{\mathrm{H}}\boldsymbol{x},\tag{22}$$

where the resolution matrix (\mathbf{R}_{H}) (Table 1) takes the form:

$$\boldsymbol{R}_{\mathrm{H}} = \boldsymbol{A}^{-g}\boldsymbol{B}, \quad \boldsymbol{B} = \begin{bmatrix} \boldsymbol{G} \\ \boldsymbol{0} \end{bmatrix},$$
 (23)

if the truncated form A^{-t} is used, then Eq. (23) becomes (An 2012):

$$\boldsymbol{R}_{\mathrm{H}} = \boldsymbol{A}^{-\mathrm{t}}\boldsymbol{G},\tag{24}$$

if A^{-t} is not truncated from A^{-g} and instead calculated from Eq. (13), which employs Tikhonov regularization of the form λL , then $R_{\rm H}$ becomes (e.g., Aster et al. 2005; Barmin et al. 2001; Boschi 2003; Crosson 1976):

$$\boldsymbol{R}_{\mathrm{H}} = (\boldsymbol{G}^{\mathrm{T}}\boldsymbol{G} + \lambda^{2}\boldsymbol{L}^{\mathrm{T}}\boldsymbol{L})^{-1}\boldsymbol{G}^{\mathrm{T}}\boldsymbol{G}.$$
 (25)

The three resolution matrices for Example 1, R_D and R_H in the inversion using the regularization matrix $I(R_H(I))$ and R_H in the inversion using the regularization matrix $L^1(R_H(L^1))$, are shown in Fig. 3a-c.

If *G* is a full-column rank matrix and no regularization is used, then \underline{x}_{D} equals \underline{x} , and R_{D} , R_{I} , and R_{H} are the same as the identity matrix. Otherwise, \underline{x}_{D} and \underline{x} are different, and the three resolution

				labl	e I	
				Resolution	matrices	
	Symbol	Problem	Condition	Projection	Equation	Significance
Direct	\pmb{R}_{D}	Linear	6	$x \to \overline{x}^{\mathrm{D}}$	(19)	Slope of linear relation $R: x \to \underline{x}$ if only G is used; observational
res. matrix Regularized	R_1	Inverse	в, С	$\overline{x} \uparrow x$	(21)	constrainability on \underline{x} or \underline{x}_D , upper bound of resolution from given G Should be <i>I</i> ; \underline{x} is unique for given G and C
Hybrid Hybrid Fas matrix	$R_{ m H}$		G, C	$\frac{x}{r}$	(23),	Slope of linear relation $R: x \to \underline{x}$ if G and C are used
Complete res. matrix	\pmb{R}_{C}	Linear or nonlinear	G or $G_{\rm J}, C, \underline{d}$, etc	$\frac{1}{x}$	(2)	Slope of Linear projection approximation of $r:x \to x$; same as $R_{\rm H}$ if only C and C are considered on D if only C is considered.
	$\boldsymbol{R}_{\mathrm{C+1}}$				(55)	O and C are considered, of AD if only O is considered Slope of Linear regression approximation of $r:x \to x$; same as R_C , but offset arrows are isolated
Statistic res. matrix Res. slope matrix at \underline{x}^{i-1} Res. slope matrix from \mathbf{r}^0 to	$\begin{matrix} \boldsymbol{R}_{\mathrm{S}} \\ \boldsymbol{R}_{\mathrm{J}}^{i} \\ \boldsymbol{R}_{\mathrm{J}}^{1-i} \end{matrix}$	Nonlinear inverse	G_1, C, x^{i-1} $G_2, C, \overline{x}^{i-1}$	$\Delta \mathbf{x}^i \to \Delta \mathbf{x}^i$	(2) (61), (62) (67)	Gauge the solution approximation of \mathbf{R}_{c} ; provides reliable resolution length Slope (Jacobian matrix) of $r:x \to x$ at \underline{x}^{r-1} Slope of $r:x \to x$ from x^{0} to x^{i}
was apply must we want we way			s, 0, 10		(10)	

matrices differ. Regularization needs to ensure that A is a full value of B is still as

On Resolution Matrices

A is a full-column rank matrix, such that \mathbf{R}_{I} is still an identity matrix. The practical solution is \underline{x} , as opposed to \underline{x}_{D} , due to regularization, and the resolution matrix for the $x \rightarrow \underline{x}$ projection is \mathbf{R}_{H} , as opposed to \mathbf{R}_{D} . \mathbf{R}_{D} has been paid little attention in previous regularized inversions in this case. However, \mathbf{R}_{D} is still very important for understanding the reliability of the solution in the inversions, as explained in the "Resolvability and constrainability from the resolution matrix" section.

Even though $R_{\rm D}$ and $R_{\rm H}$ are often different, they are both commonly called the model resolution matrix, which may confuse readers. The notations suggested by An (2012) for the matrices are adopted here for clarity, where $R_{\rm D}$ is the direct resolution matrix, $R_{\rm I}$ is the regularized resolution matrix, and $R_{\rm H}$ is the hybrid resolution matrix.

2.4. Properties of the Resolution Matrices

2.4.1 $R_{\rm D}$ from only the Observation Matrix

Truncated SVD of *G* allows R_D to be written in the form (e.g., Aster et al. 2005; Jackson 1972; Wiggins 1972):

$$\boldsymbol{R}_{\mathrm{D}} = \boldsymbol{V}_{\rho} \boldsymbol{V}_{\rho}^{\mathrm{T}}, \qquad (26)$$

where V_{ρ} (= { $v_{i,j}$ }_{$m \times \rho$}, ρ = rank(*G*)) is an unitary matrix with right singular vectors of *G*. Equation (26) indicates that $R_{\rm D}$ is a Gram matrix (or Gramian), which can be created by multiplying a matrix with its own transpose, as in Eq. (26).

The Gram matrix (e.g., Gentle 2007), $\mathbf{R}_{\rm D}$ (e.g., Fig. 3a), is symmetric (Table 2), with its rank (rank($\mathbf{R}_{\rm D}$)) equal to both its trace (trace($\mathbf{R}_{\rm D}$)) and the rank of \mathbf{V}_{ρ} (rank(\mathbf{V}_{ρ}) = ρ) (Eq. (26)). Here rank($\mathbf{R}_{\rm D}$) equals ρ because rank(\mathbf{V}_{ρ}) is equal to rank($\mathbf{G}(\rho)$). The diagonal elements in $\mathbf{R}_{\rm D}$ are nonnegative, but the off-diagonal elements in $\mathbf{R}_{\rm D}$ can be negative, with the exception that either $\mathbf{V}_{\rho}^{\rm T}$ is a fullcolumn rank matrix or \mathbf{V}_{ρ} is a full-row rank matrix. However, \mathbf{V}_{ρ} for an underdetermined problem is not a full-row rank matrix because ρ is smaller than m, such that the off-diagonal entries of $\mathbf{R}_{\rm D}$ of an underdetermined problem can be negative.





The direct resolution matrix $\mathbf{a} R_D$, and hybrid resolution matrices $\mathbf{b} R_H(I)$ and $\mathbf{c} R_H(L^1)$ for the inverse problem in Fig. 1a. A regularization parameter (λ) of one is used in the inversions. $R_H(I)$ and $R_H(L^1)$ are for the inversions that employ I and L^1 regularization matrices, respectively. **d**, **e** The 48th row and 48th column vectors of the matrices. The x_{48} parameter is constrained by the fifth observation, which overlaps with the sixth observation (Fig. 1a); the parameters from x_{40} to x_{60} are influenced by these two observations. The entries in row $r_{48,*}$ of R_D (**d**) are zeros, with the exceptions of $r_{48,41}-r_{48,55}$ (positive), and $r_{48,56}-r_{48,60}$ (negative)

A negative $r_{i,j}$ entry appears in R_D (Table 2) when some columns in two rows of G are like in a band matrix with the entries (0 and nonzero numbers a to d) below:

$$\begin{bmatrix} \dots & a & \dots & b & \dots & 0 & \dots \\ \dots & 0 & \dots & c & \dots & d & \dots \end{bmatrix}.$$
 (27)

Matrix	Inverse problem	Properties
R _D	Linear or linearized nonlinear	Gramian (e.g., symmetric, with nonnegative main-diagonal
$\boldsymbol{R}_{\mathrm{H}}(\lambda \boldsymbol{I})$	Linear	elements (but not the off-diagonal elements))
$\boldsymbol{R}_{\mathrm{JH}}^{i}(\lambda \boldsymbol{I})$	Linearized nonlinear	
$\boldsymbol{R}_{\mathrm{H}}(\lambda \boldsymbol{L}^{n})$	Linear	One-row-sum matrix; off-diagonal elements can be negative
$R_{\rm JH}^{i}(\lambda L^{n})$	Linearized nonlinear	(unlike stochastic matrix)

 Table 2

 Properties of typical resolution matrices

entries in G^{-g} related with the nonzero entries *a* and *d* in *G* have different signs from those in *G*, causing the negative $r_{i,j}$. This happens in *G* when \underline{x}_i and \underline{x}_j are constrained by different observations, which also constrain another parameter in common. For example, x_{48} and x_{58} are respectively constrained by the fifth and sixth observations (Fig. 1a), but the two observations also constrain the solution parameters $\underline{x}_{52}-\underline{x}_{55}$ in common. Consequently, $r_{48,58}$ in R_D is negative (Fig. 3a).

2.4.2 $\mathbf{R}_{\rm H}$ with Uniform Regularization Using $\lambda \mathbf{I}$

When uniform zeroth-order Tikhonov regularization $(\lambda L^0 \text{ or } \lambda I)$ is used, SVD of *G* allows the general inverse A^{-t} (Eq. (13)) to be written as below (e.g., Menke 2012):

$$\boldsymbol{A}^{-t} = \boldsymbol{V}(\boldsymbol{S}\boldsymbol{S} + \lambda^2 \boldsymbol{I})^{-1} \boldsymbol{S} \boldsymbol{U}^{\mathrm{T}}, \qquad (28)$$

where U is a unitary matrix with left singular vectors of G, S is a nonnegative diagonal matrix with singular values of G (s_i). The resolution matrix $R_{\rm H}(\lambda I)$ (Eq. (25)) can be written as:

$$\boldsymbol{R}_{\mathrm{H}}(\lambda \boldsymbol{I}) = \boldsymbol{V}(\boldsymbol{S}\boldsymbol{S} + \lambda^{2}\boldsymbol{I})^{-1}\boldsymbol{S}\boldsymbol{U}^{\mathrm{T}}\boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^{\mathrm{T}},$$

$$= \boldsymbol{V}(\boldsymbol{S}\boldsymbol{S} + \lambda^{2}\boldsymbol{I})^{-1}\boldsymbol{S}\boldsymbol{S}\boldsymbol{V}^{\mathrm{T}}$$

$$= \boldsymbol{V}\boldsymbol{F}\boldsymbol{V}^{\mathrm{T}} = \boldsymbol{V}_{\rho}\boldsymbol{F}_{\rho}\boldsymbol{V}_{\rho}^{\mathrm{T}}$$
(29)

where F_{ρ} (= diag($f_1, f_2, ..., f_{\rho}$)) (Aster et al. 2005) is truncated F and with the positive constants (f_i):

$$f_i = \frac{s_i^2}{s_i^2 + \lambda^2},\tag{30}$$

The positive diagonal matrix F_{ρ} can be written as a product of E_{ρ} (= diag($f_1^{1/2}, f_2^{1/2}, ..., f_{\rho}^{1/2}$)) and E_{ρ}^{T} . Equation (29) can then be written as:

$$\boldsymbol{R}_{\mathrm{H}}(\lambda \boldsymbol{I}) = \boldsymbol{V}_{\rho} \boldsymbol{E}_{\rho} \boldsymbol{E}_{\rho}^{\mathrm{T}} \boldsymbol{V}_{\rho}^{\mathrm{T}} = (\boldsymbol{V}_{\rho} \boldsymbol{E}_{\rho}) (\boldsymbol{V}_{\rho} \boldsymbol{E}_{\rho})^{\mathrm{T}}, \quad (31)$$

where:

$$\boldsymbol{V}_{\rho}\boldsymbol{E} = \{v_{i,j}f_{j}^{1/2}\}.$$
(32)

Equation (31) indicates that $\mathbf{R}_{\rm H}(\lambda \mathbf{I})$ is a Gram matrix (Table 2), like $\mathbf{R}_{\rm D}$. However, when a spatial variant regularization of the form λWI is used, the matrix $\mathbf{R}_{\rm H}(\lambda WI)$ cannot be written in the form of matrix multiplication with its own transpose, and is therefore not Gramian.

 $\mathbf{R}_{\rm H}(\lambda \mathbf{I})$ (e.g., Fig. 3b) is symmetric, like $\mathbf{R}_{\rm D}$. $R_{\rm H}(\lambda I)$ has a rank that is equal to rank $(V_{\rho}E_{\rho})$ (Eq. (31)). The diagonal entries of $\mathbf{R}_{\rm H}(\lambda \mathbf{I})$ are nonnegative, but the other entries can be negative. Equation (30) indicates that $f_i^{1/2}$ is in the range (0,1). Equation (32) indicates that a given column (e.g., the j^{th} column) of $V_{\rho}E_{\rho}$ equals the same column vector of V_{ρ} multiplied by $f_i^{1/2}$. All of the entries in $V_{\rho}E_{\rho}$ will be closer to zero than V_{ρ} , as $f_i^{1/2}$ is a positive value that is less than one. All of the entries in $R_{\rm H}(\lambda I)$ therefore have weaker intensities than those in $R_{\rm D}$, but they possess similar intensity patterns, as observed in comparisons of Fig. 3a and b, and Fig. 3d and e. Furthermore, trace($\mathbf{R}_{\rm H}(\lambda \mathbf{I})$) is smaller than trace(\mathbf{R}_{D}), such that the resolvability of x decreases after regularization using λI , as explained in the "Resolvability and constrainability from the resolution matrix" section.

2.4.3 **R**_H Using a Zero-Row-Sum Regularization Matrix

The matrices of derivative regularizations, e.g., the high-order Tikhonov regularization matrices λL^n and λWL^n (n > 0), are a zero-row-sum $n_s \times m$ matrix (S^0) ($n_s > 0$), such that

$$S^0 \mathbf{1} = \mathbf{0},\tag{33}$$

where $\mathbf{1} (= \{1\}_{m \times 1})$ is a vector with all elements one and $\mathbf{0}$ is a zero vector. Regardless of the number (n_s) of rows of S^0 , a relation below:

$$(\boldsymbol{G}^{\mathrm{T}}\boldsymbol{G} + \lambda^{2}(\boldsymbol{S}^{0})^{\mathrm{T}}\boldsymbol{S}^{0})\boldsymbol{1} = \boldsymbol{G}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{1} + \lambda^{2}(\boldsymbol{S}^{0})^{\mathrm{T}}(\boldsymbol{S}^{0}\boldsymbol{1})$$
$$= \boldsymbol{G}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{1} + \lambda^{2}(\boldsymbol{S}^{0})^{\mathrm{T}}\boldsymbol{0} \quad ,$$
$$= \boldsymbol{G}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{1}$$
(34)

exists. If the regularization matrix (C, Eq. (10)) is a zero-row-sum matrix (S^0), the Eq. (34) allows a relation on the resolution matrix $R_{\rm H}$ (Eqs. (24) and (25)):

$$\begin{aligned} \boldsymbol{R}_{\mathrm{H}} \mathbf{1} &= \boldsymbol{A}^{-\mathrm{t}} \boldsymbol{G} \mathbf{1} \\ &= (\boldsymbol{G}^{\mathrm{T}} \boldsymbol{G} + \lambda^{2} (\boldsymbol{S}^{0})^{\mathrm{T}} \boldsymbol{S}^{0})^{-1} \boldsymbol{G}^{\mathrm{T}} \boldsymbol{G} \mathbf{1} \\ &= (\boldsymbol{G}^{\mathrm{T}} \boldsymbol{G} + \lambda^{2} (\boldsymbol{S}^{0})^{\mathrm{T}} \boldsymbol{S}^{0})^{-1} (\boldsymbol{G}^{\mathrm{T}} \boldsymbol{G} + \lambda^{2} (\boldsymbol{S}^{0})^{\mathrm{T}} \boldsymbol{S}^{0}) \mathbf{1}^{'} \\ &= \mathbf{1} \end{aligned}$$

$$(35)$$

therefore, when regularization matrix is a zero-rowsum matrix (S^0), the resolution matrix R_H is a onerow-sum matrix S^1 (while $S^1\mathbf{1} = \mathbf{1}$). This can be simplified as:

$$\begin{bmatrix} \boldsymbol{G} \\ \boldsymbol{S}^0 \end{bmatrix}^{-g} \begin{bmatrix} \boldsymbol{G} \\ \boldsymbol{0} \end{bmatrix} = \begin{bmatrix} \boldsymbol{G} \\ \boldsymbol{S}^0 \end{bmatrix}^{-t} \boldsymbol{G} = \boldsymbol{S}^1, \qquad (36)$$

where **0** is a zero matrix.

As the high-order Tikhonov regularization matrices λL^n and λWL^n (n > 0) (e.g., flatness and smoothness) are both a zero-row-sum S^0 matrix. Therefore, $R_H(\lambda L^n)$ and $R_H(\lambda WL^n)$ are S^1 matrices (i.e., all of the rows in $\mathbf{R}_{\mathrm{H}}(\lambda \mathbf{L}^{n})$ and $\mathbf{R}_{\mathrm{H}}(\lambda \mathbf{WL}^{n})$ sum to one; Table 2). All of the rows in $\mathbf{R}_{\mathrm{H}}(\mathbf{L}^{1})$ in Fig. 3c sum to one, as shown in Fig. 4a.

The resolution matrix $R_{\rm H}(S^0)$ is similar to a stochastic matrix (a square matrix with non-negative elements and each row summing to 1), but the entries in $\mathbf{R}_{\rm H}$ (or \mathbf{S}^{1}) can be negative due to the same reason above for $R_{\rm D}$. S^0 matrix (is often band matrix) or the combined matrix of S^0 with G (Eq. (36) or A in Eq. (10)) often includes two rows like Eq. (27), i.e., the S^1 regularization matrix always yields two parameters (e.g., x_1 and x_{11}) that are constrained by the same observation (the first row of G), with one of them (x_{11}) also being constrained by one of the other observation or regularization row (the tenth row of L^1 in Eq. (S2)). The two rows also constrain a third parameter (x_{10}) . The entries in A^{-t} and A^{-g} related with parameters often have different signs from G, and the $r_{1,11}$ entry in $\mathbf{R}_{\rm H}(\mathbf{L}^1)$ (Fig. 3c) is consequently negative. Therefore, the application of L^n regularization can cause more negative entries in $R_{\rm H}(L^n)$ (e.g., Fig. 3c) than R_D (Fig. 3a), and then R_H is not a stochastic matrix.

A one-row-sum matrix $R_{\rm H}(S^1)$ implies that one is an eigenvalue of the projection $R_{\rm H}$, such that there exists an equation:

$$\boldsymbol{R}_{\mathrm{H}}\boldsymbol{x}_{\mathrm{p}} = 1\boldsymbol{x}_{\mathrm{p}},\tag{37}$$

Eqs. (22) and (37) indicate that a medium (\mathbf{x}_p) can be fully resolved and equal to a solution $(\underline{\mathbf{x}}_p = \mathbf{R}_H \mathbf{x}_p)$, even though \mathbf{x}_p is not the medium that is currently measured. It is impossible for all of the row sums in $\mathbf{R}_H(\lambda I)$ to equal one, thereby demonstrating that



Figure 4 **a** Row-vector sums and **b** main-diagonal elements of the resolution matrices $R_{\rm D}$ and $R_{\rm H}$ in Fig. 3

higher-order Tikhonov regularization is superior to damping (or λI) regularization from the viewpoint of row sums.

2.4.4 **R**_H Using Mixed Regularization

If the regularization is mixed using higher-order regularizations (e.g., L^2 and L^3), then the new regularization matrix is still a matrix S^0 , and R_H is a matrix S^1 . Otherwise, if it is mixed using damping and a higher-order L^n (n > 0) (e.g., Sigloch 2011; Tewarson 1977), then the mixed regularization matrix C is neither a diagonal matrix nor a zero-row-sum matrix. Therefore, the new R_H is neither a symmetric (Gram) matrix nor a one-row-sum matrix.

2.5. Significances of the Resolution Matrices

2.5.1 Row Vector = Content Function of the Medium

Equation (2) highlights that a solution parameter \underline{x}_i can be expressed as a weighted sum of all of the model parameters in \mathbf{x} (or $[x_1, x_2, ..., x_m]^T$):

$$\underline{x}_i = \sum_{j=1}^m r_{i,j} x_j, \tag{38}$$

where the $r_{i,j}$ entry of matrix **R** plays a role in weighting the jth model parameter x_j in the summation. The $r_{i,*}$ row vector of **R** therefore acts like an averaging vector (Backus and Gilbert 1968), with the entries in $r_{i,*}$ representing the accurate contents (or contributions) of all of the medium parameters to (or in) the ith solution parameter x_i . Therefore, $r_{i,*}$ can

also be termed the content (or contribution) function of the medium in x_i (Table 3).

2.5.2 Column Vector = Spreading Function of the Medium

The $r_{i,j}$ entry signifies the contribution of the medium parameter x_j to \underline{x}_i , such that all of the entries in the j^{th} column vector ($r_{*,j}$) (Table 3) correspond to the contributions of the j^{th} model parameter (x_j) to either all of the parameters in \underline{x} or the spread of x_j into the parameters. This has been considered the Green's function (also termed the point spread function (Smith 1997), or impulse response function) of x_j to \underline{x} .

2.5.3 Significances of the Matrices

The matrices $\mathbf{R}_{\rm D}$, $\mathbf{R}_{\rm I}$, and $\mathbf{R}_{\rm H}$ are slopes of the linear projection from \mathbf{x} to \mathbf{x} , but have different significances (Table 1) on the transformation. Regularization at least makes matrix \mathbf{A} full rank (invertible); i.e., $\mathbf{R}_{\rm I}$ should be an identity matrix. If the $\mathbf{x} \rightarrow \mathbf{x}$ projection matrix $\mathbf{R}_{\rm I}$ is an identity matrix, then a unique solution (\mathbf{x}) can be obtained from a given \mathbf{G} and \mathbf{C} . Otherwise, some of the model parameters remain poorly constrained, and further regularization should be employed.

 $R_{\rm D}$ is only produced from G (Eq. (19)). Therefore, $R_{\rm D}$ represents the effects and contributions of <u>x</u> from the given observations, regardless of whether regularization is used or not.

 $R_{\rm H}$ only reflects the $x \to \underline{x}$ projection, with no consideration of other factors (e.g., errors) in the

Vectors of resolution matrix									
Significan	nce			Reliable resolution length	More				
Vector	ctor Elem				sensitive				
Row, <i>r</i> _{<i>i</i>,*}	Content or contribution function	$r_{i,i}$ $\Sigma r_{i,*}$	x_i resolvability > 1: overestimated \underline{x}_i < 1: underestimated \underline{x}_i	Stable row vectors	Observation				
Column, $r_{*,j}$	Spreading function	$ r_{i,i} - r_{i,ii} $	$\sim \underline{x}_i - \underline{x}_{ii} $	Observation-constrained Parameters	Regularization				

Table 3

M. An

matrix. $\mathbf{R}_{\rm H}$ can therefore evaluate the reliability of $\underline{\mathbf{x}}$. Furthermore, the construction of $\mathbf{R}_{\rm H}$ as a mixture of the observations (\mathbf{G}) and regularization (λC in \mathbf{A}) (Eq. (23)) means that it reflects some combination of the observational and regularization effects in the solution. The variations or differences between $\mathbf{R}_{\rm H}$ and $\mathbf{R}_{\rm D}$ therefore represent the effects of regularization on the inversion, as $\mathbf{R}_{\rm D}$ only reflects the observational effects on the solution.

All three matrices, especially $R_{\rm H}$ and $R_{\rm D}$, are therefore essential for understanding the inversion and its result, such as the resolvability of x, the uniqueness and reliability of \underline{x} , and the effects of regularization.

2.5.4 Column Vector Variations Due to Regularization Changes

The role of regularization on the projection from x to \underline{x} can be revealed via a comparison of the resolution matrices $R_{\rm D}$ and $R_{\rm H}$. One entry $(r_{i,j})$ of either $R_{\rm D}$ in Eq. (19) or $R_{\rm H}$ in Eq. (24) can be written as:

$$r_{i,j} = \sum_{k} u_{i,k} g_{k,j}, \tag{39}$$

where $u_{i,j}$ represents an entry of either G^{-g} or A^{-t} , and $g_{k,j}$ is an entry of G. $R_{\rm H}$ (Eq. (24)) can be considered $R_{\rm D}$ (Eq. (19)), with the variation in $u_{i,j}$ obtained by replacing G^{-g} with A^{-t} . Equation (39) indicates that if the $g_{*,j}$ column vector of G is given, then the variation of $u_{i,j}$ only influences the $j^{\rm th}$ column vector ($r_{*,j}$) of R. Therefore, the addition of regularization (C) in A (Eq. (10)) allows the $j^{\rm th}$ column vector in $R_{\rm H}$ to be considered a function of the $j^{\rm th}$ column (with no relationship arising among the other columns) in $R_{\rm D}$. Regularization essentially changes the spread functions from $R_{\rm D}$ to $R_{\rm H}$, such that the spread function is sensitive to regularization (Table 3).

These column vector variations due to regularization changes are well-illustrated by comparing $\mathbf{R}_{\rm D}$ and $\mathbf{R}_{\rm H}(\mathbf{L}^1)$ (Fig. 5c and d, respectively) of a linear inverse problem example (Example 2) with threepoint observations (\mathbf{G} in Fig. 5a) and 50 unknowns in \mathbf{x} (Fig. 5a and b). The \mathbf{L}^1 regularization matrix and $\lambda = 1$ are used in the inversion. If a column in $\mathbf{R}_{\rm D}$ (e.g., $\mathbf{r}_{*,5}$) is all zeros (Fig. 5c), then the corresponding column vector in $\mathbf{R}_{\rm H}$ ($\mathbf{r}_{*,5}$) (Fig. 5d) will be all zeros. Otherwise, the tenth column in $\mathbf{R}_{\rm D}$ has one nonzero entry ($r_{10,10}$), and the tenth column vector in $\mathbf{R}_{\rm H}(\mathbf{L}^1)$ ($\mathbf{r}_{*,10}$) (Fig. 5c and d) has nonzero entries around the $r_{10,10}$ entry. The variations between the column in $\mathbf{R}_{\rm D}$ and that in $\mathbf{R}_{\rm H}$ are due to regularization. Similar results can be found via a comparison of $\mathbf{R}_{\rm H}(\mathbf{L}^1)$ and $\mathbf{R}_{\rm D}$ (Fig. 3) for Example 1 (Fig. 1).

Comparisons of the column vectors of $\mathbf{R}_{\rm D}$ and $\mathbf{R}_{\rm H}$ for the same parameter x_j can reveal the regularization effect, as the variations in the spread functions from $\mathbf{R}_{\rm D}$ to $\mathbf{R}_{\rm H}$ are due to regularization. Here, the *j*th column vector in $\mathbf{R}_{\rm H}$ is only related to the *j*th column in $\mathbf{R}_{\rm D}$, such that the relationship between the relative magnitudes of neighboring entries in a row vector of $\mathbf{R}_{\rm D}$ may be preserved in $\mathbf{R}_{\rm H}$. A given row vector in $\mathbf{R}_{\rm H}$ may exhibit a similar pattern or curve to that in $\mathbf{R}_{\rm D}$. For example, the curves of the $\mathbf{r}_{48,*}$ row vectors in $\mathbf{R}_{\rm D}$ and $\mathbf{R}_{\rm H}$ (Fig. 3d) are similar, but those of the $\mathbf{r}_{*,48}$ column vectors (Fig. 3e) are very different.

2.6. Do the Projection Matrices Represent Practical Projections?

The projection $r:x \to \underline{x}$ (Eq. (1)) includes the effects of all of the factors (e.g., uncertainty in observation *d* and prior information *c* (Menke 2015), Eq. (12)) in the process from *x* to \underline{x} , but R_D , R_I , and R_H do not. The resolution estimated from R_H can be unrealistically higher than that obtained via synthetic tests (Pilkington 2016), which may be due to the limitation of the projection matrix.

2.6.1 R_H Versus Observational Errors

The observational errors δd influence the solution <u>x</u>. When δd is considered, Eq. (2) becomes:

$$\underline{x} = R(\delta d)x, \tag{40}$$

where $R(\delta d)$ represents R as a function of δd , with $R(\delta d = 0)$ equaling R for the error-free case.

Inserting Eq. (6) into Eq. (15) then yields a projection that is similar to the form in Eq. (17):

$$\underline{\boldsymbol{x}} = \boldsymbol{R}_{\mathrm{H}}\boldsymbol{x} + \delta\boldsymbol{x}_{\mathrm{d}},\tag{41}$$

b)



Figure 5

Resolution matrices for linear inverse problem Example 2. **a** The observation matrix **G**. The observations consist of only three points at parameters x_{10} , x_{30} , and x_{32} . Only one entry in each row in **G** is nonzero. **b** A synthetic model and its corresponding solution, which was constrained using first-order Tikhonov regularization and $\lambda = 1$. **c**, **d** Resolution matrices R_D and R_H . R_D has only three nonzero entries. R_H has more nonzero entries, but they are in the same columns (10, 30, and 32) as those in R_D

Eq. (41) shows that $\mathbf{R}_{\rm H}$ is independent of $\delta \mathbf{x}_{\rm d}$, which represents the effect of $\delta \mathbf{d}$ in the solution. However, the error ranges can be used to weight the data in an inversion that employs a weighted least squares (WSL) approach; $\mathbf{R}_{\rm H}$ determined via WSL inversion is slightly different to the above $\mathbf{R}_{\rm H}$, but it is also not $\mathbf{R}(\delta \mathbf{d})$. Therefore, $\mathbf{R}_{\rm H}$ does not include the effect of observational errors.

2.6.2 Resolution Matrix of the Full Process from x to x?

Is there a matrix that reflects all of the factors in the $x \rightarrow \underline{x}$ process? If \underline{x} is the result of a process that incorporates all of the factors, then the resolution matrix R that is directly inverted from x and \underline{x} via Eq. (2) reflects all of the factors. However, the inversion scheme to determine such a matrix is often impossible in geophysics. The true model for a region (x) may be never known; otherwise, the matrix R for a region with known x is redundant and unnecessary.

Pure Appl. Geophys.

Equation (2) represents a transformation from the real model x to the corresponding solution \underline{x} , or a process via the projection $r:x \to \underline{x}$ (or R). This approach is independent of x and can be isolated from the practical true model x and practical solution \underline{x} . If x is not the true medium but rather a synthetic model, then the inversion of R is possible.

Recovery tests, e.g., checkerboard tests (Lévěque et al. 1993), that employ a synthetic model with a specific structure are frequently used to retrieve a qualitative resolution. The output solution \underline{x} of a synthetic test that employs an input model with a random structure x also contains resolution information (An 2012). An (2012) suggested that the statistical resolution matrix R_S (Table 1) can be determined by statistically comparing a limited number of input synthetic random models x and their An accurate resolution matrix that includes all of the factors in the complete $x \rightarrow \underline{x}$ projection process is suggested in this section.

3.1. Method

Equation (2) has to be reorganized for the inversion of \mathbf{R} because \mathbf{R} is an $m \times m$ matrix, and both \mathbf{x} and $\underline{\mathbf{x}}$ are $m \times 1$ vectors. The extended form of Eq. (2) is:

$$\begin{cases} \underline{x}_{1} = r_{1,1}x_{1} + r_{1,2}x_{2} + \cdots + r_{1,m}x_{m} \\ \underline{x}_{2} = r_{2,1}x_{1} + r_{2,2}x_{2} + \cdots + r_{2,m}x_{m} \\ \vdots \\ \underline{x}_{m} = r_{m,1}x_{1} + r_{m,2}x_{2} + \cdots + r_{m,m}x_{m} \end{cases}$$
(42)

where $r_{i,j}$ is the element at the *i*th row and *j*th column of **R**. The equation is reorganized as:

$$\begin{cases} \underline{x}_{1} = x_{1}r_{1,1} + x_{2}r_{1,2} + \cdots + x_{m}r_{1,m} \\ \underline{x}_{2} = & x_{1}r_{2,1} + \cdots + x_{m}r_{2,m} \\ \vdots & & \ddots \\ \underline{x}_{m} = & x_{1}r_{m,1} + \cdots + x_{m}r_{m,m} \end{cases}$$
(43)

correspondent output solutions \underline{x} via a Gaussian function approximation for each row vector.

The output solution \underline{x} includes all of the known factors in the $x \rightarrow \underline{x}$ process, with R_S including all of these factors. However, the matrix is only approximate and not necessarily accurate. If a large number of x and \underline{x} are given, then an accurate and complete resolution matrix can be obtained on the basis of Eq. (2), as discussed in the following section.

3. Resolution Matrices of the Complete Process

The compact form of Eq. (43) is:

$$\underline{x} = Xr, \tag{44}$$

where X is a band matrix that is composed of x^{T} vectors and r is a vectorization of R^{T} :

$$\boldsymbol{X} = \begin{bmatrix} \boldsymbol{x}^{\mathrm{T}} & & \boldsymbol{0} \\ & \boldsymbol{x}^{\mathrm{T}} & & \\ & & \ddots & \\ \boldsymbol{0} & & & \boldsymbol{x}^{\mathrm{T}} \end{bmatrix}_{m \times m^{2}}$$
(45)

and:

$$\mathbf{r} = \operatorname{vec}(\mathbf{R}^{\mathrm{T}}) = \begin{bmatrix} r_{1,1} & r_{1,2} & \cdots & r_{1,m} & r_{2,1} & \cdots & r_{2,m} & \cdots & r_{m,1} & \cdots & r_{m,m} \end{bmatrix}^{\mathrm{T}}.$$
 (46)

Application of the projection to a random synthetic model \mathbf{x}^k outputs the corresponding solution $\underline{\mathbf{x}}^k$. The model \mathbf{X}^k (a band matrix for \mathbf{x}^k , Eq. (45)) and $\underline{\mathbf{x}}^k$ still satisfy Eq. (44). One can therefore obtain Nsolutions ($\underline{\mathbf{x}}^1$, $\underline{\mathbf{x}}^2$, ..., $\underline{\mathbf{x}}^N$) by performing the same projection for N different random synthetic models (\mathbf{X}^1 , \mathbf{X}^2 ,..., \mathbf{X}^N), and then all of the solutions can be used to construct a new equation from N equations, following Eq. (44):

$$\{\underline{\boldsymbol{x}}\} = \{\boldsymbol{X}\}\boldsymbol{r},\tag{47}$$

where:

$$\{\underline{x}\} = \begin{bmatrix} \underline{x}^{1} \\ \underline{x}^{2} \\ \vdots \\ \underline{x}^{N} \end{bmatrix} \text{ and } \{X\} = \begin{bmatrix} X^{1} \\ X^{2} \\ \vdots \\ X^{N} \end{bmatrix}, \quad (48)$$

the extended form of Eq. (47) is shown in Eq. (S4) in the supporting information. One synthetic model (X^k) produces *m* equations, as outlined in Eq. (43), with Eq. (47) including $N \times m$ equations for *N* synthetic models. If Eq. (47) is constructed of m^2 independent equations, then *r* will be uniquely resolvable via:

$$\boldsymbol{r} = \{\boldsymbol{X}\}^{-1}\{\underline{\boldsymbol{x}}\},\tag{49}$$

a resolution matrix R is then obtained by converting the vector r back into a matrix.

The new matrix **R** is obtained via either Eq. (49) or (2) without any approximation. The synthetic solutions (\underline{x}^k) are the result of the complete $x \to \underline{x}$ process with all of the factors. Therefore, the resultant **R** reflects all of the factors (various errors, simplification, etc.) in the complete process, and is termed the complete resolution matrix, which is denoted **R**_C (Table 1).

This extraction of $\mathbf{R}_{\rm C}$ from random synthetic input models and output solutions is similar to that proposed by An (2012). An (2012) focused on the extraction of a reliable resolution length from a small number of \mathbf{x}^i and \underline{x}^i to construct the approximate resolution matrix $\mathbf{R}_{\rm S}$ (Table 1). Here, an accurate resolution matrix ($\mathbf{R}_{\rm C}$) is directly inverted without approximation. Various procedures, such as linear and nonlinear inverse problems—and also noninverse problems (An 2012)—can be implemented to obtain $\mathbf{R}_{\rm C}$ (Fig. 6) and $\mathbf{R}_{\rm S}$, as they are isolated from the $\mathbf{x} \to \mathbf{x}$ process. For example, $\mathbf{R}_{\rm S}$ can be obtained for kriging and minimum curvature gridding (Chiao et al. 2014). The extraction of $\mathbf{R}_{\rm C}$ by Eq. (47) is a linear regression of the relationship between \mathbf{x}^k and \mathbf{x}^k (Fig. 2a), such that $\mathbf{R}_{\rm C}$ can then be considered a linear approximation of $r:\mathbf{x} \to \mathbf{x}$ for either the nonlinear inverse or non-inverse problem.

3.2. Equation Simplification

The ability to resolve $\mathbf{R}_{\rm C}$ via Eq. (49) requires the inversion of a large matrix with $\geq m^2$ rows and m^2 columns. However, the calculation can be simplified.

Equations (42) and (43) state that the *i*th solution parameter \underline{x}_i^k of the solution \underline{x}^k can be written as:

$$\underline{x}_{i}^{k} = \boldsymbol{R}_{i} \boldsymbol{x}^{k} = (\boldsymbol{x}^{k})^{\mathrm{T}} \boldsymbol{R}_{i}^{\mathrm{T}}, \qquad (50)$$

where \mathbf{R}_i is the *i*th row of \mathbf{R} ($\mathbf{r}_{i,*}$). All of the equations for the parameters from x_i^1 to x_i^N form:

$$\{\underline{x}_i\} = \{\boldsymbol{x}^{\mathrm{T}}\}\boldsymbol{R}_i^{\mathrm{T}},\tag{51}$$

where:

$$\{\underline{x}_i\} = \begin{bmatrix} \underline{x}_i^1 \\ \underline{x}_i^2 \\ \vdots \\ \underline{x}_i^N \end{bmatrix} \text{ and } \{\mathbf{x}^T\} = \begin{bmatrix} (\mathbf{x}^1)^T \\ (\mathbf{x}^2)^T \\ \vdots \\ (\mathbf{x}^N)^T \end{bmatrix}.$$
(52)

Equation (51) is actually the same as Eq. (S4), but it only contains the rows for the solution parameter \underline{x}_i . Equation (51) can then be used to invert the *i*th row of $R_C(R_i)$ from:

$$\boldsymbol{R}_{i}^{\mathrm{T}} = \{\boldsymbol{x}^{\mathrm{T}}\}^{-1}\{\underline{x}_{i}\}, \qquad (53)$$

the other rows of $R_{\rm C}$ can also be inverted using Eq. (53).

The inversion of $\mathbf{R}_{\rm C}$ via Eq. (53) is easier than that via Eq. (49). Equation (49) has an (left) inverse of the matrix {X} with $N \times m$ rows and m^2 columns. However, Eq. (53) has an (left) inverse of the much smaller $(N \times m)$ matrix { $\mathbf{x}^{\rm T}$ }. Furthermore, the inverse of $\{\boldsymbol{x}^{\mathrm{T}}\}$ for $\boldsymbol{R}_{i}^{\mathrm{T}}$ in Eq. (53) is the same for the calculation of the other rows (e.g., $\boldsymbol{R}_{j}^{\mathrm{T}}$), and is then directly used in the calculation for all of the rows of $\boldsymbol{R}_{\mathrm{C}}$.

We derived $\mathbf{R}_{\rm C}$ for Example 1 (Fig. 1) using the same regularization to obtain $\mathbf{R}_{\rm H}(\mathbf{L}^1)$ (Fig. 3c) without considering observation errors ($\delta \mathbf{d}$) for comparison. The solutions (e.g., $\underline{\mathbf{x}}^k$) for 100 (N = 100) different synthetic random models (e.g., \mathbf{x}^k) were resolved in step 1 (Fig. 6) of the calculation. The resultant $\mathbf{R}_{\rm C}$ (Fig. 7b) is the same as $\mathbf{R}_{\rm H}$ (Fig. 3c) for the linear inverse problem. However, $\mathbf{R}_{\rm C}$ can reflect all of the factors in a practical study, whereas $\mathbf{R}_{\rm H}$ only reflects the effects of \mathbf{G} and regularization.

3.3. Resolution Matrices with Error Effects

All of the measurements and processing include errors which influence the solution \underline{x} and then the $x \rightarrow \underline{x}$ projection. R_C can include the effects of various (quantifiable and unquantifiable) errors and additional prior information (such as C and c in Eq. (10)). For example, system simplification may cause error in solution but is often difficult to be quantified. If \underline{x}^k is obtained through the simplification, the resulted R_C will reflect the error related with the simplification. However, for the sake of comparison, examples on the effects of quantified errors in data are given below.

When the observational errors δd are considered, the equation with $R(\delta d)$ (Eq. (40)) is of the same form as Eq. (2). Therefore, the above method (Fig. 6) for obtaining the complete resolution matrix R_C via either Eq. (49) or (53) on the basis of Eq. (2) can be used to obtain $R(\delta d)$ (Eq. (40)). One pair of models (x^k, \underline{x}^k) serves as an independent measurement for R in this method. However, the addition of more factors in the processing than those contained in G and C make the relationship between x and \underline{x} more complex, such that a larger number (> m) of model pairs is often necessary to obtain a reliable R_C .

The traditional resolution matrix $R_{\rm H}$ (Eq. (41)) cannot include the effect of observation error, then equals R(0) or $R(\delta d = 0)$. The resolution matrix ($R_{\rm C}$ or $R(\delta d)$) with the effect of observation error is different from $R_{\rm H}$. An equation with their difference can be obtained from Eqs. (40) and (41):

$$\delta \underline{x} = (R(\delta d) - R_{\rm H})x, \qquad (54)$$

Eq. (54) indicates the difference $(\mathbf{R}(\delta d) - \mathbf{R}_{\rm H})$ or $(\mathbf{R}(\delta d) - \mathbf{R}(\mathbf{0}))$ is a projection matrix on the solution error $(\delta \underline{x})$.



Figure 6

Flowchart for obtaining the complete resolution matrix (\mathbf{R}_{C}) from synthetic random models and solutions. \mathbf{r} is the vector form of \mathbf{R}_{C}





Resolution matrix $\mathbf{R}_{\rm C}$ for the 1-D inverse problem in Fig. 1. **a** Example synthetic (true) random model (input model) and its corresponding solution (output model) for the inverse problem. **b** Matrix $\mathbf{R}_{\rm C}$ or \mathbf{R}_{-of} . The resolution matrix $\mathbf{R}_{\rm C}$ is calculated from 100 pairs of random input and output models, as in (**a**). Flatness regularization, with $\lambda = 1.0$ used. \mathbf{R}_{-of} is the resolution matrix after the offset errors are isolated (Eq. (55)). **c**, **d** Two row vectors and **e**, **f** two column vectors of $\mathbf{R}_{\rm C}$ in (**a**). Red lines in (**c**-**f**) are Gaussian approximations of the vectors around the diagonal entry. The row and column vectors for the 48th parameter can be represented by Gaussian function curves (**c** and **e**, respectively), whereas those for the 95th parameter are far from Gaussian curves (**d** and **f**, respectively). The column vector ($\mathbf{r}_{*,95}$) (**f**) is all zeros. The widths of the Gaussian approximation curves (**c**-**e**) can represent the widths of the vector curves

Random errors exist in practical studies. The influence of random errors on the solution diminishes as the number of observations increases. However, if observations are limited, then the average effect of the random errors appears as a regular systematic error; we therefore only test systematic errors here. A stronger regularization (larger λ) is normally required for inversions with observational errors. However, we used the same λ that was applied in the above error-free inversion for comparison.

Two main types of systematic errors, offset and scale factor errors, are tested here. Offset observational errors in a linear inversion will introduce offset errors in \underline{x} (Fig. 2b). The $x \to \underline{x}$ process with offset errors in \underline{x} can be better represented by Eq. (17) than Eq. (40). Equation (17) has more variables (\mathbf{R} and δx_{of}) than Eq. (40), such that the process with offset errors is more complex and requires a larger number of model pairs to obtain $\mathbf{R}(\delta d)$. If we still invert for $\mathbf{R}(\delta d)$ via Eq. (40) using the same number (N = 100) of model pairs (x^k , \underline{x}^k) in Fig. 7, then the resultant



Figure 8

a Row and **b** column vectors of the resolution matrices for a process that incorporates the offset observational errors calculated via Eq. (40) using 100 (N = 100) and 200 (N = 200) model pairs. The row vector of the matrix obtained from 100 model pairs is unstable in the row vector (**a**). The column vector (**b**) is less influenced by the errors because the L^1 regularization used here mainly affects the column vectors via a smoothing process, whereas the row vectors are minimally affected

matrix is somewhat unstable (Fig. 8a). The column vector is somewhat stable (Fig. 8b) due to regularization (L^1) because flatness regularization directly influences the column vectors, but not the row vectors. $R(\delta d)$ generally becomes stable when a larger number (e.g., N = 200) of model pairs (Fig. 8a) is used.

Scale factor observational errors in a linear inversion will introduce scale factor errors in \underline{x} . The $x \rightarrow \underline{x}$ process with scale factor errors in \underline{x} can be well represented by Eq. (40), with $R(\delta d)$ equaling $R_{\rm H}$ multiplied by the factors related to the scale factor errors.

An offset error of 2.0 s ($\delta d_1 = 2.0$ s) and a scale factor error of 20% ($\delta d_1 = 0.2d_1 = \sim 2.5$ s) in the first observation data (d_1), which have a mean of ~ 12.5 s for all of the synthetic observations, are considered. Two hundred (N = 200) model pairs are used, with the resultant matrices $R(\delta d)$ shown in Figure S1. The matrix differences, $R(\delta d) - R_H(L^1)$, which reflect the observational error effects in the projection from **x** to **x**, are shown in Fig. 9c and d.

These two types of observational errors yielded similar errors (δx_d) in the solution parameters $(\underline{x}_{1}-\underline{x}_{25})$ in $\underline{x}(L^1)$ (Fig. 9a, b), resulting in similar summation curves for the 1st-25th content vectors in $R(\delta d)$ (Fig. 9e, f). However, the errors produced different effects in the projection matrices $R(\delta d)$. The row vector of $R(\delta d)$ for a solution parameter with errors (e.g., \underline{x}_5) in Fig. 9c is very different from that in Fig. 9d. The addition of the scale factor error δd_1 (= 0.2 d_1) only caused resolution matrix changes in the 1st-10th columns (Fig. 9d) (i.e., only the contributions of x_1 - x_{10} , which are constrained by \underline{d}_1 , are influenced, as expected). However, the offset error δd_1 (= 2.0) caused changes in all of the columns (Fig. 9c). Therefore, the spread functions are sensitive to different types of errors.

3.4. Offset Error Isolation

The $\mathbf{x} \to \underline{\mathbf{x}}$ process with offset errors in $\underline{\mathbf{x}}$ ($\delta \mathbf{x}$) (Fig. 2b) can be better represented via linear regression (Eq. (17)), whereby the offset errors are isolated from either $\underline{\mathbf{x}}$ or $\mathbf{R}(\delta d)\mathbf{x}$. When the offset errors $\delta \mathbf{x}_{of}$ (= $[o_1, o_2, ..., o_m]^T$) are isolated, the obtained resolution matrix is denoted as either \mathbf{R}_{-of} or $\mathbf{R}_{-of}(\delta d)$. Equation (17) can be written as:

$$\underline{\mathbf{x}} = \mathbf{R}_{-\mathrm{of}}(\delta d)\mathbf{x} + \delta \mathbf{x}_{\mathrm{of}} = \mathbf{R}_{+1}\mathbf{x}_{+1}, \qquad (55)$$

where $\mathbf{R}_{+1} = [\mathbf{R}_{-\text{of}} \delta \mathbf{x}_{\text{of}}], \mathbf{x}_{+1} = [\mathbf{x}^{T} \ 1]^{T}$. The extended form of Eq. (55) is:

$$\begin{bmatrix} \underline{x}_{1} \\ \underline{x}_{2} \\ \vdots \\ \underline{x}_{m} \end{bmatrix} = \begin{bmatrix} r_{1,1} & r_{1,2} & \cdots & r_{1,m} & o_{1} \\ r_{2,1} & r_{2,2} & & r_{2,m} & o_{2} \\ \vdots & & \ddots & & \vdots \\ r_{m,1} & r_{m,2} & \cdots & r_{m,m} & o_{m} \end{bmatrix} \begin{bmatrix} x_{1} \\ x_{2} \\ \vdots \\ x_{m} \\ 1 \end{bmatrix}.$$
(56)

Equation (55) is of the same form as Eq. (17), such that the $m \times (m + 1)$ matrix \mathbf{R}_{+1} can be inverted using model pairs (e.g., $\underline{\mathbf{x}}^k$ and \mathbf{x}^k) via the above procedure for \mathbf{R}_C (Fig. 6). The obtained \mathbf{R}_{+1} is denoted as \mathbf{R}_{C+1} (Table 1). As expected, the resultant matrix $\mathbf{R}_{-of}(\delta d)$ for the above case with offset errors is



Figure 9

Complete resolution matrices for processes considering (*left*) a 2.0 offset error and (*right*) 20% scale factor error in the first observation data (\underline{d}_1) for the inverse problem in Fig. 1a. **a**, **b** Example synthetic model and corresponding solution. **c**, **d** Resolution matrix difference $\overline{E} (= R(\delta d) - R(0))$. **e**, **f** Row sums of $R(\delta d)$. $R(\delta d)$ is the projection matrix estimated by Eq. (40), which incorporates the observational errors in \underline{d}_1 . R(0) is the same as in Fig. 7b. PE = parameters constrained by erroneous observation \underline{d}_1

the same as $\mathbf{R}(\delta d = 0)$ (Fig. 7b), and the resultant $\delta \mathbf{x}_{of}$ (not shown here) equals the offset errors in $\underline{\mathbf{x}}$.

3.5. Properties of the Complete Resolution Matrix

Unlike R_D (from G only) and R_H (from G and C), a complete resolution matrix can be obtained from any combination of all of the factors in a study (Table 1). If the synthetic solution \underline{x}^k is obtained from a generalized inversion using error-free observations, then \mathbf{R}_C is equal to \mathbf{R}_D , with both matrices characterizing the same properties. If regularization is used during the processing, then \mathbf{R}_C is equal to \mathbf{R}_H , with both matrices characterizing the same properties. Therefore, the properties of \mathbf{R}_C vary based on the considered factors in the $\mathbf{x} \to \underline{\mathbf{x}}$ process.

3.6. Utilities of Resolution Matrices

Resolution matrices (e.g., $R_{\rm C}$) are a quantitative indicator not only for solution but also for other factors in a study system. All the factors (e.g., solution, observation, and regularization) can be appraised by the matrices.

Resolution matrix has been widely used to appraise solution reliability in an inverse problem (e.g., Aki et al. 1977; Aster et al. 2005; Backus and Gilbert 1970; Menke 2012; Tarantola and Valette 1982; Wiggins 1972). The diagonal entry $r_{i,i}$ reflects whether x_i can be resolved or how much of x_i is contained in x_i and then has received considerable attention for several decades (e.g., Aki et al. 1977; Day-Lewis et al. 2005; Wiggins 1972). Resolution spread estimated from resolution matrix can quantify the degree of departure of R from an identity matrix, and then the goodness of the model (Backus and Gilbert 1970; Menke 2012, 2015). However, resolution length, discussed in the "Resolution length" section, is widely used in practical model appraisals at present. Furthermore, the matrices $R_{\rm S}$ and $R_{\rm C}$ can also be applied to evaluate solution stability. In an unstable inversion, a small variation in observation causes a large change in solution, then $R_{\rm S}$ and $R_{\rm C}$ calculated from unstable solutions from random models are sensitive to the instability.

The resolvability and constrainability of \underline{x} under given observation, which is the most important information in a study, controls the solution reliability and must be taken into account to select parameterization and regularization. They can be quantitatively retrieved from resolution matrices, which is explained in the "Resolvability and constrainability from the resolution matrix" section.

The influences of regularization and errors on solution can also be evaluated by resolution matrices. The matrix $\mathbf{R}_{\rm D}$ reflects model projection under given observation. $\mathbf{R}_{\rm H}$ reflects the projection with the combination of observation and regularization. The difference, $\mathbf{R}_{\rm H} - \mathbf{R}_{\rm D}$, reflects the regularization influence on solution, which is discussed above in the "Significances of the resolution matrices" section. If an error exists in the process to obtain $\mathbf{x}, \mathbf{R}_{\rm C}$ will include the effect of the error. The difference, $\mathbf{R}_{\rm C}$

 $-R_{\rm D}$, reflects the error's effect, which is discussed in the Sects. 3.3 and 3.4.

4. Resolution Length

The smallest possible feature that can be detected is an important constraint in the model. The feature size is generally called the (spatial) resolution or resolution length (or width). Following the suggestion of Lebedev and Nolet (2003), the resolution length is defined here as the half size of the feature. This section focuses on how to obtain the resolution length from a resolution matrix.

4.1. Content Extent Versus Resolution Length

Equation (38) indicates that one solution parameter generally comes from the weighted averages of all of the medium parameters, with the entries from a row vector of \mathbf{R} being the weights or contents. The smallest resolvable feature represented by \underline{x}_i or the resolution length at \underline{x}_i should therefore be estimated from the row vector.

The medium parameters with high contents/ weights in \underline{x}_i mainly determine \underline{x}_i , such that the feature represented by \underline{x}_i is related to the high-content segment (e.g., $r_{48,41}$ to $r_{48,55}$ in Fig. 7c) in the rowvector curve. However, as the feature is represented by \underline{x}_i , the resolution length is not defined by the highvalue segment extent, but rather the distances from \underline{x}_i to the parameters at the segment borders. In Fig. 7c, the resolution length for \underline{x}_{48} is not the half distance from x_{41} ($r_{48,41}$) to x_{55} ($r_{48,55}$), but it is instead related to the distances from x_{48} ($r_{48,48}$) to x_{41} ($r_{48,41}$) and x_{55} ($r_{48,55}$); similarly, the resolution length for \underline{x}_{95} is not the half distance from x_{86} to x_{90} , but it is instead related to the distances from x_{95} to x_{86} and x_{90} (Fig. 7d).

In general, a parameter (e.g., x_i) and its neighbors should provide a large contribution to the average (the solution parameter \underline{x}_i) (Jackson 1972), and the contributions ($r_{i,j}$; j = 1, m) should decrease quickly with increasing distance from x_i to the other parameter (x_j). Therefore, each row of the resolution matrix ($r_{i,j}$; j = 1, m) can be approximated as either a Gaussian-shaped function (e.g., An 2012; Fichtner and Trampert 2011; Nolet 2008) or a cone (Barmin et al. 2001). The row-vector curve may be similar to the shape of a Gaussian function (e.g., $r_{48,*}$; Fig. 7c), but it can be quite different (e.g., $r_{95,*}$; Fig. 7d). In general, the width of the Gaussian function approximation to the row-vector curves can represent the resolution length (distance from x_i to the borders of the high-content segment) (Fig. 7c and d). For example, the resolution lengths in Fig. 7c and d are 4 km for x_{48} and 9 km for x_{95} , respectively.

4.2. Estimation from a Row or Column Vector?

The resolution length may also be estimated from a column vector (e.g., Alumbaugh and Newman 2000; Smith 1997). $\mathbf{R}_{\rm D}$ is symmetric, such that the resolution lengths estimated from its $i^{\rm th}$ row and column vectors are the same. However, regularization directly influences the spread functions (or column vectors), such that most of the resolution matrices with regularization (e.g., $\mathbf{R}_{\rm H}$) are asymmetric. Therefore, the resolution lengths estimated from the $i^{\rm th}$ column and row vectors can be different.

The resolution lengths that are estimated from the column and row vectors are largely the same or similar when Tikhonov regularization is applied. The resolution matrix $\mathbf{R}_{\rm H}(\lambda \mathbf{I})$ is symmetric, such that the lengths from its row and column vectors are same. Higher-order Tikhonov regularizations $(\lambda \mathbf{L}^n)$ yield a smoother column vector $(\mathbf{r}_{*,48} \text{ in Fig. 7e})$ than its corresponding row vector $(\mathbf{r}_{48,*} \text{ in Fig. 7e})$, and then the lengths from the two vectors are largely similar, as previously confirmed (Miller and Routh 2007; Pilkington 2016), but with exceptions.

The column vector cannot provide a valid resolution length for a parameter constrained by no observations (e.g., \underline{x}_{95} in Fig. 1a, b). The column vector for the parameters in $\mathbf{R}_{\rm H}$ (e.g., $\mathbf{r}_{*,95}$ in Fig. 7f) is all zeros, but the corresponding row vector is not ($\mathbf{r}_{95,*}$ in Fig. 7d), with a row-vector sum for $\mathbf{R}(\mathbf{S}^0)$ (e.g., $\mathbf{R}_{\rm H}(\mathbf{L}^n)$) that is equal to one. The row vector can provide a reliable resolution length in this case, whereas the column vector cannot (Table 3).

The row vector may be unstable and possess strong oscillations when large errors exist (e.g., $r_{5,*}$ in Fig. 8a, c). The high-amplitude oscillations around the diagonal entry of the row vector may imply an

In summary, the resolution length should generally be taken from the row vectors based on the definition of the resolution length (Table 3). However, special cases (e.g., observation errors) may yield an unstable row-vector curve that may in turn influence the estimated resolution from this vector. Therefore, it is advised to simultaneously extract the resolution length from the row and column vectors of the resolution matrix to ensure the resolution length is accurately defined.

4.3. Resolution Estimations that are not from R

The spatial resolution in seismic tomography studies is widely estimated via visual inspection of the restoration of the synthetic structure (e.g., checkerboard tests) (Feng and An 2010; Lévěque et al. 1993; Thurber and Ritsema 2009), as illustrated in Fig. 10a-c. If the checker size can be recovered at a given location, then the resolution length at that location in the final result is at least the same as the checker size. This method is powerful and easily realized. However, the resolution length is qualitative, not quantitative. Furthermore, the recovered (x) and synthetic checkerboard pattern model (x) in one test are equivalent to one pair of x and x. Several tests cannot produce sufficient model pairs to provide fully resolution information, as explained in the "Resolution matrices of the complete process" section.

The quantitative resolution length can still be retrieved when no resolution matrix is given or needed. The output solution \underline{x} of a synthetic test using an input model with a random structure x contains the resolution information (An 2012). Quantitative resolution information can be retrieved via a number of approaches, including a comparison of many x and \underline{x} (An 2012), cross correlation of x and \underline{x} (Trampert et al. 2013), and autocorrelation of \underline{x} (Fichtner and Leeuwen 2015). The resolution lengths in Fig. 10d were obtained using the An (2012) method on the basis of limited pairs of random synthetic models and



Figure 10

Example resolution lengths for a given tomography study of Rayleigh-wave dispersion at a period of 50 s. The figures are edited from Figs. 13 and 14 in Ma et al. (2014). **a–c** The resolution lengths are estimated using checkerboard tests, with the recoveries of three specific checker sizes inspected. The statistical resolution lengths in (**d**) are estimated from synthetic random models

5. Resolvability and Constrainability from the Resolution Matrix

Several essential questions arise when the real model x cannot be fully resolved in x for an ill-posed problem. For example, how much information from x can be reflected in the solution \underline{x} , or what is the resolvability of x_i (or the content of x_i in x_i)? How much information from \underline{x} is controlled by the observations? What is the constrainability (constraining status) of an individual solution parameter x_i under given observations? These questions are not only essential for understanding the reliability of the solution, but also instrumental in providing basic information to guide the improvement of the study system. These questions are essentially centered on the relationship between x and x, such that their answers can be derived from the $x \rightarrow x$ projection/ resolution matrix.

The $r_{i,j}$ entry in **R** represents the accurate content or contribution of the j^{th} model parameter x_i to the i^{th} solution parameter x_i . Therefore, the entries of the *i*th content vector $(\mathbf{r}_{i,*})$ and their sum $(\Sigma \mathbf{r}_{i,*})$ are indicators of the resolvability of x_i and the constrainability of x_i . As previously mentioned, $R_{\rm H}$ reflects the combined effects of the observation matrix and regularization during the $x \to x$ process, whereas $R_{\rm D}$ only reflects the observational effects. The reliability of the practical solution has therefore been evaluated via $R_{\rm H}$, but not $R_{\rm D}$. However, the constrainability of the solution parameter \underline{x}_i under given observations can be better obtained from $R_{\rm D}$ than from $R_{\rm H}$ for the same reason. The factors rather than the observations and regularization matrices can also influence the reliability of x_i ; this cannot be reflected by R_D and $R_{\rm H}$, but can be by $R_{\rm C}$.

5.1. Resolvability Defined by the Main-Diagonal Element

Equation (38) indicates that the main-diagonal element $r_{i,i}$ reflects the content (or contribution) of the real model parameter x_i in (to) its counterpart \underline{x}_i . This entry reflects whether x_i can be resolved or how much of x_i is contained in \underline{x}_i . $r_{i,i}$ can therefore be considered the resolvability of x_i (Table 3) and has received considerable attention for several decades (e.g., Aki et al. 1977; Day-Lewis et al. 2005; Wiggins 1972).

The main-diagonal element $r_{i,i}$ (e.g., Fig. 4b for Example 1) may take one of the following values:

- $r_{i,i} = 1$ (as illustrated in Fig. 11a). The curve shape of the elements in the *i*th row vector is a delta function, where all of the elements are zero, except $r_{i,i}$. In this case, \underline{x}_i equals x_i , which means that \underline{x}_i is well constrained and x_i can be fully resolved. If $r_{i,i}$ is in \mathbf{R}_D , then the parameter x_i is fully resolvable under the given observations.
- $r_{i,i} = 0$ (Fig. 11b). This case indicates that x_i makes no contribution to its counterpart \underline{x}_i and is unresolvable. For example, parameter \underline{x}_{95} in Example 1 is not constrained by an observation (Fig. 1a, b), such that $r_{95,95}$ in both R_D and R_H is zero (Figs. 3 and 4b), and \underline{x}_{95} in \underline{x}_D equals zero (Fig. 1c).
- *r_{i,i}* ∈ (0,1). *x_i* partially contributes to its counterpart <u>*x_i*</u> and is partially resolvable. If *r_{i,i}* in *R*_D belongs to (0,1), then <u>*x_i*</u> shares the observation with other parameters (e.g., <u>*x_j*); *x_i* therefore contributes to both its counterpart <u>*x_i*</u> and <u>*x_j*. For example *r*_{1,1} = 0.1 in *R*_D in Example 1 (Fig. 4b), which indicates that <u>*x*₁</u> shares observation 1 with <u>*x*₂</u> (Fig. 1a); *x₁* therefore contributes to both <u>*x*₁</u> and <u>*x*₂</u>.
 </u></u>

The main-diagonal element $r_{i,i}$ in $\mathbf{R}_{\rm D}$ reflects the resolvability of x_i under a given observational condition, such that the sum of all of the maindiagonal elements, or trace($\mathbf{R}_{\rm D}$), can be considered the resolvability of the model vector \mathbf{x} . As $\mathbf{R}_{\rm D}$ is a Gram matrix, trace($\mathbf{R}_{\rm D}$) equals the number of independent observations (rank(\mathbf{G})). $\mathbf{R}_{\rm D}$ will therefore be an identity matrix, and \mathbf{x} can be fully resolved when either trace($\mathbf{R}_{\rm D}$) or rank(\mathbf{G}) equals the number of model parameters (m).



Figure 11

Illustration of the four constraining statuses for the *i*th solution parameter \underline{x}_i based on the elements of the *i*th row vector $\mathbf{r}_{i,*}$

5.2. Deviation from the Expectation Given by the Row-Vector Sum

Equation (38) indicates that \underline{x}_i equals the weighted sum of all of the parameters in x, and $\sum r_{i,*}$ (or $\sum_i R$) is the sum of all of the weighted $r_{i,*}$. If R is a stochastic matrix, then a sum of one means that \underline{x}_i reflects the true average of x (Nolet 2008) and lies at least within the extremes of x. For example, if $\sum_1 R_D = 1$ (Fig. 4a) and $r_{1,*}$ in R_D are positive (Fig. 3a), then \underline{x}_1 in \underline{x}_D is a good representative of x_1 (Fig. 1c). However, the assumption is often false, as the entries in R can be negative (Menke 2015) (i.e., R is often not a stochastic matrix). For example, $\sum r_{1,*}$ in $R_H(L^1)$ in Example 1 equals one (Fig. 4a), but the parameter \underline{x}_1 in $\underline{x}(L^1)$ (Fig. 1c) deviates from the expected average. The polarity of $r_{i,*}$ must therefore be considered when $\sum_i(R)$ is used to judge the reliability of x_i .

While $\Sigma_i \mathbf{R} = 1$ does not necessarily correspond to the perfectness of \underline{x}_i , the deviation of $\Sigma_i \mathbf{R}$ from one is a good indicator of the deviation of \underline{x}_i from the true model average (Table 3). A comparison of Figs. 4a and 1c indicates that an overestimated (larger than the model average) parameter \underline{x}_i (\underline{x}_{54} in \underline{x}_D and $\underline{x}(I)$ in Fig. 1c) corresponds to $\Sigma_i R > 1$ (e.g., $\Sigma_{54} R_D$ and $\Sigma_{54} R_H(I)$ in Fig. 4a), and an underestimated (smaller than the model average) parameter (\underline{x}_{48} in \underline{x}_D and $\underline{x}(I)$ in Fig. 1c) corresponds to $\Sigma_i R < 1$ (e.g., $\Sigma_{48} R_D$ and $\underline{x}(I)$ in Fig. 1c) corresponds to $\Sigma_i R < 1$ (e.g., $\Sigma_{48} R_D$ and $\underline{x}_{48} R_H(I)$ in Fig. 4a).

5.3. Difference Between Neighboring Parameters

If the parameter x_i is partially resolvable $(r_{i,i} \in (0,1))$, then the solution parameter \underline{x}_i will include content $(r_{i,ii} > 0)$ from neighboring parameters (e.g., x_{ii}), and the $r_{i,i}$ and $r_{i,ii}$ entries in the *i*th row vector of **R** can reflect the similarity between \underline{x}_i and \underline{x}_{ii} .

When two neighboring parameters \underline{x}_i and \underline{x}_{ii} (e.g., \underline{x}_{60} and \underline{x}_{61} ; Fig. 1a, b) are constrained by unrelated observations, $r_{i,i} < 1$ and $r_{i,ii} = 0$ (Fig. 11c), as observed for $r_{60,61}$ and $r_{60,60}$ in either \mathbf{R}_D or $\mathbf{R}_H(\mathbf{I})$ (Fig. 3a and b). In this case, x_{ii} does not contribute to \underline{x}_i (Eq. (38)), and x_i does not contribute to \underline{x}_{ii} . It is possible to discriminate the difference

between x_i and x_{ii} in \underline{x}_i and \underline{x}_{ii} from the unrelated observations, as \underline{x}_{61} in either \underline{x}_D or $\underline{x}(I)$ is obviously different from \underline{x}_{60} (Fig. 1c).

When two neighboring parameters \underline{x}_{ii} and \underline{x}_i (e.g., \underline{x}_{13} and \underline{x}_{14} , which are constrained by the second observation in Fig. 1a, b) are constrained by the same or related observations, both $r_{i,i}$ and $r_{i,ii}$ (ii = i - 1 or i + 1; Fig. 11d) (e.g., $r_{13,13}$ and $r_{13,14}$ in either \mathbf{R}_D or $\mathbf{R}_H(\mathbf{I})$; Fig. 3a and b) are in the range (0,1). In this case, \underline{x}_i has contents from both x_{ii} and x_i . \underline{x}_{ii} also has contents from both x_{ii} and x_i because of the symmetry of \mathbf{R}_D . Consequently, \underline{x}_i and \underline{x}_{ii} cannot be discriminated, and their difference is often related to the difference $\Delta r_{i,ii}$:

$$\Delta r_{i,ii} = |r_{i,i} - r_{i,ii}|. \tag{57}$$

The difference between \underline{x}_i and \underline{x}_{ii} in \underline{x}_D is often very large when $\Delta r_{i,ii}$ is large, even if the real model parameters x_i and x_{ii} are the same. When $\Delta r_{i,ii}$ is small, the difference is also small, even if x_i and x_{ii} are quite different. When $\Delta r_{i,ii}$ is zero, \underline{x}_i is often equal to \underline{x}_{ii} . The solution parameters \underline{x}_{13} and \underline{x}_{14} in either \underline{x}_D or $\underline{x}(I)$ are quite different (Fig. 1c), even though the real model parameters x_{13} and x_{14} are almost the same. $\Delta r_{i,ii}$ is therefore a good indicator of the difference between \underline{x}_i and \underline{x}_{ii} (Table 3) if $r_{i,i}$ is less than one, even though this difference has no relation to the difference between x_i and x_{ii} .

5.4. Short Summary on Constrainability

In summary, the constrainability of a solution parameter can be quantitatively evaluated from its content vector in a resolution matrix (Table 3). The main-diagonal element $(r_{i,i})$ can be considered the resolvability of x_i . If \mathbf{R}_D is used, then $r_{i,i}$ values of 0, 1, and (0,1) mean that x_i is unresolvable, fully resolvable, and partially resolvable, respectively, for given observations. The deviation of the content vector sum $(\Sigma \mathbf{r}_{i,*})$ from one can be considered an indicator of the deviation from the model expectation. Values of $\Sigma \mathbf{r}_{i,*} > 1$ and < 1 mean that \underline{x}_i is overestimated and underestimated, respectively. In the case where $\Sigma \mathbf{r}_{i,*} = 1$ and all of the elements $\mathbf{r}_{i,*}$ are nonnegative, \underline{x}_i is the model true average. The difference between $r_{i,i}$ and $r_{i,ii}$ ($\Delta r_{i,ii}$) is a reflection of the difference between \underline{x}_i with \underline{x}_{ii} for a partially resolvable parameter x_i ($r_{i,i} \in (0,1)$). Large and small $\Delta r_{i,ii}$ values correspond to large and small differences between two neighboring parameters, respectively, although the solution difference has no relation to the

difference between the true medium parameters.

5.5. Constrainability from $R_{\rm H}$

Practical studies that employ regularization do not provide $R_{\rm D}$, but rather $R_{\rm H}$. Estimation of the parameter constrainability under a given observation from $R_{\rm H}$ is necessary in this case to determine how much information in the solution is from observation rather regularization. As mentioned above, $R_{\rm H}(\lambda I)$ is the most similar matrix to $R_{\rm D}$, making it a good alternative to $R_{\rm D}$ for evaluating the constrainability. However, most of the main-diagonal elements in $\mathbf{R}_{\rm H}(\lambda \mathbf{I})$ are smaller than those in $\mathbf{R}_{\rm D}$. The row vector sum of $R_{\rm H}(\lambda I)$ may therefore be smaller than that of $R_{\rm D}$, such that its deviation from one cannot be used to evaluate the underestimated parameters. $R_{\rm H}(\lambda I)$ can also be used to evaluate the well constrained parameters, as the curve shape of their row vectors is still a delta function, even though $r_{i,i}$ may not be one.

 $\mathbf{R}_{\mathrm{H}}(\mathbf{L}^{n})$ (n > 0; e.g., $\mathbf{R}_{\mathrm{H}}(\mathbf{L}^{1})$ in Fig. 3c) is significantly different from R_D (Fig. 3a). With the exception of the unconstrained parameters, the main-diagonal elements of $\mathbf{R}_{H}(\mathbf{L}^{1})$ ($r_{i,i}$ and $r_{ii,ii}$) for two neighboring parameters $(x_i \text{ and } x_{ii})$ that are constrained by the same observation can be different (Fig. 4b). The row-vector sum of $\mathbf{R}_{\rm H}(\mathbf{L}^1)$ for any single parameter equals one (Fig. 4a). Therefore, the main-diagonal elements $r_{i,i}$ and row-vector sum $\Sigma_i \mathbf{R}_{\rm H}(\mathbf{L}^1)$ cannot be used to evaluate the constrainability of the parameter x_i . However, the all-zero column vectors of $\mathbf{R}_{\rm H}(\mathbf{L}^n)$ (n = 0, 1, 2) for the unconstrained parameters are the same as those in $R_{\rm D}$, such that the unconstrained parameters and unconstrained neighbors can be evaluated from $\mathbf{R}_{\mathrm{H}}(\mathbf{L}^{n})$. Furthermore, the curve shape of the row vectors of $\mathbf{R}_{\rm H}(\mathbf{L}^n)$ (n = 0, 1) is similar to that of $\mathbf{R}_{\rm D}$ (Figs. 1f and 3). The mutual relationship between two neighboring parameters can therefore be roughly evaluated using $\Delta r_{i,ii}$ of $\mathbf{R}_{\mathrm{H}}(\mathbf{L}^{n})$.

5.6. Resolution Upper Bound

 $R_{\rm D}$ is only related to G (Eq. (19)), with no relationship to the observation data \underline{d} with observational errors (δd) and other factors, whereas $R_{\rm H}$ is composed of both G and C. Regularization adds artificial constraints to make the solution appear more rational. However, various factors, including observational errors, regularization, and the instability of the solution can decrease (but not increase) the resolution reflected by $R_{\rm D}$. Therefore, the resolution derived from $R_{\rm D}$ marks the upper bound of resolvability (Table 1).

While repeated observations can improve the precision of the solution by improving the precision in the observation data (\underline{d}), they cannot increase either rank(G) or the number of independent observations. Repeated observations therefore have no effect on R_D and cannot improve the upper bound of resolution. The only way to improve the upper bound of resolution is to increase the independent observations (or rank(G)).

5.7. What is a Perfect Inversion?

A perfect inversion, or a perfect constrainability, corresponds to a resolution matrix that equals the identity matrix I (e.g., Jackson 1972; Menke 1989). However, this rule is only applicable for direct resolution matrix $R_{\rm D}$, not the other resolution matrices (e.g., $\mathbf{R}_{\rm H}$ or $\mathbf{R}_{\rm C}$). A main-diagonal element $r_{i,i}$ of one in \mathbf{R}_{D} means that x_{i} can be fully resolved in \underline{x}_{i} , thereby implying that a perfect inversion has identity matrix of $R_{\rm D}$. $R_{\rm D}$ always equals I except when the problem is under-determined. However, $R_{\rm H}$ includes regularization, and regularization is largely employed when the constrainability of the solution parameters under a given observational condition is imperfect, (i.e., trace($\mathbf{R}_{\rm D}$) is often smaller than *m* or the inverse of G is unstable). Therefore, $trace(\mathbf{R}_{H})$ is always smaller than trace($\mathbf{R}_{\rm D}$), regardless of the quality of the regularization, such that $R_{\rm H}$ will never be an identity matrix I. Therefore, the perfectness of the inversion cannot be judged based on the degree of similarity between $R_{\rm H}$ and I. The perfectness of a study using regularization should be judged based on the amount of valid information in G that is reflected in the solution, as different regularizations yield different solutions.

6. Resolution Matrices in Nonlinear Inversions

The relationship between x and x is different from that between x (or x) and d (or d). The relationship between x and x $(r:x \to x)$ of a nonlinear problem can be nonlinear, but can also be linear, such as when true model x is perfectly resolved (x = x = I x = r(x)). This indicates, the $x \to x$ projection is a different relation with the problem but relates with the ability of solving the problem. The nonlinear inverse problem cannot be described by either Eq. (6) or (9), such that neither R_D (Eq. (19)) nor $R_{\rm H}$ (Eq. (24)) can be obtained for the problem. Anyway, Eq. (2) remains a valid linear projection approximation of the nonlinear equation (Eq. (1)), such that $R_{\rm C}$, which is directly obtained from Eq. (2), can still be provided, regardless of the method used to solve the problem.

A nonlinear inverse problem can be solved via either a global optimization or linearized method. If the problem is resolved using linearized iterative methods (e.g., Aster et al. 2005; Bourgeois et al. 1989), such as Newton's method, then a resolution matrix can be obtained after each iteration (Jackson 1972) from either Eq. (19), Eq. (24), or a similar equation. However, the resolution matrix is not any of the above-mentioned resolution matrices (\mathbf{R}_D , \mathbf{R}_H , \mathbf{R}_C , \mathbf{R}_I , and \mathbf{R}_S), but rather one of three new resolution matrices that are specifically constructed for nonlinear inverse problems.

An iterative inversion is designed to invert for the model perturbation $\Delta x^i (= x - \underline{x}^{i-1})$ of the reference model \underline{x}^{i-1} at the *i*th iteration on the basis of the first-order approximation of the inverse problem (Eq. (4)):

$$\Delta \boldsymbol{d}^{i} = \boldsymbol{G}_{\mathrm{I}}^{i} \Delta \boldsymbol{x}^{i}. \tag{58}$$

where $\Delta d^i = \underline{d} - g(\underline{x}^{i-1})$ and G_J^i is a Jacobian matrix with the partial derivatives of d taken with respect to x at \underline{x}^{i-1} . Equation (58) is a linear equation, such that the solution of the perturbation Δx^i ($\Delta \underline{x}^i$) can be obtained from G_J^i and/or C using the same methods that were employed to obtain the model solution \underline{x} from G and/or C in the above linear inversions (e.g., Eq. (11)). However, the inverted solution $(\Delta \underline{x}^i)$ is a model perturbation, not the model. The inverted model (solution) after the *i*th iteration \underline{x}^i is \underline{x}^{i-1-} $+ \Delta \underline{x}^i$, which is the reference model at the next inversion iteration. This is the general procedure of Newton's method.

A surface-wave inversion to constrain the S-wave velocity structure (Example 3) (Fig. 12a) is synthetized here to illustrate the three resolution matrices that can be implemented in a linearized nonlinear inversion. The employed inversion is a typical nonlinear inversion approach in geophysics that is widely applied to elucidate 1-D and three-dimensional sedimentary, crustal, and lithospheric Earth structures (e.g., Feng and An 2010; Knopoff 1972; Snoke and James 1997; Wiggins 1972; Xia et al. 1999); therefore, the details of the nonlinear relationship between x and d are not explained here. First-order Tikhonov regularization has been widely used in this inversion approach, although it prevents the correction of bad discontinuities in the reference model (An 2020). A regularization approach that is adapted to the reference models suggested by An (2020) can overcome this problem and lead to a rapidly convergent iterative inversion; this regularization approach is used here. The regularization parameter was set to $\lambda = 0.01$ after a series of tests that explored the tradeoff between the misfit and model flatness. The synthetic observation (d) and predictions (Fig. 12a) and partial-derivative matrix $G_{\rm J}^{i}$ (Eq. (58)) were calculated using the surf96 program (Herrmann 2013). Given a reference model at the first iteration (the starting model), the model solutions after the first to fifth iterations (x^i) and their fits to the reference model are shown in Fig. 12a. The solution \underline{x}^5 after the fifth iteration is nearly the same as x^4 after the fourth iteration, which implies that the inversion converged around \underline{x}^4 . \underline{x}^4 is, therefore, considered the final solution.

6.1. Linear Approximation of $r: x \rightarrow x$

The $R_{\rm C}$ calculation for a nonlinear inversion is the same as that for the above linear inversions, whereby

only the synthetic input models (x) and their corresponding solutions (\underline{x}) are used. This calculation, which is based on either Eq. (43) or (47), is in fact a linear regression of x and \underline{x} (i.e., $R_{\rm C}$ is a linear approximation of $r: \underline{x} \to \underline{x}$) (Table 1) (Fig. 2a). However, the relationship between \underline{x} and x is often nonlinear due to the complexity in the $x \to \underline{x}$ process for a nonlinear problem. The calculation of a reliable $R_{\rm C}$ therefore requires more pairs of input/output models. Furthermore, the resultant $R_{\rm C}$ will somewhat depend on the synthetic random input models. If they are closer to the practical medium, then the resultant $R_{\rm C}$ will be more realistic.

6.2. Projection of Solution Improvement After an Iteration

As Eq. (58) is of the same form as Eq. (6), the process from $\Delta \mathbf{x}^i (= \mathbf{x} - \underline{\mathbf{x}}^{i-1})$ to $\Delta \underline{\mathbf{x}}^i (= \underline{\mathbf{x}}^i - \underline{\mathbf{x}}^{i-1})$ after the *i*th iteration can be expressed in a form similar to Eq. (22) to represent the $\mathbf{x} \to \underline{\mathbf{x}}$ process (e.g., Jackson 1972; Wiggins 1972):

$$\Delta \underline{x}^{i} = R^{i}_{J} \Delta x^{i} \tag{59}$$

or:

$$\underline{\boldsymbol{x}}^{i} - \underline{\boldsymbol{x}}^{i-1} = \boldsymbol{R}_{\mathrm{J}}^{i}(\boldsymbol{x} - \underline{\boldsymbol{x}}^{i-1}), \qquad (60)$$

where \mathbf{R}_{J}^{i} denotes the resolution matrix $\mathbf{R}:\Delta \mathbf{x}^{i} \rightarrow \Delta \underline{\mathbf{x}}^{i}$. Equation (59) is of the same form as Eq. (2), such that \mathbf{R}_{J}^{i} (denoted \mathbf{R}_{JD}^{i}) can be obtained via Eq. (19):

$$\boldsymbol{R}_{\mathrm{JD}}^{i} = (\boldsymbol{G}_{\mathrm{J}}^{i})^{-\mathrm{g}} \boldsymbol{G}_{\mathrm{J}}^{i}. \tag{61}$$

When the regularization matrix C is used, R_J^i (denoted R_{JH}^i) can be obtained via Eq. (24):

$$\boldsymbol{R}_{\rm JH}^i = (\boldsymbol{A}_{\rm J}^i)^{-t} \boldsymbol{G}_{\rm J}^i, \qquad (62)$$

where A_J^i is the combination of G_J^i and C, which is similar to how A is the combination of G and C in Eq. (10). The matrix R_{JD}^i (or R_{JH}^i) is often denoted R previously, but it possesses a different significance than the resolution matrix $R:x \to \underline{x}$.

The first-order Taylor expansion of Eq. (1) at \underline{x}^{i-1} is:

$$\underline{\boldsymbol{x}}^{i} = \underline{\boldsymbol{x}}^{i-1} + \boldsymbol{J}_{r}(\underline{\boldsymbol{x}}^{i-1})(\boldsymbol{x} - \underline{\boldsymbol{x}}^{i-1}), \qquad (63)$$

where $J_r(\underline{x}^{i-1})$ denotes the Jacobian matrix (or the gradient) of the projection $r: x \to x$ at reference





Example resolution matrices for the solutions after iteration 4 of Newton's iterations of the surface-wave dispersion inversion to determine the 1-D S-wave velocity structure. **a** Synthetic true model (**x**) and solutions (\underline{x}^i) for the 1-D S-wave velocity structure after the *i*th iteration. Synthetic observations (\underline{d}) and predictions $d(\underline{x}^i)$ of the surface-wave dispersion curves for solution \underline{x}^i . **b** Complete resolution matrix R_C for solution \underline{x}^i (i = 4). **c** R_{IH}^{-4} , which is the resolution matrix for solution improvement just after iteration 4. **d** Difference between R_{IH}^{-4} and $R_{IH}^{1\rightarrow 4}$, which is the resolution matrix of the solution improvement on the starting model. Gaussian widths (*red dashes*) in (**b–d**) correspond to the resolution length

model \underline{x}^{i-1} . Equation (63) has the same form as Eq. (59). Therefore, R_J^i is exactly the Jacobian matrix $J_r(\underline{x}^{i-1})$ (Table 1). Practically, R_J^i represents the

projection from $\mathbf{x} - \underline{\mathbf{x}}^{i-1}$ to $\underline{\mathbf{x}}^i - \underline{\mathbf{x}}^{i-1}$ (Eq. (59)) (i.e., the projection of the solution improvement on the reference model ($\underline{\mathbf{x}}^{i-1}$) just after the *i*th iteration) (e.g.,

 \mathbf{R}_{J}^{1} illustrated in Fig. 13b). The matrix \mathbf{R}_{J}^{i} in Example 3 (e.g., \mathbf{R}_{JH}^{4} just after the fourth iteration (Fig. 12c)) is slightly different than \mathbf{R}_{C} (Fig. 12b).

6.3. Projection of the Solution Improvement up to an Iteration

The inversion after each iteration is represented by Eq. (59), but an application needs one more iteration. The solution improvement from the k^{th} to i^{th} iterations can be expressed as:

$$\underline{\boldsymbol{x}}^{i} - \underline{\boldsymbol{x}}^{k-1} = \boldsymbol{R}_{\mathrm{J}}^{k \to i} (\boldsymbol{x} - \underline{\boldsymbol{x}}^{k-1}), \qquad (64)$$

where:

$$\Delta \boldsymbol{R}_{\mathrm{J}}^{k \to i} = \Delta \boldsymbol{R}_{\mathrm{J}}^{i} + \Delta \boldsymbol{R}_{\mathrm{J}}^{k \to (i-1)} - \Delta \boldsymbol{R}_{\mathrm{J}}^{i} \Delta \boldsymbol{R}_{\mathrm{J}}^{k \to (i-1)}. \quad (65)$$

If k = i just after the *i*th iteration of a given inversion, then Eq. (64) should be the same as Eq. (59). The matrix $\mathbf{R}_{J}^{i \to (i-1)} = \{0\}$, which means that $\mathbf{R}_{J}^{i \to i} = \mathbf{R}_{J}^{i}$. Therefore, if k = 1, then Eq. (64) becomes:

$$\underline{\boldsymbol{x}}^{i} - \underline{\boldsymbol{x}}^{0} = \boldsymbol{R}_{J}^{1 \to i} (\boldsymbol{x} - \underline{\boldsymbol{x}}^{0}), \qquad (66)$$

where:

$$\boldsymbol{R}_{\mathrm{J}}^{1\to i} = \boldsymbol{R}_{\mathrm{J}}^{i} + \boldsymbol{R}_{\mathrm{J}}^{1\to (i-1)} - \boldsymbol{R}_{\mathrm{J}}^{i}\boldsymbol{R}_{\mathrm{J}}^{1\to (i-1)}.$$
 (67)

The matrix $\mathbf{R}_{J}^{1 \to i}$ represents the projection from $\Delta \mathbf{x} \ (= \mathbf{x} - \underline{\mathbf{x}}^{0})$ to $\Delta \underline{\mathbf{x}} \ (= \underline{\mathbf{x}}^{i} - \underline{\mathbf{x}}^{0})$, which is a projection of the solution improvement on the starting model $\underline{\mathbf{x}}^{0}$ after *i* iterations $(\underline{\mathbf{x}}^{i} - \underline{\mathbf{x}}^{0})$ (Fig. 13b). This is different from the gradient of $r: \mathbf{x} \to \underline{\mathbf{x}}$ at $\underline{\mathbf{x}}^{i} \ (\mathbf{R}_{J}^{i})$, as $\mathbf{R}_{J}^{1 \to i}$ is the slope of $r: \mathbf{x} \to \underline{\mathbf{x}}$ from $\underline{\mathbf{x}}^{0}$ to $\underline{\mathbf{x}}^{i}$ (Table 1). The magnitude difference between $\mathbf{R}_{JH}^{1 \to i}$, which is obtained using \mathbf{R}_{JH}^{1} , ..., \mathbf{R}_{JH}^{4} , and \mathbf{R}_{JH}^{4} (Fig. 12d) is remarkable. The matrix $\mathbf{R}_{J}^{1 \to i}$ better represents the solution improvement in the inversion than \mathbf{R}_{J}^{i} .

6.4. Four Types of Resolution Matrices in a Linearized Inversion

The matrices $\mathbf{R}_{J}^{i} : \Delta \mathbf{x}^{i} \to \Delta \underline{\mathbf{x}}^{i}$ and $\mathbf{R}_{J}^{1 \to i} : \Delta \mathbf{x} \to \Delta \underline{\mathbf{x}}$ in a linearized inversion can also be obtained via the \mathbf{R}_{C} calculation. If the synthetic perturbations $\Delta \mathbf{x}^{i} (= \mathbf{x} - \underline{\mathbf{x}}^{i-1})$ and solution $\Delta \underline{\mathbf{x}}^{i}$ after the *i*th iteration are used as the true model \mathbf{x} and corresponding solution $\underline{\mathbf{x}}$, respectively, then the resultant \mathbf{R}_{C} (denoted \mathbf{R}_{JC}^{i}) should be the same as $\mathbf{R}_{JH}^{i} : \Delta \mathbf{x}^{i} \to \Delta \underline{\mathbf{x}}^{i}$. If the synthetic perturbations $\Delta \mathbf{x}^{(i-1)}$ and corresponding solution $\Delta \mathbf{x}^{1-i}$.





Illustration of resolution matrices in a linearized inversion. The model \mathbf{x} contains one parameter (x_1) . **a** True model (\mathbf{x}) and solution at each iteration. **b** Projections with resolution matrices \mathbf{R}_C or $\mathbf{R}_C^{\ i}$ (for $r:\mathbf{x} \to \underline{\mathbf{x}}^i$), $\mathbf{R}_C^{\ 1}$ (for $r:\mathbf{x} \to \underline{\mathbf{x}}^1$), $\mathbf{R}_J^{\ 1}$ at first iteration (slope of $r:\mathbf{x} \to \underline{\mathbf{x}}$ at $\underline{\mathbf{x}}^0$), and $\mathbf{R}_J^{\ 1\to i}$ after i iterations (slope of $r:\mathbf{x} \to \underline{\mathbf{x}}$ from $\underline{\mathbf{x}}^0$ to $\underline{\mathbf{x}}^i$)

 $(= \frac{\mathbf{x}^{i} - \mathbf{x}^{0}}{\mathbf{x}^{1-i}})$ are used, then the resultant \mathbf{R}_{C} (denoted \mathbf{R}_{JC}) is the same as $\mathbf{R}_{JH}^{1 \to i}$.

Furthermore, if the synthetic \mathbf{x} and corresponding solution $\underline{\mathbf{x}}^i (= \mathbf{x}^{i-1} + \Delta \underline{\mathbf{x}}^i)$ after the i^{th} iteration are used in the $\mathbf{R}_{\rm C}$ calculation, then the resultant $\mathbf{R}_{\rm C}$ (denoted $\mathbf{R}_{\rm C}^{i}$) (e.g., $\mathbf{R}_{\rm C}$ and $\mathbf{R}_{\rm C}^{-1}$ in Fig. 13b) represents the projection from \mathbf{x} to $\underline{\mathbf{x}}^i (= \mathbf{x}^{i-1} + \Delta \underline{\mathbf{x}}^i)$ just after the iteration ($\mathbf{R}^i: \mathbf{x} \to \underline{\mathbf{x}}^i$). This is one more new type of resolution matrix that may appear in a linearized iterative application. If $\underline{\mathbf{x}}^i$ is final solution $\underline{\mathbf{x}}, \mathbf{R}_{\rm C}^i$ is written as $\mathbf{R}_{\rm C}$.

In summary, completed resolution matrix for the $x \rightarrow \underline{x}$ process can be obtained from the linear approximation of $r:x \rightarrow \underline{x}$, regardless of the method used to solve a given nonlinear inverse problem. However, a linearized iterative application can have four classes of resolution matrices (Tables 1 and 2, Fig. 13b), $R_{\rm C}$, $R_{\rm C}^{i}$, $R_{\rm J}^{1\rightarrow i}$, and $R_{\rm J}^{i}$, which represent the $x \rightarrow \underline{x}$, $x \rightarrow \underline{x}^{i}$, $(x - \underline{x}^{0}) \rightarrow (\underline{x}^{i} - \underline{x}^{0})$, and $(x - \underline{x}^{i-1}) \rightarrow (\underline{x}^{i} - \underline{x}^{i-1})$ projections, respectively. **R** is a linear approximation of the operator *r*, whereas $R_{\rm J}$ is the Jacobian matrix of *r*. The resolution matrix $R_{\rm JH}^{i}$, which is often provided in the literature, reflects the solution improvement just at the *i*th iteration. The matrix $R_{\rm JH}^{1\rightarrow i}$ reflects the solution improvement of the solution after all *i* iterations from starting model to the solution \underline{x}^{i} .

The surface-wave dispersion inversion in Example 3 highlights that even though the magnitudes of $R_{\rm C}$, $R_{\rm JH}^4$, and $R_{\rm JH}^{1\to4}$ (Fig. 12b–d) are obviously different, their magnitude patterns (that of $R_{\rm JH}^{1\to4}$ is not shown here) are similar. The resolution lengths estimated from the three matrices (Fig. 12d) are also similar. Therefore, the resolution lengths retrieved from either $R_{\rm JH}^{1\to4}$ or $R_{\rm JH}^4$ in a surface-wave dispersion inversion are also acceptable if $R_{\rm C}$ cannot be given.

7. Conclusion

Here, we reviewed previous resolution matrices and their applications to clarify the properties of resolution matrices in linear and nonlinear inversions that implement zeroth- and higher-order Tikhonov regularizations. We explained how to use the resolution matrix to understand both the resolvability of the medium parameters and the constrainability of the solution parameters. Furthermore, we suggested a new resolution matrix, the complete resolution matrix, which reflects all of the factors in a study system. This new matrix, which is able to overcome many of the limitations encountered by previous matrices, can be broadly applied in linear and nonlinear (inverse and non-inverse) problems. This study is designed to assist the reader in fully understanding both the concept and application of a resolution matrix and in recognizing how to appropriately appraise a solution and understand the relationship between the solution and all of the factors in the study. These matrix suggestions can guide the reader in improving the study system.

Acknowledgements

This paper is dedicated to the memory of the author's father, An Feng-Lou, who passed away during the preparation of this manuscript. This work was funded by the National Natural Science Foundation of China (Grants 41574049 and 41974051), Basic Research Foundation of Chinese Academy of Geological Sciences (JKY202217), and the China Geological Survey Program (DD20221643-2). I thank two anonymous reviewers for their constructive comments and suggestions. All of the figures were generated using Generic Mapping Tools (Wessel and Smith 1991).

Author contributions M.A. completed the manuscript.

Funding

This work was funded by the National Natural Science Foundation of China (Grants 41574049 and 41974051) and the China Geological Survey Program (DD20221643-2).

Data availability statement

No practical data is used in this paper.

Declarations

Conflict of interest The author has no relevant financial or non-financial interests to disclose.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

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

References

- Aki, K., Christoffersson, A., & Husebye, E. S. (1977). Determination of the three-dimensional seismic structure of the lithosphere. *Journal of Geophysical Research*, 82(2), 277–296. https://doi.org/10.1029/JB082i002p00277
- Alumbaugh, D. L., & Newman, G. A. (2000). Image appraisal for 2-D and 3-D electromagnetic inversion. *Geophysics*, 65(5), 1455–1467. https://doi.org/10.1190/1.1444834
- An, M. (2012). A simple method for determining the spatial resolution of a general inverse problem. *Geophysical Journal International*, 191(2), 849–864. https://doi.org/10.1111/j.1365-246X.2012.05661.x
- An, M. (2020). Adaptive Regularization of the Reference Model in an Inverse Problem. *Pure and Applied Geophysics*, 177(10), 4943–4956. https://doi.org/10.1007/s00024-020-02530-z
- Aster, R. C., Borchers, B., & Thurber, C. H. (2005). Parameter Estimation and Inverse Problems. Burlington: Academic Press.
- Backus, G., & Gilbert, F. (1968). The resolving power of gross earth data. *Geophysical Journal of the Royal Astronomical Society*, 16, 169–205.
- Backus, G., & Gilbert, F. (1970). Uniqueness in the Inversion of Inaccurate Gross Earth Data. *Philosophical Transactions of the Royal Society B*, 266(1173), 123–192. https://doi.org/10.1098/ rsta.1970.0005
- Barmin, M. P., Ritzwoller, M. H., & Levshin, A. L. (2001). A fast and reliable method for surface wave tomography. *Pure and Applied Geophysics*, 158, 1351–1375.
- Benning, M., & Burger, M. (2018). Modern regularization methods for inverse problems. Acta Numerica, 27, 1–111. https://doi.org/ 10.1017/S0962492918000016

- Boschi, L. (2003). Measures of resolution in global body wave tomography. *Geophysical Research Letters*, 30(19), 1978. https:// doi.org/10.1029/2003gl018222
- Bourgeois, A., Jiang, B. F., & Lailly, P. (1989). Linearized inversion: A significant step beyond pre-stack migration. *Geophysical Journal International*, 99(2), 435–445. https://doi.org/ 10.1111/j.1365-246X.1989.tb01700.x
- Chevrot, S., Villaseñor, A., Sylvander, M., Benahmed, S., Beucler, E., Cougoulat, G., Delmas, P., de Saint Blanquat, M., Diaz, J., Gallart, J., Grimaud, F., Lagabrielle, Y., Manatschal, G., Mocquet, A., Pauchet, H., Paul, A., Péquegnat, C., Quillard, O., Roussel, S., ... Wolyniec, D. (2014). High-resolution imaging of the Pyrenees and Massif Central from the data of the PYROPE and IBERARRAY portable array deployments. *Journal of Geophysical Research*, *119*(8), 6399–6420. https://doi.org/10.1002/ 2014JB010953
- Chiao, L.-Y., Chen, Y.-N., & Gung, Y. (2014). Constructing empirical resolution diagnostics for kriging and minimum curvature gridding. *Journal of Geophysical Research*, 119(5), 3939–3954. https://doi.org/10.1002/2013JB010364
- Constable, S. C., Parker, R. L., & Constable, C. G. (1987). Occam's inversion: A practical algorithm for generating smooth models from electromagnetic sounding data. *Geophysics*, 52(3), 289–300. https://doi.org/10.1190/1.1442303
- Craven, P., & Wahba, G. (1978). Smoothing noisy data with spline functions. *Numerische Mathematik*, 31(4), 377–403. https://doi. org/10.1007/BF01404567
- Crosson, R. S. (1976). Crustal structure modeling of earthquake data 1. Simultaneous least squares estimation of hypocenter and velocity parameters. *Journal of Geophysical Research*, 81(17), 3036–3046. https://doi.org/10.1029/JB081i017p03036
- Day-Lewis, F. D., Singha, K., & Binley, A. M. (2005). Applying petrophysical models to radar travel time and electrical resistivity tomograms: Resolution-dependent limitations. *Journal of Geophysical Research*, 110(B8), B08206. https://doi.org/10.1029/ 2004JB003569
- Engl, H. W., Hanke, M., & Neubauer, A. (2000). Regularization of Inverse Problems. Amsterdam: 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
- Fichtner, A., & Leeuwen, T. V. (2015). Resolution analysis by random probing. *Journal of Geophysical Research*, 120(8), 5549–5573. https://doi.org/10.1002/2015jb012106
- Fichtner, A., & Trampert, J. (2011). Resolution analysis in full waveform inversion. *Geophysical Journal International*, 187(3), 1604–1624. https://doi.org/10.1111/j.1365-246X.2011.05218.x
- Foulger, G. R., Panza, G. F., Artemieva, I. M., Bastow, I. D., Cammarano, F., Doglioni, C., Evans, J. R., Hamilton, W. B., Julian, B. R., Lustrino, M., Thybo, H., & Yanovskaya, T. B. (2015). What lies deep in the mantle below? *Eos.* https://doi.org/ 10.1029/2015EO034319
- Gentle, J. E. (2007). Matrix Algebra: Theory, Computations, and Applications in Statistics. New York: Springer.
- Golub, G. (1965). Numerical methods for solving linear least squares problems. *Numerische Mathematik*, 7(3), 206–216. https://doi.org/10.1007/BF01436075
- Golub, G., & Kahan, W. (1965). Calculating the Singular Values and Pseudo-Inverse of a Matrix. *Journal of the Society for*

Industrial and Applied Mathematics Series B Numerical Analysis, 2(2), 205–224. https://doi.org/10.1137/0702016

- Golub, G. H., & Reinsch, C. (1970). Singular value decomposition and least squares solutions. *Numerische Mathematik*, 14(5), 403–420. https://doi.org/10.1007/BF02163027
- Hansen, P. C. (1992). Analysis of Discrete Ill-Posed Problems by Means of the L-Curve. SIAM Review, 34(4), 561–580. https://doi. org/10.1137/1034115
- Herrmann, R. B. (2013). Computer Programs in Seismology: An Evolving Tool for Instruction and Research. *Seismological Research Letters*, 84(6), 1081–1088. https://doi.org/10.1785/ 0220110096
- Jackson, D. D. (1972). Interpretation of Inaccurate, Insufficient and Inconsistent Data. *Geophysical journal of the Royal Astronomi*cal Society, 28(2), 97–109. https://doi.org/10.1111/j.1365-246X. 1972.tb06115.x
- Katamreddy, S. H., & Yalavarthy, P. K. (2012). Model-resolution based regularization improves near infrared diffuse optical tomography. *Journal of the Optical Society of America A*, 29(5), 649–656. https://doi.org/10.1364/JOSAA.29.000649
- Knopoff, L. (1972). Observation and inversion of surface-wave dispersion. *Tectonophysics*, 13(1–4), 497–519. https://doi.org/10. 1016/0040-1951(72)90035-2
- Lawson, C. L., & Hanson, R. J. (1995). Solving Least Squares Problems. Philadelphia: Society for Industrial and Applied Mathematics.
- Lebedev, S., & Nolet, G. (2003). Upper mantle beneath Southeast Asia from S velocity tomography. *Journal of Geophysical Research*, 108(B1), 2048. https://doi.org/10.1029/2000JB000073
- Levenberg, K. (1944). A method for the solution of certain nonlinear problems in least squares. *Quarterly of Applied Mathematics*, 2, 164–168.
- Lévěque, J.-J., Rivera, L., & Wittlinger, G. (1993). On the use of the checker-board test to assess the resolution of tomographic inversions. *Geophysical Journal International*, 115(1), 313–318. https://doi.org/10.1111/j.1365-246X.1993.tb05605.x
- Lin, Y.-P., Zhao, L., & Hung, S.-H. (2014). Full-wave multiscale anisotropy tomography in Southern California. *Geophysical Research Letters*. https://doi.org/10.1002/2014GL061855
- Lütkenhöner, B., de Peralta, G., & Menendez, R. (1997). The resolution-field concept (in eng). *Electroencephalography and Clinical Neurophysiology*, 102(4), 326–334. https://doi.org/10. 1016/s0013-4694(96)96590-6
- Ma, Z., Masters, G., Laske, G., & Pasyanos, M. (2014). A comprehensive dispersion model of surface wave phase and group velocity for the globe. *Geophysical Journal International*, 199(1), 113–135. https://doi.org/10.1093/gji/ggu246
- Menke, W. (1989). *Geophysical data analysis: Discrete inverse theory* (revised). San Diego: Academic Press.
- Menke, W. (2012). *Geophysical data analysis: Discrete inverse theory* (3rd ed.). San Diago: Academic Press.
- Menke, W. (2015). Review of the Generalized Least Squares Method. Surveys in Geophysics, 36, 1–25. https://doi.org/10. 1007/s10712-014-9303-1
- Miller, C. R., & Routh, P. S. (2007). Resolution analysis of geophysical images: Comparison between point spread function and region of data influence measures. *Geophysical Prospecting*, 55(6), 835–852. https://doi.org/10.1111/j.1365-2478.2007.00640. x

- Moore, E. H. (1920). On the Reciprocal of the General Algebraic Matrix (Abstract). Bulletin of American Mathematical Society, 26, 394–395.
- Morozov, V. A. (1984). *Methods for Solving Incorrectly Posed Problems*. New York: Springer.
- Nolet, G. (2008). A breviary of seismic tomography: imaging the interior of the earth and sun. Cambridge: Cambridge University Press.
- Penrose, R. (1955). A generalized inverse for matrices. Mathematical Proceedings of the Cambridge Philosophical Society, 51(3), 406–413. https://doi.org/10.1017/S0305004100030401
- Pilkington, M. (2016). Resolution measures for 3D magnetic inversions. *Geophysics*, 81(2), J1–J9. https://doi.org/10.1190/ geo2015-0081.1
- Pogue, B. W., McBride, T. O., Prewitt, J., Österberg, U. L., & Paulsen, K. D. (1999). Spatially variant regularization improves diffuse optical tomography. *Applied Optics*, 38(13), 2950–2961. https://doi.org/10.1364/AO.38.002950
- Ren, Z., & Kalscheuer, T. (2020). Uncertainty and resolution analysis of 2D and 3D inversion models computed from geophysical electromagnetic data. *Surveys in Geophysics*, 41(1), 47–112. https://doi.org/10.1007/s10712-019-09567-3
- Sanny, D. R., Prakash, J., Kalva, S. K., Pramanik, M., & Yalavarthy, P. K. (2018). Spatially variant regularization based on model resolution and fidelity embedding characteristics improves photoacoustic tomography. *Journal of Biomedical Optics*, 23(10), 100502.
- Sigloch, K. (2011). Mantle provinces under North America from multifrequency P wave tomography. *Geochemistry, Geophysics, Geosystems.* https://doi.org/10.1029/2010GC003421
- Smith, S. W. (1997). The Scientist and Engineer's Guide to Digital Signal Processing. Pasadena: California Technical Pub.
- Snoke, 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.
- Tan, L. (2017). 3—Generalized inverse of matrix and solution of linear system equation. In L. Tan (Ed.), A Generalized Framework of Linear Multivariable Control (pp. 38–50). Oxford: Butterworth-Heinemann.
- Tarantola, A., & Valette, B. (1982). Generalized nonlinear inverse problems solved using the least squares criterion. *Reviews of Geophysics*, 20(2), 219–232. https://doi.org/10.1029/ RG020i002p00219
- Tewarson, R. P. (1977). Use of smoothing and damping techniques in the solution of nonlinear equations. *SIAM Review*, 19(1), 35–45.
- Thurber, C. H., & Ritsema, J. (2009). Theory and observations— Seismic tomography and inversion methods. In B. Romanowicz & A. Dziewonski (Eds.), *Treatise on geophysics: Seismology and structure of the earth* (pp. 323–360). Amsterdam: Elsevier.
- Tikhonov, A. N. (1963). On the solution of ill-posed problems and the method of regularization. *Doklady Akademii Nauk*, 151(3), 501–504.
- Trampert, J., Fichtner, A., & Ritsema, J. (2013). Resolution tests revisited: The power of random numbers. *Geophysical Journal International*, 192(2), 676–680. https://doi.org/10.1093/gji/ ggs057
- Varah, J. M. (1973). On the numerical solution of Ill-conditioned linear systems with applications to Ill-posed problems. *SIAM Journal on Numerical Analysis*, 10, 257–267.

- Wessel, P., & Smith, W. H. F. (1991). Free software helps map and display data. *Eos, Transactions of the American Geophysical Union*, 72, 441. https://doi.org/10.1029/90EO00319
- Wiggins, R. A. (1972). The general linear inverse problem: Implication of surface waves and free oscillations for Earth structure. *Reviews of Geophysics*, 10(1), 251–285. https://doi.org/ 10.1029/RG010i001p00251
- Xia, J., Miller, R. D., & Park, C. B. (1999). Estimation of nearsurface shear-wave velocity by inversion of Rayleigh waves. *Geophysics*, 64(3), 691–700.
- Yao, Z. S., Roberts, R. G., & Tryggvason, A. (1999). Calculating resolution and covariance matrices for seismic tomography with the LSQR method. *Geophysical Journal International*, 138(3), 886–894. https://doi.org/10.1046/j.1365-246x.1999.00925.x

(Received September 6, 2022, revised October 24, 2022, accepted December 6, 2022, Published online December 23, 2022)