Direct stress field estimation through waveform matching

Wenhuan Kuang$^1$ and Jie Zhang$^1$

$^1$Department of Geophysics, University of Science and Technology of China, Hefei, Anhui 230026, P. R. China.

First author: Wenhuan Kuang (kuangwh@ustc.edu.cn)
Corresponding author: Jie Zhang (jzhang25@ustc.edu.cn)

Key words: Geomechanics; Waveform Inversion; Seismicity and Tectonics.

Summary

Conventionally, the routine workflow of stress field estimation from seismic data consists of two steps: focal mechanism inversion and stress inversion. This two-step workflow suffers from the cumulative uncertainties of both the focal mechanism inversion process and stress inversion process. To mitigate the cumulative errors, a few previous studies have put efforts to directly estimate the stress field using P-wave polarities. In this study, we develop a new approach to directly estimate tectonic stress fields with better accuracy through waveform matching. This new approach combines the two steps into a one-step workflow to mitigate the cumulative uncertainties through the physical relationship between a stress field and the recorded
waveforms. This method assumes a homogeneous stress field in space in the local source region and that the fault slip occurs in the direction of the resolved shear stress acting on the fault plane. Under these assumptions, the stress pattern that generates the theoretical waveforms that best fit the waveforms observed is directly retrieved as the true stress field. The merits of the new approach include that this approach can mitigate the cumulative uncertainties suffered from the conventional two-step workflow and does not require determination of the focal mechanisms of each event; thus, this method is applicable to datasets with few stations. Synthetic tests with and without noise are conducted to demonstrate the performance and merits of this method. Then, the new approach is applied to a real dataset from central Oklahoma between March 2013 and March 2016. The resulting stress pattern is consistent with that estimated from previous studies examining the same region. These applications show the benefits and validity of the new approach.

1 Introduction

Conventionally, local and regional tectonic stress fields are determined from a group of well-constrained earthquake focal mechanisms (Zoback 1992; Hardebeck & Shearer 2003; Tan & Helmberger 2007; Zoback 2007; Hardbeck 2010, 2014; Zhang et al. 2014; Kuang et al. 2017; Alt & Zoback 2017). Several methods such as Michael’s method (Michael 1984, 1987), Angelier’s method (Angelier 2002), and Gephart and Forsyth’s method (Gephart & Forsyth, 1984) have been proposed for the determination of stress patterns from a set of focal mechanisms of earthquakes. These kinds of methods are usually realized in two steps. The first step is to determine the focal mechanisms of medium to large earthquakes. Then, the second step is to apply a stress inversion process to the focal mechanisms that are determined from the first step to
estimate the stress field. However, when the focal mechanisms cannot be precisely determined, or there are great ambiguities in the focal mechanisms determined (such as only a few stations are available), the stress field estimated from the conventional two-step workflow has high uncertainties. Furthermore, the two-step inversion workflow may include cumulative uncertainties because each implementation of the two steps accumulates errors from each inversion process.

To mitigate the cumulative errors from the two-step workflow, in several previous studies, efforts to skip the focal mechanism inversion and directly estimate the stress field have been made. Rivera & Cisternas (1990), Horiuchi et al. (1995), and Abers & Gephart (2001) proposed approaches toward finding the best stress field to explain the P-wave first-motions observed from multiple events. In these approaches, the stress field can be found when only a few P-wave polarities are obtained from a single event, as long as there is an adequate number of events.

Furthermore, Iwata (2018) expands their work to incorporate the spatial heterogeneity of a stress pattern without determining the focal mechanisms, where the spatial pattern is directly inferred from P-wave first-motions.

Methods based on P-wave first motions have the merit of low computational cost and can be applied to small earthquakes where a fair number of identifiable first motions are available. However, the input data of this kind of methods is only the simple sign of up or down first motion polarity of the recorded waveform, thus a large amount of first motions are needed to reliably constrain the stress field. The requirement of a large amount of available first motions might not always be satisfied in some cases (Hardebeck and Shearer 2003), especially in cases where the recording stations are very sparse. On the contrary, the full waveform carries substantial information over P-wave first motion polarities, hence adding more waveform
amplitude information can remedy the scarcity of first motions (Tan and Helmberger 2007). Thus introducing full waveform information other than using simple P-wave first motion polarities has the potential to better constrain the stress field in cases where only few recording stations are available.

Considering that full waveforms can provide more constraints for the stress field, in this study, we develop a new approach to directly estimate the stress field based on waveform matching. This new approach combines the two-step workflow into a one-step workflow to mitigate the cumulative inversion errors. This approach first builds the direct physical relationship between the stress field and the waveforms recorded. Then, a neighborhood algorithm (Sambridge 1999a; Sambridge 1999b) is adopted to directly invert the four parameters of the stress field through waveform matching. The stress field that generates the theoretical waveforms that best fit the observed waveforms is determined to be the true stress field. The new method is examined for both synthetic and real data to demonstrate the merits and validity of the method. Therefore, to better illustrate the new approach, the main body of this article is organized into the following three parts: theory and methods, two synthetic tests with and without noise, and real data application.

2 Theory and Methods

2.1 General workflow comparison

To better compare the new approach with the conventional workflow, we show a general flowchart comparison between the conventional two-step workflow and the new workflow in
Fig. 1. The input and output are recorded waveforms and four stress parameters respectively, and are the same for both the conventional and new workflows.

The conventional workflow involves two steps. The first step is the focal mechanism inversion, which usually uses the P-wave first motion method (Hardebeck & Shearer 2002, 2003), P/S ratio method (Song & Helmberger 1998; Tan & Helmberger 2007), classic CAP method (Zhao & Helmberger 1994; Zhu & Helmberger 1996), or other waveform matching method (Zhang et al. 2014). The second step is the stress inversion (Michael 1987; Angelier 2002; Gephart & Forsyth 1984), which uses the inverted focal mechanisms from the first step. Each step usually results in uncertainties from the inversion process, for example, the focal mechanism inversion results from different researchers can have errors of 5 to 10 degrees (Hardebeck & Shearer 2002; Shelly et al. 2016) in the strike, dip and rake angles due to the station coverage, signal-to-noise ratio, velocity uncertainties, or different inversion schemes. The stress inversion process also has uncertainties because of the linearization assumption or the true fault plane selection (Michael 1984, 1987). Since the second step is based on the inverted focal mechanism results from the first step, the final results usually accumulate inversion uncertainties from both inversion steps.

The new workflow involves only one inversion process. Through reasonable geomechanical assumptions, we can build the direct physical relationship between the stress field and the waveforms recorded. The new workflow contains the following parts. For a given sets of initial stress fields, we first predict the rake angle for a group of fault orientations that satisfy the stress field. Then, we simulate the theoretical waveforms and compare the theoretical waveforms with the observed waveforms using an objective function. Thus, a single inversion loop is constructed to invert the four stress parameters. If the stress field can generate the best-fitted theoretical waveforms with the observed data, then the inversion loop is terminated and the current stress
field is estimated as the true stress field. Otherwise, the inversion regenerates the stress field and the loop continues until the exit condition is satisfied.

2.2 Slip angle prediction

To estimate the relative stress pattern in a local region, we assume that (1) the tectonic stress is uniform (homogeneous) in the source region, (2) the slip vector points in the direction of the resolved shear stress on the fault, which is also called the Wallace-Bott hypothesis (Wallace 1951; Bott 1959). The relative stress pattern contains the orientations of the three principal axes (S1>S2>S3, where S1, S2, and S3 are the maximum, intermediate, and minimum principal stresses, respectively) and the value of Bishop’s stress ratio Φ (Φ = (S2-S3) / (S1-S3)), which defines the relative magnitudes of the three principal stresses, following Bishop (1966). The orientations of the three axes are represented by three Euler angles a, b, and c (the definition of three Euler angles: a = trend of S1, b = -plunge of S1, and c = rake S2. However, if S1 is vertical, a = trend of SHmax - π/2, b = -trend of S1, and c=0, following Zoback 2007). Conventionally, with only slip data, the deviatoric stress tensor can be estimated, and the isotropic component cannot be determined (Michael 1984). The relative magnitudes between the three principal stresses, which is defined as parameter Φ, can be determined. Therefore, the goal in this study is to estimate the three Euler angles and the parameter Φ. Assuming the direction of fault slip is along the shear traction, we have

\[
s = \frac{\tau(n, \sigma)}{|\tau(n, \sigma)|}
\]
where \( s \) is fault slip and \( \tau(n, \sigma) \) is shear traction, \( n \) is the normal vector of the fault plane, and \( \sigma \) is the deviatoric stress tensor. The equation can be linearized if we assume that the shear stresses are equal on all slipping faults (Michael 1984),

\[
A'\Sigma = SV^i
\]  

(2)

where \( A \) is the fault orientation parameters (i.e., strike and dip), \( \Sigma \) is the stress tensor, \( SV \) is slip vector, and \( i \) denotes the event index. Thus, the deviatoric stress tensor is formulated in a linearized equation and can be inverted with known focal mechanisms (i.e., strike, dip and rake). This method is the conventional approach of stress inversion based on inverted focal plane mechanisms.

In this study, to directly estimate the stress pattern through waveform matching, we first need to build the forward physical relationship between the stress field and the waveforms recorded.

Supposing the three Euler angles, \( \Phi \), and the orientation of the fault plane are given, following the second assumption, the expected slip direction on the fault can be theoretically estimated. The stress tensor is a second-order tensor. Following Zoback (2007), in a principal coordinate system, the stress tensor \( S \) at depth is denoted,

\[
S = \begin{bmatrix}
S_1 & 0 & 0 \\
0 & S_2 & 0 \\
0 & 0 & S_3 \\
\end{bmatrix}
\]  

(3)

Assuming three Euler angles of \( a, b, \) and \( c \) that denote the orientations of the three principal stresses are given, we can first transform the stress tensor into a geographical coordinate system,

\[
S_g = R_f^* S R_f 
\]  

(4)
where $S_g$ denotes the stress tensor in the geographical coordinate system. Then, in the geographic coordinate system, the stress tensor can be projected onto an arbitrarily orientated fault plane,

$$S_f = R_2 \ast S_g \ast R_1$$  \hspace{1cm} (5)$$

where $S_f$ denotes the stress tensor in the fault plane coordinate system. The detailed form of rotation matrix $R_1$ and $R_2$ can be found in supplementary material or refer to Zoback (2007). Thus, by analyzing the stress vector on the fault plane (the stress tensor is projected onto the fault plane to derive the stress vector), which is denoted as $S_f(3,1)$ and $S_f(3,2)$, we can predict the slip direction under a given stress field.

2.3 Objective function

Assuming that the double couple source model is applicable, conventionally, the theoretical waveforms can be formulated as,

$$P = SR(str, dip, rake) \ast G$$  \hspace{1cm} (6)$$

where SR denotes the source term and G denotes Green’s function. Since the slip angle can be directly predicted with a known fault plane and stress tensor, this equation can be rewritten as,

$$P = SR(str, dip, a, b, c, \Phi) \ast G$$  \hspace{1cm} (7)$$

Thus, with precalculated Green’s functions, the stress field is directly linked to the waveforms recorded. In a small study area, though the focal mechanisms may vary, all the earthquake events should share the same stress pattern. Hence, an obvious benefit of the formulation is that in cases where only a few stations are available and the focal mechanisms cannot be well-constrained, the proposed method can constrain the local stress pattern if there are enough events because all the
events share the same stress pattern and each event can individually provide consistent constraints to the stress field.

With formula (7), we can directly determine the stress pattern by evaluating the fit between the theoretical and recorded waveforms. The fault orientation parameters (strike and dip) are unknown and need to be searched.

To evaluate the goodness-of-fit of the target stress parameters, an objective function that measures the fit between the theoretical waveforms and observed waveforms is constructed using a cross-correlation criteria:

\[ \text{obj} = \{ P_{\text{syn}} \otimes P_{\text{obs}} \}_{\text{max}} \]  

where \( P_{\text{syn}} \) and \( P_{\text{obs}} \) are the theoretical and observed waveforms, respectively. The symbol \( \otimes \) represents the cross-correlation operation. In the operation, the cross-correlation coefficients are computed for each pair of a theoretical and an observed waveform at individual station. Each pair of the coefficients is maximized for each station and then the average of the maximized coefficients at all stations is computed to derive the final correlation coefficient. The time lags of vertical and radial components are tied to be the same, but allow different time lag for tangential component to account for possible azimuthal anisotropy (Zhu \\& Helmberger 1996). The sum of the two coefficients for the vertical and radial components and the one for the tangential component are maximized by changing the time lags. Then the sum of the maximized sum of the two coefficients and the maximized one for the tangential component is divided by three. Therefore, the objective function measures the match level between all the theoretical and observed waveforms. The stress field that can maximize the objective function is estimated as the true stress pattern. In implementation, since the magnitudes of events may not be accurately
determined, a normalization is applied to both the theoretical and the observed waveforms based on the maximum amplitude.

2.4 Inversion algorithm

To estimate the four parameters of the stress pattern, we adopt the neighborhood algorithm (Sambridge 1999a; Sambridge 1999b) to invert the target stress parameters. The neighborhood algorithm initially searches the global parameter space with a sparse grid and subsequently iteratively generates the neighborhood model space using the rank of the objective function (Sambridge 1999a); thus, this algorithm can produce a globally optimized solution. In this study, the neighborhood algorithm is very suitable for our multiparameter inversion problem and thus is selected for inverting the three Euler angles and the parameter Φ. Since the orientation of the fault is unknown, when each trial stress pattern is tested, we go through all the fault orientations and select the fault plane with the maximum objective function. To guarantee the uniform sampling over the focal sphere, referring Abers & Gephart’s (2001) work, our search scheme is to utilize a suite of fault plane poles which evenly samples all fault orientations, utilizing the spherical mesh algorithms of Baumgardner and Frederickson (1985). Through the search of focal plane orientation and the neighborhood algorithm, each trial stress pattern is directly evaluated for the best match level between the theoretical waveform and the observed waveforms.

Thus, after several iterations through the neighborhood algorithm, the stress pattern that generates the theoretical waveforms that best fit the observed waveforms is retrieved as the true stress pattern. The ranges for the three Euler angles a, b, and c are 0°-360°, 0°-90°, and 0°-90°, respectively. The value of parameter Φ ranges from 0 to 1. After testing, for the initial parameterization, the intervals for a, b, c, and Φ are set as 45°, 22.5°, 22.5°, and 0.2, respectively. Therefore, a total of 640 (8*4*4*5) possible solutions are searched in the first iteration. For each
subsequent iteration, the model space is regenerated based on the ranking of the objective functions. After ranking, the top 40 solutions are selected as seeds to reproduce a new solution space with smaller intervals for later iterations. The reproduced number of possible solutions for subsequent iteration is kept as 640.

For each trial stress pattern, a number of theoretical waveforms should be modeled. Therefore, for both synthetic tests and real data applications, Green’s functions are precalculated with a known velocity model. Green’s functions are computed through the elastic wave modeling of a point earthquake source in a multilayered half-space using the Thompson-Haskell propagator matrix technique (Zhu & Rivera 2002). In this study, for both synthetic tests and real data applications, the theoretical and recorded waveforms are both filtered between 10 and 25 seconds before matching.

To evaluate the performance of the method, we conducted two synthetic tests and a real data application. The first synthetic test is noise free (we assume that the event locations and the velocity model are accurate) to test the validity of the method. The second synthetic test includes errors in both the event locations and the velocity model. The two synthetic tests are used to show the merits and sensitivity of the proposed new approach. Then, we apply the method to a real dataset from central Oklahoma and examine the results with other previous studies in the same region to show the validity of the method.

### 3 Synthetic Tests

Fig. 2 shows the data setup for the synthetic tests. Ten events and three recording stations are designed. The size of the synthetic tests is approximately 100 km in both the X- and Y-directions.
The blue dots are true event locations that are used to generate the “observed” data for the synthetic tests. The black triangles are three recording stations. The true stress pattern is set as 60°, 10°, 80°, and 0.7 for parameters a, b, c, and Φ value, respectively. Thus this is a strike slip stress regime (Further, a normal faulting stress regime case is also tested in supplementary materials). This true stress pattern is used to generate the “observed” data in the synthetic tests. The fault orientations are randomly assumed for each event. For the first synthetic test, the true event locations and the true velocity model (solid lines in Fig. 3) are used. The first synthetic test examines the validity of the method. For the second synthetic test, the event locations are randomly perturbed in both horizontal and depth direction with a maximum value of 2 km, as shown by the red dots in Fig. 2. The P- and S-wave velocities in each layer are randomly perturbed with a maximum value of 5 percent, as shown by the dashed lines in Fig. 3. The second synthetic test with location and velocity errors examines the stability and merits of the method. In addition to these differences, both synthetic tests have the same true stress pattern and the same “observed” data.

3.1 Synthetic test 1

Based on the proposed new workflow, we implement the first synthetic test for noise free data using the neighborhood algorithm. Fig. 4 depicts the inversion results of the four stress parameters for the first synthetic test. A total of 12 iterations are implemented. The evolution of the three Euler angles and Φ value with each iteration are shown in the subplots. The dots in each subplot denote the trial seeds of the regenerated parameters. The inverted stress parameters of each iteration are color-coded by the values of the objective functions. The red dashed lines show the true values of each parameter that is used to generate the “observed” data. With increasing number of iterations, the stress parameters gradually converge to the global true solution, and the
values of the objective functions gradually increase to the maximum value. Therefore, for the noise free data, we can directly determine the stress pattern through waveform matching using the proposed method. For each iteration, 640 sets of stress parameters are maintained.

Fig. 5 shows, for one event, the comparisons between the theoretical and the “observed” waveforms of one set of the initial stress parameters (Fig. 5a) and those of the inverted final stress parameters (Fig. 5b). The “observed” data are colored blue and the theoretical waveforms are colored red. The waveforms are aligned based on the maximum cross-correlation. The figure shows that when the stress field is not correct, the predicted slip direction of the faults cannot produce waveforms with a good fit. After the inversion, as shown in Fig. 5b, the waveforms can be fitted well with the correctly predicted slip direction by the true stress pattern. The blue waveforms in Fig. 5b are fully covered by the red waveforms, since the waveforms perfectly match for the noise free synthetic test.

Since the four stress parameters are simultaneously inverted, there may be certain trade-offs between these parameters in the inversion process. To better understand the trade-offs, we conduct a sensitivity analysis for the four parameters using the dataset of synthetic test 1. The sensitivity analysis is done by perturbing one parameter while the other three parameters are held constant at true values. Fig. 6 shows the sensitivity of each parameter in the subplot. Euler angle a and Euler angle c have very good resolution and should be estimated well. The value of Φ has moderate resolution, and Euler angle b has low resolution compared with the other three parameters.

To better measure the inversion errors of the results, additionally, we introduce Kagan angle analysis (Kagan 1991; 2007) to evaluate the uncertainties of the inverted stress field. Using the 640 solutions of the stress fields in the last iteration (from NA algorithm), we calculate the
distribution of Kagan angles as shown in Fig. 7. For noise-free data, most of the 640 Kagan angles of the last iteration range from 0° to 2°, indicating a fairly reasonably estimation resolutions of the proposed method.

To further investigate the parameter resolution with respect to different stress regime, in supplementary materials, we change the stress regime from strike slip to normal faulting and implement another synthetic test (see supplementary material) with the same geometry setup. The true stress pattern is set as 30°, 80°, 10°, and 0.7 for parameters a, b, c, and Φ value, respectively. Thus this is a normal faulting stress regime. For normal faulting case, the inversion (Fig. S1) can also gradually converge to the true stress field similar to strike slip case (Fig. 4). However, the parameter resolutions from the normal faulting case (Fig. S2) show both common and different resolutions with strike slip case (Fig. 6). The common pattern is that, Euler angles a, c, and parameter Φ have fair resolution in both strike slip and normal faulting cases. Euler angle a presents a strong local minimum for both strike slip and normal faulting cases (Fig. 6a and Fig. S2a). The difference is that, for strike slip case, Euler b presents less resolution than normal faulting case (Fig. 6b and Fig. S2b).

3.2 Synthetic test 2

We then implement the second synthetic test with location and velocity errors to test the performance of the method. For the second synthetic test, the same true stress pattern of strike slip faulting and same “observed” data as those of strike slip case in the first synthetic test are used. However, when conducting the inversion, we add random errors to event locations (red dots in Fig. 2) and velocity model (dashed lines in Fig. 3) to simulate a real scenario. Even though in reality the location errors sometime depend on the velocity errors, the detailed discussion of this dependency will be out of the scope in this paper, thus, for simplicity, we
assume randomly perturbed errors in both location and velocity. Fig. 8 shows the inversion results of the stress pattern for the second synthetic test. Similar to Fig. 4, the red dashed lines show the true values of each parameter. With maximum location errors of 2 km and maximum velocity errors of 5 percent within each layer, the inversion can still gradually converge to the global solution with remarkable precision. The inverted a, b, c, and Φ values are 61.16°, 7.22°, 80.56°, and 0.69, respectively. The differences between the resolved stress pattern and the true stress pattern for a, b, c, and Φ are 1.16°, 2.78°, 0.56°, and 0.01, respectively. The estimation errors in Euler angle a and Euler angle c are smaller than the estimation error in Euler angle b, which is consistent with the sensitivity analysis results of strike slip case (Fig. 6) in the previous section where Euler angle b presents lower resolution. The estimation error in Φ is only 0.01, which shows a very good resolution.

As a comparison, we conduct the conventional two-step workflow to show the merits of the proposed new approach. First, the focal mechanisms are inverted using a grid search algorithm based on waveform matching. The inverted focal mechanisms are shown in Fig. 9, where the black beach balls are true focal mechanisms and the red beach balls are conventionally inverted focal mechanisms. Visually, the black and red beach balls appear to be the same, which means that the focal mechanism inversion has achieved very good results. However, the results are subtly different. Table 1 shows the true fault plane values and the inverted fault plane values, where we can see that there are small differences in the strike, dip, and rake angles. These differences are estimation errors from the first step in the two-step workflow and will affect the final estimation of the stress field. Next, we conduct the stress inversion using the inverted focal mechanisms. The stress inversion code is a widely used program from Zoback’s group (2014, Stanford) and is based on Michael’s method (1984, 1987). The inverted stress results are shown
in black lines in Fig. 8. The results of conventional two-step approach (Fig. 8 black lines) are 68.5°, 8.1°, 71.5°, and 0.78 for the four parameters, with differences of 8.5°, 1.9°, 8.5°, and 0.08 compared with true values. They deviate more from the true values compared with results using the proposed method, especially for Euler c and parameter Φ. Thus, using the conventional two-step workflow, the inversion errors are larger than those of the new approach. The high uncertainties are accumulated errors and come from both the focal mechanism inversion process and stress inversion process. However, we believe the errors mainly result from the estimation errors of the strike, dip, and rake angles in the first inversion process because subtle estimation errors in the rake angle can lead to considerable stress uncertainties; the stress inversion relies on the assumption that the slip vector points in the direction of the resolved shear stress on the fault (the second assumption in section 2.1). Using the proposed method, the noisy waveforms in the second test will also have impact on the final results compared with noise-free test, but it shows less impact than using the conventional workflow. Through the second synthetic test, we conclude that for data with errors in the event locations and velocity model, we can better estimate the stress pattern by using the proposed new workflow than by using the conventional two-step workflow.

Fig. 10 shows, for the second synthetic test, comparisons between the theoretical and “observed” waveforms of one example event using one set of the finally estimated stress patterns. The “observed” data are colored blue and the theoretical waveforms are colored red. While the initial stress patterns cannot produce theoretical waveforms fitting the “observed” data, after inversion, the best estimated stress pattern can generate a theoretical waveform that fits well with the “observed” data.
In these two synthetic tests, only three recording stations are used. Conventionally, with errors in
the event locations and velocity model, the focal mechanisms determined may present great
ambiguity, and the conventional two-step workflow would result in high uncertainty. However,
in the proposed approach, even though in certain real cases, we may have only a few recording
stations, all the recorded full waveforms can provide consistent constraints for estimating the
stress field. The second synthetic test suggests this merit of the proposed waveform-matching-
based approach. Thus, this new approach can be applied to datasets with a few stations and a
large number of events.

4 Application to Real Data: Central Oklahoma Dataset

After the synthetic tests, the proposed method is applied to a real dataset taken from central
Oklahoma recorded from March 2013 to March 2016. The tectonic stress field in central
Oklahoma is overall characterized by strike-slip faulting with uniform stress orientations
(McNamara et al. 2015; Alt & Zoback 2017). Therefore, this dataset supports the first
assumption, in which we assume that the stress field is overall uniform within the source region.
We select this real dataset to test the method because the tectonic stress field in this region has
already been well-studied, and thus, we can evaluate the validity and performance of the
proposed method.

The dataset analyzed here is from the temporary Nanometrics Research Network
(https://doi.org/10.7914/SN/NX), which was operated by Nanometrics Seismological
Instruments. The seismic data recorded by this network are available and accessible at the
Incorporated Research Institutions for Seismology (IRIS) online Data Management Center
In total, 27 earthquake events with magnitudes greater than 3.5 that occurred in this region during the deployment period of this seismic network are processed in this study (Fig. 11). The largest earthquake occurred in December 2013 with a magnitude of 4.7. Sixteen available recording stations that are close to the region studied are used. The event locations are initially from the catalog of the IRIS online system. To prepare the precalculated Green’s functions, a 1D western U.S. velocity model that best fits the observed local and regional P-wave travel times (Herrmann et al. 2011; McNamara et al. 2015) is used to calculate the three-component theoretical waveforms (solid lines in Fig. 3). Green’s functions are precalculated with the known event locations and velocity model.

Fig. 12 shows the inversion results of the four stress parameters for this real data application. Similar to the synthetic tests, the evolution of the three Euler angles and Φ with each iteration are shown. The inverted stress parameters of each iteration are color-coded by the values of the objective functions. A total of 12 iterations are implemented. The red dashed lines show the values of each stress parameter that are taken from a previous study (Alt & Zoback, 2017), which are referred here as the true stress pattern. Even though in the first few iterations, certain stress parameters (Euler angle a, c, and Φ) present local maximums, after six iterations, these parameters gradually converge to the global maximum. The values of the objective functions gradually increase to the maximum with increasing number of iterations. According to a previous study based on borehole measurements in this region, the maximum principal stress axis is slightly inclined from the horizontal plane with a direction around N80°-85°E (Alt & Zoback 2017). From the detailed results of their study, the epicenters of the analyzed events in our study are distributed only over the southern part of area 3 and the northern part of 5 of their study. The stress field is overall consistent strike-slip stress regime with the SHmax azimuth at about N85°E.
in central Oklahoma (north Oklahoma/South Kansas presents strike-slip faulting mixed with normal faulting). Thus for validation purpose, we derive the reference values by averaging their results in area 3 and area 5. The average values of a, b, c, and Φ in this region from their results are about 84°, 2.25°, 89.0°, and 0.76, respectively (dashed red lines in Fig. 12). From our proposed new approach using seismic data, the inverted a, b, c, and Φ are 86.04°, 4.68°, 87.31°, and 0.77, respectively. The differences between the estimated stress pattern of the proposed method and that of the previous study for a, b, c, and Φ are 2.04°, 2.43°, 1.69°, and 0.01, respectively. These results are essentially consistent. Therefore, the real data application proves the validity of the proposed new approach; we can directly estimate these four stress parameters through waveform matching.

Fig. 13 shows, for one event with the real data, comparisons between the theoretical and the recorded waveforms using one set of the finally inverted stress patterns. The recorded data are colored blue, and the theoretical waveforms are colored red. For the initial parameterization, the predicted slip direction on all faults cannot produce theoretical waveforms that fit well with the recorded data. After the inversion converges to the global solution, the recorded waveforms fit the theoretical waveforms generated by the final stress pattern very well.

Once the stress pattern is estimated, the orientations of the faults that generate the best matched theoretical waveforms are returned and retrieved. The slip direction can be predicted by the estimated stress pattern. Thus, the focal mechanisms can be retrieved as a byproduct for quality control purposes or other uses. Fig. 14 shows the focal mechanisms retrieved by the best inverted stress pattern. The sizes of the beach balls are based on the magnitudes. Overall, the results are characterized by strike-slip faulting in this region, and the nodal planes steeply dip. For most focal mechanisms, the direction of one nodal plane of each event shows a favorable slipping
angle of approximately 30° to the maximum horizontal stress orientation estimated
(approximately N85°E).

5 Discussion

The essential idea of the proposed approach is building a direct physical relationship between a
stress field and waveforms. Once the forward problem is built, the neighborhood algorithm is
adopted to solve the inversion problem. The determination of focal mechanisms is not needed in
one-step workflow. Thus, based on full waveform matching, the new method can be applied in
cases where only a few stations are available. In conventional two-step workflow, the focal
mechanism inversion is done for individual event, thus they do not have a global constrain for all
events at a time. Moreover, the estimated uncertainties in stress parameters in some cases might
differ substantially (Balfour et al. 2005; Hardebeck & Hauksson 2001) and might be attributed to
the selection of the nodal plane of each focal mechanism (Carlson et al. 2018). Conversely, the
proposed new workflow and other one-step methods can mitigate the dilemma of how to choose
the actual fault plane. Additionally, one-step approach can provide such a global constrain for all
the events to share the same stress pattern, thus we think one-step workflow has the potential to
reduce the cumulated errors of two-step workflow.

In the proposed method, since we are using the full waveform information, the computational
cost of such a waveform based method is heavier than that of the P-wave first motion methods
(due to the computation of waveforms). Thus, one disadvantage of the proposed method is that it
would require more computation time to finish the inversion. Other disadvantages of the
proposed new methods include the difficulty of quantitative evaluation of the uncertainties in
inverted stress field, the need of a reasonably velocity model, and good signal to noise ratio of the recorded data.

For validation purpose of the new method, in this study we apply the method to median natural earthquakes using low frequency full waveforms. Theoretically, the method can be extended to use body waves (short-period waves of small earthquakes), as long as the body waves have fair signal-to-noise ratios and can be reasonably modelled. But considering the difficulties in modeling full waveforms of small earthquakes (ML<3) with insufficient velocity model, in current practice it is still challenging to apply the proposed method to small earthquakes whereas the conventional methods with the two-step workflow and the P-wave methods are sometimes applicable.

A homogeneous stress state is assumed in the current study. For large earthquake, the stress field could be very complicated due to the dynamic rupture process/effects. Especially during the rupture process of large earthquake, the stress field is not only changing spatially but also has time-dependent changes, and the point source assumption might also be not valid, which makes both the conventional methods and the proposed method difficult to apply. Alternatively, for cases of spatial stress complexity, if with adequate earthquakes, we can divide the whole study area into subareas to study the detailed stress field, similar to the strategy done by Alt & Zoback (2017). Further, based on waveform matching, an inversion strategy for spatial stress field estimation with an objective smoothness constraint (Carlson et al. 2018; Iwata 2018) is needed to be developed/extended.

Regarding the geophysical implication from the sensitivity analysis results of the four parameters, it is clear that the Euler angle a has high resolution in both strike-slip faulting and normal faulting cases. Recalling the definition of Euler angle a (it is defined as the trend of S1 in strike-
slip faulting case and the trend of $\text{SHmax}-\pi/2$ in normal faulting case), in both cases Euler angle $\alpha$ is related to the azimuth of horizontal maximum stress. Thus the sensitivity analysis results indicate that the azimuth of horizontal maximum stress tends to have high resolution using the proposed method and it should be well constrained in real data applications. Additionally the $\Phi$ values in both test results only present mild resolution, which indicates the determination resolution of the relative magnitude of the three principal stresses is not very high. Other two parameters show variable resolutions, thus are difficult to draw conclusion.

In the synthetic tests, the introduced errors in location and velocity generally follow the uncertainties in real scenario. For low frequency recordings, the waveforms are insensitive to local heterogeneities, a 1D velocity model is applicable in most cases. However, if the velocity model is severely inaccurate, both the conventional focal mechanism inversion and the proposed method might tend to fail, because they both depend on the full waveform modelling. Thus its successful application will depend on the accuracy of the modelled waveforms using pre-determined event locations, depth, and/or the velocity model. In the proposed method, regarding the frequency content used, the assumption of layered media is a satisfactory approach and that the waveforms are expected to have a good signal-to-noise ratio in that frequency range.

Nevertheless, with the rapid development in computational efficiency, for future improvement the event location and depth can be simultaneously inverted together with the stress parameters to mitigate the effect of errors in event location and depth.
6 Conclusions

In this study, we propose a new approach to directly estimate a stress field based on waveform matching. This new approach can mitigate the cumulative errors of the conventional two-step workflow. In this new approach, the stress field can be estimated without focal mechanism determination. The proposed method is also applicable to datasets with only a few stations. Both synthetic tests and an application to real data are designed to examine the performance of this new method. Two synthetic tests are designed to show the merits and advantages of the proposed method over the conventional two-step workflow, considering realistic errors in the event locations and velocity model. The synthetic results show that the stress field can be better inverted with the new workflow than with the conventional two-step workflow because the direct estimation approach can mitigate cumulative errors. The application of the method to a real dataset from central Oklahoma reveals stress parameters consistent with the results of previous studies examining the same region. Thus, through both synthetic tests and a real data application, we demonstrate the benefits and validity and the proposed new approach.

Acknowledgments

We thank the financial support of The National Key Research and Development Program of China with grant number 2018YFC1504003-01, thank the support of National Earthquake Reporting Research Program of China Earthquake Administration with the grant number 201408018. The waveform data used in this study, from 16 seismic stations in seismic network NX at Oklahoma, are archived and available for download from the IRIS Data Management Center (DMC). Additionally, we greatly thank Prof. Haijiang Zhang for helpful discussions.
References


Figure 1. Flowchart showing the comparison between conventional two-step workflow vs. the new direct inversion workflow.
Figure 2. The geometry setup for synthetic tests. Three recording stations (black triangles) are used for the synthetic tests. The true event locations are denoted in blue dots and the perturbed event locations are denoted in red dots. (a) Map view. (b) Depth view. Totally 10 events are used.
Figure 3. The P- and S-wave velocity models used in this study. The solid red and blue lines are the true velocity model. The dashed red and blue lines are the perturbed velocity model. The true velocity comes from previous study for central Oklahoma. Thus it is subsequently used for the following real data application.
Figure 4. The inversion results for the first synthetic test (noise free test) using neighborhood algorithm. The four stress parameters are color-coded by the values of objective function. The theoretical and “observed” data are normalized based on maximum amplitude, thus the maximum value of objective function is 1 when it is a perfect match. The color of each dot of each parameter denotes the average objective function over all possible trial stress fields of varying other three parameters. Dashed red lines are the true values of each stress parameter. (a) Evolution with iterations for Euler angle a. (b) Evolution with iterations for Euler angle b. (c) Evolution with iterations for Euler angle c. (d) Evolution with iterations for parameter $\Phi$. 
(a),

![Graphs showing waveforms for Radial, Transverse, and Vertical components over time and receiver index.](https://academic.oup.com/gji/advance-article-abstract/doi/10.1093/gji/ggaa034/5710936)
Figure 5. Waveform comparison of an example event between the theoretical (red) and “observed” (blue) waveforms under different stress field for the first synthetic test. (a) Waveform comparison using one set of the initial stress patterns. (b) Waveform comparison using the finally inverted stress pattern. Please note, since this is a noise free synthetic test, the blue waveforms are fully covered by red waveform in (b).
Figure 6. The sensitivity analysis for the four stress parameters for the proposed approach. This analysis is done by perturbing one parameter while setting true values to the other three parameters. (a) The sensitivity distribution for Euler angle a. (b) The sensitivity distribution for Euler angle b. (c) The sensitivity distribution for Euler angle c. (d) The sensitivity distribution for parameter $\Phi$. 
Figure 7. The Kagan angel analysis for the inverted stress fields of the last iteration using the proposed approach. Totally 640 stress fields in the last iteration is included in this analysis.
Figure 8. The inversion results for the second synthetic test (with location and velocity errors).

Dashed red lines are the true values of each stress parameter. The color of each dot of each parameter denotes the average objective function over all possible trial stress fields of varying other three parameters. Dashed black lines are the results from conventional two-step workflow. (a) Evolution with iterations for Euler angle a. (b) Evolution with iterations for Euler angle b. (c) Evolution with iterations for Euler angle c. (d) Evolution with iterations for parameter Φ.
Figure 9. The comparison between true focal mechanisms and the inverted focal mechanisms using conventional method through waveform matching. Upper black beach balls are true focal mechanisms and lower red beach balls are inverted focal mechanisms. Visually they are almost the same with minor differences.
Figure 10. Waveform comparison of an example event between the theoretical and “observed” waveforms for the second synthetic test using the finally inverted stress pattern.
Figure 11. Map showing the seismic data and recording stations used for real data application.

16 recording stations (black triangles) are available during the deployment period for this dataset. The cataloged event locations are denoted in red dots. Totally 27 events with magnitudes larger than 3.5 are used.
Figure 12. The inversion results for the real data application. The stress parameters are color-coded by the values of objective function. The theoretical and recorded data are normalized based on maximum amplitude, thus the maximum value of objective function is 1 when it is a perfect match. The color of each dot of each parameter denotes the average objective function over all possible trial stress fields of varying other three parameters. Dashed red lines are the results from previous study. (a) Evolution with iterations for Euler angle a. (b) Evolution with iterations for Euler angle b. (c) Evolution with iterations for Euler angle c. (d) Evolution with iterations for parameter Φ.
Figure 13. Waveform comparison of one example event between the theoretical and recorded waveforms for the real data application using the finally estimated stress pattern.
Figure 14. The predicted focal mechanisms using the estimated stress field. The sizes are plotted based on magnitudes. Overall the focal mechanisms show consistent strike-slip faulting in central Oklahoma.
Table 1. True focal mechanisms vs. inverted focal mechanisms using conventional grid search method via waveform matching for the second synthetic test. Here for better comparison, only one nodal fault plane for each event is shown while the other nodal plane can be easily calculated using the known nodal plane.

<table>
<thead>
<tr>
<th>Event ID</th>
<th>True strike</th>
<th>True dip</th>
<th>True rake</th>
<th>Inverted strike</th>
<th>Inverted dip</th>
<th>Inverted rake</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>210</td>
<td>75</td>
<td>156</td>
<td>213</td>
<td>76</td>
<td>154</td>
</tr>
<tr>
<td>2</td>
<td>50</td>
<td>55</td>
<td>126</td>
<td>43</td>
<td>50</td>
<td>112</td>
</tr>
<tr>
<td>3</td>
<td>10</td>
<td>45</td>
<td>174</td>
<td>15</td>
<td>45</td>
<td>172</td>
</tr>
<tr>
<td>4</td>
<td>70</td>
<td>25</td>
<td>84</td>
<td>70</td>
<td>25</td>
<td>85</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>90</td>
<td>-175</td>
<td>21</td>
<td>85</td>
<td>-164</td>
</tr>
<tr>
<td>6</td>
<td>190</td>
<td>85</td>
<td>175</td>
<td>179</td>
<td>85</td>
<td>174</td>
</tr>
<tr>
<td>7</td>
<td>270</td>
<td>85</td>
<td>22</td>
<td>270</td>
<td>85</td>
<td>25</td>
</tr>
<tr>
<td>8</td>
<td>30</td>
<td>65</td>
<td>166</td>
<td>33</td>
<td>65</td>
<td>163</td>
</tr>
<tr>
<td>9</td>
<td>50</td>
<td>85</td>
<td>-165</td>
<td>41</td>
<td>85</td>
<td>-164</td>
</tr>
<tr>
<td>10</td>
<td>50</td>
<td>85</td>
<td>-165</td>
<td>43</td>
<td>75</td>
<td>-164</td>
</tr>
</tbody>
</table>