Earth and Planetary Science
https://journals.nasspublishing.com/index.php/eps

ARTICLE

How can the Choice of Integration Method Optimize the Earthquake Finite Fault Simulation: A Case Study of 2004 M_w 6.1 Parkfield Earthquake

Ameneh HoushmandViki [1*], ***Anooshiravan Ansari*** [2]

[1]*Geophysics, International Institute of Earthquake Engineering and Seismology, Tehran 19537-14453, Iran*
[2]*Earthquake Engineering, International Institute of Earthquake Engineering and Seismology, Tehran 19537-14453, Iran*

ABSTRACT

Earthquake finite fault simulations confront time-consuming and complex calculations. Therefore, finding methods that result in rapid calculations generalized with sufficient accuracy is predominantly necessary. Different methods of finite fault surface integrals for the 2004 M_w 6.1 Parkfield earthquake have become the subject of the current study to achieve a fast and accurate calculation of earthquake finite fault simulation. Calculations are performed considering fault elements carried out with constant and variable source parameters, while interpolation functions can also be considered. The investigations implemented in this research show that changing the conventional trapezoidal integration method into Gaussian integration on optimal element size could reduce the estimated time of calculations. The surface integral can be done only on one Gaussian point, while the required time for calculations can decrease considerably. To simplify the complex structure of Green's function calculations, a constant Green's function can be assumed in a half-space, with a time shift corresponding to the arrival time of the fault rupture representing the Green's function for other sub-faults.

Keywords: Rapid and Accurate Simulation; Parkfield Earthquake; Surface Integral; Source Parameter; Interpolation

*CORRESPONDING AUTHOR:
Ameneh HoushmandViki, Geophysics, International Institute of Earthquake Engineering and Seismology, Tehran 19537-14453, Iran; Email: a.hooshmandviki@iiees.ac.ir

ARTICLE INFO
Received: 6 January 2025 | Revised: 21 March 2025 | Accepted: 28 March 2025 | Published Online: 3 April 2025
DOI: https://doi.org/10.36956/eps.v4i1.1665

CITATION
HoushmandViki, A., Ansari, A., 2025. How can the Choice of Integration Method Optimize the Earthquake Finite Fault Simulation: A Case Study of 2004 Mw 6.1 Parkfield Earthquake. Earth and Planetary Science. 4(1): 52–65. DOI: https://doi.org/10.36956/eps.v4i1.1665

1. Introduction

Earthquake rupture can be analyzed through two types of models: kinematic and dynamic. Kinematic models focus on the fault's slip without linking it to the stresses that induce it, whereas dynamic models take into account the entire fracture process, connecting fault slip to the stresses in the fault area. In kinematic earthquake rupture, the slip vector is defined as the coordinates on the fault plane and the time. This approach simplifies the determination of the associated elastic displacement field [1].

The representation theorem in the context of seismic waves generated by faults allows us to model and analyze how energy is radiated from a buried fault during slip. It can be summarized in the following key steps:

If the earth is modeled as an elastic solid, the displacement field resulting from a point dislocation can be considered a Green's function for the earthquake faulting problem. The displacement field at all points in the earth due to an arbitrary slip distribution on a fault is expressed as an integral over the fault surface, incorporating the slip distribution with Green's function. The slip distribution introduces linearity into the integral, allowing it to be derived to solve a linear problem where the recorded ground motion at the earth's surface is taken as data [2].

$$u_i(x,t) = \int_{-\infty}^{+\infty} d\tau \iint_{\Sigma} [u_j(\varepsilon,\tau)] c_{jkpq} G_{ip,q}(x,t,\varepsilon,\tau) v_k d\Sigma \qquad (1)$$

The first integral in equation 1 is a convolution in time. The second is over the fault surface Σ. Σ is illustrated as a plane striking along an arbitrary axis with a specific dip. The vector v_k, gives the unit normal to the fault surface. c_{jkpq} is the elastic constant. $G_{ip,q}$ is the derivative of G known as the Green's function. Green's function utilized in the representation theorem represents the medium's response to a point force in the absence of discontinuities. This Green's function can be employed to satisfy any boundary condition on the fault surface; thus, the elastodynamic equations need only be solved once. The representation theorem is tailored to the scenario where the stress field remains continuous across the fault surface, while only the displacement field is permitted to be discontinuous.

The process of simulating earthquake ground motion has significant implications for seismological applications. For example, solving forward problems forms the basis of analyzing inverse problems. Forward problems are typically well-posed, possessing a unique solution that remains stable against small variations in the initial values; in contrast, inverse problems are often ill-posed and non-unique. Non-linear inverse problems represent a particularly challenging subset of inverse problems, where numerous forward problems may be in progress. Ultimately, we need to develop methods to reduce the extensive computation time associated with forward problems. In this research, we explore various simulation integration methods that lead to faster and more accurate solutions.

The practice of numerically simulating time series can be traced back to the contributions of Hartzel [3] and Irika [4]. The simulation methods have been expanded to incorporate stochastic representations of source and path effects [5], theoretical full waveform Green's functions [6], or various combinations of these techniques. Hartzel et al. [7] provided a detailed comparison of several kinematic simulation methods. When employing hybrid methods to simulate earthquake ground motion, it is common to differentiate between high-frequency and low-frequency components, specifically around 1 Hz. The low-frequency component is usually modeled using deterministic methods, which provide a theoretically sound depiction of fault rupture and wave propagation. In contrast, high-frequency components are typically generated using stochastic models due to their inherently random characteristics.

2. Finite Fault Integration

To solve the forward problem of equation 1, we should first discuss various methods of solving a numerical integration. Numerical integrations can be computed using a variety of techniques. The fundamental challenge in numerical integration is to calculate an approximate solution to an integral. The simplest method of this type is to let the approximated function be constant (a polynomial of zeroth order) that passes

through the points known as midpoints. This is called the midpoint rule or rectangle rule.

A typical technique for approximating calculations is trapezoidal integration. This method employs linear approximation (a polynomial of order one) and converts the function into a trapezoidal series before calculating the integral.

Simpson's rule is another method of integration based on a polynomial of order two. In Simpson's Rule, we use parabolas for approximation. This proves to be very efficient since it's generally more accurate than the other numerical methods.

In complex functions that cannot be integrated easily, instead of calculating the integral analytically, the value of the function at certain points is multiplied by weighting factors, and the values at the determination points are summed with each other. The Gaussian integration method is an approximate method that is done by approximating a series of Gaussian points with specific weights.

In finite fault forward problems, the temporal and spatial distribution of source parameters play an important role in ground motion simulations, they can be used to predict ground motions for engineering design purposes and to study the effects of complex earth structures. The source parameters in each sub-fault can have a constant or variable distribution, and as a result, integration on the entire surface of the fault creates the ground motion on the earth's surface. Integration can be done on a zeroth order approximation known as constant distribution while variable distribution relies on greater order approximated functions known as interpolation functions. A constant distribution is insufficient to describe the fault failure process, so an interpolation function could be used to describe the variability of source parameters.

In corresponding research, Custadio et al. [8] used a linear interpolation function with first order to invert the 2004 Parkfield earthquake. Liu and Archuleta [9] introduced a type of linear interpolation function to investigate the slip changes on each sub-fault of the 1989 Loma Prieta earthquake. Babuska and Suri [10] discussed that interpolation accuracy could be increased by increasing the order of the interpolation function while Liu and Archuleta [9] discussed that If a higher-order interpolation is employed, synthetic seismic waveforms from a sub-fault are no longer linearly related to the slip amplitudes of each corresponding node. The higher-order interpolation will also impose a stronger spatial correlation among the source parameters. Liu and Archuleta [9] showed that Numerical methods with higher-order interpolation make the inversion converge significantly more slowly. Bi-cubic spline interpolation requires over five times the CPU time compared to linear interpolation for the same problem.

3. Finite Fault Discretization and Parameterization

To solve forward problems in finite fault solutions, it is necessary to perform a discrete shape of the fault plane. As a result, the fault surface is divided into sub-faults, and each sub-fault is parameterized by a specific amount of slip, rise time, and rupture time along the strike and dip angles. Olson and Apsel [11] provided a method to estimate the surface integral using constant elements, where locations within each element were assumed to experience the same slip over a specified time shift. By assuming constant slip among each sub-fault, the slip function on the fault surface can be expressed using the following equations:

$$s(x,t) = \sum_{j=1}^{J} X_j(x) \sum_{k=-K}^{K} s_{jk} P_k(x,t)$$

$$X_j(x) = \begin{cases} 1 & if\ x\ in\ j^{th}\ subfault \\ 0 & else \end{cases} \tag{2}$$

The first sum is over J cells, the second sum defines the slip within j^{th} sub-fault . the vector s_{jk} is the slip direction at the j^{th} cell at the k^{th} time point and has two components on the plane. $P_k(x,t)$ contains the time dependence of the k^{th} slip. By substituting equation 2 in equation 1 we obtain equation 3:

$$u_i(x,t) = \sum_{j=1}^{J} \sum_{k=-K}^{K} s_{jk} . g_j^i(x, t+\delta t) \tag{3}$$

where $g_j^i(x,t+\delta t)$ is the Green's function for the j^{th} cell

at the position x in the i^{th} Component direction. since s_{jk} and $g_j^i(x,t+\delta t)$ are two-component vectors, the total number of scalar parameters is $2J(2K+1)$. It is possible to express equation 3 in matrix form and solve the forward problems accordingly.

4. Finite Fault Interpolation

Interpolation functions can be developed in one, two, or even three dimensions These functions are obtained by formulating finite element method (FEM) equations. The studied interpolation functions in this research are in two dimensions and are used for finite elements that are continuous with zeroth order. In this sense the only zero-order derivative of the displacement field at the boundaries of each element is continuous. The purpose of FEM is limited to the medium divided into several finite elements following equations of algebraic rules and solving a system of algebraic equations is the approximate solution of a forward problem.

Finite elements could have different shapes. Rectangular elements are convenient for use in modeling a regular geometric space. Liu and Archuleta [9] used quadrilateral elements that do not have to be rectangular or equal in area to model the variability of simulation parameters. If the quadrilateral element is rectangular, then linear interpolation can be applied directly.

The state of the rectangular element is a rectangle with four nodal points. It is assumed that the length and the width of the rectangle be parallel to the Cartesian coordinates and the numbering of the nodal points Conventionally should be counter-clockwise (**Figure 1**).

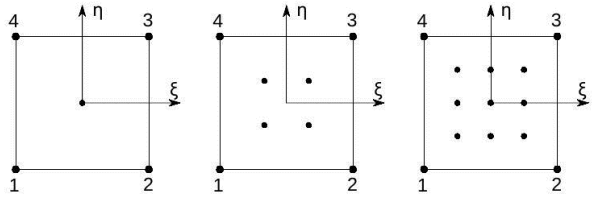

Figure 1. Quadrilateral elements described by 4 nodal points in a two-dimensional environment including 1 integral point (left), 4 integral points (middle), and 9 integral points (right). ξ and η are local coordinates.

In such elements we deal with four points and necessarily confront four degrees of freedom for our algebraic equation and any desired parameter on the element can be described by a polynomial with four terms as below:

$$\phi(\xi,\eta) = a_0 + a_1\xi + a_2\eta + a_3\xi\eta \tag{4}$$

The relation of this function is complex and can be solved by selecting an appropriate coordinate. We consider η and ξ as local coordinates (shown in **Figure 1**) and accordingly achieve a new equation as below:

$$\begin{aligned}\phi(\xi,\eta) = N_1(\xi,\eta)\phi_1 + N_2(\xi,\eta)\phi_2 + \\ N_3(\xi,\eta)\phi_3 + N_4(\xi,\eta)\phi_4\end{aligned} \tag{5}$$

$\phi(\xi,\eta)$ represents the interpolation parameter and $\phi_1,\phi_2,\phi_3,\phi_4$ represent parameters at the nodal points. N_1,N_2,N_3,N_4 are known as shape functions described below:

$$\begin{aligned}N_1(\xi,\eta) &= \frac{1}{4}(1-\xi)(1-\eta)\\ N_2(\xi,\eta) &= \frac{1}{4}(1+\xi)(1-\eta)\\ N_3(\xi,\eta) &= \frac{1}{4}(1+\xi)(1+\eta)\\ N_4(\xi,\eta) &= \frac{1}{4}(1-\xi)(1+\eta)\end{aligned} \tag{6}$$

Assuming linear changes of the parameters inside the rectangular element, the coordinates (x,y) of any arbitrary point inside the element are obtained by the following equations:

$$\begin{aligned}x^e(\xi,\eta) &= \sum_{i=1}^{4} N_i(\xi,\eta)x_i^e\\ y^e(\xi,\eta) &= \sum_{i=1}^{4} N_i(\xi,\eta)y_i^e\end{aligned} \tag{7}$$

e introduces the element index. N_i is the shape function related to nodal points.

Similarly, any simulation parameters within the element can be acquired through linear interpolation of the nodal values of the elements as follows:

$$m^e(\xi,\eta) = \sum_{i=1}^{4} N_i(\xi,\eta)m_i^e \tag{8}$$

Where $m^e(\xi,\eta)$ could include various simulation parameters such as the slip value, rake angle, rupture velocity, and rise time. Within each element or subfault, it is feasible to simulate the waveform radiated from rupture through a series of point dislocations that

cover the fault. **Figure 2** illustrates the discretization of a finite fault surface model. In this context, the source parameters are represented by nodal points (black stars). The Green's functions are computed within each sub-fault at the points indicated by white triangles, and the surface integration of equation 1 is performed at the black dots. $N\xi$ points along the strike and $N\eta$ points along the dip are assigned within each element at equal distances from one another Following this discretization procedure, equation 1 can be reformulated as equation 3, where J represents the total number of integration points (black dots in **Figure 2**). Assuming that all elements or sub-faults have rectangular shapes, the local coordinate (ξ, η)within each sub-fault can be defined as:

$$\xi_j = \frac{2j-1}{N_\xi} - 1 \qquad j = 1 \ldots N_\xi$$
$$\eta_k = \frac{2k-1}{N_\eta} - 1 \qquad k = 1 \ldots N\eta \tag{9}$$

The source parameters may exhibit continuous variation throughout the entire fault, although their values are not smooth at the nodal points. Furthermore, the radiated wave from each sub-fault depends on the source parameters of four nodes within each sub-fault, rather than relying solely on one parameter as is the case with constant elements.

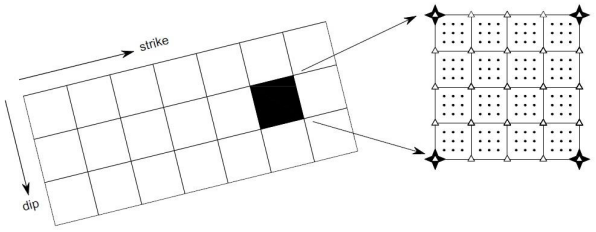

Figure 2. Sample model of discretization of the fault surface.

5. An Overview of the 2004 M_w 6.1 Parkfield Earthquake

The 2004 M_w 6.1 Parkfield earthquake is a well-documented event that fulfills all the necessary criteria for this inquiry. It occurred in the Parkfield area of California's central coast at 17:15:24 (UTC) on September 28, 2004. This earthquake produced a significant amount of strong ground motion data, with over 40 recording stations located within 32 km of the epicenter. The California Geological Survey (CGS) and the U.S. Geological Survey (USGS) captured the strong motions of the event using accelerometers [12].

To analyze the kinematic finite-fault characteristics of the 2004 M_w 6.1 Parkfield earthquake, Custodio et al. [8] Applied a non-linear global inversion to various subsets of near-source ground motion data, resulting in a slip model with amplitudes of less than 0.65 m in two areas along the fault. Ji [13], determined a varying slip distribution through low-frequency waveform inversion and identified two specific regions of interest. Dreger et al. [14] presented a two-dimensional source model using near-field GPS records, which aligned with the distribution of 39 aftershocks that occurred within 48 hours of the initial earthquake. Johanson et al. [15] simultaneously analyzed InSAR and GPS data to study the co-seismic and post-seismic patterns. Additionally, Mendoza and Hartzell [16] calculated the co-seismic slip using both synthetic Green's functions based on a flat-layered 1D velocity model and empirical Green's functions derived from a M_w 5.0 aftershock.

In this survey, the source model of the 2004 M_w 6.1 Parkfield earthquake reported by Custodio [8] is utilized to solve several forward problems (**Figure 3**). The chosen model is a strike-slip fault with a strike of 140 degrees, a dip of 87 degrees, and a length of 40 km.

We utilized strong motion records from 26 CGS stations surrounding the epicenter, which offer comprehensive azimuthal coverage (**Figure 4**). The epicenter is positioned at a latitude of 35.81 degrees and a longitude of -120.37 degrees [17]. For the simulation analysis, we analyzed the velocity waveforms for comparison, filtering all waveforms and Green's functions with a 4th-order Butterworth band-pass filter featuring a linear phase response in the frequency range of 0.16 to 1.00 Hz.

In order to compute the synthetic seismograms for each station, we must calculate the Green's function at any point on the fault surface and then convolve it with the specific source parameter at that point. The Green's functions are computed using the AXITRA code [18] and the frequency-wavenumber method of Bouchan [19]. The 1D velocity structure is employed in the computation of the

Green's functions [20] (**Table 1**).

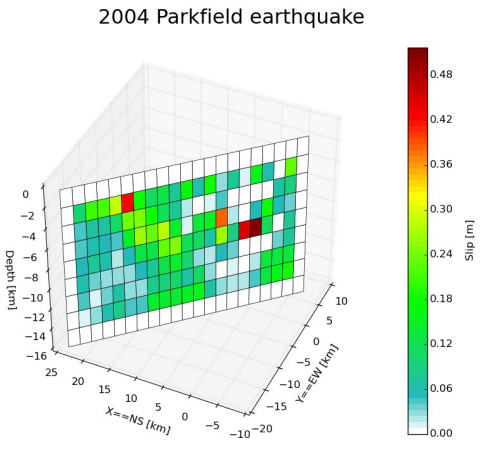

Figure 3. 2004 M_w 6.1 Parkfield slip model.

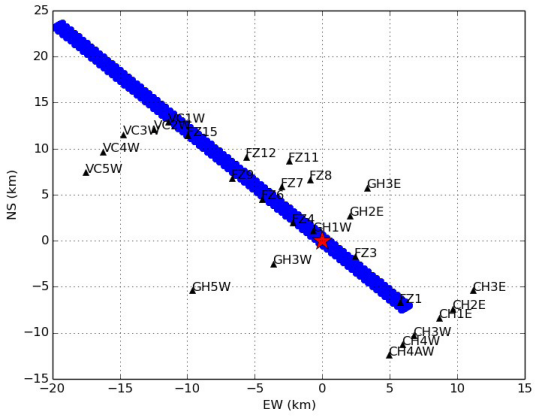

Figure 4. Location of the 2004 M_w 6.1 Parkfield earthquake and 26 stations utilized in this study. The epicenter is indicated by a red star. The blue line represents an overview of the San Andreas fault.

Table 1. P-wave and S-wave velocity structure.

Thickness (m)	V_p (m/s)	V_s (m/s)	Density (kg/m3)	Q_p	Q_s
1000	1900	1000	2000	70	35
1000	3400	1700	2300	270	160
1000	4600	2400	2300	450	260
1000	5100	3100	2700	500	300
1400	5600	3600	2700	550	350
13300	6300	3600	2800	600	350
20000	6800	3600	2800	680	360

6. TF Goodness-of-Fit Criteria for Quantitative Comparison of Time Signals

Comparing two signals with significantly different envelope phases is often crucial. Such comparisons are typically used to assess the validity of new theoretical models or analytical methods. Kristekova et al. [21] introduced an extension of the Time-Frequency (TF) misfit criteria that enables a quantitative and objective comparison of observed and synthetic time signals. The TF representation allows for the analysis of spectral content at any given moment and the time history at any frequency. To assist in comparing the two signals, Kristekova et al. [21] developed a method for analyzing the envelopes of the TF representation and their associated phases. By evaluating the differences in envelope and phase at specific TF points, they proposed a normalized TF misfit criterion that ranges from zero—which indicates the least similarity—to ten, which shows the highest level of agreement between the two signals. **Figure 5** displays the numerical values for the goodness-of-fit (GOF) criteria, along with verbal classifications for the overall signal comparison.

The envelope-phase misfits are primarily useful for comparing relatively close envelopes-phases in the simulation process, particularly when there are too many simulated waveforms to present, the framework proposed by Kristekova et al. [21] is more robust and objective from both mathematical and signal processing perspectives and is employed for making comparisons between observed and simulated time histories in this study.

Goodness-of-Fit	
Numerical value	Verbal value
10	exellent
9	
8	good
7	
6	fair
5	
4	poor
3	
2	
1	
0	

Figure 5. Discrete goodness-of-fit numerical and verbal values.

7. Finite Fault Simulation for the 2004 Parkfield Earthquake

In this section various surface integral calculations containing: constant, linear, and Gaussian approximations are examined to solve the forward problem of the 2004 Parkfield earthquake. The main purpose of this survey is to investigate the accuracy and calculation time derived from each method and reach the most accurate and rapid solution. We also discuss how changing the element size can affect our results. To measure the accuracy of our solution the GOF value presented by Kristekova et al. [21] has become our benchmark.

7.1. Constant and Linear Integration

To calculate the mentioned forward problem, the fault surface discretization and parametrization must be done as discussed in Section 2. According to **Figure 2**, Green's function for each sub-fault is evaluated on the points shown with white triangles, and the source parameters are known on nodal points (black stars). Green's functions are calculated separately and multiplied by the source parameters, and the forward problem is determined using a constant or linear integration on the fault surface. While the constant values of source parameters on nodal points are used in calculation we consider a constant (zeroth order) integration (P = 0). The second approach arises when the source parameters are not constant anymore and to display the variability of source parameters some internal points are considered in each sub-fault (**Figure 2**). If the source parameters pursue a linear function (interpolation function) adopted from internal points then the solution is known as linear integration (P = 1). There can be various configurations of this type depending on the number of integrated points (**Figure 6**).

In our inquiry, the fault surface is discretized into 21 × 9 sub-faults measuring 1.9 × 1.7 (km) in the strike and dip angles, respectively, which is the size that Custodio [8] considered for inversion studies.

In our initial analysis, a constant element is taken into account. We consider 21 × 9 sub-faults measuring 1.9 × 1.7 (km) in the strike and dip angles, along with a coarser arrangement of 7 × 3 sub-faults measuring 5.7 × 5.1 (km) in size in the strike and dip directions.

Simulation No. 1 and 2 in **Table 2** and **Table 3** respectively show the GOF envelope and phase values obtained by constant integration. The results show that a higher GOF is achieved when 21 × 9 sub-faults are presented. Consequently, simulation No. 1 represents a more accurate dimension for the sub-fault rather than simulation No. 2.

In the second approach, different simulation cases with various interior points are considered using linear interpolation functions. Additionally, each sub-fault can be further discretized for a finer Green function calculation. **Figure 6** shows various configurations of Green's functions and interior points. Green's functions are evaluated on points marked with white triangles within each sub-fault, and the surface integration is calculated on black dots.

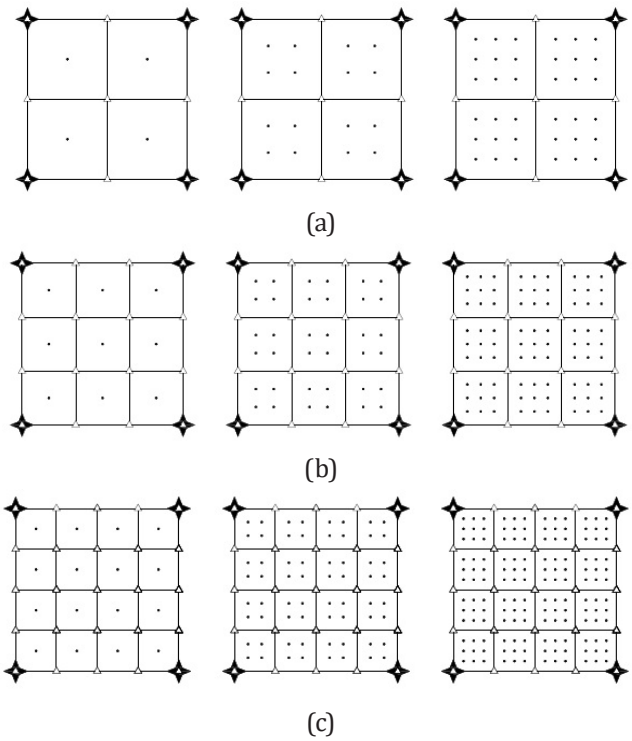

Figure 6. Each sub-fault is divided into elements that represents a finer grid spacing for simulation computation. The elements contain 1, 4, and 9 internal points from left to right, respectively. In sub-figure (**a**) the sub-fault is divided into 4 elements; in sub-figure (**b**), into 9 elements; and in subfigure (**c**), into 16 elements. The white triangles indicate the position for Green's function calculation and are referred as Green's points.

It can be seen from **Table 2** and **Table 3**, that the GOF value increases as the sub-fault size decreases. Looking at the results of linear elements (P = 1), we see a kind of trade-off between the size of the sub-faults and the number of Green's functions in each sub-fault. However, comparing simulation cases 3 and 4 in **Table 2** and **Table 3**, the effect of sub-fault size is more significant. In simulation case 3, the initial sub-fault size is 5.7 × 5.1 (km) along strike and dip respectively, with 16 elements of Green's function and 9 integration points and the distance between Green's points is smaller in Case 3 while Case 4 has a higher GOF value, more sub-faults (21 × 9), and only one element and one integration point. This can be linked to the size of the asperities on the fault. Consequently, the dimensions of sub-faults should be sufficiently small to capture the variability of slip on the fault surface, particularly at the locations of asperities.

Simulation cases 4–15 illustrate the effects of varying the number of Green's and integration points within each sub-fault. It is evident that when the sub-fault size is fixed at 1.9 × 1.7 (km), increasing the number of Green's and integration points does not significantly alter the value of GOF. For instance, by examining simulation cases 4 and 15, it's evident that the number of Green's elements in the latter case 15 is 16 times greater, and the number of integration points is 9 times larger. Meanwhile, the GOF values in the EW direction range from 4.67 to 4.84, reflecting approximately 4% of variability.

In fact, from a computational perspective, this is not a significant case, while the computational cost of the analysis rises considerably as the number of Green's points in each sub-fault increases. It is noteworthy that the most intensive and computation-heavy aspect of finite fault modeling is the calculation of Green's function. The results in **Table 2** and **Table 3** show that increasing the number of Green's points in each sub-fault does not significantly improve the simulation accuracy and, consequently, we should search for an acceptable criterion that defines an optimal computational element size in calculations.

At a given frequency, there should be an appropriate relationship between fault element size, seismic

Table 2. Goodness-of-fit verbal value comparison on EW, NS, and UP components.

	sim. no.	N_s^e	N_d^e	D_s^e (m)	D_D^E (M)	N_s^g	N_d^g	$N\xi$	$N\eta$	GOF^{EW}	GOF^{NS}	GOF^{UP}
P = 0	1	7	3	5700	5100					1.22	1.15	1.43
	2	21	9	1900	1700					3.02	3.22	2.15
P = 1	3	7	3	5700	5100	4	4	3	3	3.97	3.76	3.24
	4	21	9	1900	1700	1	1	1	1	4.70	4.65	4.60
	5	21	9	1900	1700	1	1	2	2	4.67	4.61	4.50
	6	21	9	1900	1700	1	1	3	3	4.68	4.62	4.54
	7	21	9	1900	1700	2	2	1	1	4.80	4.67	4.46
	8	21	9	1900	1700	2	2	2	2	4.78	4.67	4.47
	9	21	9	1900	1700	2	2	3	3	4.78	4.68	4.48
	10	21	9	1900	1700	3	3	1	1	4.85	4.69	4.47
	11	21	9	1900	1700	3	3	2	2	4.83	4.69	4.48
	12	21	9	1900	1700	3	3	3	3	4.84	4.69	4.47
	13	21	9	1900	1700	4	4	1	1	4.83	4.68	4.49
	14	21	9	1900	1700	4	4	2	2	4.83	4.69	4.49
	15	21	9	1900	1700	4	4	3	3	4.84	4.69	4.49
Gauss	16	21	9	1900	1700	1	1	1	1	4.99	4.80	4.56
point	17	21	9	1900	1700	1	1	2	2	5.00	4.76	4.49
	18	21	9	1900	1700	1	1	3	3	4.90	4.71	4.47

N_S^E, N_D^E: the number of sub-faults along the strike and dip of the fault respectively; D_S^E, D_D^E: the size of the sub-faults along the strike and dip of the fault respectively; N_S^G, N_D^G: the number of elements at each sub-fault used for Green's function calculation along the strike and dip of the fault respectively; $N\xi$, $N\eta$: the number of integration points in each element along the strike and dip of the fault respectively; GOF^{EW}, GOF^{NS}, GOF^{UP}: average of goodness-of-fit verbal values among 26 stations along EW, NS, UP components respectively; P: the order of interpolation. P = 0 refers to the constant element and P = 1 corresponds to the linear element.

Table 3. Goodness-of-fit phase value comparison on EW, NS, and, UP components.

	sim. no.	N_s^e	N_d^e	D_s^e (m)	D_D^E (M)	N_s^g	N_d^g	$N\xi$	$N\eta$	GOF^{EW}	GOF^{NS}	GOF^{UP}
P = 0	1	7	3	5700	5100					4.90	4.76	4.48
	2	21	9	1900	1700					4.97	5.10	4.55
	3	7	3	5700	5100	4	4	3	3	4.62	4.97	4.59
	4	21	9	1900	1700	1	1	1	1	5.57	5.64	4.62
	5	21	9	1900	1700	1	1	2	2	5.57	5.60	4.57
	6	21	9	1900	1700	1	1	3	3	5.57	5.62	4.59
P = 1	7	21	9	1900	1700	2	2	1	1	5.85	5.75	4.63
	8	21	9	1900	1700	2	2	2	2	5.83	5.74	4.66
	9	21	9	1900	1700	2	2	3	3	5.84	5.74	4.67
	10	21	9	1900	1700	3	3	1	1	5.87	5.79	4.58
	11	21	9	1900	1700	3	3	2	2	5.85	5.76	4.61
	12	21	9	1900	1700	3	3	3	3	5.86	5.77	4.61
	13	21	9	1900	1700	4	4	1	1	5.86	5.75	4.65
	14	21	9	1900	1700	4	4	2	2	5.86	5.74	4.67
	15	21	9	1900	1700	4	4	3	3	5.86	5.75	4.66
Gauss	16	21	9	1900	1700	1	1	1	1	5.99	5.84	4.67
point	17	21	9	1900	1700	1	1	2	2	6.08	5.96	4.64
	18	21	9	1900	1700	1	1	3	3	6.11	6.01	4.64

NS^E, ND^E : the number of sub-faults along the strike and dip of the fault respectively; DS^E, DD^E: the size of the sub-faults along the strike and dip of the fault respectively; NS^G, ND^G: the number of elements at each sub-fault used for Green's function calculation along the strike and dip of the fault respectively; $N\xi$, $N\eta$: the number of integration points in each element along the strike and dip of the fault respectively; GOF^{EW}, GOF^{NS}, GOF^{UP}: average of goodness-of-fit verbal values among 26 stations along EW, NS, UP components respectively; P: the order of interpolation. P = 0 refers to the constant element and P = 1 corresponds to the linear element.

wave propagation, and data; Therefore, the selection of suitable element dimensions can be influenced by the following factors:

1-The data has been filtered up to 1 Hz, therefore the elements should be able to cover the 1 Hz frequency range as well. There must be a good compatibility between the smallest data wavelength and the smallest dimensions of the grid so that the wavelength of the data must be greater than the dimensions of the grid [11]. The maximum data frequency is 1 Hz, which corresponds to a period of 1 second, so the minimum wavelength of the data can be obtained by the relation: λmin = v/fmax where v is the wave velocity and fmax is the maximum frequency. Considering the different values of P wave velocity, the minimum value of wavelength also changes, as shown in **Table 4**. If we select the minimum velocity (1900 m/s), we achieve the minimum wavelength, which is 1.9 km. This quantity is compared with the dimension of the corresponding length of sub-faults in our study. By choosing the maximum wave velocity (6800), the selected element dimension for width (1.7 km) is proportional to 0.25 of the P wavelength (6800 m). Note that the P wave velocity is higher and includes higher frequencies. As a result, a ratio of the P wavelengths can be considered as a suitable criterion for choosing the optimal dimensions of sub-faults. The selected sub-fault dimension is about 5% of the fault length along the strike and 10% of the fault width along the dip and this factor can also be used as a criterion in sub-fault division.

2-Velocity models represent the wave velocity in different layers and indicate the maximum frequency that can be read and consequently limit the wavelengths that can be simulated. The velocity model considered for the 2004 Parkfield earthquake, covers the frequency range of 1 Hz and higher as well. A more accurate identification of the velocity model allows higher frequency modeling.

Table 4. Minimum wavelength for different velocities.

Thickness (m)	V_p (m/s)	λmin (m)
1000	1900	1900
1000	3400	3400
1000	4600	4600
1000	5100	5100
1400	5600	5600
13300	6300	6300
20000	6800	6800

7.2. Gaussian Integration

In this section, it is assumed that the integration is done on Gaussian points. In simulations No. 16, 17, and 18 of **Table 2** and **Table 3**, values of the GOF envelope and phase obtained by the Gaussian integration method considering 1,4,9 Gaussian points, are presented. The obtained values reveal considerable accuracy.

While the forward problem is a time-consuming process, we need to decrease the calculation time required in integration; Therefore, we have studied a great comparison between the speed of the calculations performed in our simulation process, which covers simulation numbers 4–18. In **Table 5**, a report of time calculations of Green's function, surface integration, and total simulation process is shown.

The processing personal computer is equipped with a 5-core CPU, 1 GB of RAM, and a 500 GB storage capacity. For our calculations, we utilized MATLAB as the programming language. The average computing time provides a general idea of the computational cost for the integration processes and the standard deviation indicates the consistency of the execution times. According to **Table 5**, we computed the average of the total time of calculations and the standard deviation related to them. The average total time used for simulations is approximately 104.05 minutes, and the standard deviation is approximately 63.32 minutes. This result suggests a notable variation in execution times, emphasizing both the potential for optimization and the need for careful consideration of computational consistency.

Results show that choosing Gaussian points has increased the speed of calculations considerably and the higher GOF value presents a better compatibility in modeling. Simulation No. 16 of **Table 5**, whose calculation speed is signed with a red asterisk, shows that the surface integral can be done on only one Gaussian point at a higher speed, and consequently, the necessary time for calculations decreases 10 times more in comparison with simulation number 15, where 9 non-Gaussian points with 16 Green's elements were used. Hence, it's possible to do the integration on only one Gaussian point on the appropriate sub-fault size and reach a more accurate and rapid solution.

7.3. Uncertainty Quantification in Simulation

Understanding the influence of variability in source parameters and Green's functions is key to assessing the robustness of the simulation analysis. Sampling the probability distributions of results based on the uncertainties is an essential step in uncertainty quantification. To evaluate the effect of variability in source parameters and Green's functions, the mean value and standard deviation of GOF values of EW components are calculated.

There are uncertainties in the number and size of sub-faults, which vary from 7 × 3 subfaults with a size of 5.7 × 5.1 (km) to 21 × 9 sub-faults with a size of 1.9 × 1.7 (km). In this case, simulations number 1 and 2 are compared, considering constant values for source parameters and Green's functions. The calculated mean of GOF values for the EW components is approximately 2.12, with a standard deviation of 1.27, which is considerable. Therefore, Selecting the appropriate size for simulation would be a big target.

In another case we consider variables for Green's functions and source parameters, the number of elements related to Green's functions varies from 1 to 16 in each subfault, and the number of integration points related to source parameters varies from 1 to 9 in each element. Considering simulations 4, 5, and 6, there is one element in each sub fault and the integration points vary from 1 to 4 and 9 points respectively, the calculated mean of GOF values for the EW components is achieved at approximately 4.68, with a standard devia-

Table 5. A timetable of simulation calculations from simulation 4–18.

	sim. no.	N_s^e	N_d^e	D_s^e (m)	D_D^E (M)	N_s^g	N_d^g	$N\xi$	$N\eta$	G_time (min)	int_time (min)	A_TIME (MIN)	Ratio
P = 1	4	21	9	1900	1700	1	1	1	1	37.76	00.15	38.00	6.69
	5	21	9	1900	1700	1	1	2	2	37.76	00.55	39.87	6.38
	6	21	9	1900	1700	1	1	3	3	37.76	03.15	41.96	6.06
	7	21	9	1900	1700	2	2	1	1	82.32	01.34	83.66	3.04
	8	21	9	1900	1700	2	2	2	2	82.32	05.00	87.32	2.91
	9	21	9	1900	1700	2	2	3	3	82.32	10.42	93.74	2.71
	10	21	9	1900	1700	3	3	1	1	138.80	03.35	142.15	1.79
	11	21	9	1900	1700	3	3	2	2	138.80	11.43	150.23	1.69
	12	21	9	1900	1700	3	3	3	3	138.80	24.45	163.25	1.55
	13	21	9	1900	1700	4	4	1	1	206.99	07.00	213.99	1.18
	14	21	9	1900	1700	4	4	2	2	206.99	20.00	226.99	1.12
	15	21	9	1900	1700	4	4	3	3	206.99	47.55	254.54	1
Gauss	16	21	9	1900	1700	1	1	1	1	28.14	00.07	28.21	9.02*
point	17	21	9	1900	1700	1	1	2	2	28.14	00.25	28.39	8.96
	18	21	9	1900	1700	1	1	3	3	28.14	01.53	29.67	8.57

N_S^E, N_D^E : the number of sub-faults along the strike and dip of the fault respectively; D_S^E, D_D^E: the size of the sub-faults along the strike and dip of the fault respectively; N_S^G, N_D^G: the number of elements at each sub-fault used for Green's function calculation along the strike and dip of the fault respectively; $N\xi$, $N\eta$: the number of integration points in each element along the strike and dip of the fault respectively; GOF^{EW}, GOF^{NS}, GOF^{UP}: average of goodness-of-fit verbal values among 26 stations along EW, NS, UP components respectively; P: the order of interpolation. P = 0 refers to the constant element and P = 1 corresponds to the linear element; G_time: the time required for Green's function calculation; int_time: the time required for integration; A_time: G_time + int_time; Ratio: the ratio between each simulation and simulation 15.

tion of approximately 0.015. The very small standard deviation suggests a high level of consistency in the GOF across simulations, indicating that the simulation approach is robust under the tested parameters. Since all values are close, it appears that there are no drastic fluctuations. While we consider simulations 4, 7, 10, and 13, the number of elements increases from 1 element to 4, 9, and 16 elements respectively, and the integration points related to source parameters are considered 1. In this case, the calculated mean of GOF values is approximately 4.8, with a standard deviation of approximately 0.067. The standard deviation is higher than when only the number of integration points changed. This result shows that the effect of increasing the number of Green's functions has more influence than increasing the number of integration points related to source parameters. In the next step simulation number 4 with 1 element and 1 integration point and simulation number 16 with 1 element and 1 Gaussian integration point are considered. In this case, the mean of GOF values is achieved at approximately 4.84 with a standard deviation of approximately 0.2 which points out how considering a Gaussian point could affect the simulation result.

7.4. Quantitative Study of Green's Function in the Simulation

While the calculation of Green's function is a very sensitive and time-consuming part of fault modeling, we are having a short inquiry in this area.

To gain a clearer understanding of how the variability of Green's functions affects the simulation results, this study compares the velocity waveforms of the moment tensor elements of Green's functions for two-point sources. As shown in **Figure 7**, it is assumed that A and B are point sources situated within a specific sub-fault area in a half-space.

As shown in **Figure 8**, the moment tensor elements of Green's function are quite similar to one another, indicating the validity of the assumption of a constant Green's function, with only time shifts among each sub-fault. The waveforms in **Figure 8** are adjusted

to better highlight the similarity of the signals.

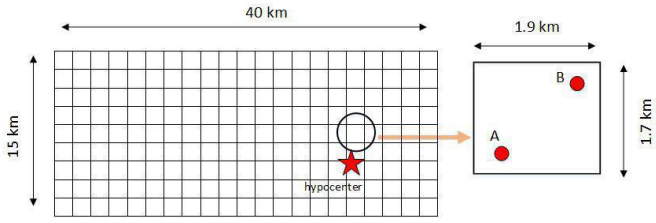

Figure 7. Two specific point sources have been identified to compare Green's function components. The red star shows the location of the epicenter.

According to the distribution of Green's points and interpolation points within each sub-fault, if the size of the sub-fault is sufficiently small and falls within the limits of the asperity size and the resolution of the velocity model, it can be assumed that a constant Green's function exists for every sub-fault located in a layer, with only a time shift dependent on the arrival time of the failure. Consequently, Green's functions at various points of a sub-fault can be calculated rapidly in this manner.

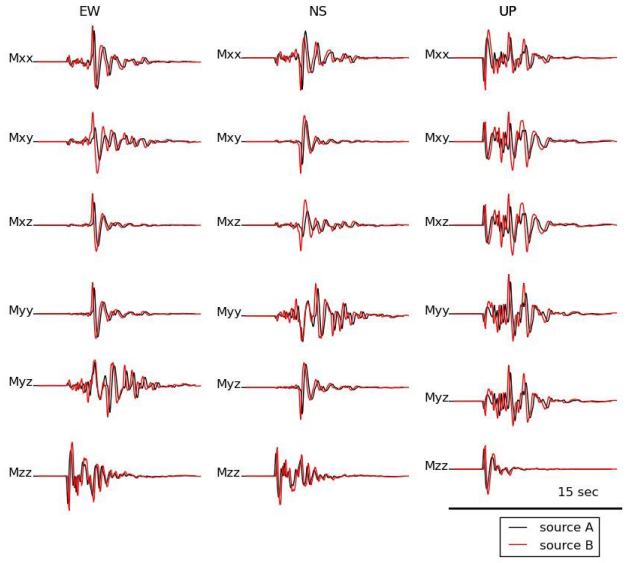

Figure 8. Velocity waveform of Green's function moment tensor elements for the CH1E station. The left-handed Figures correspond to the EW component, the middle figures belong to the NS component, and the right figures pertain to the UP component.

8. Discussion and Conclusions

Parameterization and discretization of finite fault surfaces significantly impact the results of finite fault simulations. Various simulation analyses indicate that employing an appropriate mesh size and enhancing the order of interpolation have substantial effects on finite fault simulations.

While the uncertainty quantification of results relies on the importance of choosing a suitable sub-fault size, a multiple of P wavelength can serve as an effective criterion for the sub-fault dimension. The findings of this study demonstrate that selecting 5% of the fault length in the strike direction and approximately 10% of the fault width in the dip direction can be an effective criterion for determining an appropriate sub-fault dimension. The results indicate that Gaussian integration on the finite fault surface yields a more accurate and faster calculation in earthquake simulations. The surface integral can be performed at just one Gaussian point with increased speed, thereby significantly reducing the time required for calculations. Additionally, the findings suggest that if the size of the sub-fault is sufficiently small relative to the size of asperities and the resolution of the medium's velocity model, it is a valid assumption to treat the Green's function as constant in a half-space and compute the remaining Green's functions solely by applying a time shift corresponding to the rupture time. This would represent a significant advancement, considering the considerable computational demands of calculating Green's functions in finite fault inversions and forward modeling.

To adapt methods from the Parkfield earthquake to other seismic events, it's crucial to understand the target's specific characteristics, such as magnitude, geometry, and geology. Larger earthquakes require a coarser mesh in numerical models due to their broader seismic wave propagation, while smaller earthquakes can use larger element sizes without losing significant detail. However, it is essential to balance the element size with the resolution needed to capture specific features of the simulation, such as fault behavior and local site effects. In regions where these features are critical to the simulation, smaller elements may still be necessary to capture the details accurately. The performance of Gaussian integration might be less sensitive to element size changes in straightforward geometrical scenarios. For more homogeneous geological settings, larger element sizes may suffice, leading to a more

straightforward application of Gaussian integration with potentially greater computational efficiency.

Author Contributions

The conceptualization of this research was done by A.H. The methodology and software development were carried out through the contributions of A.H. and A.A. Validation was performed by A.A. The analysis and investigation were conducted by A.H. Writing—original draft preparation and review/editing—were done by A.H. Visualization, supervision, and project administration were contributions from both authors. All authors have read and agreed to the published version of the manuscript.

Funding

This research received no external funding.

Institutional Review Board Statement

Ethical approval was not required for this study as it did not involve human participants or animals.

Informed Consent Statement

Informed consent was obtained from all subjects involved in the study.

Data Availability Statement

Data supporting the calculations of this article can be found at: EQuake-RC [Online], Available from: http://equake-rc.info/srcmod; CESMD [Online], Available from: https://www.strongmotioncenter.org

Acknowledgments

We sincerely acknowledge the Strong Motion Center for providing easy access to the data, and the SIV-Project investigation of Equake-RC for their valuable support.

Conflicts of Interest

The authors disclosed no conflict of interest.

References

[1] Udias, A., 1999. Principle of Seismology. Cambridge University Press: Cambridge, UK. pp. 29–40.

[2] Aki, K., Richards, P.G., 2002. Quantitative seismology, 2nd Ed. University Science Books: Sausalito, CA, USA. pp. 37–41.

[3] Hartzell, S., 1978. Earthquake aftershocks as greens functions. Geophysical Research Letters. 5(1), 1–4.

[4] Irikura, K., 1978. Semi-empirical estimation of strong ground motions during large earthquakes. Bulletin of the Disaster Prevention Research Institute. 33, 63–104.

[5] Boore, D., 1983. Stochastic simulation of high-frequency ground motions based on seismological models of the radiated spectra. Bulletin of the Seismological Society of America. 73(6A), 1865–1894.

[6] Zeng, Y., Anderson, J.G., Yu, G., 1994. A composite source model for computing realistic synthetic strong ground motions. Geophysical Research Letters. 21(8), 725–728.

[7] Hartzell, S., Harmsen, S., Frankel, S.L., 1999. Calculation of broadband time histories of ground motion: Comparison of methods and validation using strong-ground motion from the 1994 Northridge earthquake. Bulletin of the Seismological Society of America. 89(6), 1484–1504. DOI: https://doi.org/10.1785/BSSA0890061484

[8] Custodio, S., Liu, P., Archuleta, R., 2005. The 2004 M_w 6.0 Parkfield, California, earthquake: Inversion of near-source ground motion using multiple data sets. Geophysical Research Letters. 32(23), L23312. DOI: https://doi.org/10.1029/2005GL024417

[9] Liu, P., Archuleta, R.J., 2004. A new nonlinear finite fault inversion with three-dimensional Green's functions: application to the 1989 Loma Prieta, California, earthquake. Journal of Geophysical Research. 109(B2), B02318. DOI: https://doi.org/10.1029/2003JB002625

[10] Babuska, I., Suri, M., 1994. The p and h-p version of the finite element method, basic principles, and properties. Society for industrial and applied mathematics. SIAM Review. 36(4), 578–632. DOI: https://doi.org/10.1137/1036141

[11] Olson, A., Apsel, R., 1982. Finite faults and inverse theory with applications to the 1979 Imperial Valley earthquake. Bulletin of the Seismological Society of America. 72(6), 1969–2001.

[12] Custodio, S., 2007. Earthquake rupture and ground-motions: The 2004 Mw 6 Parkfield earthquake [PhD thesis]. Santa Barbara, CA: University of California. pp. 11–13.

[13] Ji, C., 2004. Slip history the 2004 (M_w 5.9) Park-

field earthquake (Single-Plane Model). Available from: https://www.tectonics.caltech.edu/slip_history/2004_ca/parkfield2.html (Accessed 10 January 2020).

[14] Dreger, D.S., Gee, L., Lomband, P., et al., 2005. Strong ground motions: Application to the 2003 M_w 6.5 San Simeon and 2004 M_w 6 Parkfield earthquakes. Seismological Research Letters. 76.

[15] Johanson, I.A., Fielding, E.J., Rolandone, F., et al., 2006. Coseismic and Postseismic slip of the 2004 Parkfield earthquake from space-geodetic data. Bulletin of the Seismological Society of America. 96(4B), S269–S282. DOI: https://doi.org/10.1785/0120050818

[16] Mendoza, C., Hartzell, S., 2008. Finite-Fault Analysis of the 2004 Parkfield, California Earthquake Using Pnl Waveforms. Bulletin of the Seismological Society of America. 98(6), 2746–2755. DOI: https://doi.org/10.1785/0120080111

[17] Thurber, C., Roecker, S., Roberts, K., et al., 2003. Earthquake locations and three-dimensional fault zone structure along the creeping station of the San Andreas fault near Parkfield, CA: preparing for SAFOD. Geophysical Research Letters. 30(3), 1112.

[18] Cotton, F., Coutant O., 1997. Dynamic stress variations due to shear faults in a plane-layered medium. Geophysical Journal International. 128(3), 676–688. DOI: https://doi.org/10.1111/j.1365-246X.1997.tb05328.x

[19] Bouchon, M., 1981. A simple method to calculate Green's functions for elastic layered media. Bulletin of the Seismological Society of America. 71(4), 959–971. DOI: https://doi.org/10.1785/BSSA0710040959

[20] Thurber, C., Zhang, H., Waldhauser, F., et al., 2006. Three-dimensional compressional wave speed model, earthquake relocations, and focal mechanisms for the Parkfield, California, region. Bulletin of the Seismological Society of America. 96(4B), S38–S49. DOI: https://doi.org/10.1785/0120050825

[21] Kristekova, M., Kristek, J. Moczo, P., 2009. Time-frequency misfit and goodness-of-fit criteria for quantitative comparison of signals. Geophysical Journal International. 178(2), 813–825. DOI: https://doi.org/10.1111/j.1365-246X.2009.04177.x