# Misfit Criteria for Quantitative Comparison of Seismograms

by Miriam Kristeková, Jozef Kristek, Peter Moczo, and Steven M. Day

Abstract   We have developed and numerically tested quantitative misfit criteria for comparison of seismograms. The misfit criteria are based on the time-frequency representation of the seismograms obtained as the continuous wavelet transform with the analyzing Morlet wavelet. The misfit criteria include time-frequency envelope and phase misfits, time-dependent envelope and phase misfits, frequency-dependent envelope and phase misfits, and single-valued envelope and phase misfits.

We tested properties of the misfit criteria using canonical signals. The canonical signals, taken as the reference signals, were specifically amplitude, phase shift, time shift, and frequency modified to demonstrate the ability of the misfit criteria to properly quantify the misfits and recognize the character and cause of the misfits between the reference and modified signals. In all cases the misfit criteria properly quantified and characterized the misfits.

The misfit criteria were also calculated for four different numerical solutions for a single layer over half-space (the SCEC LOH.3 problem) and the reference FK solution. The misfit criteria provided useful insight into the misfits between individual numerical solutions and the reference solution.

The standard *RMS* misfit matches the single-valued envelope misfit only in the case of a pure amplitude modification of the signal. In all other cases *RMS* considerably overestimates the misfits and does not characterize them.

## Introduction

It often is useful or necessary to compare seismograms—for example, the seismogram calculated by a tested method against a reference solution (which is exact or independent) or a synthetic against a real record. In many articles two seismograms are simply displayed together. Although the simple visual comparison of two seismograms can be useful in some cases, it is obvious that it cannot provide proper quantification and characterization of the difference between the seismograms.

In some articles (e.g., Aoi and Fujiwara, 1999) the misfit of two seismograms is shown by using a difference seismogram defined as

$$D(t) = s(t) - s_{REF}(t), \quad (1)$$

where $s(t)$ is the tested seismogram, $s_{REF}(t)$ is the reference seismogram, and $t$ is time. Though $D(t)$ shows a time-dependent difference between two seismograms, it is clear that it can provide very misleading information. The simplest example is a pure time shift of two identical signals—$D(t)$ can be very large without any indication of the reason for, and character of, the difference.

Sometimes it is necessary to investigate and show dependence of the misfit between two solutions on some important parameter(s) as, for example, epicentral distance,

Poisson's ratio, grid spacing, timestep. In such cases it is reasonable to characterize the misfit by a proper single-valued integral quantity. A simple integral criterion, say, misfit *MD*, corresponding to the difference seismogram $D(t)$ may be defined as

$$MD = \frac{\sum_t |s(t) - s_{REF}(t)|}{\sum_t |s_{REF}(t)|}. \quad (2)$$

A more commonly used misfit criterion (e.g., Geller and Takeuchi, 1995) is the *RMS* (root mean square) misfit defined as

$$RMS = \sqrt{\frac{\sum_t |s(t) - s_{REF}(t)|^2}{\sum_t |s_{REF}(t)|^2}}. \quad (3)$$

It is clear from the three preceding definitions that $D(t)$, *MD*, and *RMS* quantify a difference between two seismograms without having the property of recognizing what causes the difference. In other words, they are unable to properly char-

acterize it. Still the question is whether they can really properly quantify it. Kristek *et al.* (2002) used single-valued (integral) envelope and phase misfits based on the envelope and phase of the analytical signal. The criteria are applicable if the signals to be compared are sufficiently simple.

Considering some time signal as a reference, it is clear that some modifications of the signal can be more visible and understandable in the time domain, some in the frequency domain. Some modifications can change only/mainly amplitudes or envelope, some others can change only/mainly phase. At the same time, the most complete and informative characterization of a signal can be obtained by its decomposition in the time-frequency plane, that is, by its time-frequency representation (TFR). The TFR enables us to see the time evolution of the spectral content. Therefore, it seems quite natural to define misfit criteria based on the TFR, that is, time-frequency dependent criteria. From the time-frequency signal or misfit representation it is then easy to obtain time- or frequency-dependent quantities by projecting the TFR onto either of two domains. It is also possible to naturally define single-valued quantities based on the TFR.

The importance of having reasonable misfit criteria has recently been underlined by the SCEC (Southern California Earthquake Center) and SPICE (Seismic wave Propagation and Imaging in Complex media: a European network) code validation projects (e.g., Day *et al.*, 2003; Moczo *et al.*, 2005; Igel *et al.*, 2005). In particular, the goal of the SPICE Code Validation is to create a long-term, interactive web-based platform for detailed comparison and testing methods and computer codes for the numerical modeling of seismic-wave propagation and earthquake motion.

## Time-Frequency Misfit Criteria

The continuous wavelet transform (CWT) of signal $s(t)$ is defined by

$$CWT_{(a,b)} \{s(t)\} = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} s(t) \, \psi^* \left(\frac{t-b}{a}\right) dt, \quad (4)$$

where $t$ is time, $a$ is the scale parameter, $b$ is the translational parameter, and $\psi$ is the analyzing wavelet. The scale parameter $a$ is inversely proportional to frequency $f$. Consider an analyzing wavelet with a spectrum that has zero amplitudes at negative frequencies. Such a wavelet is an analytical signal and is called the progressive wavelet. The Morlet wavelet,

$$\psi(t) = \pi^{-1/4} \exp(i\omega_0 t) \exp(-t^2/2), \quad (5)$$

with the parameter $\omega_0 = 6$ is an example. With a relation $f = \omega_0/2\pi a$ between the scale parameter $a$ and frequency $f$, the TFR of signal $s(t)$ can be defined as

$$W(t, f) = CWT_{(a,b)} \{s(t)\}; \quad a = \omega_0/2\pi f, \quad b = t. \quad (6)$$

Let $W_{REF}(t,f)$ be the TFR of the reference signal $s_{REF}(t)$, $W(t, f)$ be the TFR of signal $s(t)$, and $N_T$ and $N_F$ be the numbers of time and frequency samples in the time-frequency (TF) plane, respectively. (For the continuous wavelet transform and Morlet wavelet see, e.g., Daubechies, 1992; Holschneider, 1995.)

We define a local TF envelope difference,

$$\Delta E \, (t, f) = |W \, (t, f)| - |W_{REF}(t, f)|, \quad (7)$$

and a local TF phase difference,

$$\Delta P \, (t, f)$$
$$= |W_{REF}(t, f)| \, \frac{\{Arg[W \, (t, f)] - Arg \, [W_{REF}(t, f)]\}}{\pi}. \quad (8)$$

Having the local TF envelope and phase differences, we can define envelope and phase misfits dependent on both time and frequency: time-frequency envelope misfit (*TFEM*)

$$TFEM \, (t, f) = \frac{\Delta E \, (t, f)}{\max_{t,f} \, (|W_{REF}(t, f)|)}, \quad (9)$$

time-frequency phase misfit (*TFPM*)

$$TFPM \, (t, f) = \frac{\Delta P \, (t, f)}{\max_{t,f} \, (|W_{REF}(t, f)|)}, \quad (10)$$

*TFEM* $(t, f)$ characterizes the difference between the envelopes of the two signals as a function of time and frequency. Analogously, the *TFPM* $(t, f)$ characterizes the difference between the phases of the two signals as a function of time and frequency. Both differences are normalized with respect to the maximum absolute TFR value of the reference signal.

Assume, for example, a multiplication of the entire signal by 1.05, that is, the same 5% relative modification of the signal amplitude at each time of the signal. It follows from definition (9) that *TFEM* $(t, f)$ will not be constant along the time axis because the same relative change does not mean the same value of the envelope difference (7). The latter also depends on the absolute value of the local amplitude.

Assume, for example, change of the signal's phase by 5%. Though the change applies to the entire signal, *TFPM* $(t,f)$ will not be constant along the time axis. This is because the definition of the phase difference (8) also includes $|W_{REF}(t, f)|$, that is, the local absolute TFR value of the reference signal.

These two features are clearly demonstrated in the numerical examples.

In some cases it may be useful to see the misfit between two signals as a function of (only) time. Such a misfit can be naturally defined as a projection of the TF misfit onto the

time domain. Thus, the time-dependent misfits are defined as follows: time-dependent envelope misfit (*TEM*)

$$TEM\ (t)\ =\ \frac{\langle \Delta E\ (t, f)\rangle_f}{\max_t\ (\langle |\ W_{\mathrm{REF}}(t, f)|\rangle_f)}, \qquad (11)$$

time-dependent phase misfit (*TPM*)

$$TPM\ (t)\ =\ \frac{\langle \Delta P\ (t, f)\rangle_f}{\max_t\ (\langle |\ W_{\mathrm{REF}}(t, f)|\rangle_f)}, \qquad (12)$$

where

$$\langle \Theta(t, f)\rangle_f\ =\ \frac{1}{N_F} \sum_f \Theta(t, f), \qquad (13)$$

with $\Theta$ representing either of $\Delta E$, $\Delta P$, $|W_{\mathrm{REF}}|$.

Analogously, in some cases it may be useful to see the misfit between two signals as a function of (only) frequency. Such a misfit can be naturally defined as a projection of the TF misfit onto the frequency domain. The frequency-dependent misfits are defined as follows: frequency-dependent envelope misfit (*FEM*)

$$FEM\ (f)\ =\ \frac{\langle \Delta E\ (t, f)\rangle_t}{\max_f\ (\langle |\ W_{\mathrm{REF}}(t, f)|\rangle_t)}, \qquad (14)$$

frequency-dependent phase misfit (*FPM*)

$$FPM\ (f)\ =\ \frac{\langle \Delta P\ (t, f)\rangle_t}{\max_f\ (\langle |\ W_{\mathrm{REF}}(t, f)|\rangle_t)}, \qquad (15)$$

where

$$\langle \Theta(t, f)\rangle_t\ =\ \frac{1}{N_T} \sum_t \Theta(t, f), \qquad (16)$$

with $\Theta$ representing either of $\Delta E$, $\Delta P$, $|W_{\mathrm{REF}}|$.

In addition to the preceding time-frequency, time, and frequency misfit criteria, it is also reasonable to have a single-valued measure of the envelope or phase misfit between two compared signals. The single-valued (say, integral) envelope misfit can be defined as

$$EM\ =\ \sqrt{\frac{\sum_f \sum_t |\Delta E(t, f)|^2}{\sum_f \sum_t |W_{\mathrm{REF}}(t, f)|^2}}. \qquad (17)$$

Similarly, single-valued phase misfit will be

$$PM\ =\ \sqrt{\frac{\sum_f \sum_t |\Delta P(t, f)|^2}{\sum_f \sum_t |W_{\mathrm{REF}}(t, f)|^2}}. \qquad (18)$$

As already pointed out, *TFEM* $(t, f)$ and *TFPM* $(t, f)$ are normalized with respect to the maximum absolute TFR value of the reference signal. Consequently, the time-dependent and frequency-dependent criteria also are scaled with respect to the corresponding maxima. In principle, we could define all criteria normalized with respect to the local $(t, f)$, $(t)$, or $(f)$ values. Some reasons for the adopted normalization are: (1) If one part of the seismogram reaches some values and the rest of seismograms one or more order smaller values, then, in most applications, it is much more important to see the meaningful misfits instead of the entire detailed anatomy of difference between two signals (e.g., the same local level of misfit due to some uniform amplitude or phase modification no matter how large the local values of the signals are). (2) In the comparison of three-component seismograms with two nonzero components and one component with one or more orders smaller values, the global normalization defined with respect to maximum from all three components is the only reasonable one for comparison of the two seismograms. (3) Except of the simplest signals and modifications, the locally normalized criteria are almost "unreadable" or difficult to interpret. (4) For the smallest values of $|W|$, the local normalization leads to numerical division of two very small values and this may cause numerical problems.

In some specific problems, one can be interested in seeing detailed structure of misfits without "recognizing" significant and insignificant amplitudes of seismograms. In such a case the locally normalized criteria can be used—though with possible problems labeled 3 and 4 previously. The locally normalized criteria are listed in the Appendix.

Because some authors use the *RMS* misfit defined by equation (3) to quantify the difference between two seismograms, we will also calculate this misfit in the numerical tests.

Note that the wavelet transform was used for comparison of seismograms by Ji *et al.* (1999) in their finite-fault inversion. Ji *et al.* (1999) used a single-valued criterion based on differences of wavelet coefficients calculated for the observed and synthetic seismograms. In other words, they compared seismograms in the wavelet domain.

## Testing Signals

We consider three signals for testing the previously defined misfit criteria. Signal *S*1 is defined by

$$S1\ =\ A_1(t\ -\ t_1)\exp[-2(t\ -\ t_1)] \cdot$$
$$\cos[2\pi f_1(t\ -\ t_1)\ +\ \varphi_1\pi] \cdot H(t\ -\ t_1). \qquad (19)$$

Signal *S*2 is defined by

$$S2 = A_2 \exp[-2(t - t_2)^2] \cdot$$
$$\cos[2\pi f_2(t - t_2) + \varphi_2\pi]. \quad (20)$$

Here, $H(t)$ means the Heaviside step function. A superposition of the two signals, $S1 + S2$, defines the third signal. In the numerical calculations, the following values of parameters were used: $A_1 = 4$, $t_1 = 2$ sec, $f_1 = 2$ Hz, $\varphi_1 = 0$, $A_2 = 1$, $t_2 = 3.5$ sec, $f_2 = 3$ Hz, $\varphi_2 = 0$.

$S1$ is a harmonic carrier with a sudden onset and decaying amplitude. Its amplitude spectrum has a peak at 2 Hz. $S2$ is Gabor signal, that is, a harmonic carrier with a Gaussian envelope. Its amplitude spectrum, with a peak at 3 Hz, is relatively narrow compared with that of signal $S1$. The relatively broader spectrum of signal $S1$ is due to the sudden amplitude onset of the beginning of the signal. The three signals, $S1$, $S2$, and $S1 + S2$, as well as their spectra are shown in Figure 1. The TFRs of the signals are shown in Figure 2.

## Amplitude and Phase Modifications of the Signals

To test the previously defined misfit criteria, we define canonical modifications of the reference testing signals $S1$, $S2$, and $S1 + S2$.
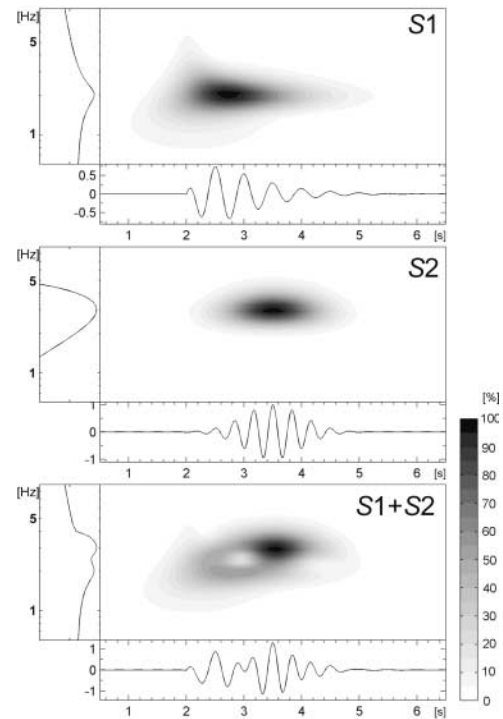


Figure 2. TFR of the reference signals $S1$, $S2$, $S1 + S2$ obtained as the modulus of the CWT of the signals.
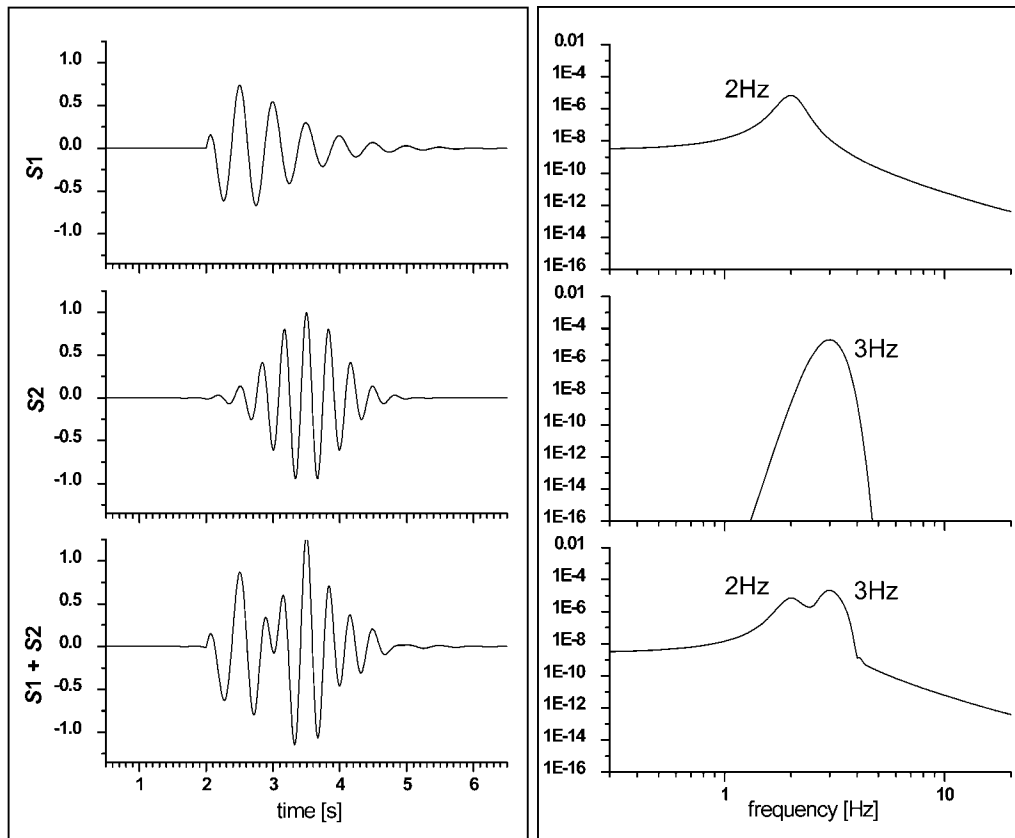


Figure 1. The reference signals $S1$, $S2$, and $S1 + S2$ and their Fourier power spectra.

### Amplitude Modification

Let $s(t)$ be the signal. The modified signal $am05(s(t))$ is defined as

$$am05(s(t)) = 1.05 \cdot s(t). \tag{21}$$

Similarly,

$$am10(s(t)) = 1.10 \cdot s(t), \quad am20(s(t)) = 1.20 \cdot s(t). \tag{22}$$

These definitions mean that, for example, $am05(S1)$ is the modified signal obtained by the 5% amplification of the entire signal $S1$.

### Phase-Shift Modification

Let $s(t)$ be the signal. Its analytical signal can be expressed as $\hat{s}(t) = A(t) \exp[i\varphi(t)]$, with $A(t)$ being the amplitude and $\varphi(t)$ being the phase of the analytical signal. Then the modified signal $pm05(s(t))$ is defined by

$$pm05(s(t)) = \text{Re}[A(t)\exp(i\varphi(t) + 0.05 \, i \, \pi)]. \tag{23}$$

Similarly,

$$
\begin{aligned}
pm10(s(t)) &= \text{Re}[A(t)\exp(i\varphi(t) + 0.10 \, i \, \pi)], \\
pm20(s(t)) &= \text{Re}[A(t)\exp(i\varphi(t) + 0.20 \, i \, \pi)].
\end{aligned}
\tag{24}
$$

These definitions of the phase-shift modification mean that, for example, $pm05(S1)$ is the modified signal obtained by increasing the phase by 5% of $\pi$ in signal $S1$.

Then, for example, $TFEM\text{-}am05(S1 + S2)$ means the TFEM between the reference $S1 + S2$ signal and the modified signal $am05(S1 + S2)$. An other example is $FPM\text{-}am05(S1) + S2$ that means the FPM between the reference $S1 + S2$ signal and the modified signal $am05(S1) + S2$, in which only the $S1$ component of the composed signal is amplitude modified. Here we omitted the dependence of the signals on time.

## Misfits for the Amplitude-Modified Signals

We considered 5, 10, and 20% amplitude modifications of signals $S1$, $S2$, and $S1 + S2$. We also considered 5, 10, and 20% amplitude modifications of the $S1$ component in the composed signal $S1 + S2$. Table 1 lists all the reference and modified signals for which we calculated the misfit criteria (9)–(15) and (17)–(18). The misfits between reference signal $S1$ and the amplitude-modified signals $am05(S1)$, $am10(S1)$, and $am20(S1)$ are shown in Figure 3a. We can see that the distribution of nonzero values of $TFEM(t, f)$ in the $(t, f)$ plane corresponds to the distribution of nonzero values of TFR of the reference signal. In other words, the shape of the area with nonzero values of $TFEM(t, f)$ in the $(t, f)$ plane corresponds to the shape of the area with nonzero

values of TFR of the reference signal. Moreover, the maximum value of the $TFEM(t, f)$ exactly equals the percentage of the amplitude modification for each of the three considered levels (5, 10, and 20%). Its position in the $(t, f)$ plane exactly corresponds to the position of the maximum value of the TFR; this is correct because this is the position at which the absolute amplitude difference between the reference and modified signal has the largest value. The statements on the maximum value and its position are also true for the time-dependent misfit $TEM(t)$ and frequency-dependent misfit $FEM(f)$. The single-valued envelope misfit $EM$ also exactly equals the percentage of the amplitude modification for each of the three considered levels.

At the same time, the phase misfits $TFPM(t, f)$, $TPM(t)$, $FPM(f)$, and $PM$ are all zero: the phase misfits correctly reflect the fact that there is no phase modification of the reference signal.

The $RMS$ misfit exactly equals the single-valued envelope misfit $EM$ in all three levels of the amplitude modification.

What we have said about the misfits for $am05(S1)$, $am10(S1)$, and $am20(S1)$ is also true in the misfits for $am05(S2)$, $am10(S2)$, and $am20(S2)$ (not shown here), and misfits for $am05(S1 + S2)$, $am10(S1 + S2)$, and $am20(S1 + S2)$. The misfits for the latter case are shown in Figure 3b.

Figure 3c shows the envelope and phase misfits between the reference signal $S1 + S2$ and modified signals $am05(S1) + S2$, $am10(S1) + S2$, and $am20(S1) + S2$. Note that only the $S1$ component of the composed reference signal $S1 + S2$ is amplitude modified. We can see in Figure 3c (and by comparison with Fig. 2) that the shape of the area with nonzero values of $TFEM(t, f)$ in the $(t, f)$ plane corresponds to the shape of the area with nonzero values of TFR of the $S1$ component, which is correct. The maximum values of the envelope misfits are proportional to the percentage of the amplitude modification. The maximum values cannot be equal to the percentage of the amplitude modification because the $S1$ component contributes to the TFR less than the $S2$ component (see Fig. 2). Also note the alternating negative and positive values of $TFEM(t, f)$ along the contact of the modified $S1$ and nonmodified $S2$, as well as nonzero values of $TFPM(t, f)$ along the contact. These are because the amplitude modification of only the $S1$ component changed both the envelope and phase of the composed signal. The sign of the envelope and phase differences alternates along the time and frequency axes.

As in all the previous cases, the $RMS$ misfit exactly equals the single-valued envelope misfit $EM$ in all three levels of the amplitude modification.

## Misfits for the Phase-Shift-Modified Signals

According to definitions (23) and (24) we considered 5, 10, and 20% phase-shift modifications of signals $S1$, $S2$, and $S1 + S2$. We also considered 5, 10, and 20% phase-shift modifications of the $S1$ component in the composed signal

Table 1

Reference and Amplitude-Modified Signals Used for Calculation
of the Misfit Criteria

| Reference Signals | Amplitude-Modified Signals | | |
|---|---|---|---|
| $S1$ | $am05(S1)$ | $am10(S1)$ | $am20(S1)$ |
| $S2$ | $am05(S2)$ | $am10(S2)$ | $am20(S2)$ |
| $S1 + S2$ | $am05(S1 + S2)$ | $am10(S1 + S2)$ | $am20(S1 + S2)$ |
| $S1 + S2$ | $am05(S1) + S2$ | $am10(S1) + S2$ | $am20(S1) + S2$ |

$S1 + S2$. Table 2 lists all the reference and modified signals for which we calculated the misfit criteria (9)–(15) and (17)–(18). The misfits between reference signal $S1$ and the phase-shift-modified signals $pm05(S1)$, $pm10(S1)$, and $pm20(S1)$ are shown in Figure 4a. The shape of the area with nonzero values of $TFPM(t, f)$ in the $(t, f)$ plane corresponds to the shape of the area with nonzero values of TFR of the reference signal. The maximum value of the $TFPM(t, f)$ exactly equals the percentage of the phase-shift modification for each of the three considered levels (5, 10, and 20%). The position of the maximum value in the $(t, f)$ plane exactly corresponds to the position of the maximum value of the TFR. The statements on the maximum value and its position are also true for the time-dependent misfit $TPM(t)$ and frequency-dependent misfit $FPM(f)$. The single-valued phase misfit $PM$ also exactly equals the percentage of the phase-shift modification for each of the three considered levels.

At the same time, the envelope misfits $TFEM(t, f)$, $TEM(t)$, $FEM(f)$, and $EM$ are all zero: the envelope misfits correctly reflect the fact that there is no amplitude modification of the reference signal. Note that this is the symmetric situation with respect to the phase misfits for the amplitude-modified signals in the previous section.

The $RMS$ misfit is approximately three times larger than the single-valued phase misfit $PM$ in all three levels of the phase-shift modification. This means that $RMS$ approximately three times overestimates the level of the phase-shift modification. This obviously is due to the definition of the $RMS$ misfit, which can only sense local difference between two signals no matter what is the cause of the differences.

What we have said about the misfits for $pm05(S1)$, $pm10(S1)$, and $pm20(S1)$ is also true in the case of the misfits for $pm05(S2)$, $pm10(S2)$, and $pm20(S2)$ (not shown here), and misfits for $pm05(S1 + S2)$, $pm10(S1 + S2)$, and $pm20(S1 + S2)$. The misfits for the latter case are shown in Figure 4b.

Figure 4c shows the envelope and phase misfits between the reference signal $S1 + S2$ and modified signals $pm05(S1) + S2$, $pm10(S1) + S2$, and $pm20(S1) + S2$. In the modified signals only the $S1$ component of the composed reference signal $S1 + S2$ is phase shift modified. The shape of the area with nonzero values of $TFPM(t, f)$ in the $(t, f)$ plane corresponds to the shape of the area with nonzero values of TFR of the $S1$ component (see Fig. 2), which is correct. The maximum values of the phase misfits $TFPM(t, f)$, $TPM(t)$, $FPM(f)$, and $PM$ are proportional to the percentage of the phase-shift modification. The maximum values are not equal

to the percentage of the phase-shift modification because the $S1$ component contributes to the TFR less than the $S2$ component (see Fig. 2).

Both $TFPM(t, f)$ and $TFEM(t, f)$ show alternating negative and positive values along the contact of the modified $S1$ and nonmodified $S2$. This is because the phase-shift modification of only the $S1$ component changed both the envelope and phase of the composed signal. The sign of the envelope and phase differences alternates along the time and frequency axes. Compared with the nonzero alternating values of $TFPM(t, f)$ in the case of the amplitude modification of only the $S1$ component (the previous section), the absolute values of the alternating-sign misfits are larger here because the relative change of the envelope due to phase-shift modification here is larger than the relative change of the phase due to amplitude modification in the former case (the previous section).

As in the previous examples of the phase-shift-modified signals, the $RMS$ misfit is approximately three times larger than the single-valued phase misfit $PM$ in all three levels of the phase-shift modification. This obviously is not affected by the nonzero values of the single-valued envelope misfit $EM$.

## Examples of the Time-Shift and Frequency Modifications and the Corresponding Misfits

To additionally illustrate the capability of the defined misfit criteria to quantify differences between a reference and some other signal, we also considered time-shift and frequency modifications of the reference signal.

We calculated the misfits between the reference signals and signals obtained by simple shifting along the time axis. For example, $tm1/120(S1)$ means the same signal as the reference one but delayed in time by 1/120 sec. Table 3 lists all modifications for which the misfits were calculated. The misfits between reference signal $S1$ and the time-shift-modified signals $tm1/120(S1)$, $tm1/60(S1)$, and $tm1/30(S1)$ are shown in Figure 5a. The time shift causes mainly a change of the phase but also a change of the envelope with respect to the reference signal. As expected, the absolute values of the phase misfits are clearly larger than those of the envelope misfits. The shape of the area with nonzero values of $TFPM(t, f)$ in the $(t, f)$ plane corresponds to the shape of the area with nonzero values of TFR of the reference signal. The maximum values of the envelope and phase misfits are directly proportional to the level of the time shift. At each frequency, the constant time delay of the entire signal causes a relative negative difference of the envelope to the left (along the time axis) of the signal's center and a relative positive difference of the envelope to the right of the signal's center. At each frequency, the projections of the negative and positive $TFEM(t, f)$ values cancel each other and result in zero $FEM(f)$, which is correct.

The $RMS$ misfit is approximately three times larger than the single-valued phase misfit $PM$ in all three levels of the time shift. Simple visual comparison of the reference and
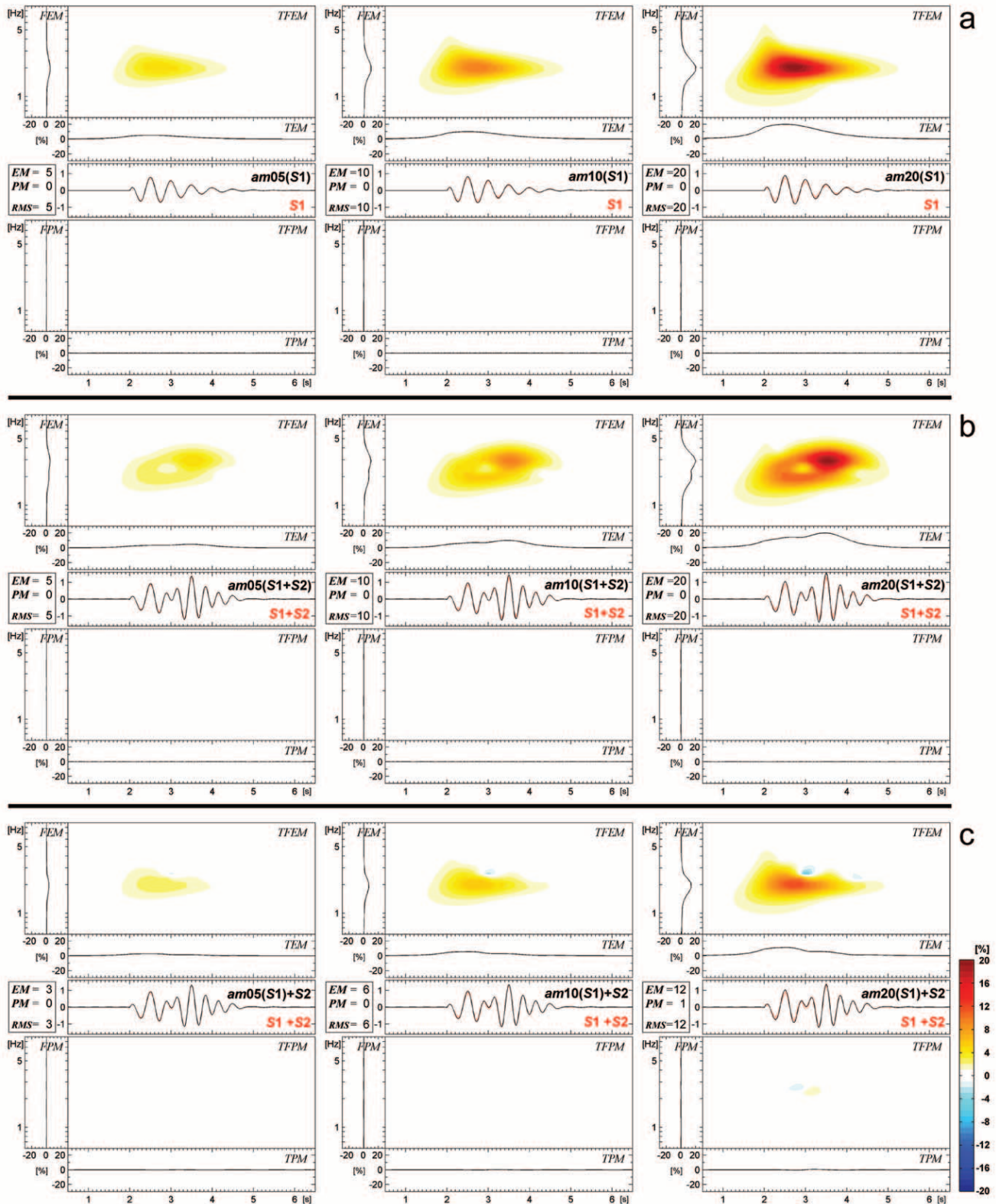
Figure 3. (a) Misfits between the reference signal $S1$ and modified signals $am05(S1)$, $am10(S1)$, and $am20(S1)$ (Middle) Reference and amplitude-modified signals, values of the single-valued envelope misfit $EM$, phase misfit $PM$, and $RMS$ misfit. (Top) Time-frequency envelope misfits $TFEM(t, f)$, time envelope misfits $TEM$ $(t)$, and frequency envelope misfits $FEM$ $(f)$. (Bottom) Time-frequency phase misfits $TFPM$ $(t, f)$, time phase misfits $TPM$ $(t)$, and frequency phase misfits $FPM$ $(f)$. (b) The same for $S1 + S2$ and modified signals $am05(S1 + S2)$, $am10(S1 + S2)$, and $am20(S1 + S2)$. (c) The same for $S1 + S2$ and modified signals $am05(S1) + S2$, $am10(S1) + S2$, and $am20(S1) + S2$.

Table 2

Reference and Phase-Shift-Modified Signals Used for Calculation of the Misfit Criteria

| Reference Signals | Phase-Shift-Modified Signals | | |
|---|---|---|---|
| $S1$ | $pm05(S1)$ | $pm10(S1)$ | $pm20(S1)$ |
| $S2$ | $pm05(S2)$ | $pm10(S2)$ | $pm20(S2)$ |
| $S1 + S2$ | $pm05(S1 + S2)$ | $pm10(S1 + S2)$ | $pm20(S1 + S2)$ |
| $S1 + S2$ | $pm05(S1) + S2$ | $pm10(S1) + S2$ | $pm20(S1) + S2$ |

modified signals in Figure 5a suggests that the *RMS* over-estimates the level of the phase-shift modification.

We do not show here the misfits for $tm1/120(S2)$, $tm1/60(S2)$, $tm1/30(S2)$, $tm1/120(S1 + S2)$, $tm1/60(S1 + S2)$, and $tm1/30(S1 + S2)$. They would lead to the similar statements.

Figure 5b shows the envelope and phase misfits between the reference signal $S1 + S2$ and modified signals $tm1/120(S1) + S2$, $tm1/60(S1) + S2$, and $tm1/30(S1) + S2$. In the modified signals only the $S1$ component of the composed reference signal $S1 + S2$ is delayed in time with respect to the reference signal. The shape of the area with nonzero values of $TFPM(t, f)$ in the $(t, f)$ plane corresponds to the shape of the area with nonzero values of TFR of the $S1$ component (see Fig. 2), which is correct. The maximum values of the phase misfits $TFPM(t, f)$, $TPM(t)$, $FPM(f)$, and $PM$ are proportional to the level of the time delay.

Both $TFPM(t, f)$ and $TFEM(t, f)$ show alternating negative and positive values along the contact of the modified $S1$ and nonmodified $S2$. This is because the time delay of only the $S1$ component changed both the envelope and phase of the composed signal. The sign of the envelope and phase differences alternates along the time and frequency axes. The absolute values of the alternating-sign $TFEM(t, f)$ misfits are larger than those of $TFPM(t, f)$ here because the relative change of the envelope due to time delay of only the $S1$ component of the signal here is larger than the relative change of the phase (as simple visual comparison of the reference and modified signals also suggests). Another consequence of the latter fact is that single-valued envelope and phase misfits, $EM$ and $PM$, are approximately equal, which is very different from the situation shown in Figure 5a.

As in the previous examples, the *RMS* misfit is approximately three times larger than the single-valued phase misfit $PM$ in all three levels of modification.

In the last canonical example we show misfits between the reference signal $S1$ and modified signals obtained by a small change of the signal's dominant frequency. Modified signals $fm 1.5(S1)$, $fm3.0(S1)$, and $fm6.0(S1)$ are obtained by the 1.5, 3.0, and 6.0% increase of the dominant frequency $f_1$ in definition (19), respectively. Figure 5c shows the reference and modified signals and the calculated misfits. As it is obvious from Figure 5c, the increase of the dominant frequency causes phase misfit in the time-frequency and time and frequency domains. The maximum values of the phase misfits are proportional to the level of the frequency modi-

fication. The position of the maximum $TFPM(t, f)$ along the frequency axis is determined by the dominant frequency. The position of the maximum along the time axis is determined by two opposing factors: the larger local envelope value is shifting the maximum to the left, whereas the larger phase difference is shifting it to the right along the time axis.

It is easy to understand the patterns of the envelope misfit. The increase of the signal's dominant frequency causes positive $TFEM(t, f)$ and $FEM(f)$ at frequencies larger than the dominant frequency. At the same time it causes negative $TFEM(t, f)$ and $FEM(f)$ at frequencies smaller than the dominant frequency. The maximum values of the envelope misfits are proportional to the level of the frequency modification. Relatively small values of $TEM(t)$ are caused by relatively small change of the signal's envelope in the time domain, as suggested by simple visual comparison of the reference and modified signals.

Again, the *RMS* values are approximately three times larger than the *PM* values.

## Application to the Numerical Solutions of the SCEC LOH.3 Problem

The SCEC code validation project (Day *et al.*, 2003) compared 3D wave-propagation codes for a hierarchy of test problems, ranging from simple point-source problems in canonical earth structures (e.g., layer over half-space) to propagating ruptures in complex 3D representations of Los Angeles Basin geology. Figure 6 shows the results for the layer over half-space test with anelastic attenuation ($Q$ is frequency independent), with point-dislocation source (Gaussian time function, spread 0.05 sec), Problem LOH.3. Three solutions were calculated by the FD codes. They are labeled as UCSB (Olsen, 1994), UCBL (Larsen and Grieger, 1998), and WCC1 (Graves, 1996). CMUN is the FE solution (Bao *et al.*, 1998); FK the frequency-wavenumber solution using a modification of the method of Apsel and Luco (1983). These calculations were done with codes that were under ongoing development, and the results were used to facilitate subsequent enhancements to some of the codes, especially with respect to the treatment of anelastic attenuation. The waveforms analyzed here are therefore not representative of the final performance of these methods, but rather are presented as examples of how the proposed misfit measures can provide valuable assessment and guidance during the develop-

Table 3

Reference and Time-Shift-Modified Signals Used for Calculation of the Misfit Criteria

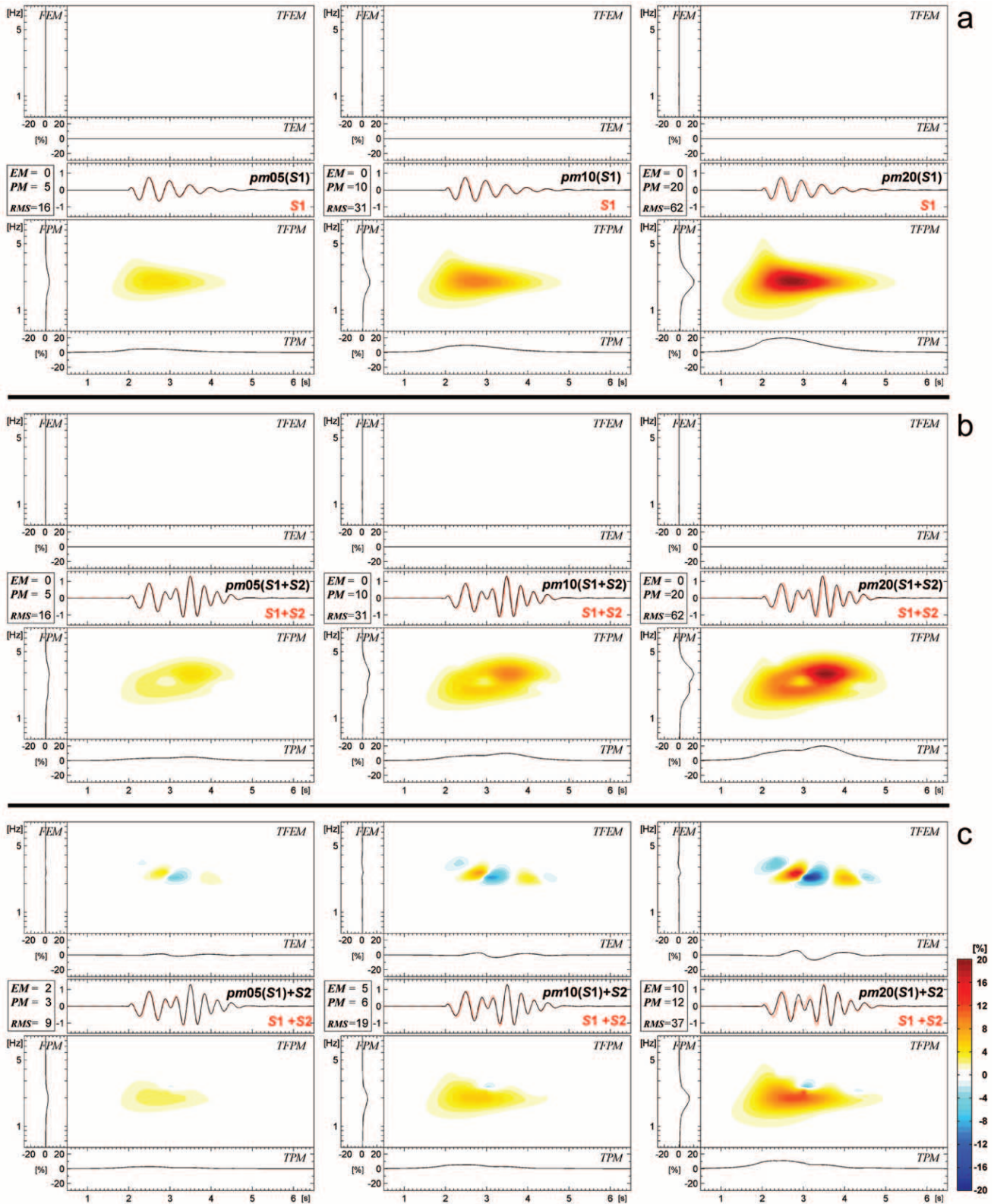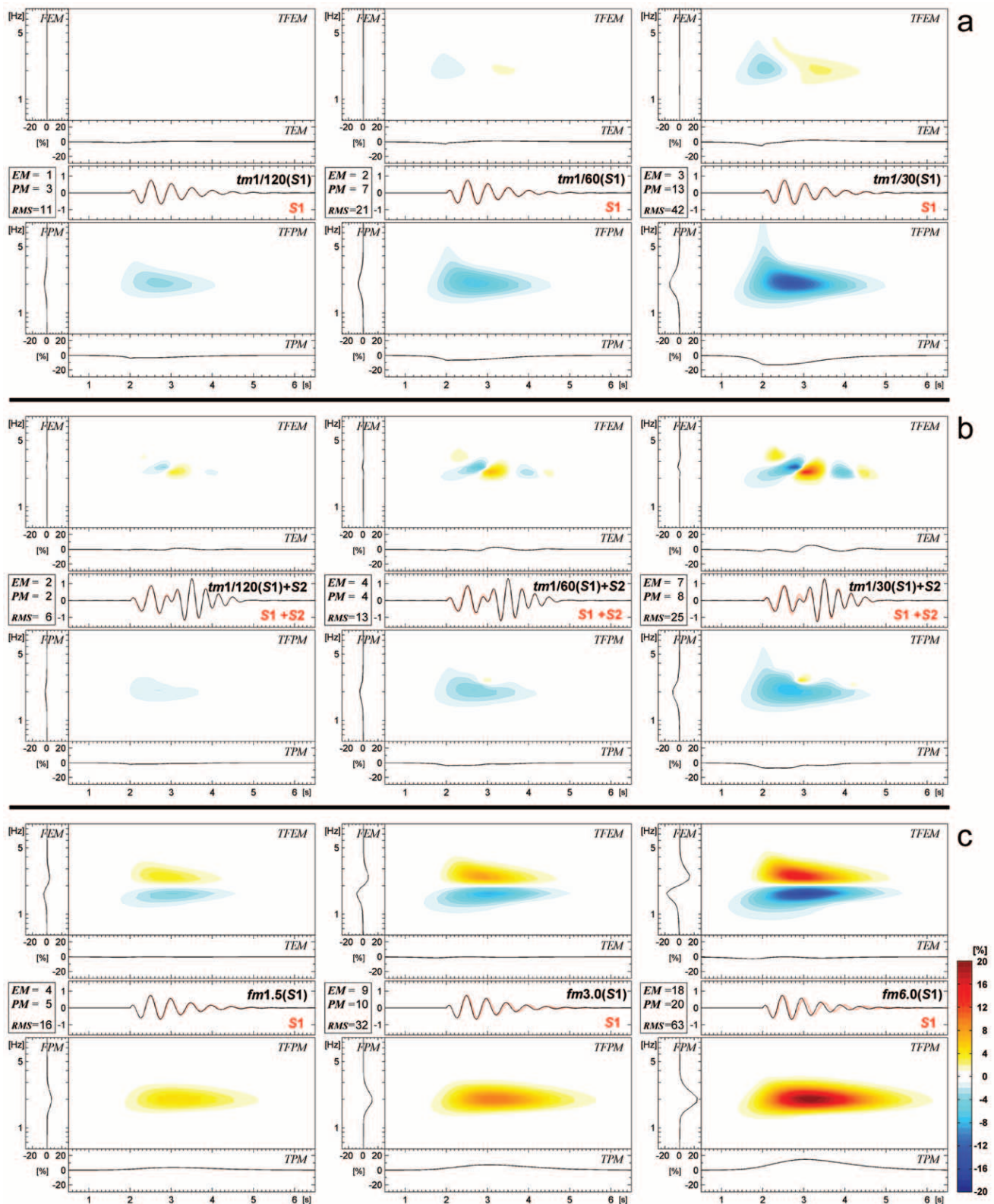| Reference Signals | Time-Shift-Modified Signals | | |
|---|---|---|---|
| $S1$ | $tm1/120(S1)$ | $tm1/60(S1)$ | $tm1/30(S1)$ |
| $S2$ | $tm1/120(S2)$ | $tm1/60(S2)$ | $tm1/30(S2)$ |
| $S1 + S2$ | $tm1/120(S1 + S2)$ | $tm1/60(S1 + S2)$ | $tm1/30(S1 + S2)$ |
| $S1 + S2$ | $tm1/120(S1) + S2$ | $tm1/60(S1) + S2$ | $tm1/30(S1) + S2$ |

Figure 4.   (a) Misfits between the reference signal $S1$ and modified signals $pm05(S1)$, $pm10(S1)$, and $pm20(S1)$. (Middle) Reference and phase-shift-modified signals, values of the single-valued envelope misfit $EM$, phase misfit $PM$, and $RMS$ misfit. (Top) Time-frequency envelope misfits $TFEM$ $(t, f)$, time envelope misfits $TEM(t)$, and frequency envelope misfits $FEM(f)$. (Bottom) Time-frequency phase misfits $TFPM(t, f)$, time phase misfits $TPM(t)$, and frequency phase misfits $FPM(f)$. (b) The same for $S1 + S2$ and modified signals $pm05(S1 + S2)$, $pm10(S1 + S2)$, and $pm20(S1 + S2)$. (c) The same for $S1 + S2$ and modified signals $pm05(S1) + S2$, $pm10(S1) + S2$, and $pm20(S1) + S2$.
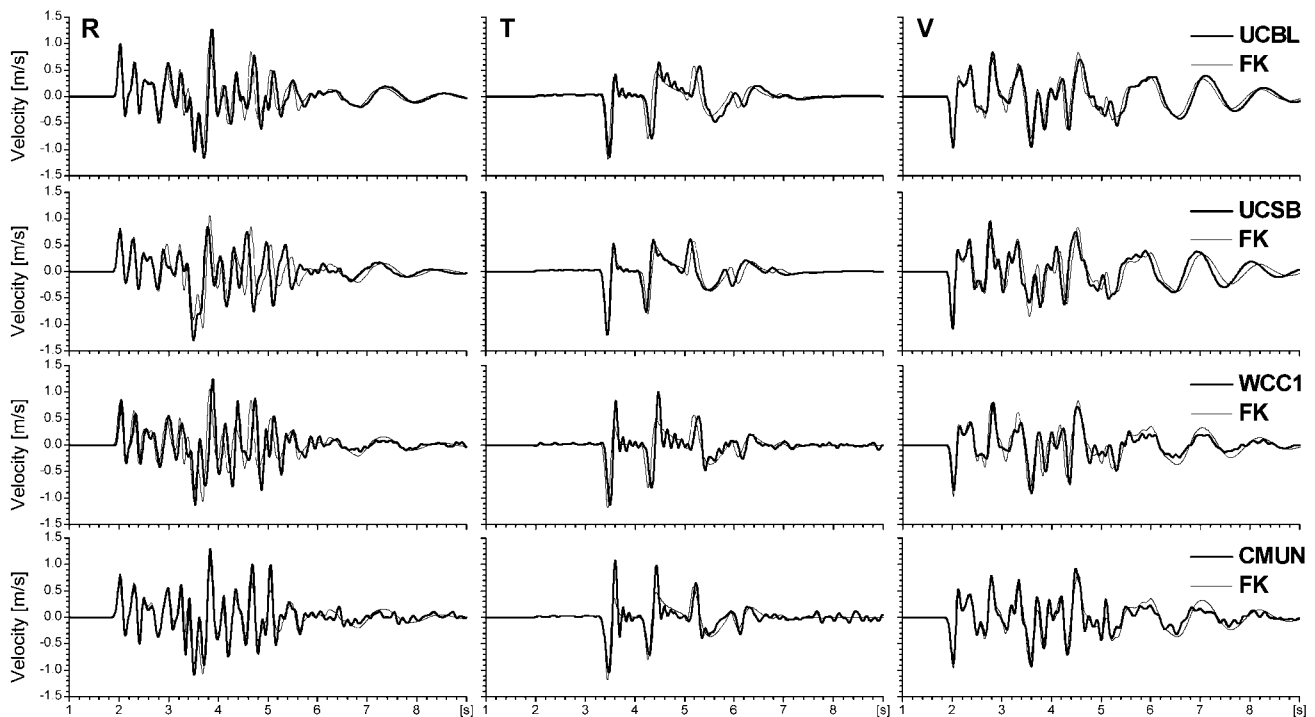
**Figure 5.** (a) Misfits between the reference signal $S1$ and modified signals $tm1/120(S1)$, $tm1/60(S1)$, and $tm1/30(S1)$. (Middle) Reference and time-shift-modified signals, values of the single-valued envelope misfit $EM$, phase misfit $PM$, and $RMS$ misfit. (Top) Time-frequency envelope misfits $TFEM$ $(t, f)$, time envelope misfits $TEM$ $(t)$, and frequency envelope misfits $FEM$ $(f)$. (Bottom) Time-frequency phase misfits $TFPM$ $(t, f)$, time phase misfits $TPM$ $(t)$, and frequency phase misfits $FPM$ $(f)$. (b) The same for $S1 + S2$ and modified signals $tm1/120(S1) + S2$, $tm1/60(S1) + S2$, and $tm1/30(S1) + S2$. (c) The same for $S1$ and modified signals $fm1.5(S1)$, $fm3.0(S1)$, and $fm6.0(S1)$.

Figure 6. Synthetic velocity seismograms obtained by three FD codes (UCSB, UCBL, and WCC1), one FE code (CMUN), and the frequency-wavenumber code (FK) for the SCEC Problem LOH.3. The FK solution is taken as a reference. Columns show radial (left), transverse (middle), and vertical (right) components, respectively, at distance of 10 km.

ment process. For a detailed description and evaluation of the numerical test for the LOH.3 problem see the report by Day *et al.*, (2003).

Here we focus on the formal quantification of the misfits between the tested FD/FE solutions and reference FK solution using the misfit criteria developed and tested previously. The time-frequency misfits can provide a useful insight into the differences between individual solutions and reference solution. Seeing the structure of the misfits and their quantification will be useful also because, given the relative simplicity of the problem (single layer over half-space), one can find the scatter of solutions a somewhat surprising.

Figure 7a shows the envelope and phase misfits between the UCBL and FK solutions, whereas Figure 7b shows the misfits between the UCSB and FK solutions. Both the UCBL and UCSB give reasonable phases at frequencies below 1 Hz, mainly at the radial and transversal components. The phases are not so good at frequencies above 2 Hz. Almost systematic negative phase misfits for UCBL and positive phase misfits for UCSB can be noticed in the velocity seismograms themselves. However, their quantification with respect to time and frequency can be only seen in the calculated phase misfits.

There is obvious envelope misfit for the two solutions even for frequencies below 1 Hz. The envelope misfit for UCBL becomes larger for frequencies above 3 Hz. For the UCSB solution, the envelope misfit becomes larger for frequencies above 2 Hz, mainly at the radial and vertical components.

The average (from the three components) *EM* and *PM* values for the UCBL solution are 16% and 21%. The average *EM* and *PM* values for the UCSB solution are 20% and 24%. We do not evaluate *RMS* because the upper-frequency limit in the test was set up as 5 Hz, and *RMS* would account for the entire velocity seismograms which do have nonnegligible spectral content above 5 Hz.

Figure 8a shows the envelope and phase misfits between the WCC1 and FK solutions, whereas Figure 8b shows the misfits between the CMUN and FK solutions. Figure 8b clearly shows that, overall, the CMUN solution gives the best phases. At the same time, the WCC1 solution gives the best phases for frequencies below 2 Hz and the worst phases for frequencies above 3 Hz.

Overall, both the WCC1 and CMUN solutions give worse envelopes than the UCBL and UCBS solutions do. The envelope misfits are negative for frequencies below 1 Hz and positive for frequencies above 2 Hz. This agrees well with the fact that both the WCC1 and CMUN incorporated a simplified model of the attenuation (analogous to mass-proportional Rayleigh damping; see, e.g., Graves, 1996) which suppresses motion at lower frequencies and enhances motion at higher frequencies.
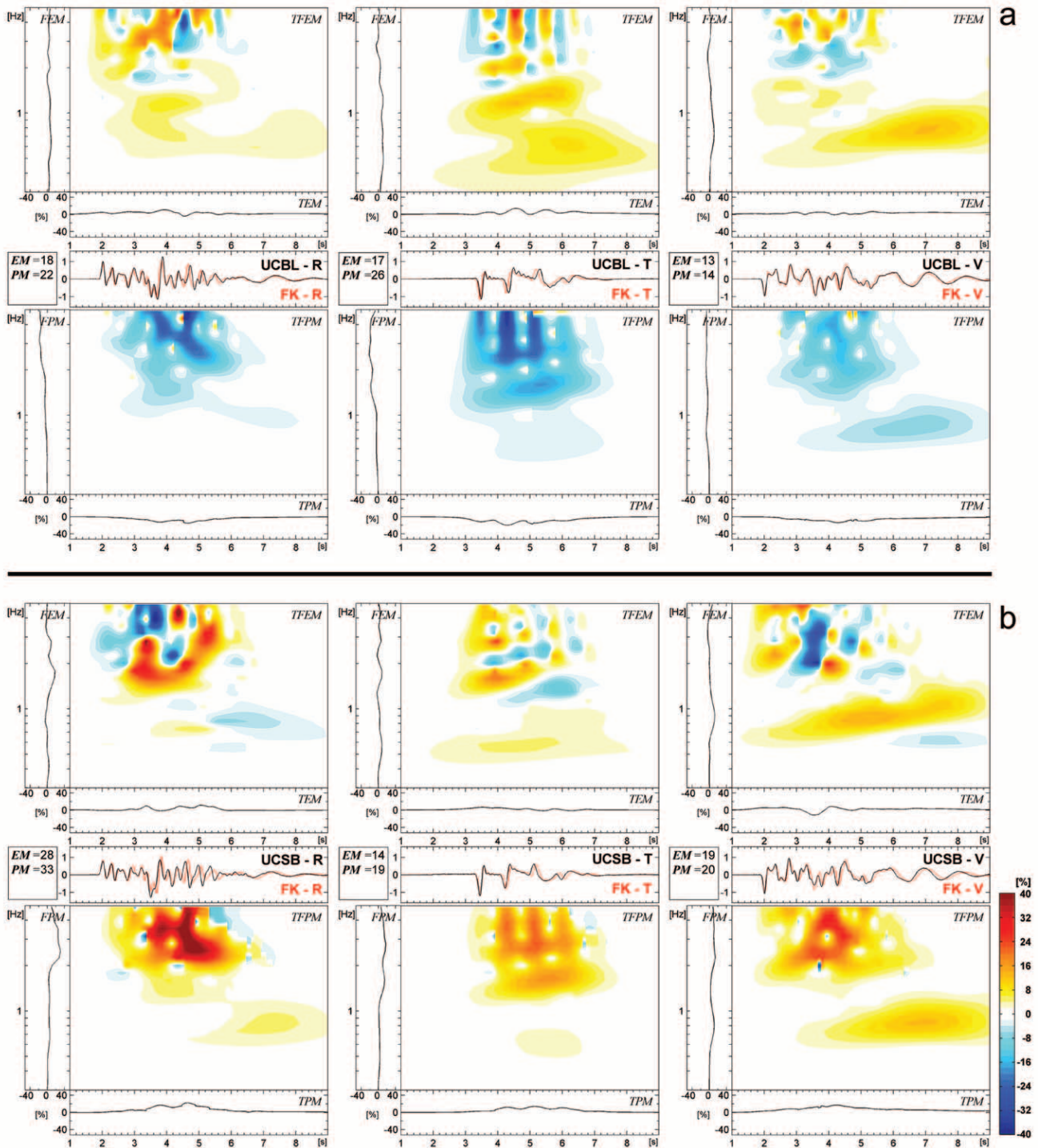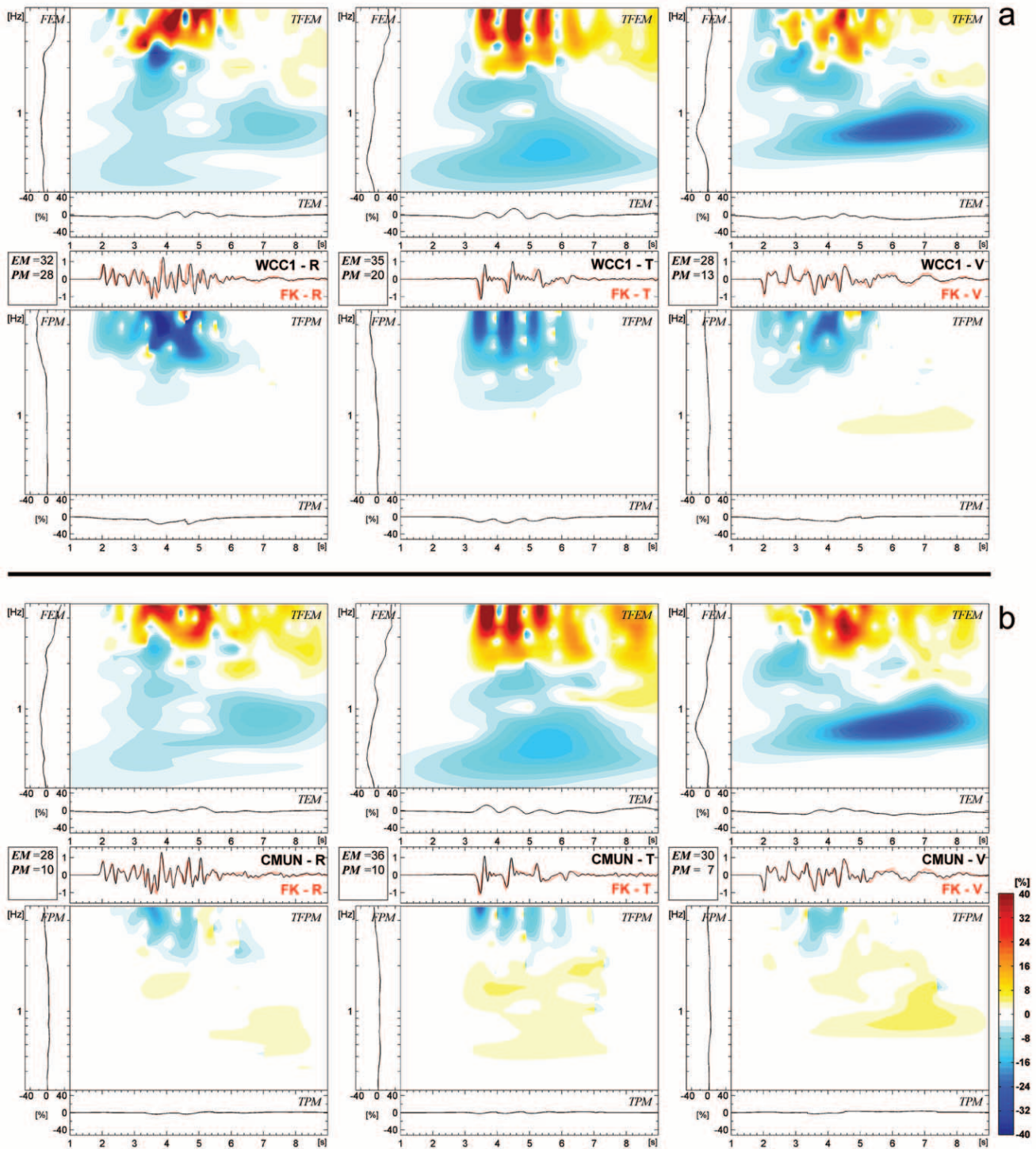
Figure 7. (a) Misfits between the UCBL solution and reference FK solution. (Left) Radial component. (Center) Transversal component. (Right) Vertical component. (Middle row) Reference and UCBL seismograms, values of the single-valued envelope misfit *EM* and phase misfit *PM*. (Top) Time-frequency envelope misfits *TFEM*$(t, f)$, time envelope misfits *TEM*$(t)$, and frequency envelope misfits *FEM*$(f)$. (Bottom) Time-frequency phase misfits *TFPM*$(t, f)$, time phase misfits *TPM*$(t)$, and frequency phase misfits *FPM*$(f)$. (b) The same for the UCSB solution.

Figure 8.    (a) Misfits between the WCC1 solution and reference FK solution. (Left) radial component. (Center) Transversal component. (Right) Vertical component. (Middle row) Reference and WCC1 seismograms, values of the single-valued envelope misfit *EM* and phase misfit *PM*. (Top) Time-frequency envelope misfits $TFEM(t,f)$, time envelope misfits $TEM(t)$, and frequency envelope misfits $FEM(f)$. (Bottom) Time-frequency phase misfits $TFPM(t,f)$, time phase misfits $TPM(t)$, and frequency phase misfits $FPM(f)$. (b) The same for the CMUN solution.

The average *EM* and *PM* values for the WCC1 solution are 32% and 20%. The average *EM* and *PM* values for the CMUN solution are 31% and 9%.

## Conclusions

We have developed and numerically tested quantitative misfit criteria for comparison of seismograms. The misfit criteria include

- Time-frequency envelope $TFEM(t, f)$ and phase $TFPM(t, f)$ misfits,
- Time-dependent envelope $TEM(t)$ and phase $TPM(t)$ misfits,
- Frequency-dependent envelope $FEM(f)$ and phase $FPM(f)$ misfits,
- Single-valued envelope *EM* and phase *PM* misfits.

The misfit criteria are based on the TFR of the seismograms obtained as the continuous wavelet transform with the analyzing Morlet wavelet.

We tested properties of the misfit criteria using canonical signals and also relatively complicated synthetics obtained by several numerical methods. The canonical signals, taken as the reference signals, were specifically amplitude, phase shift, time shift, and frequency modified to demonstrate the ability of the misfit criteria to properly quantify and recognize the character and cause of the misfit between the reference and modified signal.

Partial conclusions from the numerical tests for pure amplitude modification are:

- The maximum values of the envelope misfits are proportional to the level of the modification; they exactly equal the percentage of the modification if the entire signal in the time domain is multiplied by some (constant) coefficient.
- In the latter case, the all phase misfits are zero because there is no phase modification of the reference signal.
- The *RMS* misfit exactly equals the single-valued envelope misfit *EM*.

Partial conclusions from the numerical tests for pure phase-shift modification are:

- The maximum values of the phase misfits are proportional to the level of the phase-shift modification; they exactly equal the percentage of the modification if the time-dependent phase of the analytical signal is shifted by the same value at each time.
- In the latter case, all envelope misfits are zero because there is no amplitude modification of the reference signal.
- The *RMS* misfit is approximately three times larger than the single-valued phase misfit *PM*.

We also calculated the misfits between the reference signals and modified signals obtained by simple shifting along the time axis or by a small change of the signal's dominant frequency. In all cases the misfit criteria properly quantified and characterized misfits between the reference and modified signals.

Application to SCEC test problem LOH.3 (single layer over half-space with attenuation) demonstrated that the single-valued envelope and phase misfit criteria successfully capture and quantify the principal visible discrepancies between the numerical solutions and the FK reference solution. By putting phase and amplitude differences, respectively, on a quantitative basis, the proposed single-valued misfit metrics provide a scheme for rating different numerical methods according to their relative suitability for specific applications. Furthermore, the time-frequency-dependent, time-dependent and frequency-dependent misfit functions provide valuable additional descriptions of the discrepancies, together with insights into their origins. In particular, these misfit functions clearly exhibit bandwidth limitations on phase and amplitude accuracy, sign changes in the phase misfits between different numerical solutions, and spectral biases due to approximations to the constant-$Q$ attenuation operator, all of which can be traced to the characteristics of the respective numerical solution methods.

In contrast, the standard *RMS* misfit matches the single-valued envelope misfit (*EM*) only in the case of a pure amplitude modification of the signal. In all other cases, *RMS* considerably overestimates misfits compared with *EM* and *PM*. In contrast with *RMS*, the more precise and complete characterization provided by *EM* and *PM* makes the latter more effective for assessing the applicability of numerical solution methods to specific applications. Suitability for phase-sensitive applications such as surface-wave dispersion or travel-time studies, for example, can be assessed by giving higher weight to *PM* than to *EM*, whereas suitability for amplitude-sensitive applications such as strong-motion simulation might be better assessed by giving higher weight to *EM* than to *PM*.

The Fortran95 program package TF-MISFITS is available at www.nuquake.eu/Computer_Codes/.

## Acknowledgments

## References

Aoi, S., and H. Fujiwara (1999). 3D finite-difference method using discontinuous grids, *Bull. Seism. Soc. Am.* **89,** 918–930.

Apsel, R. J., and J. E. Luco (1983). On the Green's functions for a layered half-space. Part II, *Bull. Seism. Soc. Am.* **73,** 931–951.

Bao, H, J. Bielak, O. Ghattas, L. F. Kallivokas, D. R. O'Hallaron, J. R. Shewchuk, and J. Xu (1998). Large-scale simulation of elastic wave propagation in heterogeneous media on parallel computers, *Comput. Methods Appl. Mech. Eng.* **152,** 85–102.

Daubechies, I. (1992). *Ten Lectures on Wavelets*, Society for Industrial and Applied Mathematics, Philadelphia.

Day, S. M., J. Bielak, D. Dreger, R. Graves, S. Larsen, K. Olsen, and A. Pitarka (2003). Tests of 3D elastodynamic codes: Final report for Lifelines Project 1A02, Pacific Earthquake Engineering Research Center, Berkeley, California.

Geller, R. J., and N. Takeuchi (1995). A new method for computing highly accurate DSM synthetic seismograms, *Geophys. J. Int.* **123,** 449–470.

Graves, R. W. (1996). Simulating seismic wave propagation in 3D elastic media using staggered-grid finite differences, *Bull. Seism. Soc. Am.* **86,** 1091–1106.

Holschneider, M. (1995). *Wavelets: An Analysis Tool*, Clarendon Press, Oxford.

Igel, H., R. Barsch, P. Moczo, J.-P. Vilotte, Y. Capdeville, and E. Vye (2005). The EU SPICE Project: a digital library with codes and training material in computational seismology (abstract), *EOS Trans. AGU* **86,** no. 52, Fall Meet. Suppl., S13A–0179.

Ji, C., D. J. Wald, and D. V. Helmberger (2002). Source description of the 1999 Hector Mine, California earthquake; Part I: wavelet domain inversion theory and resolution analysis, *Bull. Seism. Soc. Am.* **92,** 1192–1207.

Kristek, J., P. Moczo, and R. J. Archuleta (2002). Efficient methods to simulate planar free surface in the 3D 4th-order staggered-grid finite-difference schemes, *Stud. Geophys. Geod.* **46,** 355–381.

Larsen, S., and J. Grieger (1998). Elastic modeling initiative, Part III: 3-D computational modeling, *Soc. Exp. Geophys.*, Expanded Abstracts **17,** 1803.

Moczo, P., J.-P. Ampuero, J. Kristek, M. Gális, S. M. Day, and H. Igel (2005). The European Network SPICE Code Validation (abstract), *EOS Trans. AGU* **86,** no. 52, Fall Meet. Suppl., S13A–0180.

Olsen, K. B. (1994). Simulation of three-dimensional wave propagation in the Salt Lake Basin, *Ph.D. Thesis*, University of Utah, Salt Lake City, Utah, 157 pp.

## Appendix

Locally normalized time-frequency envelope misfit

$$TFEM_L(t, f) \; = \; \frac{\Delta E\ (t, f)}{|W_{\mathrm{REF}}(t, f)|} \tag{A1}$$

Locally normalized time-frequency phase misfit

$$TFPM_L(t, f) \; = \; \frac{\Delta P\ (t, f)}{|W_{\mathrm{REF}}(t, f)|} \tag{A2}$$

Locally normalized time-dependent envelope misfit

$$TEM_L(t) \; = \; \frac{\langle \Delta E\ (t, f)\rangle_f}{\langle |W_{\mathrm{REF}}(t, f)|\rangle_f} \tag{A3}$$

Locally normalized time-dependent phase misfit

$$TPM_L(t) \; = \; \frac{\langle \Delta P\ (t, f)\rangle_f}{\langle |W_{\mathrm{REF}}(t,f)|\rangle_f} \tag{A4}$$

Locally normalized frequency-dependent envelope misfit

$$FEM_L(f) \; = \; \frac{\langle \Delta E\ (t, f)\rangle_t}{\langle |W_{\mathrm{REF}}(t, f)|\rangle_t} \tag{A5}$$

Locally normalized frequency-dependent phase misfit

$$FPM_L(f) \; = \; \frac{\langle \Delta P\ (t, f)\rangle_t}{\langle |W_{\mathrm{REF}}(t, f)|\rangle_t} \tag{A6}$$

Geophysical Institute
Slovak Academy of Sciences
Dubravska cesta 9
845 28 Bratislava, Slovak Republic
  (M.K.)

Faculty of Mathematics, Physics and Informatics
Comenius University
Mlynska dolina F1
842 48 Bratislava, Slovak Republic
  (J.K., P.M.)

Department of Geological Sciences
San Diego State University
San Diego, California 92182
  (S.M.D.)