Third International Symposium on the Effects of Surface Geology on Seismic Motion
Grenoble, France, 30 August - 1 September 2006
Paper Number: 65

QUANTITATIVE MISFIT CRITERIA FOR COMPARISON OF SEISMOGRAMS

Miriam KRISTEKOVA[1], Jozef KRISTEK[2,1], Peter MOCZO [2,1], Steven M. DAY[3]
1 Geophysical Institute, Slovak Academy of Sciences, Bratislava, Slovakia
2 Faculty of Math., Physics and Informatics, Comenius University, Bratislava, Slovakia
3 Department of Geological Sciences, San Diego State University, San Diego, USA

ABSTRACT - Misfit criteria based on the time-frequency representation (TFR) of seismograms have been developed and numerically tested. The TFR is obtained using the continuous wavelet transform. The criteria include time-frequency, time-dependent, frequency-dependent, and single-valued envelope and phase misfits. The misfit criteria were numerically tested using canonical signals (taken as the reference signals) and their modifications. The ability of the criteria to quantify and characterize misfits between the reference and modified signals was examined using pure amplitude, phase-shift, time-shift, and frequency modifications of the reference signals. Except pure amplitude modification, RMS (root mean square) considerably overestimates the misfits and does not characterize them. The criteria are illustrated using synthetics for the Grenoble Valley.

1. Introduction

Often two seismograms are compared by simply showing them together. In some papers a difference between tested $s(t)$ and reference $s_{REF}(t)$ seismograms, $D(t) = s(t) - s_{REF}(t)$, is used. It is obvious that $D(t)$ can provide very misleading information. For example, in the case of 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. A commonly used single-valued misfit criterion is the RMS (Root Mean Square) misfit defined as

$$RMS = \sqrt{\sum_t \left| s(t) - s_{REF}(t) \right|^2 \Big/ \sum_t \left| s_{REF}(t) \right|^2} \quad . \tag{1}$$

It is clear from the above definitions that $D(t)$ and RMS quantify a difference between two seismograms without having the property of recognizing and characterizing the difference. Still the question is whether they can really properly quantify it.

Some modifications of a (reference) signal can be more visible and understandable in the time domain, other in the frequency domain. Whereas one modification changes only/mainly amplitudes or envelope, some other change only/mainly phase. Given the variety of aspects, recall that the complete characterization of a signal can be obtained by its time-frequency representation (TFR). The TFR enables to see time evolution of the spectral content. Therefore, it seems quite natural to define misfit criteria based on the TFR. The importance of reasonable misfit criteria has been recently 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). The goal of the SPICE Code Validation is a long-term interactive-web-based platform for detailed comparison and testing methods and computer codes for the numerical modeling in seismology.

2. Time-frequency Misfit Criteria

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

$$CWT_{(a,b)}\left\{s(t)\right\} = \left|a\right|^{-1/2} \int_{-\infty}^{\infty} s(t)\ \psi^{*}\left(t-b\right)/a\ dt\ , \tag{2}$$

where t is time, a scale parameter, b translational parameter. As an analyzing wavelet we take Morlet wavelet $\psi(t) = \pi^{-1/4} \exp(i\,\omega_0\, t) \exp\left(-t^2/2\right)$ with $\omega_0 = 6$, which is an analytical signal. The TFR of signal $s(t)$ can be defined as

$$W\left(t, f\right) = CWT_{(a,b)}\left\{s(t)\right\}; \quad a = \omega_0/2\pi f\ ,\ b = t\ . \tag{3}$$

For the continuous wavelet transform and Morlet wavelet see, e.g., Daubechies (1992) and Holschneider (1995). Let $W_{REF}\left(t, f\right)$ be the TFR of reference signal $s_{REF}\left(t\right)$, $W\left(t, f\right)$ TFR of signal $s\left(t\right)$, and N_T and N_F the numbers of time and frequency samples in the time-frequency (TF) plane, respectively. We follow Kristekova et al. (2006) and define:

local TF envelope difference

$$\Delta E\left(t, f\right) = \left|W\left(t, f\right)\right| - \left|W_{REF}\left(t, f\right)\right|\ , \tag{4}$$

and a local TF phase difference

$$\Delta P\left(t, f\right) = \left|W_{REF}\left(t, f\right)\right| \left\{\mathrm{Arg}\left[W\left(t, f\right)\right] - \mathrm{Arg}\left[W_{REF}\left(t, f\right)\right]\right\}/\pi\ , \tag{5}$$

time-frequency envelope misfits

$$TFEM\left(t, f\right) = \frac{\Delta E\left(t, f\right)}{\max_{t,f}\left(\left|W_{REF}\left(t, f\right)\right|\right)}, \qquad TFEM_L\left(t, f\right) = \frac{\Delta E\left(t, f\right)}{\left|W_{REF}\left(t, f\right)\right|}\ , \tag{6}$$

time-frequency phase misfits

$$TFPM\left(t, f\right) = \frac{\Delta P\left(t, f\right)}{\max_{t,f}\left(\left|W_{REF}\left(t, f\right)\right|\right)}, \qquad TFPM_L\left(t, f\right) = \frac{\Delta P\left(t, f\right)}{\left|W_{REF}\left(t, f\right)\right|}\ . \tag{7}$$

$TFEM\left(t, f\right)$ and $TFEM_L\left(t, f\right)$ characterize the difference between the envelopes of the two signals as functions of time and frequency. Analogously, $TFPM\left(t, f\right)$ and $TFPM_L\left(t, f\right)$ characterize the difference between the phases of the two signals as a function of time and frequency. In the globally normalized misfits, local TF envelope and phase differences for a given $\left(t, f\right)$ are normalized with respect to the maximum absolute TFR value of the reference signal. In the locally normalized misfits, the local differences are normalized with respect to the local absolute TFR value of the reference signal.

Assume, e.g., 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. While $TFEM_L\left(t, f\right)$ will be constant, $TFEM(t,f)$ will not be constant because the same relative change does not mean the same value of the envelope difference (4). In other words, while $TFEM_L(t,f)$ only reflects the 'structure' of the modification, $TFEM(t,f)$ also recognizes 'large and small' amplitudes in the reference signal. Similarly, in the case of change of the signal's phase by 5%, $TFPM_L(t,f)$ will be constant because the change applies to the entire signal, while $TFPM(t,f)$ will not be constant along the time axis.

If we want to see the misfit between two signals as a function of (only) time, we can project the TF misfit onto the time domain:

time-dependent envelope misfits

$$TEM(t) = \frac{\langle \Delta E(t,f)\rangle_f}{\max_t\left(\langle |W_{REF}(t,f)|\rangle_f\right)}, \qquad TEM_L(t) = \frac{\langle \Delta E(t,f)\rangle_f}{\langle |W_{REF}(t,f)|\rangle_f}\ , \tag{8}$$

time-dependent phase misfits

$$TPM(t) = \frac{\langle \Delta P(t,f)\rangle_f}{\max_t\left(\langle |W_{REF}(t,f)|\rangle_f\right)}, \qquad TPM_L(t) = \frac{\langle \Delta P(t,f)\rangle_f}{\langle |W_{REF}(t,f)|\rangle_f}\ , \tag{9}$$

where $\langle \Theta(t,f)\rangle_f = \sum_f \Theta(t,f) \Big/ N_F$ with Θ representing either ΔE, ΔP, $|W_{REF}|$.

Analogously, projection onto the frequency domain gives the frequency-dependent envelope and phase misfit:

$$FEM(f) = \frac{\langle \Delta E(t,f)\rangle_t}{\max_f\left(\langle |W_{REF}(t,f)|\rangle_t\right)}, \qquad FEM_L(f) = \frac{\langle \Delta E(t,f)\rangle_t}{\langle |W_{REF}(t,f)|\rangle_t}\ , \tag{10}$$

$$FPM(f) = \frac{\langle \Delta P(t,f)\rangle_t}{\max_f\left(\langle |W_{REF}(t,f)|\rangle_t\right)}, \qquad FPM_L(f) = \frac{\langle \Delta P(t,f)\rangle_t}{\langle |W_{REF}(t,f)|\rangle_t}\ , \tag{11}$$

where $\langle \Theta(t,f)\rangle_t = \sum_t \Theta(t,f) \Big/ N_T$ with Θ representing either ΔE, ΔP, $|W_{REF}|$.

Single-valued envelope or phase misfits between two signals can be defined as

$$EM = \sqrt{\frac{\sum_f\sum_t |\Delta E(t,f)|^2}{\sum_f\sum_t |W_{REF}(t,f)|^2}}\ , \qquad PM = \sqrt{\frac{\sum_f\sum_t |\Delta P(t,f)|^2}{\sum_f\sum_t |W_{REF}(t,f)|^2}}\ . \tag{12}$$

These single-valued misfits will be compared with the RMS misfit defined by eq. (1).

3. Reference signals and their canonical modifications

The properties of the defined misfit criteria were numerically demonstrated using three signals,

$$S1 = A_1(t-t_1)\exp[-2(t-t_1)]\cdot\cos[2\pi f_1(t-t_1)+\varphi_1\pi]\cdot H(t-t_1)\ , \tag{13}$$

$$S2 = A_2\exp[-2(t-t_2)^2]\cdot\cos[2\pi f_2(t-t_2)+\varphi_2\pi]\ , \tag{14}$$

and their superposition, $S1$+$S2$. $H(t)$ is the Heaviside step function. $S1$ is a harmonic carrier with a sudden onset and decaying amplitude. Its amplitude spectrum has a peak at 2 Hz. $S2$ is Gabor signal with relatively narrow spectrum peaked at 3 Hz. The three signals, $S1$, $S2$, and $S1$+$S2$, their spectra as well as their TFR are shown in Fig. 1.
The amplitude modification: $am05(s(t)) = 1.05 \cdot s(t)$, $am20(s(t)) = 1.20 \cdot s(t)$.
The phase-shift modification: $pm05(s(t)) = \mathrm{Re}\left[A(t) \exp\left(i\varphi(t) + 0.05\, i\, \pi \right) \right]$, $pm20(s(t)) = \mathrm{Re}\left[A(t) \exp\left(i\varphi(t) + 0.20\, i\, \pi \right) \right]$. Here, $A(t)$ and $\varphi(t)$ are the amplitude and the phase of the analytical signal $\hat{s}(t) = A(t) \exp[i\,\varphi(t)]$. (We show only some of tested modifications.)

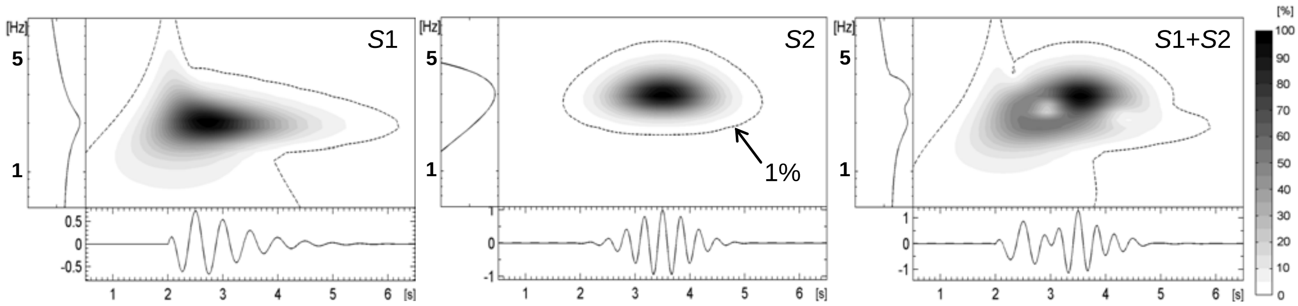

Figure 1. Reference signals S1, S2, S1+S2, their Fourier power spectra, and time-frequency representations (moduli of the continuous wavelet transforms of the signals).

4. Misfits for the Amplitude-modified and Phase-shift-modified Signals

The globally and locally normalized misfits between reference signal $S1$+$S2$ and the amplitude-modified signal $am05(S1+S2)$ are shown Fig. 2a. Similarly, Fig. 3a shows the misfits for $am20(S1+S2)$. The shape of the area with nonzero values of $TFEM(t, f)$ and $TFEM_L(t, f)$ in the (t, f)-plane corresponds to the shape of the area with nonzero values of TFR of the reference signal (values equal or larger than 1% of the TFR maximum, see Fig. 1). The maximum of $TFEM(t, f)$ equals the level of the amplitude modification in both cases (5 and 20%). Its position in the (t, f)-plane corresponds to the position of the maximum of the TFR; this is correct because this is the position, where the absolute amplitude difference between the reference and modified signal has the largest value. The statements on the maximum 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 in both cases. As expected, the locally normalized $TFEM_L(t, f)$ is constant over the area with nonzero values of TFR of the reference signal (small deviations from the correct value close to the area's border are due to numerical inaccuracies of divisions by small values).

All the phase misfits are zero: the phase misfits correctly reflect the fact that there is no phase modification of the reference signal.

Fig. 2b shows the misfits between the reference signal $S1$+$S2$ and modified signal $am05(S1)+S2$ in which only the S1-component was 5%-amplitude modified. Similarly, Fig. 3b shows the misfits for $am20(S1)+S2$. The shape of the area with nonzero values of $TFEM(t, f)$ and $TFEM_L(t, f)$ in the (t, f)-plane correctly corresponds to the shape of the area with nonzero values of TFR of the $S1$ component. The maxima of the envelope misfits are proportional to the level of the amplitude modification. They 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. 1. The alternating negative and positive values

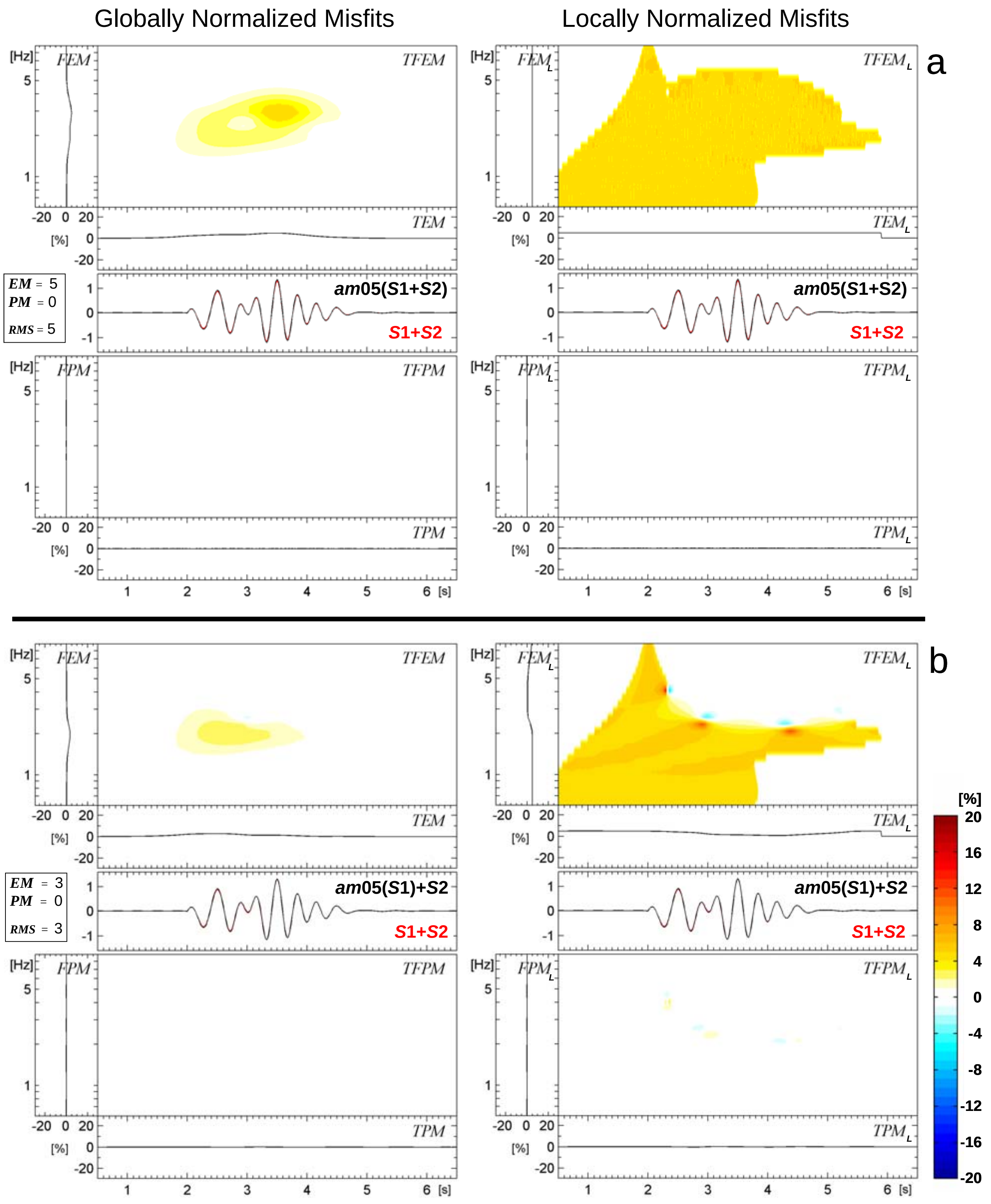

Figure 2. (a) Misfits between the reference signal $S1+S2$ and modified signal $am05(S1+S2)$. Middle row: reference and amplitude-modified signals, values of the single-valued envelope misfit EM, phase misfit PM, and RMS misfit. Upper row: Time-frequency envelope misfits $TFEM(t,f)$ and $TFEM_L(t,f)$, time envelope misfits $TEM(t)$ and $TEM_L(t)$, and frequency envelope misfits $FEM(f)$ and $FEM_L(f)$. Bottom row: Time-frequency phase misfits $TFPM(t,f)$ and $TFPM_L(t,f)$, time phase misfits $TPM(t)$ and $TPM_L(t)$, and frequency phase misfits $FPM(f)$ and $FPM_L(f)$. (b) The same for $S1+S2$ and modified signal $am05(S1)+S2$.

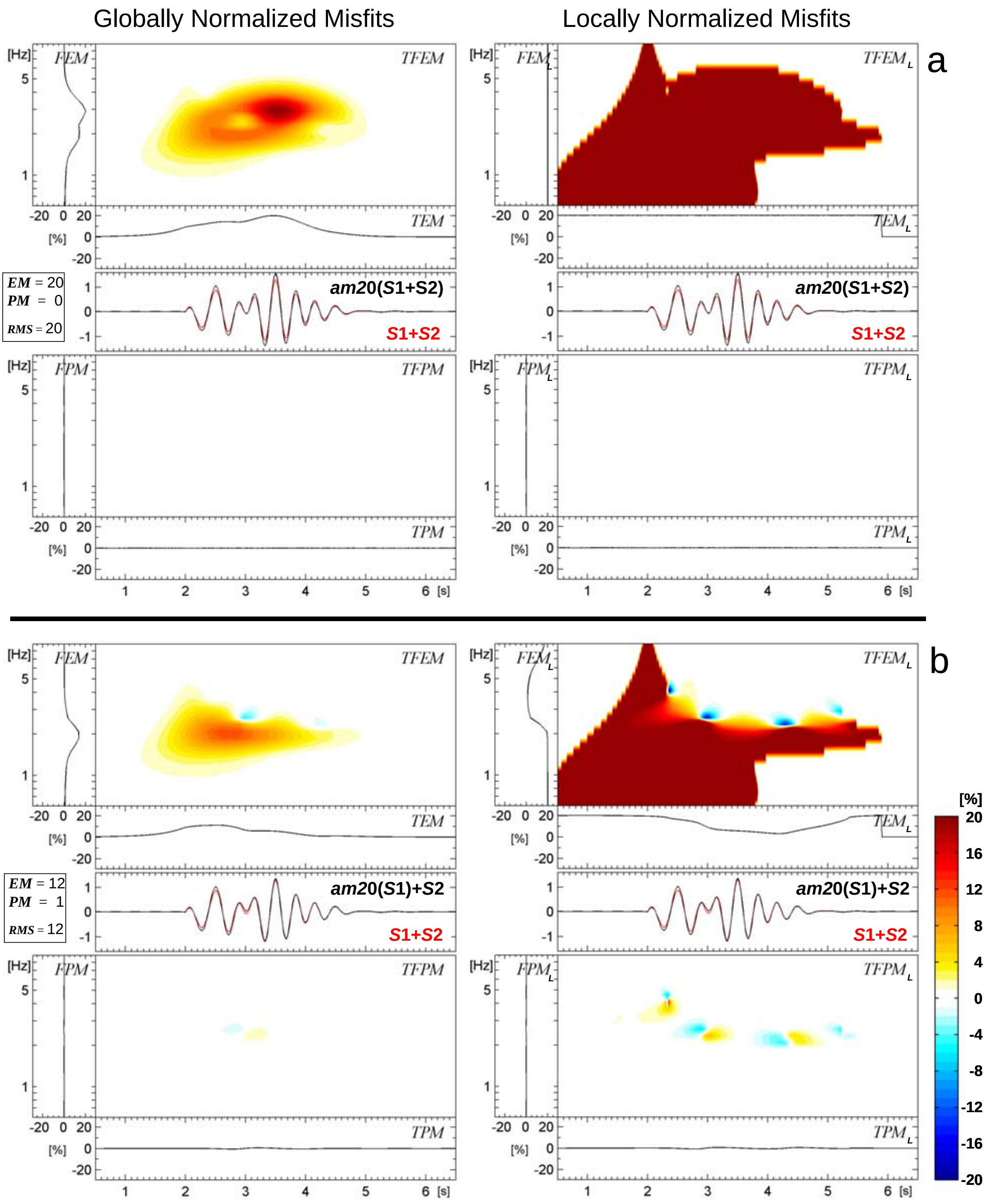

Figure 3. (a) Misfits between the reference signal $S1+S2$ and modified signal $am20(S1+S2)$. Middle row: reference and amplitude-modified signals, values of the single-valued envelope misfit EM, phase misfit PM, and RMS misfit. Upper row: Time-frequency envelope misfits $TFEM(t,f)$ and $TFEM_L(t,f)$, time envelope misfits $TEM(t)$ and $TEM_L(t)$, and frequency envelope misfits $FEM(f)$ and $FEM_L(f)$. Bottom row: Time-frequency phase misfits $TFPM(t,f)$ and $TFPM_L(t,f)$, time phase misfits $TPM(t)$ and $TPM_L(t)$, and frequency phase misfits $FPM(f)$ and $FPM_L(f)$. (b) The same for $S1+S2$ and modified signal $am20(S1)+S2$.

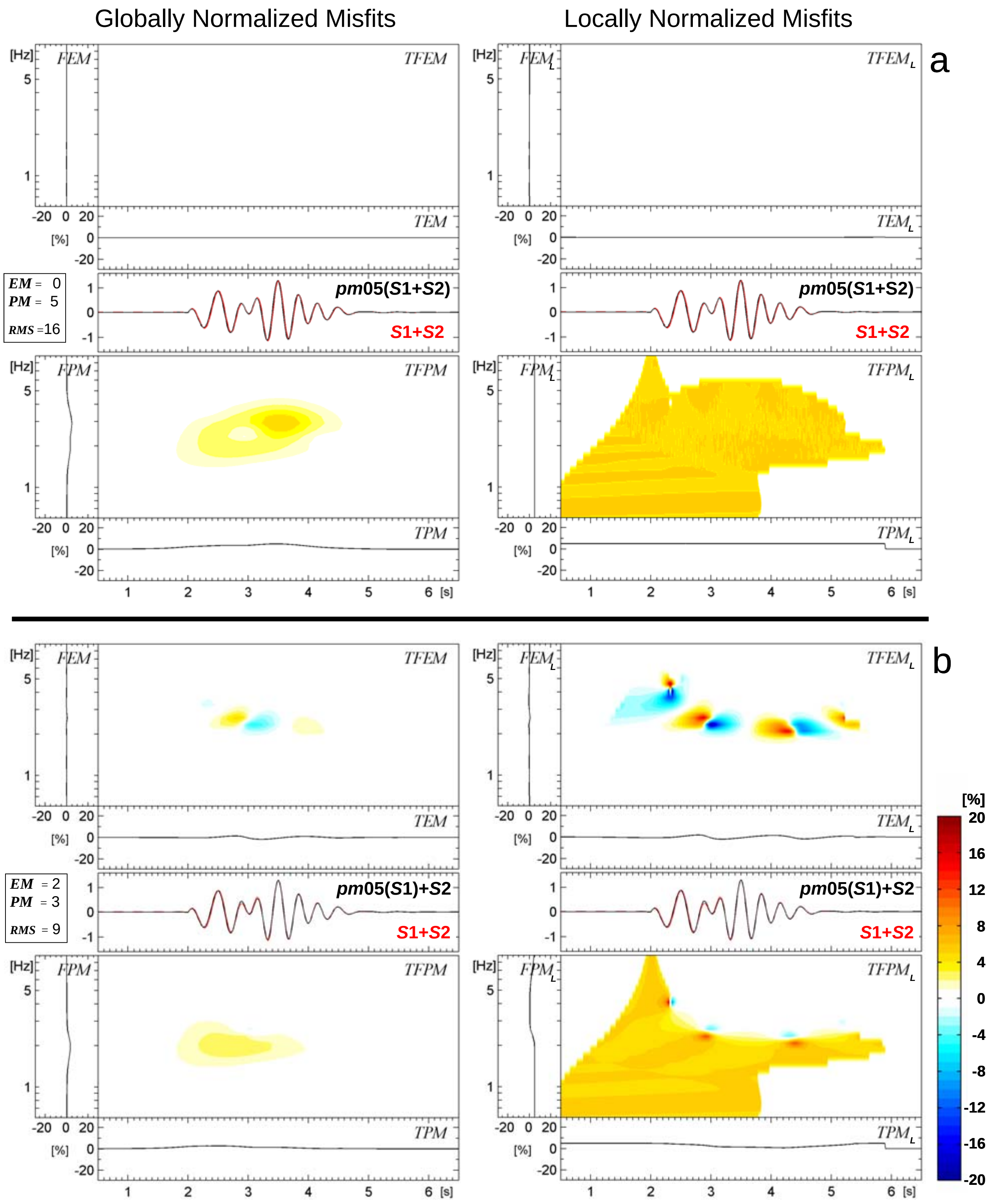

Figure 4. (a) Misfits between the reference signal $S1+S2$ and modified signal $pm05(S1+S2)$. Middle row: reference and amplitude-modified signals, values of the single-valued envelope misfit EM, phase misfit PM, and RMS misfit. Upper row: Time-frequency envelope misfits $TFEM(t,f)$ and $TFEM_L(t,f)$, time envelope misfits $TEM(t)$ and $TEM_L(t)$, and frequency envelope misfits $FEM(f)$ and $FEM_L(f)$. Bottom row: Time-frequency phase misfits $TFPM(t,f)$ and $TFPM_L(t,f)$, time phase misfits $TPM(t)$ and $TPM_L(t)$, and frequency phase misfits $FPM(f)$ and $FPM_L(f)$. (b) The same for $S1+S2$ and modified signal $pm05(S1)+S2$.

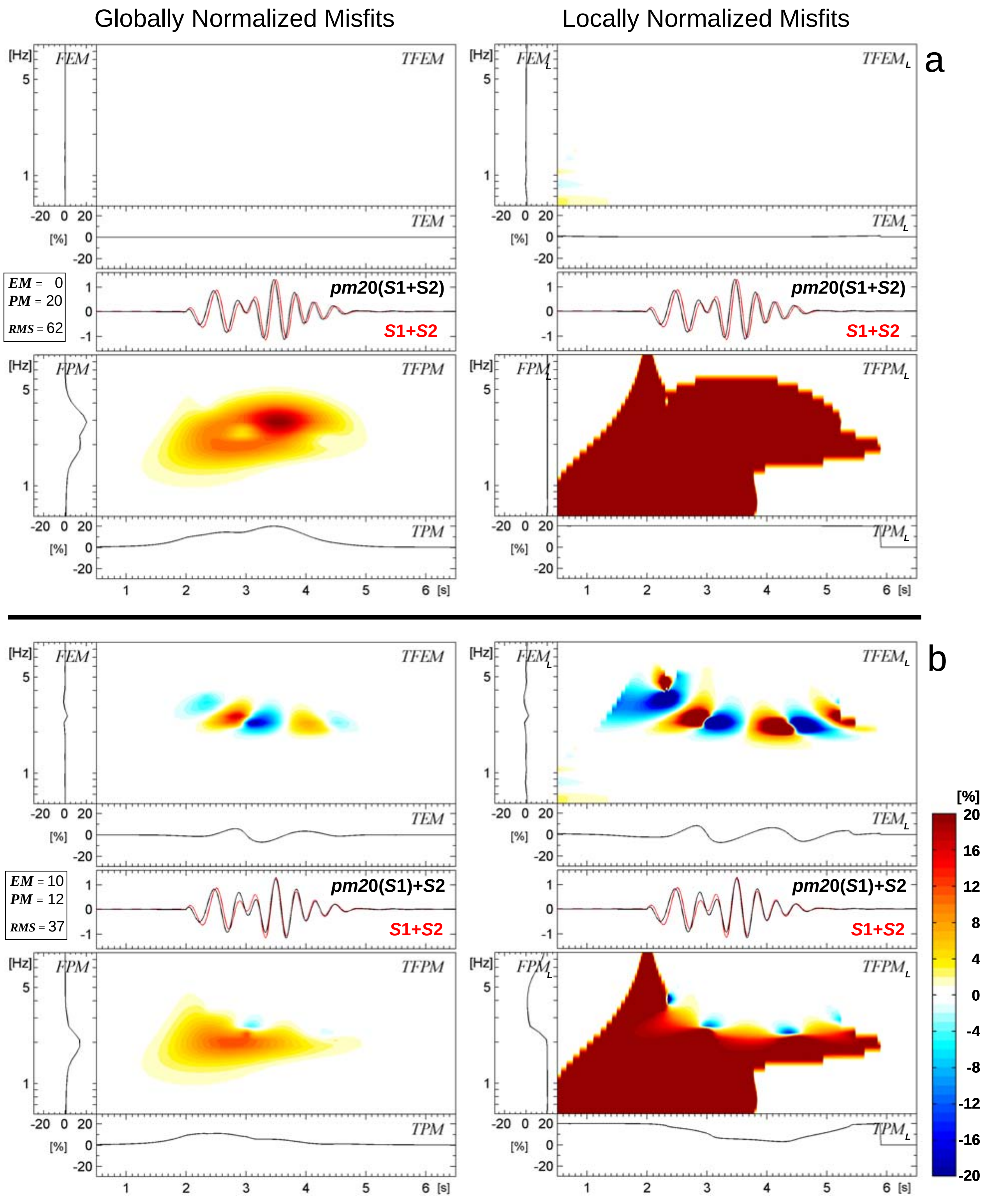

Figure 5. (a) Misfits between the reference signal $S1{+}S2$ and modified signal $pm20(S1+S2)$. Middle row: reference and amplitude-modified signals, values of the single-valued envelope misfit EM, phase misfit PM, and RMS misfit. Upper row: Time-frequency envelope misfits $TFEM(t,f)$ and $TFEM_L(t,f)$, time envelope misfits $TEM(t)$ and $TEM_L(t)$, and frequency envelope misfits $FEM(f)$ and $FEM_L(f)$. Bottom row: Time-frequency phase misfits $TFPM(t,f)$ and $TFPM_L(t,f)$, time phase misfits $TPM(t)$ and $TPM_L(t)$, and frequency phase misfits $FPM(f)$ and $FPM_L(f)$. (b) The same for $S1{+}S2$ and modified signal $pm20(S1)+S2$.

of $TFEM(t, f)$ and $TFEM_L(t, f)$ along the contact of the modified $S1$ and non-modified $S2$, as well as nonzero values of $TFPM(t, f)$ and $TFPM_L(t, f)$ along the contact, are due to the fact that 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.

The RMS misfits equal the single-valued envelope misfit EM in all above cases.

Fig. 4a shows the misfits between the reference signal $S1$+$S2$ and 5% phase-shift modified signal $pm05(S1+S2)$. Similarly, Fig. 5a shows the misfits for $pm20(S1+S2)$. The shape of the area with nonzero values of $TFPM(t, f)$ and $TFPM_L(t, f)$ in the (t, f)-plane corresponds to the shape of the area with nonzero values of TFR of the reference signal. The maximum of $TFPM(t, f)$ equals the percentage of the phase-shift modification in both cases (5% and 20%). Its position in the (t, f)-plane corresponds to the position of the maximum value of the TFR. The statements on the maximum 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 equals the percentage of the phase-shift modification in both cases. The locally normalized phase misfits are constant over the area with nonzero values of TFR of the reference signal. All the envelope misfits are zero. This correctly reflects the absence of amplitude modification.

Fig. 4b shows the misfits between the reference signal $S1$+$S2$ and modified signal $pm05(S1)+S2$ in which only the $S1$-component was 5% phase-shift modified. Similarly, Fig. 5b shows the misfits for $pm20(S1)+S2$. The shape of the area with nonzero values of $TFPM(t, f)$ and $TFPM_L(t, f)$ in the (t, f)-plane correctly corresponds to the shape of the area with nonzero values of TFR of the $S1$ component (see Fig. 1). The maxima of the phase misfits are proportional to the level of the phase-shift modification. They are not equal to the level because the $S1$-component contributes to the TFR less than the $S2$ component, see Fig. 1.

Similarly to the case of $am05(S1)+S2$ and $am20(S1)+S2$, $TFPM(t, f)$ and $TFEM(t, f)$ show alternating negative and positive values along the contact of the modified $S1$ and non-modified $S2$. Compared to the nonzero alternating values of the phase misfits in the case of $am05(S1)+S2$ and $am20(S1)+S2$, the absolute values of the alternating-sign envelope 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.

In all cases of the phase-shift modifications, the RMS misfit is approximately three times larger than the single-valued phase misfit PM. 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.

5. Illustration of TF Misfits in the Case of Synthetics for the Grenoble Valley

Within the Numerical Benchmark of 3D Ground Motion Simulation for the Valley of Grenoble, French Alps, we simulated ground motion in the viscoelastic model, as required by organizers. We also simulated motion in a perfectly elastic model in order to see the effect of attenuation. Figure 6 shows globally normalized misfits between synthetics for the viscoelastic and elastic (reference) models. Though Q-values are relatively large, the effect of attenuation is considerable. Blue color in the upper panel characterizes decrease

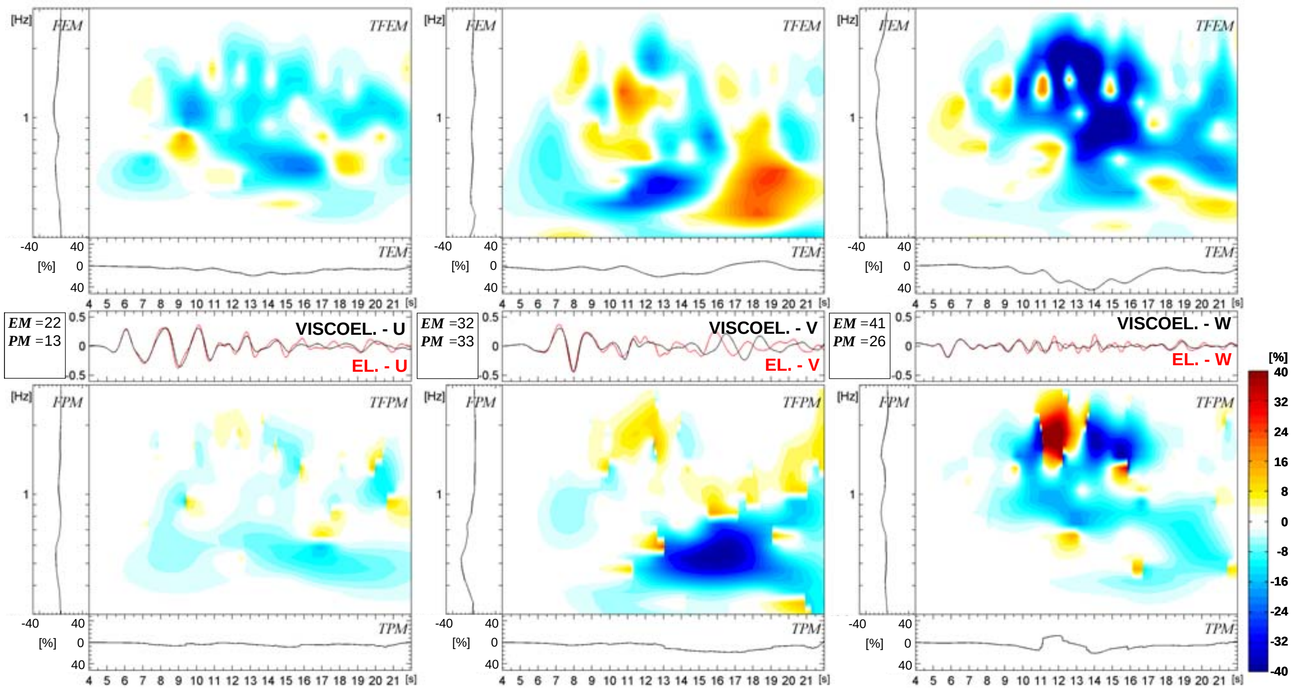

Figure 6. Misfits between synthetics simulated at R28-receiver for an elastic (reference) and viscoelastic models in the ESG2006 Grenoble Valley benchmark, the S1 case

of amplitudes. Orange color at later times of the V-component characterizes amplitude change due to time shift caused by velocity dispersion.

6. Conclusions

Numerical tests demonstrated that in all cases the developed misfit criteria properly quantified and characterized the misfits between reference and modified signals. Usual RMS misfit properly quantifies misfit only in the case of pure amplitude modification.

7. Acknowledgement

This work was supported by the EU Marie Curie research program SPICE (Seismic Wave Propagation and Imaging in Complex media: a European Network).

8. References

Daubechies, I. *Ten Lectures on Wavelets*, SIAM, Philadelphia, (1992)

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.

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

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, *Eos Trans. AGU* 86(52), Fall Meet. Suppl., Abstract S13A-0179.

Kristekova, M., J. Kristek, P. Moczo, and S. M. Day (2006). Misfit Criteria for Quantitative Comparison of Seismograms. *Bull. Seism. Soc. Am.*, in press.

Moczo, P., J. P. Ampuero, J. Kristek, M. Gális, S. M. Day, and H. Igel (2005). The European Network SPICE Code Validation, *Eos Trans. AGU* 86(52), Fall Meet. Suppl., Abstract S13A-0180.