Development of a multi-phase dynamic ray-tracing code

T.A. Stabile[1], R. De Matteis[2], A. Zollo[1]

[1] *Dipartimento di Scienze Fisiche, Università degli Studi di Napoli Federico II (RISSC-Lab), Napoli, Italy*

[2] *Dipartimento di Studi Geologici ed Ambientali, Università degli Studi del Sannio, Benevento, Italy*

Abstract: We here propose a method for rapid, high-frequency seismogram generation that makes use of an algorithm to automatically generate an exhaustive set of seismic phases that produce an appreciable amplitude on the seismogram. The method uses a hierarchical order of rays and seismic phases generation, taking into account some existence constraints for a ray-path and some physical constraints. To compute synthetic seismograms, the COMRAD code (from the Italian: "COdice Multifase per il RAy-tracing Dinamico") uses as its core a dynamic ray-tracing code. To validate the code, we have computed in a layered medium synthetic seismograms using both COMRAD and a code which computes the complete wavefield by the discrete wavenumber method. The seismograms are compared according to a time-frequency misfit criteria based on the continuous wavelet transform of the signals. The comparison shows that the ray-theory seismogram is enough complete and moreover, the time for the computing of the synthetics using the COMRAD code (truncating the ray series at the 10^{th} generation) is 3-4-fold less than that needed for the Axitra code (to a frequency of 25 Hz).

Introduction

Over the last decades, the calculation of synthetic seismograms has become a useful tool in seismological research, and a wide variety of techniques have been developed. Forward modelling, which is the generation of synthetic seismograms, represent an important part of many seismological studies, such as seismic tomography or the kinematic inversion of source parameters. The solution of an inverse problem requires the repeated solving of the forward problem, so that speed is the most stringent condition that the method must supply.

It is possible to obtain the solution of the wave equation in a short time by using approximate high-frequency methods (Cerveny, 2001). The final solution of the elastodynamic equation is composed of elementary body waves that correspond to the various rays connecting the source to the receiver. Although the computation of ray synthetic seismograms are only approxima-

te and the ray method can fail in certain situations, the high-frequency methods are preferable to direct numerical methods for many applications, both for the shorter computing times and for the full interpretability of the seismograms. The problems arise when we want to use a synthetic seismogram as similar to the real one as possible in the high-frequency approximation, such that the seismogram should be relatively complete, although it is not necessary that it contains every feature of the full elastic wave field. It is becoming necessary to select from all of the rays connecting the source to the receiver only those that produce an appreciable amplitude on the seismogram. The problem of generating a comprehensive set of rays was approached for the first time by Hron (Hron, 1971, 1972, Hron et al., 1986) for layered media, by grouping individual rays into families of kinematic equivalents. Afterwards, Clarke (1993 a, b) developed a technique for computing synthetic seismograms based on a ray-generation algorithm that involved the symbolic manipulation of complete wavefield expressions from reflectivity theory, which were truncated to produce a finite ray series.
In this study, we propose a new technique for the rapid definition of an exhaustive set of rays that is based on the hierarchic generation of strings that describe the ray paths and the phase types. The string generation is subjected to physical constraints that are related to the propagation medium and the source-receiver geometry. The ray sets will represent the input of a kinematic or dynamic ray-tracing algorithm (i.e. Cerveny and Hron, 1980; Farra and Madariaga, 1987; Virieux, 1991; Snieder and Spencer, 1993). In particular, the technique developed has been implemented in the multiphase dynamic ray-tracing code (COMRAD) that uses as its core the dynamic ray-tracing code provided by Farra (1987).

Method

Our goal is to carry out an algorithm that rapidly generates an exhaustive number of seismic-phases to calculate an high frequency seismogram as complete as possible. The amplitudes, the raypaths and the travel times of the seismic-phases are computed by the dynamic ray-tracing code provided by Farra and Madariaga (1987). In this section we describe our method and the discretization of the propagation model.
It is possible to use an arbitrary medium as long as it is discretized by M ordered elements between a free surface and an half space. Each element should have an arbitrary shape and it is characterized by the following properties:
– Vp_i: compressive wave velocity (P wave);
– Vs_i: shear wave velocity (S wave);
– ρ_i: density;
– Qp_i: P-wave quality factor (optional);
– Qs_i: S-wave quality factor (optional);

- T_i: top surface of the element;
- B_i: bottom surface of the element (= T_{i+1}: top surface of the next element);
- i=1,...,M: index of the element.

where the elastic wave velocities, densities and quality factors should be functions of spatial coordinates.
Both the source and the receivers take the index i of the element in which they are included. Each ray goes from the source to the receiver and is divided into L curves if it passes through L elements. The numerical string of a ray is composed of L numbers, and each one is the index of the element crossed by the ray. For a ray there are also 2^L phase numerical strings, because in each element crossed by the ray the wave can have two different polarization (P wave or S wave). A numerical string for a phase is composed of L numbers and each one can be 1 (to indicate a P wave) or 2 (to indicate an S wave).
An example of how the numerical strings for the direct ray should be and for all of its possible phases is shown in Figure 1. Here, the elements of the medium are horizontal, parallel layers, while the source and the receiver are inside the third and first layers (elements), respectively.

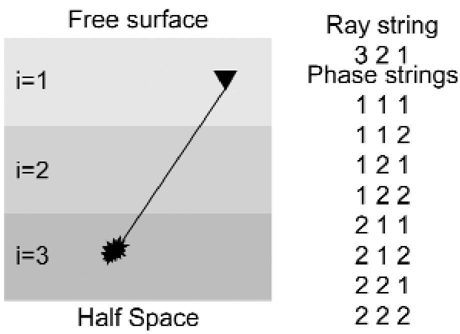

Fig. 1. Example of the ray and phase strings for a medium with three elements, where the source is in the third element and the receiver is in the first element.

Once fixed the length L of the numerical ray strings there are M^L ray strings of the same length if the medium is discretized by M elements. However, not all of the ray strings describe a real ray, because there are some existence constraints for a ray path:

a) The ray starts at the source;
b) The ray arrives at the receiver;
c) The ray path is bordered by the free surface and the half space;
d) If a ray crosses the top surface of the i^{th} elements, it must go to the $(i-1)^{th}$ element;

e) If a ray crosses the bottom surface of the i^{th} elements, it must go to the $(i+1)^{th}$ element;
f) If the receiver is on the free surface, the ray must come from below;
g) If the source is on the free surface, the starting direction of each ray cannot be up.

To generate only the ray strings that describe real ray paths, we use a hierarchic order of ray generation, where the root is the index of the element that contains the source, and the number of the generations is the length of the numerical strings of the rays. The hierarchic order of ray generation satisfies itself for the a, b, c, d and e constraints. Using the additional f and g constraints, we discard some branches inside the numerical string tree. An example of hierarchic ray generation taken to the fourth generation for the same case as Figure 1, although putting the receiver on the free surface, is shown in Figure 2. In this case, we have two ray strings (3-2-1 and 3-3-2-1) to the fourth generation. If we put the receiver inside the first element instead of on the free surface, we will have the additional ray string 3-2-1-1 because there is no f constraint yet, and consequently the ray can arrive at the receiver also from above.

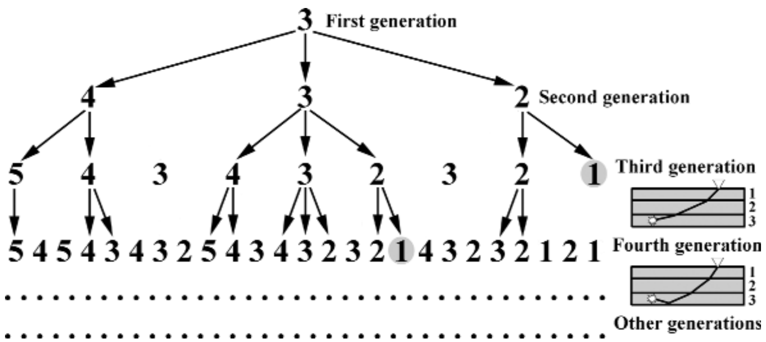

Fig. 2. Hierarchical ray strings generation using all of the a to g constraints. There are only two ray strings until the fourth generation because the receiver is on the free surface. If the receiver is inside the first element, there will also be the string 3-2-1-1.

Since the number of phase strings becomes greater and greater if the number of generations grows up, and consequently the computing time rises, it is possible to introduce some propagation constraints in order to reduce the number of phases. For this reason, we created two additional constraints: the first is on the maximum number of reflections in each element (selected by the RIFMULT parameter), and the second is a phase selection based on their expected amplitude values (selected by the PHS parameter).

METHOD VALIDATION

The higher the number of generations we consider, the higher the number of phases the core must calculate. To make the best choice on the number of generations, we have to compare the calculation time with the complexity we need for the synthetics. We compute synthetic seismograms for two receivers respectively at 1 km and 30 km distance from the epicentre of an explosive source. We use a crustal velocity model (Bernard and Zollo, 1989), where the elements are considered as horizontal parallel layers. The model is described in Table 1, while the results of the simulations are given in Table 2.

Tab. 1. The crustal velocity model used for our simulations.

Interfaces	Depth (km)	Vp (km/s)	Vs (km/s)	ρ (g/cm^3)
1	0	2.30	1.33	2.2
2	3	5.30	3.06	2.3
3	7	6.00	3.46	2.4
4	10	6.28	3.63	2.6
5	20	6.54	3.78	2.8

The source depth is 4 km, and both of the receivers are at 1 m in depth. For this source-receiver geometry, the source is in the second layer, the receivers are in the first layer and the first real phases start at the second generation. In Table 2, for each generation, the time needed by both the COMRAD and core programmes for the calculations can be seen, along with the total number of phases and the RMS. This has been calculated between the synthetic seismogram S_{REF} obtained at the previous generation and the seismogram S obtained for that generation, for both of the receivers and for both of the X and Z components. The amplitudes are normalized with respect to their maximum amplitudes. The RMS is defined as follows:

$$RMS = \sqrt{\frac{\sum_t \left|S(t) - S_{REF}(t)\right|^2}{\sum_t \left|S_{REF}(t)\right|^2}} \qquad (1)$$

where t is the time.
From Table 2, it is clear that the RMS is lower and lower as a function of the number of generations. Moreover, starting from the RMS calculated between the 9th and 10th generations, the RMS decrease rate at 30 km distance becomes

Tab. 2. Results obtained for when the source is in the second layer and the receivers are in the first layer. The RMS is calculated between the seismograms calculated at the i-1th and ith generation.

N. GEN.	COMRAD TIME	CORE TIME	N. PHASES	RMS CompX 1 km	RMS CompZ 1 km	RMS CompX 30 km	RMS CompZ 30 km
2	0m 00.05s	0m 00.06s	2	-	-	-	-
3	0m 00.05s	0m 00.15s	10	1.037	0.992	4.628	4.789
4	0m 00.05s	0m 00.29s	34	0.108	0.104	0.980	0.932
5	0m 00.05s	0m 01.13s	114	0.107	0.103	0.262	0.340
6	0m 00.05s	0m 04.42s	370	0.027	0.144	0.440	0.494
7	0m 00.05s	0m 16.84s	1266	0.026	0.028	0.398	0.399
8	0m 00.05s	1m 07.45s	4210	0.008	0.010	0.186	0.252
9	0m 00.11s	4m 18.82s	14706	0.008	0.010	0.131	0.220
10	0m 00.27s	16m 18.15s	49522	0.002	0.004	0.087	0.109
11	0m 01.18s	1h 00m 23.47s	174450	0.002	0.004	0.077	0.092
12	0m 03.39s	3h 40m 56.38s	590194	0.001	0.002	0.037	0.058

stable for both the X and Z components. This means that if we want both complete body wave seismograms and short computing times, we can truncate the ray series at the 9th or almost at the 10th generation for this study. We can also reduce the number of phases using the RIFMULT and the PHS parameters discussed in the previous paragraph.

To better understand if our choice is appropriate, we can calculate the complete wavefield for the case under study by the Axitra programme, which uses the discrete wavenumber method developed by Bouchon (Bouchon, 1981; Coutant, 1989). Afterwards, we can compare the Axitra results with those obtained using the COMRAD programme stopped at the 10th generation, where the amplitudes of the synthetics are expressed in velocity and are normalized with respect to their maximum amplitude. The Axitra computation time (up to the frequency of 25 Hz) is 53 minutes and 41.6 seconds, 3-4-fold longer than the COMRAD+core computation time (see Table 2). For a quantitative comparison between the synthetics, we use the misfit criteria in time and frequency that was developed by Kristekova et al. (2006), calculating both the time frequency envelope misfit (TFEM) and the time frequency phase misfit (TFPM). The TFEM is calculated according to the following equation:

$$TFEM(t,f) = \frac{\Delta E(t,f)}{\max_{t,f}(|W_{REF}(t,f)|)} \tag{2}$$

while the TFPM is calculated according to equation (3):

$$TFPM(t,f) = \frac{\Delta P(t,f)}{\max_{t,f}(|W_{REF}(t,f)|)} \tag{3}$$

where t is the time, f is the frequency, $\Delta E(t,f)$ is the local time-frequency envelope difference, $\Delta P(t,f)$ is the local time-frequency phase difference, and $W_{REF}(t,f)$ is the time-frequency representation of the reference signal $S_{REF}(t)$ based on the continuous wavelet transform.

In Figure 3, the TFEM and the TFPM for the Z component of the receiver at 30 km distance (as worst case scenario) are shown. We also overlap the two synthetic seismograms in Figure 4, cutting the synthetics from six to 20 seconds (with respect to the origin time).

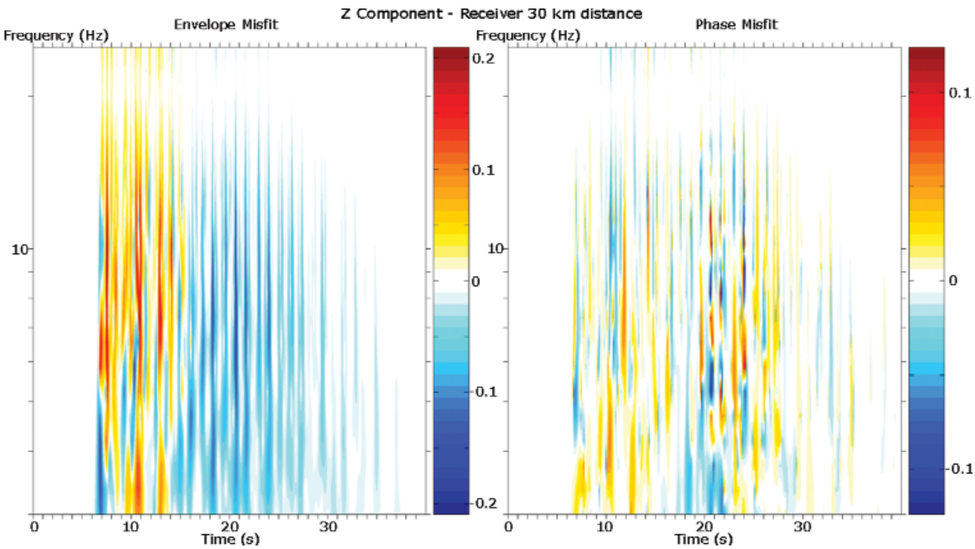

Fig. 3. The TFEM (left) and TFPM (right) plots for the vertical components of a receiver at 30 km distance from the source.

From the TFEM plot in Figure 3, and looking also at Figure 4, we can see that there are differences in the amplitudes for some of the seismic phases, but only for two of them (8 s, 11 s) there is an envelope misfit of about 20%. Moreover, from the TFPM plot in Figure 3, and again looking also at Figure 4, we can see that the phase misfit is less than 5% until 17 s, after which time it increases to about 10% because of the coda waves, as calculated by Axitra and not by COMRAD due to the truncation of the ray-series.

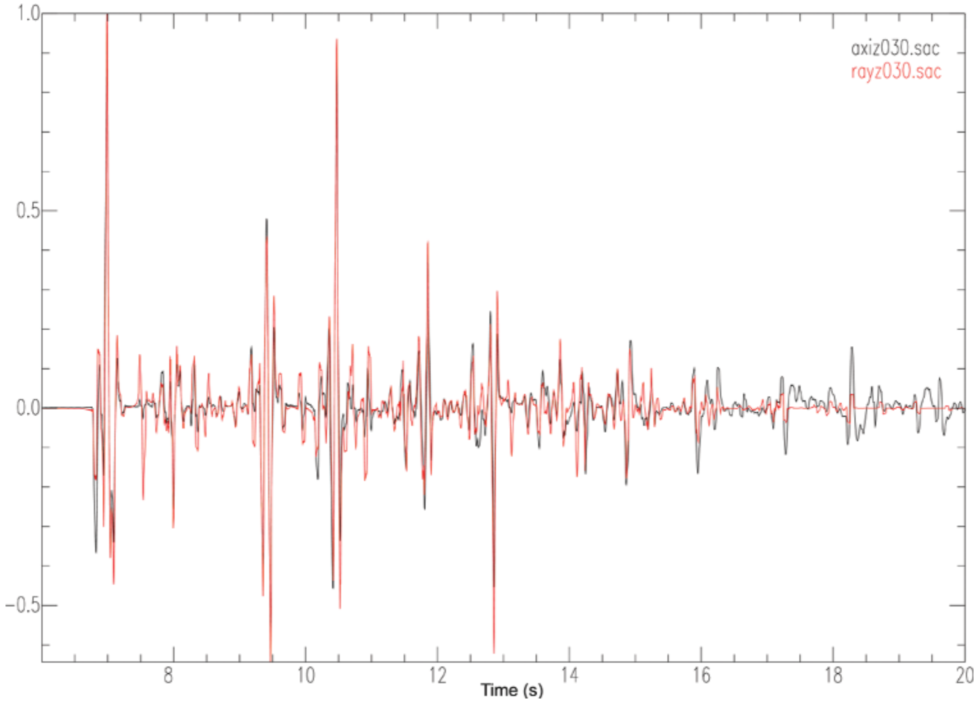

Fig. 4. The two synthetic seismograms for velocity computed by Axitra (in black) and COMRAD (in red). The amplitudes are normalized with respect to the maximum amplitude of each seismogram.

Structure of the Comrad.f code

The Comrad.f code is the Fortran77 version of multiphase code that uses as its core the dynamic ray-tracing code developed by Farra (Farra and Madariaga, 1987) to compute synthetic seismograms, travel times and ray paths. It is accompanied by the bash script multiphase.sh, which optimizes the whole calculation process and runs both the COMRAD and core programmes.
The structure of the code is represented in Figure 5, where we outline its most important features. The code requires the following input data:

– a file (creation.inp) which contains the medium information;
– a file (xxxx.dis, where xxxx represents a four characters word) which contains the source-receiver geometry;
– the maximum number of generations (Lmax), which truncates the ray series. The user can choose to use the default value (Lmax=Lmin+2*m);
– the multiple reflections parameter (RIFM), which allows the ray to have a finite number of reflections in each element. The user can choose to use the default value (RIFM=8);

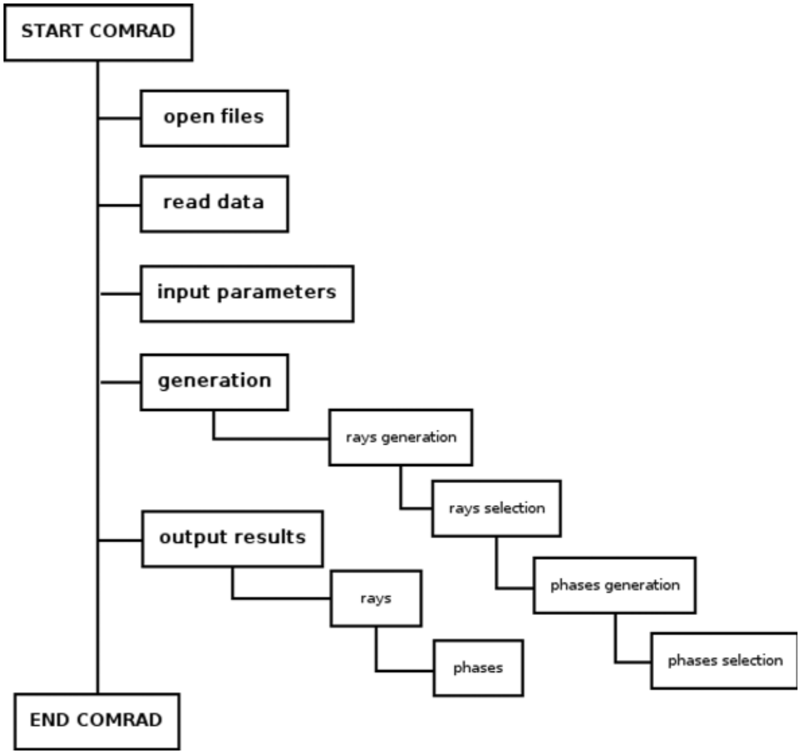

Fig. 5. Block scheme of the comrad.f computer code.

- the phase selection parameter (PHS), which discards phases with negligible amplitudes. The user can choose to use the default value (calculated on the approximated amplitude value of the direct P wave and fixed at 0.1% of this value);
- the initial polarization of the phases (P), which allows the user to use an explosive source (only P waves from the source) or a non-explosive source (both P and S waves from the source).

Using the hierarchical generation method with the constraints described previously, the code generates the ray and phase strings and creates the following output data:
- a file (kernel.inp) which contains the input parameters for the core;
- two files (xxx1.dis and xxx2.dis, where xxx represents a three character word) inside which the comrad.f code will write all of the ray and phase strings to be used by the core programme for the computation of the synthetics. The first file describes rays where their starting direction is up, and the second, where it is down.

The existence of the two xxx1.dis and xxx2.dis files assures that the calculations of both the P/P\ and P\P/ phases (and similar) are described by the same ray and phase strings if the source and the receiver are both in the same element (i.e. the ray string is 1-1 both for P/P\ and P\P/ phases if the source and the receiver are in the first element of the medium). Moreover, if the source is on the free surface, the xxx1.dis does not exist because of the g constraint.

Conclusions

We have developed a method for rapid high frequency seismogram generation. The method is used in a code (Comrad.f) to compute high frequency synthetic seismograms, which uses as its core the dynamic ray-tracing code developed by Farra and Madariaga (1987).
We have numerically tested the results in two steps. First, generation by generation, we evaluated both the computing time and the RMS between the COMRAD synthetics obtained in each generation and those obtained in the previous generation, to understand what the best choice is to have both an exhaustive seismogram and a short computing time. Using the crustal velocity model described in Table 1 and an explosive source at 4 km in depth, we found that the best choice is to stop the ray series at the 10^{th} generation. After fixing the number of generations, we compared the COMRAD-derived synthetic seismograms with Axitra-derived synthetic seismograms, the latter being a programme that calculates the complete wavefield. We computed the seismograms in velocity and we normalized the amplitudes with respect to the maximum amplitudes. The numerical tests were carried out by the quantitative misfit criteria developed by Kristekova (2006).
The comparison of seismograms computed by the COMRAD and Axitra programmes is good, in particular if the source-receiver distance is short. In this study we have shown the misfits for a receiver 30 km distant from the source, where we have more differences. We demonstrate that the envelope misfit is about 20% only for the 8 s and 11 s seismic phases, due to the differences in amplitudes. Although there are also amplitude differences for some other phases, the envelope misfit is less than 10%-15%. Moreover, for the receiver at 30 km distance, the coda waves are not so well computed by the COMRAD code, as is clear from the value of the phase misfit, which is about 10% when there are coda waves computed by the Axitra code and not by the COMRAD code. Finally, we compared the computation times for both of the programmes, and we can be sure that COMRAD (with 10 generations) is 3-4-fold faster than Axitra (up to a frequency of 25 Hz). If we want to compute the synthetics at higher frequencies, the Axitra code needs more time, while the computation time of the COMRAD code remains the same.

References

Bernard, P. and Zollo A., 1989. The Irpinia (Italy) 1980 Earthquake: detailed analysis of a Complex Normal Faulting. J. Geophys. Res., 94, No. B2, 1631-1647.

Bouchon, M., 1981. A simple method to calculate Green's functions for elastic layered media. Bull. Seism. Soc. Am., 67, 259-277.

Cerveny, V., 2001. Seismic Ray Theory, Cambridge University Press, Cambridge.

Cerveny, V., and Hron, F., 1980. The ray series method and dynamic ray-tracing system for three-dimensional inhomogeneous media. Bull. Seis. Soc. Am., 70, 47-77.

Clarke, T.J., 1993. The complete ordered ray expansion -I. Calculation of synthetic seismograms. Geophys. J. Int., 115, 421-434.

Clarke, T.J., 1993. The complete ordered ray expansion -II.Multiphase body wave tomography. Geophys. J. Int., 115, 435-444.

Coutant, O., 1989. Program of Numerical Simulation AXITRA. Res. Report LGIT, Grenoble, in French.

Farra, V. and R. Madariaga, 1987. Seismic waveform modeling in heterogeneous media by ray perturbation Theory. J. Geophys. Res., 92, No. B3, 2697-2712

Hron, F., 1971. Criteria for selection of phases in synthetic seismograms for layered media. Bull. Seism. Soc. Am., 61, No.3, 765-779

Hron, F., May B.T., Covey J.D. and Daley P.F., 1986. Synthetic seismic section for acoustic, elastic, anisotropic inhomogeneous layered media, Geophysics, 51, 710-735

Hron, F., 1972. Numerical methods of ray generation in multilayered media, in Methods in Computational Physics, 12, Academic Press, New York.

Kristekova, M., Kristek J., Moczo P. and Day S.M., 2006. Misfit criteria for quantitative comparison of seismograms. Bull. Seism. Soc. Am., 96, No. 5,1836-1850

Snieder, R., and Spencer, C., 1993. A unified approach to ray bending, ray perturbation, paraxial ray theories. Geophys. J. Internat., 115, 456-470.

Virieux, J., 1991. A fast and accurate ray tracing by Hamiltonian perturbation, J. Geophys. Res., 96, 579-594.