Crustal and upper mantle structures of Makran subduction zone, SE Iran by combined surface wave velocity analysis and gravity modeling

Somayeh Abdollahia,⁎, Vahid Ebrahimzade Ardestani, Hermann Zeyen, Zaher Hossein Shomalia

a Institute of Geophysics, University of Tehran, P.O. Box 14155-6466, Tehran, Iran
b GEOPS, Univ. Paris-Sud, CNRS, Université Paris-Saclay, 91405 Orsay, France
c Department of Earth Sciences, Uppsala University, Uppsala, Sweden

ABSTRACT

The inversion of Rayleigh wave group velocity dispersion curves is challenging, because it is non-linear and multimodal. In this study, we develop and test a new Rayleigh wave dispersion curve inversion scheme using the Shuffled Complex Evolution (SCE) algorithm. Incorporating this optimization algorithm into the inverse procedure not only can effectively locate the promising areas in the solution space for a global minimum but also avoids its wandering near the global minimum in the final stage of search. In addition, our approach differs from others in the model parameterization: Instead of subdividing the model into a large number of thin layers, we invert for thickness, velocities and densities and their vertical gradients of four layers, sediments, upper-crust, lower-crust and upper mantle. The proposed inverse procedure is applied to non-linear inversion of fundamental mode Rayleigh wave group dispersion curves for shear and compressional wave velocities. At first, to determine the efficiency and stability of the SCE method, two noise-free and two noisy synthetic data sets are inverted. Then real data for Makran region in SE Iran are inverted to examine the usage and robustness of the proposed approach on real surface wave data. In a second step, we applied 3D Gravity Modeling based on surface wave analysis results to obtain the density structure and thickness of each layer. The reason for using both types of data sets, is that gravity anomaly has a bad vertical resolution and surface wave group velocities are good for placing layer limits at depth, but they are not very sensitive to densities. Therefore, using gravity data increases the overall resolution of density distribution. In a final step, we used again the SCE method to invert the fundamental mode Rayleigh wave group dispersion curves based on the gravity results. Gravity results like thicknesses and sediment densities have been used to constrain the limit of search space in the SCE method. Results show a high shear and compressional velocity under the Gulf of Oman which reduce to the North of the Makran region. The Moho depth of the Oman Gulf is about 18–28 km and it increases to 46–48 km under the Taftan-Bazman volcanic-arc. The density image shows an average crustal density with maximum values under the Gulf of Oman decreasing northward to the Makran.

1. Introduction

The aim of this study is to calculate the seismic wave velocity variations and density structure of the crust and upper mantle in the Makran area, SE Iran. Our data are Rayleigh wave group velocity dispersion curves and gravity anomalies. In recent surveys, Rayleigh wave group velocities have been used as a tool to estimate shear and compressional wave velocities and density variation at different depths to characterize the crust and upper mantle structure. However, as is the case for most other geophysical optimization problems, inversion of Rayleigh wave dispersion curves, is typically a highly non-linear, multi-parameter, and multimodal inversion problem (Julia et al., 2000). Local optimization methods like matrix inversion or conjugate gradients, are prone to being trapped by local minima, and their success depends heavily on the choice of a good initial model and the accuracy of the partial derivatives. Therefore, global optimization methods that can

⁎ Corresponding author.
E-mail address: sabdollahi@ut.ac.ir (S. Abdollahi).

First author, Abdollahi made substantial contributions to conception, acquisition of data, analysis and interpretation of data as her PhD thesis.

All other authors made substantial contributions to conception, analysis and interpretation of data as well; participated in drafting the manuscript and revising it critically and also give final approval of the version to be submitted and any revised version.

https://doi.org/10.1016/j.tecto.2018.10.005
Received 15 May 2018; Received in revised form 28 September 2018; Accepted 4 October 2018
Available online 06 October 2018
0040-1951/ © 2018 Elsevier B.V. All rights reserved.
overcome this limitation are particularly interesting for surface wave analysis (Song et al., 2012). The SCE algorithm is a global optimization strategy well capable of dealing with problems characterized by a large numbers of local minima and/or maxima, without the need of calculating any gradient or partial derivative information, especially for addressing problems for which the objective functions are not differentiable, stochastic, or even discontinuous. The SCE method has recently been used and tested to optimize complex mathematical problems of irregular, multi-parameter, and multimodal objective functions (Zhao and Zhang, 2012; Song et al., 2012; Kan et al., 2017). However, few attempts have been made to use the method for real geophysical data, especially for non-linear inversion of surface wave data. In this work, we implemented and tested a new inversion scheme for Rayleigh wave group velocity dispersion curves based on the SCE approach. The calculation efficiency and stability of the proposed inverse procedure are tested on two synthetic models and a real data set in the Makran region, SE Iran. Surface wave dispersion curves are primarily sensitive to seismic shear wave velocities. Theoretically, the dispersion curve is a non-linear function of shear wave velocity, compressional wave velocity and density of the media (Bucher and Smith, 2013). However, it has been proven that the sensitivity to the P-wave velocity is significantly smaller than the sensitivity to the shear wave velocity (Aki and Richards, 2002). Also the sensitivity for density is smaller than the one for the shear wave velocity, which is why shear velocity variations are usually the prime model parameters in inversion studies of surface wave dispersion (Bache et al., 1978; Tanimoto, 1991). On the other hand, gravity data may be used to better constrain the density distribution. They have good lateral resolution, but bad vertical resolution. In this way, surface wave dispersion curves, which are good for placing layer limits at depth, and gravity data, which are sensitive to lateral density variations give complementary information and increase the overall resolution of crustal and upper mantle structures. Therefore, we propose here to carry out a sequential inversion of wave dispersion curves and gravity data. We expect to obtain a three-dimensional density model with well constrained discontinuities in addition to S-wave velocity distribution.

2. Geological and geophysical setting

The 1000 km long Makran mountains (Fig. 1) form the southeastern end of the Arabian–Eurasian Plate boundary, where the oceanic crust of the Arabian plate (Oman seafloor) is subducting northward beneath the Makran belt since the early Cretaceous. The E-W trending Makran is located between two, nearly N-S directed transform fault systems. To the west, the dextral Minab fault separates the Makran subduction zone from the Zagros continent–continent collision zone and to the east, the sinistral Chaman fault system separates it from the Indian continent (Fig. 2). This tectonic setting triggers seismicity and volcanic activity (Byrne et al., 1992; Farhoudi and Karig, 1977). The subduction zone of Makran is segmented into a western and an eastern part with different seismic and tectonic characteristics (Byrne et al., 1992; Zarifi, 2006). In the western part of Makran, the subduction is occurring with a steeper dip than in the eastern part. In the Makran subduction zone, some deep earthquakes (depths up to 150 km) have been observed, showing the position of the north-dipping subducting oceanic crust of the Arabian plate. The convergence rate of the Arabian–Eurasian plates is increasing from west to east in the Makran (Nilforoushan et al., 2003). Compared with the Zagros suture zone, the continental collision part of the Arabian–Eurasian Plate boundary, the Arabian plate subducts at a higher convergence rate in the Makran subduction zone.

Three major volcanic centers in the north of Makran from west to east are the Bazman, Taftan and Sultan volcanoes, respectively, extending from WSW to ENE over a distance of 400–500 km. This volcanic belt is the southeasternmost and youngest manifestation of arc-related volcanism in Iran, which began with the Eocene–Miocene Magmatic Arc in NW and central Iran, and extends into western Pakistan where the existence of fumaroles and sulfataras shows that volcanism is still active. Two major tectonic trends separate the three volcanic centers: the eastern border fault zone of the Lut Block between Bazman and Taftan and the north-eastern border of the eastern Iranian ranges between Taftan and Sultan. The Taftan volcano is located in the Sistan suture zone and its basement rocks include ophiolite, flysch and volcanic rocks (mostly Upper Cretaceous to Eocene). North of Bazman, andesites of Miocene–Eocene age are partly buried beneath the volcano. Carboniferous and Permian metamorphic rocks form the oldest basement at Bazman. The Sultan resembles very much the Taftan volcano with its andesitic compositions and a corresponding strong erosion (Richards, 2014). An Eocene mélange makes up part of the volcano’s basement.

As mentioned above, the level of seismicity in Makran is low and increases from west to east. The western and eastern parts of the Makran subduction zone have different seismic and tectonic characteristics. Recent and active deformation in Sistan is dominated by right-lateral strike-slip and thrust faults related to the indentation of Iran by the Arabian shield. The seismic activity in the Taftan-Bazman volcanic-arc, is also quite weak. In 1979, several right-lateral moderate-sized earthquakes occurred in the Sistan Suture Zone, between the Lut and Helmand blocks. This seismic activity indicates that the Sistan Suture Zone may be defined as the boundary between western and eastern Makran (Byrne et al., 1992). To the east, the distance of the
volcanic-arc and fore-arc setting increases, and this suggests that the slab dip decreases to the East (Byrne et al., 1992; Zarifi, 2006; Manaman et al., 2011). Eastern Makran has experienced most of its seismic activity in its eastern part near the Chaman Fault (Zarifi, 2006).

According to the litho-stratigraphic and deformation evolution, five main E-W striking structural units are separated by major thrust zones in the Makran zone (McCall, 1985; McCall, 2002; Grando and McClay, 2006; Dolati, 2010). These sub-divisions are as follows (Fig. 2):

1) North Makran: Ophiolitic-sedimentary assemblage zone, 2) Inner Makran: Tertiary allochthonous turbiditic assemblage zone, 3) Outer Makran: Neogene autochthonous sediment succession zone, 4) Coastal Makran: Late Miocene-Pliocene sediment succession zone and 5) Offshore Makran: Submarine shale diapiric - imbricate fan system.

The North Makran domain is bounded by the Jazmuran depression which is considered as a fore-arc basin opened at the southern rim of the Lut block as a consequence of the Makran subduction (McCall and Kidd, 1982; Burg et al., 2012). In contrast, to the Inner, Outer and Coastal Makran domains, which resulted from a northward subduction that was established since Eocene times, the North Makran domain preserves remnants of the pre-Eocene geodynamic history (Fig. 2).

Few geophysical studies have been conducted in the Makran subduction zone most of which have been limited to the global tomography surveys with low resolution and to shallow seismic studies to investigate the sedimentary structure of the Makran belt. Seismic data across the onshore part of the Iranian Makran are not available but profiles across the coastal Pakistani Makran (Ellouz-Zimmermann et al., 2007) and the offshore Makran are available (White and Klitgord, 1976; White and Louden, 1982; Fowler et al., 1985; Minshull et al., 1992; Kopp et al., 2000). Wide-angle seismic lines show an oceanic crustal thickness of about 9 km to the south of the trench. The oceanic crust is covered by > 7 km of undeformed sediments at the front of the wedge. The total thickness of deformed and undeformed sediments reaches > 10 km near the coast line (Minshull et al., 1992; Kopp et al., 2000).

Across the western Makran, the Moho map resulting from partitioned waveform inversion indicates that crustal thickness under the Oman seafloor and Makran fore-arc is about 25–30 km (Manaman et al., 2011). Beneath the Makran highlands, the Moho depth is suddenly increasing towards the north, which indicates the starting location of underthrusting of the Arabian oceanic crust under the Iranian plateau. The Moho deepening continues to reach its maximum value of almost 48–50 km, where the subducting plate bends below the Taftan-Bazman volcanic-arc (Manaman et al., 2011). The deepest part of the Moho is found around the Taftan volcano (Fig. 1). In eastern Makran, the crustal thickness is also increasing from the fore-arc setting to the volcanic-arc, where the maximum Moho depth is about 40 km. The Moho map clearly depicts the western edge of the Makran subduction zone, where the Minab fault marks the boundary between the thick continental crust of the Arabian Plate and the thin oceanic crust of Oman sea (Manaman et al., 2011). Also, Molinari and Morelli (2011) found that the average range of the Moho depth increases in the Makran region from S to N from 25 to 45 km. The shear-wave velocity images of the upper mantle across the Makran subduction zone depict a high-velocity anomaly under the Oman seafloor, which is subducting under the entire zone of Makran belt (Manaman et al., 2011). These authors show that the high-velocity slab of the Arabian plate subducts northward beneath the low-velocity overriding lithosphere of the Lut block in the western Makran and of the Helmand block in the eastern Makran.
3. Data

Our dispersion velocity data set consists of fundamental-mode Rayleigh wave group velocities at periods of 16 s, 20 s, 24 s, 30 s and 40 s for each point on a grid of 1° by 1° in Makran between 53–66°E and 23–30°N (Fig. 3). These dispersion observations were taken from a surface wave tomographic study in the Makran region (Abdetedal et al., 2015) in which 12 months of continuous ambient noise data from January 2009 through January 2010, recorded at broadband seismic stations, were analyzed. Group velocities of the fundamental mode Rayleigh wave dispersion curves were obtained from the empirical Green’s functions. Multiple-filter analysis was used to obtain group velocity variations at periods from 10 to 50 s. The authors indicate that the 16 s map, are sensitive down to 10 km depth approximately, and 30 s and 40 s maps are most sensitive down to a depth of 60 to 80 km. Thus, the use of the dispersion observations from the periods of 16 s to 40 s allows us to constrain the shear velocities at the depth range of 10 to 80 km (Abdetedal et al., 2015). These considerations are based on sensitivity kernels of group velocity as a function of period for the shear velocity measurements. The checkerboard test results of the observed data and the ray-path coverage for periods of 16, 20, 30, and 40 s suggest that the resolution is fairly good for most periods in central Iran and indicates that the pattern and absolute amplitude values were well recovered. However, for the easternmost and southernmost areas of the region, due to a smaller station density, the path coverage is not dense enough, and since most waves travel in parallel, the resolution is rather low, and the inversion process does not recover well the absolute amplitudes in these areas.

The gravity data used for inversion come from freely accessible global free-air gravity data with a resolution of 2.5 x 2.5 arc-minutes (http://bgi.omp.obs-mip.fr/), which are extracted and then projected onto a rectangular grid in Cartesian coordinates with 20 km by 22 km grid spacing (Fig. 4).

The free-air gravity anomaly over the Taftan-Bazman volcanic-arc and the onshore part of the Makran accretionary prism is positive with values above +100 mGal, reflecting the denser volcanic rocks and the dense subducting slab. The trench is marked by a negative anomaly which is weaker than in usual subduction zones reflecting the fact that the trench is also hardly visible in the topography (Fig. 1), since it is overridden by one of the most impressive accretionary prisms on Earth. The anomaly stays slightly negative as usual over oceanic crust and becomes strongly positive in Oman.

4. Shuffled Complex Evolution (SCE) method

The SCE method is a metaheuristic for global optimization, proposed by Duan et al. (1992). In the SCE method, the population is divided into sub-populations, named complexes. The inversion algorithm is composed of several cycles of population optimization. At every cycle, an internal optimization mechanism, called Competitive Complex Evolution (CCE), makes evolve each complex over several iterations. After each cycle of CCE iterations, the complexes are recombined to
the main population. Then, new segmentation and partitioning creates new complexes and will thus shuffle the population and complexes.

The SCE algorithm, in fact, has been created by combining the evolutionary capabilities of the genetic algorithm and the Controlled Random Search (CRS) (Price, 1977; Price, 1983). So it can be also classified, to some extent, in the category of the Memetic Algorithms. In Memetic algorithms, each individual acquires (according to the Lamarck theory of evolution) attributes and useful features, whereas in genetic algorithms, characters and capabilities, are inherited from parents. In the SCE method, transmission of these properties is done by local search within a complex. This algorithm is based on a synthesis of four concepts that have proved successful for global optimization: 1) Combination of deterministic and probabilistic approaches; 2) Systematic evolution of a “complex” of points spanning the parameter space, in the direction of global improvement; 3) Competitive evolution; and 4) Complex shuffling. The synthesis of these elements makes the SCE method effective and robust, and also flexible and efficient (Price, 1977; Duan et al., 1992; Duan et al., 1993; Sorooshian et al., 1993). The search begins with a randomly selected complex of points spanning the entire feasible space. A large enough number of points will help to ensure that the complex contains information regarding the number, location, and size of the major regions of interest. The implementation of an implicit clustering strategy helps to concentrate the search in the most promising regions identified by the initial complex (Duan et al., 1993).

As mentioned above, the SCE method contains several probabilistic and deterministic components that are controlled by some algorithmic parameters. For the method to perform optimally, these parameters, m, p, \( p_{\text{min}} \), \( q \), \( \alpha \) and \( \beta \), must be chosen carefully. p is the number of complexes that may vary during the inversion process and m is the number of points (models) in a complex. So, the total number of contemporaneous models (sample size) is computed by \( s = p \times m \). \( p_{\text{min}} \) is the minimum number of complexes required in the population; The members of each complex are distributed into \( q \) sub-complexes; \( \alpha \) is the number of consecutive offsprings generated by each sub-complex and \( \beta \) is the number of evolution steps taken by each complex before re-shuffling. A sub-complex is constructed by the CCE optimization mechanism according to a probability distribution selecting \( q \) points from the complex (community) randomly. The probability distribution is specified such that the best point (i.e. the point with the best objective function value) has the highest chance of being chosen to form the sub-complex, and the worst point has the least chance. Based on previous studies, by selecting \( m = 2n + 1 \), where \( n \) is the number of parameters to be optimized, and varying the number of complexes (p), the SCE algorithm provided better results overall than by increasing the m value alone. So, to get the global optimum for the parameters in the SCE method, some default values were selected as follows (Sorooshian et al., 1993):

1. The number of points in a complex, \( m = 2n + 1 \);
2. The number of points in a sub-complex, \( q = n + 1 \);
3. The number of consecutive offspring generated by each sub-complex, \( \alpha = 1 \);
4. The number of evolution steps taken by each complex, \( \beta = m \).

Numerical studies to investigate the proper selection of the parameters found that, no matter what values of \( p \), \( p_{\text{min}} \), \( \alpha \), and \( \beta \), were chosen for optimization, the SCE method was consistently able to find the ‘true’ parameters, provided that a sufficiently large value for \( p \), the number of complexes, was given. However, it was recommended that \( p_{\text{min}} \) should be equal to \( p \) (i.e. \( p \) is constant during the inversion process) as well and it is also recommend that the default value of \( 2n + 1 \) should be used for \( \beta \) and the default value for \( \alpha \) should be set to one (Sorooshian et al., 1993).

The philosophy behind the SCE approach is to treat the global search as a process of natural evolution. The sampled points constitute a population. The population is partitioned into several communities (complexes), each of which is permitted to evolve independently (i.e., search the space in different directions). After a certain number of generations, the communities are forced to mix, and new communities are formed through a process of shuffling. This procedure enhances
survivability by sharing the information (about the search space) gained independently by each community. The SCE method is designed to improve the best features of the CRS method (i.e., global sampling, complex evolution) by incorporating the powerful concepts of competitive evolution and complex shuffling. Both help to ensure that the information contained in the sample is efficiently and thoroughly exploited. They also help to ensure that the information set does not degenerate. These properties enable the SCE approach to have very good global convergence features over a broader range of global optimization problems (Duan et al., 1992).

### 4.1. Surface wave velocity analysis based on the SCE method

Generally, dispersion curves are inverted in 1D using a large number of thin layers considering a constant velocity in each layer. The method is well established (e.g., Herrmann and Ammon, 2007; Herrmann, 2013), but has the inconvenience that it may be difficult to locate geological layer limits such as the base of a sedimentary basin or the crust-mantle boundary (Moho) based on the inversion results due to often relatively smooth transitions in the resulting models. We chose therefore to use another approach. The area of interest is subdivided into rectangular columns of constant size in E–W (X) and N–S (Y) direction. In depth (Z), each column is subdivided into four layers: sediments, upper and lower crust and mantle. In each layer and each column, we consider velocities and densities to have a linear vertical gradient. We are thus looking for the layer thickness, average P-wave velocity (Vp) and compressional velocity by the ratio of compressional to shear wave velocity (Vp/Vs) which we considered constant within each layer, eliminating in this way three more unknowns, the vertical gradients of the compressional wave velocity. Finally, we fixed the density at the lithosphere-asthenosphere boundary (LAB) to 3200 kg/m³ (e.g., Lachenbruch and Morgan, 1990), by which the total number of unknowns per data point reduces to 24. In addition, we replaced the compressional velocity by the ratio of compressional to shear wave velocity (Vp/Vs) which we considered constant within each layer, eliminating in this way three more unknowns, the vertical gradients of the compressional wave velocity. Finally, we fixed the density at the lithosphere-asthenosphere boundary (LAB) to 3200 kg/m³ (e.g., Lachenbruch and Morgan, 1990), by which the total number of unknowns per data point reduces to 24.

In order to calculate the group velocities correctly, the different layers have to be subdivided into a number of sub-layers. To estimate the most appropriate thickness for those sub-layers we repeated the calculations for different numbers of sub-layers. For each layer, the result for U was calculated subdividing the layer into 1 to 20 sub-layers. For a small number of sub-layers, the result depends on the number of layers, whereas for larger numbers from a certain number on, the result stays stable. This number determines the maximum sub-layer thickness. With these calculations, we obtained as maximum values 1 km for the sedimentary sub-layers, 2.5 km for the upper crustal sub-layers, 3 km for those of the lower-crust and 10 km for the mantle sub-layers.

### 4.2. Synthetic data inversion based on the SCE method

In order to show the utility of the SCE inversion for 1D Rayleigh wave dispersion curves, we applied this method to two noise-free and their corresponding noisy synthetic 1D models that are shown in Tables 1 and 2. We fixed the lithosphere depth to 150 km and lithosphere-asthenosphere boundary (LAB) density to 3200 kg/m³ during the inversion. So, to get the global optimum for the model parameters in the SCE method, we used the following parameters: the number of

\[
\begin{align*}
&\left\{ \frac{\partial U_1}{\partial x}, \frac{\partial U_2}{\partial x}, \frac{\partial U_3}{\partial x}, \frac{\partial U_4}{\partial x}, \frac{\partial U_5}{\partial x}, \frac{\partial U_6}{\partial x}, \frac{\partial U_7}{\partial x}, Ti \\
&\frac{\partial V_1}{\partial x}, \frac{\partial V_2}{\partial x}, \frac{\partial V_3}{\partial x}, \frac{\partial V_4}{\partial x}, \frac{\partial V_5}{\partial x}, \frac{\partial V_6}{\partial x}, \frac{\partial V_7}{\partial x}, \frac{\partial V_8}{\partial x}, \frac{\partial V_9}{\partial x}, \frac{\partial V_{10}}{\partial x}, Ti \\
&16s, 20s, 24s, 30s, 40s, i = 1: 5, n = 1: 4 \right\}
\end{align*}
\]
complexes, \( p = 8 \), the number of points in a complex, \( m = 41 \), the number of points in a sub-complex, \( q = 21 \), the number of consecutive offspring generated by each sub-complex, \( \alpha = 1 \) and the number of evolution steps taken by each complex, \( \beta = 41 \). The number of iterations (the number of cycles) was selected as 1000. In total 528,070 models have been calculated for test 1 and 546,563 models for test 2.

Data misfit of the resulting model parameters is defined here as the square root of the difference between measured and calculated Rayleigh wave group velocities.

For the noise-free datasets, all obtained models with a misfit of < 0.001 km/s have been plotted for every test (green lines in Figs. 5 and 8) to indicate the range of possible models explaining the data within a
Fig. 6. Histogram of model parameters for models with $E < 0.001$ km/s in test 1 (the horizontal axes have been scaled to the search space). a) Thickness of sediments (km), b) thickness of upper crust (km), c) thickness of lower crust (km), d) Vp in the sediment layer (km/s), e) Vp at the top of upper crust (km/s), f) Vp at the bottom upper crust (km/s), g) Vp at the top of lower crust (km/s), h) Vp at the bottom of lower crust (km/s), i) Vp at the top of mantle lithosphere (km/s), j) Vp at the bottom of mantle lithosphere (km/s), k) Vp/Vs ratio in sediments, l) Vp/Vs ratio in upper crust layer, m) Vp/Vs ratio in lower crust layer, n) Vp/Vs ratio in mantle lithosphere, o) density of the sediments (g/cm$^3$), p) density at the top of upper crust (g/cm$^3$), q) density at the bottom of upper crust (g/cm$^3$), r) density at the top of lower crust (g/cm$^3$), s) density at the bottom of lower crust (g/cm$^3$) and t) density at the top of mantle lithosphere (g/cm$^3$). The red circle shows the real value of synthetic model parameters. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
small uncertainty. The numbers of selected models were 179,123 for test 1 and 169,578 for test 2. Figs. 5 to 7 show the results of this method for test 1 and Figs. 8 to 10 show the results of this method for test 2. The histograms of all model parameters with misfits of < 0.001 km/s have been shown in Fig. 6 for test 1 and Fig. 9 for test 2. The best model with the smallest misfit for test 1 (9.931788e−05 km/s) and for test 2 (5.246228e−05 km/s) have been extracted in Tables 1 and 2. The minimum data misfits as a function of iterations are illustrated in Figs. 7 and 10. It is observed that for both tests, the misfit values significantly decrease after the first iterations, and then gradually converge to approximately zero.

As mentioned above, for noise-free synthetic data, shear wave velocities, compressional wave velocities and thickness of each layer can be fairly well resolved by the SCE algorithm. However, real surface wave velocities are always noisy. Addition of noise often introduces complications and therefore, the robustness of an inversion algorithm can be judged only after applying it to noisy data. Abdetedal et al. (2015) reports maximum uncertainties of Rayleigh wave group velocities of ± 0.1 km/s. Therefore we perturbed the synthetic Rayleigh wave group velocities of the two synthetic models by adding white Gaussian noise with a sigma of 0.1 km/s. Inversion results of the SCE method are illustrated in Figs. 11 and 12 respectively. The best models found for both tests with noisy data sets have a misfit of about 0.015 km/s (see comparison of noise-free and noisy data results in Tables 1 and 2).

Evidently, one single example of noisy data will not give as optimum model the real model. In order to be able to judge the ability of the procedure to recover the models, we created a set of 100 data sets from the original one having the same noise characteristics, did the inversion for each model and analyzed the best models. As should be expected, their average is near to the synthetic model, showing that the algorithm works well also with noisy data (blue lines in Figs. 11 and 12).

---

**Fig. 7.** Minimum data misfit as a function of iterations in the SCE algorithm for test 1.

**Fig. 8.** Results of the synthetic test 2 of the SCE algorithm. Red line is synthetic model, dark blue line is the best solution, green lines are all the results with a data misfit of < 0.001 km/s and cyan lines are the lower and upper bounds of the search space. a) Compressional wave velocity profiles with respect to depth, b) shear wave velocity profiles with respect to depth and c) density velocity profiles with respect to depth. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
Fig. 9. Histogram of model parameters for models with $E < 0.001\ \text{km/s}$ in test 2 (the horizontal axes have been scaled to the search space). a) Thickness of sediments (km), b) thickness of upper crust (km), c) thickness of lower crust (km), d) $V_p$ in the sediment layer (km/s), e) $V_p$ at the top of upper crust (km/s), f) $V_p$ at the bottom of upper crust (km/s), g) $V_p$ at the top of lower crust (km/s), h) $V_p$ at the bottom of lower crust (km/s), i) $V_p$ at the top of mantle lithosphere (km/s), j) $V_p$ at the bottom of mantle lithosphere (km/s), k) $V_p/V_s$ ratio in sediments, l) $V_p/V_s$ ratio in upper crust layer, m) $V_p/V_s$ ratio in lower crust layer, n) $V_p/V_s$ ratio in mantle lithosphere, o) density of the sediments (g/cm³), p) density at the top of upper crust (g/cm³), q) density at the bottom of upper crust (g/cm³), r) density at the top of lower crust (g/cm³), s) density at the bottom of lower crust (g/cm³), and t) density at the top of mantle lithosphere (g/cm³). The red circle shows the real value of synthetic model parameters. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
5. Gravity modeling

As mentioned above, the dispersion curve is a non-linear function of shear and compressional wave velocity and density of the media. However, the sensitivity to shear and compressional wave velocity is greater than the sensitivity to density. So, we have done gravity modeling to increase the resolution of the density model.

The gravitational effect of a rectangular column with a linear vertical density variation can be calculated using the analytical formula of Gallardo-Delgado et al. (2003):

$$\Delta g = \frac{G \rho_0}{\pi} \left[ x \ln(y + r) + y \ln(x + r) - z \frac{\arctan(xy)}{2} \right] \left[ \frac{y_1}{y_1} + \frac{y_2}{y_1} \right]$$

$$+ \frac{G \rho}{\pi} \left[ -x \frac{\ln(r + z)}{2} + \frac{y^2}{2} \frac{\arctan(zr)}{y^2} \right]$$

$$+ \frac{y^2}{2} \frac{\arctan(zr)}{y^2} \left[ \frac{x_1}{x_1} + \frac{x_2}{x_1} \right]$$

$$r = \sqrt{x^2 + y^2 + z^2}, \quad p(x) = \rho_0 + yz, \quad G = \text{gravitational const}$$

$$: 6.6726 \times 10^{-11} \text{ m}^3 \text{ s}^{-2} \text{ kg}^{-1}$$

where \((x_1,x_2)\) is the distance between a measurement point and the two vertical N-S faces of the column, \((y_1,y_2)\) is the distance between a measurement point and the two vertical E-W faces of the column and \((z_1,z_2)\) is the distance between a measurement point and the upper and lower boundary of the column. Summation over all columns gives the 3D gravitational effect of the model. The model parameters that are to be found in the gravity inversion procedure are thickness, and the average crustal density in each column.

The area of interest, Makran, is also here subdivided into rectangular columns of constant size 1° by 1° (100 km by 110 km) in E-W (X: 53°–66°) and N-S (Y: 23°–30°) direction. In depth (Z), each column is subdivided into four layers: sediment, upper crust, lower crust and mantle. The distance between the gravity data points is 0.2° by 0.2° (20 km by 22 km). The data vector is thus:

$$d^T = [d_1, ..., d_{Ng}] = [\Delta \rho_0, ..., \Delta \rho_{Ng}]$$

where \(Ng\) is the total number of data points. The inverse problem involves minimizing the following cost function (Zeyen and Pous, 1993):

$$CF = \varepsilon_1 + \varepsilon_2 = (G(p - d)^T C_p^{-1}(G(p - d)) + \lambda(p - p_0)^T C_p^{-1}(p - p_0)$$

$$= \varepsilon_1 + \varepsilon_2$$

where \(G(p - d)^T C_p^{-1}(G(p - d))\) and \(\lambda(p - p_0)^T C_p^{-1}(p - p_0)\) are the data misfit and the second term to the distance of the final model from the initial model containing, if available, a priori information. Minimizing this cost function ultimately takes the following iterative form (Menke, 2012):

$$p^{k+1} = p^k + (G^T C_p^{-1} G + \lambda C_p^{-1})^{-1}(G^T C_p^{-1} \Delta d)$$

The program minimizes two \(\varepsilon\) parameters, the relative importance of which may be controlled by the user. The first term corresponds to the data misfit and the second term to the distance of the final model from the initial model containing, if available, a priori information. Minimizing this cost function ultimately takes the following iterative form (Menke, 2012):

$$p^{k+1} = p^k + (G^T C_p^{-1} G + \lambda C_p^{-1})^{-1}(G^T C_p^{-1} \Delta d)$$

where \(Ad\) is the vector of difference between the measured data and those calculated with model parameters \(p^n\) and \(C_p\) is the variance matrix of the data, containing on the diagonal the squared uncertainty of each data point. \(p\) is a vector with the model parameters, \(p_0\) is the vector of initial parameters and \(C_p\) the variance matrix of the parameters which has on its diagonal the uncertainties (variability) \(\sigma_p^2\) of the parameters. For parameters that are constrained by a priori information, \(\sigma_p^2\) will be smaller than for parameters that are unconstrained. \(k\) is the number of iterations and \(G\) is the Fréchet derivative matrix. The damping factor \(\lambda\) controls the overall importance of data adjustment (\(\varepsilon_d\)) with respect to distance of calculated and initial model parameters (\(\varepsilon_p\)). If \(\lambda\) is small, data adjustment controls the inversion process and parameters may change more freely, which may introduce stability problems if the initial parameter set is too far away from the
optimum set. On the other hand, if $\lambda$ is large, the parameters are forced to stay near their initial values at the expense of a good data fit (Motavalli-Anbaran et al., 2013; Zeyen et al., 2005; Zeyen and Pous, 1993). In our algorithm, $p_0$ is updated in every iteration step with the values obtained from the last iteration. In this way, the parameters are not necessarily forced to stay near the initial values, except for those having a small $\sigma_p$, i.e. where prior knowledge is available (Motavalli-Anbaran et al., 2013; Zeyen and Pous, 1993).

Using this approach, the density variation and thickness distribution of the different lithosphere layers were obtained which are shown in Fig. 15. Density of the asthenosphere is fixed (3200 kg/m$^3$) and its base is defined at a constant depth of 150 km as for the surface-wave model.

Table 3 shows the data and model parameter uncertainties used for inversion. As previously mentioned, for model parameters which are constrained by a priori information, $\sigma_p$ will be smaller than for unconstrained parameters (Table 3).

6. Real data inversion based on the SCE method constrained by the gravity results

In a final step, we used the results of gravity modeling to constrain densities and sediment thicknesses by performing a second round of surface wave dispersion inversion with the SCE algorithm. Results for each point of the grid are shown, for all parameters, in Fig. 16 whereas

**Fig. 11.** Inversion results of test 1 using the SCE method. (a) Noise-free synthetic model (red line) and best model resulting from the SCE inversion method (blue line) with error bars (black color). (b) Noisy synthetic model (red line), the best model resulting from one example of single inversion (cyan line) and the average best result (blue line) with their corresponding uncertainties (black error bars). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
Fig. 17 shows the misfit of these results with those of the first SCE inversion. Finally, the Moho geometry, average crustal shear wave velocity and average crustal density for the Makran region are shown in Fig. 18.

7. Results and discussion

The Makran accretionary wedge is an excellent example to study the geology and structures of an active convergent plate boundary involving subduction of the oceanic lithosphere. The E-W trending Makran accretionary wedge extends between the Minab dextral transform fault (to the west) and the sinistral Chaman transform fault (to the east) with a width of 300–350 km. More than half of the accretionary wedge is exposed on land (McCall and Kidd, 1982; Burg et al., 2012; Dolati, 2010). The elevation of the E-W trending mountain range rises from the coast up to 1800 m in the north, along the southern border of the Jazmurian Depression. As mentioned above, the crustal thickness of the Makran region is not well known and there have been few studies of the deep crustal structure of the upper mantle in this area.

In order to test our new inversion algorithm, we created synthetic data with and without noise, shown in Tables 1 and 2. We also employed histograms of the best final models for both tests to demonstrate the performance of the SCE algorithm. We observe that the behavior of the histograms shows quite a good agreement with respect to the values
Fig. 13. The final model based on the SCE inversion. a) Thickness of sediments (km), b) Vs in the sediment layer (km/s), c) Vp in the sediment layer (km/s), d) density of the sediments (kg/m³), e) thickness of upper crust (km), f) Vs of upper crust (km/s), g) Vp of upper crust (km/s), h) density of upper crust (kg/m³), i) thickness of lower crust (km), j) Vs of lower crust (km/s), k) Vp of lower crust (km/s), l) density of lower crust (kg/m³), m) Moho depth (km), n) Vs of mantle lithosphere (km/s), o) Vp of mantle lithosphere (km/s) and p) average density of mantle lithosphere (kg/m³).
of the corresponding model parameters (Figs. 6 and 9).

The misfit evolution can be observed in Figs. 7 and 10. The misfit values significantly decrease and converge to approximately zero, which suggests that the algorithm had completed the exploration for a global minimum. As illustrated in Fig. 7, after iteration 170 to 1000, the misfit is about $10^{-5}$ km/s and all models are similar to each other and the best model does not change anymore.

Fig. 13 reports the best solutions of the implementation of the SCE strategy for real data. Inversion of surface waves alone gives quite patchy results, because the inversion is done in 1D, independent for every column. We found that the patchiness in our results is not related to the number of inversions, but is a consequence of the 1D inversion which gives completely independent results for each column and shows the effect of noise in the data. Although the misfit decreases in the real-data model after 100 iteration, the patchiness does not increase.

Our data cannot cover the sedimentary layer in detail where it is much thinner than 10 km as predicted by Abdetal et al. (2015) results, as before mentioned. Therefore, the results presented in Fig. 14 may be influenced by the underlying basement. However, the upper crust and lower crust velocities are also quite patchy but resolved better than the sediment layer. In contrast, constraints by the gravity model, which is in 3D, not only allow better density models but also give a geologically more coherent, smoother model without increasing considerably the data misfit.

Abdetal et al. (2015) have found the Rayleigh wave with the shortest period of 16 s has a fair sensitivity to the top 10 km, and the wave with the longest period of 40 s has a peak sensitivity at about 60 km depth and fair sensitivity up to about 80 km. Thus, using the dispersion curves from 16 to 40 s periods allows us to constrain shear velocities from 10 km to 80 km depth. It is thus not surprising that the obtained model parameters for the sediment layer are not close to the synthetic models even for the noise-free data. Whereas in the crust and upper mantle, results are very close to the synthetic models, especially for S-wave velocities, except for the limit between upper and lower crust where the velocity contrasts are small and difficult to obtain for any inversion algorithm. Densities are less well resolved by single surface wave inversion, as should be expected (Figs. 5 and 8).

As mentioned above, the lower and upper bounds of the search space were approximated based on the global crustal model in both of the noise-free and noisy data sets. In fact, we used wider search space boundaries to simulate more realistic cases where no a priori information is available in the SCE modeling. As we expected, the velocities are well recovered, whereas the densities differ from the synthetic models that can be seen from errors in Tables 1 and 2.

Since, the real surface wave velocities are always noisy, we cannot obtain the true model parameters just by the SCE method. Therefore, we applied 3D gravity modeling using the surface wave analysis results for thickness of upper crust and lower crust, to obtain thickness and

![Fig. 14. Data misfit for real data inversion (calculated data minus observed data) for Rayleigh wave group velocity in Makran at periods of a) 16 s, b) 20 s, c) 24 s, d) 30 s and e) 40 s.](image-url)
Fig. 15. The final thickness and density model obtained from gravity inversion. a) Thickness of sediments (km), b) density of the sediments (kg/m³), c) thickness of upper crust (km), d) density of upper crust (kg/m³), e) thickness of lower crust (km), f) density of lower crust (kg/m³), g) Moho depth (km) and h) average density of mantle lithosphere (kg/m³).

Table 3

Inversion parameters used for gravity inversion.

<table>
<thead>
<tr>
<th>Data</th>
<th>Gravity (mGal)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uncertainty</td>
<td>5 mGal</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Thickness (m)</th>
<th>Density (kg/m³)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \sigma_p )</td>
<td>1000 m (500 m for sediment)</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>500 m (100 m for sediment) with prior information</td>
<td>10 with prior information</td>
</tr>
</tbody>
</table>
Fig. 16. The final model based on the SCE inversion. a) Thickness of sediments (km), b) Vs in the sediment layer (km/s), c) Vp in the sediment layer (km/s), d) density of the sediments (kg/m³), e) thickness of upper crust (km), f) Vs of upper crust (km/s), g) Vp of upper crust (km/s), h) density of upper crust (kg/m³), i) thickness of lower crust (km), j) Vs of lower crust (km/s), k) Vp of lower crust (km/s), l) density of lower crust (kg/m³), m) Moho depth (km), n) Vs of mantle lithosphere (km/s), o) Vp of mantle lithosphere (km/s) and p) average density of mantle lithosphere (kg/m³).
density variations of sediment, upper crust and lower crust layers (Fig. 15). Then, the resulting model of gravity inversion was then used to constrain densities and sediment thicknesses in the surface wave modeling based on the SCE method (Figs. 16 and 18), resulting in a smoother and geologically more meaningful model. In fact, to reduce the errors of the model parameters, we used the result of the gravity modeling to limit the search space in the SCE method.

For the noisy synthetic data; it is clear that one single example of noisy data will not give a optimum model the real model. In fact this is not a problem of the method but of the noisy data. In order to be able to judge the ability of the procedure to recover the models, we created a set of 100 noisy data sets from the original one with the same noise characteristics, did the inversion for each of them and analyzed the best resulting models. As should be expected for a working inversion algorithm, the average of these models resulted close to the synthetic model. Evidently, this procedure is only applicable to synthetic data, however it shows that the algorithm gives the expected results also with noisy data. As the best model resulting from one example of single inversion (cyan line) and the average best result (blue line) with their corresponding uncertainties have been shown in Figs. 11 and 12. The final model of the single inversion differs clearly more from the original model than the average model, in particular for densities in the upper crust and mantle.

As we expected, data misfit of the final model increases in comparison with the first SCE unconstrained inversion are larger (Fig. 17). This is an effect of the gravity constraint, which imposes some lateral smoothness of the model, which, however, does not imply that the final model would not be less meaningful geologically, since the final model explains two independent data sets. Also, based on the checkerboard test results of the Rayleigh wave group velocity tomography results, the resolution is low for the east and south of the region, which makes the inversion process not recover the absolute amplitudes well in this region (Abdetedaletal., 2015). So as we expected data misfit are larger in these parts of our region.

Since, the Makran trench is deeper than the other part of the ocean basin, the sediments flooring the trench have low density. The mass deficiency of the water and sediments in the trench cause a strong negative free-air anomaly. This anomaly parallels the trench and has an amplitude of $-150$ mGal which are generally consistent with the previous studies (Lowrie, 2007). The average crustal density is highest densities in the Oman Sea as expected for the oceanic crust as well.

It has been found that the undeformed sediments thickness is $>7$
km at the front of the Makran wedge and the total thickness of deformed and undeformed sediments becomes > 10 km near the coastline (Minshull et al., 1992; Kopp et al., 2000). Our results show a good agreement with these previous studies (Fig. 16a).

In the final results, we have found that Moho depth is 35–45 km beneath the Makran region, 18–28 km for the Oman Gulf, and 46–48 km for the Taftan-Bazman volcanic-arc that is in good agreement with recent studies by Kopp et al. (2000), Niazi et al. (1980), Dehghani and Makris (1984), Manaman et al. (2011), Smith et al. (2013) and Entezar-Saadat et al. (2017).

The shear velocity images of upper-mantle structure across the Makran subduction zone depict a high-velocity anomaly under Oman seafloor and lower velocities in the on-shore part. Also the crust under the fore-arc, volcanic-arc and back-arc settings of Makran subduction zone is determined by low-velocity structures which is consistent with the previous studies (Manaman et al., 2011; Abdetedal et al., 2015) and may be due to high temperatures related to the volcanic activity. The subducting slab is not visible underneath the onshore areas, certainly due to the data resolution limited to the uppermost 80 to 100 km.

Many of the prominent features in our results are consistent with the known geological structures. The Strait of Hormuz is considered a transition between the Zagros collision and the Makran oceanic subduction (Regard et al., 2010). A sharp transition boundary between the low- and high-velocity zones with an N-S trend is clearly depicted at the Minab fault system (see Fig. 9a, b) which is considered to represent the boundary between the continental collision zone between Eurasia and Arabia to the west and the oceanic subduction between Eurasia and the Oman Sea to the east (White and Ross, 1979). A pronounced low-velocity anomaly in the crust, extends in WSW-ENE direction east of the Minab fault, which is attributable to volcanic-arc and back-arc settings of the Makran region and the Bazman and Taftan volcanoes (see Fig. 2a, b).

8. Conclusions

We implemented a new method of iterative, sequential inversion of gravity anomaly and Rayleigh wave group velocity data. With this method, we established a model of the crust and upper mantle structure of the Makran region which is one of the largest accretionary wedges on the globe, formed by the convergence between the Eurasian and the Arabian Plates. The main results of the modeled parameters are as follows (Figs. 16 and 18):

- The Oman Gulf is characterized by a Moho depth of about 18–28 km in the southern part, increasing northward towards the Iranian coast to 35 km. Further North, the Moho depth increases to 35–40 km below the Jazmurian basin and 46–48 km beneath the Taftan-Bazman volcanic-arc which is compatible with earlier studies in the region.
- The shear velocity images of the upper crust and lower crust display a high-velocity anomaly under the Gulf of Oman with an oceanic crust, decreasing northward towards the Makran region with a continental crust.
- The density image of the region shows that the average crustal density is relatively high under the Gulf of Oman with an oceanic crust, decreasing northward to Makran with a continental crust.

Acknowledgements

The work was initiated during a 9 months research stay at Université Paris-Saclay, Département des Sciences de la Terre, Orsay, France. Ministry of Science, Research and Technology of Iran acknowledged for the financial support of the first author’s sabbatical leave at Université Paris-Saclay, Département des Sciences de la Terre in the period of December 2014 to June 2015. Figures were created with GMT (Wessel and Smith, 1995). Thanks to Bureau Gravimétrique
International (BGI) for gravity data. Thanks to M. Abdetedal for generously providing us with her PhD research results on the Rayleigh wave dispersion curve data.

References


