This is a peer-reviewed, accepted author manuscript of the following research article: Parastatidis, E, Pytharouli, S, Stankovic, L, Stankovic, V, Shi, P & Hildyard, MW 2023, 'Multichannel Coherency Migration grid search (MCMgs) in locating microseismic events by a surface array', Geophysical Journal International.

Multichannel Coherency Migration grid search (MCMgs) in locating microseismic events recorded by a surface array

E. Parastatidis[1], S. Pytharouli[1], L. Stankovic[1], V. Stankovic[1], P. Shi[2], M. W. Hildyard[3]

[1] *University of Strathclyde, Glasgow, U.K.*

[2] *Swiss Seismological Service, ETH Zurich, Switzerland.*

[3] *University of Leeds, Leeds, U.K.*

Received 2022 December 16

SUMMARY

Microseismic monitoring has been used in geo-energy related activities, such as shale-gas exploitation, mining, deep geothermal exploitation, geotechnical and structural engineering, for detecting and locating fractures, rock failures, and micro-earthquakes. The success of microseismic monitoring depends on reliable detection and location of the recorded microseismicity. Multichannel Coherency Migration (MCM) is a detection and location waveform migration based approach which does not require phase picking, identification, and association and performs well on noisy data. Its caveat is a high computational cost, which impedes its application of MCM on large datasets or for real-time monitoring. To address this issue, we propose an improved approach, the Multichannel Coherency Migration grid search (MCMgs), by introducing an adaptive grid optimization technique. Based on results from synthetic and real data, we show that MCMgs reduces the computation time up to 64 times. In addition, MCMgs generates multiple maximum coherency values with various grid sizes instead of a single (maximum) coherency value that links to a single grid point and size, thus resulting in more accurate locations. Our simulation results on different deployment geometries demonstrate that MCMgs is effective even with a small number of recordings available - a minimum of seven. We conduct a sensitivity analysis to assess how the detectability of events is affected

by the spatial arrangement of the deployed monitoring array. If a limited number of seismometers are available for deployment, our analysis favors a patch array deployment geometry. We show that twelve seismometers deployed at a patch array geometry can have similar detection and localisation capability as a large rectangular array of more than 100 seismometers but at a much lower computational and deployment cost.

Key words: Microseismic – time-series analysis – computational seismology – location algorithm – grid search – earthquake monitoring – patch array.

1 INTRODUCTION

Microseismic monitoring technology used for passive seismic monitoring of subsurface processes has evolved rapidly in recent years (Li, L. et al.,2020; Shi, P. et al.,2022; Zhu, T.2022). Limitations and challenges still exist however, especially concerning near real-time results from analysis of microseismicity recordings with low signal-to-noise ratio (SNR). Microseismic monitoring is commonly performed in noisy environments e.g., near pumps or drilling equipment, during hydraulic stimulation, near heavy traffic, and in other environments with high levels of anthropogenic noise.

Many different approaches have been proposed for seismic event localisation, such as the linearized travel time inversion (Geiger L.1912), travel time-based non-linear global grid search, and double-difference relative location method (Li, L. et al.,2020). For the non-linear global grid search approach, different global search methods have been developed, including genetic algorithms (Kennett, B. et al.,1992) and Monte Carlo (Sambridge, M. et al.,1992), to search over the target space. These methods, known as picking or ray-based methods, use phase arrival times and locate seismic events by searching for the minimum misfit between the theoretical and observed travel-times through either a linearized travel-time inversion or a grid search (Li, L. et al.,2020). In the case of weak microseismic events or short source-to-receiver distances (a few hundred meters to 1-2 km), picking phase arrival-times is challenging as the SNR is low and/or P- and S-phase are not well separated. In such instances, only one phase can be reliably picked but with uncertainties both in terms of timing (the onset is often emergent rather than impulse) and phase identification

(Diehl, T., et al.,2012). Hypocentral location results based on a single phase have large location uncertainties, especially concerning the hypocentral depth. A methodology that does not depend on phase picking for the location process overcomes this problem.

Waveform-based methods such as time reverse imaging (TRI) and full waveform inversion (FWI), satisfy this constraint as they calculate the source location through migration and stacking processes similar to the methods used in active seismic monitoring (Li, L. et al.,2020). To be effective, they all require one or more of the following: long computation times, availability of large number of seismic recordings per seismic event, accurate velocity models (Zhu, T.2014) (Li, L. et al.,2020; Shi, P. et al.,2019a) (Larmat, C. et al.,1992; Steiner, B. et al.,2008). All the above render their application less user friendly, and prohibiting for real-time projects.

Migration-based approaches on the other hand, are typical waveform-based methods that fo-cus on reconstructing the seismic source as a discrete grid point based on waveform travel times (Grigoli, F. et al.,2014). The approach of Multi-channel Coherency Migration (MCM) (Shi, P. et al.,2019a; Shi, P. et al.,2019b), instead of directly migrating the waveforms, stacks the coherency of waveforms and associates the source location to the point with the maximum coherency value (Shi, P. et al.,2019a; Shi, P. et al.,2019b; Li, L. et al.,2020)). It calculates the coherencies between waveforms from different pairs of stations. MCM can be accurate even for data with low SNR and does not require a detailed (1D) velocity model, a favourable attribute for the case of shallow microseismic events. Because MCM stacks waveform coherency from station pairs whose number is proportional to the square of available stations, the more recordings available, the better the estimated hypocentral location. However, to achieve a location accuracy of the order of tens of meters or less, a dense monitoring network is necessary, which results in additional computation time. Despite this, the algorithm is still more computationally efficient than the traditional FWI location methods but less computationally efficient from location algorithms that are based on phase picking (Shi, P. et al.,2019a; Shi, P. et al.,2019b; Li, L. et al.,2020).

In this paper, we address the issue of computational complexity. We name the proposed improved version of the MCM algorithm Multichannel coherency migration grid search (MCMgs). We use synthetic data generated using the numerical modelling code WAVE3D (Hildyard, M. W.,

2007) which can simulate a more realistic waveform in cracked media to test and validate the effectiveness of MCMgs. We then demonstrate the application of MCMgs on real microseismic events recorded during the 2021 M_L 6.3 Elassona-Larisa earthquake sequence.

2 METHODOLOGY

(Shi, P. et al.,2019b) showed that the computation time required by the MCM algorithm increases linearly with the number of grid points. As a result, the larger the considered volume and/or the smaller the spatial interval between the grid points, the more time is needed to locate a seismic event. Near-real-time monitoring cannot be combined with fine grid spacing due to increased computation times, and better resolution for the seismic event locations requires post-processing. Other factors affecting the computation time are: the number of processor cores (N_c), the number of origin times related to the number of sampling points in the chosen window frame and the full length of the waveform (N_t), the number of recording stations (N), the number of imaging/grid points (N_s) based on the grid and a coefficient related to computer architecture k through the linear relationship (Shi, P. et al.,2019b) :

$$t = k \times N_s \times N_t \times N \times (N-1)/N_c. \quad (1)$$

Assuming that the same processor is used then N_c and k are constant. For a fixed waveform length N_t is constant, and for the same network deployment topology the number of stations does not change. The computation time is then proportional to the number of imaging points N_S in the grid. We propose to use the collapsing grid approach to dynamically reduce the number of grid points during the migration location process the number of grid points, and thus reducing the computation time of the MCM algorithm without affecting its performance as it concerns the resolution of the hypocentral location.

2.1 The collapsing grid as part of the MCM localisation process

Location methodologies have successfully incorporated grid search approaches with collapsing grid iterations (Nelson, G.D. et al.,1990; Lomax, A. et al.,2001; Wang, H. et al.,2016). Such ap-

proaches identify the grid point at which the difference between the observed and calculated travel times is minimised and the origin time that better fits to that location in the grid is calculated. At the next step, the grid search space is changed to a finer grid (a smaller spatial interval between grid points) around that point and a new best-fit location is calculated.

In order to reduce the computation time required by MCM to calculate a hypocentral location at a specific spatial resolution, we split the process into several iterations starting with a large search space and a large spatial interval (distance between consecutive grid points). Based on the result of the initial iteration, the search space becomes gradually smaller and the resolution is increased by decreasing the spatial interval between the grid points.

To define the search area for the first iteration we transform the waveforms recorded by each seismometer (station) to the time-frequency domain. Based on the maxima of the dominant frequencies on the power spectrogram we define an approximate time for the P and S wave arrivals at each station. The algorithm then calculates the time differences between the stations with the earliest and latest arrival times and compares these times with the theoretical travel times from each grid point. The minimum and maximum coordinates of these grid points define the search volume for the first iteration.

At a next step, we set a maximum time for each iteration. Since the number of stations, the processor cores, and the number of origin times are known, we can estimate the number of grid points N_s from eq.1. The initial spatial interval between grid points Δx along each of the three directions can then be defined as:

$$\Delta x^{\frac{1}{3}} = \frac{V}{N_s - 3}, \qquad (2)$$

where V is the initial search volume. Then we execute the first iteration. For a cubic search volume, there is a maximum location error linked to the grid element size and is equal to:

$$D = \frac{\Delta x}{2}\sqrt{3}, \qquad (3)$$

where D is the distance between the centre of gravity of the cubic grid element and each of the nodes, and Δx is the spatial interval between grid points (nodes) in any of the three directions.

The grid should collapse around the location calculated in the first iteration. However, this is

not without implications. By incorporating the collapsing grid approach within the MCM workflow we violate one of the basic criteria of MCM; the distance between consecutive grid points cannot be larger than half the wavelength of the seismic wave (Shi, P. et al.,2019a). It should be noted that this limitation is not referring to solving the wave equation numerically and avoiding numerical dispersion of the synthetic waveforms. This is a limitation to ensure correct execution of cross-correlation of the waveforms in space and time. It also ensures that if the hypocentral location changes as much as the associated spatial error (see eq. (3)), this new hypocentral location is still included within the search volume of the new iteration. From preliminary analysis, we find that violating the distance criterion results in multiple local maxima, i.e., neighbouring grid points with similar high coherency values, instead of a distinct single point with a maximum coherency value. These points lie within an average distance three times the D value given by eq.(3). Hence, any of these grid points could be a potential hypocentral location. MCM is programmed to choose the point with the maximum coherency value, which does not necessarily consist the most correct or accurate solution (hypocentral location). To overcome this, we propose an alternative to the traditional collapsing grid approach. Our alternative approach is based on the probability density function (PDF) (Lomax, A. et al.,2001). Once we obtain an initial hypocentral location from the first iteration, that location is set as Point Zero. Next, we "filter" the calculated locations (grid points) and the corresponding origin times based on their coherency values: Any location and origin time with coherency above the 0.99 quantile of the total combined number of locations and origin times, are stored to be used in the PDF calculation. For these locations, we calculate the distance from Point Zero along the North, East, and Vertical (Depth) (N, E, D) direction. Next, we calculate the PDF for each of the above North, East and depth values

$$f(x|\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}, x \in R, \qquad (4)$$

where x is the distance of the point of local maximum coherency from Point Zero along each direction separately (either N, E or D), μ is the mean value of the distances of local maxima points from Point Zero and σ is the standard deviation. The maximum values in each direction define the dimensions of the new search space to which the grid will collapse. The next iteration starts with a reduced spatial interval between the grid points related to the new space and the maximum

iteration time we have defined originally. This process is repeated several times until we reach the desired resolution. Using the PDF to calculate the grid search volume for each next iteration improves the final location error of MCMgs, compared to the first iteration location error, as the grid collapses. The entire workflow of the proposed MCMgs is summarised in Figure 1a. Figure 1b shows an example of the grid collapsing around the area with higher coherencies.

3 DATA

For this study we use synthetic data which we create using numerical modeling and a synthetic source of two Ricker wavelets with opposite polarity and the same amplitude, rise time, and dominant frequency of 15 Hz (see also Figure A1 in supplementary material for a full description). All models are homogeneous isotropic cuboids, with density 2740 kg/m^3 and wave velocities of 6020 m/s and 3254 m/s for P and S wave, respectively. The boundary conditions are viscous at all sides of the model (apart from the top surface that is set as a free surface) and absorb most energy of the wave. Some of the energy of the wave is still reflected back. We set the model run-time to less than the time needed for the reflections to reach the stations and as a result, boundary reflections are not included in the generated waveforms. We use 3 groups of models. Group 1 consists

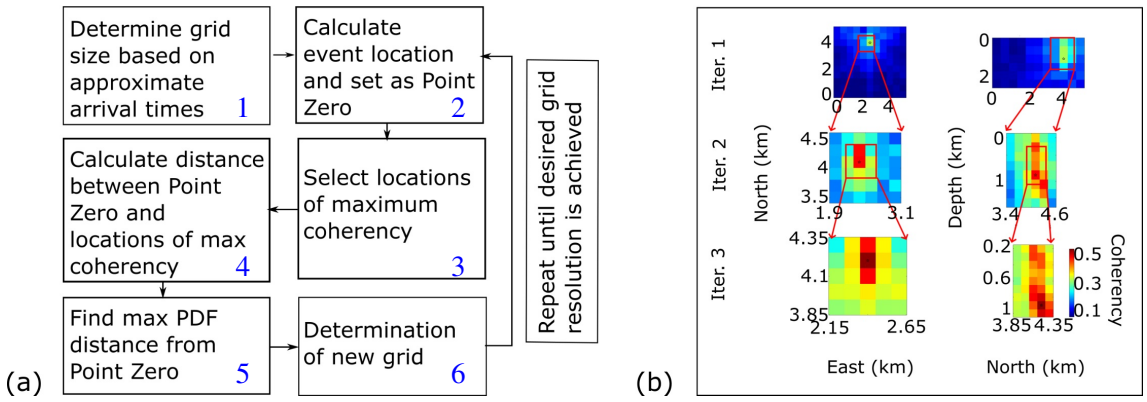

Figure 1. (a) MCMgs workflow. (b) Graphical explanation of migration and collapsing grid method in MCMgs.Left column shows the epicentral plane (North-East), the right column is a cross-section along the North direction (depth profile). The event is located at 4 km North 2 km East and 1 km depth. Top row shows the initial focused area (Iteration 1). The middle row is Iteration 2 based on the initial location. Bottom row is Iteration 3 and final computed event location. The colour scale shows the coherency, linearly normalised between 0 and 1. A larger version of this figure is provided in supplementary material (Figure A2)

of 20 models with epicentral locations inside the station array grid. The dimensions of Group 1 models are 5 km $\times$ 5 km $\times$ 2 km (N, E, D) (Figure 2a). We position the seismic source at 4 different depths; 1200 m, 900 m, 600 m, and 300 m from the free surface. For each depth, we try 5 different epicentral locations inside the station array grid (see Figure 2a, red circles). Group 2 is the same as Group 1 but with the epicentral locations outside the array grid, considering again 4 different depths (1200 m, 900 m, 600 m, and 300 m) and 5 epicentral locations for each depth. The dimensions of this set of models are 5 km $\times$ 7.5 km $\times$ 2 km (N, E, D) (Figure 2b). Group 3 makes use of a more realistic seismic source. For this group, we created a slipping fracture model. The dimensions of the model are 4.5 km $\times$ 4.5 km $\times$ 4 km and its material properties are the same as for the previous models. We simulate a mining excavation near a pre-existing fracture, a case frequently faced by mining/tunnel engineers and geologists (Figure 2c). We initially solve a static model with defined principal stresses close to failure based on Mohr-Coulomb criteria. As the excavation advances and meets the condition for failure, the model is solved in dynamic mode and the fracture slips creating a seismic event. A detailed description of this model is provided in (Hildyard, M. W. et al.,1995; Napier, J.A.L. et al.,1997). It should be noted that the seismic source is in this case not a single point, as in the models in Groups 1 and 2, but a surface where the rupture initiates. The initial rupture has an ellipsoidal shape with major and minor axes equal to 75 m and 6 m, respectively. We test 4 models in Group 3, the only change between them being the depth of the initial rupture (3500 m, 2500 m, 2000 m, and 1000 m). The models of all three Groups use 9 monitoring stations (seismometers) set at a grid array geometry on the top surface of each model (as shown in Figure 2), recording the velocity history in three dimensions. The distance between consecutive stations is 1.8 km (Figure 2). To the synthetic velocity histories of all 44 models (20 models from Group 1, 20 from Group 2 and 4 from Group 3), we add normally distributed random noise. Two different SNR values were tested for each of the 44 models for the three Groups: SNR=20 (practically almost noise-free) and SNR=1. The SNR was calculated as:

$$SNR = (\frac{RMS(A_{signal})}{RMS(A_{noise})})^2, \qquad (5)$$

where RMS stands for Root Mean Square and A is the amplitude of the seismic waveform. Figure 3a shows examples of the waveforms generated by models from Groups 1 and 3 to which we have

added noise such that SNR = 1 and SNR = 20. For our analyses, we used an office workstation with 4 cores (8 threads) and 32 GB of RAM which gives approximately 4 GB on each core. Due to access issues to the office workstation, the models of Group 2 and 3 were run on a personal laptop with 2 threads and 8 GB of RAM.

4 PERFORMANCE OF MCMGS WITH SYNTHETIC DATA

4.1 Reducing computation time

For the models in Groups 1 and 2, the search space for the first iteration covered the full dimensions of the model used for the creation of the synthetic data. To comply with the distance criterion mentioned previously, the interval between grid points should be for our models less than 108 m which served as our spatial resolution target. To reach this resolution we split the localisation process into three Iterations starting with a coarse grid and collapsing to grids with smaller spatial intervals each time. Then we compare both the location results and the computation time with results obtained with MCM from a single iteration that uses the smallest spatial interval (100 m).

Tables 1, 2 and 3 shows the computation time and median values of the location errors obtained when applying MCMgs over three iterations (with the grid becoming finer as the number

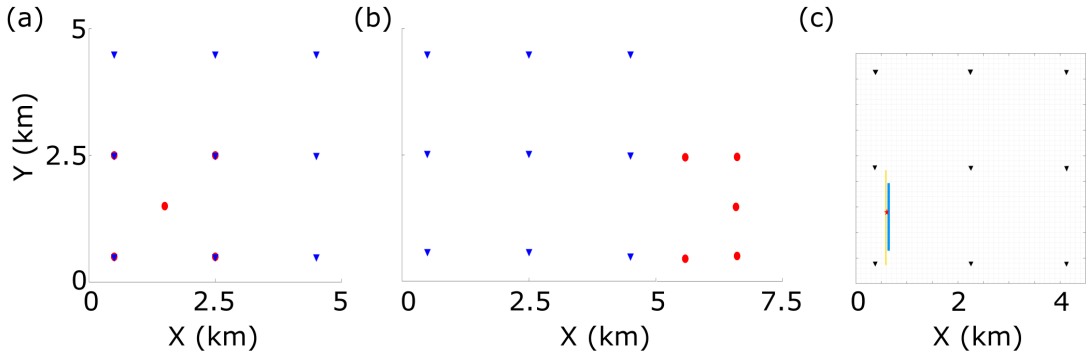

Figure 2. Plan view of models in (a) Group 1, (b) Group 2 and (c) Group 3. The interval between stations (blue triangles) is 1.8 km. We test 5 different epicentral locations (red ellipses) (a) inside the grid array and (b) outside the grid array. Group 3 models also have 9 stations (black triangles), a vertical plane fracture (yellow line) and an excavation (tunnel whose longitudinal axis is shown as the blue line). The trace of the rupture on the XY horizontal plane does not change for the models of Group 3, only its depth.

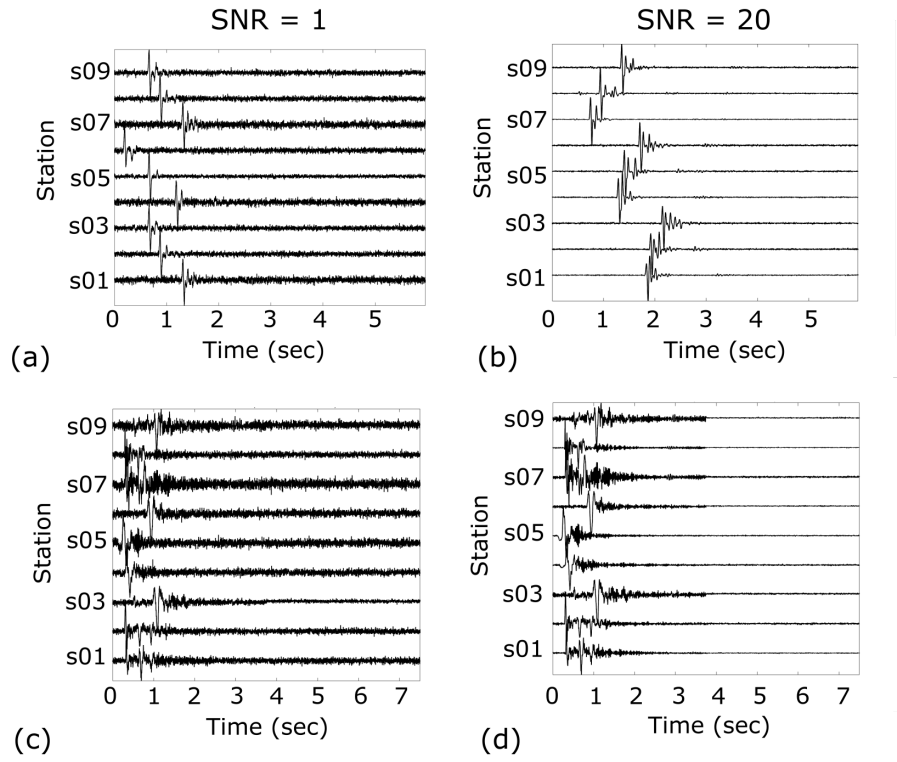

Figure 3. Representative synthetic waveforms derived using a double Ricker wavelet source from models of Group 1 with (a) with SNR=1, (b) SNR = 20. And representative synthetic waveforms from the models of Group 3 with (c) SNR = 1 and (d) SNR = 20.

of iterations increases), against a single iteration with a fine grid since start for the models of all Groups. These results are based on synthetic data with SNR = 1 and 9 stations. For Group 1, the total computation time (3 x 10s = 30 s) was reduced up to 64 times when using the MCMgs compared to the initial version of MCM (see Table 1). This stands for Group 2 models too, while for Group 3 the difference in total computation times between MCMgs and MCM is slightly smaller (MCMgs is 62 times faster than MCM). Computation times for Group 2 and 3 are very different compared to Group 1 because we had to use a different computer with lower specifications. Nevertheless, it is evident that the localisation process is accelerated the same amount when using MCMgs compared to MCM irrespective of the computational strength of the computer used. The computation times in Tables 1, 2 and 3 include the time required for the computations of the travel times table in every iteration; the travel times table is recomputed every time the grid gets finer. The computation times are the same for every iteration in Group 1 and respectively, in Group 2

and 3, because as explained previously, we define the number of grid points on each iteration and as described in above in eq. 1 the N_s controls the computation time.

Figure 4a presents the summary statistics (as box plots) for the models of Group 1 (seismic source within the array grid) and Group 2 (seismic source outside the array grid) for both SNR = 1 and SNR = 20. In Figure 4c are also shown the summary statistics for the results of the localisation of the 31 real seismic events from the 2021 Elassona-Larisa seismic sequence. No summary statistics are provided for the models of Group 3 because these were only 4 and as such not enough for the statistical quantities to be meaningful. The median values of the location errors for Group 3 are provided in Table 3. They follow the same pattern as Groups 1 and 2. The location error is larger when the seismic source is outside the array grid compared to that when the source is within the array grid. This is expected due to poorer azimuthal coverage for Group 2 models. The location error has larger values for the models of Group 3 that have a more complicated source compared to that of models in Groups 1 and 2, but still acceptable and within the same order of magnitude as the error in the models of Group 2 (source outside the array grid). Overall, the median value of the errors both in the horizontal (epicentres) and in the vertical (depth) direction decrease with the number of iterations for Groups 1 and 2 and for the real data from Larisa. In some cases, it appears that results, when waveforms of SNR = 20 are used, have larger errors compared to those obtained from waveforms of SNR = 1. For example, in Iteration 3 for Group 1 models and Iteration 2 for Group 2 models (see Figure 4 a and b). At first glance, this appears to be a paradox and unexpected. However, this happens due to the numerical model itself (numerical noise and small reflections from free surface and viscous boundaries). This noise is of very low amplitude (2-3 orders of magnitude than the actual signal) but large enough to produce high coherency values when running the MCMgs. When the noise level is higher, the numerical noise is suppressed and does not influence the final result.

4.2 Minimum number of stations required for localisation

MCM can identify the event location even in a noisy environment with SNR=1 (Shi, P. et al.,2019a; Shi, P. et al.,2019b). In practice, a microseismic event is often not recorded by all available re-

Table 1. Summary statistics of location errors and computation times for MCMgs and MCM for SNR=1 and nine stations for the models of Group 1 (Figure 2a).

	Distance between grid points (m)	Search volume (km^3)	Median epicentral location error (m)	Median depth location error (m)	Computation time (s)
1st iteration	500	108	223	200	10
2nd iteration	200	1.5	100	100	10
3rd iteration	100	0.13	0	55	10
MCM (Single pass)	100	108	0	55	1920

Table 2. Summary statistics of location errors and computation times for MCMgs and MCM for SNR=1 and nine stations for the models of Group 2 (Figure 2b).

	Distance between grid points (m)	Search volume (km^3)	Median epicentral location error (m)	Median depth location error (m)	Computation time (s)
1st iteration	500	108	412	300	210
2nd iteration	200	1.5	71	300	210
3rd iteration	100	0.13	50	187	210
MCM (Single pass)	100	108	71	150	40000

Table 3. Summary statistics of location errors and computation times for MCMgs and MCM for SNR=1 and nine stations for the models of Group 3 (Figure 2c).

	Distance between grid points (m)	Search volume (km^3)	Median epicentral location error (m)	Median depth location error (m)	Computation time (s)
1st iteration	500	125	242	1315	240
2nd iteration	200	1.2	201	1015	240
3rd iteration	100	0.12	156	600	240
MCM (Single pass)	100	125	50	50	45300

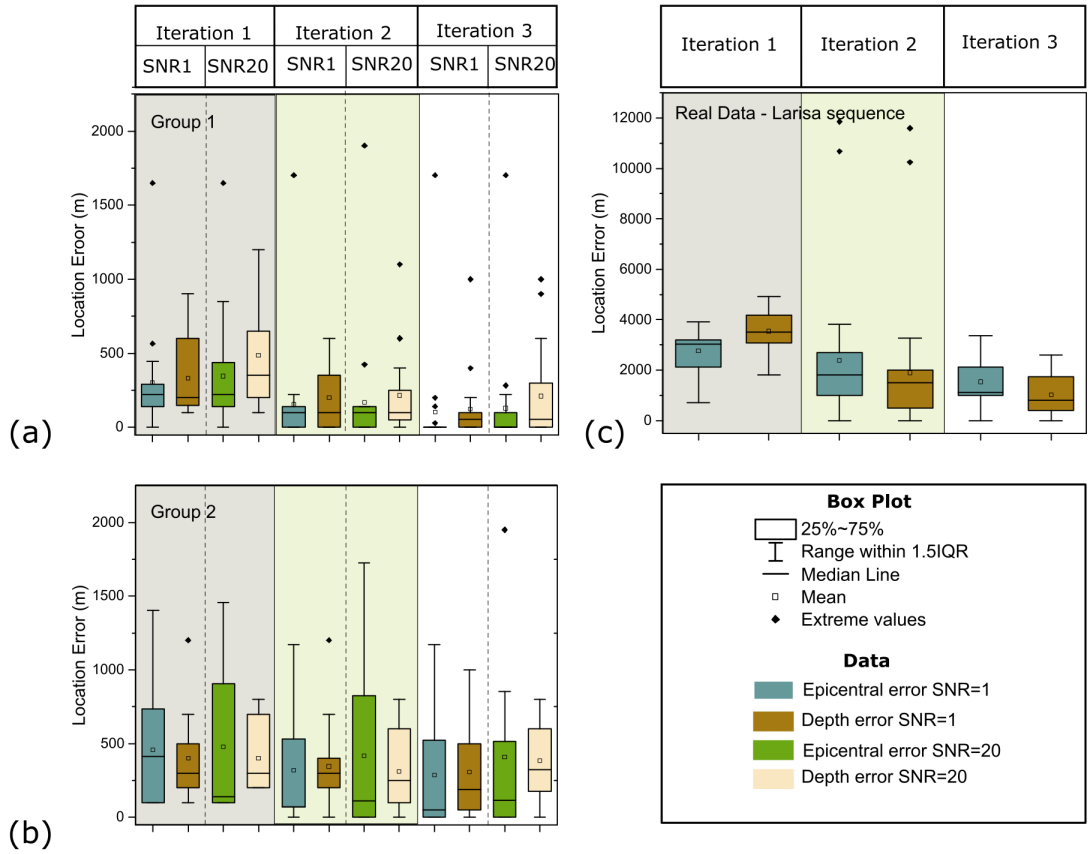

Figure 4. Box plots of errors in epicentral location and depth for (a) the models of Group 1, (b) the models of Group 2 and (c) the 31 real seismic events from the 2021 Elassona-Larisa sequence

ceivers due to noise and signal attenuation. We investigate the number of minimum recordings that are required to accurately locate a microseismic event using MCMgs. For this, we use the results of a single model from Group 1 and post-process the output to "mute" different receivers. We add random noise to the waveforms and test two different levels of noise (SNR=20 and SNR=1) producing in total 240 cases with the number of receivers reduced from 9 to 6 and noise levels increasing from noise-free to SNR=1 (20 different locations, 3 noise levels and 4 array geometries). The median location error calculated when using 8 receivers instead of 9 did not change from 400 m. When the number of receivers is reduced to 6 the location error is more than 4 times larger (see Table 4, error after 3rd iteration). For our models and grid geometry, this gives a threshold of a minimum of 7 recordings to locate an event within 400 m accuracy for a single iteration. This error is reduced further to 100 m when using 3 iterations within MCMgs. The different levels of noise had no impact on the location error for the models used.

Table 4. Example of median location error for different number of receivers and SNR=1 for all the different models tested.

Number of receivers	Error after 1st iteration (m)	Error after 2nd iteration (m)	Error after 3rd iteration (m)
9	400	200	55
8	400	200	100
7	400	200	100
6	800	600	450

4.3 Effect of array size and geometry on MCMgs results

We test the location accuracy achieved with MCMgs of an array with few seismometers (i.e., 12) deployed at a patch geometry versus a large array (= 122 seismometers) deployed at square grid geometry. Reducing the number of receivers results in dramatically reducing (on the order of N^2) the computation time and hardware cost. Patch array geometries have been previously presented in the international literature, e.g.(Eisner, L. et al.,2010; Maxwell, S. C. et al.,2010; Duncan, P. M. et al.,2010; Zimmer, U.,2011; Eisner, L. et al.,2011; Yaowen, B. et al.,2017; McClellan, B. et al.,2018). The main advantage of the patch design is that the receivers, which are close to one another, provide recordings that can be used for de-noising purposes and thus increasing the SNR (Eisner, L. et al.,2010; McClellan, B. et al.,2018). For these reasons we chose the patch array geometry to test the performance of MCMgs. We run numerical simulations for the magnitude sensitivity of the array to detect small events. Then we use synthetic data to compare the accuracy of the patch array against a larger grid array using MCMgs for locating the events. We use a larger array with 122 receivers in a square geometry at a 200 m spacing between receivers (Figure 5). The patch array used in this work consists of 12 receivers divided into 3 groups of small aperture (200 m), and a larger distance among the groups (1730 m) (Figure 6). We test the theoretical magnitude sensitivity of each deployment geometry using the InSite software. Magnitude sensitivity refers to the minimum theoretical magnitude of an event that can be detected at each point of the monitoring space. To calculate this value the parameters used are as defined in (Havskov, J. et al.,2010): the Q factor, the density of the medium, the sensitivity threshold of the receivers, the

SNR, and the wave velocities within the medium. The magnitude scale used is the moment magnitude. In our tests, we consider SNR=1 and Q-factor=75 in a homogeneous isotropic medium with 4.5 km/s and 2.5 km/s, as P and S wave velocities, respectively. For the sensor sensitivity we use 5.685×10^{-9} m/s which is the sensitivity of a commercial seismometer, commonly used for microseismic monitoring applications. The results of the magnitude sensitivity analysis are presented in Figures 5 & 6. This difference is acceptable considering the very small magnitude of the seismic event and how much smaller the number of surface receivers, used in the patch array geometry, is. Therefore, using a smaller number of surface stations, installed in a patch array design, can provide a cost-effective volume coverage compared to that provided by a large, 122-station array: A smaller number of receivers translates to reduced hardware cost and computation time (approximately 10 times lower) for data processing.

We also investigate the ability of MCMgs to accurately locate a seismic event recorded by arrays of the two aforementioned geometries. The purpose of these two models is to provide an example of the efficiency of the patch array design not only on the detectability of the events but also on locating the event. The models for the generation of synthetic data for this part of the analysis are homogeneous isotropic and their properties described in the previous sections. The size of the models is a cuboid with dimensions 4 km $\times$ 4 km $\times$ 2 km (N, E, D). For both arrays we have used the same single theoretical source location at a depth of 0.8 km and epicentral location 0.2 km west of the center of the model. We found that the location error of the two arrays is the same for all iterations. The major difference is the error on the origin time of the event which is higher for the coarse grid but it gets to zero value as we collapse the grid in iterations 2 and 3. The iteration time for the large rectangular array is about 10 times larger compared to the patch array. The magnitude of the location error between the patch and large arrays is identical for the size of this specific array. In order to understand if the aperture of the array has an impact on the location error, we increase linearly the distances among the receivers for both arrays. We repeat the same analysis for an array with a larger spacing (400 m) between neighbouring receivers for the rectangular large array and 4 km aperture, from centre to centre among groups on the patch array and 3 km between the central receivers of each group. The error differences between the two array

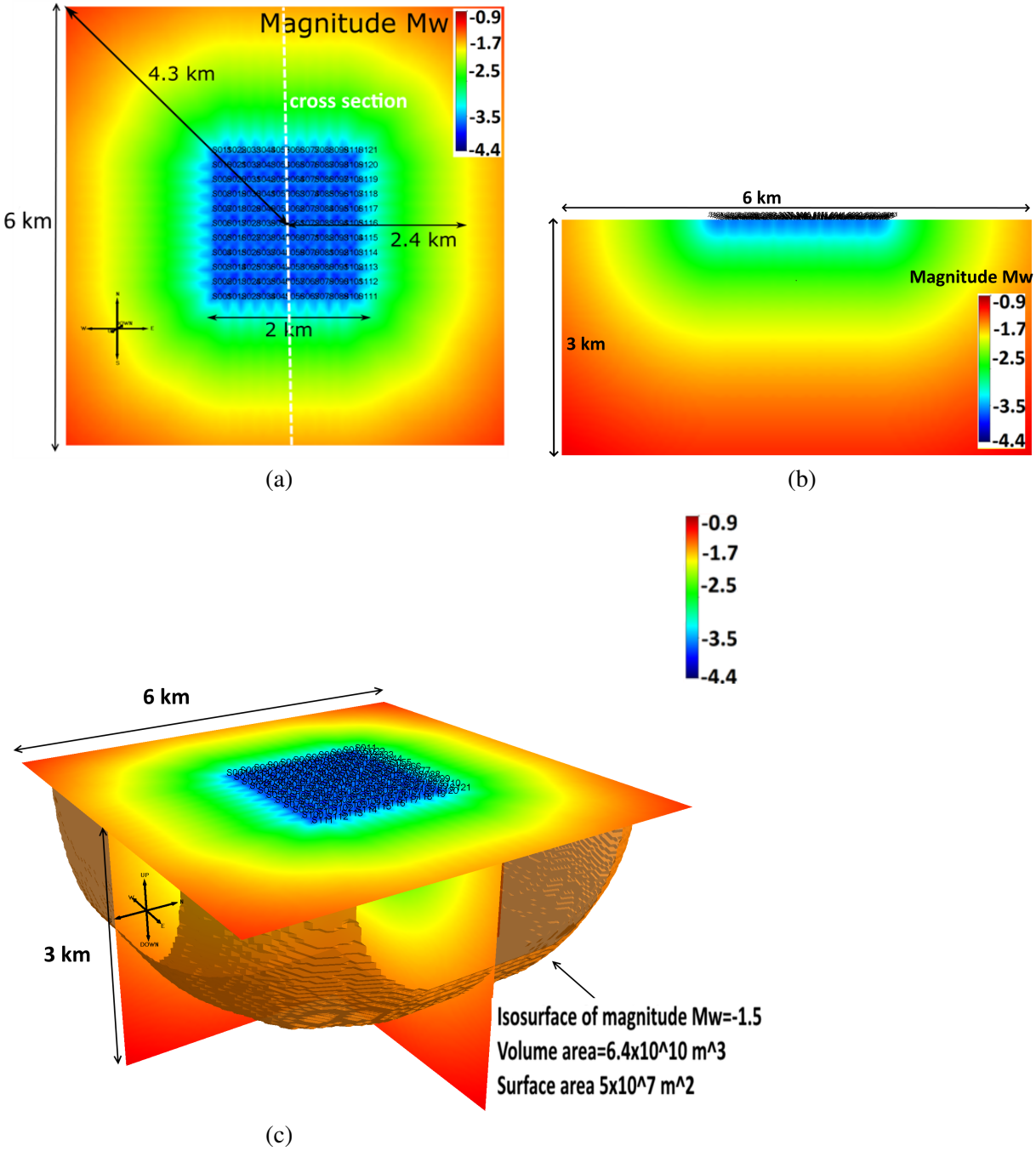

Figure 5. Plan view (a), cross-section North-South (b) and side view (c) of the magnitude sensitivity results for a surface array with 122 receivers in a grid of 200 m and the iso-surface of Mw-1.5 magnitude. The array can detect a signal of Mw-1.5 for depth up to 2200 m in minimum three receivers within a total rock mass volume of 64 km^3.

geometries (large array and patch array) are the same as previously, the results are summarised in Table 5. Consequently, the aperture of the array is not changing the effectiveness of the patch

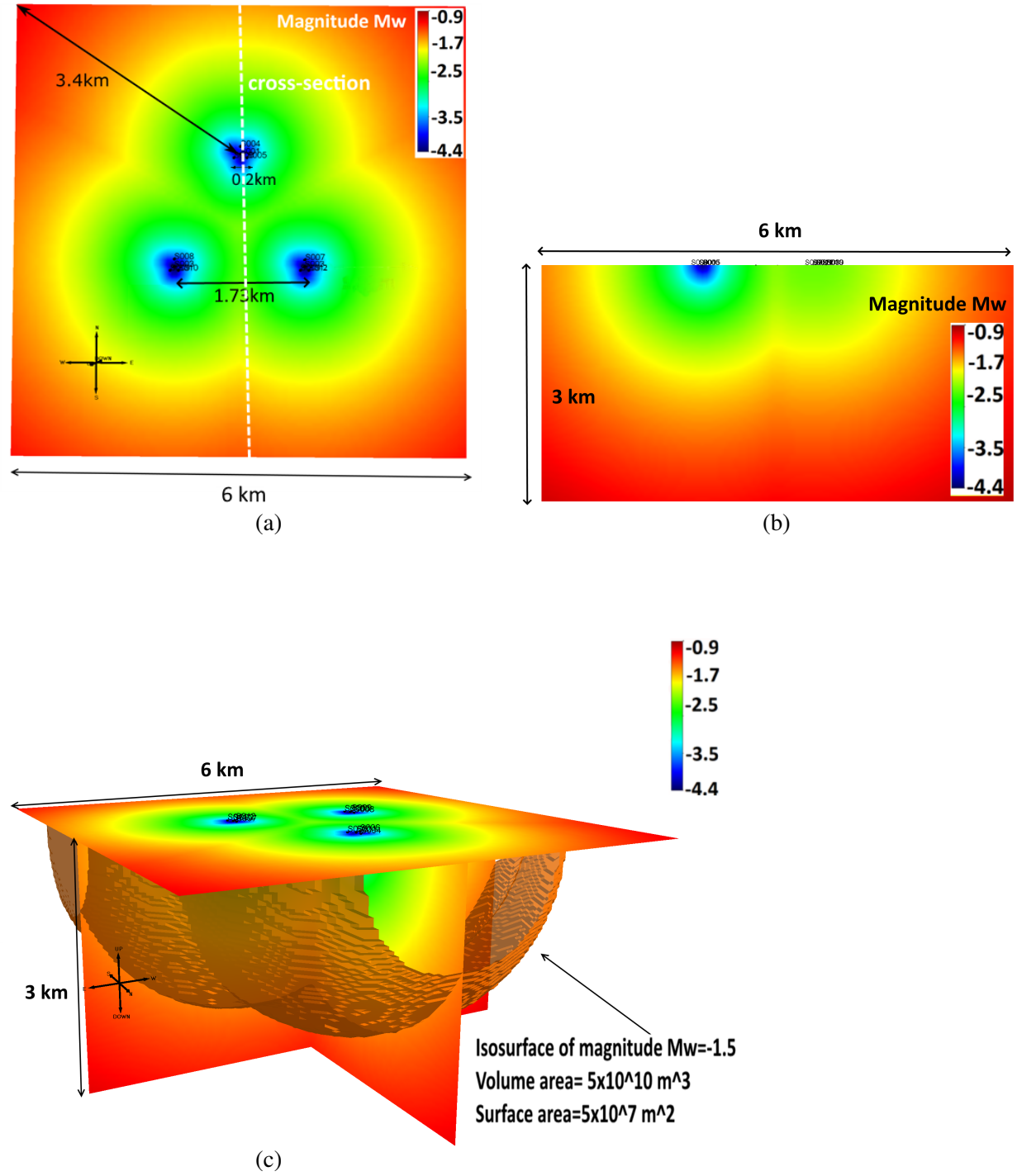

Figure 6. Plan view (a), cross-section North-South (b) and side view (c) of the magnitude sensitivity results for the patch array design with 12 surface receivers divided into 3 groups of small aperture (200 m), and a larger distance among the groups (1730 m) and the iso-surface of Mw-1.5 magnitude. The array can detect a signal of Mw-1.5 for depth up to 2100 m in minimum three receivers within a total rock mass volume of 50 km^3.

array against the large array. We show further evidence of the efficiency of MCMgs with real microseismicity data and support our findings on computational efficiency.

5 PERFORMANCE OF MCMGS WITH REAL DATA

The MCM algorithm has already been tested and proved effective in terms of localisation with real data as opposed to synthetic data (Shi, P. et al.,2019a; Shi, P. et al.,2019b). Here, we provide results obtained by MCMgs using a microseismicity data set from Greece and focus on the time efficiency of our method. The data we use were recorded by the National Observatory of Athens (NOA) network in Greece, and made available by the Aristotle University of Thessaloniki ((EIDA; Karakostas, V. et al.,2021)). The monitoring network consists of 13 seismometers, seven of which are part of the permanent network and 6 were deployed in the area after a major earthquake M_L=6.3 from 03/03/2021 in the area of Elassona-Larisa. We use a 24-hour segment of continuous recordings on 17/03/2021. The event catalogue published by NOA for this day contains detected and automatically located events using the HYPOINVERSE method. To locate the events we used the recordings from the 27 channels closest to the area of interest (nine 3-component receivers), while NOA has used data from up to 82 channels. For the location of each event, we cut the recordings into 45 s windows around the origin time of the event based on the NOA catalogue (15 s before the origin time and 30 s after) of each event and MCMgs is applied. The origin time is provided in the available waveforms. A detailed velocity model of the area, consisting of 5 layers

Table 5. Summary statistics of location errors and computation times for MCMgs and MCM patch array and large grid array (Figure 5 and 6).

	Maximum array aperture (m)	Minimum distance between stations	Median epicentral location error (m)	Median depth location error (m)	Computation time (s)
Patch array 1	1940	100	100	100	70
Large array 1	2000	200	100	100	760
Patch array 2	3880	200	110	200	70
Large array 2	4000	400	110	200	760

up to the depth of 29 km as described in (Karakostas, V. et al.,2021), is used. The average run-time of the MCMgs to locate an event using three iterations each time with 100 m target resolution was 60 s (using the workstation).

The epicentral locations of the 31 seismic events obtained by MCMgs are provided in Table A1. The local magnitudes shown in Table A1 are those from the published catalogue by NOA. We did not calculate the magnitudes of the events as part of this study. The differences between the epicentres of the catalogued events by NOA and MCMgs vary between 0 km to 3.3 km. These differences can be attributed to a number of factors, the main being the different velocity models used by NOA and by the present study. It should be noted that the differences between the epicentral locations (which fall within the error margins from both approaches) are not used here to show comparison between the methods or to support whether one is more accurate than the other since the true locations of real seismic events are never known. We use them to show the proximity in the epicentral locations obtained for the same event so that the computation time and comparisons of the number of recordings used are put into context.

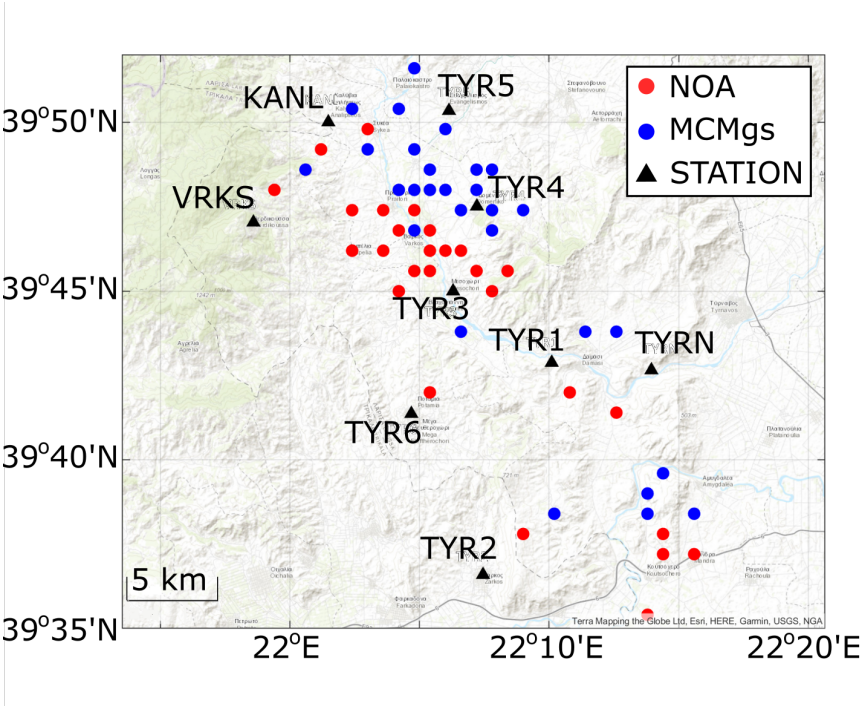

Figure 7. Epicentres of 31 events from the Larisa earthquake sequence: red circles - solutions provided by the National Observatory of Athens (NOA; red circles), blue circles - derived in this study using MCMgs

6 CONCLUSIONS

We present a location algorithm that has all the advantages of a favourable migration-based localisation method, i.e., the MCM location algorithm, but at the same time is more efficient in terms of computation time, by combining an existing method of a collapsing grid based on probability density function (PDF) (Lomax, A. et al.,2001). Faster results allow for quicker decision-making, which is important when it comes to safety and production in geo-energy projects. The MCMgs algorithm incorporates a modified collapsing grid approach that reduced computation time up to 64 times in our tests based on synthetic data while keeping the epicentral location error within an acceptable range. This renders our algorithm able to work in almost real-time using a simple workstation/laptop. For quick locations within an acceptable error (100 m), we find that 7 is the minimum number of recordings required (this error is based on the model and grid used in this study). When the number of available receivers is small, a patch array geometry is recommended. We find that 12 seismometers deployed in a patch array geometry offer a detection capability similar to a rectangular array consisting of > 100 receivers. Even though the grid array offers coverage over a somewhat larger volume, when considering the extra hardware and computation cost that such an array needs, a patch array with a significantly smaller number of receivers is a cost-effective alternative. MCMgs made possible the location of synthetic events with the same magnitude for the location error for both types of array geometries despite the large difference in the number of recordings used between the two, even for noisy recordings (SNR = 1). We locate 31 microseismic events using the MCMgs approach. The differences in the epicentral locations are within the location errors by both approaches (MCMgs and HYPOINVERSE). The MCMgs novelty is that it can provide locations in near-real time using a simple workstation with no need for an advanced computer and no phase picking which might be challenging for small magnitude and low SNR seismic events or events with short inter-event times. Further improvements on the run time of the algorithm could include the application of more advanced stochastic optimization algorithms to further speed up the processing time as described by (Li, L. et al.,2019).

ACKNOWLEDGMENTS

Prof E. Papadimitriou, Prof V. Karakostas and Dr T. Kostoglou from Aristotle University of Thessaloniki (Greece) are thanked for providing access to the microseismic dataset from the Elassona-Larisa seismic sequence used in this study.

This manuscript has significantly improved by the constructive feedback and comments of three anonymous reviewers, Editor Dr Carl Tape and Assistant Editor Fern Storey.

This work was funded by EPSRC Prosperity Partnership research and innovation programme under grant agreement EP/S005560/1.

DATA AVAILABILITY

All data used for the analysis in this paper are available through the University of Strathclyde https://doi.org/10.15129/a1d816a0-1849-4116-b510-da765fbffbdf and the code is freely available on GitHub https://github.com/eparast3/MCM_MAT/tree/MCM_gs

Epicentral locations of the 31 events used in this study as published in the catalogue by NOA can be found at http://bbnet.gein.noa.gr/HL/databases/database, using as search criteria Date: 17/03/2021, Latitude: 39.5-39.9, Longitude: 21.9-22.3. The waveforms of the 31 real seismic events were made available to us by the University of Thessaloniki. They could also be accessed upon request from http://www.orfeus-eu.org/data/eida/, using as search criteria stations: TYRN, TYR1, TYR2, TYR3, TYR4, TYR5 TYR6, VRKS, KANL and THL and date: 2021-03-17.

REFERENCES

Diehl, T., Kissling, E., Bormann, P. 2012. "Tutorial for consistent phase picking at local to regional distances. - In: Bormann, P. (Ed.) New Manual of Seismological Observatory Practice 2 (NMSOP-2)" *Potsdam : Deutsches GeoForschungsZentrum GFZ*, 75**5**, pp. 1-21.

Duncan, P. M., and L. Eisner, Reservoir characterisation using surface microseismic monitoring., *geop*, 75**5**, pp. 139.

EIDA, Accessed: 2021. [online]. Available: `http://www.orfeus-eu.org/data/eida/`.

Eisner, L., Hulsey, B. J., Duncan, P., Jurick, D., Werner, H., & Keller, W., 2010.,Comparison of surface and borehole locations of induced seismicity. *Geophysical Prospecting*, pp. 809-820.

Eisner, L. and A. De La Pena and S. Wessels and W. Barker, and W. Heigl, Why surface monitoring of microseismic events works., *Third Passive seismic Workshop. Athens: EAGE*

Geiger, L. 1912. "Probability method for the determination of earthquake epicenters from the arrival time only." *Bull. St. Louis Univ, 1912*.

Grigoli, F., Cesca, S. , Amoroso, O., Emolo, A., Zollo, A. & Dahm., T., 2013., Automated seismic event location by waveform coherence analysis., *Geophys. J Int.,* , 196**(3)**.

Grigoli, Francesco, Simone Cesca, Ortensia Amoroso, Antonio Emolo, Aldo Zollo, and Torsten Dahm. "Automated seismic event location by waveform coherence analysis." Geophysical Journal International 196, no. 3 (2014): 1742-1753.

Havskov, J. and Ottemöller, L. 2010., Routine Data Processing in Earthquake Seismology., pp. 299-326

Hildyard, M. W. 2007. "Manuel Rocha Medal Recipient Wave interaction with underground openings in fractured rock" *Rock Mechanics and Rock Engineering* 40**6**, pp. 531-561.

Hildyard, M.W., Daehnke, A., and Cundall, P.A. (1995), WAVE: A computer program for investigating elastodynamic issues in mining. *Proc. 35th U.S. Symp. on Rock Mech.* June 1995, Balkema, pp. 519-524.

Kaderli, J., McChesney, M. D., & Minkoff, S. E., 2018. A self-adjoint velocity-stress full-waveform inversion approach to microseismic source estimation., *Geophysics,* , 83**5**.

Kao, H., and S. J. Shan. 2004. The source-scanning algorithm: mapping the distribution of seismic sources in time and space. *Geophys. J Int.,* pp. 157 (2).

Karakostas, Vasileios and Papazachos, Costas and Papadimitriou, Eleftheria and Foumelis, Michael and Kiratzi, Anastasia and Pikridas, Christos and Kostoglou, Anastasios and Kkallas, Charalambos and Chatzis, Nikolaos and Bitharis, Stylianos and others. 2021 The March 2021 Tyrnavos, central Greece, doublet (Mw6. 3 and Mw6. 0): Aftershock relocation, faulting details, coseismic slip and deformation. *Bulletin of the Geological Society of Greece* 58 pp. 131–178.

Kennett, B. L., and M. S. Sambridge. 1992. Earthquake location—Genetic algorithms for teleseisms. *Phys. Earth Planet. Inter.,* 75 (1-3).

Klein, F., 2000. User's guide to HYPOINVERSE–2000, a Fortran program to solve earthquake locations and magnitudes. *US Geological Survey* Report 02–171 Version 1.0.

Larmat, C., R. Guyer, and P. Johnson. 2009. Tremor source location using time reversal: Selecting the appropriate imaging field. *Geophys. Res. L ett.,* 36 (22).

Li, L., J. Tan, B. Schwarz, F. Stanek, N. Poiata, P. Shi, L Diekmann, L. Eisner, and D. Gajewski. 2020. Recent Advances and Challenges of Waveform-Based Seismic Location Methods at Multiple Scales. *Rev. Geophys.,* 58.

Li L., J. Tan, Y. Xie, Y. Tan, J. Walda, Z. Zhao, D. Gajewski 2019. Waveform-based microseismic location using stochastic optimization algorithms: A parameter tuning workflow. *Computers & Geosciences* 124, pp.115-127

Li, J., L. Stankovic, S. Pytharouli, V. Stankovic. 2020. Automated Platform for Microseismic Signal Analysis: Denoising, Detection and Classification in slope stability studies. *IEEE Transactions on Geoscience and Remote Sensing* 59(9), pp. 7996-8006.

Lomax, A., Zollo, A., Capuano, P., Vireux, J., 2001. Precise, absolute earthquake location under Somma-Vesuvius volcano using a new three dimensional velocity model. *Geophys. J Int.,* 146(2), pp. 313-331

Maxwell, S. C., J. Rutledge, R. Jone, and M. Fehler. 2010. Petroleum reservoir charachterisation using downhole microseismic monitoring. *Geophysics,* 75 (5).

McClellan, J. H., L. Eisner, E. Liu, N. Iqbal, A. A. Al-Shuhail, and S. I. Kaka. 2018. Array processing in microseismic monitoring. *IEEE signal processing magazine* 99-111.

Napier, J.A.L., Daehnke, A., Dede, T., Hildyard, M.W., Kuijpers, J.S., Malan, D.F., Sellers, E.J., Turner, P.A. (1997), Quantification of stope fracture zone behaviour in deep level gold mines. *J. South African Inst. Min. Metall.* Vol. 97 (3), pp.119-134.

Nelson, G. D., J. E. Vidale 1990. Earthquake location by 3-D finite difference travel times. *Bull. Seism. Soc. A M.* 80 (2), pp. 395-410.

Pujol, J. 2004. Earthquake location tutorial: Graphical approach and approximate epicentral location techniques. *Seismological Research Letters* 75 (1).

Sambridge, M., and K. Mosegaard. 2002. Monte Carlo methods in geophysical inverse problems. *Rev. Geophys.,* 40 (3).

Schaff, D. P., G. H. R. Bokelmann, W. L. Ellworth, E. Zanzerkia, F. Waldhauser, and G. C. Beroza. 2004. Optimizing Correlation Techniques for Improved Earthquake Location. *Bull. seism. Soc. Am.,* 94 (2).

Shi, P., A. Nowacki, S. Rost, and D. Angus. 2019. Automated seismic waveform location using Multichannel Coherency Migration (MCM)—II. Application to induced and Coherency Migration (MCM)—II. Application to induced and. *Geophys. J Int.,* pp. 1608-1632.

Shi, P., D. Angus, S. Rost, A. Nowacki, and S. Yuan. 2019. Automated seismic waveform location using multichannel coherency migration (MCM)–I: theory. *Geophys. J Int.,* pp. 1842-1866.

Shi, P., F. Grigoli, F. Lanza, G. C. Beroza, , L. Scarabello, and S. Wiemer. 2022. MALMI: An Automated Earthquake Detection and Location Workflow Based on Machine Learning and Waveform Migration. *Seismo. Res. Lett.* pp. 2467-2483.

Steiner, B., E. H. Saenger, and S. M. Schmalholz. 2008. "Time reverse modeling of low-frequency microtremors: Application to hydrocarbon reservoir localization." *Geophys. Res. Lett.,* 35 (3).

Tarantola, A. & Valette, B., 1982. Inverse problems=quest for information, *J. Geophys.* 50 (159–170).

Yaowen Liu, Rugang Liao, Yuan Zhang, Dongwei Gao, Huaili Zhang, Ting Li, Chi Zhang, Application of

surface–downhole combined microseismic monitoring technology in the Fuling shale gas field and its enlightenment., *Natural Gas Industry B* , 4**1**, pp.62-67,

Waldhauser , F., and WL. Ellsworth . 2000. "A double-difference earthquake location algorithm: method and application to the Northern Hayward Fault,." *Bull. seism. Soc. Am.,* 90:1353-1368.

Wang, H., Li, M., Shang, X., 2016. Current developments on micro-seismic data processing. *Journal of Natural Gas Science and Engineering* 32, pp. 521-537.

Xu, J., W. Zhang, X. Chen, and Q. Guo. 2020. "An effective polarity correction method for microseismic migration-based location." *Geophysics,* 85 (4).

Zimmer, U., Microseismic design studies., *Geophysics,* , 76**6**, pp. wc17-wc25,

Zhu, T. 2014. "Time-reverse modelling of acoustic wave propagation in attenuating media." *Geophys. J Int.,* 197 (1).

Weiqiang Zhu, Alvin Brian Hou, Robert Yang, Avoy Datta, S Mostafa Mousavi, William L Ellsworth, Gregory C Beroza. 2022. QuakeFlow: a scalable machine-learning-based earthquake monitoring workflow with cloud computing, Geophysical Journal International, Volume 232, pp. 684–693