Comments by Jan Šílený on the PhD thesis “Source mechanisms of microseismic events induced by hydraulic fracturing” by František Staněk

The subject of the thesis by Mr. Staněk is highly topical. It is situated somewhere on the border between science and application, in particular between earthquake seismology and the seismology adapted to induced earthquakes, even more specifically between earthquakes and microearthquakes, nota bene those occurring during hydrofracturing for enhancement of the shale gas production. Dealing with earthquakes, we need to detect them, locate them, determine their mechanism, possibly the stress field responsible for their initiation etc. With microearthquakes related to industrial activities we need to do the same, moreover, there is additional task consisting in the necessity first to find them in the haze of eternal whispering of the environment, to dig them out of the background of seismic noise, in which – contrary to most earthquakes – we might not see them by naked eye at all. To get the information, to distinguish seismic events from random peaks of the noisy background, to locate them and even estimate their mechanism - that is a challenge with immense consequences for the practice. Mr. Staněk moved it up and did a very good job in the series of six papers gathered to form the thesis submitted. In the course of the work leading to the notable success, he benefitted from two lucky points: expertise of his supervisor Leo Eisner, who is undoubtedly one of the top experts in the field, and availability of the unique dataset of microseismicity induced during hydraulic fracturing of shale formation to exploit the natural gas, recorded by a very special surface network, which could hardly be mimicked outside the industrial environment. Then, to add a good deal of wit and hard-work – and that’s it!

Chapter 1 estimates theoretical resolution of the mechanism in terms of full moment tensor for three seismic networks used to monitor microseismicity: (1) a dense array of surface sensors distributed in a star-like configuration over the area of expected seismicity, (2) two vertical monitoring wells, and (3) set of shallow monitoring boreholes, each of them with several receivers. No data entered the game, resolving power of the network configuration for individual points of hypothetical hypocentra within the zone below the network was assessed. As a criterion, the authors used the condition number of the matrix in the core of the least-square solution of the linear inverse task. This is a good idea, as this number basically yields information on the stability of the solution, the trouble is that there is no receipt in theory concerning the threshold. In terms of comparing various parts of the zone of expected seismicity, it however yields useful information. In this respect, results obtained are valuable and provide a good insight what can be expected in different parts of the zone. The authors acknowledge that condition number is only one of the measures of the resolution for the mechanism and are mentioning other ones. I would like to add one comment to the issue, namely that the condition number is useful in general, nevertheless it tells about the resolution in general, i.e. concerning the mechanism “per se”, making no difference in individual features of the phenomenon. They used the full moment tensor to describe the mechanism and estimated which is the chance to reconstruct it in the inverse task. In fact, basically there are two pieces of information in the full moment tensor: the geometry, i.e. the orientation of the force system, and the contents of fundamental sources into which the moment tensor can be decomposed. It is a well-known fact that the resolution of both these aspects is different: while the geometry is amazingly stable, the decomposition is extremely vulnerable “to anything”. In my view, for practice it is of much more value to know more about something – one part of the solution, than
to know only a little about everything – the complete solution. The approach chosen by the authors provides the latter, in fact. My comment thus raises a more general question, namely whether in the tasks which are not well constrained (both due to limits in the station distribution as investigated in this chapter, and due to the other factors reducing the resolving power of the inverse task which are treated in chapter 2) it is reasonable to use the most comprehensive description of the mechanisms in terms of the full moment tensor. There are several alternatives providing similar versatility, basically the ability to capture both a shear slip and a volume change within the focus, which are at the same time simpler than the full moment tensor, though yielding a non-linear inverse task. I understand that within the processing chain proposed the evaluation of the mechanism must be quick, nevertheless for non-linear optimization there are fast procedures as well.

Chapter 2 investigates the stability of the inversion for the mechanism applying the observation from the surface (or effectively from the surface in the case of the array composed from shallow boreholes). Among other issues, it demonstrates excellently the amazing feature of noise reduction thanks to stacking performed implicitly within the least square solution of the inverse task. Stacking allows to gain a solution (approximate at least) from noisy records, where there would be no chance to get anything considering single records only. The level of the noise, which does not prevent to get an insight on the mechanism yet, is however the matter of debate, in my opinion. If I understood the description well, the way in which the authors perform the noise contamination of the synthetics in the synthetic experiments is rather tricky, I think. The point is that it does not simulate the processing in reality: here the noise is superimposed straight onto the synthetic amplitude and not onto the waveform which would be processed in the next step identically as the seismic record, i.e. starting with the search for the onset of the signal. In fact, the authors consider the arrival time known, regardless the signal in the noisy background would be seen or not. This is, in my view, a large simplification of the experiment, which makes the job of the inverse task much easier. The very optimistic results in Figs. 2.4 ab, displaying a very good resolution of the ISO, DC and CLVD within the mechanism up to a high noise level, should be considered with a caution.

In the flow of the microseismic data processing designed in Chapters 3 and 4, the correction of the stack function on the effect of the mechanism is crucial, as it rules the effectivity of the stacking towards reduction of a random noise. For this purpose, the mechanism needs to be estimated in each floating point within the area of expected seismicity (which may be rather large) and for a floating origin time. At the same time, determination of the mechanism is a rather delicate procedure, depending of numerous factors which may threaten the success of the job easily. I want to say that recovered mechanism may be distorted by a large error that we do not know a priori, and to be sure in the credibility of the mechanism, its confidence should be estimated. In my opinion, the way how the authors have done it in Chapter 1 is not enough. If the mechanism is not retrieved well, it predicts incorrect polarities/amplitudes, and the whole procedure of stacking fails. This is the technical point about the procedure designed. The conceptual one is the question, whether it makes a good sense to evaluate the mechanism in points which are very far from the hypocenter. My concern is what actually such mechanism represents, and which is the effect to correct the stack function to it. Does it really help? To be convinced, I would need to see some synthetic experimentation which the authors however have not performed; synthetic tests comprised in Chapter 2 deal with a fixed hypocenter.
In my view, more synthetic experiments would be needed to fully justify the statement from p.63 “Staněk et al. (2014) showed on synthetic data that it is possible to reliably invert the source mechanism even from a very noisy signal (i.e. of an individual seismogram SNR up to 0.05).” The reason for my objection is that they superimpose the noise directly on the amplitude and not on the waveform (see the comment above), thus avoiding, in fact, problems transformed into the detection and location. See also p. 84 “…and we know that the source mechanism is very robust methodology for large arrays as shown in Staněk et al. (2014). Therefore we may assume the source mechanisms used for semblance computation are correct.” Staněk et al. (2014) were not testing the effect of noise contamination on the semblance distortion.

The utilization of the image function (3.6) (and the semblance (3.7) as well) in the processing flow needs to make a choice of the thresholds for both; is this just a trial and error procedure or are there some a priori hints how to do that? And is this a compulsory step preceding the processing of any new dataset?

The idea of using the image function \( F_r(r) \) for construction of the probability density function and make advantage of it in the estimation of the location error is smart, is it however justified theoretically – is the squared \( (F_r(r)-\text{max}(F_r(r))) \) really an analogy of the likelihood function used for this purpose by Tarantola (2005)? And even if yes, I do not understand the argumentation “However, the accuracy of this approach is limited by the grid step of location zone. To overcome this problem, we transform the image function slice \( F_r(r) \) into a probability density distribution…” A rough grid step is the trouble always, in a rough grid there is no point to realize the formulas (3.10-12 and 3.13-15).

One more comment to the stacking: Waveform inversions for the mechanism involve by default waveform adjustment, i.e. the option to shift the synthetics with respect to the data a little bit back and forth and select the value with the best match. This is designed to accommodate possible small errors in the velocity model resulting in different kinematics. In the mechanism business this helps greatly. Evidently the stacking cannot use that and, in the consequence, will be more vulnerable to the effect mentioned.

The benchmark test in Chapter 3 presenting in parallel the manual and automated processing is very instructive. It yielded a difference in the depth of the events treated – is its size significant from the viewpoint of the practice? In other words, is it important to know, e.g., if the events have occurred above or below the horizontal wells, etc.? On the set of manually picked data some algorithm of a group location (JHD, DD,...) could be applied, promising an increased resolution; isn’t it worth of trying?

I cannot rest easy with the bedding plane slip model of the authors. In my view, the arguments in favor of the model are rather vague. Similarly its description. The horizontal planes are clear. So is the orientation of the vertical planes along the maximum horizontal stress. They are however described as “represented by poor continuity in the vertical direction within the more compact and uniform layer” (p. 117). According to the cartoon in Fig. C.11, they form a “block structure”; have these blocks some characteristic size? This would have to be reflected in the magnitude of the events released. The cartoon is a 2-D display, what happens along the 3rd coordinate? The model has a symmetry: left-right with respect to the vertical expanding rupture, and top-bottom as for the
upper and lower boundaries of the block. A slip on the bottom should redistribute the stress so that the top layer slips soon (implying an event of the opposite polarity), similarly on the other side in terms “left-right”. The events originated should be of similar magnitudes. Thus, doublets, possibly quadruplets. Is there a chance to check it by the observation?

Similarly I cannot rest easy with the other models of induced seismicity sketched in Fig. O.1, possibly except the model of diffusion. In my opinion, model of Rutledge et al. (2004) is merely a concept which, however, cannot work mechanically. A straight concatenation of a shear-slip and tensile crack cannot occur, because the stress evaluation on the tip of a shear fault displays the distribution of extensional and compressive stress that implies the tensile crack leaving the fault tip nearly perpendicularly. And where there are the “choke points” in this scheme, which should be represented by the microseismic events (p. 123)? Though not displayed in Fig. O.1, as an alternative the model by Eisner et al. (2010) is mentioned, based on “opening of a horizontal fracture overcoming vertical stress” with “microseismicity shearing almost vertical planes”. This, in my view, contradicts the fact of the maximum stress being just vertical. The last sketch in Fig. O.1 – d, presented to display the model of opening and closing – is just a mess. I doubt that the stress calculation would approve any part of it. The points gathered within this paragraph do not bear much relation to the topic of the thesis itself, nevertheless as a part of the text they should be presented clearly and precisely. In my opinion, there was no need to incorporate into the thesis the last chapter, the Outlook “Microseismic data interpretation – what do we need to measure first?”, and especially its final part O.2 “What can be done”, they are rather speculative contrary to the bulk of the thesis, which is well-designed and well done.

Minor issues

In Fig. 1.1 displaying condition numbers of the inversion applying the star-like surface array, I would expect an octagon in the map view, because of the symmetry with respect to each of the array wings. The result presented is more a square: why the resolution in the diagonal directions (far from the center) is worse than along the coordinate axes, if the length of all the wings is the same?

Figure 1.2 The pattern of the confidence number distribution should be symmetric with respect to the plane of both the monitoring wells. Is is the case in the map-view plot?

Figure 1.3 I am puzzled by the huge depth of well-resolved MTs (low confidence numbers) for the array of shallow boreholes: the excellent resolution extends to the depth comparable to the aperture of the array, while for the surface star-like array the resolution is good to about 1/3 of the aperture only (cf. Figs. 1.3 and 1.1, bottom parts).

p. 34 It is not clear why the authors use the omega angle – the difference between input and reconstructed orientation of the P-axis – as the measure of the deviation of the orientation of the mechanism. The usual measure is the Kagan angle (by the way, the author used it later, in the concluding chapter of the thesis).

p. 37 Dip-slip behaves worse in the experiments with the noise than the strike slip. In the Overview the authors mark this behavior even as surprising. In my opinion this is not a surprise, a worse resolution of dip-slip-type mechanisms in comparison with strike-slips has been observed earlier, and is related with the fact that the former yields smaller “variability” of the data than the latter: for a
vertical dip-slip the polarity pattern is bipolar, for an inclined one may be even unipolar (depending of the geometry of the monitoring), while for the vertical strike-slip it is quadrupolar.

Fig. 2.7 The scatter of the mechanisms for the strike slip type in the Hudson plot extends nicely along the line connecting compressional and extensional linear dipole; have you a hint for the reason? The dip slip mechanism does not display any regularity in the distribution.

Fig. 2.8 If the velocity model used to invert the noisy data is correct (contrary to the case in Fig. 2.7), the aligning of the mechanisms for the strike slip model along the linear dipole line in the Hudson plot disappears. It’s puzzling that this happens when more regularity is added into the inverse task!

p. 57, Fig. 3.6 Semblance validation criterion S>0.1 was used, but on p.51 the authors are recommending much higher values “close to 1”; this note relates to the point questioning the choice of the thresholds above.

p. 63 The sentence “A relatively high semblance means that the event is more likely real, but in general it does not mean that its mechanism can be reliably determined.” is not internally consistent, as the mechanisms are applied to evaluate the value of the semblance. In the case they are not reliable, the semblance is of little value as the detection tool.

p. 71 “Only when these amplitudes are corrected for the source radiation pattern can we achieve high semblance values. Furthermore, source mechanisms inverted for each potential location can serve as a prediction for the selection of traces with higher SNR and eliminating traces that do not contribute to the stack due to low signal.” Again this relates to my hesitation concerning the (unknown) reliability of the mechanisms evaluated without checking its quality.

p. 72-73 ditto: The correction of observed amplitudes by the radiation pattern is critically dependent on the accuracy of the mechanism determining, especially its nodal lines/planes.

p. 74 Two-step procedure of eliminating noisy records: “…and afterwards to remove noisy traces we select the receivers with small difference between observed and expected amplitudes (i.e. small difference between $ar$ and $br$ in equation 4.3)”. My question is what is used for the $ar$ – some estimate of the level of the signal? The two-step procedure is described twice (p. 74 vs. p. 77) in slightly different way (or it seems so, at least), the former seems to be more convincing, the latter is doubtful. In the 2nd step records strongly differing from the predicted amplitudes are discarded. This however may work improperly if the predicted amplitude is not OK by an inexact model. Then, it is different from the data even without noise, and elimination of the channel is undesirable.

p. 78, Fig. 4.4 The semblance for the synthetics with noise level 1 is smaller than for the field data: it means that the field data are of better quality than the synthetics with the lowest level applied?

p. 80, Fig. 4.5 The threshold lines for the STA/LTA are mostly not visible.

p. 80 Speculation about the detection using combining the stack, STA/LTA and semblance: again the trouble with selecting the thresholds for all the three functions – any hint? The choice $S>0.17$ has been done based on the synthetic tests, i.e. the Fig. 4.4, but there does not seem to be resolution enough for such a detail ($S$ between 0.1 and 0.2 – OK, but why just 0.17?)}
p. 80, Fig. 4.6 “Fig. 4.6 shows that with a lower semblance threshold we get events with high location error and we also get random source mechanism orientation inconsistent with orientation of strong events.” Fig. 4.6 obviously documents the field data processing, then the criterion of a consistency of the mechanisms of strong and weak events need not be very helpful – they may be different depending on the size of the tectonics.

p. 84 “The most useful receivers in this sense are those that are consistent with some type of seismic source mechanism...” This is a very misleading formulation.

p. 88 “Locations and source mechanisms allow building many reservoir models such as discrete fracture network, stimulated rock volume or propped volume.” Are the last two items really reservoir models?

p. 88 “...non-shear components usually result from noise in data or mismodeling (Staněk and Eisner, 2013). Therefore, we use the shear component of the source mechanism.” To my knowledge, stress inversion algorithms work with the shear part of the mechanism by default, regardless the (un)certainty of the non-shear components.

Fig. 5.1 In the Group 2 there are less reliable mechanisms (p. 91), nevertheless the jackknife test indicates a smaller error in the stress retrieval than for the Group 1!

p. 92 “The minimum horizontal stress \( \sigma_3 \) is the most stable. ... Perhaps, only the elongated uncertainty of the vertical stress may be an indication of the spatial or temporal variation of the stress field.” If \( \sigma_3 \) is stable, both \( \sigma_1 \) and \( \sigma_2 \) must exhibit some elongation as the triple is perpendicular mutually.

p. 92 “An interesting observation is that strike-slips and dip-slips are consistent with a single stress field.” How much the strike slips are consistent, it is not documented: deviations of the predicted stress vector vs. slip vectors yielded by the mechanisms would have to be presented for individual events.

p. 92 “Another surprising result is the fact that a large number of weak events ... do not cause any significant deviation or scatter in the inverted stress field orientation.” This depends of the number of jackknife trials and their “statistics” – how numerous are the trials with minimum number of events. Fig. 5.1 hints that with Group 2 there were not markedly more trials than with Group 1, regardless the notable difference in number of events in each of them.

p. 100 The statement “less steeply dipping planes” would need some quantification.

p. 107 evaluation of the Green’s function by rays; reference to Staněk et al. (2014), i.e., p. 29, but there not documented: which program package was used?

p. 108 “The strike-slip events seem to have their shear components oriented more favorably toward the regional stress field.” This is not obvious form Fig. C.5. Strike angles of the strike-slips seem to coincide with the regional stress, which does not imply a favorable orientation (it is about 30°, depending of the friction on the fault). If they are so tightly coinciding with the regional stress, the pattern “left-lateral vs. right-lateral” should be checked for them, if this fits
the property “strike < reg. stress azimuth vs. strike > reg. stress azimuth”. From the plot on the right of Fig. C.5 this cannot be easily assessed.

p. 110, Fig. C.6 It would be desirable to present also signs of the ISO and CLVD components, because they are useful indicators of their “confidence”. For sources that physically represent an opening/closing within the focus they are of the same sign, whereas opposite signs hints they are probably spurious.

p. 112 Discussion “FM vs. DC”: misfit amounting 2% hints that the difference is small, however the number 2 is obviously a random measure. To be more correct, the F-test should be performed.

p. 118 “Woodford shale and generally all successfully producing shale formations are considered overpressured reservoirs (e.g., Agrawal, 2009); therefore, the minimum horizontal stress is close to the vertical stress.” Is this consistent with Fig. 5.1? The maximum stress is vertical there.

p. 118 “The shear motions may also occur along the two faces of hydraulic fractures due to the pressure applied to materials with slightly different stiffness parameters, resulting in stress discontinuity.” This means also elsewhere than the arrows are marked in the cartoon Fig. C.11?

Typos


p. 29 extend → extent

p. 39 Fig. 3.7 → Fig. 2.7

p. 45 “3d of data” - it means three days of data?

p. 49 eq (1) → eq (3.1)

p. 84 methodology → methodology; semplance → semblance; recievers → receivers

p. 89 constrain → constraint

p. 90 “…consisting of almost thousand vertical geophone groups installed around the wellhead in ten lines…”


p. 123 extend → extent; An Alternative model → An alternative model

p. 125 to trace (chemical or other the injected fluids → to trace (chemical or other) the injected fluids

p. 126 bullet No. 4 is a mistyping

p. 139 stick split → stick slip (?)
Závěr:

Disertační práce plně prokazuje předpoklady autora k samostatné tvořivé práci. Doporučuji proto udělit mu po její úspěšné obhajobě vědecký titul PhD.

V Praze dne 31.7.2018

RNDr. Jan Šílený, CSc.
GFÚ AV ČR