General comment on the Manuscript
A Multi-Technology Analysis of the 2017 North Korean Nuclear Test
by Gaebler et al.

The paper summarizes different analysis and observations concerning the September 2017 nuclear event in North Korea. It illuminates the event from the perspectives given by the different geophysical methods, that are used for monitoring of nuclear explosions.

First, the authors give an introduction on the need of monitoring nuclear explosions. The largest part of the paper is dedicated to seismological observations and modeling, which comprise 5 subsections. One section is about infrasound observations and modeling, one section is on the surface deformation and the last section on propagation of radionuclides. Except for the section on surface deformations, all sections comprise both, theoretical modeling as well as analysis based on observations.

The described seismological methods provide a refined epicenter location, an depth estimate, moment tensor solution as well as a yield estimate from the retrieved magnitude. The topography effect has been modeled, showing an focusing of the seismic energy to smaller slownesses. Infrasound signals have been modeled and detected, surface displacement retrieved from DinSAR-method and atmospheric modeling performed to explain detections of radionuclides.

The paper gives a nice overview on all different aspects. However, in some details it appears to be not mature but has a lot of potential for improvements. The paper, especially the seismological/modeling analysis, appears to be incoherent in some aspects. It seems to be composed of pieces which are put one next to each other but which do not take into account the results from each other . All individual aspects are interesting and definitively worth publishing.

One can resolve most problems without additional analysis, but by discussion and mentioning the discrepancies, which arise due to the different focus of the different methods. For example the depth estimate of the moment tensor inversion is 2 km, while the pP depth-phase depth estimate is 600 m. This is no surprise, since both methods have different foci, but it should be mentioned and discussed in more detail. As other example, the topography effect is modeled and discussed qualitatively, but the influence on the depth phase estimate is not investigated. By discussing the modeling more quantitatively, one could possibly make arguments on slowness and frequency ranges, for which this focusing effect is relevant. It should be discussed if this has an effect on the other described methods (moment tensor solution, depth estimate, magnitude estimation by regional vs. teleseismic phases.) At other points, the paper should display the observations, that the described analysis are based on. For example the input parameter for the double difference method in section 2.1 could be displayed instead of the coherency value, which is not so relevant for the analysis. As other example it would be nice to show the wave-forms, that the magnitude estimations in section 2.4 is based on, to demonstrate the differences in high- and low-frequency magnitude estimations.

In general I would appeal to the authors to keep in mind the reproducibility/traceability of their results. To achieve this, the level of details in the description has to be appropriate (which, of cause, can also be achieved by referring the reader to the relevant references).

In conclusion, I would recommend the paper for publication after MAJOR REVISIONS. Reviewing this paper was quite some work, but also very interesting. All comments are meant to be constructive in order to improve the manuscript. In general, I have no doubt in the relevance of this work, which is why it appears to me to be worth the work that still should be done.

List of Issues (sorted chronologically)
Citations from the paper are italic. Sometimes I corrected or commented in the citations directly. Most of the time my comment/suggestion are enclosed, indented with a dot and in red upright fonts.

Abstract

Page 1

Line 9:

- Please clarify:
  Do you mean Seismological investigations of depth phases?
  Then it must be 0.6 km below surface instead of 0.8 km (otherwise it is in contradiction with the results in 2.2) In this case I would include "of depth phases" as well in the text.

Line 16:

- ehance -> enhance

Section 1

Page 2:

Line 3:

- *It was opened* -> CTBT was opened (since in the sentence before you talk about CTBT and CTBTO)

Line 10:

- Place a comma after *At the time of this study*

Page 3:

- Line 11 – 25: Paragraph should be resorted and corrected (shift one sentence and 3 typos):

  In this context, new methodical approaches are introduced to improve the depth estimate from teleseismic observations and to quantify uncertainties in the non-isotropic source component. ~~Radionuclide monitoring demonstrates the importance of atmospheric transport modelling (ATM) to avoid over-interpretation of variations in 133 Xe concentrations.~~ The seismological study retrieves an independent absolute location based on a combination between seismological and remote sensing data. The relative location between the six North Korean nuclear tests is obtained by means of waveform cross correlation time lag data. The event depth is estimated for the first time by a joint inversion of source time function (STF) and depth phase waveform modeling observed at small aperture, high-frequency arrays in ~~teleseimsic~~teleseismic distances. A full waveform moment tensor inversion (MTI) is applied and compared to results of previous explosions on North Korea. The source time overshoot and peak amplitude is compared to traditional body wave magnitudes (mb) estimates. Additional~~l~~, the effect of topography on peak amplitudes is estimated by 2D-waveform

This phrase has to be modified. You do not invert the depth phase waveform modeling.

You do joint inversion of the STF and the waveform composition of direct and depth phase to retrieve the depth.

Unknown Author
10/16/2018 13:25

modeling to estimate the possible range of explosive yield. Infrasound observations and modeling is used to understand the earth-atmosphere coupling and the propagation of infrasound from the North Korean test site;. The analysis of satellite based remote sensing data is important to improve the absolute location of shallow sources as well as to quantify the secondary mass movement effects at the surface. Radionuclide monitoring demonstrates the importance of atmospheric transport modeling (ATM) to avoid over-interpretation of variations in 133 Xe concentrations. While radionuclide monitoring provides the only direct evidence of nuclear explosions, it is demonstrated by careful modeling how difficult it is to interpret such data and that early claims of causal anomalies have possibly been over-interpreted.

Section 2.1

Page 4:
Line 15:

- I do not understand the relevance of Figure 2 for the described analysis. It shows the similarity of the events. However, as input parameter for the double difference method one needs the lag times. It would be good to find a way to show the lag time data, for the different events, maybe in a similar way as the correlation coefficient is shown.

Section 2.2

Line 29:

The analysis of the time lag of near-source, surface-reflected P-phases, so-called depth phases, can potentially help in such a case, because they only depend on the depth and the P-wave velocity in layer above the source.

- I found this explanation misleading. In teleseismic distances one cannot resolve a depth phase or a depth phase lag time in the seismograms. And looking at your array beams one cannot identify any separated depth phase, since the source time function is much longer than the lag time. What you are doing is modeling the wavelet under the assumption that is consist of source time function (STF) and an reflected phase, which resembles the inverted and time shifted STF. However, in the paragraph it sounds like you are able to identify the depth phases by applying beamforming.

- A very similar approach as been applied in the section ARRAY BEAM MODELING of Cesca et al (2017). Please cite it here. In that publications of the 2016 nuclear explosion, arrays from different azimuths had been used (ASAR, GERES and PDAR). Why those data were not utilized for the inversion shown in this paper?

Page 5:

Figure 3:

- The figure is not displayed right in the SE-manuscript (`se-2018-102.pdf`). However I found a proper displayed one in `se-2018-102-manuscript-version1.pdf`. Please make sure, that it is right in the final version.

- It would be good to show the residual fit to deeper depths, since it seems like the residual decreases again from 900 to 1000m. Also 2 km is the depth estimate from MTI and this result should be in the range of the refined method. So best would be to show the residual curve down to at least 4 km.

Line 5:

- What is *the best moment tensor solution*? How is it determined? Is it the solution derived later in Section 2.3? Please clarify in the text.

Line 6:

- The abbreviation STF is used only in the introduction. For better readability I would define it here again as "source time function (STF)"

Line 12 ff:

It is assumed that all moment tensor components Mjk have the same time dependency, which is described as normalized STF m(t) with m(t → ∞) = 1. The waveform of far-field displacement pulses are controlled by the time derivative of m(t), which is declared as moment rate function ṁ(t). For a ~~The~~ P-wave from an earthquake, ṁ(t) ~~has~~is a single-sided pulse.

In the figure caption of Fig.3 you call m(t) moment function. Here you call it normalized STF. Please be consistent.
Unknown Author
10/16/2018 14:39

Please give a citation for this statements
Unknown Author
10/16/2018 14:45

Please give a citation or explanation for this statement.
Unknown Author
10/16/2018 14:48

Line 16:

- Green function -> Green's function (2 times)

Line 23:

The convolution in (3) can be ~~equated and~~ written in discrete form,...

Page 6:

Line 5, 7 and 9:

- Green function -> Green's function

Line 11 ff:

AK135 (Kennett et al., 1995) is used for the mantle, while different crustal models for source and station region are taken from CRUST2.0 (Bassin et al., 2000). Intrinsic attenuation for P-waves is set to 5000, since otherwise high frequencies are damped out at teleseismic distances. The sampling frequency is 20 Hz. The grid search depth phase modeling has been applied previously to different cases of induced seismicity (Dahm et al., 2007) and the 2016 nuclear explosion (Cesca et al. 2017), but the simultaneous STF inversion ~~is~~was implemented for the first time for this study.

Is this realistic? Please comment why you are allowed to do that. If not the question arises, what are you inverting for?
Unknown Author
10/16/2018 14:54

Line 20 ff:

Interesting are the retrieved moment rate and moment function, which show a clear double pulse and overshoot, respectively (Figure 3c). Such an overshoot is not expected for the rupture process of tectonic earthquakes, where the moment rate and moment functions of the P-wave are single sided pulses and monotonously increasing functions, respectively.~~,~~ For nuclear explosions, however, this feature is ~~, but~~ commonly observed. ~~for nuclear explosions. It~~ This can be explained by at least a partial collapse of the explosion cavity immediately after the explosion. In this case, the final moment is only 23 % of the peak moment in Figure 3c. The inversion of long-period waves led to a seismic moment M0 of 2.33×1017 Nm, representing the time after the overshoot. The peak seismic moment is thus estimated in the range of M0,peak ≈1.02×1018 Nm. The associated (peak) low and high frequency moment magnitudes (MW) are MW =5.55 and MW =5.97, respectively, and can explain the large difference between the long period MW and the high-frequency mb estimates which emerged during the magnitude estimation described in Subsection 2.4.

Section 2.3

Page 7:

Line 11:

The best solution, found at a depth of around 2 km, has a scalar moment of 2.33×1017 Nm, equivalent to a MW of 5.55.

This is inconsistent with section 2.2, where the best depth is at 600 m. Comment on that.
Unknown Author
10/18/2018 14:24

Figure 5:

- How the shown wave forms are selected? Do you only show those stations where the fit is good? Why do you show less wave forms in Fig. 7 compared to Fig. 5?
  It would be good to show the same example stations in both cases or explain why you select the given stations. Please highlight the selected stations on a station map (e.g. in Fig. 1).

Figure 6:

- It would be good arrange the Figure in two parts a) and b), a) showing the decomposition of the Moment tensor and b) showing the source type diagram. (same in Fig.8)
- I would find it very useful to explain roughly the Hudson-diagram, since it might be common sense for experts on moment tensor decomposition, however the paper is for a

broader audience. For example all moment tensor solutions (beach balls) in the Hudson diagram represent double-couple solutions, while the best fit solution is mainly isotropic. Could you please comment on the representation and how one should read it?

- How was the ensemble of moment tensor solution retrieved? Do you use a statistical Monte-Carlo-type inversion? Give some more details and/or references in order to present reproducible results.

Figure 7: (see Fig.5)

Figure 8: (see Fig. 6)

Section 2.4

- You do not show any seismic section of the event. This is a pity. Here would be a good position to include one (maybe compare it to with waveforms of a previous test) and show the peak values that you are using to estimate mb. Show the 15 IMS stations you are using here on a map or give at least a reference, where one can find them.

Page 8:

As no particular magnitude-yield relation has been approved so far for the North Korean test area, the latter relation by Bowers et al. (2001) is used in this study, as it supposedly most accurately represents the geological conditions at the test site.

- Can you please comment on why the geological setting is supposed to be similar to Nova Zemlya? More important, is the North Korean test area known to consist of “wet hard rock”, rather than of “dry unconsolidated rock”? It changes the interpretation dramatically (by one order of magnitude). Give some references.

Section 2.5

Figure 10: The separation of P- and S-waves (and displaying in the same figure) is very beautiful.

The absolute transmitted energy is in all cases the same. A focusing/defocusing effect to certain slowness ranges is modeled. The conclusion is, that a focusing of the energy to small slownesses might result in an overestimation of the magnitude.

- For the discussion of this point I would ask the authors to be more quantitative. Give numbers: For which slowness range the energy is increased/decreased. Is it relevant only for teleseismic records or also for regional ones. Could a focusing to small slownesses also lead to an underestimation of the magnitude when using local/regional stations?

- Please discuss on the frequency ranges that are affected by the topography: For which frequency ranges the focusing effect is relevant? The lens (concave mirror) of the mountain has a range of approx. 2 km. For a P-wave velocity of 5.7 km/s wavelengths of the same dimension have a frequency of 2.85 Hz. Thus, I would expect, that the focusing is effective only for high frequencies (> 2.5 Hz). Will this still affect the magnitude estimation from teleseismic records?

- You show the topography effect in this chapter, however, you do not consider the topography in any of your analysis in the other chapters, especially for modeling the depth phases. Would the topography effect also effect the depth estimation by using Pp phases? Please comment on that, since otherwise the paper appears to be incoherent.

Page 9

Lines 6ff:

In general, this numerical modeling gives indications for: (1) Clear infrasonic signals, because surface reflections with higher amplitudes correspond to transmitted amplitudes with higher amplitudes yielding a transmission coefficient greater than 2 (see Section 3).

Higher than what?
Unknown Author
10/18/2018 21:04

Where in section 3 are you taking about transmission coefficients?
Unknown Author
10/18/2018 21:04

- For me, it is not clear how you relate the pure elastic modeling to the elasto-acoustic transmission coefficient.

Section 3

Page 9:

Line23:

Figure 11a highlights the waveform beam of the Russian 4-element infrasound array I45RU (denoted as the co-located seismic station USRK in Figure 1a)

- Please give here or in the next sentence the distance of I45RU to the test-site.
- It is surprising, that the It phase shows the strongest amplitudes. Is this in agreement with the modeling?

Figure 11:

- The color code of the modeled ray travel-time is not chosen well to show differences in travel time of It and Is2.

Lines 11-16:

The eigenrays show that the stratospheric phases are about 170 s and 70 s faster than the thermospheric phase (caused by a higher effective sound speed in the stratosphere and the longer wave path for the thermospheric phase).
The signal attenuation indicates that only a small portion of signal energy is ducted in the stratosphere caused by partial reflections from gravity wave variations of the stratospheric mean background. This leads to higher attenuation in the stratospheric duct and thus stronger signal amplitudes in the thermospheric duct, which corresponds to the observed waveforms.

I can not see this in Fig 11 (see above)
Unknown Author
10/18/2018 18:13

Please explain this point in a bit more detail. This is interesting. Why is the amplitude of It even larger than Is2?
Unknown Author
10/18/2018 18:23

Lines 18-20:

Apart from the strong epicentral surface movement, infrasonic signatures were also identified from seismo-acoustic coupling and the assumed cavity collapse associated to the eight minute subsequent aftershock.

This is interesting. In 400km distance? Can you show these data? (maybe put it in the supplement)
Unknown Author
10/18/2018 18:18

Section 4

Page 11

Lines 13-14:

Figure 12 shows the surface deformation ~~restricted to~~at the test site area ~~after~~during the 2017 test ~~for~~restricted to pixels with ~~a~~ coherence values ~~of~~ greater than 0.25.

Figure 12:

- Displacement scale should have more numbers. (You only show >10 and <10). Please add at least a 0 and +/-5.
- Please add the legend, explaining the colors of numbered circles, which is present in Fig 1b), to avoid misinterpretation of these colors with displacement values.

Line 18-20:

Due to the incidence (43°) and lock angle (ENE) of the sensors and the calculated slope and aspect angle (20 to 27° facing SW, 10° facing NE) based on Shuttle Radar Topography Mission data, 30 to 80 % of the vertical measured displacements in the area are detected.

30-80% is a very wide range. I do not understand the statement. Please clarify.
Unknown Author
10/18/2018 18:40

Line 20-27: (two comments + two typos)

The resulting displacement map clearly shows an area of subsidence of up to 10 cm around 3 km north of the tunnel entrance (compare Figure 1b), and an area with clear uplift of up to 10 cm west of the Mt. Mantap peak. DInSAR processing of C-Band Sentinel data and TerraSAR-X data for the 2017 test did not reveal acceptable quality measures regarding coherence in the central part of Mt. Matap, whereas for the tests in 2016 Sentinel data did show excellent results (Hartmann et al., 2017; Wei, 2017). A study by Wang et al. (2018) shows additional results based on amplitude images from TerraSAR-X data and calculates 0.5 m subsidence in the center and 3.5 m of horizontal displacement.

I could not find the tunnel entrance on Fig 12 of Fig 1b. Please mark it in Figure 12
Unknown Author
10/18/2018 18:52

Please comment, why your processing of TerraSAR-X did not work here, while Wang et al was able to calculate the displacements.
Unknown Author
10/18/2018 18:58

Lines 27-32:

To validate the displacement maps of the DInSAR analysis of the nuclear test, Pleiades data sets from August 26th and from September 8th 2017 were processed to reveal surface characteristics *related to test (Figure 13). Change detection analysis show numerous landslides activated during the test and aftershocks.*

- How does the results validate each other? For the reader it is not clear how they are related. Please show the locations of the purple patches also in map view to make the results comparable.

Line 32-33:

As a result of the processing of the ALOS-2 data, an area of around 3×4 km2 can be delineated, where surface movement rates range between -10 and 10 cm.

- Is this result of relevance for the conclusions or discussion? Can you show these results? Where is this area located?

Section 5

Page 12

During October three peaks containing five samples with 133 Xe activity concentrations between 0.5 and 1 mBq/m3 were measured at the station RN58, which went back to operation in between, at the days indicated by the forward simulations.

- Please describe the data that you want to explain by modeling at the begin of the section. Both data-sets with positive detections should be described, the one from South Korea that you want to disprove coming from the test site, as well as the one from RN58, that might be explained by a late release from the test site.
- What is the exact timing of the detections at RN58?
- Why did you choose October 4th and not any other day in October to run your simulations?

Section 6

Page 13:

Line 14:
The depth phase modeling of the P-waveform was performed at 2.5 Hz and found a centroid depth between 400 m and 800 m, which is within the expected resolution of about half the wavelength.

- What do you want to state here?
  Seismograms have been filtered between 0.5 and 2.5 Hz (See caption Fig. 3). You do not show the spectral content of the signal. Clearly 2.5 Hz is only the upper limit. If 400 m is half a wavelength, lambda=800 and c=lambda*f = 800*2.5=2 km/s for the upper limit and c=800*0.5=400m/s for the lower limit. This is too low for P-waves velocities, so I do not get the point here.

Page 14:

Line 2 ff: (2 typos, one inclusion)
Following the main event, around eight minutes later, a large aftershock of local magnitude mL 4.1 was detected. MTI analysis of this event shows a clear negative isotropic part which leads to the conclusion that this ~~is~~ event is caused by at least a partial collapse of the cavity formed by the main explosion.

Section 7

Page 15:

Line 22:

• *Explosive character of the September 3rd 2017 North Korean event is confirmed by cross correlation and MTI analysis.*

- How does cross correlation confirm the explosive character of the event?

Line 26 f.:

• *The estimated yield of the nuclear device is certainly smaller than the largest documented yield ever achieved by a boosted fission device and is therefore still compatible with a fission only device.*

- On the other hand: would it also be compatible with a "small" fusion bomb?

Line 28-29:

• *Strong surface deformations (±10 cm) are observed in an area of 3×4 km2. Furthermore multiple landslides as well as a number of aftershocks were observed in the aftermath of the test.*

- This is the observation. What is here the conclusion? → either skip this point or add a conclusion to the observation.

Page 16

Lines 3-4:

• *Due to the non-operating radionuclide station RN38, ~~N~~no immediate measurements of radionuclides related to the test in September were observable. ~~, but~~ However, later occurrences of radionuclides are consistent with a delayed leakage from the test site in October.*

Line 5:

• *The test site might have been strongly stressed and shattered and thus ~~might be~~ rendered useless for further test activities.*