Reviewer A Comments

Realtime selection of **optimal Source Parameters** using **ground motion envelopes**

A- Topic: (As i understand)
Seismic networks may have multiple independent algorithms to locate events (EEW and/or "standard" locations systems)
The authors are proposing an approach different from some network adhoc rules used to qualify or prefer a solution from another.
This method combines (1) a comparison of Predicted Vs Observed event waveform envelopes peak amplitudes with (2) their cross-correlation fit.
The method proposed is assumed to provide:
(a) an independent single solution assessment for real-time alerting discrimination specifically on mislocated, quarry or non existing events.
(b) a general/independent comparison of location algorithms solutions with regards to a reference model of predicted envelopes for an area at a seismic station.
The authors test and discuss the results over a sample of large events, quarry blasts and "false" or mislocated detections within the Swiss national network.
B- General comments
This paper is interesting, has some content and I could go through and understand the approach and the motives.
i had a bit more difficulties with some figures-tables.
I would propose this paper for "revisions". According to me, this manuscript would benefit from
(a) Some minor additions or clarifications in some paragraph and parts
(b) Some clarity, simplification or changes in the figures.
(c) Some more discussions/conlusions about for instance requirements/limitations such a data/metadata quality of a seismic monitoring networks/systems to operate the method and/or simplification or improvement of the envelope calibration.
(e) One maybe suggestion in the text/conclusion etc: planning future test would it be worth to experiment with "usual/normal" seismicity?
(f) The method is assumed possible to be ported to other seismic networks. How does the method handle depth? Should that be tested later on?

Question more general: Were all or some of the large events used to model the predicted envelopes?

Comments/suggestions/questions:
if the authors agrees and the reviewer understood…

L28: Add "..rapidly estimate *developing* ground shaking.. " for clarity ?
L46/50: Perhaps more emphasize about the pro/cons from multiple detections process frm a seismic network?
L54-58: Minson et al. description for clarity perhaps splitting into 2 sentences?
The method is described as complex.
L62: Minson et al is adapted/modified by the authors:
Perhaps slightly more emphasize about the benefits of using a "complex" method and turn it into a

more “simple” one.
L66 perhaps name *“both”* solution origin process (finder etc..) ?
L72-76: The motives could be earlier in the paragraph?
L80: Perhaps precise location “blunder”?

L84: Perhaps Clarify the steps: the author started with a “first step” statement, though does not clearly state the next ones.
L90 Question: Is there one max amplitude per 1 second interval ?
L99 equation (2) has a typo : should be A 😊
L106 I guess it is …“good *amplitude* fits for certain …”.
L107 Perhaps an example of wrong magnitude/localization relevant to the sentence?
Figure 1 Legend: add perhaps “observed *peak* amplitude”?
L112: Should it be *0-1* instead of *0-100*?
L112/116 Perhaps a figure or a display to *highlight* the content in parenthesis, that may deserve a sentence by itself.

L144: Needs clarification (scautoloc/Scanloc and or nonlinloc). I reckon that is scautoloc and or scanloc that are using 6 picks for a solution.
L145: I kind of missed the SED location score reference -Can this be clarified?
Figure 2: Are there are other stations from CH outside Switzerland in the North?

L165: Perhaps add more context about not disturbing the origin time, if this is relevant.
L177: The Author defines a good GOF (>50) and a bad GOF (<50). Perhaps clarify briefly somewhere. As well the choice of 55 for instance.
L187/188 Perhaps to be clarified – I might haven’t completely understood the second statement.
L197/211 Could ld Table/figure5 be simplified or split a little?

L226-227 Perhaps clarify the time – (provided a dense network I guess)

L212-225: If relevant: I d guess form the results that quarry blast and others may have different envelopes versue modelled local tectonic event. This could be evoked in the text.

L238-240: The Author may discuss the balance between simplicity and complexity of methods in an operational environment later in the conclusion if relevant.

L241-247 Question For the discussion: Why wouldn’t the author impose a severe SNR selection/ GOF threshold criteria per station (less data but good)? Or Do you need all the stations or a subset of them to establish a good solution ?

A bit short
L299-301: Would the author detail/complement the discussion about “regionalization” and “predicted onset time”?

Supplement

In the figures: in the figures/tables. Is it horizontal distance change?

Figure S6: Is velocity unit ok? Is s6a necessary or S6a is a good example and s6b is a bad example ? Perhaps simplity that figure?

L32: GMM should be ground motion model

L62: Perhaps the motive of P-P ratio to be stated at the beginning of the scaling?

Biblio Paper

L319: typo « J. Andrews »

Reviewer B Comments

This is a good paper proposing a methodlogy to use envelopes to judge the accuracy of earthquake early warning (EEW) algorithms and to make intelligent decisions when comapring EEW parameters fromt wo or more algorithms. I have no major concerns and I applaud the authros for well-done research.

Minor comments which authors might consider addressing:

(i) I think a cartoon/demo figure of waveforms and envelopes would be useful early on in the text to calrify the method. The paper in its current form is aesthetically somewhat spartan. I think an example of a real eathtquake and an envelope and an non-earthquake and an envelope would help the reader to follow what is being done

(ii) I think most of what is in the supplement is improtant methods and would cosnider moving it to the main text. How the evnelopes from Cua (2005) (a tragedy this was never peer reviewed by the way) is interesting and fundamental to the paper. It does not belong tucked away on Zenodo. Maybe not all the figures need be moved but certainly long parts fo the text along with a few key figures.

(iii) The Achilles heel of this is of course large events. Even if Georgia Cua used data up to M7.3 there are precious few records of M6.5+ events in that database, never mind M7.5+ or M8+ so how this methodlogy would generalize to those large and very large events is not clear. I'm not suggesting you need to have an answer tot his but acknowledging it int he dsciussion would be nice.

Other very minro things int he annotated PDF.

I enjoyed your paper. Thank you.

Response to Reviewers

Dear Editor and the Reviewers,

We thank you for the constructive comments which helped us to improve our manuscript. We have taken into account all the suggestions of the editor and the reviewers. Please find our detailed responses to each point raised by the reviewers and the editor below. Our answers to the points are in red. Furthermore, we also send the manuscript (and the supplement) with changes implemented, with a version in which they are highlighted (added text in blue, removed text in red).

The Authors

Editor:

I am pleased to say that I have now received two peer-review reports for your manuscript. Both reviewers found the work very interesting and are supportive of it being published as a Research Article in Seismica. However, they suggest that some revisions are needed before publication.

- In particular, I agree with both reviewers that it would help to have some additional clarity and details in the method description, perhaps by moving some of the information in the supplement to the main text.
  A: We have moved some of the information from the supplement into the main text. We explained our choice in the answer to Reviewer B.
- I also agree that some discussion about how large events are or would be handled would be helpful.
  A: We have added a short discussion about large events to the manuscript, as explained in the answer to Reviewer B.

Reviewer A:

We thank the reviewer for the constructive suggestions to improve our manuscript. We split them into those coming from the email and those coming from the accompanying *A-Comments_SeismicaReview2023_JozinovicEtAl.docx* file. The detailed answers to the suggestions with the implemented changes are listed below:

Email issues:

(a) Some minor additions or clarifications in some paragraph and parts
(b) Some clarity, simplification or changes in the figures.
A: We assume that the reviewer was referring to the issues in the docx file, which are answered below.

(c) Some more discussions/conlusions about for instance requirements/limitations such a data/metadata quality of a seismic monitoring networks/systems to operate the method and/or simplification or improvement of the envelope calibration.
A: We have discussed the limitations of the envelope prediction in lines 295-305. We agree with the reviewer that the errors in timing/signal quality/metadata could strongly affect the results of the method. We have added the following to the discussion: "*We expect that the errors in timing/signal quality/metadata could strongly affect the results of the method. The test of the influence of these errors on the results of the method will be made during real-time implementation and testing of the method, where unplanned errors can occur.*"

(e) One maybe suggestion in the text/conclusion etc: planning future test would it be worth to experiment with "usual/normal" seismicity?
A: As also stated in answer to the question (c) we would like to test the method in a real-time implementation to understand all its power and limitations, which could come from different factors (noise levels, predicted arrival times, real-time trigger information, station weighting, etc.). We have added a sentence to the end of the discussion section: "*Given all the unknowns just described, the real-time implementation of the method will also allow us to understand the performance of the method during times of normal (i.e. low magnitude) seismicity*."

(f) The method is assumed possible to be ported to other seismic networks. How does the method handle depth? Should that be tested later on?
A:. We are using hypocentral distance when obtaining the predicted envelopes (in the Cua, 2005 the depth was set to a constant 3 km), therefore we are currently taking depth into

account. The depth should not affect the method itself if it was modelled through the predicted envelopes. Therefore, other seismic networks could take depth into account by selecting the appropriate envelope prediction method.

Question more general: Were all or some of the large events used to model the predicted envelopes?
A: Yes. More precisely they were used in the methods we use for the scaling of the Cua (2005) envelopes. Some events were used for the GMM relationship calculation, and all of them were used for the S-P ratio calculation. However, the number of large events is minuscule compared to the total number of events used, so we assume there was little or no information leakage in the results. Furthermore, the article focuses on the method itself, and we did not strive for perfection when predicting the envelopes.

Appended docx file issues:

Manuscript

- L28: Add "..rapidly estimate developing ground shaking.. " for clarity ?
- L54-58: Minson et al. description for clarity perhaps splitting into 2 sentences?The method is described as complex.
- L145: I kind of missed the SED location score reference -Can this be clarified?
- L99 equation (2) has a typo : should be A ��
- L106 I guess it is …"good amplitude fits for certain …".
- L112: Should it be 0-1 instead of 0-100?
- L226-227 Perhaps clarify the time – (provided a dense network I guess)

Supplement

- L32: GMM should be ground motion model
- L62: Perhaps the motive of P-P ratio to be stated at the beginning of the scaling?

A: We thank the reviewer for the constructive suggestions. They have been implemented in the article.

Q: L46/50: Perhaps more emphasize about the pro/cons from multiple detections process frm a seismic network?
A: We have added the following to better explain the point: "*Providing solutions from a few stations allows small events to be identified and the earliest solutions for large events. Allowing frequent updates as more phase picks arrive means better accuracy can be achieved as the energy from the seismic event progresses across the seismic network*"

Q: L62: Minson et al is adapted/modified by the authors: Perhaps slightly more emphasize about the benefits of using a "complex" method and turn it into a more "simple" one.
A: We have explained the differences between the methods in the text. However, we prefer not to criticise the previous work, and therefore we do not make any changes to the text.

Q: L66 perhaps name "both" solution origin process (finder etc..) ?
A: The word "both" was used to name the uses of the method (selecting the best solution and checking if it reaches a pre-defined threshold). We removed the word "both" to avoid confusion.

L72-76: The motives could be earlier in the paragraph?
A: We think that it is more natural to state the specific motives at the end of the introduction. We already stated the general motivation in previous sentences (lines 54-64)

Q: L80: Perhaps precise location “blunder”?
A: We have now changed the sentence to say: “*suppressing processing blunders from significant errors in automated locations or significantly elevated magnitudes that are regular issues in seismic network monitoring*”

Q: L84: Perhaps Clarify the steps: the author started with a “first step” statement, though does not clearly state the next ones.
A: We removed the “first step” statement as there is no preferred order of doing the processing (i.e. do we first obtain the observed or the predicted envelopes) or the calculations (i.e. do we first calculate A(S,t) or C(S,t))

Q: L90 Question: Is there one max amplitude per 1 second interval ?
A: Yes - now clarified in Figure 1.

Q: L107 Perhaps an example of wrong magnitude/localization relevant to the sentence?
Figure 1 Legend: add perhaps “observed peak amplitude”?
A: We have added an example to the sentence: “*(e.g. for an M 5.5 earthquake at 50 km distance the predicted PGV is 0.0054 m/s and for an M 5 earthquake at 25 km distance the predicted PGV is 0.0055 m/s)*.”
We have modified “observed (o) and predicted (m) values” to “observed (o) and predicted (m) peak values”.

L112/116 Perhaps a figure or a display to highlight the content in parenthesis, that may deserve a sentence by itself.
A: Figure 2, referenced in the previous sentence, highlights the content of the text which explains the function A(s,t). We think that the current text and Figure is sufficient.

Q: Figure 2: Are there are other stations from CH outside Switzerland in the North?
A: The reviewer correctly points out that there are stations outside of Switzerland in South Germany. However, all the active stations used in the study are already plotted in the Figure (see https://networks.seismo.ethz.ch/en/networks/ch/).

L144: Needs clarification (scautoloc/Scanloc and or nonlinloc). I reckon that is scautoloc and or scanloc that are using 6 picks for a solution.
A: The statement as written in the text is correct, solutions are generated through scautoloc using a minimum of 6 picks.

L165: Perhaps add more context about not disturbing the origin time, if this is relevant.
A: We have added the following to the article: “*We do not perturb the origin time, though origin time errors have a strong impact on the GOF. Algorithms that do not provide strong constraint on the origin times can be penalised by this metric, and further work is required to address this.*

Q: L177: The Author defines a good GOF (>50) and a bad GOF (<50). Perhaps clarify briefly somewhere. As well the choice of 55 for instance.
A: In line 210 we describe the choice of GOF > 55 as the threshold for selecting a solution acceptable in the current study. The reviewer is correctly pointing out that we mention values of GOF > 50 as high fit. Therefore we changed the sentence in line 177 to now state: "*relatively high fit values (close to or higher than 55)...*" instead of "relatively high fit values (higher than 50)"

Q: L187/188 Perhaps to be clarified – I might haven't completely understood the second statement.
A: We reformulated the sentence to be clearer as follows:
"*Furthermore, for solutions with significant distance errors, the observed increase in the goodness-of-fit value at longer time windows is also a consequence of the inclusion into the GOF calculation of 1) more distant stations for which the predicted amplitudes are often close to the noise level (Figure S6a), and 2) the stations for which the true and false epicentre can be at a similar distance (Figure S6b), e.g. a station halfway between the true and false epicentre*"

Q: L197/211 Could ld Table/figure5 be simplified or split a little?
A: We have split the Figure 5 table into two parts to make it more readable.

Q: L212-225: If relevant: I d guess from the results that quarry blast and others may have different envelopes versue modelled local tectonic event. This could be evoked in the text.
A: We agree with the reviewer. It is especially clear for regional and teleseismic events. We added the following to the text: " *The predicted envelopes do not include event type. However, the efficacy of the method on different event types is demonstrated in Figure 7.*"

Q: L241-247 Question For the discussion: Why wouldn't the author impose a severe SNR selection/ GOF threshold criteria per station (less data but good)? Or Do you need all the stations or a subset of them to establish a good solution ?
A: Using a sort of station quality metric could be useful to lower the influence of noise from the stations. However, noise also carries important information that a certain station has not recorded any earthquake signal (e.g. for false alarms from non-existing events where it's crucial to know that all the stations with predicted earthquake ground motions have recorded noise). However, we do impose some selection criteria allowing us to select only the relevant stations, as noted in lines 141-142: "*In practice, at any given time t, we only use stations where the predicted or observed earthquake ground-motion envelopes are non-zero*"

L238-240: The Author may discuss the balance between simplicity and complexity of methods in an operational environment later in the conclusion if relevant.
This is a fair point. Our proposed method would be relatively complex to program efficiently in a realtime environment, and requires sophisticated realtime processing and computing power. Yet on the other hand, it will remove false local earthquake detections and false magnitudes, even for EEW. It will also select a preferred solution between similar location / magnitude pairs, and can operate effectively in realtime and with little number of stations and

only seconds of data. This has a significant benefit to operational systems and seem worth the effort to realise. We have added the following to the conclusion section: "*The incorporation of the method into a real-time environment brings more challenges beyond just the calculation of the goodness-of-fit. However, the method can bring significant benefits to operational (EEW and earthquake monitoring) systems, justifying the effort needed to implement it*."

Q: A bit short L299-301: Would the author detail/complement the discussion about "regionalization" and "predicted onset time"?
A: The discussion about these two topics has already been written in the discussion (lines 295-300 and 313-314). In the conclusion we just summarise the key points, so we prefer to keep it as is.

Supplement:
Q: In the figures: in the figures/tables. Is it horizontal distance change?
A: Yes. We have changed "*magnitude-distance perturbation*" to "*magnitude-epicentral distance perturbation*" in both the supplement and the manuscript text.

Figure S6: Is velocity unit ok? Is s6a necessary or S6a is a good example and s6b is a bad example ? Perhaps simplity that figure?
A: Yes, the velocity unit is correct. Both figures are necessary as they show two scenarios, described in the manuscript in lines 205-209. We have added the same description to the figure caption.

Biblio Paper

- L319: typo « J. Andrews »
  A: We made an error by citing the Zenodo repository. We have now changed the text to only contain the citation referenced in the GitHub repository of the module: "*Massin, F., Clinton, J., & Böse, M. (2021). Status of Earthquake Early Warning in Switzerland. Frontiers in Earth Science, 9. https://doi.org/10.3389/feart.2021.707654*"

Reviewer B:

We thank the reviewer for the constructive suggestions to improve our manuscript. We split them into those coming from the email and those coming from the annotated manuscript file. The detailed answers to the suggestions with the implemented changes are listed below:

Email issues:

(i) I think a cartoon/demo figure of waveforms and envelopes would be useful early on in the text to calrify the method. The paper in its current form is aesthetically somewhat spartan. I

think an example of a real eathtquake and an envelope and an non-earthquake and an envelope would help the reader to follow what is being done
Answer: Thank you for this suggestion - we have added a figure with an example of a real earthquake (Figure 1a) and a false alarm (Figure 1b) to the article.

(ii) I think most of what is in the supplement is improtant methods and would cosnider moving it to the main text. How the evnelopes from Cua (2005) (a tragedy this was never peer reviewed by the way) is interesting and fundamental to the paper. It does not belong tucked away on Zenodo. Maybe not all the figures need be moved but certainly long parts fo the text along with a few key figures.
A: Thanks for this comment. Way back in 2005, Cua submitted a paper describing these relations that was rejected by reviewers. Even though these continue to be widely used, when they are included in a journal article, they often receive a hard time from reviewers - one of the key reasons why we decided to leave them in the supplement is that we wanted to avoid these discussions for this paper.

The envelopes were intentionally not a main part of the originally submitted manuscript. In practice, as we describe in the supplement, in order to optimise the fit, the predicted envelopes required a series of adjustments to the original Cua (2005) relationship. This discussion can distract from the methodology which is the main point of the paper. In fact, the envelopes could be predicted using the Cua (2005) relationships (adapted or not) or in some other way (e.g. creating envelopes from synthetic seismograms). The choice of the envelope prediction relationship is up to the user who wishes to apply this method. Following the reviewer's suggestion, we have inserted a concise description of the envelope prediction into the manuscript and refer the reader to the supplement for more details:
"We calculate the predicted envelopes following the Cua (2005) relationship, using magnitude, hypocentral distance and site class (either "rock" or "soil"). To attribute the site class to the stations used in this study, when available, we used the EC8 ground types (Code, 2005) available from the SED stations website (Zurich, 1983). EC8 ground types are categorised as rock for EC8 ground types A and B) or soil for all other EC8 ground types and as default in the absence of EC8 type. The relationship then outputs P and S waveform envelopes, which start at earthquake origin time with a specified duration that matches the time window of the available station waveform. The Cua (2005) predicted envelopes were calibrated using data from southern California, which consisted of about 30,000 records (vertical and horizontal acceleration, velocity, and displacement) from 70 southern California earthquakes ($2 \leq M \leq 7.3$) recorded within 200 km from the earthquake source. However, for the subset of Swiss earthquakes we used, we observed (Figure S7 in the electronic supplement) that these predicted envelopes often do not fit the observed shaking well and visual checks showed a systematic overpredicting of the observed shaking. Therefore we decided to scale the predicted envelopes using the GMM developed by Cauzzi et al. (2015) for Switzerland (see also Edwards and Fäh, 2013). This approach reduced the difference in peaks between observed and predicted envelopes (Figure S7 in the electronic supplement). While this significantly improved the overall envelope fit, we still found that the maxima of the P-waves (especially at the closest stations) were higher in the observed data. We then further adapted the predicted envelopes using a station-specific S-P scaling and multiplication of the P-wave amplitude with an ad-hoc scaling factor. Further details are given in the Electronic Supplement."

(iii) The Achilles heel of this is of course large events. Even if Georgia Cua used data up to M7.3 there are precious few records of M6.5+ events in that database, never mind M7.5+ or M8+ so how this methodlogy would generalize to those large and very large events is not clear. I'm not suggesting you need to have an answer tot his but acknowledging it int he dsciussion would be nice.
A: We agree with the reviewer that the generalisation of the method to large events is highly dependent on the predicted envelopes. Some (successful) tests of the Cua (2005) relationships on large events (Chi-Chi dataset) have been done (Yamada and Heaton, 2008), but it would probably be required to do more tests to confirm the usefulness of the relationship for large events. Therefore our outlook for future work is to develop a new envelope prediction relationship/ML model that would tackle some of the issues with the predicted envelopes. This is, however, far outside of the scope of the article. We have added the following to the discussion section of the manuscript: "*Furthermore, it is unclear how well the envelope prediction relationships apply to large (bigger than M 6.5) earthquakes, which could affect the goodness-of-fit values for those events. Some preliminary tests on this topic have been done (Yamada and Heaton, 2008), but more extensive testing is required to confirm the results. To make the method more general, our next steps in improving the method will include developing a more general envelope prediction method developed on a global earthquake dataset with a significant representation of large events (ideally uniform across magnitudes).*"

PDF issues:

L91-92: Why? This feels crucial to the paper and should be in the main text.
A: Answered in point (ii)

L153: Consider showing plots/examples of the envelopes compared to real data.
Answer: We have added the figures for an example of a real earthquake (Figure 1a) and a false alarm (Figure 1b) to the article.

L193-196: Would you then set different thresholds for different time windows?
A: Setting different thresholds for different time windows could improve the robustness of the method. This, and other configuration choices, should be decided after a careful real-time analysis of the performance of the method. This would require implementing the method in real-time monitoring software (SeisComP) which we are currently working on. It will, however, take a significant amount of time and the design choices will be reported in subsequent manuscripts that will follow the real-time implementation of the method.

Reviewer A Comments

Thanks for the answer to the reviewers an editors as well as the explanations and the changes provided
I accept the updates and clarification and I reckon that the manuscript
is ready and suitable for publication from my point on view.

best regards

thanks.