Nat. Hazards Earth Syst. Sci. Discuss.,
doi:10.5194/nhess-2016-153-RC2, 2016

Natural Hazards and Earth System Sciences
Discussions

Interactive comment on "Tsunami arrival time detection system applicable to discontinuous time-series data with outliers" *by* J.-W. Lee et al.

Anonymous Referee #2

Received and published: 30 June 2016

General comments: A novel tsunami arrival time detection system is proposed in the manuscript. The system consists of three separate modules: outlier detection algorithm, gap filling algorithm, and tsunami detection algorithm. System is tested on sea level time series measured at Ulleung-do (an island off shore of Korea) tide gauge instrument during March 2011.

Described tsunami detection method is interesting and worthy of further development and testing. However, before manuscript can be published I have two major comments: (1) Detection system is calibrated on too short (and too few) time series, namely only one month of sea level data including only one tsunami event (regardless of strength of this event), and this is simply not enough to test or calibrate such a system. Authors are aware of this short coming and say that it might be overcome by either using longer measured time series which include more tsunami events or by using synthesized time

Printer-friendly version

Discussion paper

series. I believe authors should follow their own suggestion. Certainly, there are a number of tide gauge stations in the World which have longer records and have more tsunami events in them, perhaps even in a nearby Japan. Or if authors are not able to obtain this data, then they should do their tests on synthesized time series, following some of the papers they quote.

(2) Manuscript is not very clearly written. With all the abbreviations, flow charts, tables without explanations , it is hard to understand some of suggested algorithms. Authors should make their text more clear. More specific comments on this will follow.

Specific comments: A number of terms are not explained when first introduced but only afterwards (and some never), e.g. event period is mentioned in the abstract, and several times after it, but not explained until results section.

In introduction what are: short-time outliers, long-term outliers, short gaps, long gaps? I understood after reading the manuscript that these values (length of time steps) are later defined through calibration for specific herein described tsunami detection system. However, in Introduction, when dealing with values from literature, you should write a range of these values used by other authors. It would also be good to say that for this particular tsunami detection system, values will be determined through calibration.

Also in Introduction, what are soft computing techniques?

Still in Introduction, there are two contradictory statements: (1) "A low probability exists for tsunamis to occur in the East Sea" (2) "This can be used to detect weak tsunami signals that are common in the Ulleung-do surge data" What is correct then? If there are more tsunami signals in the Ulleung-do surge data, why not incorporate longer time series with these signals into your analysis?

Figure 1. I suggest adding bathymetry contours (perhaps coloring the figure?) and also pointing to Yamato rise mentioned in introduction.

Figure 2. Resolution should be increased. Why are sea levels showed with dots? I

Printer-friendly version

Discussion paper

think it's better to just use line. Also, it would be nice to add a zoomed in window showing Tokohu Earthquake period.

In methods, you again refer to long and short gaps without defining them or saying that they would be defined later.

In general, idea behind your process should be more clearly presented. Why do you remove outliers and fill in gaps when these algorithms are not used during the event? I assume to be able to compare event period time series to time series from previous time steps - but then this should be clearly stated. Also, what exactly do you do with outliers, remove them and then fill the gaps? I believe so, but this is not clearly written.

In 2.1. Outlier detection algorithm, entire chapter is highly difficult to follow. Try to simplify while still keeping the most important points. Likewise, Figure 4 is also very difficult to comprehend. I understand that it is a code flow chart - but perhaps here you could put a more simple version, and put this one (alongside with a code) to supplementary material? I.e. if your code is not described in some other paper. If yes, I do not see a need for a complicated figure and code.

In 2.2. Gap-filling algorithm, I have similar comments as for 2.1., although it is written a bit more clearly, and figure is more understandable and helps follow text (but still not good enough!). There are also a number of abbreviations - so it is easy to get lost. Some of this abbrevations are, I believe, not explained: what is SW, what is EPFM, what is SWEP? what is h1, h2, h3, h4, t1, t2.

As I understand, you basically copy search data to gap window - but before that you make sure that you fit first and last point of search data to first-1 and last+1 point of the gap? If so, this can be clearly written.

In sentence "A predefined length of points (N_inter) from the last poing... to create the linear interpolation..." does this mean that you linearly interpolate data by using the least square method? Or something else?

In 2.3. Tsunami detection algorithm you present Table 1. This table is completely unclear. I guess that first two columns are related to DART, second two to SLOPE, and last two to TIDE algorithm. This is nowhere written (should be in the table caption). Some of abbreviations in Table 1. are defined in text but most of them not. What is tIS, tG, tGTide, tTide, and so on?... I assume some parameters related to DART, SLOPE and TIDE equations. But if so, these equations should be written and explained in the manuscript. Also, are all of these calibrated values or general values related to method or a mix?

How is Tsunami Detection Index divided into five levels if you have only three tsunami detection algorithms? What are these five levels?

Figure 6. is also really difficult to follow. I'd say if all of algorithm you use (including outlier and gap filling algorithm) are from previous papers, there is no need in including such a complicated version of Figures 4, 5 and 6. Something simpler would do, or even omitting figures.

Related to Tables, none of them are very clear or fully explained. In Table 3, what does it mean that windowsize is 2 (two of what? Points, hours?...), or that npastdata is 100 (100 of what?).

In Table 4, why is search window located 3-14 days before actual gap?

In Results, you say that yellow alarm is triggered outside of the event period. How is this alarm triggered if you are not in the event period? And thus (from Figure 1) no tsunami detection algorithm should be activated?

In Discussion, I believe your method which is triggered only when there is an event, would be extremely difficult to use during events which have not-easily detectable sources (like meteotsunamis, landslide, ...). You can elaborate further.

Technical corrections:

Page 5. line 22. "start and end points" instead of "end points".

Interactive comment on Nat. Hazards Earth Syst. Sci. Discuss., doi:10.5194/nhess-2016-153, 2016.

Printer-friendly version

Discussion paper