International Conference on Tsunami Warning (ICTW)
Bali, Indonesia, November 12-14, 2008

A NEW MULTI-SENSOR INVERSION APPROACH FOR FAST AND RELIABLE NEAR-FIELD TSUNAMI EARLY WARNING

JÖRN BEHRENS[1], ALEXEY ANDROSOV[1], SVEN HARIG[1], FLORIAN KLASCHKA[1], LARS MENTRUP[1] and WIDODO S. PRANOWO[1,2]

[1] *Tsunami Modeling Group, Alfred Wegener Institute for Polar and Marine Research, Bremerhaven, Germany*
[2] *United Nations University – Institute for Environment and Human Security (UNU-EHS), Bonn, Germany*
(Joern.Behrens@awi.de)

ABSTRACT: The inversion of real time sensor data from diverse sources to obtain a situation perspective in short time is a hard problem. By utilizing an analog forecasting method in combination with a simultaneous evaluation of multiple sensor measurements allows for fast and accurate situation awareness. While the traditional decision matrix approach leads to frequent false warnings, the multi-sensor inversion may give more reliable results, once measurements are available. In this presentation we will describe our approach and prove its suitability by means of a benchmark experiment.

1. INTRODUCTION

In order to give timely and spatially accurate forecasts of tsunami impact within very short time, it is necessary to invert from measurement data to the tsunami scenario. In other words, the following question needs to be answered:

Given a number of measurements, what is the corresponding tsunami situation?

Current tsunami early warning systems (TEWS) employ different approaches of different complexities, which will be reviewed subsequently. The simplest warning approach is based on seismic information alone. In a decision matrix, several parameters are checked; among them, the location to be under or close to the sea, the magnitude to be larger than a certain pre-defined threshold, etc. Once the criteria are fulfilled a warning is issued. This approach is employed in the current operational warning center of BMG in Jakarta as well as in many other warning centers.

A slightly more complex approach includes simulation results to assess the tsunami situation. From the seismic parameters, a possible earthquake induced initial uplift is derived. This uplift may then be used to perform a forward computation. However, due to timing constraints, most often the simulation results are derived from a database of pre-computed tsunami scenarios. In this case, the pre-computed scenarios contain information about the corresponding earthquake parameters. Either the closest matching scenario is taken as a forecast or a (linear) combination of several closely matching scenarios is derived. This approach works for a linearized problem formulation, where scenario computations are only taken to near shore control points. This approach is employed for example in the database developed by ITB or in the Japanese TEWS operated by JMA [1].

The most advanced TEWS operational to date employ inversion methods to derive refinements of initial uplift distributions. These inversions are computed in a linearized setting from deep ocean pressure gauge readings. By utilizing a linear relation between wave heights at known sensor locations and corresponding uplift values at a priori defined fault planes, corrections to the source can be computed by solving a linear inverse problem. This method works well in the far field, where the assumption of linear wave behavior holds and it has proved its reliability in the Pacific Warning Center equipped by NOAA [5].

For near field tsunami warning with highly non-linear wave characteristics, high sensitivity on the source distribution and rupture area, and extremely short forecast times, this approach fails. It is the aim of this paper to introduce the new multi-sensor inversion approach, developed in the GITEWS project to overcome these deficiencies. In this approach, the following assumptions are made:

1. The situation is observed by such diverse sensors as seismic sensors (giving seismic parameters), GPS sensors (giving earth crust deformation vectors), and wave gauges (giving water elevation data) is physically consistent. Therefore, the sensors observe one physical system from different perspectives.
2. Pre-computed scenarios are physically consistent representations of a possible real situation. Scenarios are simplified virtual physical systems, but their resemblance to reality is accurate within data uncertainty. Thus, virtual sensor data extracted from scenarios can be interpreted as observations of a consistent physical system.

Now, from these assumptions it follows that a comparison of measured sensor data with virtual sensor data, extracted from scenarios, gives information about the resemblance of a pre-computed scenario with the real situation. If the scenario data match well to the sensor observations, the scenario represents the real situation well.

This type of forecasting method has been used in the atmospheric sciences for a long time and is known as analog forecast [3]. An analog to the real situation is looked for and serves as an extrapolation to future times or uncovered spatial areas. The novelty in the approach proposed in this paper is the simultaneous use of multiple values. This is necessary in order to diminish uncertainty in the different measurements. To give an example: In the September 12, 2007 Mw 8.4 Earthquake west of Bengkulu, according to current knowledge the epicenter was located at (4.517°S, 101.382°E), while the area of maximum slip distribution was around (3.2°S, 100.9°E), approx. 150 km to the North [2]. Simulations based solely on the seismic parameters would give misleading forecasts. Once, gauge measurements and/or GPS dislocation measurements become available (usually after only a few minutes), a precise inversion can be made.

In order to apply this methodology, we need to answer the following questions, which will be the outline of this article:

1. How can we define a generalized distance between scenario and reality, given a multitude of different measurements?
2. How can we model and evaluate uncertainty in the measurement (for example the uncertainty in the earthquake epicenter, that does not necessarily coincide with the rupture area center)?
3. How does this method work, given an example earthquake?
4. And finally, what can be deduced as a result?

2. DISTANCE METRIC FOR MULTI-SENSOR SELECTION

It is our task to derive a generalized distance measure to compare multiple sensor values with scenario data. In this presentation we will consider three different groups of sensors, providing six different types of measurements:

Table 1. Sensor groups and measurement types.

Sensor group	Measurement type
Seismic System	epicenter location (lon, lat)
	Magnitude (Mw)
	Depth (in km)
Gauge System	Arrival time at gauge i (t_i in sec.)
	Wave height time series at gauge i ($ssh_i(t)$ in m)
GPS System	Dislocation vector at GPS reciever j (as 3D-vector in $(lon,lat,z)_j$)

2.1 Individual distance metrics

For each of the measurement types given in table 1, we will first define an individual distance metric. For the epicenter location this is the two-norm (Euclidean distance), denoted by d_{loc}:

$$\begin{aligned} d_{\mathrm{loc}} &= \|(\lambda,\phi)^{\mathrm{sensor}} - (\lambda,\phi)^{\mathrm{scenario}}\|_2 \\ &= \sqrt{(\lambda^{\mathrm{sensor}} - \lambda^{\mathrm{scenario}})^2 + (\phi^{\mathrm{sensor}} - \phi^{\mathrm{scenario}})^2}. \end{aligned} \tag{1}$$

For the magnitude M we chose the simple difference d_{M}, as well as for the depth D:

$$d_M = |M^{\mathrm{sensor}} - M^{\mathrm{scenario}}|, \quad d_D = |D^{\mathrm{sensor}} - D^{\mathrm{scenario}}|. \tag{2}$$

Arrival time t at any given gauge position is again treated by a simple difference:

$$d_t = |t^{\mathrm{sensor}} - t^{\mathrm{scenario}}|. \tag{3}$$

Time series are slightly more complicated. We employ the simplest metric here, the one-norm and denominate the distance metric by d_{ts}, where the index i stands for the time-step counter:

$$d_{\mathrm{ts}} = \sum_i |l_i^{\mathrm{sensor}} - l_i^{\mathrm{scenario}}|. \tag{4}$$

Finally we need to compare the GPS dislocation vectors to obtain d_{GPS}. Here we can employ the Euclidean distance again:

$$\begin{aligned} d_{\mathrm{GPS}} &= \|(\delta_x,\delta_y,\delta_z)^{\mathrm{sensor}} - (\delta_x,\delta_y,\delta_z)^{\mathrm{scenario}}\|_2 \\ &= \left[\sum_{i\in(x,y,z)} \left(\delta_i^{\mathrm{sensor}} - \delta_i^{\mathrm{scenario}}\right)^2\right]^{\frac{1}{2}}. \end{aligned} \tag{5}$$

2.2 Modeling uncertainty and noise

In order to take noise in the data into account as well as modeling the uncertainty not related to measurement error we introduce uncertainty radii, which complement the above defined individual distance metric. The radii are introduced individually for each measurement type and can be adjusted correspondingly. We will demonstrate this with the epicenter location.

We know that the epicenter location is not closely related to the center of the rupture area from the example in the introduction. Therefore, if only seismic parameters are available, then all scenarios with a rupture zone center located in a certain distance to the measured epicenter are equally likely to happen and therefore indistinguishable. The location distance metric for all these scenarios should be zero. Thus, we define an uncertainty radius U and modify the metric in the following way:

$$d_{\mathrm{loc}} = \max\left(\|(\lambda,\phi)^{\mathrm{sensor}} - (\lambda,\phi)^{\mathrm{scenario}}\|_2 - |U|, 0\right). \tag{6}$$

In other words, all scenarios with rupture area centers closer to the measured epicenter than U obtain a d_{loc} value of zero.

We still need to specify U. Since the dislocation between epicenter and rupture center is larger for large earthquakes we employ the following formula for U:

$$u = \max(M-6, 1), \quad U = u\cdot(u-1)\cdot C_s. \tag{7}$$

The first part of equation (7) says that an uncertainty radius is only considered for earthquakes of magnitude 7 or larger, while the second part tells us that the uncertainty radius increases quadratically with the size of the earthquake. C_s is a scaling factor, which can be tuned.

Similar uncertainty values are employed for each individual metric. The actual values for the uncertainty are tuning parameters, which need to be adjusted according to experience with the system.

2.3 Scaling to unified values

In order to be able to compare the very different values, we need to scale them. An appropriate scaling will discriminate small differences and it will project all values to a non-dimensional value range in the unit interval. The scaling function used here is given by:

$$\mathrm{scal}(d) = \arctan(\nu_d \cdot d) \cdot \frac{2}{\pi} \tag{8}$$

ν_d is a scaling factor, which has to be chosen such that the first approximately 25% of the distance metric interval is mapped to the interval [0,0.8]. To give an example: if we assume that the distance in unit magnitude d_M can lie in the range of $0<d_M<10$, then we want the distance values of 0 up to 2.5 to map to scaled values between 0 and 0.8. The value of ν_d can be fixed in the system. From now on, all different distances will be assumed to be in the unit interval (after applying the scaling function).

2.4 Combined metrics

After including an uncertainty model and scaling to the unit interval with accentuation on small distances, the individual metrics need to be combined. This is achieved with a weighted sum approach. Let us denote by $\hat{d}_\star$ the scaled metric corresponding to the measurement type $\star \in \{\mathrm{M}, \mathrm{loc}, \mathrm{D}, \mathrm{t}, \mathrm{ts}, \mathrm{GPS}\}$. For the gauge and GPS sensor groups, we need to consider individual sensors, while for the seismic sensor group we will use aggregated data (derived by the seismic system from individual seismometer sensors). Therefore, we first derive group metrics by using weighted sums for each sensor group. Let the index j denote the sensor counter in each group. Then we define the group metric d_group (where group stands for either gauge arrival time t, gauge time series ts or GPS group) by

$$d_\mathrm{group} = \sum_{j \in (\mathrm{sensor\ IDs})} w_j \cdot \hat{d}^j_\mathrm{group}. \tag{9}$$

The weights w_j can be used to switch off non-functional sensors or to indicate unreliable sensors by giving them less weight. Each individual sensor has its own weight. The set of weights needs to be tuned, based on experience and experiments.

In order to derive the mismatch value, we combine the group metrics again by a weighted sum approach:

$$\mathrm{mismatch} = \sum_{i \in (\mathrm{M,loc,D,t,ts,GPS})} W_i \cdot d_i. \tag{10}$$

Now, the weights W_i are used to emphasize the importance of each individual sensor group. We weight the direct wave measurements (by gauges) more relevant than the pure seismic information, for example. Note that all weights are normalized, that means we have

$$\sum_{j \in (\mathrm{sensor\ IDs})} w_j = 1, \quad \text{and} \quad \sum_{i \in (\mathrm{M,loc,D,t,ts,GPS})} W_i = 1. \tag{11}$$

This requirement together with the scaling function guarantees that the group metrics as well as the mismatch are normalized, in other words take values between 0 and 1. A mismatch value of 0 means perfect match of scenario with real measurements (given the data uncertainty). A value of 1 indicates maximum distance between real measurements and scenario values.

The data uncertainty may lead to a large number of indistinguishable scenario mismatches. Assume that only seismic data is available for the matching process. Assume further that the earthquake is

large and that the scenario database population density is high (i.e. a large number of scenarios is contained in the database). Then all the scenarios with their rupture center in the radius of uncertainty of the measured epicenter are assigned a mismatch of zero in the location measurement type.

3. RELIABILITY EVALUATION

In the previous section we have achieved to define a generalized distance – the mismatch – to measure the matching of a scenario with the real situation measured at a number of sensors different in type and location. Now, we will deal with situations in which either not all sensors deliver measurement data (which will usually be the case), or where not all sensor values can be compared, since they are not supported by the scenario. For example, a few minutes after an earthquake, we can assume that the seismic parameters are available, but only one or two of the gauge sensors have received a signal and only a few of the GPS sensors measured a deviation from the reference position. Additionally, we may consider scenarios in the database that include GPS deviations and others that do not include these values (e.g. scenarios based on Okada's fault plane mechanics usually do not involve a holistic view of earth crust deformation).

3.1 Reliability value

In order to take missing measurements into account, we simply set the weights corresponding to the missing values to zero. Note that we need to re-normalize the weight set by

$$w_\star^k = \frac{w_\star^k}{\sum_{j\in(\text{sensor IDs})} w_\star^j},\ \forall k. \tag{12}$$

The above formula holds for the sensor weights, but an analog formula holds for the group weights W_i. Now, in order to assess the reliability of a scenario matching result, we can compare the number of total sensors (with their weights) to the number of sensors with relevant data. As an example, if 10 sensors were available to measure the wave arrival, but only two of them had been reached, after a few minutes, the reliability or data quality would not be very high (approx. 20% of the theoretically optimal condition, where all 10 sensors provide relevant data). On the other hand, if all sensors had contributed data (after a longer time period), the reliability or data situation would be perfect. Even though the mismatch value could be worse than in the first case (it might be easy to find a scenario that matches the arrival time at two positions quite well, but to match all ten positions would be more difficult), the second situation would give much more accurate situation awareness. With this reasoning, we define the reliability

$$R = \frac{\sum_{i\in(\text{available groups})} W_i \cdot \left(\sum_{j\in(\text{available sensors})} w_i^j\right)}{\sum_{i\in(\text{all groups})} W_i \cdot \left(\sum_{j\in(\text{all sensors})} w_i^j\right)}. \tag{13}$$

Since the weights are normalized the sum over all sensor weights is always 1. Therefore, eq. (13) is equivalent to

$$R = \sum_{i\in(\text{available groups})} W_i \cdot \left(\sum_{j\in(\text{available sensors})} w_i^j\right). \tag{14}$$

To conclude: It is a simple renormalization procedure and weight manipulation operation to account for missing measurements. Additionally, the weights give insight into the reliability of the matching.

3.2 Skill value

Now, we will consider the situation, where a given set of measurements cannot be complemented by scenario data. It could be that the scenario's computational domain does not cover the whole affected area or a sensor group might not be available in the scenario (e.g. time-series or GPS values).

We can employ the same methodology as above and set the weights, which cannot be matched to zero. Again a renormalization has to be carried out. In fact this step has to be preceded by the missing sensor weight manipulation. We noted in the previous section that the mismatch might be smaller for smaller numbers of data items in the comparison. Therefore, we need to quantify the matching quality again.

Since missing values in a scenario is a quality indicator for each individual scenario, we call this value the skill of a scenario. Again, the weights are non-zero only for those values that are truly used. The ratio of weight sum of used sensors over total available sensors gives an indication of the skill. An example: if we put weight W 0.25 on the seismic sensor group, 0.5 on the gauge group and 0.25 on the GPS group and we have scenarios that contain GPS values and others that do not. Then – given a complete set of sensor measurements – the skill for the scenarios with GPS would be 1 while the skill for the scenarios without would be 0.75.

Formally, we define the skill:

$$\begin{aligned} S &= \frac{\sum_{i\in(\text{used groups})} W_i \cdot \left(\sum_{j\in(\text{used sensors})} w_i^j\right)}{\sum_{i\in(\text{available groups})} W_i \cdot \left(\sum_{j\in(\text{available sensors})} w_i^j\right)} \\ &= \frac{\sum_{i\in(\text{used groups})} W_i \cdot \left(\sum_{j\in(\text{used sensors})} w_i^j\right)}{R}. \end{aligned} \tag{15}$$

We have now derived three values characterizing a scenario match:

1. The mismatch describes the distance of a scenario to the real measurements.
2. The reliability describes the data quality or reliability. If this value is high, a large number of theoretically available sensor values could be assessed and the data situation appears to be reliable.
3. The skill describes the matching quality in terms of data available in the scenario. If the skill is large (close to 1) almost all available data could be complemented by scenario data. Therefore a scenario with high skill value is regarded to be more relevant than one with a low skill.

Note that these values are not unique. It may be that a scenario with low skill might be the one to match reality best, mainly by chance. It may well be that higher reliability leads to worse mismatch. However, this is not a problem, since the absolute mismatch value is not physically motivated and gives no absolute indication on the matching quality.

3.3 Indiscriminate matching lists

In order to deal with larger numbers of indiscriminate mismatch values for situations in which not a lot of sensor values is available, we introduce a threshold based mismatch list generation procedure. In order to quantify the number of possible scenarios, we look at the ordered spectrum of mismatches (fig. 1). The situation shows that approx. 60 scenarios are indistinguishable (zero-mismatch and zero difference). We plot also the differences from one scenario to the next scenario (the gradient of the spectrum). This plot indicates that there are jumps in the mismatch spectrum. In other words, a number of scenarios are similar in their mismatch and after a jump the next level of similarly

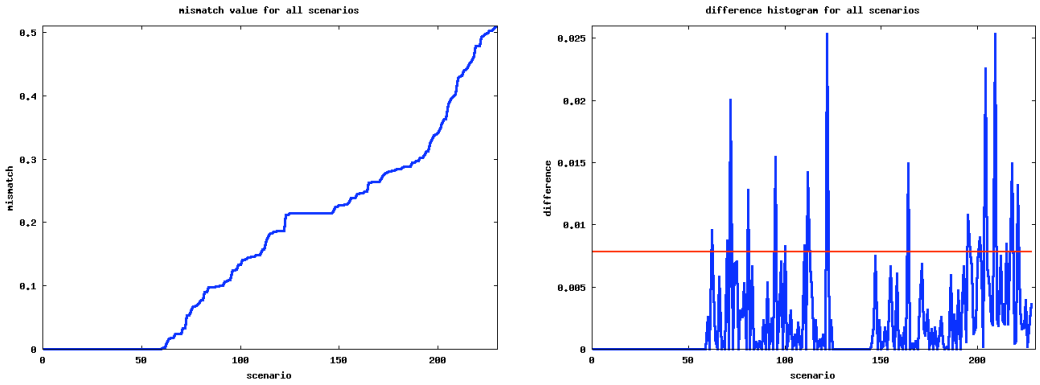

Figure 1. Spectrum of mismatches (left) and difference |mismatch(i)-mismatch(i-1)| (right). The red line indicates the cutoff for significant changes.

mismatching scenarios is reached. In order to maintain certain variability in the results and to account for uncertainties in the measurements, the list of returned scenarios is cut off, where the difference value excesses the cutoff value (three times the mean) for the third time. In this example, 72 scenarios form the set of possible events, given only the seismic information.

4. BENCHMARK EXPERIMENT

In order to test our approach, a benchmark experiment has been defined. We assume a Magnitude 8.0 earthquake close to Padang, West Sumatra, with its epicenter at (99.93°E, 1.96°S). The source distribution is inhomogeneous and depicted in figure 2 (left). We assume an instrumentation as in figure 2 (right), where (under optimal conditions) the seismic system is working, the GPS system is represented by the current SuGAr continuous GPS network stations [4], and the gauge system is represented by the two existing German deep ocean wave gauges (buoys). All "measurement data" have been computed independently by A. Babeyko's group at GFZ Potsdam, including the two benchmark coastal gauge readings in Padang and Bengkulu (fig. 3).

It is important to note that a scenario exists in the database, which matches the benchmark gauges very well, and which also matches the arrival and wave height time series at the buoys very well, even though it does not rely on an inhomogeneous slip distribution (see fig. 4).

The aim of our exercise is to find a scenario in the database that based on the matching, represents the benchmark gauges as accurate as possible. The chief officer on duty (COOD) in the warning center would then be able to give a qualified warning. Let us first consider the traditional case, where only seismic information is available. Taking the uncertainty into account, the simulation system finds approx. 70 scenarios indistinguishable. The corresponding mareograms (gauge timeseries) are plotted in figure 5. A warning decision based on the worst case would be appropriate in Padang, but in

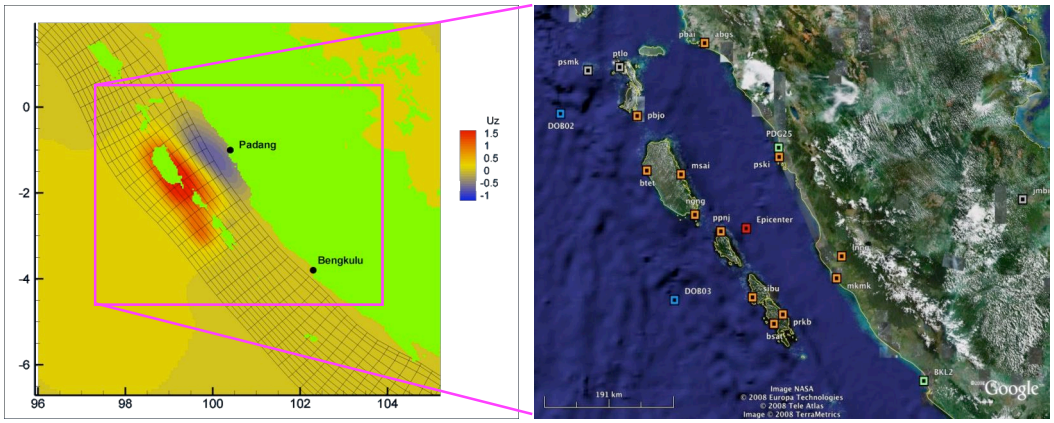

Figure 2. Initial uplift (left) and instrumentation (right) for the benchmark experiment. Seismic parameters (epicenter in red), wave gauges (blue), GPS stations (orange) and benchmark stations (green). Stations not used are marked in gray.

Bengkulu this would lead to a false warning.

In contrast to this, if also GPS and buoy data readings could be included, all scenarios give distinct mismatch values. We will still get a list of probable matching scenarios from our threshold based indiscriminate matching list (fig. 6). However, a worst case warning would lead to accurate forecasts in both benchmark positions. Additionally, the best matching scenario (red line) represents almost the optimal solution in the database. In fact, it can be shown that a selection based only on gauge and seismic data would give the optimal solution.

4.1 Conclusions

In this article we introduced the methodology for the GITEWS multi-sensor selection. It is embedded in an analog forecasting paradigm for tsunami early warning in the near-field regime. We demonstrated how the multi-sensor approach can effectively reduce uncertainty in short time and in the near-field, where non-linear wave effects cannot be neglected. The approach leads to accurate and qualified warning products, which is the subject of an other paper with U. Raape et al. (DLR).

The approach takes noise and data uncertainty into account and is robust against these effects. However, tuning of thresholds and weights needs to be performed with real life data streams. This will be the task of the operational testing phase of the GITEWS system.

5. ACKNOWLEDGMENTS

The authors gratefully acknowledge funding through the GITEWS Project, provided by BMBF under contract no. 03TSU01. Data for the benchmark experiment were generously provided by A. Babeyko.

6. REFERENCES

[1] Furumoto, A. S., Tatehata, H., and Morioko, C.(1999). "Japanese Tsunami Warning System", Science of Tsunami Hazards, Vol. 17, 85-105.

[2] Lorito, S., Romano, F., Piatanesi, A., and Boschi, E. (2008). "Source process of the September 12, 2007, Mw 8.4 southern Sumatra earthquake from tsunami tide gauge record inversion", *Geophys. Res. Lett.,* Vol. 35, No. L02310, 1-6.

[3] Hope, J. R. and Neumann, C. J. (1970). "An Operational Technique for Relating the Movements of Existing Tropical Cyclones to Past Tracks", *Monthly Weather Review,* Vol. 98, No. 12, 925-933.

[4] The Sumatran cGPS Array (SuGAr). http://www.tectonics.caltech.edu/sumatra/sugar.html, 2008.

[5] Wei, Y., Bernard, E. N., Tang, L., Weiss, R., Titov, V. V., Moore, C., Spillane, M., Hopkins, M., and Kanoglu, U. (2008). "Real-time experimental forecast of the Peruvian tsunami of August 2007 for U.S. coastlines", *Geophys. Res. Lett.,* Vol. 35, No. L04609, 1-7.

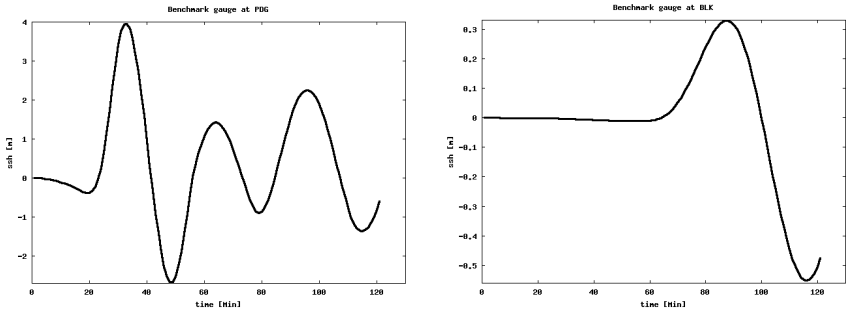

Figure 3. Benchmark gauge time series in Padang (PDG, left) and Bengkulu (BKL, right).

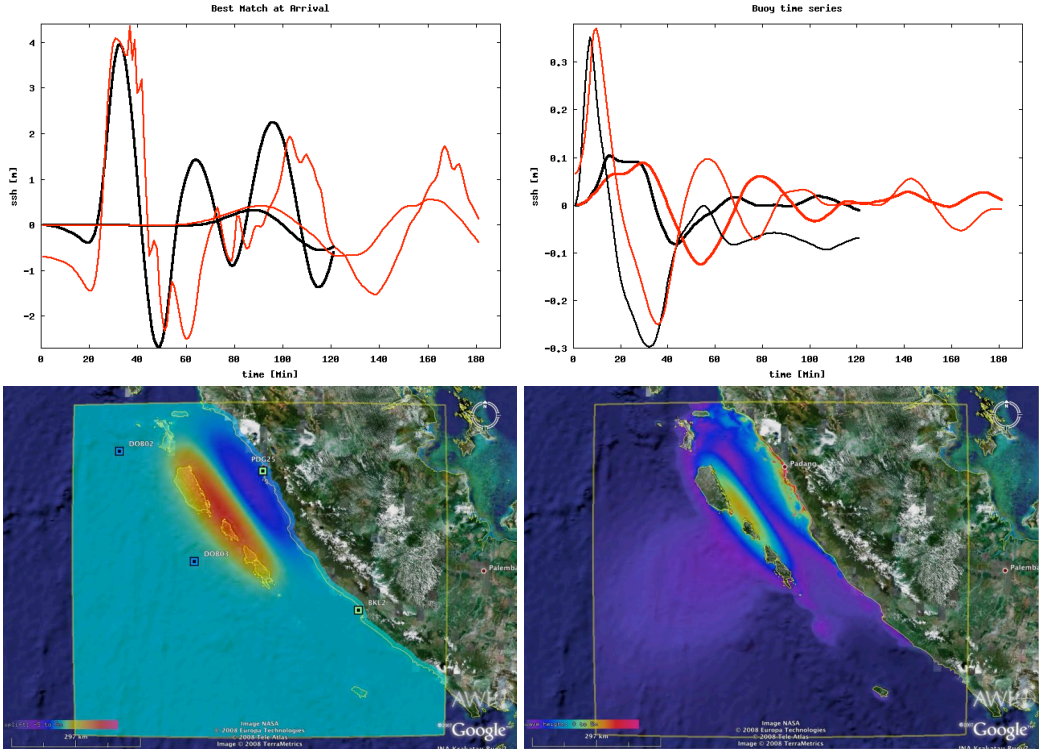

Figure 4. Manually selected best matching scenario: benchmark gauge time series [upper left; true solution (black), scenario solution (red), both gauges PDG and BKL are in one chart], buoy time series [upper right, both buoys in one chart, DOB02 (bold)], initial uplift (lower left) and maximum wave height (lower right).

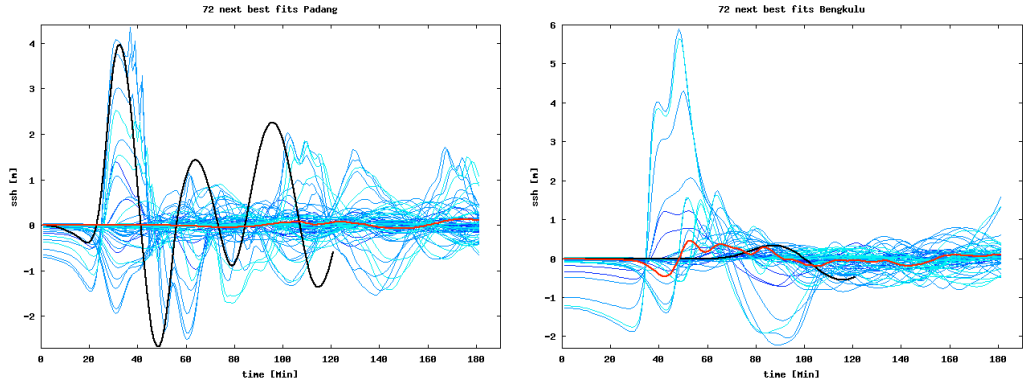

Figure 5. Selection results with only seismic information known (descriptions as in fig. 6).

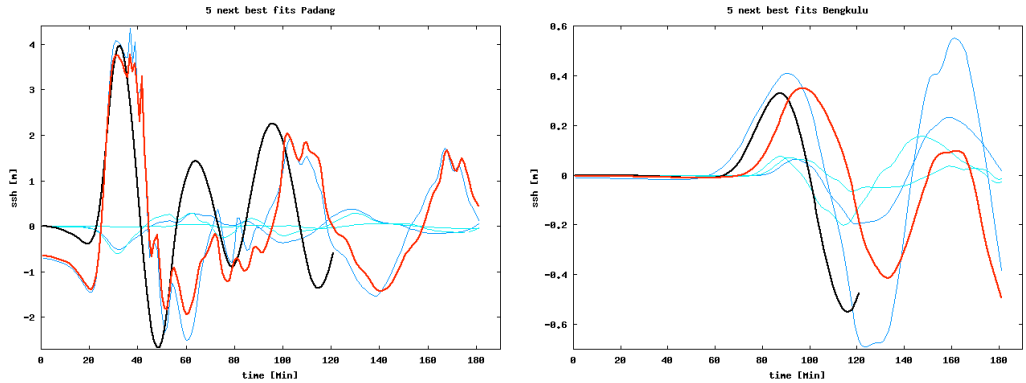

Figure 6. Matching result including all sensor groups (optimal conditions): benchmark gauge time series at PDG and BLK (upper left and right; true solution (black), best match scenario solution (red), next best matches (blue)).