Bezhaev A.Yu.[1], Marchuk An.G. [1], Titov V.V. [2]

ESTIMATION OF INITIAL ELEVATION IN THE EXTENDED TSUNAMI SOURCES ON THE BASE OF DEEP WATER WAVE MEASUREMENTS.

[1] Institute of Computational Mathematics and Mathematical Geophysics, SB RAS, Novosibirsk, Russia
[2] NOAA Center for Tsunami Research, NOAA/PMEL - UW/JISAO, Seattle, USA

Abstract. A scheme for the automated tsunami source reconstruction on the basis of deep-ocean tsunami wave recordings was developed. The initial water displacement in the source area is represented as a linear superposition of several “unit” sources. The initial value problem for the tsunami waves propagation from each of the unit sources in the whole of the north-eastern part of the Pacific ocean has been calculated using the Method of splitting tsunami (MOST). Using the database that includes all these calculated wave time series, a special algorithm and application software has been developed. The module effectively determines the amplification coefficients for unit sources, which makes it possible to approximate the shape of a vertical displacement of the sea surface over the tsunami source area. The algorithm in question is based on minimization of a calculated difference between the measured marigram(s) and a linear combination of pre-calculated synthetic marigrams. The method was tested against historical data of the 1996 Andreanov tsunami.

Acknowledgments

The work supported by CRDF Grant RUG1-2801-NO-06 and Grant SB RAS 2006-113. This study is partially funded by the Joint Institute for the Study of the Atmosphere and Ocean (JISAO) under NOAA Cooperative Agreement No. NA17RJ1232, Contribution #___.

1. Measurement system DART and FACTS database

In the late 90-s of the XX-th century, technical facilities for precise deep ocean recording of tsunami waves appeared. This makes it possible to obtain tsunami marigrams free of noise and disturbances from a near coastline event (such as a dramatic increase of dispersion and partial reflection from the continent slope). For the experiment, five deep-water recording stations

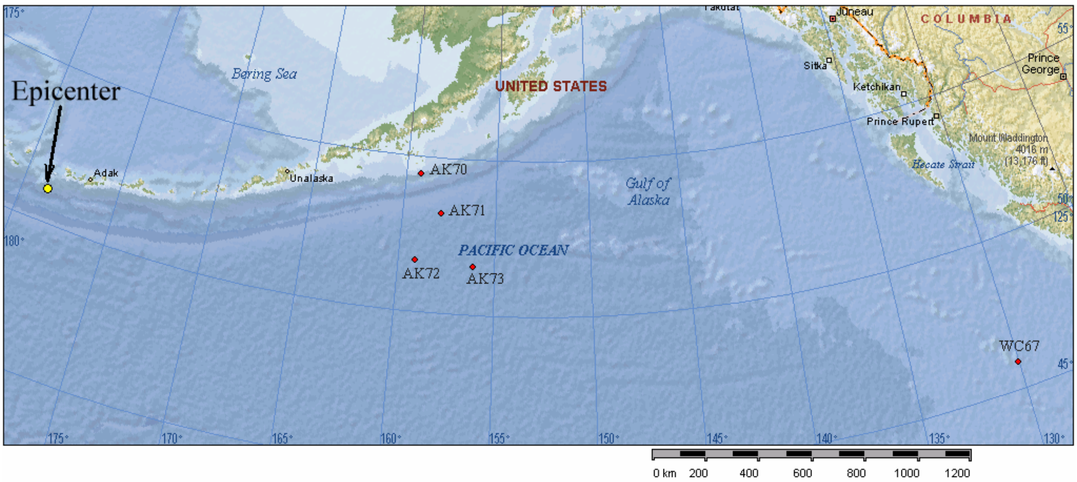


Figure 1. Experimental position of the five deep water recording stations and the epicenter of the 1996 Andreanov tsunamigenic earthquake

(AK70, AK71, AK72, AK73, WC67) were arranged in 1996 at the North Pacific (Figure 1). Telemetric facilities were not included in these stations, therefore the data were recorded at separate magnetic disks. After a few months, all the stations were extracted and the records were decoded. All the stations produced the records of waves from rather a weak Andreanov tsunami. The idea had arises to reconstruct the true source parameters through such marigrams. To do this, it is necessary to calculate the synthetic waves propagation from various model sources around the epicenter, obtained from the seismic data. Originally, the idea was to create a database of time series of the calculated waves initiated by tsunami sources typical of a given subduction zone. It is supposed that any tsunami source along the Aleut-Alaskan zone could be represented as a linear superposition of several "unit" sources. The field of surface displacements of such a unit source is shown in Figure 2.

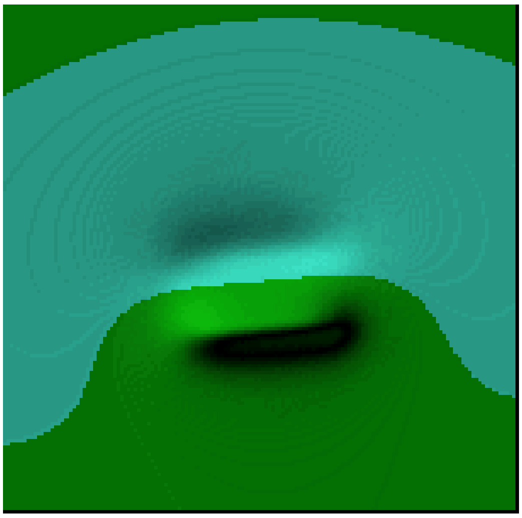


Figure 2. Surface displacement in the unit source

Based on the energy analysis and the crust faults location, it was decided to use 50 of such unit sources along the Aleutian-Alaska tsunamigenic zone. The number of sources was chosen according to the length of this zone and the average distance between tectonic faults in this area. Centers of these sources are displaced along the whole of Aleut islands chain, arranged in two lines (one row is closer and the other is farer from the deep-water trench, see Figure 3).

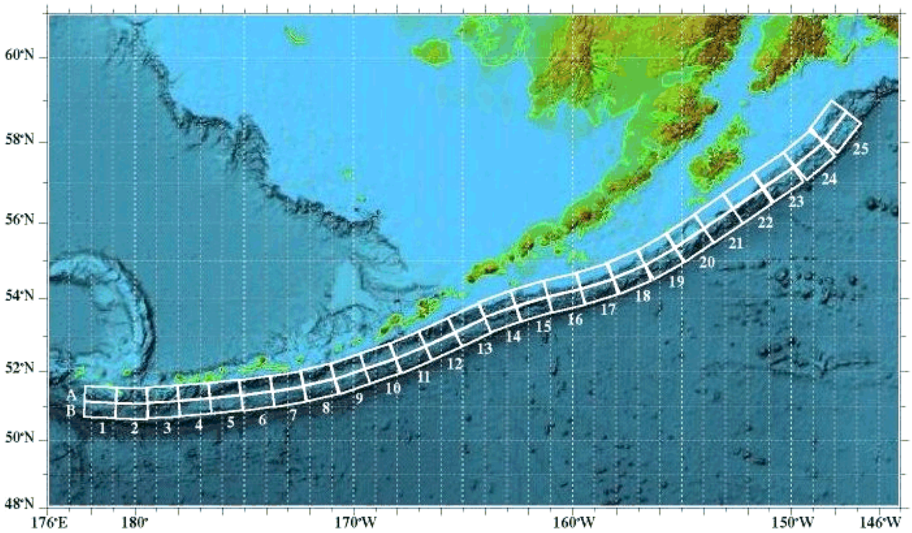


Figure 3. A set of the unit sources along the Aleutian islands and Alaska

Based on the initial displacement at a source, typical of the zone (see Figure 2), the tsunami waves propagation from each of the unit sources in the whole of the north-eastern part of the Pacific Ocean was calculated using the Method of Splitting Tsunami (MOST) [1]. Collection of all these calculated time series has been arranged as database. In particular, synthetic marigrams were obtained at all the mesh nodes, closest to real positions of the above-mentioned recording stations.

Based on this database, the first version of the Andreanov tsunami source was determined [2]. The real (measured) wave signals are displayed in Figure 4. By manual varying of weight (in fact, amplitudes) coefficients of the selected unit sources, a set of amplitudes in four unit sources was obtained. Selection of these sources was made based on an analysis of the first arrival times. Combination of waves from the unit sources with such amplitudes generates the signals at the recording points (Figure 4) similar to the measured marigrams.

Having gained this successful experience, the PMEL (the Pacific Marine Environmental Laboratory, Seattle, WA, the USA) developed the Deep- ocean Assessment and Reporting Tsunami (DART) system and placed it at six locations around the Pacific [3]. In Figure 5, the location of the DART stations is shown (dark icons). There are also indicated the proposed sites of future installations (light icons). The DART system obtains high-quality data of tsunami amplitudes in the open ocean.

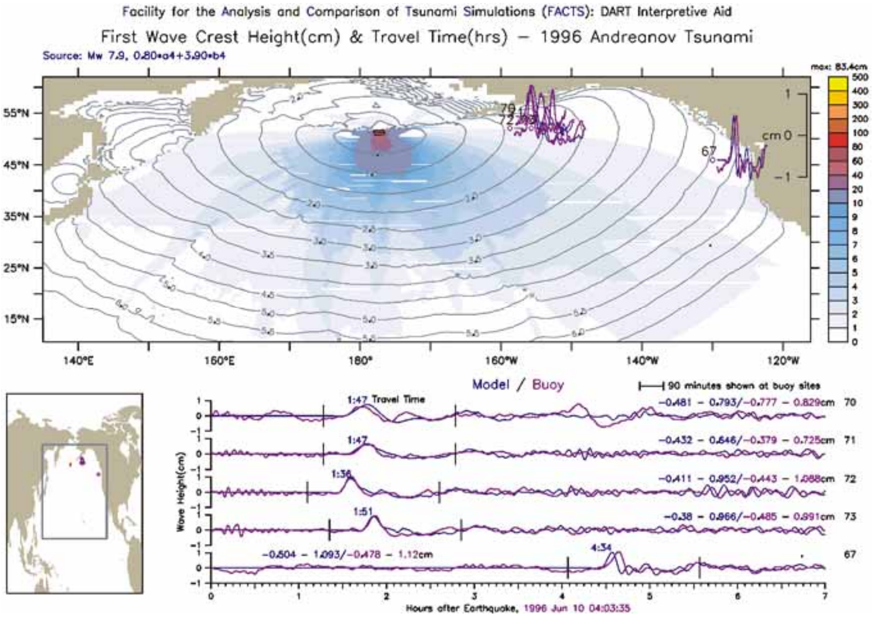


Figure 4. Wave series and computed marigrams at the recording points using manually selected coefficients

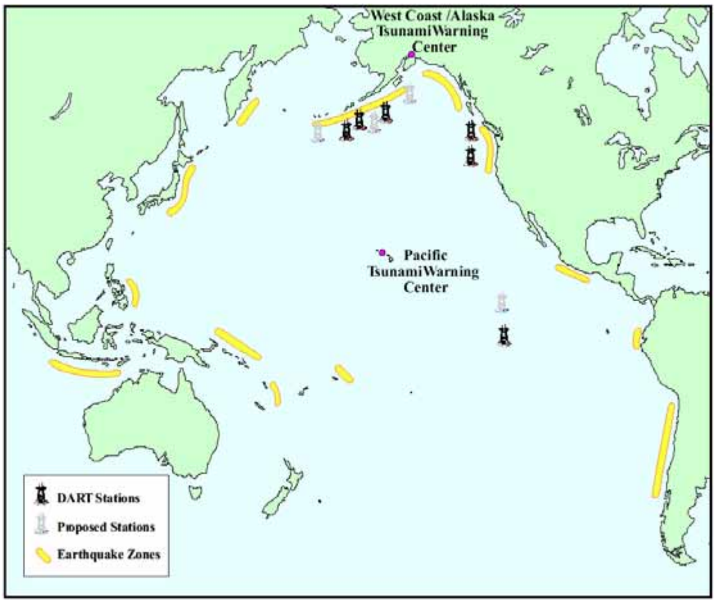


Figure 5. Location of installed (dark icons) and planned (light icons) DARTS stations

This data can be used to reconstruct the true tsunami source parameters. With this object in view, the PMEL has created a database of pre-calculated time series of tsunami waves from (unit) sources over subduction zones around the Pacific. The solutions from the database can be combined using the developed Web interface (called FACTS – Facility for the Analysis and Comparison of Tsunami Simulation), to reconstruct tsunami sources recorded by the DART stations.

2. Automated system for source parameters determination using the FACTS Database

The above-mentioned FACTS database contains time series (synthetic marigrams) from all the 50 unit sources of the Aleut-Alaskan tsunamigenic zone at all grid-points over the whole of the computational area, including the location points of the modern DART system. In order to operate with such a database in automatic mode, a special algorithm and application software were developed. This module is designed to determine the tsunami source parameters by processing the marigrams, obtained at the DARTS stations. In fact, the Module effectively determines the amplification factors for unit sources, which makes it possible to approximate the shape of a vertical displacement of the sea surface in the tsunami source area. The determination algorithm is based on minimization of a calculated difference between the measured marigram(s) and a linear combination of synthetic marigrams, taken from the FACTS database. For comparison, the time periods were selected up to one complete wave period in order to determine the source parameters in the real-time mode.

Let us describe the proposed scheme and its application to the software designed. Suppose that we have the measured records of the sea level from several, say N, DART stations. This means that the operator receives digital values of the sea level at all the stations within a specified time interval (for example, every 15 seconds). Then the operator can process N numerical sequences. The FACTS database contains synthetic marigrams at all locations of the DART stations with a similar time discretization. Varying the amplitude coefficients, the algorithm minimizes a difference between measured marigrams and a linear combination of synthetic marigrams from unit sources. Minimization is arranged at all the points of measurement. The difference is measured in L1 and L2 norms. This means that in the first case, we sum up the absolute values of differences between the measured sea level and a linear combination of calculated marigrams from the unit sources on the selected time interval (for example, during the first period of the recorded wave at each station). In the second case, the squared differences are summarized and then the square root is minimized. Mathematically the norms, which are used in the minimization process are expressed as follows:

$$\boldsymbol{l}_k(\boldsymbol{q}_1,\ldots\boldsymbol{q}_{50}) = \frac{1}{\boldsymbol{T}}\sum_{\boldsymbol{i}=1}^{\boldsymbol{T}}\left|\boldsymbol{h}(\boldsymbol{i}) - \sum_{\boldsymbol{j}=1}^{50}\boldsymbol{q}_j\boldsymbol{h}_j(\boldsymbol{i})\right| \quad (k=1,\ldots\mathrm{N}),$$

$$\boldsymbol{m}_k(\boldsymbol{q}_1,\ldots\boldsymbol{q}_{50}) = \frac{1}{\boldsymbol{T}}\sqrt{\sum_{\boldsymbol{i}=1}^{\boldsymbol{T}}\left(\boldsymbol{h}(\boldsymbol{i}) - \sum_{\boldsymbol{j}=1}^{50}\boldsymbol{q}_j\boldsymbol{h}_j(\boldsymbol{i})\right)^2} \quad (k=1,\ldots\mathrm{N}).$$

Here **T** is the number of time counts (that is, the duration of compared time period), $\boldsymbol{h(i)}$, $\boldsymbol{h_j(i)}$ represent the values of waves heights in measured and synthetic marigrams, respectively, while j indicates to the unit source number. The quantities $\boldsymbol{l}_k$ and $\boldsymbol{m}_k$ refer to minimization with respect to values of the coefficients $\boldsymbol{q}_j$ (j = 1,..,50) for each record station number $\boldsymbol{k}$, or, alternatively, for selection of such stations. In this case, the arithmetic mean of all the norms $\boldsymbol{l}_k$ or $\boldsymbol{m}_k$ is minimized. During the numerical tests it was supposed that the determined coefficients $\boldsymbol{q}_j$ in a linear combination of the unit sources are bounded from below by zero and from above +6.0 m, such a value of amplitude could be approached only as a result of a very strong underwater earthquake. In addition, such a restriction is not necessary and has been used only in order to accelerate the data processing of the Andreanov tsunami of 1996. When optimizing a set of the coefficients q_j it is important to have a rather small discretization step. The requested CPU time linearly decreases if this step is enlarged. In the case under consideration, the optimal step value happens to be at the level of 0.1–0.2 m.

3. Testing the algorithm against historical data of the 1996 Andreanov tsunami

The afore-mentioned algorithm has been numerically tested determining the source parameters for the Andreanov tsunami of 1996. The measured marigrams were available for the analysis along with the synthetic ones from the FACTS database. The earthquake epicenter and location of the recording stations AK70, AK71, AK72, AK73, and WC67 are shown in Figure 1. The station WC67 has a position close to the Eastern cost of the USA and is not shown in Figure 1. Due to a very distant location from the epicenter, the records from this station were basically excluded from the analysis of synthetic marigrams. The reason is that a synthetic tsunami wave arrived at this station some minutes earlier than the measured tsunami wave. This fact could be explained by rather a poor quality of the digital bathymetry. We would like to stress the necessity of continuation of efforts in obtaining a new digital bathymetry of the region in question and subsequent updating the FACTS database.

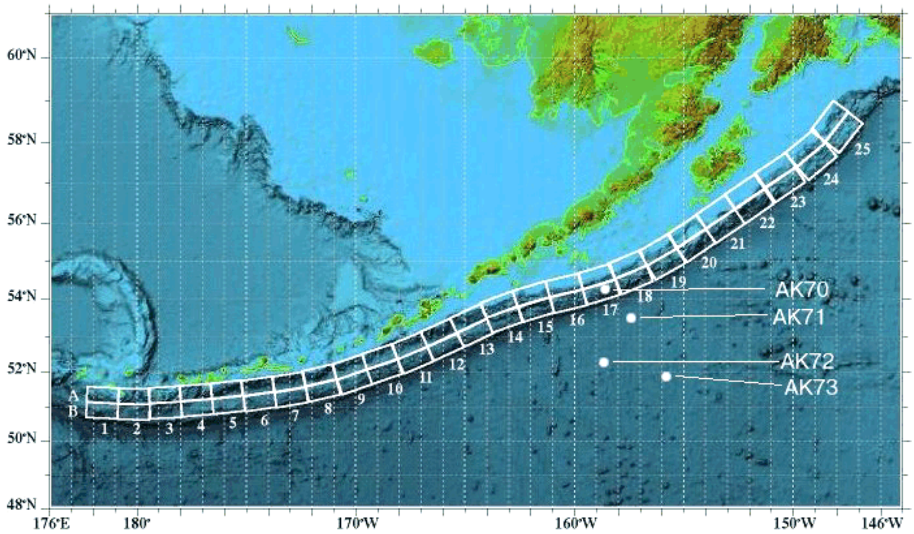


Figure 6. Location of the unit sources and the deep-ocean stations for the Andreanov tsunami of 1996

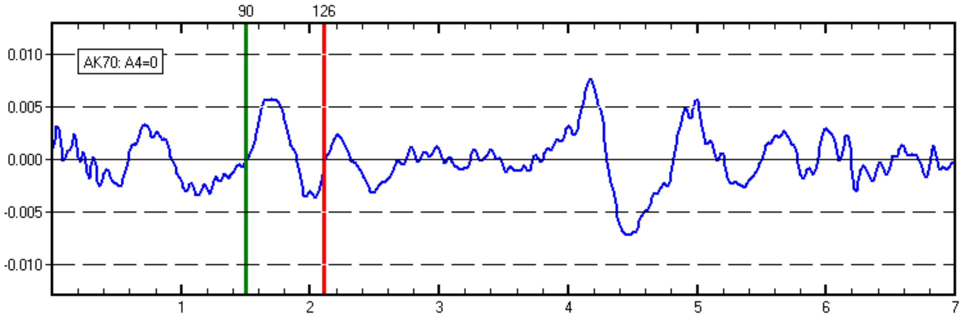


Figure 7. Wave time series (without tidal component), obtained at deep water stations

According to the seismic data, the epicenter of the Andreanov earthquake of 1996 is located near to the unit sources A4, A5, B4, B5 (Figure 6). The measured time series of the Andreanov tsunami are presented in the form of numerical sequences of the ocean level, recorded at all the five stations with a time step of 15 s. Before the minimization procedure, these data were filtered, because the FACTS database at the moment contains synthetic marigrams with a time step of 60 s. The real time series recorded by the station AK70 is displayed in Figure 7 (the tidal components excluded). In order to determine an optimal length of the time period for source parameters (amplitudes for a linear combination of the unit sources) identification, three series of optimization have been carried out. At first, the complete wave periods (both first positive and negative phases) were accounted for all the records. Then, the amplitude coefficients were obtained by processing only the parts of the real marigrams before the first maxima. Finally, 3/8 part of time series was taken into account (that is only the part of the first positive wave phase approximately from the beginning through a maximum and until the moment, when the height decreases to half the maximum). Our understanding of the first wave period on the recorded marigram is shown in Figure 7. Numerical values of these intervals are indicated in the table, which summarizes the results of the search for amplitude coefficients.

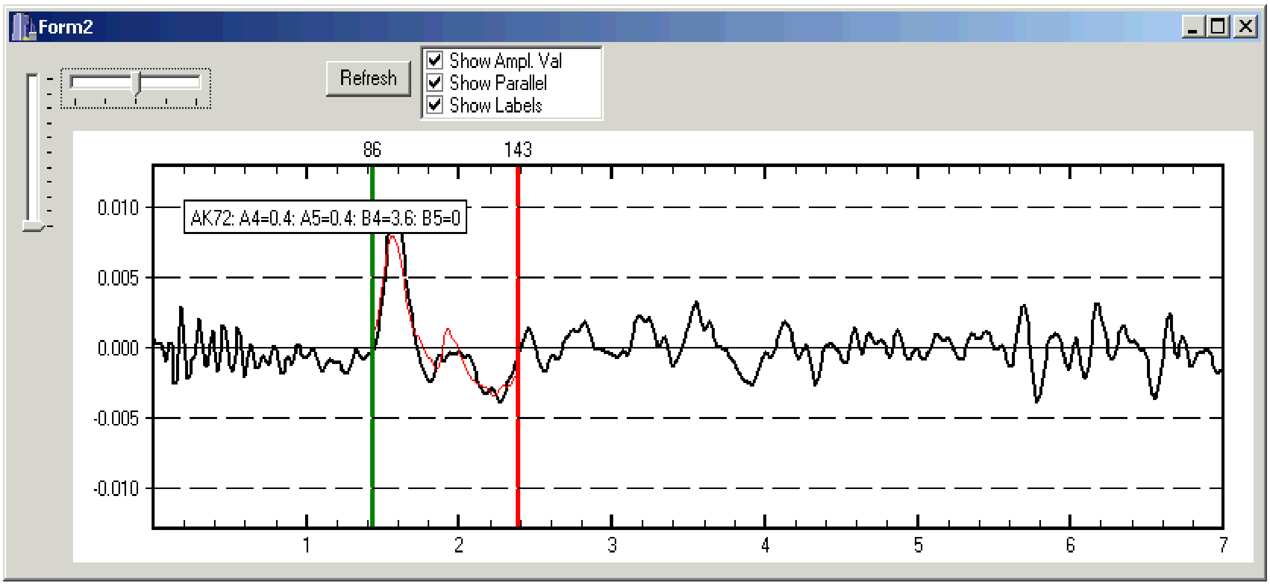
As was already mentioned, the discretization step for coefficients $\boldsymbol{q}_j$ (j =1, . . . , 50) was chosen as 0.2. Zero values of coefficients were obtained for all the unit sources, except for A4, A5, B4, and B5. Therefore, only these unit sources are mentioned in the table 1.

Table 1.

Stations in process	**Amplitudes at unit sources**				**Length of time interval, min**	**Norms values**
	A4	**A5**	**B4**	**B5**		
AK70 L2	0	0,6	3,6	0	One period	L2=0.061
L1	0	0,6	3,6	0	90-126	L1=0.046
AK71 L2	0,2	0,2	4,4	0,2	95-159	L2=0.121
L1	0,2	0,2	4,8	0		L1=0.104
AK72 L2	0,4	0,2	4,6	0	86-143	L2=0.109
L1	0,4	0	5,2	0		L1=0.087
AK73 L2	1	0,2	2,6	0,2	101-154	L2=0.193
L1	0,8	0	3,4	0		L1=0.147
WC67 L2	2,4	0	0	0,8	265-297	L2=0.268
L1	2	0	0	0,6		L1=0.204
AK (70+71) L2	0,2	0,4	3,4	0,4	One period	L2=0.108
L1	0,2	0,6	3	0,2		L1=0.084
AK (70+71+72) L2	0,2	0,4	4,2	0	One period	L2=0.116
L1	0,2	0,6	3,2	0		L1=0.096
AK (70+71+72+73) L2	0,4	0,4	3,6	0	One period	L2=0.145
L1	0,2	0	4,4	0,8		L1=0.116
AK70 L2	0	0	5,2	0,8	Before the first maximum	
L1	0	0	5,8	0,4		
AK (70+71+72+73) L2	1,4	0	2,8	0	Before the first maximum	
AK (70+71+72+73) L2	0,8	0	3,6	0,2	3/8 of period (70:90-110, 71:95-112, 72:86-100, 73:101-115)	L2=0.116
L1	1	0	3,6	0		L1=0.084
AK70 L2	0	0.4	3.8	0.4	From the signal start up to 115 min from the earthq.	
L1	0	0.6	3.6	0		
AK (70+71) L2	0	0	4.6	1	From the signal start up to 115 min from the earthq.	
L1	0	0.2	4.0	0.8		
AK (70+71+72) L2	0.2	0.2	4.6	0.2	From the signal start up to 115 min from the earthq.	
L1	0.2	0	5.0	0.4		
AK (70+71+72+73) L2	0.2	0	4.8	0.6	From the signal start up to 115 min from the earthq.	
L1	0.4	0	4.6	0.2		

Let us briefly describe the results obtained. The table contains the values of amplitude coefficients for A4, A5, B4, B5 unit sources, processing the data from one, two, three, or four DARTS stations, that is indicated in the left column along with the norm of comparison. The right column shows up numerical values of the norms of differences between the measured time series and linear combinations of synthetic ones, calculated with the amplification coefficients given in columns 2, 3, 4, and 5. The last four rows of the table demonstrate the values of amplification coefficients in the case, when data processing started 115 minutes after the earthquake. This means that a segment of each marigram, which is included into analysis, starts at the wave arrival moment and ends at the time of 115 minutes after a seismic event. This is the model of the situation, when at every time moment, the operator is able to work with marigrams of different duration.

The amplitude coefficients (calculated by optimization in L2 norm) taking the whole first period and the marigram at the station AK72, obtained from the "complex" source, is shown in Figure 8.


☑ Show Ampl. Val
☑ Show Parallel
☑ Show Labels

AK72: A4=0.4: A5=0.4: B4=3.6: B5=0

Figure 8. Comparison of a synthetic marigram from the combined source (light line), obtained by L1 norm optimization, against the real data (dark line) taking the whole first period

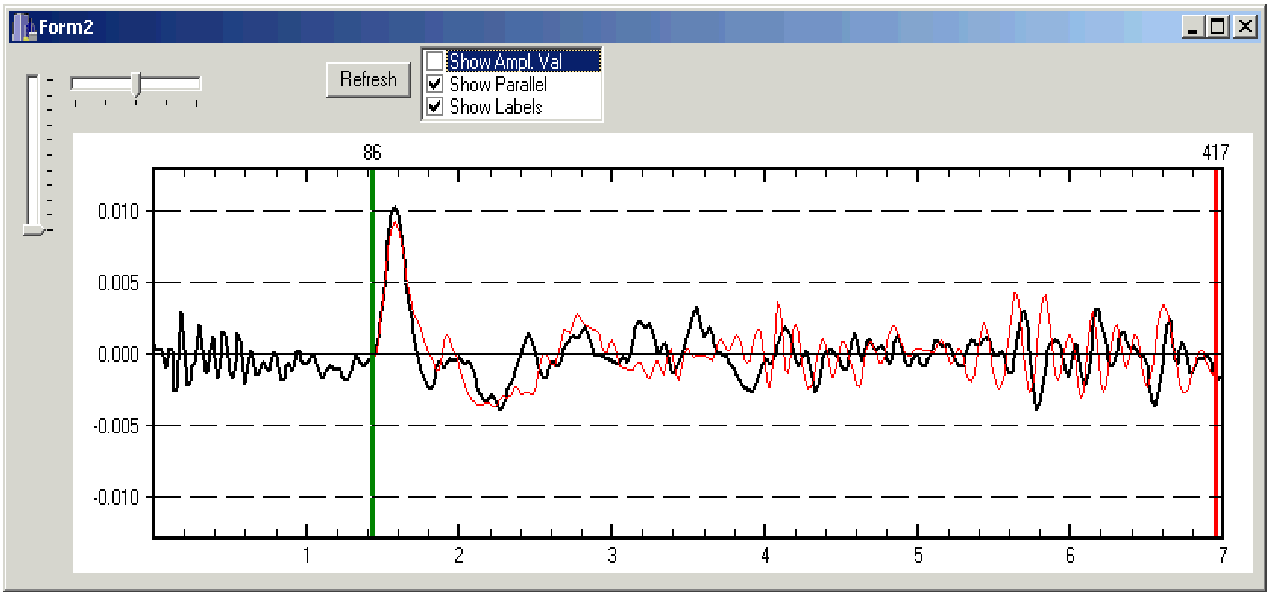
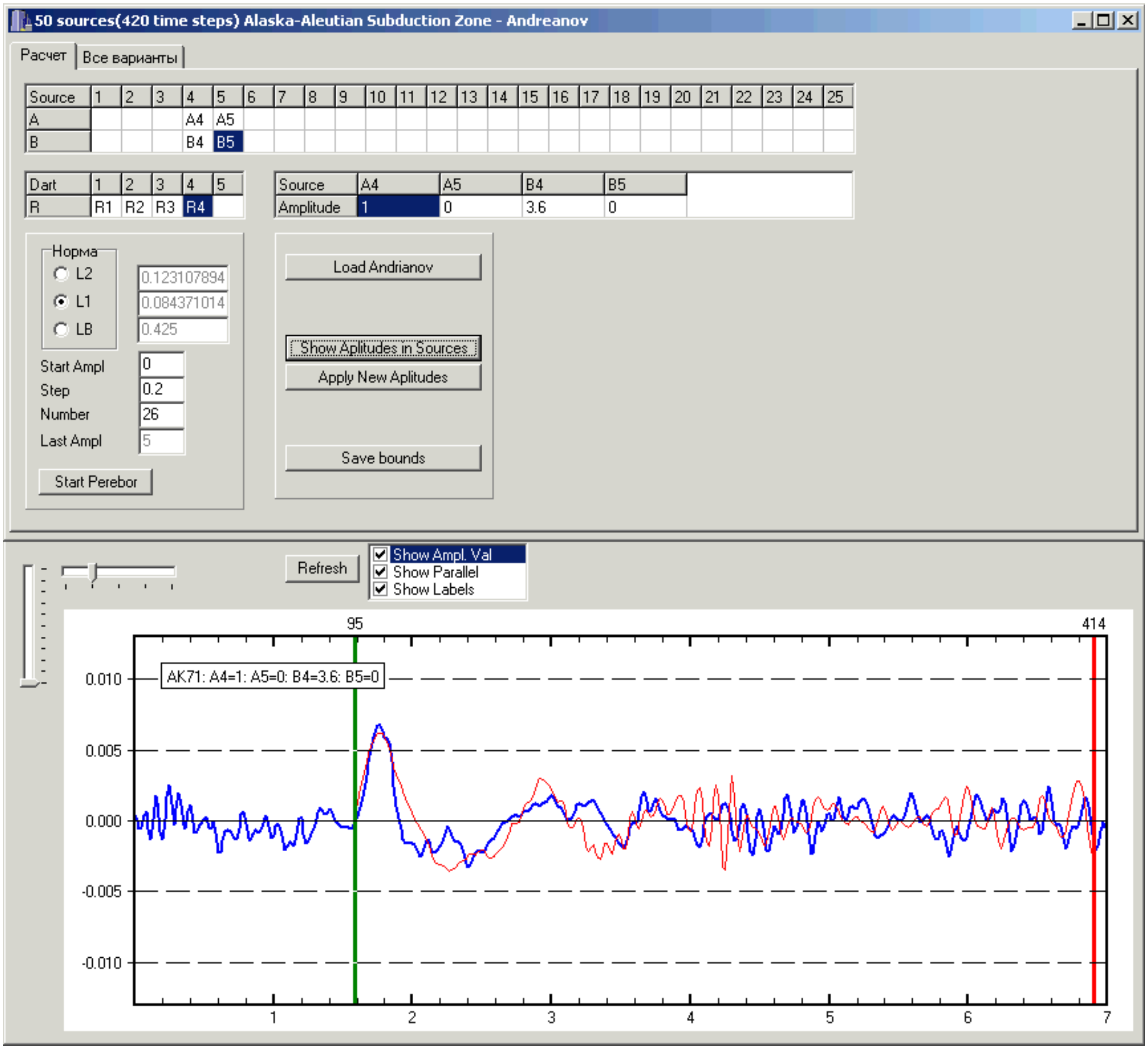

☐ Show Ampl. Val
☑ Show Parallel
☑ Show Labels

Figure 9. The same comparison as in Figure 8, accounting 3/8 of the first period data



Source	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
A				A4	A5																				
B				B4	B5																				

Dart	1	2	3	4	5
R	R1	R2	R3	R4	

Source	A4	A5	B4	B5
Amplitude	1	0	3.6	0

Норма
○ L2 0.123107894
● L1 0.084371014
○ LB 0.425

Start Ampl 0
Step 0.2
Number 26
Last Ampl 5

☑ Show Ampl. Val
☑ Show Parallel
☑ Show Labels

Figure 10. Graphic interface of the automated source definition system
(Module M1)

Figure 9 shows a comparison of a synthetic marigram from the combined source, obtained by L1 norm optimization, against the real data, accounting 3/8 of the first period data at the station AK72. Note that the application software works in real time. The source definition process takes 10–25 s using 1000 MHz CPU even in the case of four unit sources and five stations. The user interface screenshot for Module M1 is presented in Figure 10.

4. Application of the method for large (extended) tsunami sources

At small sizes of an actual source (when a linear combination includes up to six basis sources) searching of optimal coefficients can be done by only combinatory mode. That means looking trough all possible combinations of coefficients and calculating of values of a discrepancy for each of these combinations. However, if for three or four sources all process of searching of optimum coefficients takes several seconds, for six sources it is required about several minutes. The total number of processor operations at such approach increases as a constant in a power of an amount of considered basis sources. If the actual tsunami source will be very extended (as it was during the Indian ocean tsunami of 2004), process will take too much time to be usable in real-time mode. The quick algorithm for coefficients determination is proposed. Here the number of arithmetical operations required by processor linearly increases at taking into account the additional unit sources. The main idea of the method is to construct the chronological hierarchy of incoming waves (calculated time series) from basis sources. Then process of coefficients determination is carried out according to this sequence. The first couple of basis sources will be the ones with travel-time is equal (or very close) to the actual wave arrival time. On the first step the optimal coefficients for these two unit sources, which minimize the discrepancy value is to be determined (all the other coefficients are assumed to be zeroes). Then we chose another couple of basis sources that generate waves with a little bit later arrival times (to regarded DARTS station). Minimization of discrepancy is carrying out according to these two couples of unit sources. And so on. The number of arithmetic operations here linearly depends on the number of basis sources couples. For example, for Andreanov tsunami there are only two couples of such sources (A5, B5 and A4, B4). The algorithm has been tested on available records of an actual tsunami by deep water registering stations.

Conclusion

The main objective of this paper is to develop an inversion scheme that determines a tsunami source using deep-ocean tsunami records and pre- calculated wave time series from a set of the so-called “unit” sources. The method is fast and robust, which makes it appropriate for the real-time data assimilation scheme of the tsunami forecast system. The user interface for a quick definition of tsunami source parameters was developed and tested against the deep-ocean records of the weak 1996 Andreanov tsunami. Method can successfully work for extended tsunami sources, which consists from the big number of basis sources.

References

[1] Marchuk An.G., Shevchuk A.A., Titov V.V. Tsunami waves propagation along the waveguides // Bull. Novosibirsk Comp. Center. Ser. Math. Model. in Geoph. — Novosibirsk, 2003. — Iss. 8. — P.79–89.

[2] Gonzalez F.I., Titov V.V., Avdeev A.V., Bezhaev A.Yu., Lavrentiev M.M. (jr.), Marchuk An.G. Real-time tsunami forecasting: challenges and solutions // Proc. International Conf. Mathematical Methods in Geophysics–2003.— Novosibirsk: ICM&MG Publisher, 2003. — P. 225–228.

[3] Bernard E.N., Gonzalez F.I., Meining Ch., Milburn H.B. Early detection and real-time reporting of deep-ocean tsunamis // ITS 2001 Proceedings. Review of the U.S. National Tsunami Hazard Mitigation Program. Seattle, Washington, August 7, 2001, NOAA/PMEL, 7600 Sand Point Way NE, Seattle, WA 98115,
2001. — P. 85–96.