Nat. Hazards Earth Syst. Sci., 12, 1583–1606, 2012
www.nat-hazards-earth-syst-sci.net/12/1583/2012/
doi:10.5194/nhess-12-1583-2012

Natural Hazards and Earth System Sciences

Detecting the 11 March 2011 Tohoku tsunami arrival on sea-level records in the Pacific Ocean: application and performance of the Tsunami Early Detection Algorithm (TEDA)

L. Bressan and S. Tinti

Department of Physics, University of Bologna, Bologna, Italy

Correspondence to: L. Bressan (lidia.bressan@unibo.it)

Received: 26 January 2012 – Revised: 3 April 2012 – Accepted: 3 April 2012 – Published: 21 May 2012

Abstract. Real-time detection of a tsunami on instrumental sea-level records is quite an important task for a Tsunami Warning System (TWS), and in case of alert conditions for an ongoing tsunami it is often performed by visual inspection in operational warning centres. In this paper we stress the importance of automatic detection algorithms and apply the TEDA (Tsunami Early Detection Algorithm) to identify tsunami arrivals of the 2011 Tohoku tsunami in a real-time virtual exercise. TEDA is designed to work at station level, that is on sea-level data of a single station, and was calibrated on data from the Adak island, Alaska, USA, tide-gauge station. Using the parameters' configuration devised for the Adak station, the TEDA has been applied to 123 coastal sea-level records from the coasts of the Pacific Ocean, which enabled us to evaluate the efficiency and sensitivity of the algorithm on a wide range of background conditions and of signal-to-noise ratios. The result is that TEDA is able to detect quickly the majority of the tsunami signals and therefore proves to have the potential for being a valid tool in the operational TWS practice.

1 Introduction

On 11 March 2011, an offshore $M = 9.0$ seismic event struck the north-east region of Japan and generated a major tsunami that devastated the Japanese coasts and propagated across the Pacific up to transoceanic distances (Chang and Chao, 2011; Mori et al., 2011; Pollitz et al., 2011; Shao et al., 2011). For the Japanese coasts facing the source, the tsunami soon followed the earthquake and the population had very little time to react to the first warning that was based only on seismic data. However, for more distant places of the Pacific Ocean, after the near-field observations confirmed the formation of a big tsunami, information about the expected tsunami arrival times and about its expected characteristics were very important for tsunami warning.

The worldwide public and scientific awareness of the tsunami hazard and of the importance of implementing appropriate preventive measures to favour resilience of coastal communities after a tsunami disaster has increased, starting from the Sumatra tsunami in 2004. Since then, intergovernmental institutions such as IOC and individual countries facing the sea have started programs to strengthen the existing Tsunami Warning Systems (TWS) or to build new ones, which include expanding and updating sea-level monitoring networks and developing analysis of real-time sea-level data (Synolakis and Bernard, 2006; Satake and Atwater, 2007; Schindelé et al., 2008; Rudloff et al., 2009). In general, a TWS, either regional (i.e. supranational in the IOC terminology) or national, is structured to acquire and process data and to issue tsunami alerts, whereas emergency procedures after an alert issuance fall under the responsibility of local authorities. Sea-level monitoring networks are a very important module of a TWS because, nowadays, the only way to know whether a tsunami has been generated or not is to verify if tsunami waves leave a measurable signal in ocean (coastal and offshore) sensors. Of course, in addition to the binary information (tsunami: yes or no), sea-level data supply further information, such as wave amplitude and arrival time at the sea-level station, and consequently allow one to understand the level of emergency and to react as needed. Therefore, tsunami detection in sea-level records and retrieval of tsunami data play or should play a key role in any TWS,

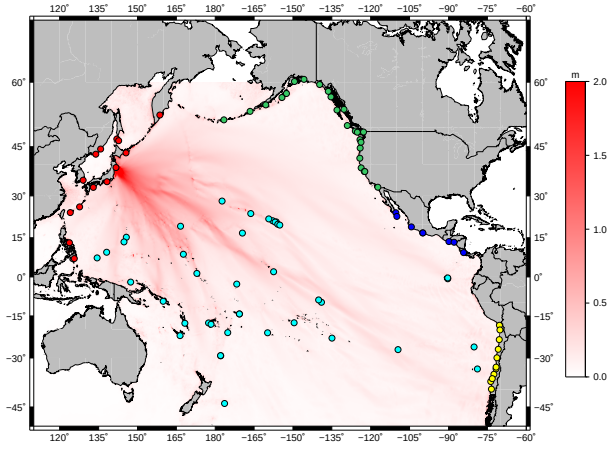

Fig. 1a. Map of the Pacific Ocean with the simulated field of the maximum sea surface elevation for the 11 March 2011 Tohoku tsunami with the tide-gauge stations considered in the analysis, from all Pacific coasts: Asia (red), North America (green), Central America (blue), South America (yellow) and Pacific Islands (light blue). The simulation (Tinti et al., 2011) is based on the initial condition by Höchner et al. (2011).

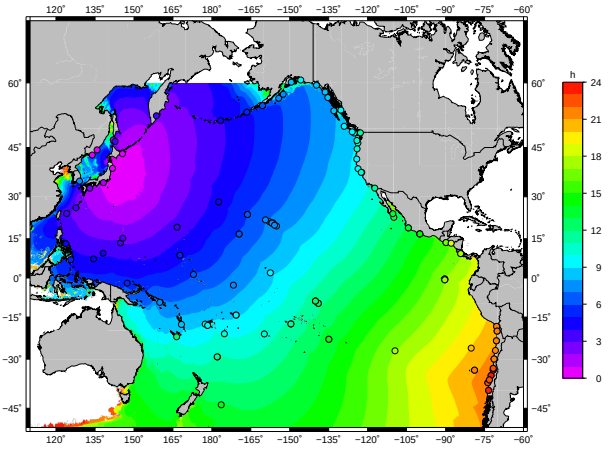

Fig. 1b. Computed tsunami travel times (TTT) for the tsunami leading wave, that reaches the farthest stations in more than 23 h. TTT is derived from numerical tsunami simulation and is defined here as the minimum propagation time for which the modulus of the sea surface elevation exceeds 2 cm in deep sea (depth larger than 1000 m). Station circles are colored according to the associated observed TTT, which makes it easier to recognize possible discrepancies between observed and computed TTT.

which is testified also by past and recent interest and studies (Mcgehee and McKinney, 1997; Mofjeld, 1997; Okal et al., 1999; Henson et al., 2006; ; Shimizu et al., 2006; Farreras et al., 2007; Casey et al., 2008; Omira et al., 2009; Šepić et al., 2009; Falck et al., 2010; Stosius et al., 2010; Beltrami, 2011; Beltrami and Di Risio, 2011; Gurgel et al., 2011).

The 2011 Japan tsunami is the first major tsunami that occurred after the fast developing phase, still ongoing, of TWSs that started after the 2004 Sumatra tsunami, and for this reason it is the best recorded major tsunami in history. We took therefore this opportunity to apply an algorithm for the automatic detection of tsunamis on sea-level records. To this purpose, we used TEDA, a Tsunami Early Detection Algorithm that was developed by the tsunami research team of the University of Bologna with the goal of detecting tsunamis and high-amplitude long-period waves (Bressan and Tinti, 2011). TEDA was devised to work real-time on a single station data and can be adapted to the specific background conditions of the station site. The calibration procedure includes the analysis of background signals and of tsunami events, if available, or of simulated tsunami signals. In this work, we applied TEDA without a proper calibration process, since we used the parameter configuration (and also some variants of it) that was found to optimize the TEDA performance for the Adak island tide-gauge station, in Alaska, USA.

2 Dataset and preliminary data analysis

The records that were processed with TEDA are the majority of the 1-min sampling sea-level data collected by the IOC from coastal, tide-gauge stations around the Pacific Ocean that are made available online at the website http://www.ioc-sealevelmonitoring.org. In total, as many as 123 sea-level signals were analyzed. These come from 97 sites, including 24 sites supplying two records and one site with three records from different types of sensors. Stations are mapped in Fig. 1a, together with the maximum sea surface elevation in the Pacific Ocean resulting from tsunami simulation modelling (Tinti et al., 2011 and Höchner et al., 2011, for the source). From the figure, it may be seen that these stations are all distributed along the coast of the Pacific (Asia, North-, Central-, and South-America) and in the mid-Pacific islands (Oceania) with the exception of Australia and New Zealand mainland. Each tide-gauge recording is denoted with the same code used in the IOC website (which is derived from the GLOSS station name) and, in case of multiple-record sites, with the additional index of the data channel (e.g. the code adak is used for the already mentioned Adak island station, and codes acaj_d1, acaj_d2 are used for the two records from Acajutla, El Salvador, station, the former from a pressure gauge and the latter from a radar gauge). The complete list of the records is given in Table 1, sorted in alphabetical order of the station code (first column), with indication of the sensor type (last column). For further information, one may refer to the IOC site.

The records selected for the analysis form a dataset of tsunami signals that come from sites differing from each other as regards the distance from the source, ranging from near- to far-field stations, and the background characteristics

Table 1. List of the stations selected for the analysis. The station code (first column) is the same used on the IOC website. The type of instrument (last column) is coded as follows in the IOC database: bwl= backup water level gauge, enc = encoder sensor, flt = float gauge, prs= pressure gauge, pwl = primary water level gauge, rad = radar gauge, wls= water level sensor. In stations with a double water level sensor, often one is assumed as the pwl and the other as the bwl. Only stations with 1-min sample rate recordings have been selected from the IOC database.

code	Country	Location	Sensor
acaj	El Salvador	Acajutla	d1: prs, d2: rad
acap	Mexico	Acapulco	d1: prs, d2: rad
acya	Mexico	Acapulco, Club De Yates	d1: flt
adak	USA	Adak, AK	d1: wls
alak	USA	Alitak, AK	d1: pwl
alam	USA	Alameda, CA	d1: pwl
anto	Chile	Antofagasta	d1: prs, d2: rad
aren	USA	Arena Cove, CA	d1: pwl
aric	Chile	Arica	d1: prs, d2: rad
asto	USA	Astoria, OR	d1: pwl
balt	Ecuador	Baltra, Galapagos	d1: prs, d2: rad
bamf	Canada	Bamfield	d1: enc
busa	Korea	Busan	d1: rad
cabo	Mexico	Cabo San Lucas	d1: enc
cald	Chile	Caldera	d1: prs
chit	New Zealand	Chatham Island	d1: prs
chrp	USA	Cherry Point, WA	d1: pwl
const	Chile	Constitucion	d1: prs, d2: rad
coqu	Chile	Coquimbo	d1: prs, d2: rad
cord	USA	Cordova, AK	d1: pwl
corr	Chile	Corral	d1: prs, d2: rad
cres	USA	Crescent City, CA	d1: pwl
davo	Philippines	Davao	d1: prs
east	Chile	Easter Island	d1: prs
elak	USA	Elfin Cove, AK	d1: pwl
fpnt	USA	Fort Point, San Francisco, CA	d1: wls
fren	USA	Tern, Fr. Frigate, HI	d1: prs, d2: rad
hana	Japan	Hanasaki	d1: enc
hens	Canada	Henslung Cove	d1: enc
hilo	USA	Hilo, HI	d1: wls
hiva	France	Hiva Oa (Marquesas, French Polynesia)	d1: pr2, d2: prs
hono	USA	Honolulu, Oahu, HI	d1: wls
iqui	Chile	Iquique	d1: prs, d2: rad
ishig	Japan	Ishigakijima	d1: prs
john	USA	Johnston Atoll	d1: prs, d2: rad
juan	Chile	Juan Fernandez	d1: prs, d2: rad
kahu	USA	Kahului, Maui, HI	d1: pwl
kant	Kiribati	Kanton	d1: rad
kaum	USA	Kaumalapau, HI	d1: prs, d2: rad
kawa	USA	Kawaihae, HI	d1: pwl
khol	Russia	Kholmsk	d1: rad
kodi	USA	Kodiak Island, AK	d1: wls
kors	Russia	Korsakov	d1: rad
kwaj	Marshall Islands	Kwajalein	d1: wls
lajo	USA	La Jolla, CA	d1: bwl
laun	El Salvador	La Union	d1: prs
lebu	Chile	Lebu	d1: prs, d2: rad
lega	Philippines	Legaspi	d1: prs, d2: rad
levu	Fiji Islands	Lautoka	d1: prs
lomb	Papua New Guinea	Lombrum Manus Is	d1: prs
lpaz	Mexico	La Paz	d1: flt

Table 1. Continued.

code	Country	Location	Sensor
mala	Palau Islands	Malakal	d1: prs
manz	Mexico	Manzanillo	d1: prs, d2: rad
midx	USA	Midway Islands	d1: bwl
naha	Japan	Naha	d1: enc
nawi	USA	Nawiliwili, Kauai, HI	d1: pwl
nkfa2	Tonga Island	NukuAlofa Harbour wharf	d1: rad
nuku	France	Nuku Hiva (Marquesas, French Polynesia)	d1: prs, d2: rad , d3: rad
numbo	France	Nouméa - Numbo (New Caledonia)	d1: rad
ofun	Japan	Ofunato	d1: enc
omae	Japan	Omaezaki	d1: enc
pagb	USA	Pago Bay, Guam	d1: wls
pago	Samoa	Pago Pago, AS (web)	d1: pwl
pagx	Samoa	Pago Pago	d1: pwl
pape	France	Papeete (Tahiti, French Polynesia)	d1: prs, d2: rad
petr	Russia	Petropavlovsk	d1: rad
preo	Russia	Preobrazheniye	d1: rad
prin	Canada	Prince Rupert	d1: enc
quepo	Costa Rica	Quepos	d1: prs, d2: rad
raro2	Cook Islands	Avatiu Rarotonga harbour wharf	d1: rad
rbct	New Zealand	Raoul Island Boat Cove	d1: prs
rfrt	New Zealand	Raoul Island Fishing Rock	d1: prs
riki	France	Rikitea (Gambier, French Polynesia)	d1: prs, d2: rad
rudn	Russia	Rudnaya Pristan	d1: rad
saip	USA	Saipan, Northern Mariana Is.	d1: prs
sanf	Chile	San Felix	d1: prs
sano	Chile	San Antonio	d1: prs, d2: rad
sant	Ecuador	Santa Cruz, Galapagos	d1: prs, d2: rad
sbea	USA	South Beach, CA	d1: wls
sdpt	USA	Sand Point, AK	d1: wls
sewa	USA	Seward, AK	d1: wls
sitk	USA	Sitka, AK	d1: wls
solo	Solomon Islands	Honiara	d1: prs
talc	Chile	Talcahuano	d1: prs, d2: rad
tara	Kiribati	Betio, Tarawa	d1: prs
tofi	Canada	Tofino	d1: enc
tosa	Japan	Tosashimizu	d1: enc
unal	USA	Unalaska, AK	d1: wls
valp	Chile	Valparaiso	d1: prs, d2: rad
vanu	Vanuatu	Port Villa, Vanuatu	d1: prs
viti	Fiji Islands	Suva Viti Levu	d1: prs
wake	USA	Wake Island	d1: wls
wint	Canada	Winter Harbour	d1: enc
wpwa	USA	Westport, WA	d1: pwl
xmas	Kiribati	Christmas	d1: prs
yaku	USA	Yakutat, AK	d1: pwl
yapi	Micronesia	Yap Island	d1: prs

that are knowingly site- and time-dependent. Several records have a duration of many days, but most are shorter and cover only the first part of the tsunami oscillations. Indeed, also one record of a gauge that broke at the arrival of large tsunami waves was included in our data set. For each record we determined by inspection the tsunami arrival time (TAT) and, where possible, the end of the tsunami oscillations. It is important to stress that the estimaton of TAT is not intended to be the exact tsunami arrival time, but rather the time when the tsunami was at first visible in the record. It is worth noting that such estimated values depend on subjective interpretation and may be strongly influenced from the level of the signal-to-noise ratio.

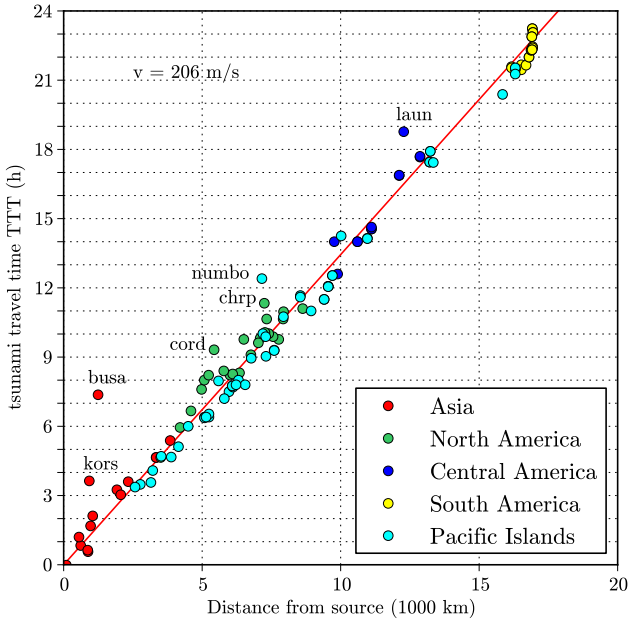

Fig. 2. Observed Tsunami Travel Times (TTT) plotted vs. distance between the tide-gauge station and the earthquake epicenter, taken at the position (38.297 N, 142.372 E) following USGS (see USGS website http://earthquake.usgs.gov/earthquakes/eqinthenews/2011/usc0001xgp/). Distances have been computed using the Clarke earth ellipsoid (coded EPSG 7008). The fitting straight line (red) slope corresponds to the tsunami speed $c = 206\,\mathrm{m\,s^{-1}}$ and the ocean depth $d = 4330$ m ($c^2 = gd$). Most of the stations are quite close to the straight line. Only a few outliers are very much delayed. They are marked with the station code (see Table 1).

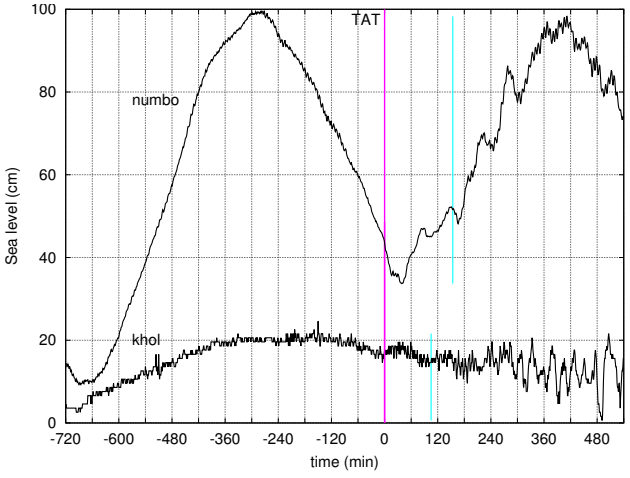

Fig. 3a. Sea level records of gauges in Nouma-Numbo, French New Caledonia (numbo), and in Kholmsk, on the south-west coast of Sakhalin island, Russia (khol). Records still include tide. Records are shown as examples of problematic identification of the TAT and of the first wave period P1. The estimated end of the first wave is indicated by a light blue line. Time is measured from the TAT.

Using TEDA terminology (Bressan and Tinti, 2011), these temporal parameters were used to define the Tsunami Interval (TI), that is the part of the signal between the TAT and the end of the tsunami oscillations or the end of the record if the record terminates before the end of the tsunami. We further introduced the Tsunami Detection Interval (TDI), that is defined as a 3-hour long interval starting from the TAT and is used for assessing the performance of the tsunami detection module of TEDA (described in the next section): if the module makes a detection within the TDI, it is considered a success, otherwise it is a failure. However, we will give more importance to detections occurring in the first hour of TDI, since in a TWS perspective the sooner the tsunami is seen, the better it is.

The Tsunami Travel Time (TTT), obtained by subtracting the earthquake origin time from TAT, is shown in Fig. 2 for all stations and listed in Table 2. It spans an interval slightly exceeding 23 h and is in good agreement with the numerical TTTs (see Fig. 1b) computed by Tinti et al. (2011), with a few exceptions. The plot of TTT vs. the distance between the station and the earthquake epicentre (USGS solution), measured along the Clarke earth ellipsoid (equatorial axis a=6378.2064 km, inverse flattening 1/f=294.9786982), shows a surprisingly very good alignment. The fitting straight line of the type $y = ax$ has a slope corresponding to an average tsunami velocity c of $206\,\mathrm{m\,s^{-1}}$ and to an average ocean depth d of 4330 m (if the relation $c^2 = gd$ is assumed, with g = gravity acceleration). Of course, several stations exhibit anticipated or delayed arrivals by an amount of 30–60 min, which is quite normal since tsunami speed is mainly governed by the bathymetry, which is quite irregular in the ocean. However, there are some stations where arrivals are much more delayed. These are found in the near-field (less than 2000 km) such as Busan (South Korea) and Korsakov (Sakhalin island, Russia), where the delay exceeds 5 h and 2 h respectively, but also in the intermediate field (between 5000–10 000 km), where delays for the stations of Cordova (Alaska, USA), of Cherry Point (Washington, USA) and Nouméa-Numbo (New Caledonia, France; see the original record in Fig. 3a) exceed 2 h, and in the far field (beyond 10 000 km), where the delay at La Union (El Salvador) is about 2 h. The main reason for such delayed arrivals is to be attributed mostly to local conditions. For instance, La Union is located inside a bay, Nouméa-Numbo is located in New Caledonia's main island facing the Coral Sea, which is protected by a coral reef. Furthermore, Cordova is located inside a fjord in Alaska, and Cherry Point, close to the border between USA and Canada, is protected by Vancouver island. However, in the near-field, the reason for delays is more related to tsunami propagation that, west of Japan, cannot be adequately represented by the straight line fit of Fig. 2. The tsunami reaches the port of Korsakov, that is located in the

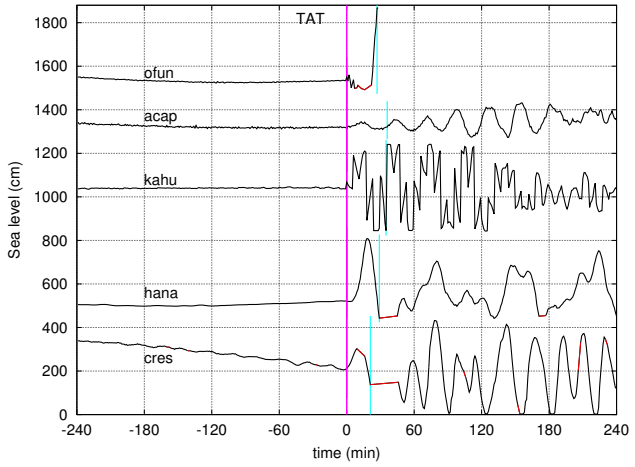

Fig. 3b. Examples of complete and incomplete records. Records are not detided and are aligned starting from TAT. The end of the first wave is indicated with a light blue line. Filled data are in red. The record of Acapulco, Mexico (acap) does not give any interpretation problem since it has no gaps and no saturation. The record of Kahului, Maui island, Hawaii (kahu) presents saturation in the negative level during the first wave oscillation, and later also for positive oscillations. The records of Hanasaki, Japan (hana) and of Crescent City, California (cres), show substantial data gaps that make it difficult to estimate both period and amplitude of the first wave. Ofunato, Japan (ofun), gives an example of a suddenly truncated record, since the tide gauge was broken by the violence of the wave. It also shows some gaps just before the large peak. The first anomaly probably refers to direct seismic effects rather than tsunami. Estimates from incomplete records are considered uncertain.

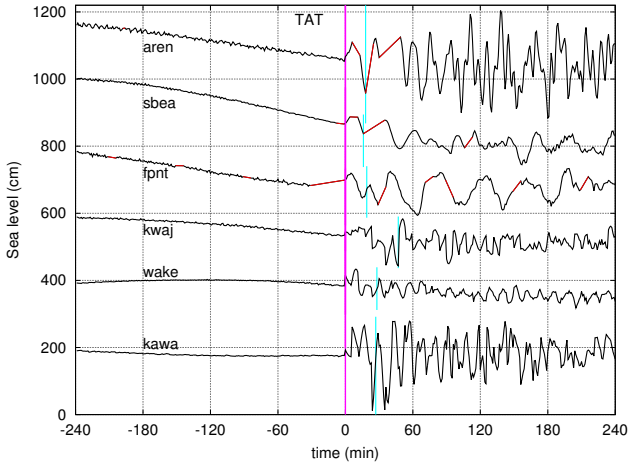

Fig. 3c. Examples of analyzed records given to show how the first tsunami wave has been characterized in case of missing data and in case of complex signals. Records are not detided and are aligned starting from TAT. The end of the first wave is indicated with a light blue line. The records of Arena Cove, California (aren), of South Beach (sbea) and Fort Point (fpnt) both in San Francisco, California, have substantial data gaps making it hard to identify TAT as well as duration and magnitude of the first wave. The records of the Kwajalein atoll, Marshall islands (kwaj), of the Wake island, USA (wake) and of Kawaihae, Hawaii (kawa) are quite complex and the first tsunami wave is modified by short-period oscillations.

Aniva bay at the southern end of Sakhalin island (Russia), after travelling around the island of Hokkaido, North Japan; similarly, to arrive at the port of Busan in South Korea, the tsunami has to go around the island of Kyushu, South Japan. It is interesting also to notice that in the record of Ofunato, in Japan (ofun), the first anomalous signal is observed just before the earthquake (TTT = −1 min, see also Fig. 3b), and is a small amplitude oscillation of difficult interpretation since probably the sample rate (1 sample min^{-1}) is too low. Furthermore, in the two Russian stations that are located on the western coast of the Sea of Japan, i.e. Preobrazheniye (preo) and Rudnaya Pristan (rudn), an anomaly is observed in the records quite soon and is probably due to the arrival of a small local tsunami, generated by the far-field effects of the vertical co-seismic deformation that also involved the floor of the Japan Sea. Here the main tsunami arrived later with stronger waves that probably reached the Sea of Japan from the Pacific Ocean passing through the Tsugaru Straits.

We characterised the first tsunami wave in the records by evaluating the period P1, the maximum (positive) sea level or wave crest M1, the minimum (negative) sea level or wave trough m1 and the wave range R1, that is defined as R1 = M1-m1. Another useful parameter is the maximum amplitude of the first wave MA1 that is defined as MA1 = max(M1,-m1). Since TEDA handles missing data by filling sea-level data gaps with linear trends as soon as a new datum becomes available, the properties and information drawn about the tsunami signal from the time series refer to the original ones with gaps interpolated. In addition, we stress that all the estimated values of these parameters characterising the tsunami signal are subjected to interpretation and uncertainties. The identification of the TAT and of the first wave was accomplished by visual inspection; first we identified TAT and the wave period P1, and then in the time interval between TAT and TAT + P1, we evaluated M1, m1 and R1 (by removing from sea-level a rough estimation of the tide). In order to account for uncertainties, we used a gross approximation for the estimated sea level values. They were discretised in intervals of 10 cm and were given the mid-value of such intervals: for example, all values from 0 to 9 cm were rounded to 5 cm, those between 10 to 19 were approximated to 15 cm, etc. Estimates of P1, MA1 and R1 may be found in Table 2. It is further observed that bad-quality data can make it difficult to estimate the above parameters. There are three main factors affecting the integrity of the records and therefore degrading their quality: (1) record truncation due to station breakdown; (2) data gaps that can also affect the beginning of the tsunami signal; (3) record saturation, cutting

Table 2. Tsunami signal characteristics and results of TEDA detection: TTT, observed tsunami travel time; MA1, maximum amplitude (positive or negative) of the first wave; R1, range of the first wave; MA(TDI) maximum amplitude (positive or negative) in the tsunami detection interval; R(TDI) tsunami range in the TDI; P1, period of the first wave; DT delay time of the secure and tsunami detection; NF, the number of false detections of the tsunami detection module. An asterisk denotes that the value is uncertain owing either to missing data or to record saturation. The tsunami travel times refer to the original earthquake source indicated by USGS at 05:46 UTC. The results of TEDA secure detection and tsunami detection modules refer to parameters configuration C4. When the secure detection module makes a detection before the tsunami arrival DT is set equal to zero. This occurs only for the two records of Constitucion (const_d1 and const_d2).

	Tsunami signal characteristics						Secure detection	Tsunami detection	
Station code	TTT (h)	MA1 (cm)	R1 (cm)	MA(TDI) (cm)	R(TDI) (cm)	P1 (min)	DT (min)	DT (min)	NF
acaj_d1	16.87	15	25	25	45	41	88	7	0
acaj_d2	16.88	15	25	25	45	41	62	missed	0
acap_d1	14.55	15	35	75	155	34	28	9	0
acap_d2	14.55	15	35	75	155	36	30	9	0
acya	14.63	25	45	85*	155*	30	22	3	0
adak	4.65	35	75	75	125	25	10	3	0
alak	7.60	5	5	5*	5*	37	missed	missed	0
alam	10.97	25	45	55	105	39	17	3	0
anto_d1	21.67	25	35	75	155	16	35	missed	0
anto_d2	21.45	15	25	75	135	33	50	16	0
aren	9.77	95*	145*	165*	325*	18	13	2	0
aric_d1	21.58	35	65	105	205	49	43	missed	0
aric_d2	21.53	35*	45*	105	195	71	45	4	0
asto	10.65	15	15	15	15	27	missed	15	0
balt_d1	17.45	75	125	95	175	46	20	14	0
balt_d2	17.45	75	115	85	175	46	21	16	0
bamf	9.82	25	35	45	85	21	23	5	0
busa	7.37	15	25	15	25	54	missed	missed	0
cabo	12.60	15	35	25	45	20	missed	5	0
cald	21.65	35	65	205	395	34	21	14	0
chit	14.25	35	55	65	125	64	72	28	0
chrp	11.33	5	5	5	15	74	missed	missed	0
const_d1	23.23	175	255	215*	365*	38	0	missed	2
const_d2	23.23	155	245	175*	335*	39	0	missed	4
coqu_d1	22.00	45	75	185	375	37	17	14	0
coqu_d2	22.00	35	65	175	355	37	18	12	0
cord	9.32	5*	5*	5*	5*	123	missed	missed	0
corr_d1	23.07	95	165	165*	285*	41	14	2	0
corr_d2	23.07	95	165	155*	285*	42	14	2	0
cres	9.88	95*	155*	245*	435*	19	7	5	1
davo	5.38	45	65	45	85	63	48	11	0
east	17.43	35	65	75	155	49*	45	5	0
elak	8.22	15	15	15*	25*	76	missed	missed	0
fpnt	10.65	45	75	75*	125*	19*	15	4	0
fren_d1	6.40	45	75	55*	105*	36	24	missed	1
fren_d2	6.37	45	65	55	95	37	26	4	0
hana	0.83	285*	365*	285*	395*	23*	12	8	0
hens	8.32	25	45	55	115	32	27	10	0
hilo	8.00	165	265	165	275	29	6	0	0
hiva_d1	12.53	125	215	145	275	29	5	1	0
hiva_d2	12.53	125	215	155	285	29	5	1	4
hono	7.50	75	145	75	145	42	10	6	0
iqui_d1	21.52	15	35	95	185	18*	17	58	0
iqui_d2	21.47	25	35	95	175	21	19	2	0
ishig	3.60	5	5	25	35	45	71	51	0
john_d1	6.40	25	25	25	35	46	missed	14	0
john_d2	6.53	15	25	15	25	39	missed	12	0

Table 2. Continued.

	Tsunami signal characteristics						Secure detection	Tsunami detection	
Station code	TTT (h)	MA1 (cm)	R1 (cm)	MA(TDI) (cm)	R(TDI) (cm)	P1 (min)	DT (min)	DT (min)	NF
juan_d1	21.53	15	15	55	105	26	55	44	0
juan_d2	21.27	15	15	55	95	45	71	60	0
kahu	7.70	195*	365*	205*	395*	35	7	1	0
kant	7.80	5	5	5	15	45*	missed	missed	0
kaum_d1	7.75	55	105	105	185	22*	8	0	0
kaum_d2	7.75	55	105	125*	215*	21*	8	0	0
kawa	7.80	165	255	165*	265*	27*	9	6	0
khol	1.68	5	5	5	5	105*	missed	missed	0
kodi	8.00	15	15	15*	25*	58	missed	missed	0
kors	3.63	5	5	15	25	45	missed	missed	0
kwaj	5.12	75	115	75	145	47	28	2	0
lajo	11.10	25	35	25*	55*	40	23	4	0
laun	18.77	5	5	5	5	35	missed	missed	0
lebu_d1	22.30	45	75	175*	285*	41	23	17	0
lebu_d2	22.30	25	55	175*	275*	43	24	18	0
lega_d1	4.65	35	55	35	55	64	40	7	0
lega_d2	4.65	35	45	35	45	65	41	9	0
levu	10.02	35	55	35	55	50	15	6	0
lomb	6.00	65	95	105	185	62	17	7	1
lpaz	14.00	5	5	15	35	47	136	missed	1
mala	4.70	5	15	15	25	47	missed	missed	0
manz_d1	14.00	85	135	125	255	29	7	missed	1
manz_d2	14.00	65	105	105	195	29	20	6	0
midx	4.67	55	105	155	255	17	10	6	0
naha	3.25	25	25	55	105	31	28	7	0
nawi	7.20	85	135	85	175	31	13	6	0
nkfa2	10.75	55	85	65	125	58	40	9	1
nuku_d1	12.07	65	115	155	295	20	7	missed	2
nuku_d2	12.07	65	115	155	315	20	7	1	0
nuku_d3	12.05	55	105	155*	305*	20	8	3	0
numbo	12.40	5	15	5	15	154	missed	missed	0
ofun	−0.02	325*	375*	335*	375*	27*	13	2	0
omae	1.20	105*	155*	145*	215*	30*	10	6	0
pagb	3.48	25	35	35	55	35	42	4	0
pago	9.30	15	15	65*	125*	24	81	7	0
pagx	9.28	15	15	65	125	29*	106	49	0
pape_d1	11.50	15	35	45	85	29	28	9	0
pape_d2	11.50	15	35	45	75	31	30	11	0

the positive and negative peaks of the oscillations, and thus compromising amplitude estimates. In addition, evaluations by visual inspection can be hard if signals are weak with respect to the background noise. Figure 3a–3c provide illustrative examples of tide-gauge records where the identification of all or some of the parameters mentioned above is problematic and gives rise to uncertainties. Figure 3a shows two records from the stations of Nouméa-Numbo, French New Caledonia, and of Kholmsk, Russia, where determining TAT and P1 by visual inspection is not easy. Figure 3b gives an example of a truncated record from Ofunato, Japan, of a saturated record from Kahului, Hawaii, and of two records with missing data from Hanasaki, Japan, and Crescent City, California. Further, Fig. 3b also shows a case of a high-quality time series from Acapulco, Mexico, where tsunami oscillations are completely recorded. Figure 3c illustrates how data gaps occurring around the beginning of the tsunami signal may affect the determination of the TAT and of the parameters of the first tsunami wave. This is shown by means of three records from California stations (Arena Cove, South Beach and Fort Point, the latter two from San Francisco). Further, by displaying records from stations in Pacific islands (Kwajalein atoll; Wake island; Kawaihae, Hawaii) we also show in Fig. 3a–3c that the first wave can be disturbed by

Table 2. Continued.

	Tsunami signal characteristics						Secure detection	Tsunami detection	
Station code	TTT (h)	MA1 (cm)	R1 (cm)	MA(TDI) (cm)	R(TDI) (cm)	P1 (min)	DT (min)	DT (min)	NF
petr	3.03	5	15	5	15	47	missed	missed	0
preo	0.57	5	5	5	15	14*	missed	missed	1
prin	9.77	15	15	15	25	114	missed	missed	0
quepo_d1	17.67	15	25	45	75	38	55	30	0
quepo_d2	17.70	15	15	45	75	37	55	29	5
raro2	11.00	15*	25*	45*	85*	40*	missed	32	0
rbct	11.67	15	15	25	45	24	102	18	0
rfrt	11.60	15	25	35	65	25	103	9	0
riki_d1	14.15	25	45	35	55	39	33	11	0
riki_d2	14.13	25	35	35	55	41	35	13	0
rudn	0.63	5	5	15	35	15*	missed	missed	0
saip	3.37	75	135	75	135	29	7	2	0
sanf	20.38	15	15	45	75	49	97	missed	0
sano_d1	22.45	35	45	85*	165*	32	31	73	1
sano_d2	22.40	25	35	85*	165*	35	35	2	0
sant_d1	17.92	75	125	205	355	33	7	1	0
sant_d2	17.92	75	115	235*	375*	33	8	2	0
sbea	10.02	25*	45*	55*	95*	16*	11	2	0
sdpt	6.67	25	45	35	65	33	26	8	0
sewa	8.22	15	15	15*	15*	62	missed	missed	0
sitk	8.27	15	35	25*	35*	79	106	11	0
solo	7.97	25	45	25	55	67	53	46	0
talc_d1	22.88	155	255	205*	365*	112	24	18	1
talc_d2	22.88	155	245	205*	355*	112	24	7	0
tara	6.40	25	45	25	55	35	26	7	0
tofi	9.62	45	95	65	115	27	11	5	0
tosa	2.12	95*	145*	95*	165*	36	8	2	2
unal	5.95	45	55	45	85	62	19	10	0
valp_d1	22.32	35	55	145*	275*	42	9	2	0
valp_d2	22.32	35	65	145*	265*	42	10	3	0
vanu	8.95	55	75	85*	145*	53	49	37	0
viti	9.88	15	35	25	45	44	36	10	0
wake	3.57	55	75	55	75	28	19	0	0
wint	9.10	55	95	55	95	42	11	5	0
wpwa	10.07	45	55	45	75	36	20	0	0
xmas	9.03	55	105	55	115	34	10	4	0
yaku	8.40	5	15	25	45	59	missed	27	0
yapi	4.08	15	35	15	35	78	missed	missed	0

shorter period oscillations, so that it can exhibit more than one local maximum and local minimum.

The plots of the period P1 and of the range R1 of the first tsunami wave vs. the distance of the station from the source are portrayed in Fig. 4. It seems that there is no specific trend in both graphs with a broad dispersion of periods and magnitudes for similar distances. This deserves some comments. As regards periods, the lack of dependence from the distance is not surprising, since it is expected that it is determined mostly by the source geometry and by the local conditions at the recording site. However, one could argue that theoretically wave dispersion may favour the appearance of long period oscillations at the tsunami front at large distances, thus causing the period of the first wave to increase. Since this is not clear from the graph (Fig. 4 left), one may take this as an evidence that dispersion was not very important for the Tohoku tsunami even for cross-oceanic propagation. As regards the wave range, a decrease with distance is expected on physical grounds since wave fronts attenuate as the effect of geometrical spreading. The computed sea surface elevation maxima, mapped in Fig. 1a, show that tsunami amplitudes are very large on the Japanese coasts that face the source (and that indeed experienced the disastrous impact) and decay strongly with distance. However, our plot shows that, in spite

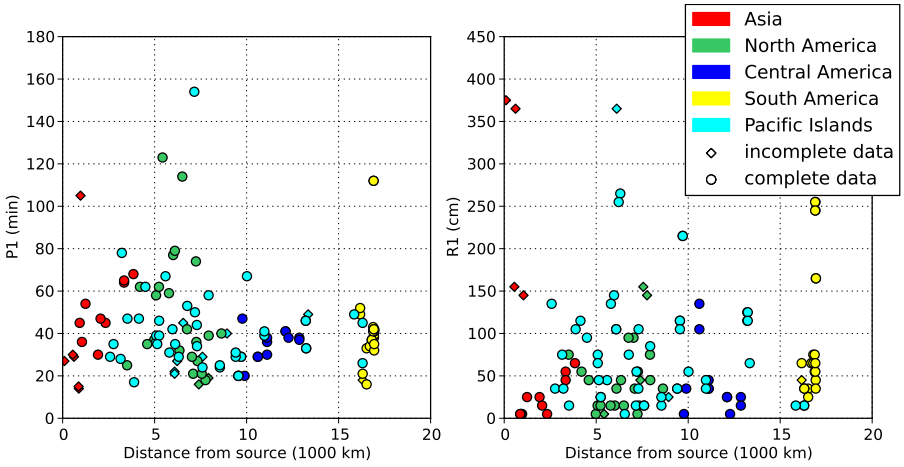

Fig. 4. Estimated period P1 (left) and range R1 (right) of the first tsunami wave plotted vs. distance from the source. Colors refer to continental regions, as in Fig. 1. Estimates from incomplete data (because of gaps or saturation in the records) are indicated with diamonds, while circles are used for all the other ones.

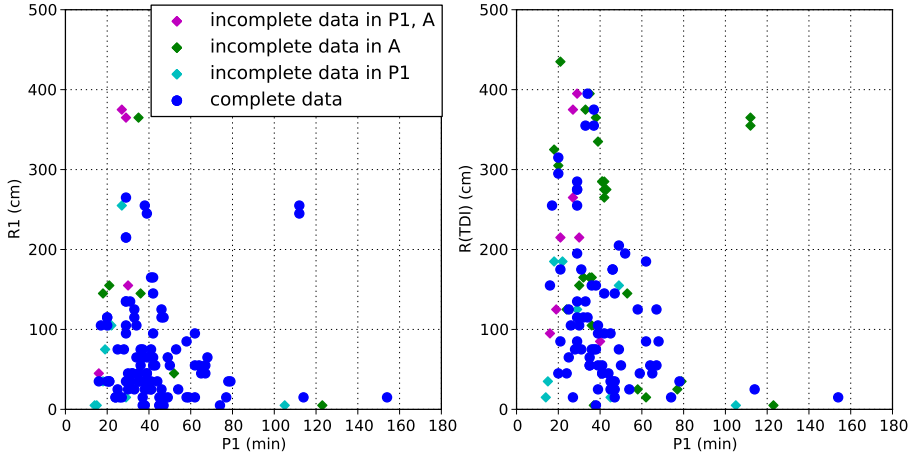

Fig. 5. Plots of the ranges R1 (left) and R(TDI) (right) in function of the period P1. In each graph diamonds are values that are uncertain in one or both of the plotted variables (owing to missing samples or record saturation), while circles are good estimates.

of this attenuation, the recorded range (which is a sort of peak-to-peak amplitude of the first wave) is quite large even for some coastal stations far from the source. We remark that our dataset includes only one Japanese station from the very near-field, that is Ofunato, where the gauge, as already mentioned, could not record the complete first wave since it broke during the attack of the first crest (Fig. 3b). Therefore the estimated range R1 = 375 cm (see Table 2) is expectedly much lower and not representative of the one of the real wave. The next Japanese stations in the database, Hanasaki (R1 = 365 cm, see also Fig. 3b) and Omaezaki (R1 = 155 cm), are located some hundreds of km away to the north and to the south, outside of the direct attack of the tsunami. Taking all this into account, and excluding from our considerations near-field stations, i.e. stations located less than 2000 km from the source that are not well covered in our database, we conclude that the graph of R1 suggests that the observed range of the first wave is more influenced by local amplification conditions than by the distance from the source, at least for costal tide-gauges. However, some observed large values of R1 may find interestingly an explanation in the propagation pattern of the tsunami that is characterised by a number of fingers or beams of higher energy determined by ocean bathymetry (see Fig. 1a). For example, Hilo (R1 = 265 cm), Kahului (R1 = 365 cm), and Kawaihae (R1 = 255 cm), in the Hawaiian islands seem to be located on the path of one of such beams. But there are also counterexamples of nearby stations (e.g. Kaumalapau, R1 = 105 cm), presumably on the

trajectory of the same beam, where the range is much lower. In the very far field, large values of R1 (around 250 cm) are found in the Chilean stations of Constitucion and of Talcahuano (see records plotted in Fig. 7).

To characterise the tsunami signal also beyond the first wave, we determined the maximum (positive) and minimum (negative) level reached by the tsunami within the 3h-long TDI, respectively denoted as M(TDI) and m(TDI), from which we further derived the maximum tsunami amplitude MA(TDI) = max(M(TDI),-m(TDI)) and the tsunami range R(TDI) = M(TDI)-m(TDI). Also these parameters have been estimated by removing a rough estimation of the tide from sea-level and they can be found in Table 2 for all the stations. Figure 5 presents the distribution of the first wave range R1 and of the tsunami range R(TDI) vs. the tsunami period P1. The ranges are spread over the interval from 0–450 cm and periods over the interval from 16–154 min. For both distributions one finds that the envelopes are curves similar to unimodal distributions, with right tail longer than the left one. The largest ranges (>250 cm) correspond to periods between 30–40 min, while ranges are quite low for periods exceeding 80 min. Exceptions are the two records from Talcahuano (see Fig. 7), Chile, where large-amplitude long-period tsunami waves were observed. As expected, since for any given record $R(\mathrm{TDI}) \geq R1$ by definition, the envelope on the right plot is higher than the one in the left plot, which also means that oscillations with the highest ranges occur often after the first wave.

3 The method: how TEDA works

TEDA is a Tsunami Early Detection Algorithm composed by two parallel detection modules devised to detect anomalous sea-level conditions: one, called "tsunami detection module", is designed to detect long-wave trains with impulsive first arrivals (typically tsunamis), while the other, called "secure detection module", is designed to identify high-amplitude long waves, including seiches and also tsunamis, that can be dangerous to people and property. The former module uses a slope-based algorithm, whereas the latter may be considered as amplitude-based, though it is built on signal slope. TEDA is designed to work at station level in real time, i.e. it makes use of functions that are updated at every new sample acquisition.

A detailed description of TEDA can be found in the paper by Bressan and Tinti (2011), where TEDA was tested and calibrated on several-year records from the US tide-gauge station of Adak Island in Alaska. Here it is enough to summarize its main features. A summary of TEDA functions acronyms and of the tsunami parameters used in this analysis is given in Table 3. The main TEDA function is the Instantaneous Signal IS(t), which is an estimation of the detided sea-level slope, computed at every time step over a time interval of length T_{IS} that includes the actual time t, i.e. $[t - T_{\mathrm{IS}}, t]$. Detiding is obtained i) first by computing the signal slope over the interval $[t - T_{\mathrm{IS}}, t]$, ii) then by computing the tide slope, which is the average of the signal slope over an interval of length T_{TIDE} ($T_{\mathrm{TIDE}} >> T_{\mathrm{IS}}$) defined as $[t - T_{\mathrm{TIDE}} - T_{\mathrm{GTIDE}}, t - T_{\mathrm{GTIDE}}]$ and iii) eventually by subtracting the tide slope from the signal slope. Notice that the calculation of the tide slope is executed on an interval that ends at the anticipated time $t - T_{\mathrm{GTIDE}}$, where the gap T_{GTIDE} should ensure that the tide evaluation is not contaminated by the anomaly to be detected. In case of missing data, TEDA suspends the update of its functions till a new sea-level datum becomes available. Then the gap is filled by a linear interpolation and the functions are calculated and updated up to the actual time t.

The TEDA "tsunami detection module" is designed for long-period wave trains that travel with a sharp leading front, i.e. for all cases when the arrival of the first wave of the train introduces an instantaneous difference from the previous background situation. It is based on the comparison of IS(t) with the Background Slope BS(t) through the control function CF(t), defined as the ratio between IS(t) and BS(t), i.e. $\mathrm{CF}(t) = |\mathrm{IS}(t)|/\mathrm{BS}(t)$. The function BS(t) is computed by taking into account the values of IS from the time interval $I_{\mathrm{BS}} = [t - T_{\mathrm{BS}} - T_G, t - T_G]$ of length T_{BS} and anticipated by the time gap T_G with respect to the actual time t. Likewise in detiding, the gap T_G serves to avoid the influence of the anomaly on the function BS(t). TEDA considers three different options to calculate BS(t), but in this work only one was used, namely the one that was found to be the most efficient for the Adak station data and that corresponds to $\mathrm{BS}(t) = \max(|\mathrm{IS}(t')|)$, with t' in I_{BS} (this option is named A3 in Bressan and Tinti's paper, 2011). TEDA triggers a "tsunami detection" if $|\mathrm{IS}(t)| \geq \lambda_{\mathrm{IS}}$ and if $\mathrm{CF}(t) \geq \lambda_{\mathrm{CF}}$ where λ_{IS} and λ_{CF} are two threshold parameters. Considering that the second inequality is equivalent to $|\mathrm{IS}(t)| \geq \lambda_{\mathrm{CF}}\mathrm{BS}(t)$, the detection condition can be expressed by $|\mathrm{IS}(t)| \geq \max(\lambda_{\mathrm{IS}}, \lambda_{\mathrm{CF}}\mathrm{BS}(t))$. If a detection occurs at time t_D, TEDA raises a tsunami state, during which detections are suspended until BS(t) returns below the value it had at the detection time, i.e. until when $\mathrm{BS}(t) \leq \mathrm{BS}(t_D)$.

The TEDA "secure detection module" is based on the function M(t), that represents a "detided filtered marigram" and is computed through a partial integration of the function IS over the interval $[t - T_{\mathrm{SD}}, t]$ of length T_{SD}. TEDA "secure detection" alerts are triggered when the function $|\mathrm{M}(t)|$ passes the threshold λ_{SD}, i.e. if $|\mathrm{M}(t)| \geq \lambda_{\mathrm{SD}}$, and it lasts for a pre-defined time interval of duration T_A after a detection. The alert interval may happen to be longer than T_A if other alerts are triggered in the meantime. For both modules, in case that the detection conditions are met during the interpolation of a gap, the detection is attributed to the end of the gap, when the first datum is available.

In order to evaluate the performance of a detection algorithm, an objective evaluation procedure has to be set up, that has to be applied of course off line, that is on records after

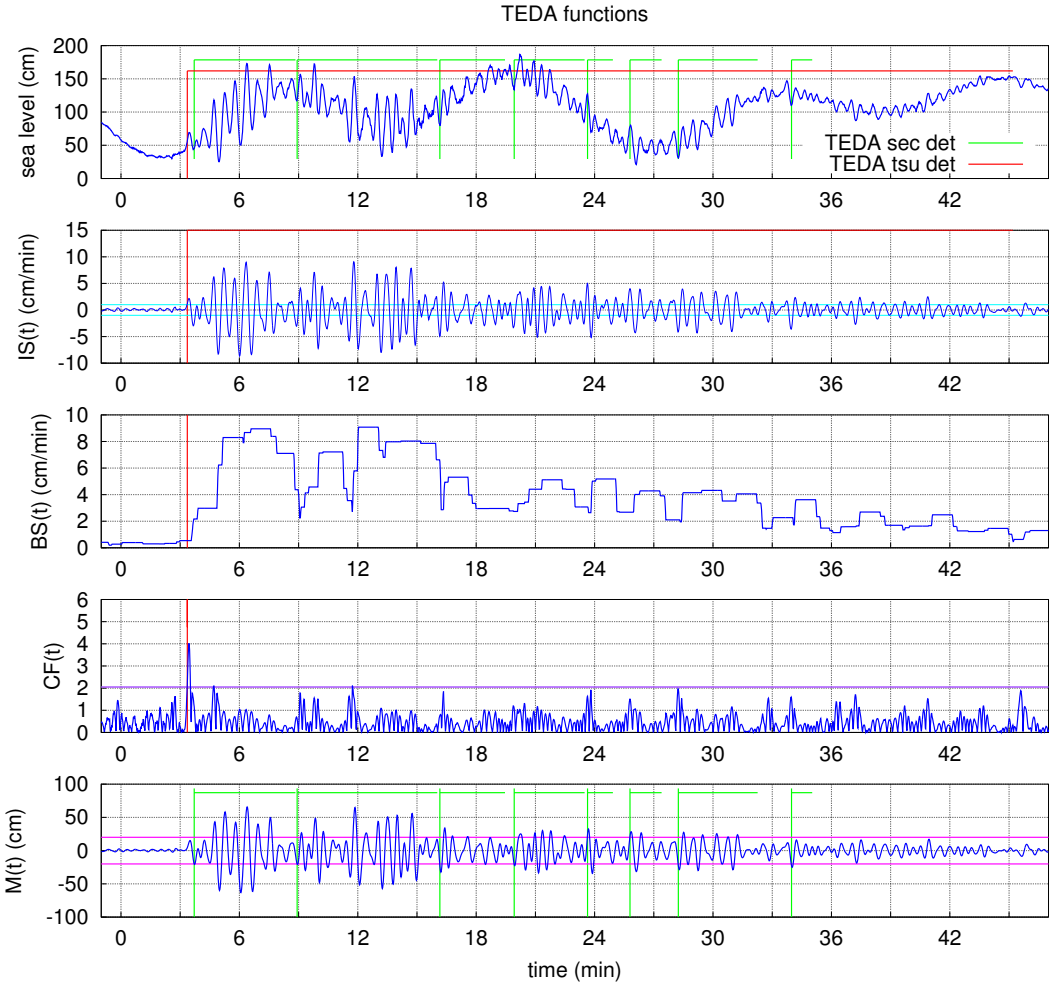

Fig. 6. TEDA functions IS(t), BS(t), CF(t) and M(t) and alerts computed for the sea-level record of Naha, Japan. Original record in the top panel. In the graphs of IS(t), CF(t) and M(t) the respective thresholds λ_{IS}, λ_{CF} and λ_{SD} are plotted as horizontal lines (fuchsia). Vertical lines mark the detection times of tsunami detection (red) and of secure detection (green) methods. Horizontal lines mark the respective tsunami state (red) and alert (green) duration. Notice that tsunami state covers all the tsunami signals until it dies out, while the alert interval has several interruptions. In all plots the time is measured from the origin time of the earthquake.

the end of the tsunami or, more generally, of the anomalous waves. As regards TEDA, the evaluation procedure is based on the computation of a number of performance indicators (Bressan and Tinti, 2011). To this purpose, the first stage is to identify the beginning and the end of the tsunami signal in each tide-gauge record, which corresponds to the determination of the tsunami arrival time TAT and of the Tsunami Interval, already defined in Sect. 2. To notice that as many as 98 records ($\sim 80\,\%$) terminate before the end of the tsunami, and therefore the TI is necessarily less than it should be, and in particular for the records of Tern, Fr. Frigate (fren_d2), of Ofunato, Japan (ofun) and of Sitka, USA (sitk), the records end before the end of the 3-h long TDI. The main TEDA indicator is the Delay Time DT that is the delay of detection with respect to the TAT. In the tsunami detection module a detection is considered a success if it occurs within the TDI, i.e. within the 3-h long time window starting at the TAT. If there is no detection or if it occurs later, then this is a failure for the algorithm. A second performance indicator is the number of false detections NF, that is of detections that may occur outside the identified tsunami interval TI. These indicators do not apply to the secure detection module, that is devised for detection of large-amplitude long-period waves. These can already affect a station in the moment when tsunami arrives and hence secure detection alerts that fall outside the TI, before or after, cannot be considered false detections. What matters in case of a large-amplitude tsunami is that the alert is "on", no matters if it is "on" due to a trigger coming from the tsunami itself or from previous dangerous long- period waves. Therefore, we consider that the tsunami is detected

Table 3. Acronyms and notations.

Acronym	Definition
	TEDA definitions:
BS(t)	Background Signal: maximum detided sea-level slope, BS(t)= max(IS(t')), t' in I_{BS}.
CF(t)	Control Function: CF(t)=IS(t)/BS(t).
DT	Delay Time of TEDA detection with respect to TAT.
I_{BS}	the time interval $I_{BS} = [t - T_{BS} - T_G, t - T_G]$, used to compute BS(t).
IS(t)	Instantaneous Signal: detided sea level slope in the time interval $[t - T_{IS}, t]$.
M(t)	secure detection function: detided filtered sea level.
NF	number of false detections by TEDA tsunami detection module.
t	actual time.
T_A	pre-defined length of the secure detection alert.
T_{BS}	length of the time interval used to compute BS(t).
t_D	time of TEDA detection.
T_G	length of the time gap used to avoid the influence of the anomaly to be detected in BS(t).
T_{GTIDE}	length of the time gap used to avoid the influence of the anomaly for the tide correction of IS(t).
T_{IS}	length of the time interval used to compute IS(t).
T_{SD}	length of the time interval used to compute the secure detection function M(t).
T_{TIDE}	length of the time interval used to compute the tide correction of IS(t).
TDI	Tsunami Detection Interval: 3-h interval after TAT considered valid for detection.
TI	Tsunami Interval : interval from TAT to the end of the tsunami oscillations/record.
λ_{IS}	threshold (together with λ_{CF}) for the detection condition of the tsunami detection module: $\lvert IS(t)\rvert \geq \lambda_{IS}$ and $CF(t) \geq \lambda_{CF}$, or equivalently $\lvert IS(t)\rvert \geq \max(\lambda_{IS}, \lambda_{CF} BS(t))$.
λ_{SD}	threshold for the detection condition of the secure detection module: $M(t) \geq \lambda_{SD}$.
λ_{CF}	threshold (together with λ_{IS}) for the detection condition of the tsunami detection module: $\lvert IS(t)\rvert \geq \lambda_{IS}$ and $CF(t) \geq \lambda_{CF}$, or equivalently $\lvert IS(t)\rvert \geq \max(\lambda_{IS}, \lambda_{CF} BS(t))$.
	Tsunami characteristics extracted from the records:
M(TDI)	Estimated maximum tsunami amplitude in TDI (without tide).
m(TDI)	Estimated minimum tsunami amplitude in TDI (without tide).
MA(TDI)	MA(TDI)= max(M(TDI), −m(TDI)), maximum amplitude (positive or negative) of the tsunami in TDI.
MA1	MA1= max(M1, −m1), maximum amplitude (positive or negative) of the first tsunami wave.
M1	Estimated maximum of the first tsunami wave (without tide).
m1	Estimated minimum of the first tsunami wave (without tide).
P1	Period of the first tsunami wave.
R1	R1= M1 − m1, range of the first tsunami wave.
R(TDI)	R(TDI)= M(TDI) − m(TDI), range of the tsunami in TDI.
TAT	Tsunami Arrival Time.
TTT	Tsunami Travel Time (from the source to the tide-gauge).

by the secure detection module if the detection occurs within the tsunami detection interval TDI and in addition if it occurs before the tsunami arrival, provided, however, that the associated alert state covers part of the TDI. Bearing this in mind, it is convenient to introduce a Delay Time DT even for the secure detection module. It was taken to be comprised between 0 and 3 h, i.e. to be less than the assumed TDI, to favor comparison between the performance of the two TEDA modules. To this purpose, when the detection occurred before TAT, implying a negative DT, the DT was forced to be zero.

4 The results of TEDA analysis

In this work, TEDA was run on 123 records of the 2011 Tohoku tsunami collected in coastal Pacific stations, and the results will be given and evaluated in the next subsections. TEDA was applied by using the best parameter configuration that was found suitable for the Adak Island, USA, station (Bressan and Tinti, 2011), i.e. with the following parameters $T_{IS} = 12$ min, $T_{BS} = 60$ min, $T_G = 16$ min, $T_{SD} = 8$ min, $T_A = 60$ min, and thresholds $\lambda_{IS} = 1.0$ cm/min, $\lambda_{CF} = 2.05$, $\lambda_{SD} = 20$ cm, that will be collectively referred to as parameters configuration C4. We have also carried out a sensitivity analysis by varying the thresholds for the tsunami detection conditions λ_{IS} and λ_{CF}, according to the values given in Table 4, for a total of 7 configurations of parameters. The results

of the analysis carried out with TEDA are illustrated in this section by first making an account of the outcomes of the secure detection method (4.1) and then of the tsunami detection method (4.2) applied with configuration C4. Results of tsunami detection using the other configurations are also commented in Sect. 4.4.

In Fig. 6 we give an example of the TEDA functions and alerts that result from applying TEDA to the record of Naha, Okinawa, Japan and in Fig. 7 we show the results of TEDA detections on a number of interesting records, on which we will make observations later in this section. We only stress now that for the stations of Acajutla, El Salvador, Constitucion and Talcahuano, Chile, two records are plotted in the graph. The record pairs of the same station are very similar to each other and are distinguishable, though hardly, only for Constitucion (const) in the scale of the figure. Interestingly, detection results are different. This will be discussed in the last Sect. 4.5 where our focus will be concentrated on the TEDA detections for the stations with records coming from more than one sensor (see Table 1). The horizontal lines in the figure mark the time interval where the alert is on for the tsunami detection module (green) and for the secure detection module (orange). Here it is only worth observing that the tsunami was seen within the first hour i) by the tsunami detection method in all records with the exception of channel 2 of Acajutla and of Constitucion (first two plots from the top), missed, and ii) by the secure detection method in all records but the one from Avatiu Rarotonga, Cook Islands (raro2), missed, and the channel 2 of Acajutla, detected after 1 h. Special consideration deserves the two records from Constitucion. In channel 1 the tsunami state (green horizontal line) is "on" well before the tsunami arrival and covers all the tsunami record including the whole tsunami detection interval TDI. This however has to be considered a false tsunami alarm for the tsunami detection module since detection does not occur inside the TI, and the case is a failure. It can be observed that even the secure detection module starts the alert before the TAT (for both channels), and continues the alert state up to covering the TDI, but in this case the performance is a success.

4.1 Detections made by using the TEDA secure detection method

The TEDA secure detection module manages to detect 99 tsunamis out of the 123 tide-gauge records in the 3-h long TDI interval, while the number reduces to 87 if we consider only detections within the first hour. The function M(t), which is computed and compared in absolute value to the threshold $\lambda_{SD} = 20$ cm to search for a detection, has the dimension of a length since it is obtained by integrating the detided Instantaneous Signal IS over a time interval of finite length $T_{SD} = 8$ min. It represents therefore the sea-level excursion during such an interval of time. To help understand better the meaning of M(t), it can be observed that if we assume a sinusoidal tsunami with period T and amplitude A and assume further that the instantaneous slope coincides with the analytical slope (i.e. $T_{IS} \ll 1$min), then it is straightforward to show that $M(t) = A[\cos 2\pi(t + T_{SD}) - \cos 2\pi t]$ and that $\max(|M(t)|) = 2A \sin(\pi T_{SD}/T)$. From this we deduce for instance that i) if the period T equals $2T_{SD}$ the peak of the function $|M(t)|$ equals the tsunami peak-to-peak height $2A$; ii) if the tsunami has period equaling $4T_{SD}$, then the maximum of $|M(t)|$ equals $\sqrt{2}A$ and if $T = 6T_{SD}$, then the maximum of $|M(t)|$ equals the tsunami amplitude A; iii) for longer-period tsunamis the $|M(t)|$ peak value is less than the tsunami amplitude (Bressan and Tinti, 2011).

All TEDA secure detection results are listed in Table 2. As a rule of thumb, one could expect that tsunamis with amplitude close or below the threshold λ_{SD} are typically missed and viceversa; tsunamis with amplitude much larger than the threshold are ordinarily detected and exceptionally missed. Indeed, of the 21 cases of tsunamis with MA(TDI)$< \lambda_{SD} =$ 20 cm, as many as 20 are missed and only 1 is detected. This occurs accidentally for the record of La Paz, Mexico, with MA(TDI) = 15 cm, for which the function $|M(t)|$ just passes the threshold. However, this is a late detection and none of them were detected within the first hour of TDI. On the other hand, of the 102 cases with MA(TDI)$\geq \lambda_{SD} = 20$ cm, as many as 98 were detected: 87 within the first hour of TDI and 11 later.

The missed and the late detections are worth considering more closely. The missed records of the Johnston atoll, USA (john_d1), of Cabo San Lucas, Mexico (cabo), and of Yakutat, Alaska (yaku) have MA(TDI)= 25 cm and their respective functions $|M(t)|$ get very close to although they do do not surpass the threshold. They present different tsunami characteristics, from long period first wave (as Yakutat, P1 = 59 min) to very short (12–13 min for Cabo San Lucas), and different amplitudes. One can observe that of the 12 records with MA(TDI)=25 cm, only these three do not trigger a detection, while the other five i.e. the two records of Acajutla, El Salvador, and the records of Ishigakijima, Japan and of Raoul Island (Boat Cove), New Zealand and the record of Sitka, USA (sitk)) present late detections. It is more interesting to understand why the tsunami in the record of Avatiu Rarotonga, Cook Islands (raro2) passed undetected though MA(TDI) = 45 cm and is well beyond the threshold. The explanation lies in the predominant period of the highest amplitude oscillations in this record (see plot in Fig. 7) that is quite short, about 6 min, and therefore these oscillations are filtered out in computing the function M(t). Taking into account the late detections, it appers that the remaining late detected records, i.e. of Chatham Island (chit), of Raoul Island (Boat Cove, rbct), New Zealand, the second record of Juan Fernandez, Chile (juan_d2), the records of Pago Pago, Samoa (pago and pagx) and of San Felix, Chile (sanf) are characterized by larger R(TDI), but are mostly detected late because the tsunami is characterized by low amplitude waves at the beginning of the signal. It is indeed important to stress

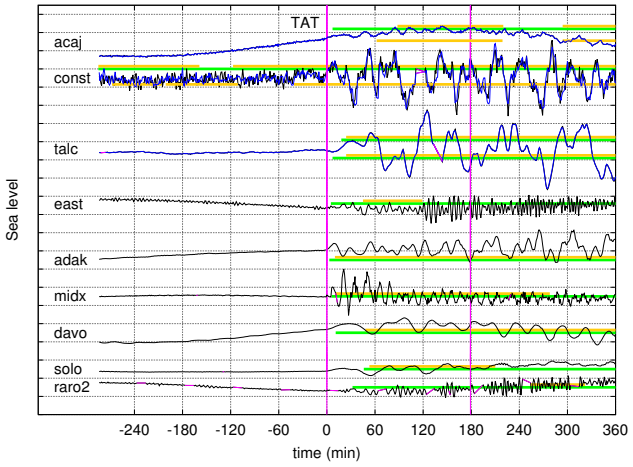

Fig. 7. Examples of tsunami records with TEDA detections. Stations are: acaj = Acajutla, El Salvador; const = Constitucion, Chile (const_d1, black; const_d2, blue); talc = Talcahuano, Chile; east = Easter Island, Chile; adak = Adak, Alaska; midx = Midway Islands, Hawaii; davo = Davao, Philippines; solo = Honiara, Solomon islands; raro2 = Avatiu Rarotonga, Cook Islands. Notice that Acajutla and Talchuano have two records by different sensors (Table 1) and that they do overlap perfectly in this graph. Horizontal lines mark the tsunami state from the tsunami detection method (green), and the alert interval from the secure detection analysis (orange). For the three stations with two records, there are two pairs of horizontal lines: the upper pair for channel 1 and the lower pair for channel 2. Vertical red lines mark the beginning and end of the 3h-long TDI. On the vertical axis, ticks are at 1 m distance.

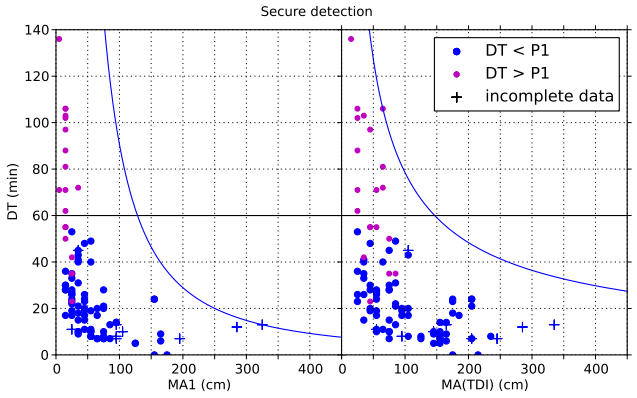

Fig. 8. Delay time DT of the secure-detection alerts vs. the maximum amplitude of the first tsunami wave MA1 (left) and of the tsunami within the TDI window MA(TDI) (right). Complete and incomplete records (affected by data gaps or magnitude saturation and marked with an asterisk in Table 1) are depicted distinctly. A gross estimate of the parameters envelope results in the following values $a = 1.79\ 10^5$, $b = 1.65$ for the graph of DT vs. MA1 and $a = 1.96\ 10^3$, $b = 0.70$ for the graph of DT vs. MA(TDI) that are displayed in Fig. 9, where DT is measured in minutes and MA1 and MA(TDI) in cm.

that the secure detections might occur only when the tsunami signal reaches a certain wave amplitude, which might happen quite late in the case of tsunamis that arrive with slowly increasing waves.

In general, one can expect that the larger the tsunami amplitude, the shorter the time for detection is. This latter observation is confirmed by Fig. 8 where the delay times of all 99 detected cases are plotted against the maximum amplitude of the first tsunami wave and of the maximum amplitude occurring within the TDI. Distributions of both graphs are quite similar and show that DT is inversely correlated to the maximum signal amplitude: expectedly, the larger is the amplitude, the shorter is the delay time, DT. It is seen that, although data spread is relevant, the envelope is well represented by a curve of the type $y = ax^{-b}$, showing an inverse correlation between the delay time and the maximum amplitude of the tsunami oscillations (see Fig. 8). From the figure, it is also possible to see that all the tsunamis detected later than 1 h were not detected in the first but in the following waves. They present very low first wave range (MA1$\leq$ 15 cm, with the exception of Chatham Island, New Zealand, with MA1 = 45 cm and P1 = 64 min) and their amplitude grows further within TDI.

It is also interesting to discuss the case of the two records of Constitucion, const_d1 and const_d2. The records are displayed in Fig. 7 and it is clear that the tsunami was a long train of very high-amplitude long-period waves. From Table 2 we get that MA(TDI)>150 cm. The records are very noisy with a long series of long-period oscillations in the background with amplitude in the range of 50–100 cm that triggered the secure detection alerts several times before the tsunami arrival. In this case, the algorithm works since it recognizes the presence of high-amplitude waves. Since the last alert before the TAT remains "on" even after the arrival of the tsunami. For these cases the delay time was considered equal to zero.

4.2 Detections made by using the TEDA tsunami detection method

The results of the TEDA tsunami detection method by using the parameter setting C4 of Table 4 are described in this section. There are 95 (77 %) records over 123 where tsunami is detected, while in 28 (23 %) cases it is missed. Only two records are detected with DT $\geq$60 min, i.e. the first record of San Antonio, Chile (sano_d1), with DT=60 min, and the second record of Juan Fernandez, Chile (juan_d2), with DT=73 min. Figure 9 displays plots of the delay times DT as a function of the maximum amplitude MA1 and range R1 of the first tsunami wave (upper graphs) and as a function of the maximum tsunami amplitude and range within TDI, i.e. MA(TDI) and R(TDI) (lower graphs). As for the secure

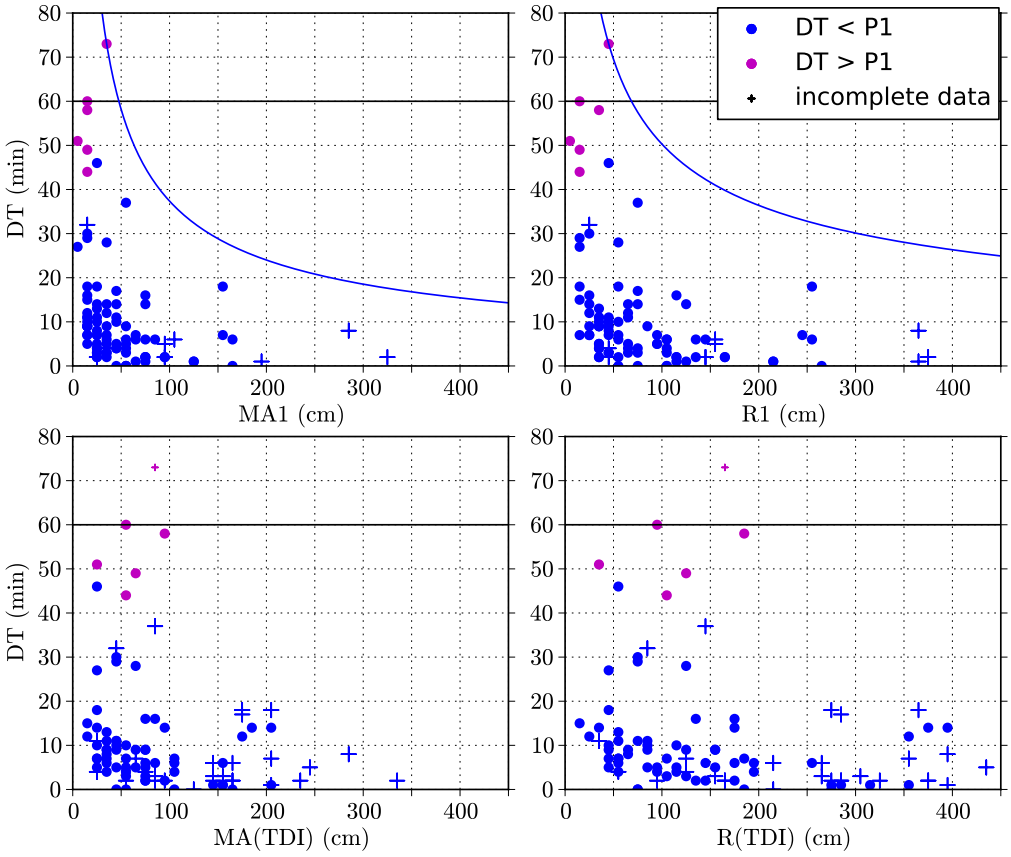

Fig. 9. Delay time DT for the tsunami detection method vs. the first wave maximum amplitude MA1 (top left) and range R1 (top right), the maximum peak amplitude (bottom left) and the range (bottom right) in TDI. Circles refer to complete records, while crosses mark records with data gaps or with saturated samples. Detections occurring on the first wave are blue, those occurring later are pink. The total number of detections is 95. For the plot of DT vs. MA1, a possible envelope is determined by the parameter values a = 7.05×10^2 and b = 0.64, while for the plot of DT vs. R1, the envelope parameters are a = 4.31×10^2 and b = 0.47.

detection case, it is seen that DTs are inversely correlated with the size of the tsunami oscillations, whatever the parameter used to quantify the size. Envelopes of the type $y = ax^{-b}$ can be estimated also for such graphs. All DT determinations may be found in Table 2. It is relevant to observe that the largest DT is 73 min, much less than the largest values found for the secure detection cases exceeding 130 min. Detections occur mostly on the first wave (blue in the graphs of Fig. 9) and there are only 6 detection cases (pink) where detections occur on the following waves. Looking at Table 2 where DTs and missed detections are displayed together with the observed tsunami size, one sees that in general, missed detections are found on the low-magnitude tsunami side, but there are also cases of missed tsunamis even when the size is quite large.

We discuss first the possible reasons for the 19 missed detections for the records of low range tsunami, i.e. with R(TDI)$\leq$ 35 cm. The 4 tsunami records with R(TDI) = 5 cm could not be detected since the signal was too small. We recall that the tsunami detection method is strictly based on signal slope estimates that are computed over intervals of duration $T_{IS} = 12$ min. A rough estimate of the function IS(t) can be made by considering a sinusoidal tsunami of the type $A \sin \omega t$. In this case one finds that IS(t)= $A(\sin \omega t - \sin \omega (t - T_{IS}))/T_{IS}$ and that $|IS(t)|$ can be maximized by $2A/T_{IS}$, i.e. by the ratio of the tsunami range $2A$ and the length T_{IS} of the computation interval for the instantaneous slope IS. If such a ratio happens to be lower than the selected threshold λ_{IS}, then the tsunami cannot be detected. This inequality contains the tsunami double amplitude 2A. By using instead the tsunami range R(TDI) as a proxy of 2A, and by applying this reasoning to our cases, we see that R(TDI)/$T_{IS} = 5/12$ cm min^{-1} < $\lambda_{IS} = 1$ cm min^{-1}, and that even taking into account the $\pm$5 cm uncertainty in range determination, one gets at most 10/12 cm min^{-1} < $\lambda_{IS} = 1$ cm min^{-1}. Another obvious way to maximize $|IS(t)|$ for a sinusoidal tsunami is $A\omega = 2A\pi/T$ where T is the wave period, which will be

approximated in our case with πR(TDI)/P1. Hence, by taking into account both the above maximizations, one gets the inequality $|\mathrm{IS}(t)| \leq 2A\min(1/T_{\mathrm{IS}}, \pi/T)$, i.e $|\mathrm{IS}(t)| \leq IS_{\max} = \mathrm{R(TDI)}\min(1/T_{\mathrm{IS}}, \pi/\mathrm{P1})$, where $IS_{\max}$ has been defined implicitly. This means that for long-period waves (more specifically $T > \pi T_{\mathrm{IS}}$) the maximum slope is more influenced by the wave period T rather than by the length of the computation interval T_{IS}. These considerations can help us when we take into account tsunami records with larger ranges. The records with R(TDI)$\geq$ 15 cm, in view only of the first inequality, could possibly be detected because 15/12 cm/min $> \lambda_{\mathrm{IS}}$. However, it is found that only two records of those with R(TDI) = 15 cm and R(TDI) = 25 cm give rise to a detection. The records are the ones from Astoria, Oregon (asto), with R(TDI) = 15 cm, and the second channel of Johnston Atoll (john_d2), with R(TDI) = 25 cm. In the case of Astoria, the tsunami period P1=27 min, and hence $IS_{\max} = 15\min(1/12, \pi/27) = 1.25\,\mathrm{cm\,min^{-1}}$, gives rise to a detection on the first wave with DT = 15 min. The record of Johnston Atoll starts with an impulsive wave, and in the group of records with R(TDI) = 25 cm is the one with the highest oscillations and the shortest period, i.e. P1 = 39 min. A reason for the missed detection can be explained on the basis of the second inequality, as in the case of the records of Cherry Point, Washington, and Nouméa-Numbo, French New Caledonia, with R(TDI) = 15 cm, where the very long period of the tsunami, (P1 = 74 and 154 min), gives πR(TDI)/P1 cm min^{-1} $< \lambda_{\mathrm{IS}}$. A further possible explanation for the missed tsunamis can be searched for in the fact that we used R(TDI) in the above inequalities to approximate the tsunami double amplitude 2A or, better said, the peak-to-peak height H resulting from consecutive oscillation extremes. These quantities are perfectly equal in a sinusoidal tsunami, but they can be quite different for very more complicated waveforms where the range can result from not consecutive peaks and is expected to be larger than H. This is the case for the Kanton island, Kiribati Republic, with a very low signal so that IS(TDI)$<$ λ_{IS}. In a similar way, for the case of Preobrazheniye, Russia, IS(TDI)$< \lambda_{\mathrm{IS}}$ because the signal is of low amplitude and quite noisy, so that the function IS acts as a filter for short period oscillations and never fulfills the first detection condition. Another more convincing explanation is that tsunami detection trigger is formed by two conditions, one based on the threshold λ_{IS} discussed so far, and one based on the threshold λ_{CF}, i.e. $|\mathrm{IS}(t)| \geq \lambda_{\mathrm{CF}}\mathrm{BS}(t)$. If the first appearance of tsunami is low-amplitude and therefore too small to trigger a detection, then the initial tsunami wave might be incorporated in the background signal, that is then increased and makes it harder to recognize the tsunami even in the subsequent oscillations. A similar explanation is valid for the records of Seward, Alaska, and of Petropavlovsk, Russia, where the long period of 62 and 47 min causes the first peak to be included in the BS function before the IS function could trigger a detection. All other records (namely Busan, Korea; Elfin Cove, Alaska; Kodiak island, Alaska; Korsakov, Russia; Malakal, Palau Islands and Prince Rupert, Canada) are characterized by very long periods and by a tsunami signal starting with low amplitudes, and are therefore missed. Of the 6 tsunami records with R(TDI) = 35 cm, half of them produce a detection. The missed record of Rudnaya Pristan is characterized by short-period ($\sim$ 6 min) waves which are too short to allow the function IS(t) to pass the threshold λ_{IS}. Instead, the first wave of the record of La Paz is too low to make IS(t)$> \lambda_{\mathrm{IS}}$ and the following slowly increasing waves do not allow the detection since the function BS(t) starts increasing together with the incoming tsunami.The same happens to the missed record of Yap Island, Micronesia, where the first maximum failed to force IS(t) to exceed the threshold λ_{IS}, and, hence, the first tsunami wave was incorporated in the background signal BS(t) making it too large. In general, the reason for a missed detection could be a combined effect of all the above explanations: the tsunami signal is not impulsive and low-amplitude.

By considering the tsunami signal characteristics MA1, R1, MA(TDI), and R(TDI), it seems in general that the tsunami detection method is not able to detect tsunamis with MA(TDI)$\leq$ 15 cm, or with MA1$\leq$ 15 cm, or with R1$\leq$ 15 cm or with R(TDI)$\leq$ 25 cm, while it detects some of the tsunami with R(TDI)$\leq$ 35 cm. In addition, it seems that TEDA tsunami detection efficiency is more affected by the first wave characteristics, i.e. by R1 and MA1, than by MA(TDI) or R(TDI), as it can be seen in Fig. 9. It can be noted that of the 19 cases of missed tsunamis mentioned so far, as many as 18 are equally missed by the secure detection method since they all are low amplitude with MA(TDI)$\leq$ 15 cm. The secure detection only detects the tsunami recorded in La Paz, Mexico, though with a large delay time (136 min) that happens to be the maximum DT occurring in our analysis (see Table 2).

There are however 9 records with missed detection that present large oscillations, with R(TDI)$\geq$ 45 cm. These are the two channels of Constitucion, Chile (const_d1, const_d2), the first channel of French Frigate, Tern Island, Hawaii (fren_d1), the first channel of Manzanillo, Mexico (manz_d1), the first channel of Nuku Hiva, Marquesas Islands (nuku_d1), the second channel of Acajutla, El Salvador (acaj_d2), San Felix, Chile (sanf), the first channel of Antofogasta, Chile (anto_d1), and the first channel of Arica, Chile (aric_d1). It is worth analyzing the reason for these missed detections. These records can be divided into two categories: records where there was an anticipated detection, i.e. a detection before the arrival of the tsunami (false alert) with a tsunami state covering also the tsunami interval, and records where no detection was really done. To the first category belong the first record of Constitucion (plotted in Fig. 7), and also the records fren_d1, manz_d1 and nuku_d1: for all of these records the detection conditions are accidentally met before the tsunami arrival due to intense sea agitation and the tsunami state remains "on" for the whole duration of the tsunami signal till the function BS(t) goes back to its

values at the detection time. We recall that during the tsunami state (i.e. as long as the alert is "on") the algorithm suspends checking detection conditions. However, a posteriori analysis shows that detection conditions occur for all these cases within the TDI. This consideration shows that for this category of cases, a different setting of the TEDA parameters (for instance, increasing the threshold λ_{IS}) could probably avoid the false detection masking the tsunami and could therefore favour a real detection to occur within the TDI. All the remnant 5 cases belong to the second category and they prove that high-amplitude long-period waves can pass undetected to the tsunami detection method of TEDA. The main reason, except for the second record of Constitucion (const_d2), is that the first arrival is not impulsive. Going more into details, we observe that the cases of acaj_d2 and sanf are low-noise records where the tsunami starts as a long wave of low amplitude. The records of anto_d1 and aric_d1 are characterized by a noisy background and by slowly increasing waves. Once the first tsunami wave is missed, it is quite unlikely that the tsunami can be detected by the tsunami detection algorithm since the tsunami signal enters in the background and hence the condition $|IS(t)| \geq \lambda_{CF}BS(t)$ is hard to fulfill. The record of const_d2 instead is quite noisy and is characterized by background waves quite similar to (but smaller than) tsunami waves that make the background increase and mask the tsunami arrival. It is worth pointing out that all these 9 cases of no-detection are indeed detected by the secure detection method, that must be seen as a complementary tool to the tsunami detection method. The corresponding delay times (see Table 2) are in increasing order: DT=0 min for const_d1 and const_d2; DT = 7 min for manz_d1 and nuku_d1; DT = 24 min for fren_d1; DT=35 min for anto_d1; DT = 43 min for aric_d1; DT = 62 min for acaj_d2 and DT = 97 min for sanf. It is interesting to notice that most of these records happen to come from sites with multiple sensors and that in the other records (i.e. fren_d2, manz_d2, nuku_d2 and nuku_d3, acaj_d1, anto_d2 and aric_d2) the tsunami was detected, but this will be discussed later in Sect. 4.5.

For the tsunami detection method, a very important performance indicator is the number of false detections. False tsunami detections undermine the usefulness and the efficiency of the detection algorithm and one of the main goals of the TEDA calibration process is to find the optimal parameters' configuration that minimizes (or eliminates) the occurrence of false detections (Bressan and Tinti, 2011). For this work we did not carried out any calibration on TEDA parameters, but we simply used the configuration C4 that was obtained by optimizing the performance for the Adak island, Alaska, station. The number of false detections, that is of detections occurring outside the tsunami interval TI, is reported in Table 2. In the dataset considered, only 15 records present false detections and most of them only 1. The fact however that false detections occur in records that are at most only a few days long is a sign that the TEDA setting used here is not adequate for those sites. We observe that if only results where no false detection occurs were considered acceptable, then the number of cases of tsunami detections (with no false detection) would diminish from 95 down to 87 with the present tsunami detection method (see Table 2). Some records give rise to false detections in the presence of normal background waves, and probably increasing the thresholds λ_{IS} and λ_{CF} could be sufficient to eliminate false detections and at the same time keeping unchanged (or increasing) the number of tsunami detections. In some other records (namely Nuku Aloga, Tonga Islands; La Paz, Mexico; the first channel of San Antonio, Chile; and the second channel of Quepos, Costa Rica), false detections are triggered by signals of difficult interpretation, of unclear origin or maybe due to erroneous data. Also in these cases, we believe that it could be sufficient to increase the thresholds λ_{IS} and λ_{CF} to avoid false alerts.

4.3 Comparison between the two TEDA detection methods

The results of the application of the secure detection method and of the tsunami detection method were illustrated in the previous subsections and some inter-comparison was already done. The first consideration regards the number of missed and detected cases. The cases missed by both methods are 18 and all were cases of small amplitude tsunamis. All the others, that are 105, were detected either by both methods (89), or only by the secure detection method (10), or only by the tsunami detection method (6). In general, large tsunamis are detected by both methods, since they arrive as an impulsive wave quite different from the previous background. When they arrive with a low-amplitude front as a series of increasing amplitude waves, they are undetected by the tsunami detection method, but captured by the secure detection method. Even small tsunamis can be detected sometimes: if the background is very low, the tsunami detection method is able to capture them.

Let us consider the delay times DTs. From Table 2 and from the few examples given in Fig. 7 we observe that the tsunami detection method identifies a tsunami before, and in many cases substantially before, the secure detection method, with a few exceptions. If we restrict our attention to the tsunami detection method, we see that most detections occur within the first 20 min of the TDI, with the number of detections tending to decrease with DT, and that the maximum DT was 73 min for the first channel of San Antonio, Chile (sano_d1). This means that there is a higher probability of detection at the very beginning of the tsunami signal, as expected from the algorithm design. A very positive result is that most detections occur at the very first wave, i.e. DT$<$P1, and within the first TDI hour.

The analysis of the application of the two methods shows that their performance is almost equivalent, although the secure detection method is slightly better with 99 vs. 95 (or 87

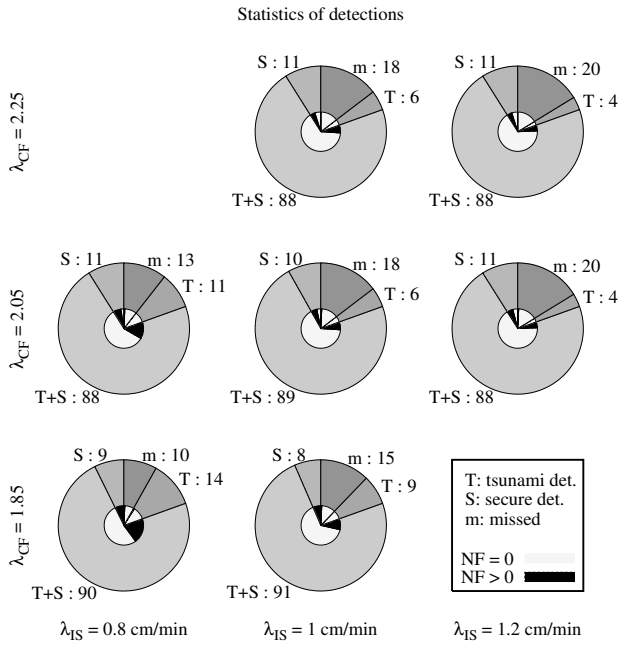

Fig. 10a. Pie charts with statistics of detections within 3-h long TDI are shown for all the 7 configurations of Table 4 ordered as in a matrix. For each pie chart the corresponding values of λ_{IS} and λ_{CF} can be read on the left and on the bottom. The following four categories are shown: (m) number of records with missed detection, (S + T) number of records with detections by both tsunami detection and secure detection methods, (S) number of records with detection only by the secure detection method, (T) number of record with detection only by the tsunami detection method. The inner part of each pie chart shows for each of the above categories if there are false detections (black) or not (white) in the records. The area of black sectors is proportional to the number of records with false detections.

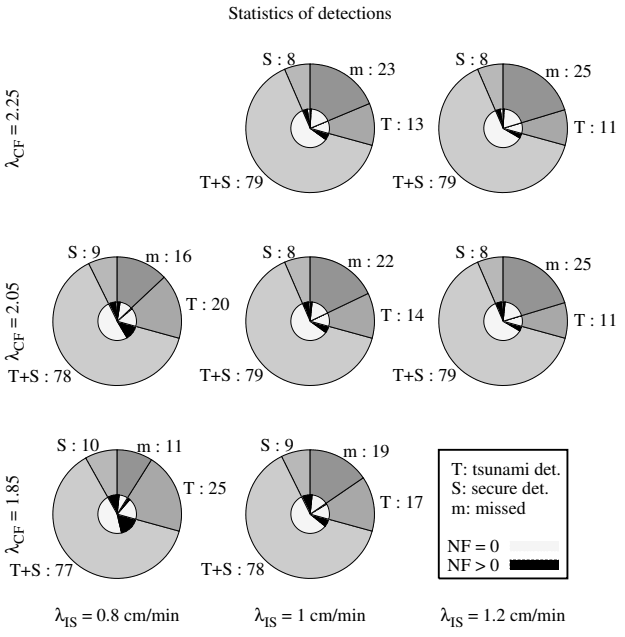

Fig. 10b. Pie charts with statistics of detections within 1-h long TDI. See also the caption of Fig. 10a. Notice how the main difference with a 3-h long TDI is the decrease of secure detection results with the consequent increase of missed records by both the secure and the tsunami modules.

if only cases with no false detection are accepted). By considering the first hour detections, the tsunami detection method becomes more effective or equivalent, by detecting 93 (or 86 if only cases with no false detection are accepted) records vs. 87 with DT < 60 min. But what is most relevant is that the joint application of the two TEDA methods results in a very efficient detection system. Combining the two methods means that a detection occurs if at least one of them makes a detection. When only one detects, the detecting method provides the delay time, while when both detect, the delay time will be the minimum of the two delay times. With this in mind, one sees easily from Table 2 that the tsunami is detected in as many as 105 records, of which 101 with DT < 60 min, and detection is of very good quality, since for most records the delay time of the combined application is very low (less than 20 min).

Table 4. Configurations of TEDA parameters and thresholds in this work. The fixed TEDA parameters are $T_{IS} = 12$ min, $T_{BS} = 60$ min, $T_G = 16$ min, $T_{SD} = 8$ min, $T_A = 60$ min, and the threshold $\lambda_{SD} = 20$ cm, while thresholds λ_{IS} and λ_{CF} vary. The first column gives the number used as identifier. Configuration C4 is the best found for the tide gauge station of Adak Island, USA.

C_n	λ_{IS}	λ_{CF}
1	0.8	1.85
2	0.8	2.05
3	1.0	1.85
4	1.0	2.05
5	1.0	2.25
6	1.2	2.05
7	1.2	2.25

4.4 Statistics of tsunami detection by varying the TEDA parameters configurations

All the previous analysis on the 123 records of the database was conducted by using TEDA in the configuration C4 already mentioned before. We conducted a sensitivity test, limited to the tsunami detection module, by slightly changing only the thresholds λ_{IS} and λ_{CF} according to the scheme of Table 4. The performance of the various configurations was

evaluated by means of the number of failures and detections in two cases: by considering TDI as previously defined as a 3-h long interval, and by reducing the length of TDI to the first hour after TAT. In Fig. 10a and 10b, results are shown with the aid of pie charts concerning the number of detections and non-detections as well as of false alerts. Following the way we used to analyze the results obtained by applying the configuration C4, the records can be divided in four categories: (S + T) records with detections by both methods (secure detection and tsunami detection), records with detection by only one method (the one (S) or the other (T)) and records where both methods fail (m). Although the sensitivity analysis does not regard the secure detection method, the frequencies in all the above categories may change from one configuration to the other since they are influenced by the results of the tsunami detection method. What cannot change in all pie charts is the total number of detections of the secure detection methods (99 for 3-h long TDI and 87 for 1-h TDI) equal to the sum of the records in categories (S+T) and (S), or equivalently the number of its failures (respectively 24 and 36), that is the sum of the records in categories (T) and (m). Looking at the pie charts, it seems that increasing the threshold λ_{CF} from 2.05 to 2.25 has little effect, while reducing it to 1.85 produces more changes. More substantial changes are found however when the threshold λ_{IS} is varied. In particular it is seen that the category of the tsunamis detected by both methods is rather stable since its frequency ranges from 88 to 91 (and from 77 to 79 for 1-h long TDI), while the number of tsunamis that none of the methods sees changes from 10 to 20 (and from 11 to 25) and, as expected, it goes along with the thresholds, i.e. it increases when both the thresholds are increased. On the other hand, decreasing the thresholds has the undesired effect of increasing the number of false alerts (represented in black in the inner part of the pie charts). It is worth observing that among the small tsunamis with R(TDI)$\leq$ 35 cm, the ones with R(TDI) = 5 cm are missed by all configurations and only in very few of them the tsunami is always detected (and with DT <60 min): these are the records of Astoria, Oregon, with R(TDI) = 15 cm, and the records of Ishigakijima, Japan, of Sitka, in Alaska, and of Johnston Atoll (channel 1), USA, all with R(TDI) = 35 cm.

The effect of changing the TEDA configurations for the 3-h long TDI can also be analyzed by means of Fig. 11 where a pie chart also including the indicator DT of the tsunami detection method is portrayed. Summarizing, the records can be partitioned into four groups. The first is the group of 15 records with no detection made by any of the 7 configurations. Then, there are 89 records where the tsunami was always detected. From these we extract the second group consisting of the 17 records where detections were triggered at exactly the same time DT, and the third group formed by the remaining 72 records where the detection time DTs were different from configuration to configuration. The fourth group is comprised of 19 records where, interestingly, changing configuration can transform a detection into a non-detection. In the inner part of the pie chart information on the number of false alerts is displayed for each of the above groups. Let us first denote the number of false alerts found for the configuration C_i by NF_i and let us further denote the maximum and the minimum value of NF_i by NF_M and NF_m respectively. Bearing this in mind, we can distinguish four categories: the one containing records with no false alerts ($NF_M = 0$, white in the graph); the one with records where all configurations see exactly the same number of false detections ($NF_M = NF_m > 0$, black in the pie chart); the one with records where all configurations see at least one false detection, but some configurations see more false detections than others ($NF_M \neq NF_m > 0$, grey in the figure); and eventually the one including those records where some configurations see false detections and some others do not ($NF_M \neq NF_m = 0$, light grey). It can be seen that for all the four groups, the frequency of these categories have the same rank, the most numerous being the one with no false detections (white), and the least numerous the one with all configurations having the same number of false detections (black).

For deeper insight into the results of the sensitivity test, it is further worth considering the statistics of DT shown in Fig. 12. In the three panels of Fig. 12, one sees the frequency distributions of DT for the various configurations, with the reference configuration C4 displayed in all graphs for comparison. Each graph includes also the frequency of the missed detections denoted as ND (standing for no-detections). It may be observed that most tsunamis are detected in the first 10 min and that only few cases are detected after 20 min. In general, it emerges that detection results seem more sensitive to varying the threshold λ_{IS} than the threshold λ_{CF}. Most importantly, it also results that changing configuration from the reference configuration C4 does not improve the performance of the tsunami detections method, since the number of missed detections may be reduced but at the expense of introducing more false alerts and vice versa. The configuration C4 can be considered satisfactorily good, but the best strategy does not seem to be the search for the best universal configuration holding for all sites of the database, but rather to adapt the TEDA operational configuration to the characteristic features of each site by a proper calibration procedure.

4.5 Sites with multiple records

In our database there are 24 sites with 2 records, and 1 with 3 records, corresponding to different sensors, usually one pressure gauge and one radar gauge. For some of the locations, the respective records look very alike, while some other records are somewhat different, especially in the background noise level. For every couple of records, the cross-correlation coefficient was calculated over the whole record and it is reported in Table 5, together with the time shift for which it was found to be maximum. All sites have coefficients larger than 0.9, and in several cases even larger than 0.99, with the only exception of the station of

Table 5. Detections for stations with multiple records. For each station pair, we report the type of instruments installed; the time shift t_s of the first series with respect to the second of the pair corresponding to the maximum cross-correlation coefficient r as well as the maximum cross-correlation coefficient itself. Moreover, for every configuration, results of detections are given with a couple of letters, one for each record. When not explicitly given, the first letter refers to channel 1 and the second to channel 2. The letter D is used to mean "detection" and the letter m to mean "missed detection". Notice that stations are sorted by decreasing values of the correlation coefficient. Out of the total of 175 combinations resulting from 25 pairs of channels analysed with 7 configurations, detection discrepancies occur in 46 cases, which is about 26 %.

site	Instrument type	t_s (min)	Correlation $r(t_s)$	C1	C2	C3	C4	C5	C6	C7
lega	prs, rad	−1	0.99990	D D	D D	D D	D D	D D	D D	D D
talc	prs, rad	0	0.99934	D D	D D	D D	D D	D D	D D	D D
acaj	prs, rad	0	0.99876	D D	D m	D D	D m	m m	m m	m m
balt	prs, rad	−1	0.99847	D D	D D	D D	D D	D D	D D	D D
quepo	prs, rad	−1	0.99800	D D	D D	D D	D D	D D	D D	D D
valp	prs, rad	0	0.99694	D D	D D	D D	D D	D D	D D	D D
riki	prs, rad	−1	0.99686	D D	D D	D D	D D	D D	D D	D D
hiva	prs, prs	0	0.99647	D D	D D	D D	D D	D D	D D	D D
juan	prs, rad	0	0.99444	D D	D D	D D	D D	D D	D D	D D
corr	prs, rad	0	0.99253	m D	m D	D D	D D	D D	D D	D D
pape	prs, rad	−1	0.99298	D D	D D	D D	D D	D D	D D	D D
john	prs, rad	−6	0.99138	D D	D D	D D	D D	D D	D m	D m
sant	prs, rad	−1	0.98928	D D	D D	D D	D D	D D	D D	D D
acap	prs, rad	0	0.98505	D D	D D	D D	D D	D D	D D	D D
lebu	prs, rad	−1	0.98135	D D	D D	D D	D D	D D	D D	D D
aric	prs, rad	0	0.97847	m D	m D	m D	m D	m D	m D	m D
nuku (d1, d2)	prs, rad	0	0.97790	m D	m D	m D	m D	D D	m D	D D
sano	prs, rad	0	0.97399	D D	D D	D D	D D	D D	D D	D D
manz	prs, rad	0	0.96240	m D	m D	m D	m D	m D	m D	m D
coqu	prs, rad	0	0.95764	D D	D D	D D	D D	D D	D D	D D
iqui	prs, rad	0	0.95045	m D	D D	m D	D D	m D	D D	m D
anto	prs, rad	−1	0.94649	D D	m D	D D	m D	m D	m D	m D
nuku (d2, d3)	rad, rad	0	0.94466	m D	D D	D D	D D	D D	D D	D D
nuku (d1, d3)	rad, rad	0	0.93936	D D	m D	m D	m D	D D	m D	D D
kaum	prs, rad	0	0.90511	D D	D D	D D	D D	D D	D D	D D
fren	prs, rad	0	0.90146	m D	m D	m D	m D	m D	m D	m D
const	prs, rad	0	0.80152	m m	m m	m m	m m	m m	m m	m m

Constitucion, Chile (plotted in Fig. 7), for which it is about 0.8. However, by looking at the Table 2, one may see that in many of these sites the tsunami exhibits different amplitudes in the different records. It is also interesting to remark that for most pairs of records, the relative time shift is confined to at most one minute (corresponding to a unit position at the given sample rate). However, for the records of the Johnston Atoll, USA, the shift of 6 min seems to be too large and rather anomalous.

The results of the tsunami detection method applied with all configurations are also shown in Table 5, where a detection is marked by the letter D and a missed detection by the letter m. That different configurations provide different results when applied to the same record is expected and was already commented in the previous subsection. The attention here is focused on the result of a configuration applied to the records of the same station. Usually, one treats these records as being equivalent, since the main reason to install more than one recording instrument in the same site is to have backup signals that can be used and processed in case the reference system has a malfunction. What is found (see Table 5) is that in most cases, there is consistency with detection (DD) or missed detection (mm) in both records of the analyzed pair, but there are also several examples of inconsistent results (about one fourth of the total), even for records with very high cross-correlation coefficients, although the higher the record correlation, the more the results are consistent. Moreover, for three stations, that is Arica, Chile, Manzanillo, Mexico and French Frigate, Tern Island, Hawaii, all the configurations provide a detection in one record and a non-detection in the other. The reason that even pairs of strongly correlated records can lead to different results for TEDA is due to the fact that the cross-correlation coefficient measures the average similarity of the records over the

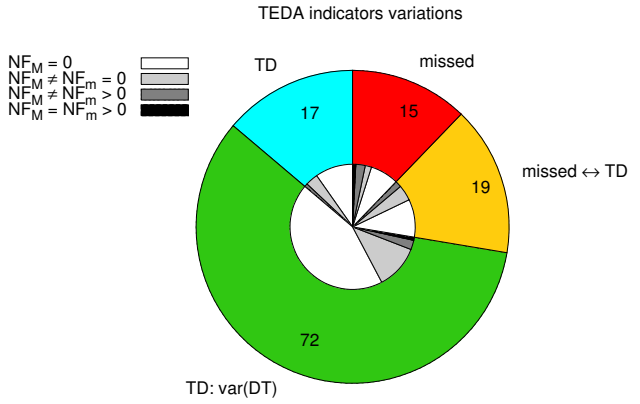

Fig. 11. Pie chart of main changes resulting from the use of the 7 configurations of Table 4. The records are grouped according to the following categories: (TD – light blue) number of records where all configurations make exactly the same detection; (TD: var(DT) – green) number of records where all configurations detect a tsunami, but with changes in DT; (missed – red) number of records where all configurations fail to detect a tsunami; (missed $\longleftrightarrow$ TD – yellow) number of records where some configurations detect a tsunami and some others do not. The inner of the pie-chart shows the statistics of false detections. Four categories are identified: ($NF_M = NF_m = 0$ – white) number of records where all configurations see no false detections; ($NF_M = NF_m > 0$ - black) number of records where all the configurations see exactly the same false detections; ($NF_M \neq NF_m > 0$ – grey) number of records where all configurations see false alerts but differing in number; ($NF_M \neq NF_m = 0$ – light grey) number of records where some configurations see false detections and some others do not.

time scale of the record length, while TEDA makes computations over much shorter time scales that are determined by the temporal parameters. The most important one in calculating the functions IS(t) and BS(t) is the parameter T_{IS} (12 min for all configurations) that serves for computing the average slope of the detided signal. So in principle, an inconsistency can occur when two records that are identical for most of the time (which ensures high correlation) differ significantly in a short time window within which the alert conditions are met in one of the two and not in the other. What happens in reality in the cases we examined is that the function $|IS(t)|$ that is checked for the detection (occurring when $|IS(t)| \geq \max(\lambda_{IS}, \lambda_{CF}BS(t))$) gets close to the alert level for both records of the pair, but it surpasses the threshold only for one record of the couple at some instant. In these cases of inconsistencies, one could suggest accepting the detection and to reject the non-detection, for a principle of precaution. In other words, when in a station there are multiple records, it can be assumed that a tsunami is detected at the station if it is detected in at least one of the records, and in case of multiple detections, the delay time of station detection is the minimum of the observed delay times.

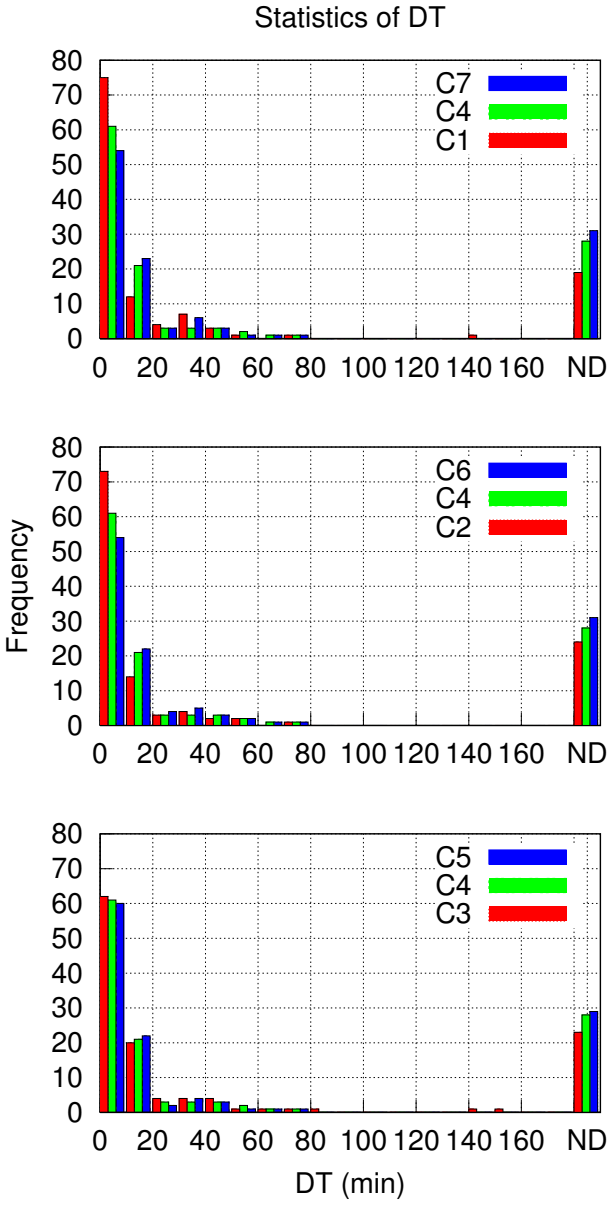

Fig. 12. Frequency histograms of the delay times DT for the 7 configurations applied in the sensitivity test. The reference configuration C4 is shown in all three graphs in green. On the right side of the graphs the frequencies of the no-detection cases (ND) are also shown.

5 Conclusions

In this work TEDA was tested on 123 tsunami records of the 11 March 2011 Tohoku tsunami taken from the Pacific region, with the configuration of the parameters (including temporal parameters and thresholds) that resulted in the optimal configuration for the station of Adak Island, Alaska (Bressan and Tinti, 2011) and that is called configuration C4

in this paper. The dataset was formed by selecting sea-level records (with one sample per minute) from the IOC repository of sea level monitoring data available at the website http://www.ioc-sealevelmonitoring.org from stations that cover a large spectrum of distances from the source and a wide range of tsunami amplitudes, from very weak to very strong wave oscillations.

All records were first analysed by visual inspection to get basic data on the tsunami arrival time, the period of the first wave, the total duration of the tsunami, and on the maximum amplitude and ranges of the oscillations, etc. One interesting finding was that tsunami travel times can be surprisingly well approximated by considering a wave propagating at a constant speed of about 206 m s^{-1}, and deviations from this crude approximation are usually small (less than 1 h) apart from few exceptions. The period of the tsunami first wave was found to differ very much from site to site and to be generally confined between 20 and 80 min, but values even exceeding 100 min were observed. However, no clear dependence was seen from the distance, as expected. What is unexpected is that interestingly, there is no clear dependence on the distance also for the tsunami range (a sort of a peak-to peak height) that appears to be strongly dominated by local response for distances larger than 2000 km. And so we found large amplitude oscillations (R1 exceeding 2 m) even at stations farther than 5000 km (see Hilo, Kahului and Kawaihae in Hawaii) or 15 000 km (Constitucion, Chile).

The application of TEDA (with the reference setting of configuration C4) to the records proved that it performed quite well. It has been seen that the secure detection method detects tsunamis with range greater than 30–35 cm and the tsunami detection method detects even smaller tsunamis, provided that they start with an impulsive wave and that there is not much noise, while it can miss even large tsunamis if they start with gradually increasing oscillations. The combination of the two methods was therefore able to detect the majority of the tsunamis with only 18 cases of missing detections, all of which concerned small tsunamis, with the exception of one record (raro2). What is also relevant is that the quality of the detection is very good, since most of detections occurred on the first wave with delay time DT shorter than 20 min, and almost all within the first hour. The sensitivity test carried out by varying the TEDA thresholds λ_{IS} and λ_{CF} showed that the results are more sensitive to varying λ_{IS} than to changing λ_{CF} and that in general decreasing the threshold values increases the number of detections but at the same time even the number of false detections. Most of the records are however stable with respect to the parameter configurations in the sense that they get a detection or a missed detection by all the 7 configurations tested. Only 19 records of the database indeed switch from detection to non-detection in passing from one configuration to the other. The conclusion is that none of the tested configurations seem to be superior to configuration C4.

The application of TEDA to the multiple records of the same station showed that one might get different results such as inconsistent detections even if records are quite similar (when measured through the cross-correlation coefficient). This suggests that in the presence of more than one record, one can assume that a detection has occurred if it has occurred in at least one of the records. In conclusion in this paper it was shown that the application of the TEDA with the same configuration (C1–C7); to records from different stations provides satisfactory results. A further improvement can be expected only after a specific calibration of TEDA, that is after selecting that set of parameters (temporal parameters as well as thresholds) which maximize the performance in a given site in view of its characteristic background and taking into account the possible tsunami characteristics. Applying extensively algorithms like TEDA that are real-time and able to provide timely detections (usually on the first tsunami wave) could be quite useful in TWS operational practice, since it could be a quick and low-cost tool for a reliable assessment of the ongoing hazard and therefore for providing essential data to authorities to make rapid decisions in critical moments of emergency.

Acknowledgements. The work of this paper is the follow up of research started under the project TRANSFER (FP6 EU Project no. 037058; http://www.transferproject.eu/); it has been completed thanks to the project TRIDEC (FP7 EU project no.258723; http://www.tridec-online.eu/).

The authors are indebted to the referees who evaluated the paper for their valuable suggestions.

Edited by: E. Pelinovsky
Reviewed by: V. K. Gusiakov and two other anonymous referees

References

Beltrami, G. M.: Automatic, real-time detection and characterization of tsunamis in deep-sea level measurements, Ocean Eng., 38, 1677–1685, 2011.

Beltrami, G. M. and Di Risio, M.: Algorithms for automatic, real-time tsunami detection in wind-wave measurements Part I: Implementation strategies and basic tests, Coast. Eng., 58, 1062–1071, 2011.

Bressan, L. and Tinti, S.: Structure and performance of a real-time algorithm to detect tsunami or tsunami-like alert conditions based on sea-level records analysis, Nat. Hazards Earth Syst. Sci., 11, 1499–1521, doi:10.5194/nhess-11-1499-2011, 2011.

Casey, K., Lim, A., and Dozier, G.: A Sensor Network Architecture for Tsunami Detection and Response, Int. J. Distrib. Sens. N., 4, 28–43, 2008.

Chang, E. T. Y. and Chao, B. F.: Co-seismic surface deformation of the 2011 off the Pacific coast of Tohoku Earthquake: Spatio-temporal EOF analysis of GPS data, Earth Planets Space, 63, 649–654, 2011.

Falck, C., Ramatschi, M., Subarya, C., Bartsch, M., Merx, A., Hoeberechts, J., and Schmidt, G.: Near real-time GPS applications

for tsunami early warning systems, Nat. Hazards Earth Syst. Sci., 10, 181—189, doi:10.5194/nhess-10-181-2010, 2010.

Farreras, S., Ortiz, M., and Gonzalez, J. I.: Steps towards the implementation of a tsunami detection, warning, mitigation and preparedness program for southwestern coastal areas of Mexico, Pure Appl. Geophys., 164, 605–616, 2007.

Gurgel, K., Dzvonkovskaya, A., Pohlmann, T., Schlick, T., and Gill, E.: Simulation and detection of tsunami signatures in ocean surface currents measured by HF radar, Ocean Dynam., 61, 1495–1507, 2011.

Henson, J. I., Muller-Karger, F., Wilson, D., Morey, S., Maul, A. G., Luther, M., and Kranenberg, C.: Strategy Geographic Positioning of Sea Level Gauges To Aid in Early Detection of Tsunami, Sci. Tsunami Hazards, 25, 173–207, 2006.

Höchner, A., Babeyko, A. Y., and Sobolev, S. V.: Geodetic Source Model for the 2011 Tohoku Earthquake and Tsunami, in: Geophys. Res. Abstr., 13, General Assembly European Geosciences Union, Vienna, Austria, 3–8 April 2011, EGU2011-14272, 2011.

Lipa, B. J., Barrick, D. E., Bourg, J., and Nyden, B. B.: HF Radar Detection of Tsunamis, J. Oceanogr., 62, 705–716, 2006.

Mcgehee, D. D. and McKinney, J. P.: Tsunami Detection and Warning Capability Using Nearshore Submerged Pressure Transducers, Case Study of the 4 October 1994 Shikotan Tsunami, in: Perspectives on Tsunami Hazard Reduction: Observations, Theory and Planning, edited by: Hebenstreit, G. T., Springer Verlag, 133–143, 1997.

Mofjeld, H. O.: Tsunami detection algorithm, only available on line at http://nctr.pmel.noaa.gov/tdadocumentation.html, 1997.

Mori, N., Takahashi, T., Yasuda, T., and Yanagisawa, H.: Survey of 2011 Tohoku earthquake tsunami inundation and run-up, Geophys. Res. Lett., 38, L00G14, doi:10.1029/2011GL049210, 2011.

Okal, E. A., Piatanesi, A., and Heinrich, P.: Tsunami Detection by Satellite Altimetry, J. Geophys. Res.-Sol. Ea., 104, B1, 599–615, 1999.

Omira, R., Baptista, M. A., Matias, L., Miranda, J. M., Catita, C., Carrilho, F., and Toto, E.: Design of a Sea-level Tsunami Detection Network for the Gulf of Cadiz, Nat. Hazards Earth Syst. Sci., 9, 1327–1338, doi:10.5194/nhess-9-1327-2009, 2009.

Pollitz, F. F., Bürgmann, R., and Banerjee, P.: Geodetic slip model of the 2011 M9.0 Tohoku earthquake, Geophys. Res. Lett., 38, L00G08, doi:10.1029/2011GL048632, 2011.

Rudloff, A., Lauterjung, J., Münch, U., and Tinti, S.: Preface "The GITEWS Project (German-Indonesian Tsunami Early Warning System)", Nat. Hazards Earth Syst. Sci., 9, 1381–1382, doi:10.5194/nhess-9-1381-2009, 2009.

Satake, K. and Atwater, B. F.: Long-term perspectives on giant earthquakes and tsunamis at subduction zones, Annu. Rev. Earth. Pl. Sc., 35, 349–374, 2007.

Schindelé, F., Loevenbruck, A., and Hébert, H.: Strategy to design the sea-level monitoring networks for small tsunamigenic oceanic basins: the Western Mediterranean case, Nat. Hazards Earth Syst. Sci., 8, 1019–1027, doi:10.5194/nhess-8-1019-2008, 2008.

Šepić, J., Denis, L., and Vilibić, I.: Real-time procedure for detection of a meteotsunami within an early tsunami warning system, Phys. Chem. Earth., 34, 1023–1031, 2009.

Shao, G., Ji, C. and Zhao, D.: Rupture process of the 9 March, 2011 Mw 7.4 Sanriku-Oki, Japan earthquake constrained by jointly inverting teleseismic waveforms, strong motion data and GPS observations, Geophys. Res. Lett., 38, L00G20, doi:10.1029/2011GL049164, 2011.

Shimizu, K., Nagai, T., Lee, J.H., Izumi, H., Iwasaki, M., and Fujita, T.: Development of Real-Time Tsunami Detection System Using Offshore Water Surface Elevation Data, in: Proc. Techno-Ocean 2006, 19th JASNAOE Ocean Engineering Symposium, Kobe, Japan, 18–20 October 2006, 24, 2006.

Stosius, R., Beyerle, G., Helm, A., Hoechner, A., and Wickert, J.: Simulation of space-borne tsunami detection using GNSS-Reflectometry applied to tsunamis in the Indian Ocean, Nat. Hazards Earth Syst. Sci., 10, 1359–1372, doi:10.5194/nhess-10-1359-2010, 2010.

Synolakis, C. E. and Bernard, E. N.: Tsunami science before and after Boxing Day 2004, Philos. T. R. Soc. A., 364, 2231–2265, doi:10.1098/rsta.2006.1824, 2006.

Tinti, S., Armigliato, A., Pagnoni, G., Zaniboni, F., Gallazzi, S. C., Bressan, L., and Tonini, R.: The 11th March 2011 Tohoku tsunami: preliminary numerical simulations of observed sea-level signals and run-up/inundation, in: Abstract Proceedings, XXV IUGG General Assembly, Melbourne, Australia, 28 June – 7 July 2011, 6062, 2011.