



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

DOTTORATO DI RICERCA IN
IL FUTURO DELLA TERRA, CAMBIAMENTI CLIMATICI E
SFIDE SOCIALI

Ciclo 36

Settore Concorsuale: 04/A4 - GEOFISICA

Settore Scientifico Disciplinare: GEO/10 - GEOFISICA DELLA TERRA SOLIDA

**TSUNAMI DETECTION STRATEGIES AND ALGORITHMS FOR OCEAN
BOTTOM PRESSURE GAUGES IN AN EARLY WARNING CONTEXT**

Presentata da: *Cesare Angeli*

Coordinatore Dottorato

Silvana Di Sabatino

Supervisore

Alberto Armigliato

Esame finale anno 2024

Abstract

The last two decades have witnessed a rapid increase of awareness regarding tsunami hazard, especially since the 2004 Sumatra event. This has been accompanied by the development of Tsunami Early Warning (TEW) systems all around the world, which includes instrumental monitoring networks and infrastructures to allow rapid communication with stakeholders and civil protection institutions. Classically, TEW systems have been based on seismic monitoring, although direct observation of travelling tsunami waves has acquired an ever more important role. The de facto standard tsunami measurement device today is the Ocean Bottom Pressure Gauge (OBPG), which measures the water column pressure at the bottom of the sea. Many TEW systems today are based on the detection in real time of tsunami waves by OBPGs that are continuously operating and transmitting data. In this thesis, we test four tsunami detection algorithms, namely Mofjeld's algorithm, tide removal with Empirical Orthogonal Functions, the Tsunami Detection Algorithm and a method based on the Fast Iterative Filtering (FIF) and IMFogram algorithms. FIF and IMFogram are data driven signal analysis tools which are used to decompose a signal into elementary oscillatory modes and obtain a time-frequency representation of the signal. We show that these techniques can be a compelling alternative to classical analysis methods, since they give equivalent results, with the added robustness of data driven methods and the ability of performing multiple operations (denoising, tide removal, bandpass filtering) in a single operation. Then, we leverage on their properties to develop a new tsunami detection algorithm. The four detection methods are tested against two datasets built from OBPG data from the NOAA's Deep-ocean Assessment and Reporting of Tsunamis (DART). The first dataset consists of month-long signals including only tides and noise, while the second includes day-long records acquired during past tsunami events. We discuss possible criteria to choose the optimal amplitude detection threshold, based on detection rates of tsunamis, seismic shaking and possible false detections. The newly developed FIF-based technique shows very promising results, both in terms of false detection rates and low optimal detection thresholds. Finally, the ability of the technique to characterize the tsunami waveshape in real time are discussed and strategies to reduce errors in operational contexts are presented.

Contents

1	Introduction	1
1.1	Global impact of tsunamis	1
1.2	Tsunami recording and applications of real time detection	3
1.3	Outline of the present work	10
2	Real-time tsunami detection algorithms	13
2.1	Mofjeld’s algorithm	13
2.2	Tide removal through Empirical Orthogonal Functions	17
2.3	Tsunami Detection Algorithm	20
2.4	The Fast Iterative Filtering and IMFogram techniques	22
2.4.1	Analysis of nonlinear and nonstationary signals	22
2.4.2	The (Fast) Iterative Filtering algorithm	24
2.4.3	Time-Frequency Representation and the IMFogram Algorithm	26
2.4.4	FIF-based Tsunami Detection Algorithm	27
2.5	Conclusions	30
3	Data-driven analysis of tsunami signals	32
3.1	Tide gauge signal from the 02/05/2020 Crete Tsunami	33
3.2	Tide gauge signals from the 06/02/2023 Eastern Mediterranean event	39
3.3	Main takeaways	48
4	Detection tests on background time series	50
4.1	NOAA’s DART® data	50
4.2	Dataset creation	52
4.3	Detection curves	53
4.4	Detection results for background signals	55

<i>CONTENTS</i>	iv
4.5 Conclusions	60
5 Detection tests on tsunami signals	77
5.1 Dataset creation	78
5.2 Detection analysis	81
5.3 Tsunami characterization from detection curves	92
5.4 Conclusions	96
Conclusion	111
Acknowledgements	114

List of Figures

1.1	Schematic representation of modern tide gauge.	3
1.2	Historical examples of tsunami waves recorded at TGs.	4
1.3	Map page of the Sea Level Monitoring Facility website.	5
1.4	Schematic example of a DART 4G OBP.	6
1.5	Examples of DART tsunami records.	7
1.6	Map page of the Deep-ocean Assessment and Reporting of Tsunamis website.	8
2.1	Example of application of the detection algorithm by Mofjeld (1997)	16
2.2	Example of Empirical Orthogonal Functions for one lunar day. . .	19
2.3	Example of FIF-based tsunami detection.	28
3.1	Epicenter of 02/05/2020 Crete earthquake and tide gauges located in Ierapetra (NOA-04) and Kasos (NOA-03).	33
3.2	Sea level record measured at the Ierapetra tide gauge.	34
3.3	Intrinsic mode functions computed using the FIF technique for the 02-05-2020 tsunami signal recorded at Ierapetra.	35
3.4	Instantaneous amplitude and instantaneous frequency for the IMFs of the Ierapetra record.	36
3.5	Comparison of EMD and FIF methods in computing the IMFs. . .	38
3.6	Comparison of IMFs computed through FIF and signals extracted with harmonic filters.	38
3.7	Epicenter of the first mainshock in the 06/02/2023 Turkey and Syria seismic sequence, and tide gauges located in Arsuz, Erdemli and Tasucu.	40
3.8	Raw data from 3 tide gauges located near the southern Turkish coast.	41
3.9	Joint FIF + IMFogram analysis of the Arsuz tide gauge record. .	44

3.10	Joint FIF + IMFogram analysis of the Erdemli tide gauge record.	45
3.11	Comparison of tsunami waves extracted from a FIF+IMFogram analysis and with bandpass filters.	46
3.12	Joint FIF + IMFogram analysis of the Tasucu tide gauge record.	47
4.1	Inconsistency of tsunami components at different time steps for the FIF-based detection method.	54
4.2	Variability range of TDA detection curves for different tidal models.	59
4.3	Raw data relative to case A in background analysis.	62
4.4	Absolute values of the FFT for detection curves obtained for background case A.	62
4.5	Detection curves for each technique for background case A.	63
4.6	Histograms of the detection curves for each technique for background case A.	64
4.7	Raw data relative to case B in background analysis.	65
4.8	Absolute values of the FFT for detection curves obtained for background case B.	65
4.9	Detection curves for each technique for background case B.	66
4.10	Histograms of the detection curves for each technique for background case B.	67
4.11	Raw data relative to case C in background analysis.	68
4.12	Absolute values of the FFT for detection curves obtained for background case BC	68
4.13	Detection curves for each technique for background case C.	69
4.14	Histograms of the detection curves for each technique for background case C.	70
4.15	Raw data relative to case D in background analysis.	71
4.16	Absolute values of the FFT for detection curves obtained for background case D.	71
4.17	Detection curves for each technique for background case D.	72
4.18	Histograms of the detection curves for each technique for background case D.	73
4.19	Raw data relative to case E in background analysis.	74
4.20	Absolute values of the FFT for detection curves obtained for background case E.	74
4.21	Detection curves for each technique for background case E.	75
4.22	Histograms of the detection curves for each technique for background case E.	76

5.1	Map of the events in Davies' catalogue	79
5.2	Statistical indicators for long signals relative to the catalogue in tab. 5.1	82
5.3	Rates of earthquake and tsunami detections in both the full and restricted datasets.	84
5.4	False detection rates for both the full and restricted datasets. . .	84
5.5	Detection score θ_1 as function of detection threshold for both the full and restricted dataset.	85
5.6	Detection score θ_2 as function of detection threshold for both the full and restricted dataset.	86
5.7	Scatterplot of maximum absolute amplitude as a function of pre- vious data availability in months for TDA detection curves. . . .	87
5.8	Raw data relative to event Pu1, DART 21415.	88
5.9	Raw data relative to event SA2, DART 44402.	88
5.10	Example of detection curves for event Pu1, DART 21415.	90
5.11	Example of detection curves for event SA2, DART 44402.	91
5.12	Example of detection curves for event Su3, DART 46404.	98
5.13	Example of detection curves for event Su3, DART 46407.	99
5.14	Example of detection curves for event KJ1, DART 46411.	100
5.15	Example of detection curves for event NH2, DART 51425.	101
5.16	Example of detection curves for event NH2, DART 52402.	102
5.17	Example of detection curves for event KT1, DART 51407.	103
5.18	Example of detection curves for event KT2, DART 32412.	104
5.19	Example of detection curves for event KJ2, DART 21414.	105
5.20	Example of detection curves for event KJ2, DART 46402.	106
5.21	Example of detection curves for event KT2, DART 51425.	107
5.22	Example of detection curves for event KT2, DART 51426.	108
5.23	Comparison between tsunami components at different times, de- tection curve, obtained by the FIF-based method, and the tsunami waveform for event KJ2, DART 52405.	109
5.24	Comparison between tsunami components at different times, de- tection curve, obtained by the FIF-based method, and the tsunami waveform for event SA3, DART 46403.	109
5.25	Comparison between tsunami components at different times, de- tection curve, obtained by the FIF-based method, and the tsunami waveform for event SA3, DART 21413.	110

List of Tables

4.1	Instrument, location, period and maximum tidal range for each background signal selected for testing.	53
4.2	Maximum and minimum value recorded in each detection curve for each technique.	56
4.3	Standard deviation of each detection curve for each technique. . .	57
5.1	Catalogue of tsunamigenic earthquakes from Davies (2019)	79

Chapter 1

Introduction

1.1 Global impact of tsunamis

In the presence of a body of water, some natural events may cause significant vertical displacements of the entire water column. Among these events, we may find offshore earthquakes, subaerial and submarine landslides, volcanic eruptions and sudden and strong atmospheric disturbances. The vertical displacement of the water body then generates propagating trains of long waves that are called tsunamis. These waves have usually amplitudes of few centimeters offshore, but due to shoaling and amplification phenomena they show remarkable increase in amplitude in the vicinity of coastal areas. For these reasons, tsunamis may represent a significant threat to coastal areas, which includes human settlements, historical and cultural heritage sites, and vulnerable or even endangered local ecosystems. The estimated losses caused by tsunamis amount to US\$280 billion in costs and more than 250 thousands casualties between 1998 and 2017 (Imamura et al., 2019). In particular, the 2004 Boxing Day tsunami, responsible for the vast majority of tsunami casualties in the aforementioned period, was a great catalyst in increasing the interest in tsunami modelling and hazard assessment (Synolakis and Bernard, 2006).

The impact that tsunamis have had in the last few decades increased the attention of public institutions and the interest of researchers in improving the understanding about the physics and phenomenology of tsunamis. For a general overview, we refer the reader to the works by Levin and Nosov (2016) and Saito (2019). However, many difficulties are still present in dealing with tsunami hazard (Behrens et al., 2021), such as:

- the complexity of triggering mechanisms and the possibility of cascading effects (e.g. an earthquake triggers a landslide, which then generates the tsunami);
- the scarcity of data, which prevents us to have a complete description of the phenomena on a purely statistical basis;
- the limits in real time forecasting methods and population alerting systems due to the large spatial scales affected in short time.

To face these difficulties, two main strategies have been developed. The first is the production of tsunami hazard and risk models, that quantify the probability that an event of a certain size, usually expressed in terms of water height at target locations, will happen within a given amount of time. Such models may be developed for different spatial scales, from global (Løvholt et al., 2014), to regional (Basili et al., 2021), to local (Volpe et al., 2019). The second strategy consists in developing so called Tsunami Early Warning (TEW) centres, i.e. institutions, usually run by multinational cooperations, that provide threat information in real time for potentially on-going tsunami events.

The most important part of TEW systems is the instrumental monitoring network. Since the majority of tsunamis are caused by earthquakes, the most commonly used instrumental data employed for TEW are seismic records. As soon as seismic data are acquired, data are processed to determine the tsunami-genic potential of the earthquake, based on empirical correlations (e.g. Lomax and Michelini, 2013), or fast algorithms for focal mechanism determination (e.g. Duputel et al., 2011). However, TEW systems based on seismic monitoring alone have significant drawbacks, namely the intrinsic uncertainties of the seismic source and the possibility of non seismic tsunamis. The former is mainly caused by the fact that the most important aspects for tsunami generation is the vertical coseismic displacement field, which is very difficult to determine in real time, especially when the epicenter is located offshore. The latter is significant because more than a quarter of tsunamis are related to non seismic origin. In locations with higher tsunami generation frequency, for example in the island of Stromboli (southern Tyrrhenian sea, Italy), local ad hoc monitoring systems are used (Selva et al., 2021a). In general, the best solution consists in sea level monitoring instruments that can measure tsunamis directly.

1.2 Tsunami recording and applications of real time detection

Historically, the first direct tsunami instrumental data were recorded by coastal Tide Gauges (TGs), which continuously measure sea level height at a fixed position. A schematic representation of a modern TG is shown in Fig. 1.1. The instruments have been used for sea level monitoring since the first half of the 19th century and self-recording variants have been available since the 1830s (Matthäus, 1972).

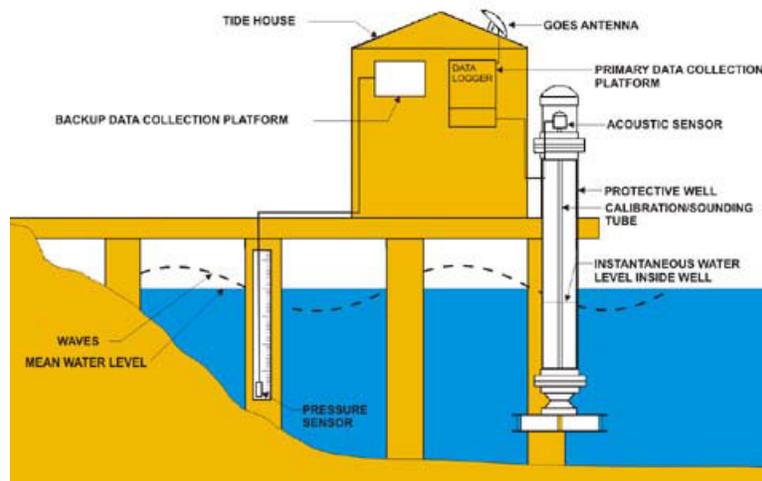


Figure 1.1: Schematic representation of modern tide gauge. Image from https://oceanservice.noaa.gov/education/tutorial_tides/tides11_newmeasure.html.

Despite the fact that TGs were deployed to monitor tides, and that old instruments had a too large sampling interval for tsunami monitoring, some of them were able to record old events, such as the one caused by the 1883 Krakatau eruption (Pelinovsky et al., 2005) and the one triggered by the 1887 Ligurian earthquake (Eva and Rabinovich, 1997). In Fig. 1.2, TG records from these events are shown. In particular, Fig. 1.2a shows the record of the Ligurian tsunami from the TG located at the Genoa harbour, while Fig. 1.2b and 1.2c are signals relative to the Krakatau event from Batavia (Jakarta) and Nagapattinam (India). A problem of TGs not suited for tsunami monitoring is the possibility of saturation, i.e. the instrument is not able to record waves over a certain amplitude. This is the case we can observe in Fig. 1.2b, where the first two

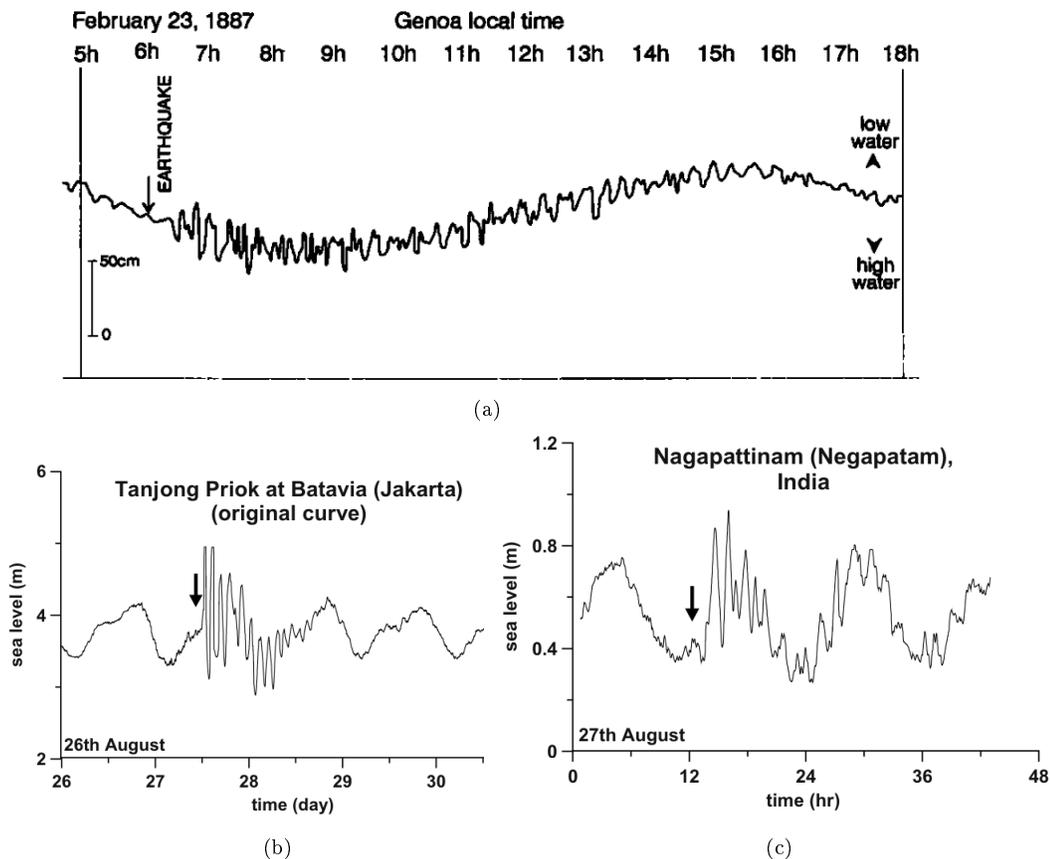


Figure 1.2: Historical examples of tsunami waves recorded at TGs. (a) record from the TG in Genoa harbour, relative to the tsunami following the 23/02/1887 Ligurian Sea Earthquake. Image from Eva and Rabinovich (1997). (b), (c) TG records for the tsunami that followed the 26/08/1883 Krakatau volcanic eruption. Image from Pelinovsky et al. (2005). Note the cut-off at 5 m amplitudes in (b).

tsunami oscillations are cut at a value of 5 m. In recent decades, the role of TGs in tsunami science has been recognized as fundamental and many instruments have been updated to better characterize tsunami events.

Nowadays, they are used both for retrospective studies (Pires and Miranda, 2001; F. Romano et al., 2016; Satake, 1987) and real time detection (Bressan et al., 2013; Lee et al., 2016). To this latter aim, UNESCO's Sea Level Station Monitoring Facility website (<https://www.ioc-sealevelmonitoring.org/>) gives access to real time sea level data from TGs managed by 175 differ-

ent data providers, through a graphical interface, shown in Fig. 1.3. However, some characteristics of TGs may be not appropriate in the context of tsunami monitoring. Firstly, they are necessarily positioned in coastal areas, thus they

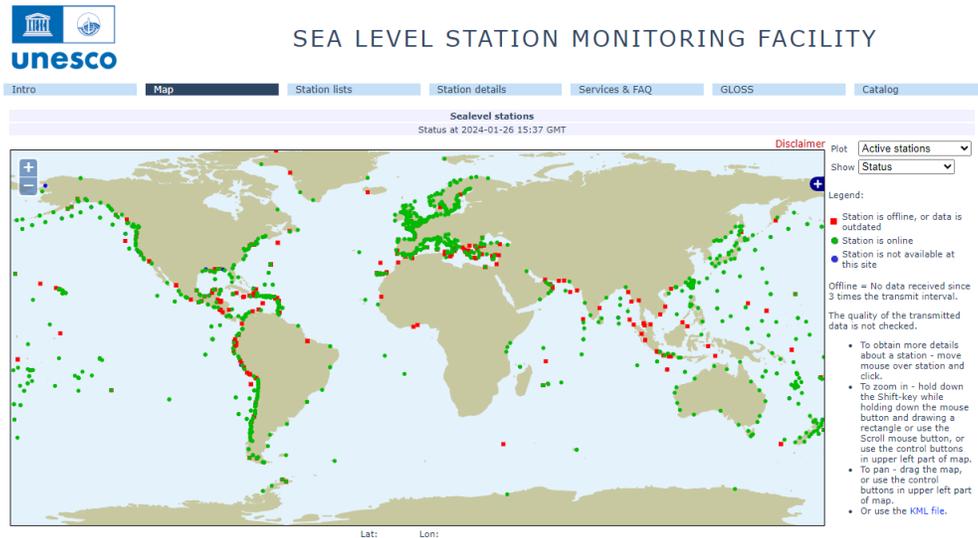


Figure 1.3: Map page of the Sea Level Monitoring Facility website (<https://www.ioc-sealevelmonitoring.org/>). The points in the map represent the stations whose data are available. Green dots indicate online stations, while red dots indicate either offline stations or stations whose data are outdated.

are able to detect sea level anomalies only when these are already close to the coast. While this is still useful to understand the evolution of the event and to warn communities further away from the source, it may be late to alert the population at the location of the TG. Secondly, the tsunami evolution at coastal locations is deeply dependent on local geomorphological and topobathymetric features. In particular, tide gauges are usually located in bays or harbours, which have normal modes that may be excited by incoming waves. This, together with oscillation modes due to the geometry of coastal shelves, is the reason of the complex response spectra observed in TG in response to incoming waves (Aranguiz et al., 2019; Rabinovich, 1997). We have to point out that in some regions, such as the Mediterranean Sea, TGs are the only available instruments in operation for tsunami warning purposes and other kinds of sensors are either absent or only experimental. Conversely, site-effects are negligible in case of measurements of tsunami waves in deep water environments. For this reason, Ocean Bottom Pressure Gauges (OBPGs) have been tested and used for

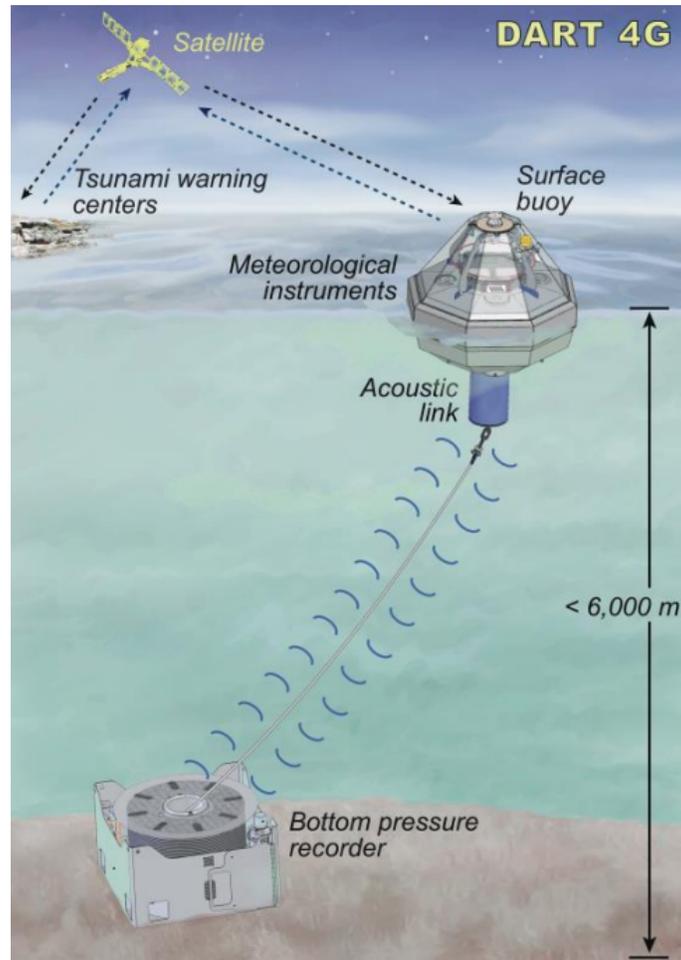
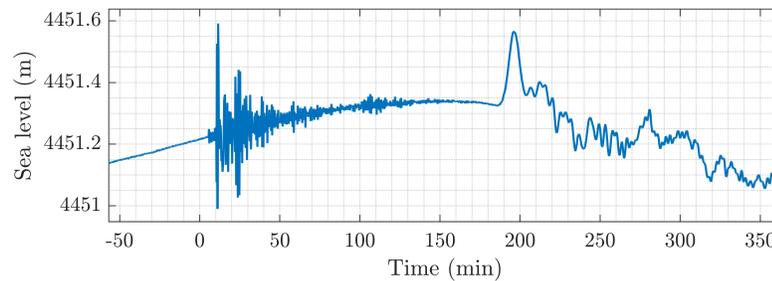


Figure 1.4: Schematic example of a DART 4G OBPG. This is the latest developed DART model. Details about DART instruments are described in later Chapters. From the DART 4G brochure, available at https://nctr.pmel.noaa.gov/Pdf/brochures/dart4G_Brochure.pdf.

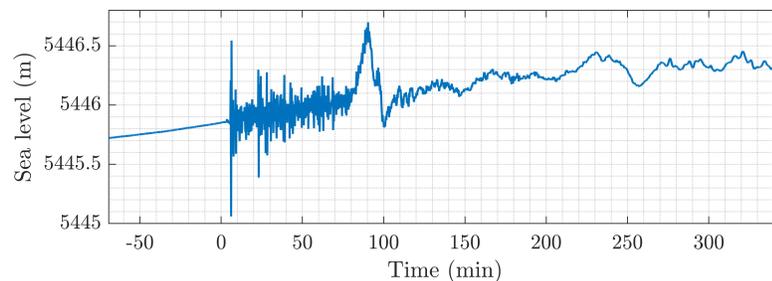
many decades. As the name implies, these instruments measure the pressure at the bottom of the water column. A schematic representation of such an instrument is shown in Fig. 1.4. The fluid layer acts as a natural filter for superficial meteoceanographical oscillations, leaving in the instrumental record only long wave phenomena, whose pressure profile is hydrostatic. Furthermore, they are not affected by nonlinearities and resonance effects that we get in TG records. Signals from OBPGs are thus a superposition of:

- tidal oscillations, in particular diurnal and semidiurnal tides, which represent the most energetic component of the signal;
- transient long wave oscillations, such as tsunamis;
- changes in bottom pressure caused by displacement of the ocean bottom.

The last case is quite significant for tsunami early warning, since OBPBs located near seismic sources will be affected by seismic shaking. Two examples of OBPBs raw tsunami records can be found in Fig. 1.5, where two signals, respectively from the 2010 Chile tsunami and the 2011 Tohoku earthquake tsunami, are shown. In both cases, long term trends related to tidal oscillations are present. Large oscillations in the periods of tens of minutes, followed by decaying codas, represent the tsunami waves. In both cases, very large high frequency seismic oscillations can be observed before the tsunami arrival.



(a)



(b)

Figure 1.5: Examples of DART tsunami records. (a) is the record from DART 32412 relative to the South American tsunami triggered by the $M_w = 8.8$ earthquake occurred on 27/10/2010 06:34:15 UTC. (b) is the record from DART 21419 relative to the Tohoku tsunami triggered by the $M_w = 9.1$ earthquake occurred on 11/03/2011 05:46:23 UTC. In both plots, time is measured in minutes since earthquake origin time

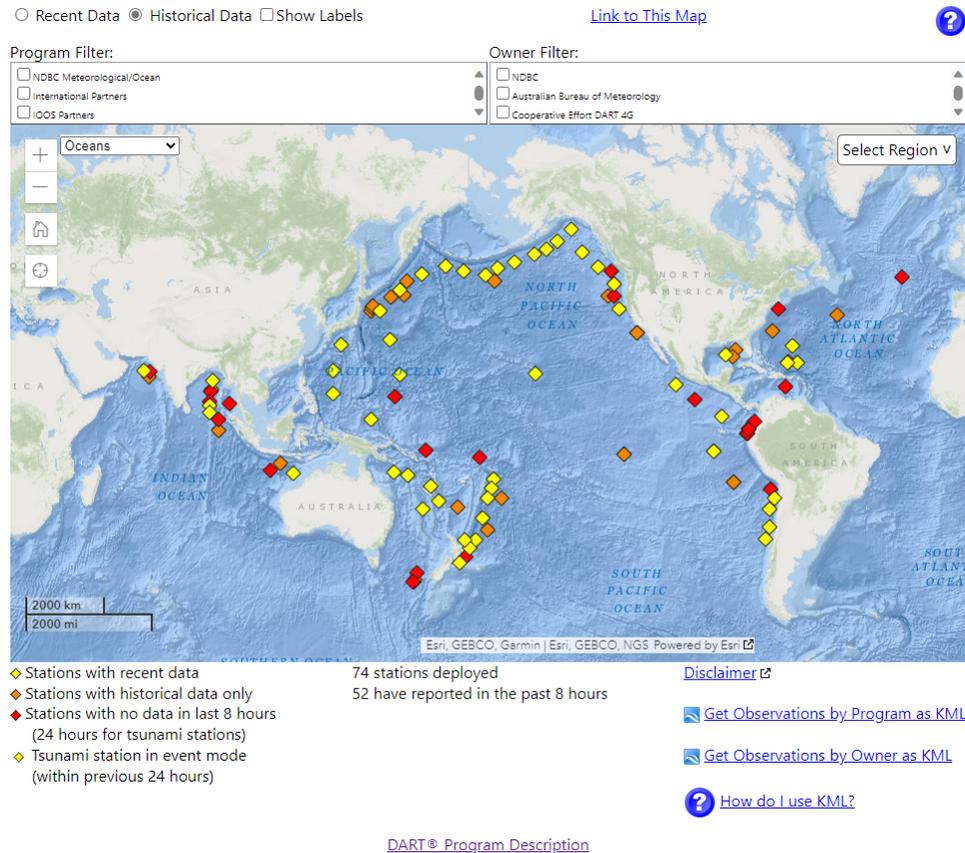


Figure 1.6: Map page of the Deep ocean Assessment and Reporting of Tsunamis website (available at <https://www.ngdc.noaa.gov/hazard/DARTData.shtml>). The points in the map represent the stations whose data are available on NOAA database, for both active and past stations.

Today, OBPGs are the standard instruments in many TEW systems. The most widely spread and used is the NOAA DART® network, which is the only global OBPG network, with instruments in the Western Indian Ocean, North Atlantic Ocean, and Pacific Ocean. Similarly to TGs, data from this network are made available through a web portal, shown in Fig. 1.6. Data from this network will be extensively used in this thesis to test tsunami detection methods, thus more details will be presented in the next Chapters. OBPGs are also used in regional dense networks, such as DONET (Kawaguchi et al., 2008) and S-NET (Mochizuki et al., 2018), installed off the coasts of Japan. These networks integrate state of the art pressure gauges with ocean bottom seismometers in

order to characterize the source and the tsunami waves as accurately as possible. Lastly, there are plans by INGV to install OBPGs off the coasts of Southern Italy (Amato et al., 2021).

The recent advancements both in detection technology and in computing power have introduced many new forecasting methods for TEW. A now classical method is a real time inversion of tsunami data from OBPGs (Percival et al., 2011; Tang et al., 2009; Titov et al., 2003). Bottom pressure data are inverted to find the optimal linear combination of precomputed elementary seismic sources, from which wave heights along different coastlines are then computed. More recently, data assimilation methods have been tested and implemented (Heidarzadeh et al., 2019; Maeda et al., 2015; Wang et al., 2019a,b, 2017, 2021). These approaches, mutated from meteorology and oceanography, allow to forecast the wave height and velocities in the near future from tsunami data, without assumptions on seismic sources. At last, a Bayesian forecasting method has been introduced by Selva et al. (2021b), where a probability distribution function is computed as soon as seismic data are acquired. Then, each time new data are available, they are used to compute a new refined probability distribution function for the expected wave heights in points of interest using Bayes' theorem. This approach has the added advantage of taking into account the uncertainties intrinsic in all the steps of data acquisition and elaboration and it could be used to take into account data from different sources.

For any of these applications, the problem of real-time tsunami detection represent a fundamental challenge. Current technology is already capable of providing invaluable informations that are used to issue timely alert messages, however many issues remain to be solved. The first issue is the choice of a tsunami detection algorithm and it is the main focus of the present work. The most used algorithm to this aim is Mofjeld's algorithm (Mofjeld, 1997), which is an incredibly computationally efficient detiding algorithm that is installed into the instruments of the DART® network. Nonetheless, the algorithm suffer from a few drawbacks, such as the inability to filter seismic shaking, which makes it unusable in the very near-field to distinguish tsunami oscillations from seismic waves. While fourth generation DART® (DART 4G) have been designed to include a specially designed filter to separate earthquake shaking from the tsunami waves, its functioning depends on obtaining data with sampling frequency of 1 Hz (Moore, 2024). However, data of these nature are still rare, given that instruments, if not externally prompted by the user, only memorize already filtered signals and that old DART models are still in use in operational contexts. Another drawback of Mofjeld's algorithm is that any anomaly

with respect to the background (tides and random noise) is distorted. Thus, the technique is not able to characterize the tsunami waveform accurately (Beltrami, 2008, 2011), which is needed for data assimilation or Bayesian forecasting methods cited before.

Other algorithms have been proposed in the literature, but, in most cases, they lack systematic tests on large dataset of real tsunami records. In the case of EOF detiding (Tolkova, 2010, 2009) and the Tsunami Detection Algorithm (Chierici et al., 2017), the present work aims to fill this gap. These techniques, contrary to Mofjeld’s algorithm, require some already acquired data in order to compute tidal models.

At last, while OBPGs are the standard instruments for tsunami detection, we should remember that they are still very expensive and difficult to install. For these reasons, alternative detection instruments have been proposed, such as wind-wave gauges (Beltrami and Di Risio, 2011; Di Risio and Beltrami, 2014), positioning of commercial ships (Foster et al., 2012), airborne instruments (Mulia et al., 2020), ionospheric sounding (Astafyeva, 2019; Occhipinti, 2011), and SMART ocean bottom cables (Howe et al., 2019), one of which has recently been installed in the Mediterranean Sea (Marinero et al., 2024).

1.3 Outline of the present work

The purpose of the present work is to test and compare algorithms for real time tsunami detection, with a special focus on algorithms suitable for OBPGs. In Chapter 2, four techniques are presented. The first is Mofjeld’s algorithm (Mofjeld, 1997), based on Newton’s forward interpolation formula. This method is of particular importance, since it is the one which instruments of the DART[®] network are equipped with. For this reason, it is considered one of the de facto standard algorithms for the purpose and it is by far the most tested tsunami detection algorithm available. It is mathematically simple and computationally efficient. However, it suffers from the drawbacks of not being able to correctly characterize the tsunami wavehape, as will be explained later in Chapter 2.

The second method consists in tidal removal by using Empirical Orthogonal Functions (Tolkova, 2009; Tolkova, 2010), where a set of basis vectors is computed from a long bottom pressure time series and then used to extract and remove tides from new data. Then, we present the Tsunami Detection Algorithm (TDA, Chierici et al., 2017), which is a sequential application of different filtering operations, i.e. tide removal with a harmonic model, spike removal and bandpass filtering. Lastly, we present a newly developed technique based on the

real time application of the Fast Iterative Filtering method (FIF, Cicone, 2020; Cicone et al., 2016; Cicone and Zhou, 2021; Stallone et al., 2020) and the IMFogram algorithm (Barbe et al., 2020; Cicone et al., 2022). The FIF method allows to decompose a signal into elementary oscillatory components called Intrinsic Mode Functions (IMFs), each with a well-defined local frequency. Then, the IMFogram algorithm may be used to obtain a time-frequency representation of the signals by computing instantaneous amplitude and instantaneous frequency of each IMF. For each technique, we present its basic mathematical structure, its needs in terms of previously acquired data and its pros and cons for their applications to pressure sensors.

Since the FIF method and the IMFogram algorithm have been developed quite recently and have not yet been applied in the context of tsunami signal analysis (which represent the most original and novel aspect of this thesis), Chapter 3 is devoted to showing how they can be used to extract informations from sea level records. This is done through the analysis of two case studies. The first is a TG signal from the 02/05/2020 Crete tsunami, chosen as representative of a tsunami signal, including the main wave, local resonance phenomena, random noise and tidal oscillations. The second case study consists in the analysis of TG signals from the 06/02/2023 small tsunami observed in the north-eastern Mediterranean sea.. The event followed the first main shock of a disastrous sequence of earthquakes occurred in Turkey and along its border with Syria, that included two events with magnitude 7.8 and 7.5 (according to USGS) respectively and many aftershocks with $M_w > 5.0$. The tsunami was probably caused by mass wasting events triggered by the first mainshock, however the number and location of slides is still debated (Heidarzadeh et al., 2023; Hu et al., 2023). It is shown how combining the FIF and IMFogram techniques we can carry out analyses similar to what is traditionally performed with classical methods (e.g. Fourier analysis or wavelet transforms), but with two added benefits. The first is that features in the time-frequency plane obtained with these techniques are better localized than with classical techniques, allowing more precise estimates of frequency content and instantaneous amplitude. The second is that the both techniques are completely data driven, allowing to perform multiple operations (denoising, detiding and extraction of tsunami waveform) as a unified process.

Then, Chapter 4 and 5 deal with the application of the four techniques to real time tsunami detection, on two different datasets, both built from DART® buoys data. In Chapter 4, the techniques are applied to 5 signals lasting one month from DART® buoys historical data, where no transient effect (seismic shak-

ing or tsunami wave) is observed. In this way, we are able to characterize the background noise of each technique's residual and to discuss the effects of the statistical distribution of such residuals. From this, we find the optimal detection thresholds for each technique to avoid false detections.

In Chapter 5, we turn the attention to testing on signals with transient oscillations. We start by describing a procedure to build a testing dataset that encompasses the wide variety of possible signals we may record in real time, based exclusively on real tsunami signals. In this way, we can avoid possible biases that may arise in using simplified waveshapes or simulated tsunami waves. Then, the detection capabilities of each techniques are quantitatively estimated, in terms of how successfully they detect tsunamis, how often they detect seismic waves and false detections occur. Criteria for choosing the optimal detection threshold for each technique are discussed. The Chapter ends with considerations regarding the characterization of the tsunami waveshape from the residuals of the techniques.

Chapter 2

Real-time tsunami detection algorithms

In this chapter, a description of four real-time tsunami detection algorithms is given. The techniques presented are Mofjeld’s algorithm (Mofjeld, 1997), detiding with Empirical Orthogonal Functions (EOFs) (Tolkova, 2009; Tolkova, 2010), Tsunami Detection Algorithm (Chierici et al., 2017) and a newly developed technique based on data driven signal analysis techniques, namely the Fast Iterative Filtering (FIF) technique (Cicone, 2020; Cicone et al., 2016; Cicone and Zhou, 2021) and the IMFogram algorithm (Barbe et al., 2020; Cicone et al., 2022). For all the techniques, the basic mathematical structure is presented and the properties are discussed, with particular attention given to application to Ocean Bottom Pressure Gauge (OBPG) data. Since FIF and IMFogram have not yet been applied to the analysis of tsunami signals, their theoretical aspects are considered in more detail than the other techniques.

2.1 Mofjeld’s algorithm

The NOAA Pacific Marine Environmental Laboratory (PMEL) developed the DART® buoys, which today are sort of standard instruments for tsunami off-shore real-time monitoring and reporting. A DART sensor is composed by a bottom pressure recorder, positioned at the seafloor, and by a surface buoy that transmits data to monitoring facilities. The two components communicate between themselves through acoustic signals. The DART program has its origin in the first experiments in direct tsunami measurements in deep oceans in the

1980s and has now provided continuous sea level monitoring for more than two decades with a variable number of sensors at any given moment in the Pacific, Northwestern Indian and North Atlantic oceans (Li, 2022; Rabinovich and Eblé, 2015).

Each DART in NOAA's network includes in its software an automatic tsunami detection algorithm developed by Mofjeld (1997). This algorithm is mathematically and computationally very efficient and, being the standard in the only global tsunami monitoring network, it has a very long history of studying and testing. The basic idea is to forecast ocean bottom pressure by extrapolating the long term trend in the data. As soon as a new pressure measurement is acquired, it is compared with the forecast. A detection is triggered if the absolute difference between the measurement and the forecast is bigger than a given threshold T .

The forecast is computed using Newton's forward extrapolation formula

$$\eta_p(t') = \sum_{i=0}^3 w_i \bar{\eta}(t - idt) \quad (2.1)$$

where η_p is the forecast bottom pressure, $\bar{\eta}(t)$ is the moving average of the pressure at time t , dt is the time between points used for extrapolation, t' is the prediction time and w_i are weights to be computed. To find the value of w_i , let us write $\eta_p(t)$ as a third degree polynomial

$$\eta(t) = a_0 + a_1(t - t_1) + a_2(t - t_1)(t - t_2) + a_3(t - t_1)(t - t_2)(t - t_3) \quad (2.2)$$

where t is the current time, t_1 , t_2 and t_3 are arbitrary the time instants at which the value of the polynomial is assumed to be known. Consistently with eq. (2.1), we choose $t_i = t - idt$, i.e. they are chosen at a constant time interval from each other. The extrapolation is made from averaged values to reduce the influence of random fluctuations and noise. The values of a_i can be found by substituting the values $(t_0 - idt, \bar{\eta}_i)$, where $\bar{\eta}_i = \bar{\eta}(t_0 - idt)$, from which

$$\begin{aligned} a_0 &= \bar{\eta}_0 \\ a_1 &= -\frac{\bar{\eta}_1 - \bar{\eta}_0}{dt} \\ a_2 &= \frac{\bar{\eta}_2 - 2\bar{\eta}_1 + \bar{\eta}_0}{2dt^2} \\ a_3 &= -\frac{\bar{\eta}_3 - 3\bar{\eta}_2 + 3\bar{\eta}_1 - \bar{\eta}_0}{6dt^3} \end{aligned} \quad (2.3)$$

If we express the prediction time t' in terms of present time t as $t' = t + p dt$, we can express the forecast as

$$\begin{aligned} \eta(t') = \eta(t + p dt) = & \bar{\eta}_0 \left(1 + \frac{11}{6} p + p^2 + \frac{1}{6} p^3 \right) + \\ & + \bar{\eta}_1 \left(-3p - \frac{5}{2} p^2 - \frac{1}{2} p^3 \right) + \\ & + \bar{\eta}_2 \left(\frac{3}{2} p + 2p^2 + \frac{1}{2} p^3 \right) + \\ & + \bar{\eta}_3 \left(-\frac{1}{3} p - \frac{1}{2} p^2 - \frac{1}{6} p^3 \right) \end{aligned} \quad (2.4)$$

The default settings of DART buoys consider a prediction time 5 min15 s after current time, $dt = 1$ h and the moving average is computed over intervals of 10 min. From these parameters, we get that $p = 0.0875$ and the weights w_i are computed from eq. (2.4) as

$$\left\{ \begin{array}{l} w_0 = +1.16818475 \\ w_1 = -0.28197559 \\ w_2 = +0.14689746 \\ w_3 = -0.03310644 \end{array} \right. \quad (2.5)$$

We note that, to the best of our knowledge, the full derivation of eq. (2.5) and the explicit expression in eq. (2.4) are not reported anywhere else in tsunami literature.

The properties of Mofjeld's algorithm for detection can be well understood by looking at an example, as the one in Fig. 2.1. In the example, the technique is applied to data from DART 21414 recorded during the 11/03/2011 Tohoku tsunami. The residual is compared with the post-processed tsunami wave from Davies (2019). The most evident characteristic of the forecast is that a few minutes after the arrival of the tsunami, it deviates from the long term trends due to tidal oscillations. In fact, since the prediction time is set at 5 min15 s after current time, eq. (2.1) uses tsunami data to estimate the long term trends. This effect should not influence the cases of waves with periods of few minutes. However, this is rarely the case for medium and large tsunamis. A possible solution has been proposed by Beltrami (2011), which consists in taking a longer prediction time. Doing this delays the influence of the tsunami waves on the tide forecast long enough to allow the characterization of the leading wave in terms of amplitude and period. However, Beltrami (2011) shows that this mod-

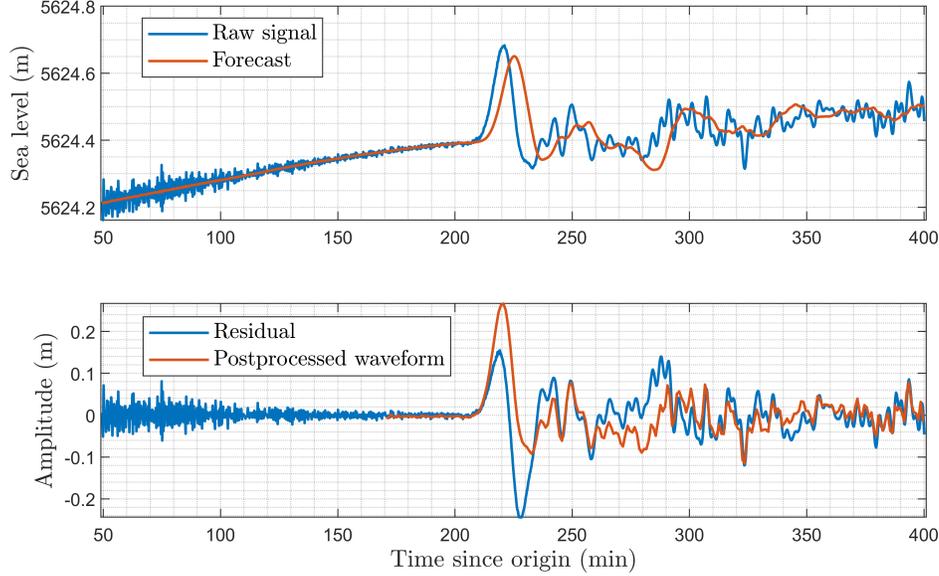


Figure 2.1: Example of application of the detection algorithm by Mofjeld (1997). The upper plot shows a comparison between the acquired sea level data and the long term forecast computed using eq. (2.4) with coefficients in eq. (2.5). The lower plot shows the residual produced by the technique and the post-processed waveform (Davies, 2019). Data relative to the 11/03/2011 Tohoku tsunami from DART 21414.

ification results in a smaller signal-to-noise ratio in the forecast, which would make smaller tsunamis undetectable.

Another characteristic of the method is that it does not filter out high frequency components. For this reason, random noise with few millimeters amplitude remain present in the residual (as will be shown in the next chapters), as do seismic waves. Modern DART 4G models have additional modules to deal with the seismic waves once they are detected, thanks also to the higher sampling frequency of 1 Hz, with respect to the 1/15 Hz that is common in older instruments.

At last, we may note that the only strong assumption of Mofjeld’s algorithm is the separation of time scales, i.e. that we want to detect a signal with typical periods of minutes from a background where hour-long oscillations dominate. Thus, the technique can be applied to instruments different from deep-ocean gauges. An example of this can be found in the work by Beltrami and Di Risio

(2011), where the detection of tsunamis from wind-wave gauges is proposed and analysed. In that case eq. (2.1) is used to remove tides from the signal.

2.2 Tide removal through Empirical Orthogonal Functions

Another site independent technique for removing long tidal oscillations from tsunameters' records has been developed by Tolкова (2009) and modified and further developed by Tolкова (2010). The technique is based on the use of Empirical Orthogonal Functions (EOFs) to analyze time series. The idea behind EOFs is to find an empirical basis to represent data in a data driven way and it is often used to study spatial patterns of oscillations of various phenomena. In the context of tsunami science, they have been applied to study resonance modes of particular topobathymetric features, such as bays (Tolкова and Power, 2011) or islands (Bellotti and A. Romano, 2017), and to the optimal placement of measuring sensors (Mulia et al., 2017).

Applications of the technique to a one-dimensional time series aim to find a basis of a given length M with which we can decompose M -long segments of the time series. The physical process represented by the time series can then be described by the coefficients of the decomposition in each segment. To find such a basis, Tolкова (2010) presents the following procedure. Starting from a long, continuous and uniformly sampled time series $\eta(t)$, we extract from it N non overlapping segments of length M . We subtract from each segment its own average and we arrange them as columns of a matrix, which can be expressed as

$$C_{ij} = \sum_{k=1}^N [\zeta(q_k + i - 1) - a_k] [\zeta(q_k + j - 1) - a_k] \quad (2.6)$$

where q_k is the index where the k -th fragment starts and a_k is the average of the k -th fragment. EOFs are then computed as the eigenvectors e_i of the matrix

$$\tilde{C}_{ij} = C_{ij} + C_{M+1-i, M+1-j} \quad (2.7)$$

A signal of length M can be represented by projecting it onto the set of vectors e_i and it is shown by Tolкова (2009) that few components are needed to accurately represent experimental tide measurements. To extract the tidal component s from a signal η , we use the formula

$$s = EE^T \eta \quad (2.8)$$

where E is the matrix whose columns are the vectors e_i . This formula acts essentially as a low-pass filter, removing characteristics of the signal that are not captured by the basis.

The works presented by Tolкова (2009) and Tolкова (2010) discuss the application of this empirical basis to the removal and prediction of tides. Here, we are interested in the former problem. Some considerations about the properties of the basis and the decomposition are interesting for tsunami detection applications and we report them in the following. For the theoretical details, we refer the reader to the already mentioned works.

For the segments, Tolкова (2010) uses basis vectors of length equal to one lunar (or tidal) day (24 h 50.4 min). For a sampling time of 15 s, common in DART buoys and used in the applications in the next chapters, this translates into $M = 5940$. This length is the shortest that allows to capture the quasi sinusoidal components related to diurnal and semidiurnal tides, which are the dominant component of OBPB records. In this work, we adopt $M = 5940$ to detide signals in real time. We note that this choice is application specific and may not be appropriate for other uses. For example, Tolкова (2009) uses three lunar day EOFs to apply the technique to short term tide amplitude prediction.

Another characteristic is that e_i tends to reproduce the periods of the diurnal and semidiurnal tides, which are more or less universal and are thus similar between different locations and instruments. In fact, tides in different locations differ mostly in their fine structure and signals of the chosen length are not long enough to resolve it. This is used to our advantage for tsunami detection, since, as it is shown by Tolкова (2010), a basis computed in a location may be used to decompose a time series from a different location or instrument with no reduction in accuracy. This fact leads us to one of the most important aspects of the technique: once EOFs have been computed from a time series, they can be used for any deep-ocean data acquired at different locations.

At last, we point out that the optimal number of basis vectors to use can be determined empirically by checking how accurate the tides computed from eq. (2.8) are. Tolкова (2010) shows that 7 components are enough to reduce the error (computed as the root mean square) to 3 mm, while traditional harmonic analysis (Pawlowicz et al., 2002) requires tens of sinusoidal components.

The application of the technique for real time tsunami detection is as follows:

- the last lunar day of data is stored;
- the signal's average is removed from the data;
- tides are extracted with eq. (2.8) and subtracted from the signal;

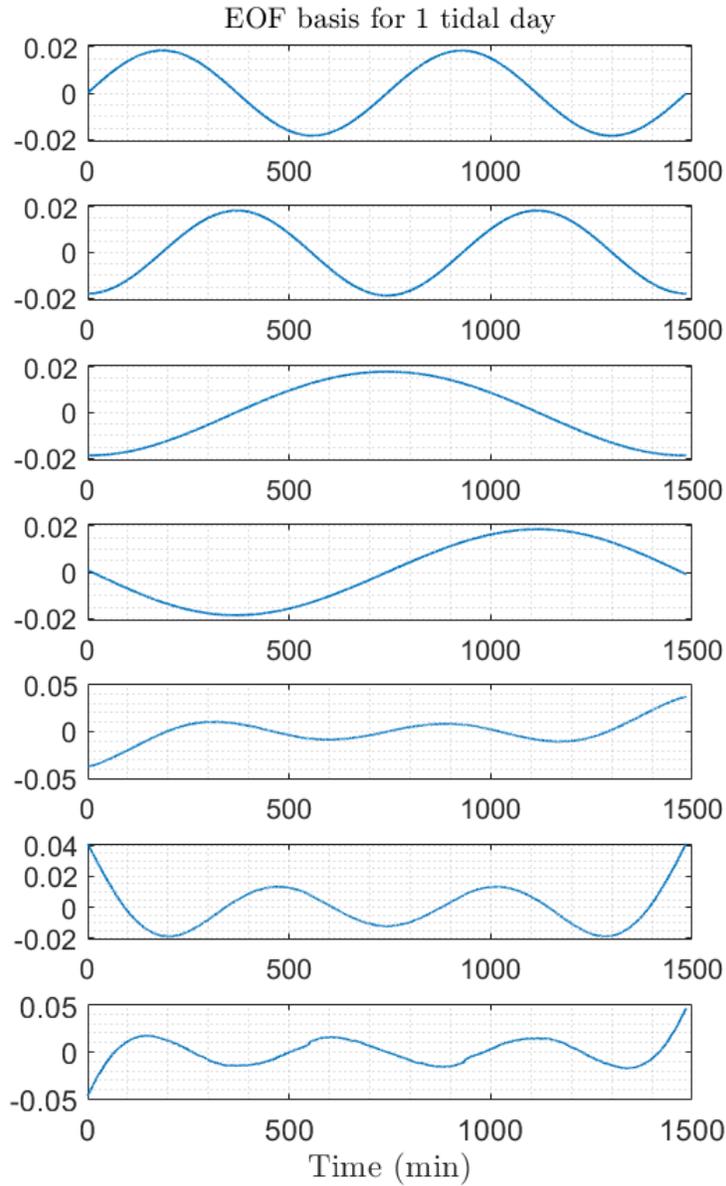


Figure 2.2: Example of Empirical Orthogonal Functions for one lunar day extracted with the technique proposed by Tolikova (2010), computed from the deployment between 06/06/2018 and 08/06/2022 of DART 46414. The vertical axis is expressed in arbitrary unit, since the technique is invariant under scaling of the basis vectors.

- a detection is triggered if the last point of the residual differs in absolute value more than a chosen threshold T .

At last, we note that the universality property previously mentioned is valid only for the longer period components, but not for the high frequency ones. Thus, the same filtering method cannot be used to filter neither random noise nor seismic shaking, as is the case for Mofjeld's algorithm.

An example of such a basis is shown in Fig. 2.2. The basis has been computed from NOAA data for DART 46414, located southeast of Chirikov Island in the Gulf of Alaska, using data from the period between 06/06/2018 and 08/06/2022. This time series is ideal to compute an EOF basis, since it presents no holes nor abrupt discontinuities. It can be noted that these basis vectors are symmetrical with respect to their central point, due to the symmetrization of the matrix in (2.7). As expected, the dominant period of the EOFs are around 1/2, 1 and 2 cycles per period, i.e. the periods of the dominant tidal oscillations.

2.3 Tsunami Detection Algorithm

Tsunami Detection Algorithm (TDA) has been introduced by Chierici et al. (2017). its core ideas are (a) adapting to operational real-time context the operations that are normally used in post processing of tsunami waveforms and (b) using a modular structure that can be modified according to specific needs. TDA is structured as a cascade of different operations: tide removal, despiking, bandpass filter, then check for detection.

Tide removal is performed by computing tides using a harmonic model (Pawlowicz et al., 2002):

$$f(t) = a_0 + \sum_{i=1}^{i=N} a_i \cos(\omega_i t + \phi_i) \quad (2.9)$$

where ω_i and ϕ_i are frequencies and phases of each component, and a_i are the tidal coefficients. To apply the technique in real-time, tidal coefficients are precomputed and, in an operational context, regularly updated.

In this work, analysis on real data are data-driven as much as possible, in order to reduce the influence of models' assumptions. For this reason, we choose to compute tidal coefficients by harmonic fit over available data at the instrument. To this aim, we use the software UTide by Codiga (2011). In each application presented in the following chapters, we select a long but variable amount of data preceding the segment of interest, which we decimate, taking

one point every 50 or 100, and then use it to compute tidal coefficients. Whenever enough data are available, removing tides computed using a tidal model results in residuals of amplitude within 4 cm. OBPGs in the DART network are regularly resurfaced for maintenance and downloading raw data, and new instruments are added to the network continuously. Due to these two factors, there are many cases where no enough preceding data are available for UTide to converge to a solution. Whenever available, between 2 and 12 months of data are used. If few data are available, it may also be the case that the algorithm converge to a solution, but this has a low accuracy. In these cases, the residual after tide removal still have amplitude of several centimeters which might not be eliminated by the subsequent filtering procedure. More details about these effects are given in Chapter 4 and 5, where applications to background signals and tsunami records respectively are given.

For the filtering module, Chierici et al. (2017) choose a Finite Impulse Response (FIR) bandpass filter, producing a filtered signal y computed as

$$y_n = \sum_{i=-N}^{i=+N} c_i f_{n-i} \quad (2.10)$$

where f_n is the signal we are filtering and N is the order of the filter. c_i is the i -th coefficient of the filter and it depends on the chosen type of filter (FIR in this case), the order N and the cut-off frequencies. These coefficients may be computed numerically with various different methods, depending on the desired characteristics and on the signal to which it is to be applied (Mousa, 2020). Often, the connection between digital filter design and convex optimization problems is employed (Boyd and Vandenberghe, 2004). In the present work, we use MATLAB's `designfilt` () function to compute the coefficients c_i given the order N , the sample rate and the cut-off frequencies. The default computational method, that consists in a constrained least square problem, is chosen. However, eq. (2.10) requires a symmetric set of samples around the n -th sample. In a real-time context, the n -th sample represents the last measurement we acquired (minus the forecast tide), thus to use eq. (2.10) in real-time, we apply a mirror boundary condition to the signal. The bandpass filter can then be written as

$$y_n = c_0 f_n + 2 \sum_{i=1}^N c_i f_{n-i} \quad (2.11)$$

By correctly choosing the frequency window, we are able to remove long term trends not removed by tidal analysis and high frequency components such as

random noise and seismic shaking. This latter case is of special interest, since the majority of tsunamis are of tectonic origin and this is the case for events analysed in the following chapters. Chierici et al. (2017) use a [2 min, 120 min] and a [4 min, 120 min] period windows, which should include the wide variety of observable tsunami periods (Mungov et al., 2013; Rabinovich, 1997; Rabinovich and Eblé, 2015). The order of the filter is set to 4000.

We note that the applications in the following chapters all have been through basic preprocessing, in order to analyse only the effects of tsunami detection algorithms. Thus, every signal has already been despiked and for this reason despiking methods, both in the context of TDA and in general, are not discussed in the present work.

2.4 The Fast Iterative Filtering and IMFogram techniques

In this section, we introduce a new tsunami detection technique which is based on the Fast Iterative Filtering and the IMFogram algorithms. Since these techniques have been developed quite recently and almost no application to tsunami science is present in the literature, some details of the technique and a few applications to tsunami time-frequency analysis are presented.

2.4.1 Analysis of nonlinear and nonstationary signals

The analysis of nonlinear and non-stationary time series can be thought as consisting of two distinct problems: the decomposition of the signal into simpler components and the individuation of the instantaneous frequency content (Cicone, 2019). Two among the most used techniques for decomposition are the Fast Fourier Transform (e.g. Brigham, 1988), which decomposes the signal as a superposition of sinusoids of different but fixed frequency, and the Wavelet Transform (e.g. Daubechies, 1992), which uses dilations and translations of a mother wavelet. Both these techniques are fundamentally linear. In fact, their applicability rests on the possibility of representing the original signal as a linear combination of a set of basis functions. Historically, the first completely data-driven technique for signal decomposition to be introduced was the Empirical Mode Decomposition (EMD) introduced by Huang et al. (1998). The technique decomposes a signal into Intrinsic Mode Functions (IMFs), which satisfy the following properties:

- the number of zero crossings and the number of relative extrema are equal or differ by one;
- the envelopes of relative maxima and relative minima are symmetric with respect to zero.

These properties ensure that each IMF has a well-defined local instantaneous frequency, which makes IMFs decompositions a good first step for time-frequency analysis. It is shown by Huang et al. (1998) that an intuitive understanding of IMFs may be as follows: while in Fourier analysis we decompose signals into components of the form $A \cos(\omega t + \phi)$, an IMF can be seen as a function in the form $A(t) \cos(\theta(t))$, also known as Amplitude Modulated - Frequency Modulated (AM - FM) function. Decompositions of a signal into IMFs can then be considered a data-driven generalization of Fourier analysis where amplitudes are allowed to vary and phase functions can evolve nonlinearly.

To decompose a signal $\{s_j\}_{j=1}^N$, we introduce the operator \mathcal{L} such that $(\mathcal{L}s)_j$ is a measure of local average of the signal at index j , and the operator \mathcal{S} defined such that $(\mathcal{S}s)_j = s_j - (\mathcal{L}s)_j$. Then, the first IMF I_1 is computed as

$$I_1 = \lim_{n \rightarrow \infty} \mathcal{S}^n s \quad (2.12)$$

and the following are defined iteratively

$$I_k = \lim_{n \rightarrow \infty} \mathcal{S}^n (s - I_1 - I_2 - \dots - I_{k-1}) \quad (2.13)$$

The result is an additive decomposition of the form

$$s = \sum_{k=1}^N I_k + r(t) \quad (2.14)$$

where $r(t)$ is the residual left from the last extraction. The process, called *sifting method* is carried out until we are left with a non-oscillatory $r(t)$ or a prescribed number of IMFs has been extracted.

In a numerical environment, the infinite recursion has to be approximated to a finite one. Usually, if I_k^j is the result of the j -th iteration in extracting the k -th IMF, a parameter δ is fixed and the iteration is carried out until

$$\frac{\|I_k^j - I_k^{j-1}\|}{\|I_k^{j-1}\|} < \delta \quad (2.15)$$

Different algorithms differ in the definition of \mathcal{L} , i.e. how the local average

is computed. In the case of the EMD method, the local average is computed as follows: local maxima and minima are interpolated with cubic splines to find an upper and lower envelope, respectively. Then, the local average ($\mathcal{L}s$) is given by the pointwise mean of the envelopes.

Despite having been successfully used in many applications in geophysics (Battista et al., 2007; Chen et al., 2012; Gómez and Velis, 2016), the EMD method has been shown to have some drawbacks: First of all, the technique is unstable to small added noise (Lin et al., 2009), which means that two signals differing only by a small amount may have very different decompositions. Mode mixing is also a common occurrence (Wu and Huang, 2009) for signals with intermittent components. In this case, separation of different time scales (i.e. different frequency content) into different IMF is not achieved. Due to these effects, the technique may result in a decomposition whose components have not real physical meaning (Huang et al., 1999).

Many EMD variants have been proposed to alleviate these problems. Usually, noise-assisted ensemble techniques have been proposed (Lang et al., 2020; Torres et al., 2011; Wu and Huang, 2009). One such example is the Ensemble Empirical Mode Decomposition (EEMD), introduced by Wu and Huang (2009). The EEMD technique works as follows. First, white noise is added to many different realization of the signal; the analysis is then carried out on the entire ensemble and the n -th IMF is obtained as the average of the n -th IMF of each decomposition in the ensemble.

2.4.2 The (Fast) Iterative Filtering algorithm

The Iterative Filtering (IF) technique, introduced by Lin et al. (2009), tackles the drawbacks of the EMD method by replacing the sifting procedure with a moving average. In particular, given a low pass filter $w(t)$, the moving average of a signal is computed as a convolution

$$(\mathcal{L}f)(x) = \int_{-l}^l f(x+t)w(t)dt \quad (2.16)$$

of the signal with a given *mask* $w(t)$ defined on a compact domain of length $2l$, where l is called mask length.

The main characteristic of the technique is that the filter length l varies during the decomposition. When extracting the k -th component, the mask is chosen such that

$$l_k = 2 \left\lfloor \xi \frac{N}{m} \right\rfloor \quad (2.17)$$

where N is the length of the analyzed sample, m is the number of relative extrema and $\lfloor \cdot \rfloor$ is the floor operator. Thus, the filter length is chosen proportionally to the density of relative maxima and minima, which can be interpreted as a measure of the high frequency content of the signal. The proportionality constant ξ usually assumes values in the interval $[1.1, 3.0]$ and can be chosen according to the characteristic of the signal (Cicone, 2019). Thus, we can direct the decomposition based on how much we want to separate various similar components.

Contrary to EMD-like method, IF can be rigorously analyzed from a theoretical point of view and many properties can be thus established. One such property is convergence, which can be guaranteed by properly choosing the mask. Let us consider a function that is symmetric, non negative, smooth, bounded and has compact support, which is known in IF literature as a *filter* (Cicone, 2020). Then, convergence can be guaranteed if we compute the masks as convolutions of a filter with itself (Cicone et al., 2016).

Another property of the IF method is that it is not subjected to mode mixing. In fact, applications to the *one or two frequencies* problem (Rilling and Flandrin, 2007) has shown that the technique is able to separate correctly sinusoidal components (Cicone et al., 2022, Cicone et al., 2024). Furthermore, an energy conservation principle can be established (Cicone et al., 2022), which shows that the frequency content of the original signal is the same as the total frequency content of its component. This is not true for EMD, which may present so-called *unwanted oscillations*, i.e. some IMFs may present non-physical oscillations that cancel each other when summed, but contribute to the frequency content of the decomposition.

Lastly, one further advantage is that the IF algorithm can be reformulated, assuming periodic boundary conditions, in a form using the Fast Fourier Transform (Cicone and Zhou, 2021). This variant, called Fast Iterative Filtering (FIF), results in great improvements in computational cost, with computing time often reduced by an order of magnitude or more (Cicone, 2020).

It has to be noted that both EMD- and IF-like techniques may introduce errors at the boundaries, as is the case even for any other signal processing techniques. In the case of IF, these errors decay exponentially moving inwards from the boundary (Cicone and Dell'Acqua, 2020). To reduce as possible the influence of possible boundary errors, the signal is extended before analysis following the procedure proposed by Stallone et al. (2020).

2.4.3 Time-Frequency Representation and the IMFogram Algorithm

The determination of the instantaneous frequency content of nonstationary signals is still a challenging problem, since there is no general consensus on the definitions of instantaneous amplitude and instantaneous frequency. In the case of nonstationary signals, it is common to adopt the Hilbert Transform (Huang et al., 1999). For a function f , its Hilbert Transform is defined as

$$(\mathcal{H}f)(x) = \frac{1}{\pi} \mathcal{P} \int_{-\infty}^{\infty} \frac{f(t)}{x-t} dt \quad (2.18)$$

where \mathcal{P} means that the integral has to be taken as a principal value. It can be shown that, for a real function $f(x)$, a complex function in the form $g(x) = f(x) + i(\mathcal{H}f)(x)$ is analytic and it can be rewritten as $g(x) = A(x) e^{i\phi(x)}$, where $A(x)$ is the instantaneous amplitude and $\phi(x)$ is the instantaneous phase, whose derivative can be interpreted as the instantaneous frequency.

Due to its property, each IMF will have only interwave modulations: the frequency may vary at each point, but will be unique at each point. For this reason, the use of the Hilbert Transform in conjunction with the EMD method has been proposed since the introduction of the latter (Huang et al., 1998). However, it has been noted (Cicone et al., 2016) that the application of Hilbert Transform may not be appropriate to retrieve local characteristics of a signal, since its definition (2.18) relies on global operators (i.e. an integration).

An alternative Time-Frequency representation suited for use with EMD- or IF-like techniques is the IMFogram (Barbe et al., 2020; Cicone et al., 2022), which relies on local definitions of frequency and amplitude. The frequency of a sinusoidal function can be computed as $f = \frac{1}{2d}$, where d is the distance between two successive zero crossings. In the case of signals with varying frequency, we can generalize this idea by computing, at each zero crossing z_i , the instantaneous frequency as $f_i = \frac{1}{2(z_{i+1} - z_i)}$. Once instantaneous frequencies at zero crossings have been found, the value at other times is found by interpolation. For the instantaneous amplitude, we take the absolute value of the signal s , individuate the relative maxima and let g be the function interpolating these maxima. The instantaneous amplitude is taken as the pointwise maximum between $|s|$ and g .

The IMFogram is usually represented as a matrix where the element A_{ij} is the instantaneous amplitude for the i -th frequency and at the j -th time. Usually, the results are averaged over an arbitrary long time window, after which we usually speak of local amplitude and local frequency. However, contrary to other

Time-Frequency representations, such as the Short-time Fourier Transform, the length of this window is completely arbitrary and its role is only to reduce the size of the output.

It can be shown that the elementwise square of the IMFogram matrix obtained from an IF decomposition with convergence parameter δ approaching zero, converges to the ordinary spectrogram. Even though the two representations have different meanings, since the IMFogram may represent the eventual non-stationarity of the signal, this property may be used to justify a posteriori the definitions introduced for local frequency and amplitude.

2.4.4 FIF-based Tsunami Detection Algorithm

The application of data-driven decomposition methods to real-time tsunami detection has been proposed by Wang and Satake (2021) and Wang et al. (2020b), based on the idea of extracting the tsunami component from an OBPG data by using the EEMD (Wu and Huang, 2009). In their work, they show that by decomposing 3 h of data with sampling time of 1 min, the result of the decomposition is

- one IMF, which they call IMF1, containing random noise and seismic shaking;
- 3 IMFs, whose sum they call IMF2, in the tsunami frequency range;
- 3 IMFs, whose sum they call IMF3, in the tidal frequency range.

The sum of the 3 middle IMFs is considered the “tsunami component” of the signal. The decomposition is performed each minute on the last 3 h of data acquired. If the last point of IMF2 is bigger than a given threshold, a detection is triggered. The technique has been tested on data from tsunamis of different size that affected Japan, i.e. the ones generated by the 1998 $M_w = 6.4$ Sanriku earthquake, by the 2011 $M_w = 9.0$ Tohoku earthquake and the 2016 $M_w = 7.4$ Fukushima earthquake. Furthermore, it has been tested on a month-long time series preceding the Fukushima event where no tsunami had been observed, to check for false detections.

However, the EEMD-based detection method as presented by Wang et al. (2020b) has a few drawbacks which limits its wider applicability. First, the fact that the tsunami is represented by the second, third and fourth IMFs is empirical. Even though these spectral properties always hold on the dataset used in the paper, it is not guaranteed that it holds true in general for any other instrument. We also note that, since each realization has different noise,

the number of IMFs in each decomposition may not be consistent among the ensemble (Torres et al., 2011). The number and spectral properties of IMFs is also dependent on the sampling time of the signal, especially in noisy environment where random oscillations are expected at all periods. Second, despite performing better than the classical EMD, the EEMD technique still presents cases of mode mixing and mode splitting quite frequently (Lang et al., 2020), which may invalidate the physical interpretations of different IMFs. Third, ensemble variants of EMD have a very high computational cost, since it needs to carry out the decomposition many times (usually in the order of hundreds). Numerical tests have shown that the EEMD technique is usually several order of magnitude more costly than IF and FIF (Cicone, 2020).

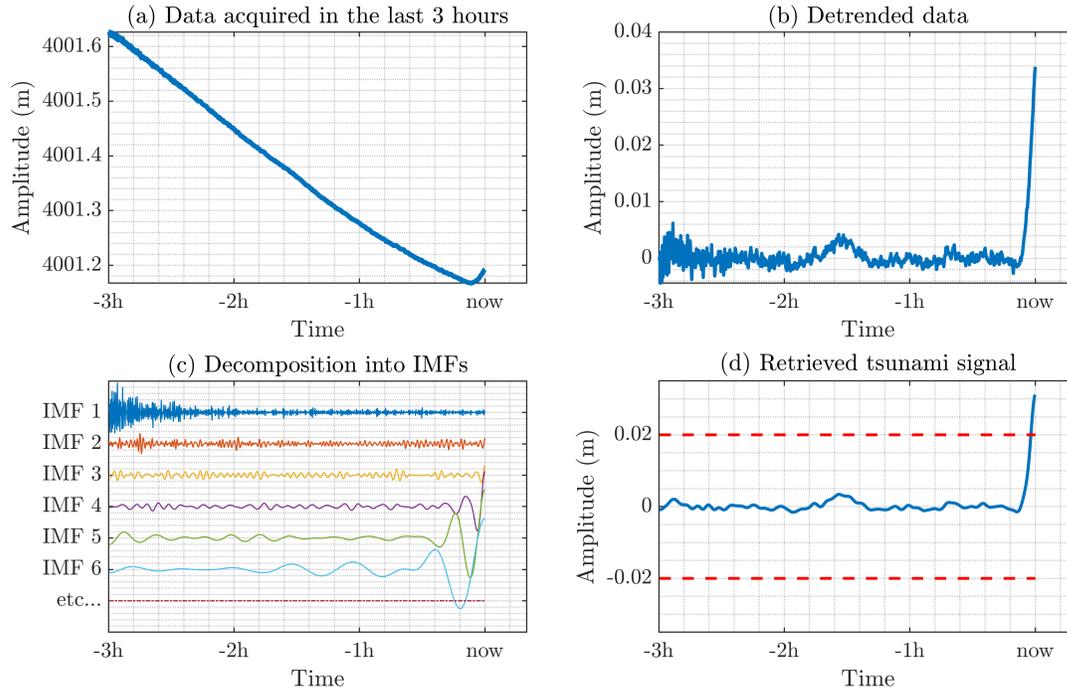


Figure 2.3: Example of FIF-based tsunami detection. At each time step: (a) we take the last 3 h of data, (b) we detrend them, (c) we decompose the detrended signal using FIF and (d) we sum the IMFs within a chosen frequency range. Data from DART 32413, during the 16/09/2015 Illapel tsunami. The generating earthquake occurred at 22:54:32 UTC, while “now” in the plot refers to 4 h43 min after origin time.

To overcome these downsides, we propose a new similar technique based on

the use of the FIF and IMFogram algorithms, which works as follows:

1. take the last 3 hours of data;
2. remove the long period trend with polynomial fit;
3. decompose the residual using FIF;
4. compute instantaneous frequency of each IMF with the IMFogram algorithm;
5. find the tsunami component by summing the IMFs with frequency within a chosen range;
6. compare the last point of the tsunami component with a chosen detection threshold.

The ability of IMFs to capture oscillations with variable frequency is one of the main motivation behind the introduction of EMD-like techniques. However, in the case of tsunami waves, that may sometimes be a disadvantage. In fact, in the case where a tsunami just arrived, as is the case in Fig. 2.3, the decomposition may produce a long term trend with variable slope instead of separating tide and tsunami. For this reason, we remove the tidal trend with a polynomial fit in step 2. In order to reduce the influence of tsunami waves on the fit, statistically robust regression (Street et al., 1988) methods are used and we choose to compute its weight as Cauchy functions. Furthermore, we will fit a degree 3 polynomial.

The FIF decomposition is then carried out using $\delta = 10^{-4}$ and $\xi = 2$. Choosing lower values for these parameters might result in a substantially larger amount of IMFs, which would not change the final result, since the selection is made based on frequency content, but it increases the computational time. Higher values for δ and ξ may not separate all the different time scales present in the detrended signal. Though, we also note that the algorithm is not very sensitive to parameter perturbations. To reduce the influence of boundary errors, we follow the procedure prescribed by Stallone et al. (2020) with a symmetric extension.

The selection of IMFs in step 5 allows us to use FIF and IMFogram as a band pass filter and the frequency window can be chosen based on the specific environment and applications. In the present work, we are mainly interested in detection of tectonic tsunamis, so we choose to consider in the tsunami band IMFs whose period in the last 30 min has an average value between 4 min and 2 h. Just as in the previous step, we note that the algorithm is not very sensible to

perturbations of this period interval and detection remains basically unchanged if the upper bound is changed to 3 h or is even removed, since tides are already been subtracted. To sum up, the chosen settings for the application of the method are

- a polynomial of degree 3 obtained with a statistically strong regression in step 2;
- FIF decomposition is applied with parameters $\delta = 10^{-4}$ and $\xi = 2$;
- components with periods between 4 min and 2 h are selected in step 5.

Once the tsunami waveform has been extracted, we check for detection with an amplitude based criterion as we do for all other techniques previously presented in the chapter. So, at each time step, a detection is triggered if the last point of the tsunami waveform is larger in absolute value than a chosen threshold.

At last, we note that the present FIF-based technique is still computationally more efficient than the EEMD-based method developed by Wang et al. (2020b). In fact, despite needing two more numerical procedures (detrending through fit and time-frequency computation from the IMFogram), the performance gain of FIF with respect to EEMD makes the detection method still much more efficient. This makes the technique easier to test on large datasets and power efficient for real-time use. The technique is also more easily applicable to newly developed or newly deployed instruments, since the choice of IMFs that represent the tsunami is based on time-frequency computation in real-time, instead of an a priori choice based on calibration on already available data.

2.5 Conclusions

In this chapter, four tsunami detection methods have been presented: Mofjeld's algorithm, EOF detiding, TDA and a new tsunami detection methods based on the Fast Iterative Filtering and IMFogram algorithms. For each technique, the necessary theoretical tools are briefly introduced, with reference to the relevant literature for details. In particular, a derivation of Mofjeld's algorithm is presented, which is, to the best of our knowledge, not available in the published tsunami literature. A digression on the FIF and IMFogram algorithm is also present, since this work represents the first extensive application of these technique in the field of tsunami science.

The techniques differ from each other in many aspects. Firstly, different amount of previous data are needed. The most demanding technique is TDA,

for which months of previous data for each instrument are necessary to compute tidal coefficients. In the case of EOF, basis vectors have to be computed from a long time series, however the same basis vectors can be used for every instrumental record. On the other hand, Mofjeld's algorithm and FIF-based detection need only the last few hours of data to work.

Another difference between techniques is the ability to filter high frequency disturbances. OBPG records show near-constant amplitude noise and sparsely distributed larger high frequency oscillations caused by seismic shaking. Both TDA and the FIF-based method have filtering capabilities built in, which allow to extract oscillations within the tsunami frequency band. This is not the case for Mofjeld's algorithm and EOF detiding, for which we should expect detection triggered by seismic waves. All of these properties will be investigated experimentally on real tsunami records in the following chapters.

One last aspect which differentiates the techniques is the computational costs of a single step. In the case of TDA and Mofjeld's algorithm, only basic arithmetic operations are needed for each time step, making them essentially instantaneous with respect to EOF detiding and FIF-based detection. In fact, the latter two require linear algebra operations at each step that, even if very optimized, are orders of magnitude more demanding than the former. These aspects will not be analysed further in this work. The reason is that all techniques in their simplest implementations (mostly through native MATLAB linear algebra functions) are already able to work faster than real time, i.e. computations for each step take less time than the sampling time of standard OBPGs. Any further optimization should take into account the specific hardware and software aspects of a particular instrument.

Chapter 3

Data-driven analysis of tsunami signals

In this chapter, two examples of tsunami signals from recent Mediterranean events are presented. Namely, the first example is from the 02/05/2020 Crete earthquake-generated tsunami, measured by the tide gauge located at Ierapetra. The second example regards the small tsunami observed after the first shock of the 06/02/2023 earthquake doublet occurred at the Turkey-Syria border. In both cases, data have been retrieved from the IOC Sea Level Monitoring Facility's website (<https://www.ioc-sealevelmonitoring.org/>). The main goal of the analysis is to show in practice how the FIF and IMFogram algorithms can be used to effectively analyse tsunami signals. More specifically, for the Crete 2020 tsunami we show that the analysis is able to produce the same results as traditional filters in a data driven way. In the case of the 2023 Turkey-Syria earthquake-generated event, it is shown that the analysis of the tide gauge records in the Turkish stations of Arsuz, Erdemli and Tasucu obtained through the decomposition into IMFs and their time-frequency analysis are consistent with what is reported in already published papers by other authors. The conclusion from both cases is that data-driven methods can be used to analyse tsunami signals giving results very similar to traditional techniques but with the added benefit of being completely data-driven and performing detiding and denoising in a single operation.

3.1 Tide gauge signal from the 02/05/2020 Crete Tsunami

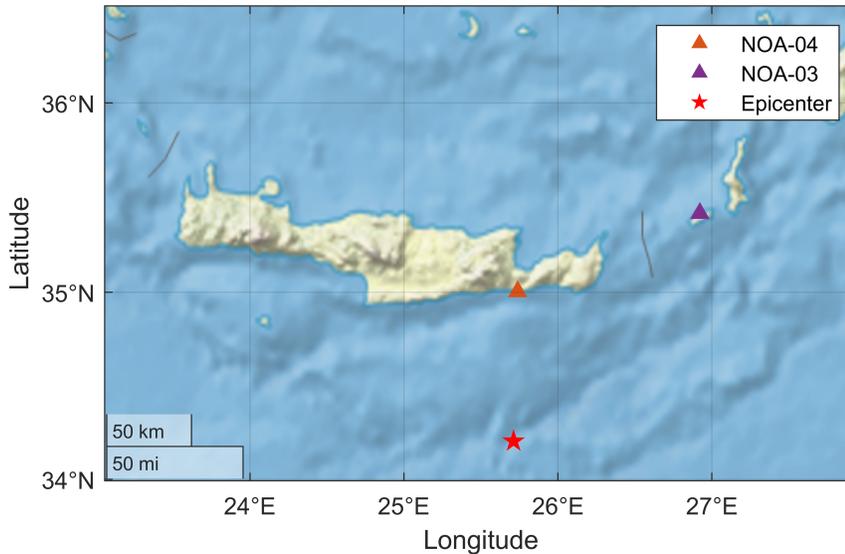


Figure 3.1: Epicenter of 02/05/2020 Crete earthquake and tide gauges located in Ierapetra (NOA-04) and Kasos (NOA-03).

A $M_w = 6.6$ earthquake occurred on 02/05/2020 at 12:51:06 UTC, about 80 km offshore from the island of Crete (Greece). Despite not being a particularly impactful event, it has nonetheless been the subject of multiple studies, both as a case study for improvements to Mediterranean Tsunami alert systems (Kalligeris et al., 2022; Papadopoulos et al., 2020; Wang et al., 2020a) and to determine the earthquake mechanism, since it was located in an area of interest from a seismotectonic point of view (Baglione et al., 2021; Heidarzadeh and Gusman, 2021). In particular, the event was located in an area where several earthquakes of similar size have been observed in the last century (Kalligeris et al., 2022) and it showed that the area might not be prepared for more serious events, both from an alert system and a population awareness point of view (see Papadopoulos et al., 2020 and Kalligeris et al., 2022 respectively). The tsunami was recorded at the NOA-03 and NOA-04 tide gauges, located in the island of Kasos and in the port of the city of Ierapetra, respectively, as shown in Fig. 3.1. Since it has been shown by Baglione et al. (2021) that the Kasos record does not

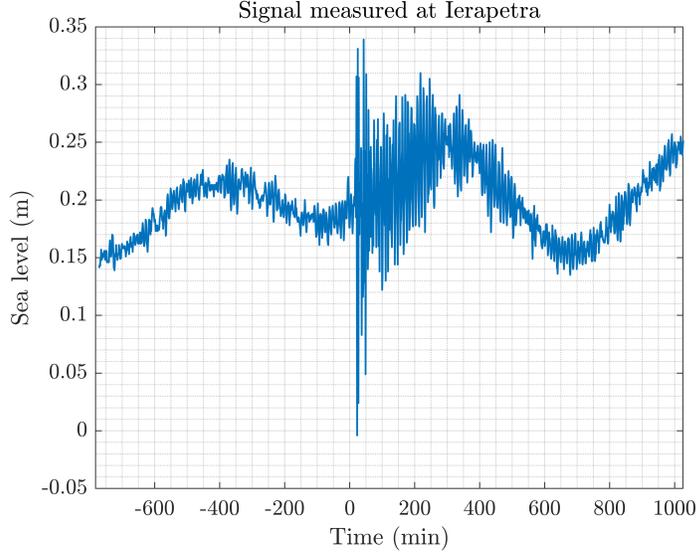


Figure 3.2: Sea level record measured at the Ierapetra tide gauge. Time is expressed in minutes since the earthquake origin time 12:51:06 UTC. Data from 00:00:00 02/05/2020 to 06:00:00 03/05/2020, with sampling time of 1 min.

contribute significantly in inversions for the source, we are going to analyse only the Ierapetra record, shown in Fig. 3.2. The signal has some typical characteristics of tsunami records, i.e. it is dominated by long term oscillations, ascribable to tides, to which high frequency oscillations are superimposed. These high frequency oscillations include random noise for the entirety of the record and a transient wavetrain starting at 16 min after earthquake origin time. These transient oscillations are made of two different phases: the first is made by a few oscillations reaching a peak-to-peak amplitude of 33 cm, while the second starts around 1 h after origin time and consists of slowly decaying oscillations reaching 15 cm peak-to-peak amplitude lasting a few hours. Many of these aspects can be better visualized from the decomposition into IMFs obtained using FIF, shown in Fig. 3.3, and the instantaneous frequencies and amplitudes for each IMFs in Fig. 3.4. From the instantaneous frequencies we can see that all the components are contained within a quite narrow frequency range, which allows us to separate them based on physical origin. The first component IMF 1, that has a median period of 4 min captures random noise and the first part of the tsunami wavetrain, consistently with other published studies (Heidarzadeh and Gusman, 2021). From the instantaneous amplitudes in Fig. 3.4, we can observe

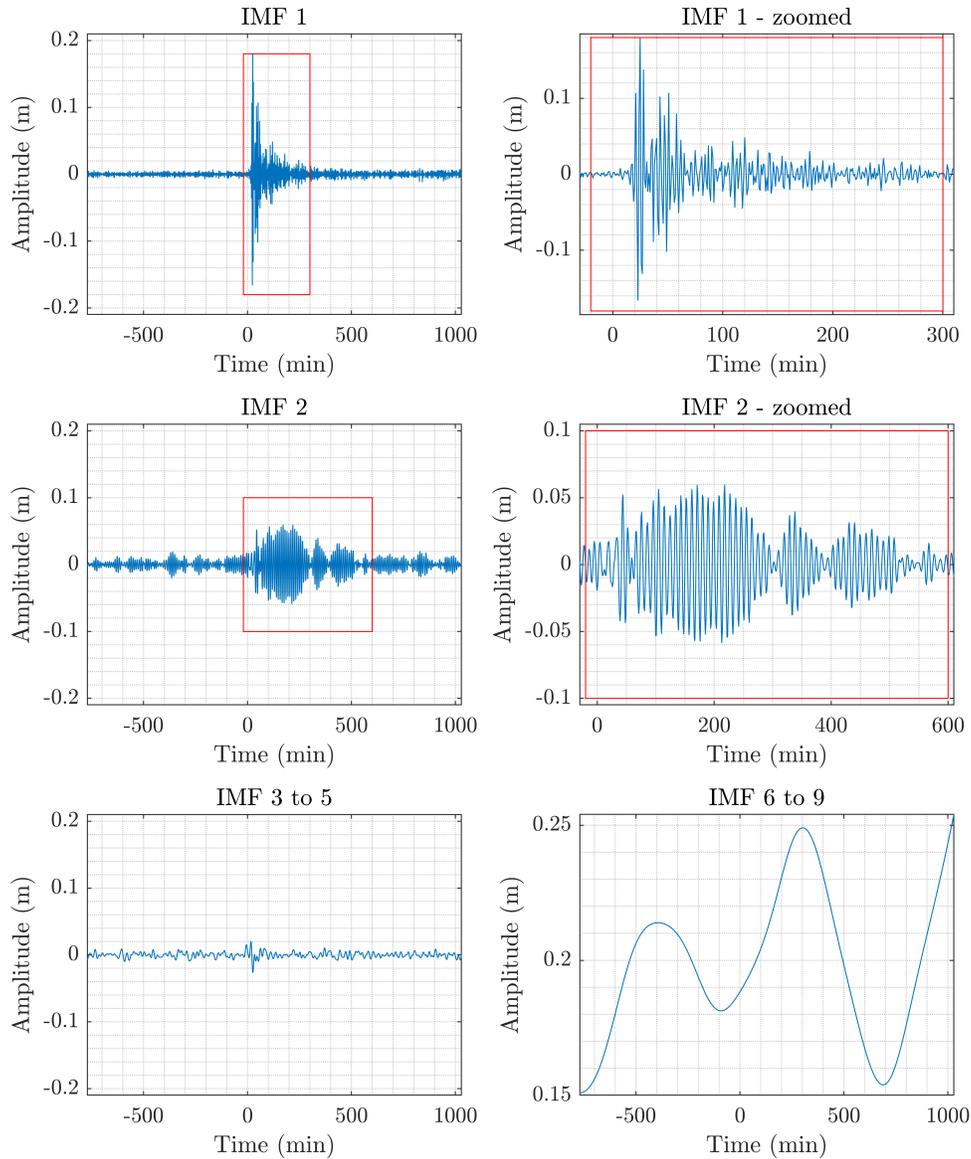


Figure 3.3: Intrinsic mode functions computed using the FIF technique for the 02/05/2020 tsunami signal recorded at Irapetra. Time is expressed in minutes since the earthquake origin time 12:51:06 UTC. IMFs 3 to 5 and 6 to 9 are summed together based on frequency considerations. Zoomed versions of IMF 1 and IMF 2 are also shown.

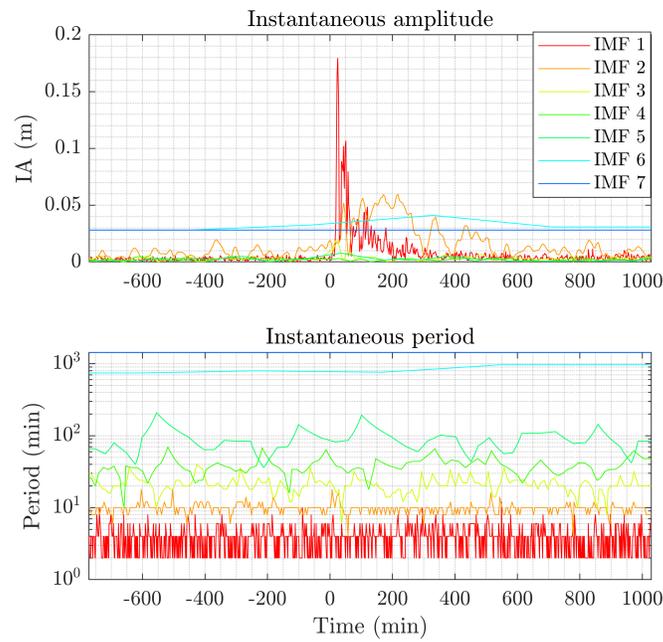


Figure 3.4: Instantaneous amplitude and instantaneous frequency for the IMFs in Fig. 3.3, relative to the 02/05/2020 tsunami signal recorded at Ierapetra. Time is expressed in minutes since the earthquake origin time 12:51:06 UTC.

two rapid high oscillations followed by a relatively fast decay. This component is the arrival of the direct wave from the tsunami and a first interaction with the Cretan coastline. The second component captures the persistent oscillations that follow the first wavetrain, with a median period of 10 min and an instantaneous amplitude of 6 cm that remain almost constant between 1 h and 4 h after origin time. These features of IMFs 1 and 2 are shown Fig. 3.3. While resonances are commonly observed in harbours after the arrival of a tsunami, this is probably not the case, since the corresponding periods are usually around tens of seconds. The oscillation captured by IMF 2 is probably the result of shelf resonance. These qualitative considerations are in line with what is presented by Kalligeris et al. (2022), where the natural oscillation frequencies of the region have been studied through numerical modelling. In that work, it is shown that the Ierapetra harbour has a fundamental oscillation period of 2.4 min, which explains why the direct wave arrival, with a period around 4 min, does not trigger resonance phenomena. On the other hand, Kalligeris et al. (2022) show that a period around 9.5 min probably corresponds to an oscillations mode of the bay to the east of the harbour, which can be observed also in the background spectra before the arrival of the tsunami. The properties of this oscillation mode correspond to what we observe in the IMF2 in Fig. 3.3, both in terms of period (around 10 min) and in the presence of clearly observable oscillations before the tsunami first arrival.

IMFs 6 to 9 have periods of several hours. Thus, they can be assumed to represent a decomposition of tides, as evident from their sum shown in Fig. 3.3. At last, the intermediate components 3 to 5 have periods of tens of minutes and very low amplitudes (within 3 mm each) and are probably long wave noise typical of coastal areas. We note that time-frequency results, especially for the transient wave train, are consistent with the analyses already presented in literature, e.g. by Heidarzadeh and Gusman (2021). We also note that period estimates from the IMFogram algorithm are more focused than what is reported by Heidarzadeh and Gusman (2021), where a similar analysis is performed through wavelet analysis. In fact, as shown by Cicone et al. (2022), the time-frequency distributions of transient signals obtained through the IMFogram algorithm are better localized in the time-frequency plane: Heidarzadeh and Gusman (2021) finds from wavelet dominant periods of 3-5 min and 5-10 min, while the IMFogram algorithm shows that the periods are located in narrower bands. To show the validity of using FIF for analysis, we compare the technique with EMD and traditional filters. We can see in Fig. 3.5 that the first two IMFs obtained through EMD have some similar characteristics to IMF1 and 2 obtained with

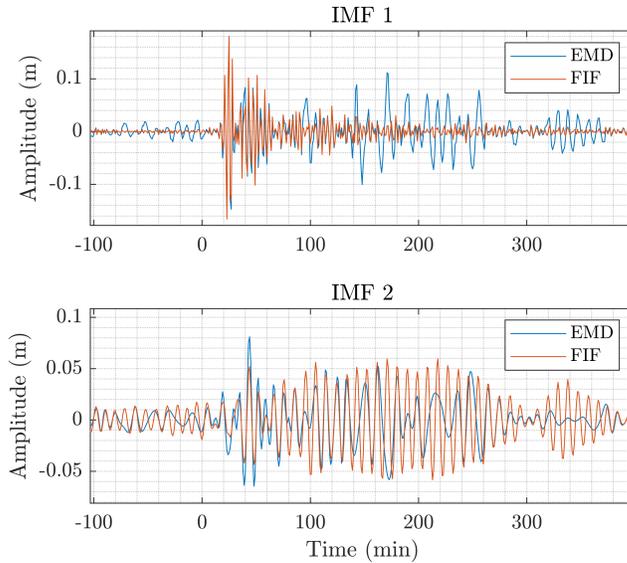


Figure 3.5: Comparison of EMD and FIF methods in computing the highest frequency IMFs for the signal in Fig. 3.2, relative to the 02/05/2020 tsunami signal recorded at Ierapetra.

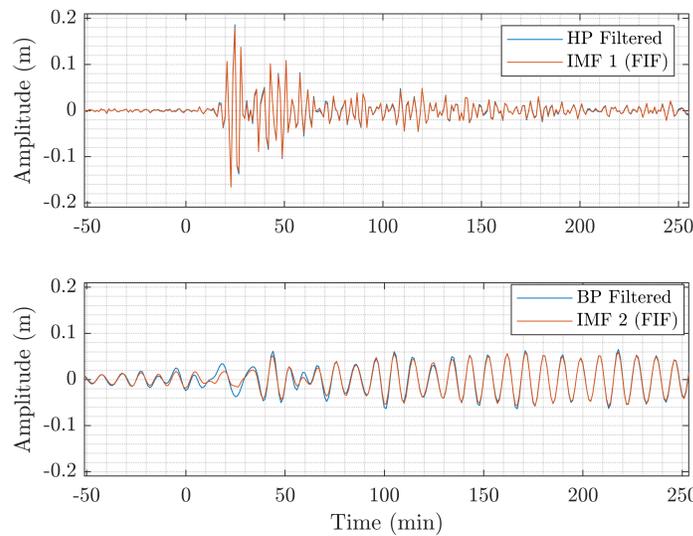


Figure 3.6: Comparison of IMFs computed through FIF and signals extracted with harmonic filters for the record in Fig. 3.2, relative to the 02/05/2020 tsunami signal recorded at Ierapetra. IMF 1 is compared with the high pass (HP) filtered signal with 1/400 Hz cutoff frequency. IMF 2 is compared with the bandpass filtered signal with cutoff frequencies at 1/2000 Hz and 1/400 Hz

FIF, but the separation between the two different kind of oscillations is not reproduced by EMD. In fact, while FIF separates contributions of different period, such as the first wavetrain from the following resonance mode, the two regimes are not clearly separated in EMD. In fact, we find in the first IMF both sparse larger oscillations between 2 h and 4 h and the direct arrival of the tsunami wave train. Viceversa, we find a much more pronounced oscillation in correspondence of the first arrival even in the second component. This exemplifies the fact that classical EMD is sometimes not able to separate different oscillation regimes.

Between EMD and FIF, we can confirm that FIF gives the expected results by comparing with what we obtain by classical filters. In Fig. 3.6, IMFs 1 and 2 are compared with a filtered version of the original signal. In particular, IMF 1 is compared with the result of a high-pass filter with cutoff frequency of 1/400 Hz and IMF 2 is compared with a band-pass filtered signal with cutoff frequencies 1/2000 Hz and 1/400 Hz. In both cases, there is very good agreement between the two waveforms, confirming that the two signals live in different frequency bands. It is important to notice that the physical separation of the 4 min and 10 min components is obtained by FIF in a totally data driven way, with no need to determine the frequency band to study in advance.

3.2 Tide gauge signals from the 06/02/2023 Eastern Mediterranean event

On 06/02/2023 the region between southern Turkey and northern Syria was hit by a devastating earthquake sequence that started with a $M_w = 7.8$ event at 01:17:34 UTC on the Eastern Anatolian Fault, and was followed by a $M_w = 7.5$ earthquake at 10:24:49 UTC (according to USGS). This doublet earthquake was also accompanied by a large number of aftershocks. The first very important aspect of this seismic sequence is the profound societal impact it had and continues to have, due to the tragic amount of victims (approximately 60 000) and the very wide spread coseismic effects that affected human settlements (Ozkula et al., 2023). On the other hand, the event is also very important from a seismological point of view. In fact, it has been shown that the main shocks had a very complex rupture process (Liu et al., 2023; Petersen et al., 2023), the study of which can help in understanding the complexity of source mechanisms.

The earthquakes were also accompanied by a variety of secondary effects, such as ionospheric disturbances (Haralambous et al., 2023), dynamic triggering of far away seismic structures with both aftershocks and seismic tremors

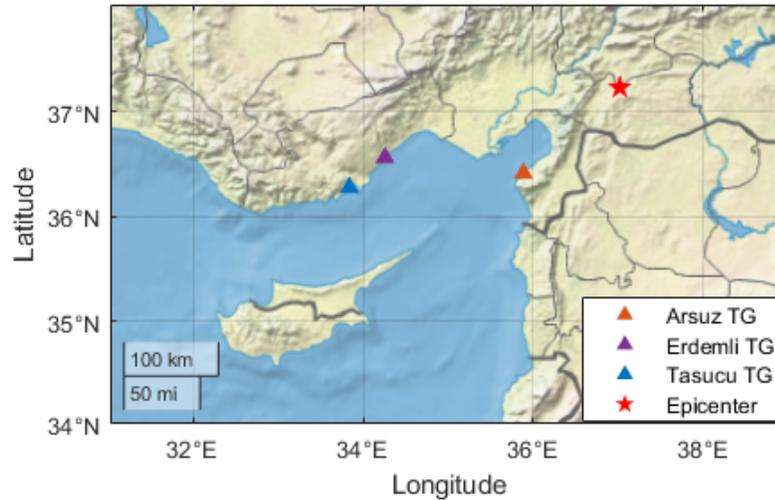


Figure 3.7: Epicenter of the first mainshock in the 06/02/2023 Turkey and Syria seismic sequence, and tide gauges located in Arsuz, Erdemli and Tasucu. Only the first shock of the sequence is shown, since it was the only one followed by a tsunami. In the legend, TG stands for Tide Gauge.

(Inbal, 2023; Inbal et al., 2023), widespread mass wasting events (Görüm et al., 2023), and tsunami waves observed at some tide gauges around the Turkish coasts, in particular in the tide gauges located in Arsuz, Erdemli and Tasucu, shown in Fig. 3.7. The tsunami itself has been and still is the subject of debate, since its records are consistent with a source contemporary with the first earthquake of the sequence, despite this having an epicenter located far inland and having a strike-slip mechanism. Even though the exact source has yet to be determined, the literature published so far suggests that the waves were generated by underwater mass wasting, though the position and number of slides is yet to be determined (Heidarzadeh et al., 2023; Hu et al., 2023). Furthermore, the situation was complicated by infragravity waves of atmospheric origin in the days following the tsunami (Medvedeva et al., 2023). At last, it should be also stressed out that, following the operating procedures of the NEAMTWS (Tsunami Warning System in the North-Eastern Atlantic, the Mediterranean and connected seas), the first earthquake of the sequence lead two Tsunami Service Providers to issue a basin wide alert. Given the characteristics of the tsunami and the location of the epicenter, questions about refining the methods employed by NEAMTWS have gained much attention again (Lorito et al., 2023).

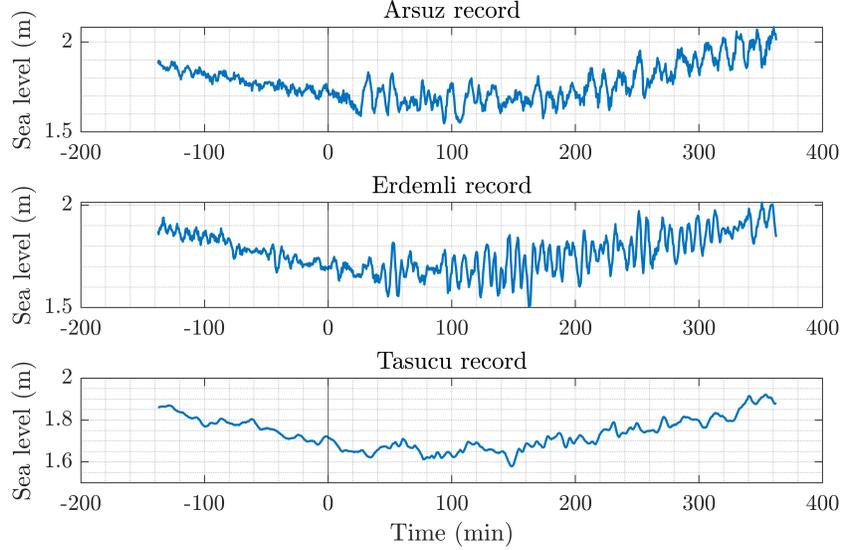


Figure 3.8: Raw data from 3 tide gauges located near the southern Turkish coast. Time is measured in minutes from origin time of the first earthquake of the sequence (01:17:34 UTC). Tsunami-related wavetrains are visible in the Arsuz and Erdemli records at 26 min and 46 min after origin time. The data have a 30 s sampling time.

Many of the mentioned characteristics of the tsunami may be inferred from a joint use of the available records. In particular, sea level oscillations compatible with a tsunami can be found by visual inspection in records of at least a couple of tide gauges, located at Arsuz, (36.416N, 35.885E) in the İskenderun bay, and Erdemli, (36.564N, 34.255E) on the southern Turkish coasts, as shown in Fig. 3.8. In order to perform a time-frequency analysis of the signals, we extract long portions of the tide gauges' data going from 12:00:00 04/02/2023 to 18:00:00 09/02/2023, in order to include the atmospheric perturbation that followed the tsunami in Arsuz. These signals are decomposed using FIF and we perform time-frequency analysis with the IMFogram algorithm. In the Arsuz record (Fig. 3.9a), both the tsunami wavetrain and the atmospheric disturbance starting a few hours later are evident. We can see that these different factors are separated from one another by the FIF decomposition. In fact, we can see from Fig. 3.9b that IMF 1, representing the highest frequency component, has a slow increase in the hours after the events, reaching the maximum variations around 2000 min (i.e. on 07/02/2023). The tsunami wavetrain is well captured by IMF

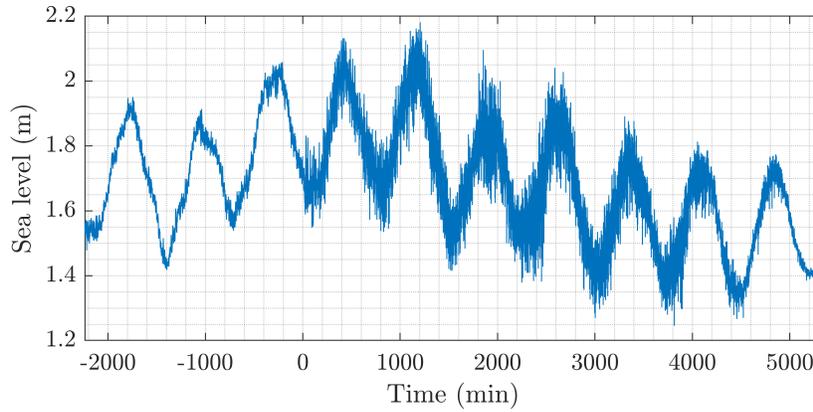
2 and 3, which arrives at 26 min, but they show different long term behaviour. In particular, IMF 2 shows oscillations lasting for the next days, probably due to the influence of the atmospheric disturbance as well, while IMF 3 decays to the pre-event noise level by the next day. All the considerations are confirmed by computing instantaneous amplitudes and frequencies with the IMFogram algorithm. As shown in Fig. 3.9c, IMF 1 has a median period of less than 2 min, capturing high frequency noise and showing the very large oscillations of the following days, with a maximum amplitude of 23 cm two days after origin time. Tsunami energy is concentrated in the components with median period of 7.6 min and 19 min. Furthermore, IMF 2 captures the maximum amplitude of 11 cm, observable by eye inspection from the raw data. IMF 4 has a median period of around 2.5 h and amplitude within 4 cm, compatible with long period noise we can expect in semiclosed basins such as the eastern Mediterranean. The remaining IMFs have frequencies in the tidal spectrum.

The analysis of the Erdemli signal gives similar results regarding the tsunami, while not showing the atmospheric oscillations observed in the Arsuz record. As shown by the decomposition in Fig. 3.10b and the amplitude-frequency computations in Fig. 3.10c, the first IMF has a median period of 2 min and an amplitude within 5 cm, which seems correlated with later tsunami oscillations. The tsunami wavetrain is again captured by IMF 2 and 3, with IMF 2 being the most energetic one. IMF 2 has a median period of 7.7 min with a maximum amplitude of 14.5 cm, while IMF 3 has a longer period around 27 min and a maximum amplitude of 6 cm; both decay within a day after the event. Both of these components lie in the relative vicinity of the periods observed for IMF 2 and 3 of the Arsuz record. In particular, the second components agree quite significantly in terms of period. On the other hand, the third components have periods that differ by 8 min. Given that their amplitude patterns is compatible with a tsunami, i.e. they increase in correspondence with IMF 2 and decay later, IMFs 3 may be related to the interaction of the tsunami with local geomorphology, though numerical simulations should be carried out to confirm it.

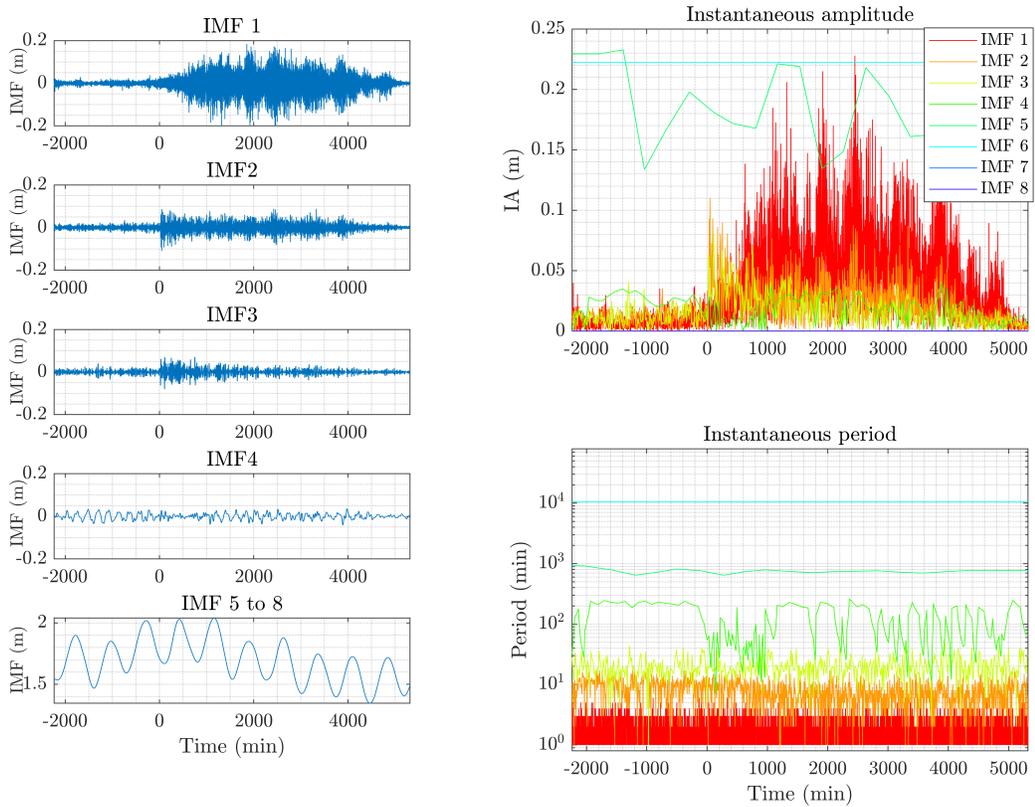
The results of these analyses can be compared easily with results in published works. Heidarzadeh et al. (2023) use wavelet analysis to find that the most energetic bands in the signals are 5.4-7 min and 17-21 min, while Hu et al. (2023) gives broader estimates as main frequency band between 2-10 min and 10-30 min, which include the main period observed by joint FIF-IMFogram analysis. A comparison with bandpass filtered signals is shown in Fig. 3.11. The period bands used are the same used by Hu et al. (2023) and the components

obtained by bandpass filtering are compared with the IMFs obtained using FIF. In particular, the extracted signal is the sum of the IMFs with median frequency within the frequency band. The waveforms agree quite well. On one hand, we confirm that the Erdemli signals has energy concentrated within the 2-10 min period band (Hu et al., 2023), since the waveforms for the two bands are similar. The opposite is true for the Arsuz record. In this case, we also observe that a lower agreement for the signals within the 2-10 min band, where the waveform obtained from FIF decomposition is smoother than the one obtained from harmonic filters. The explanation may be found in the analysis we just conducted: IMF 1, which captures noise with period close to sampling time, has a period around 2 min, i.e. it oscillates around the filter's cut-off frequency. Thus, the signal obtained by harmonic filtering *keeps in* some energy from the noise component. This does not happen using the FIF decomposition because it produces components each lying within a narrow and well-defined frequency range.

From the conclusions we just made about Arsuz and Erdemli records, we can analyse the tide gauge signal from the Tasucu instrument, located at (36.281N, 33.836E). In fact, IMF 2 of its decomposition, shown in Fig. 3.12b, shows an increase in amplitude with a maximum of 4.5 cm around 2.5 h after origin time, compatible with a late tsunami oscillations. The median period of IMF 2 is around 19 min, which lies within the tsunami frequency band observed in the previous data. However, recognizing the tsunami waveform by visual inspection is not possible in this case. The presence of the oscillations consistent with the tsunami in terms of frequency amplitude may be confirmed by the following time-frequency analysis. One of the reasons, besides the low amplitude of the signal, is that IMF 3, which has a median period of around 1 h has a comparable amplitude with what is observed for the tsunami. Remaining IMFs have periods of several hours and represent tidal oscillations.



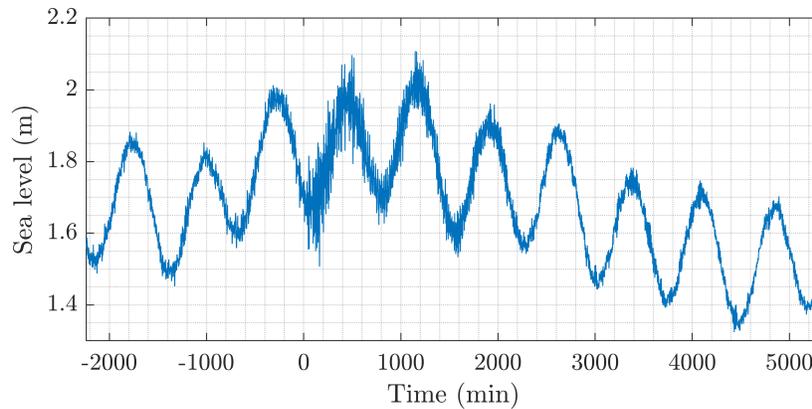
(a) Arsuz tide gauge record.



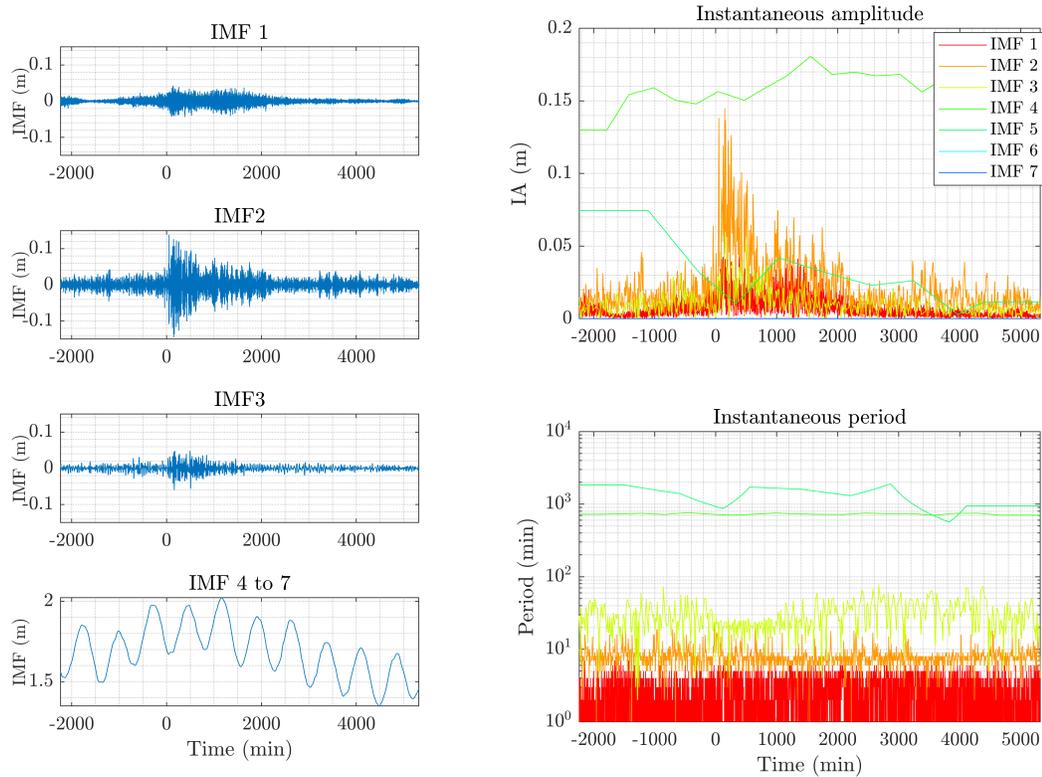
(b) IMFs obtained through the FIF technique

(c) IMFogram analysis

Figure 3.9: Joint FIF + IMFogram analysis of the Arsuz tide gauge record. (a) raw signal 12:00:00 04/02/2023 to 18:00:00 09/02/2023, (b) Intrinsic Mode Functions obtained with the Fast Iterative Filtering technique, (c) Instantaneous Amplitudes and Periods obtained through the IMFogram. Time is measured in minutes from origin time of the first earthquake of the sequence (01:17:34 UTC).



(a) Erdemli tide gauge record.



(b) IMFs obtained through the FIF technique

(c) IMFogram analysis

Figure 3.10: Joint FIF + IMFogram analysis of the Erdemli tide gauge record. (a) raw signal 12:00:00 04/02/2023 to 18:00:00 09/02/2023, (b) Intrinsic Mode Functions obtained with the Fast Iterative Filtering technique, (c) Instantaneous Amplitudes and Periods obtained through the IMFogram. Time is measured in minutes from origin time of the first earthquake of the sequence (01:17:34 UTC).

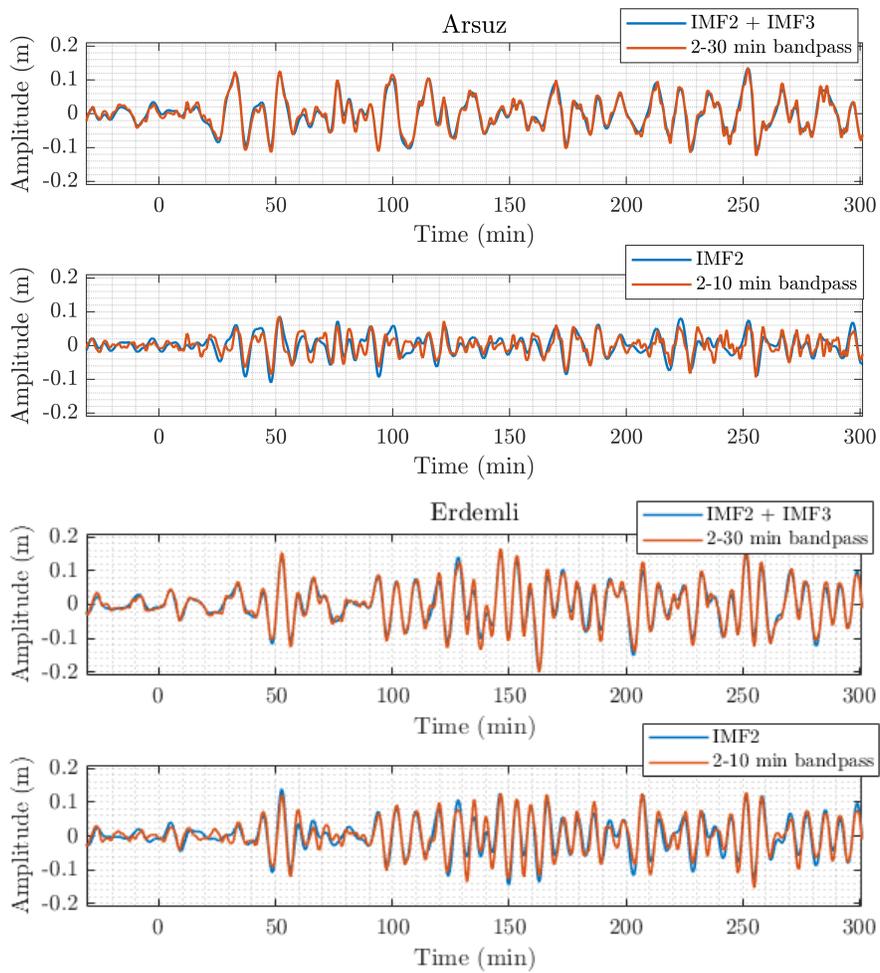
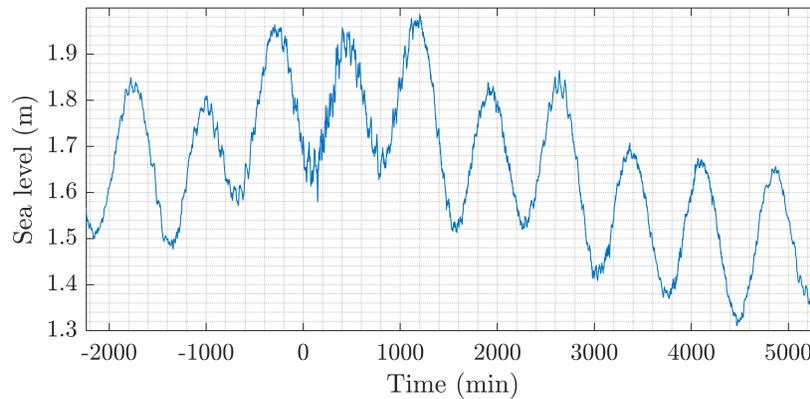
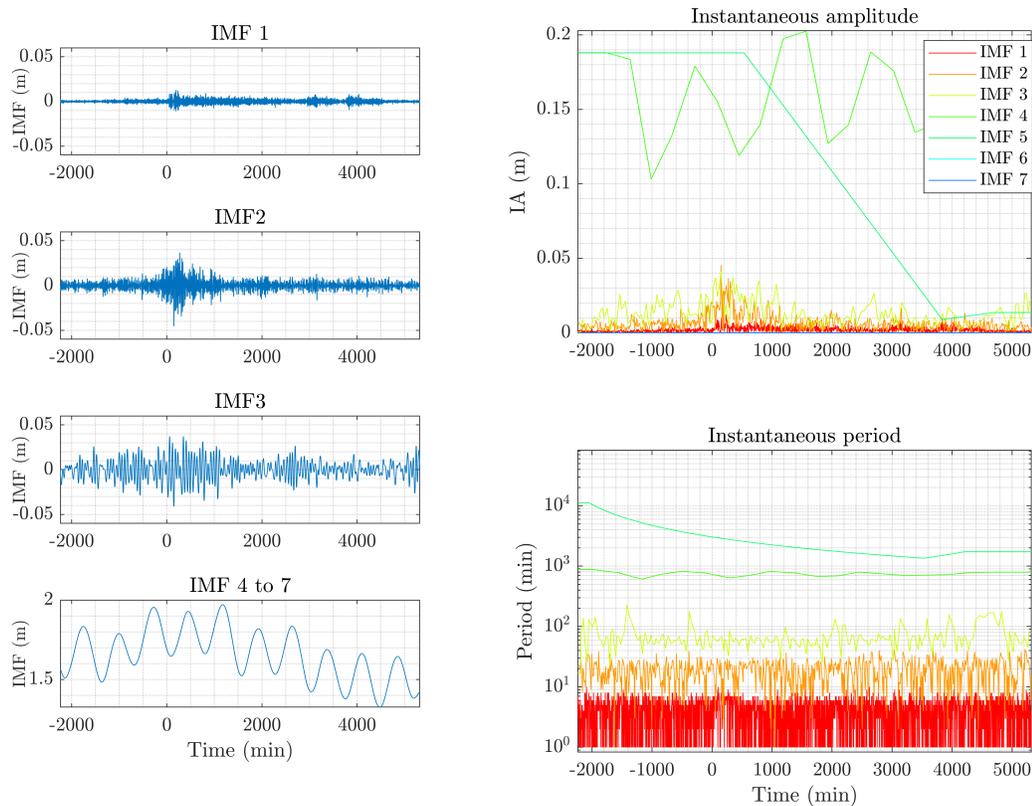


Figure 3.11: Comparison of tsunami waves extracted from a FIF+IMFogram analysis and with bandpass filters. The period windows are taken to reproduce the filtered signals by Hu et al. (2023).



(a) Tasucu tide gauge record.



(b) IMFs obtained through the FIF technique

(c) IMFogram analysis

Figure 3.12: Joint FIF + IMFogram analysis of the Tasucu tide gauge record. (a) raw signal 12:00:00 04/02/2023 to 18:00:00 09/02/2023, (b) Intrinsic Mode Functions obtained with the Fast Iterative Filtering technique, (c) Instantaneous Amplitudes and Periods obtained through the IMFogram. Time is measured in minutes from origin time of the first earthquake of the sequence (01:17:34 UTC).

3.3 Main takeaways

In this chapter, data driven techniques have been applied to tsunami records from the 02/05/2020 Crete tsunami and the 06/02/2023 small tsunami observed along the Turkish coasts. The signals have been decomposed using the Fast Iterative Filtering technique, which decomposes a signal into a variable number of oscillatory modes, called Intrinsic Mode Functions. Each IMF is locally bounded within a narrow frequency band that may change with time. After the decomposition, the IMFogram algorithm is used to obtain a time-frequency representation of the signal. Both techniques are designed to work in case of nonlinear and nonstationary time series.

In each case, the components have been analysed and we determined which ones can be traced back to the tsunami by checking if their instantaneous amplitude increase after earthquake origin time. From there, the main frequency content of the tsunami wave have been determined. Thanks to the analysis carried out on the Arsuz and Erdemli records, for the 06/02/2023 event, we have been able to determine that the tsunami has been recorded in the Tasucu record as well, where it is difficult to spot by visual inspection.

In general, the IMFogram analysis has shown that:

- a joint FIF+IMFogram analysis may be used to carry out routine tsunami data analysis with results that are in agreement with Fourier and wavelet transforms that are commonly used in the published literature;
- contrary to classical techniques, the decomposition and analysis is totally data-driven.

The last point is particularly significant in cases such as the Arsuz and Erdemli records. FIF splits high frequency noise, the tsunami signal and tides into different IMFs, allowing to denoise and detide the signal at the same time in a data driven way. Instantaneous frequency computation of each IMF with the IMFogram algorithm makes it possible to use these technique to bandpass filter a signal and we have shown how the results are in very good agreement with traditional harmonic filters. However, the data driven nature of the techniques make them more robust with respect to variations in the frequency band, due to the fact that each IMF is contained in a narrow frequency band at each given time.

The joint use of the FIF and IMFogram algorithms to analyse tsunami signals is thus able to give the same results as classical techniques but in an easier and more straightforward way, thanks to the data driven nature of the techniques.

For this reason, we believe that these techniques have the potential to enter the routine data analysis toolbox in tsunami science.

Chapter 4

Detection tests on background time series

In this chapter, applications of the four detection techniques to background signals, i.e. signals without seismic shaking or tsunami waves, are presented and discussed. First, NOAA DART® data are described in the details needed for the selection of time series to test. The reasons why we choose to DART® are multiple. Firstly, past data are made freely available by NOAA in their raw form, allowing us to test detection algorithm as if they were operating in real time. Secondly, NOAA's DART program represents the only global tsunami monitoring network, so they include instruments from very different environmental conditions. Lastly, the program has been active for decades, making a large amount of data available.

After the criteria for the choice of the testing signals are explained, the four detection techniques are all applied simulating real time operation. The results are analyzed and the four techniques are compared in terms of properties of their detection curves. The result of the analysis can then guide us in the applications to observed tsunami waveforms. For readability reasons, all the plots are moved in a dedicated appendix at the end of this chapter.

4.1 NOAA's DART® data

The Deep Ocean Assessment and Reporting of Tsunamis (DART) project is part of the U.S. National Tsunami Hazard Mitigation Program and constitutes the most widespread tsunami offshore monitoring and detection network in the

world, with instruments installed in the Northeastern Indian ocean, Northwestern Atlantic ocean, and in the Pacific Ocean. The first instruments developed by NOAA Pacific Marine Environmental Laboratory (PMEL) for the project were composed by a bottom pressure sensor and a surface buoy that allow for one-way transmission of sea level data. Four of these instruments were deployed in August 2000. A few years later, a second generation of instruments were developed, capable of two-way communication, to allow remote control. The third generation integrated the pressure gauge and the buoy into a single Easy-To-Deploy (ETD) system. Finally the most recent fourth generation added modules for automatic filtering of seismic shaking, allowing the deployment of instruments closer to subduction areas. Details about the technology and development of DART® instruments can be found at the website of the NOAA Center for Tsunami Research (<https://nctr.pmel.noaa.gov/Dart/>) and a detailed history of deep ocean tsunami measurement is presented by Rabinovich and Eblé (2015).

DART® instruments have different operating modes. Normally, the instruments are in “standard mode”: they acquire data with a sampling time of 15 s, but transmit to the surface buoy one measurement every 15 min. Instruments also come with detection software installed, which usually consists of an implementation of Mofjeld’s algorithm, as in eq. (2.4). Once a detection is triggered, the instruments enter in “event mode”, during which the instrument transmits data at a higher frequency. In particular, it transmits the full 15 s sampled data for 4 minutes, then it transmits 1 min averages for 3 hours, while also transmitting 1 min averages for the hour preceding the trigger.

DART® data are available via a web service hosted by NOAA (<https://www.ngdc.noaa.gov/thredds/catalog/enhancedCatalogWaterlevel.html>) in two forms:

- Unassessed Ocean Bottom Pressure, which are the raw data from each instrument;
- Quality-controlled, modeled, and de-tided Ocean Bottom Pressure, where post-processing operations have been applied (Mungov et al., 2013).

Both datasets are available at the highest available resolution, which may change between different instruments or deployments. In fact, despite the instruments acquiring and memorizing 4 pressure measurements per minute, these data can only be retrieved whenever the instrument is resurfaced for maintenance or substitution. In case of malfunctioning or data corruption, there may be missing

segments of data, which are filled by data transmitted to land during operation, wherever available.

4.2 Dataset creation

An important aspect of a tsunami detection technique is its behaviour whenever there are no transient signals, such as tsunamis or seismic shaking, which, in the case of OBPG records, corresponds to the case where only tides and random noise is present. Hereafter, we will refer to these signals as *background signals*. For these signals to be suitable for detection testing, a few criteria may be followed. The first is that the signal should be long enough to be tested for stability in the detection capabilities. In our cases, we choose to take signals that are one month long, in order to test for variations in the tidal regime, up to monthly constituents. Second, since we want to simulate real-time application of the techniques, we only select segments of data where full data sampled at 15 s are available. Third, we select signals that have no holes, jump discontinuities or instrumental spikes. Even though any of these artifacts can be treated by preprocessing the data, we want to reduce the influence of any preprocessing operation on the results of the analysis. Discussions on how to deal with these artifacts in real-time is outside the scope of the present analysis. Lastly, we want to test data from different geographical locations, to see if environmental factors influence the detection capabilities.

All the techniques are applied in the default version presented in Chapter 2. In particular, we use the weights in eq. (2.5) for Mofjeld’s algorithm, the basis shown in Fig. 2.2 for EOF detiding, and tidal coefficients computed with UTide and a FIR bandpass filter on the [4 min, 120 min] band for TDA. The use of precomputed tidal coefficients in the latter imposes another requirement: the signal needs to be part of long enough deployment to compute tides precisely. In each case, around 9 months of previous data have been used to compute tidal coefficients. The FIF-based technique is applied exactly as explained in Chapter 2. The basic characteristics of the selected signals, named from now on with letters A to E, are presented in Tab. 4.1, where we list instrument, location, time period and maximum tidal range within that period for each one. The signals are shown in Fig. 4.3, 4.7, 4.11, 4.15 and 4.19, respectively.

Case	DART	Location	Period	Tidal Range
A	46414	SE of Chirikov Island (Alaska, USA)	01/08/2019 - 31/08/2019	337.25 cm
B	52402	ESE of Saipan (Northern Mariana Islands, USA)	01/06/2016 - 30/06/2016	82.04 cm
C	32413	WNW of Lima (Perú)	01/01/2019 - 31/01/2019	129.62 cm
D	51407	W of Kailua-Kona (Hawaii, USA)	15/04/2022 - 14/05/2022	88.70 cm
E	21413	ESE of Tokyo (Japan)	01/06/2021 - 30/06/2021	86.02 cm

Table 4.1: Instrument, location, period and maximum tidal range for each background signal selected for testing.

4.3 Detection curves

All the techniques presented in this work are amplitude discriminating detection methods (Chierici et al., 2017), which means that a waveform in the tsunami frequency is extracted from the most recently acquired data and a detection is triggered if the last point of this component is bigger in absolute value than a given detection amplitude threshold T . In the case of Mofjeld’s algorithm, the tsunami component is extracted by subtracting the forecast trend from the raw data, while EOF detiding computes it by subtracting from data the tides found by projection on the empirical basis. In both cases, the techniques act as highpass filters. On the other hand, TDA and the FIF-based method act as bandpass filters, since they not only remove long term trends (with tidal prediction and polynomial fitting, respectively), but they also filter data in chosen frequency bands. Since the main observable for detection application is given by the last point of the processed data, the properties of the techniques can be studied by looking at the properties of the time series of these last points. Here, we define the time series of these last points as the *detection curve* of the signal. The detection is thus triggered once the detection curve has an amplitude bigger than T .

Ideally, a detection curve should be different from zero whenever a tsunami is passing at the instrument location and zero elsewhere. In reality, detection curves contain various contributions even for background signals, whose origin depends on the applied technique. The reason why detection curves can show

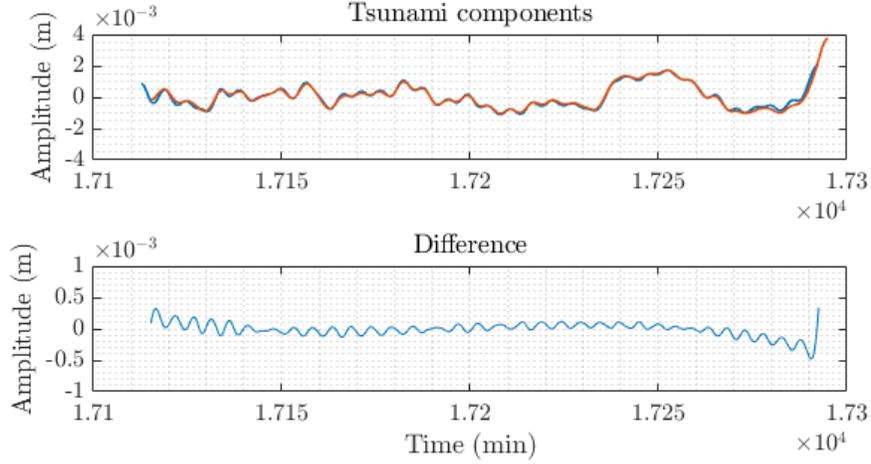


Figure 4.1: Inconsistency of tsunami components at different time steps for the FIF-based detection method. Data from the time series relative to background case A (see Tab. 4.1). Top panel: the blue curve is the tsunami component computed at step $n = 69170$, i.e. approximately 12 days after the beginning of case A; the orange curve is the tsunami component computed 10 steps later. Bottom panel: difference between the two tsunami components where the domains overlap.

such patterns is that the extracted tsunami component at each time step is not necessarily *consistent* with what is found at a different time step. To illustrate this fact, we can consider the detection steps at two successive time steps t_1 and t_2 , for which the presented techniques will produce tsunami components which largely overlap. Now, we consider a time t_0 that lies within this overlap. The aforementioned lack of consistency means the values at t_0 of the two tsunami components are not necessarily equal. Thus, we can expect oscillations, both of high and low frequency, to be present in the detection curves, even if the techniques we use include filters. We note that this inconsistency can mean also that the tsunami shape visible in a detection curve does not correspond to the real wave observed at that position. An example of this is presented in Fig. 4.1, where two tsunami components, obtained using the FIF-based detection technique, computed at a distance of 150 s. i.e. 10 steps at a 15 s time step, are compared. For any time included in the domain of both curves, their amplitude is not exactly the same. The amplitude of each point in detection curves will thus include the contributions not filtered by the technique, plus a small perturbation caused by this inconsistency. This is the reason why detection curves include

high frequency noise even for techniques that include a high frequency filter, such as TDA and the FIF-based method.

Mofjeld's algorithm is only capable to remove long term trends, so it is able to remove tides and long term pressure variations of meteoceanographical origin, but not random noise, which will be then included in the detection curve. Detection curves for EOF detiding and TDA may also include long term oscillations caused by tidal residuals. In the former case, the reason is that tides found by projecting onto the empirical basis are not able to capture the tidal fine structure that is usually site specific and has periods of 6, 8 and 12 h. In the latter case, tide computation may not remove tides in their entirety and the bandpass filter may not remove whatever oscillation is left, since we are taking a boundary point after applying a mirror boundary condition. For the same reason, TDA's detection curves may also show some residual high frequency components. At last, the new FIF-based detection may contain some long term oscillations, due to possible deviations between the polynomial fit and the long term trend, and some high frequency components due to noise and possibly seismic shaking.

In the following section, we are going to look at the detection plots for background signals. The property of detection curves in the presence of tsunami or earthquake waves is treated in the next chapter.

4.4 Detection results for background signals

The four detection techniques are applied to simulating real-time operation on the signals presented in Tab. 4.1 and the detection curves for each case per each technique are reported in the appendix of this chapter (Fig. 4.5, 4.9, 4.13, 4.17, 4.21). Some differences between the techniques can be noticed by visual immediate visual inspection.

As expected, all techniques show some oscillations, though the main periods in each of them differ quite significantly. In particular, EOF detection curves show quite coherent oscillations with almost constant amplitude throughout the entirety of the curves for all 5 cases. TDA also has some similar oscillations, but also some very long term non oscillating trends. For example, in case B, the TDA detection curve oscillates around positive values, except for the period around day 15 where it temporarily decreases to zero. Case C and D have the opposite behaviour, with the curve having mostly negative values with some occasional variations in average value on the scale of days. The most evident long term trend is found in the curve for case E, where the signal oscillates

Case	DART	EOF	TDA	FIF
A	$[-0.59, 0.51]$	$[-1.39, 1.45]$	$[-2.16, 1.75]$	$[-0.92, 1.13]$
B	$[-0.40, 0.42]$	$[-1.67, 1.85]$	$[-0.63, 1.25]$	$[-0.63, 0.71]$
C	$[-0.54, 0.58]$	$[-1.72, 1.84]$	$[-1.30, 0.75]$	$[-0.74, 0.74]$
D	$[-0.51, 0.54]$	$[-1.94, 1.93]$	$[-1.27, 0.82]$	$[-0.94, 0.97]$
E	$[-0.44, 0.43]$	$[-1.71, 1.62]$	$[-0.69, 1.46]$	$[-0.64, 0.63]$

Table 4.2: Maximum and minimum value in centimeters recorded in each detection curve for each technique. DART refers to Mofjeld’s algorithm.

around zero up to day 17, after which there is an increase of more than 0.5 cm. The Fourier spectra of the detection curves (in Fig. 4.4, 4.8, 4.12, 4.16, 4.20) also show the presence of strong oscillations in TDA’s and EOF’s detection curves. EOF detections curves have strong spectral peaks in correspondence of the 8 h period, which is a typical period observed in tidal fine structure that the technique is not able to remove. The spectral peaks of TDA are more variable among the cases, but its energy is concentrated mainly around the periods of diurnal and semidiurnal tides, indicating that imprecise tidal modelling is the biggest issue in the detection curves. We can also note that TDA’s detection curves are the only with a significant portion of energy concentrated at the zero frequency limit, due to the whole-signal long trends they have. The Fourier spectra of both Mofjeld’s algorithm and FIF have peaks near the 12 h period, although they are much weaker than in other techniques. An exception to this general trends is given by case A, where TDA has much more coherent oscillations, with taller peaks in the Fourier spectra. This seems to indicate a relation between TDA characteristics and the maximum tidal range observed at the recording location, since DART 46414 is located in the Alaskan Gulf, where high amplitude tides are observed.

The presence of very long trends in a detection curve causes an asymmetry in the distribution of the residuals. Such an asymmetry can be a problem for tsunami detection. If we consider a tsunami wave superimposed to a background similar to case B, its detectability depends on its polarity. In particular, a leading crest tsunami would be easier to detect, since the data are already skewed towards positive values, while the detection of leading trough tsunamis would be more difficult. This factor can be very well visualized in two ways. The first is by comparing the asymmetry in the maximum range observed for each detection technique, reported in Tab. 4.2. From that, it is evident that TDA is the only technique where this kind of asymmetry is significant, with a difference between the maxima and the (absolute value of the) minima differ of

Case	DART	EOF	TDA	FIF
A	0.15cm	0.41cm	0.64cm	0.24cm
B	0.10cm	0.56cm	0.27cm	0.15cm
C	0.13cm	0.61cm	0.29cm	0.17cm
D	0.13cm	0.63cm	0.30cm	0.20cm
E	0.10cm	0.57cm	0.35cm	0.14cm

Table 4.3: Standard deviation in centimeters of each detection curve for each technique. DART refers to Mofjeld’s algorithm.

about 4 mm, while the same metric is an order of magnitude smaller for other techniques.

The second way to visualize the asymmetric distribution of values is from the histogram of amplitudes (in Fig. 4.6, 4.10, 4.14, 4.18, 4.22). Mofjeld’s algorithm produces the narrowest histograms, centered around 0, i.e. detection curves for background signals obtained by this technique are the ones that are statistically closer to the ideal behaviour. The FIF-based method has a very similar behaviour, with slightly thicker tails in cases A and D. EOF and TDA both deviates from this behaviour. The former has wider histograms with thicker tails in all the selected cases, with cases C and D having an almost bimodal distribution, indicating that the signal tends to oscillate coherently during its whole length. On the other hand, TDA produces detection curves that are narrower than EOF’s ones, but they have peaks that are significantly far from zero. The very long trend present in the second part of signal E results in a slightly bimodal distribution, where neither of the maxima are located at zero.

The general considerations on the distributions of values in the detection curves done previously are confirmed by comparing the standard deviations of detection curves, reported in Tab. 4.3. Standard deviation follows a general ordering rule: Mofjeld’s algorithm produces detection curves with the lowest standard deviation, just under FIF, then followed by TDA and EOF. Case A, where tidal residuals in TDA are particularly evident, is the only one where TDA’s detection curve has a greater standard deviation than the EOF’s curve.

The general conclusions we can draw from this analysis are:

- the precision in tide prediction is the fundamental factor in detection curve for EOF and TDA;
- Mofjeld’s algorithm and FIF-based detection perform much better in terms of *closeness* to 0 of their detection curves;
- the performance in tsunami detection can be dependent on the polarity

of the wave only using TDA, while other techniques have a symmetric distribution around 0.

One thing to stress again is that TDA results in this chapter represent optimal cases, since signals have been specifically chosen to have enough preceding data (around 8 or 9 months) to obtain an accurate tidal model. We already pointed out that TDA detection curves contain residual oscillations in the tsunami frequency band (see spectral plots in Fig. 4.4 to 4.20). In the case where fewer data are used, we may obtain less precise tidal models, which would result in larger amplitude detection curves even for background cases. To show this effect, TDA detection curves have been computed using tidal models obtained from time series of different length. In particular, coefficients for the harmonic tidal model in eq. (2.9) have been computed from time series ending two weeks before the starting point of the case study (in Tab. 4.1) and starting at different times before that. The result are shown in Fig. 4.2. For each case, the maximum, minimum and mean value of each detection curve is plotted. It is immediately evident that detection curves amplitude is strongly dependent on how many data we use to compute the tidal model. In particular, the amplitude tends to stabilize within ± 2 cm when at least 6 to 7 months of data are used. Fewer data not only result in larger oscillations, but also in more asymmetric distribution, as we can see from the signal averages, i.e. the red curves, in Fig. 4.2.

However, there are two more problems we may face. The first is that it may be difficult to satisfy all the requirements. For example, in cases B, C and E the detection curves with average value closer to zero, i.e. the most symmetrical, are not the ones with the narrower variability range. We recall that asymmetric detection curves for background signals mean that the detectability of a tsunami may depend not only on its amplitude but also on its polarity. Thus, we may not be able to choose a set of tidal coefficients that minimize detection curve residuals and polarity effects at the same time.

The second problem we can face with the computation of tidal models is that, while the variability ranges initially converges around zero, there may be cases where this trend breaks for large enough time series. In case A, we can notice that, after reaching a narrow and symmetrical distribution for around 8 months of data, then both the minimum and maximum of the detection curves increase with longer time series. In case C, we can appreciate a significant jump for time series length greater than 12 months. These effects may be related to the possible presence of very long term trends in OBPB records, which, if not taken into account, make tidal coefficients computation dependent on which sections of the time series we use. Given that these trends are usually studied

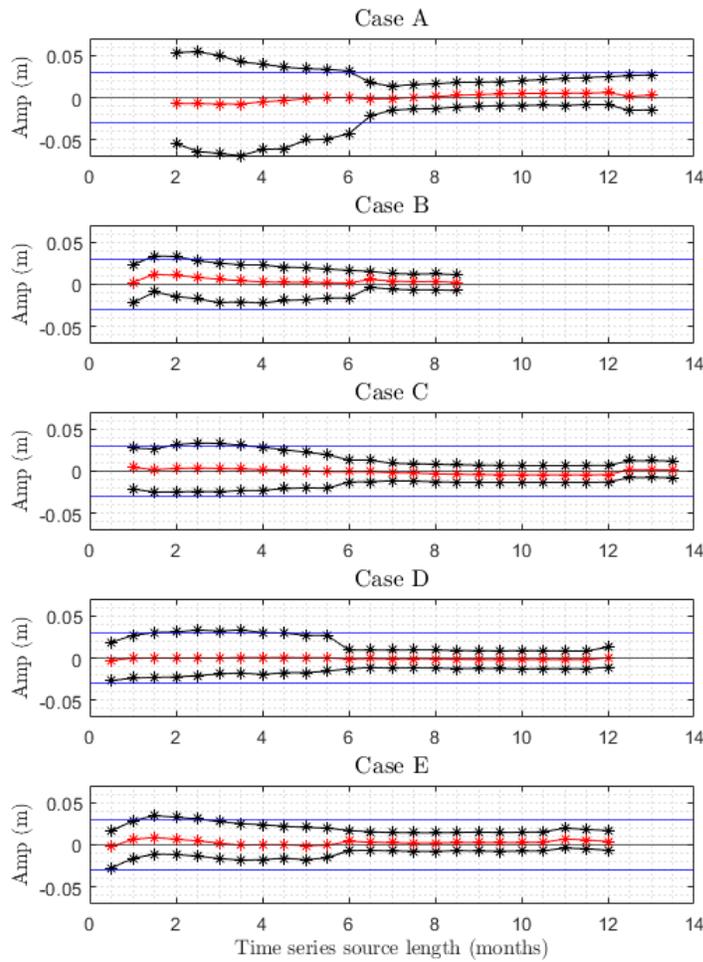


Figure 4.2: Variability range of TDA detection curves for different tidal models. For each case, the maximum and minimum (in black) and the average value (in red) of each detection curve corresponding to a different length of source data used to compute tidal coefficients. Horizontal blue lines corresponds to amplitudes of ± 3 cm. The smallest data length for which data are plotted in each case is the smallest for which UTide converges. The largest data length used correspond to the maximum amount of data available in that deployment.

and removed a posteriori (Mungov et al., 2013), no attempt to take them into account has been made here. To have background detection curves within 2 to 3 cm, we can simply propose the rule of thumb of using 7 to 9 months of previous data to compute tidal coefficients.

We can comment what the background analysis suggests us in terms of detection thresholds. The de facto standard threshold for amplitude discriminating techniques is $T = 3$ cm, by virtue of it being used in NOAA’s instruments (Mofjeld, 1997), but some studies, such as Chierici et al. (2017) and Wang et al. (2020b) consider also $T = 2$ cm. Given the analysis of background signals, we can take as a hard low limit a threshold of $T = 0.5$ cm, for which Mofjeld’s algorithm, which performed the best, would trigger detections in multiple cases. None of the techniques considered here trigger any detection for $T \geq 2.5$ cm, while for $T = 2$ cm we get false detection only with TDA in only one of the cases. Based on previous considerations, this may suggest that detection threshold should be chosen depending on tidal range, whenever a harmonic tidal model is used. At last, we note that the new FIF-based method gives no false detection even for $T = 1.5$ cm and would trigger a single false detection in case A for $T = 1$ cm. Obviously, the choice of the optimal threshold depends not only on false detection in absence of tsunamis, but also on the behaviour in presence of real events. Thus, we address the discussion in the next chapter.

4.5 Conclusions

In this chapter, the four tsunami detection methods have been tested on background signals, i.e. OBPG records when only tides and high frequency noise are present. In order to assess the stability of the findings, the five different time series have been chosen from different geographical areas and time periods, all taken from the Unassessed Ocean Bottom Pressure data collected by NOAA.

The main findings of this analysis are in the spectral and amplitude properties of the detection curve of each technique. In general, Mofjeld’s algorithm produces the detection curves *closer to zero*, in terms of maximum variability, spectral amplitudes and standard deviation. These metrics are only slightly higher for the FIF-based methods. For these techniques, we can consider a 2 cm as a candidate threshold to use to identify anomalous oscillations in the detection curves. The threshold may be brought down to 1.5 cm for the FIF-based method and 1 cm for Mofjeld’s algorithm, according to the presented tests. More rigorous criteria to choose a detection threshold will be discussed in the following chapter.

Both TDA and EOF detiding show much more pronounced oscillations, as evident from their Fourier spectra. Since the observed spectral peaks have periods that can be related to tidal periods, we can conclude that these techniques perform worse in terms of long term oscillation removal. However, the detection curves have amplitude almost always within 2 cm, which allows to use them for tsunami detection in an effective way.

At last, TDA is the only technique which presents detection curves with a significantly asymmetric amplitude distribution. This aspect, which is related to the difference between measured tides and computed tidal coefficients, may render the technique more or less effective depending on the polarity of the tsunami wave.

Appendix - Plots

In this appendix, plots for the background signals analysed and referenced in this chapter are presented. For each case we present

- the raw signal as measured by the sensor;
- the detection curves for each of the four techniques;
- the modulus of the Fast Fourier Transform of each detection curve;
- the histogram of amplitudes for each detection curve.

Each technique is mentioned by a shortened name or acronym. In particular, Mofjeld's algorithm is referred to as DART, detiding with Empirical Mode Decomposition as EOF, Tsunami Detection Algorithm as TDA and the new FIF-based technique as simply FIF.

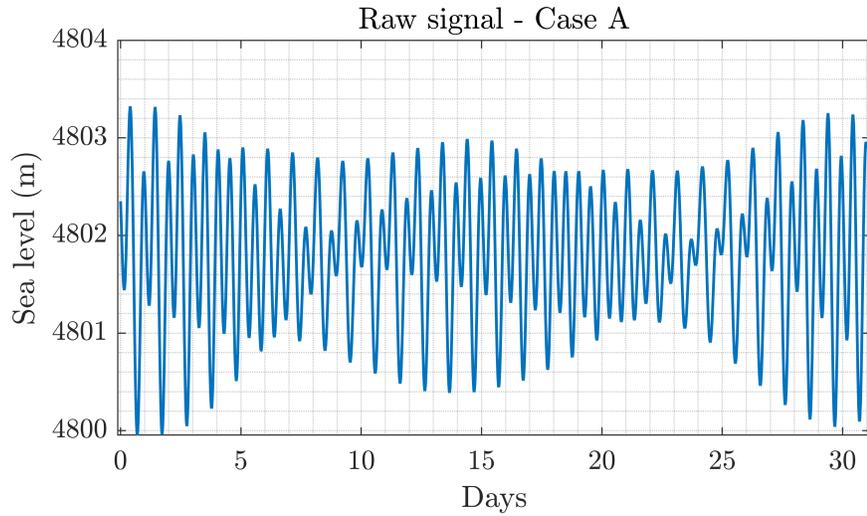


Figure 4.3: Raw data relative to case A in background analysis. Data from DART 46414, between 01 /08/2019 and 31/08/2019.

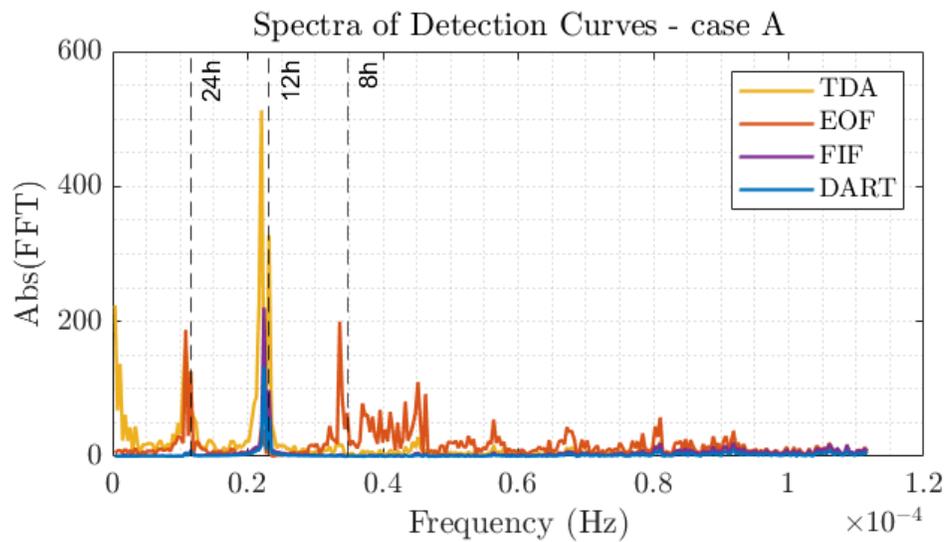


Figure 4.4: Absolute values of the FFT for detection curves obtained for background case A. DART refers to Mofjeld's algorithm.

Detection curves - Case A

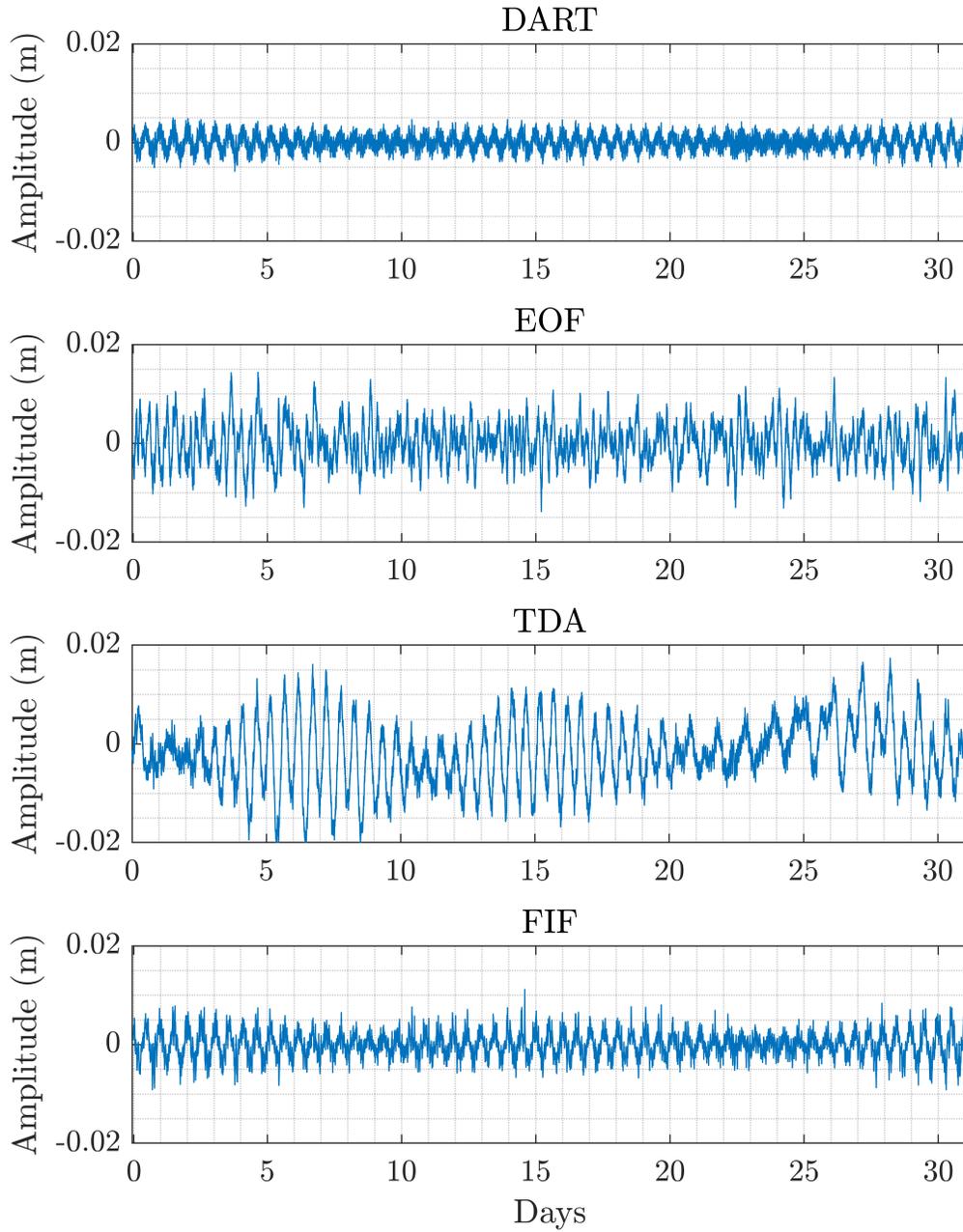


Figure 4.5: Detection curves for each technique for background case A. DART refers to Mofjeld's algorithm.

Histograms - Case A

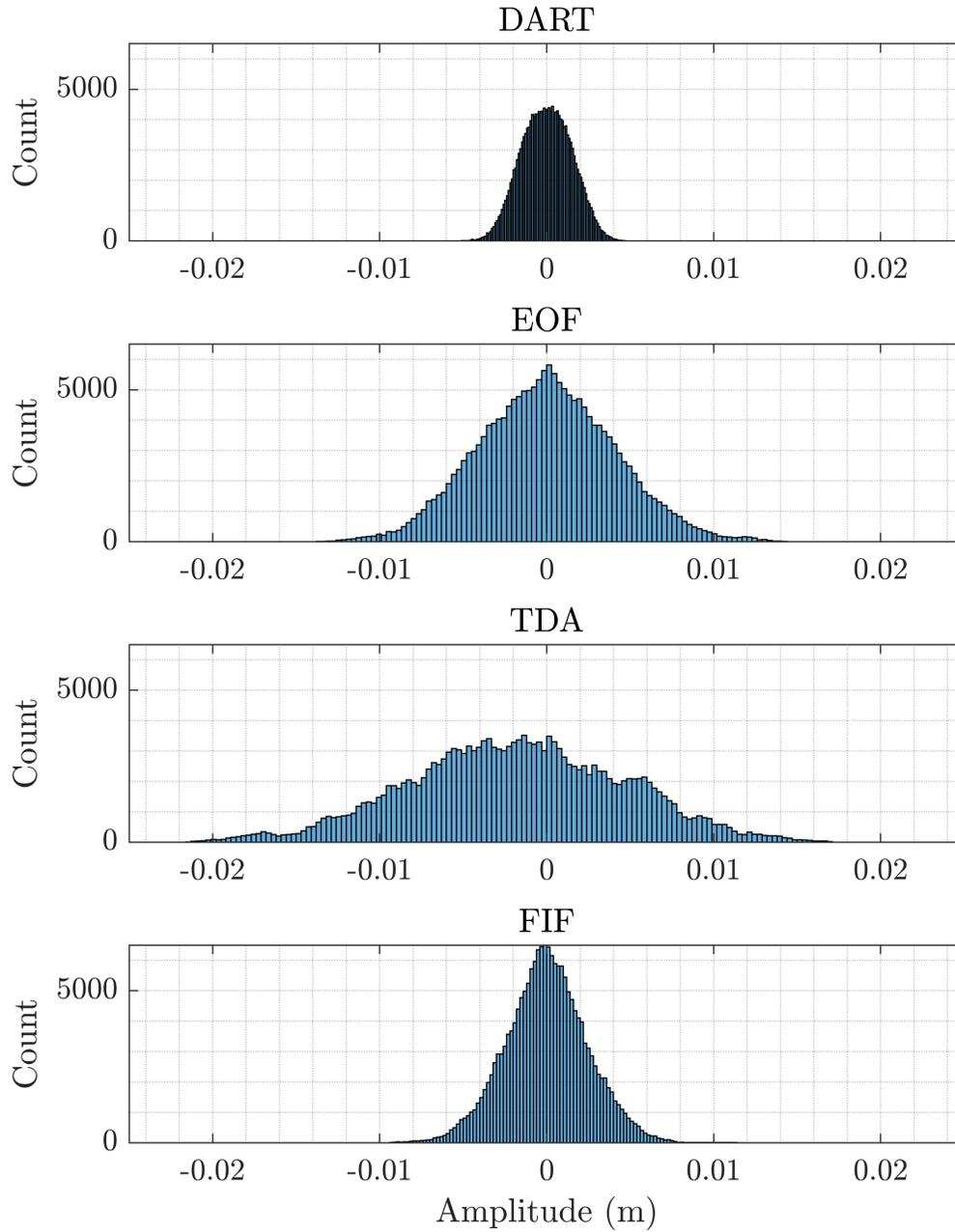


Figure 4.6: Histograms of the detection curves for each technique for background case A. DART refers to Mofjeld's algorithm.

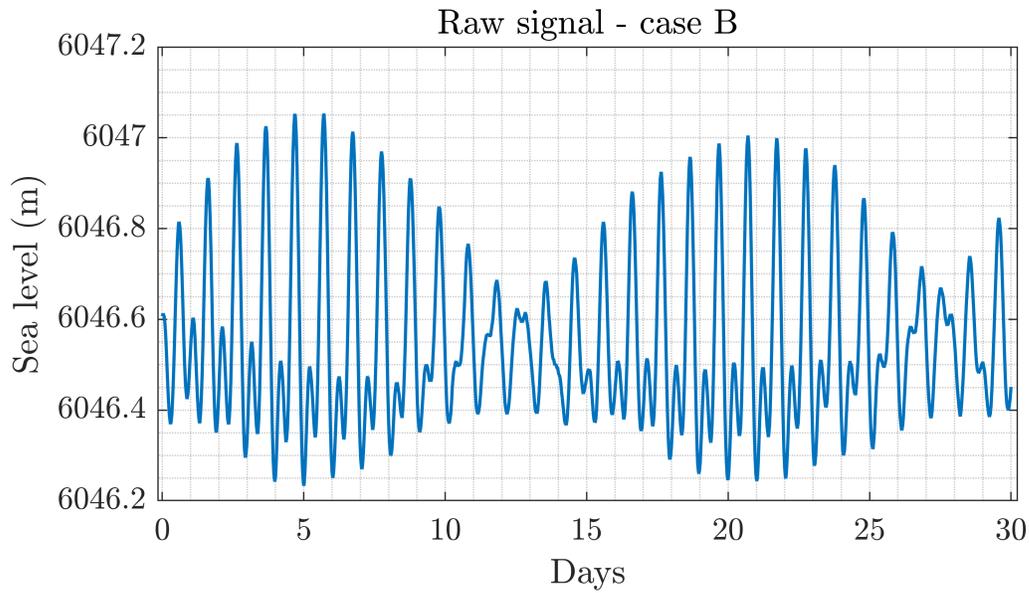


Figure 4.7: Raw data relative to case B in background analysis. Data from DART 52402, between 01/06/2016 and 30/06/2016.

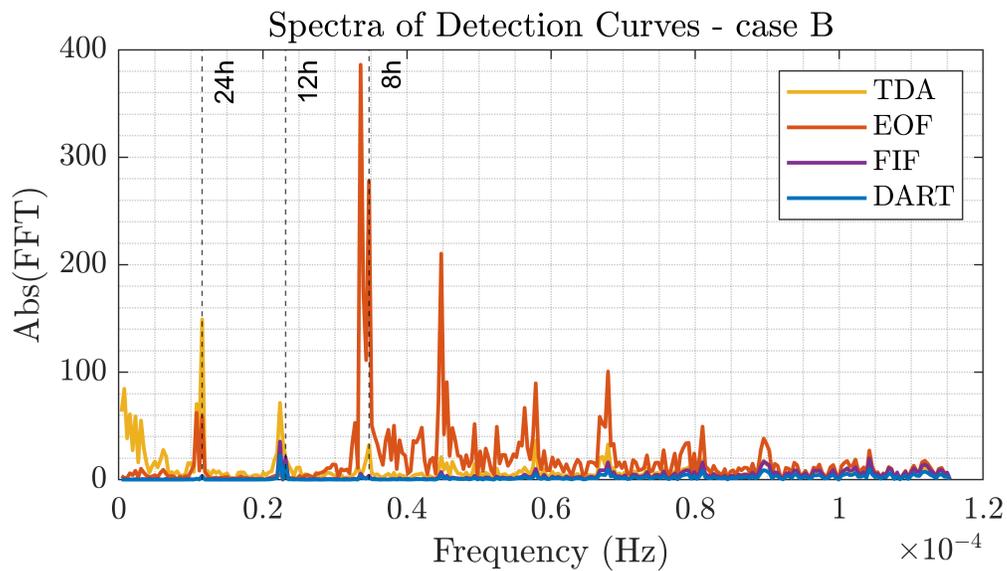


Figure 4.8: Absolute values of the FFT for detection curves obtained for background case B. DART refers to Mofjeld's algorithm.

Detection curves - case B

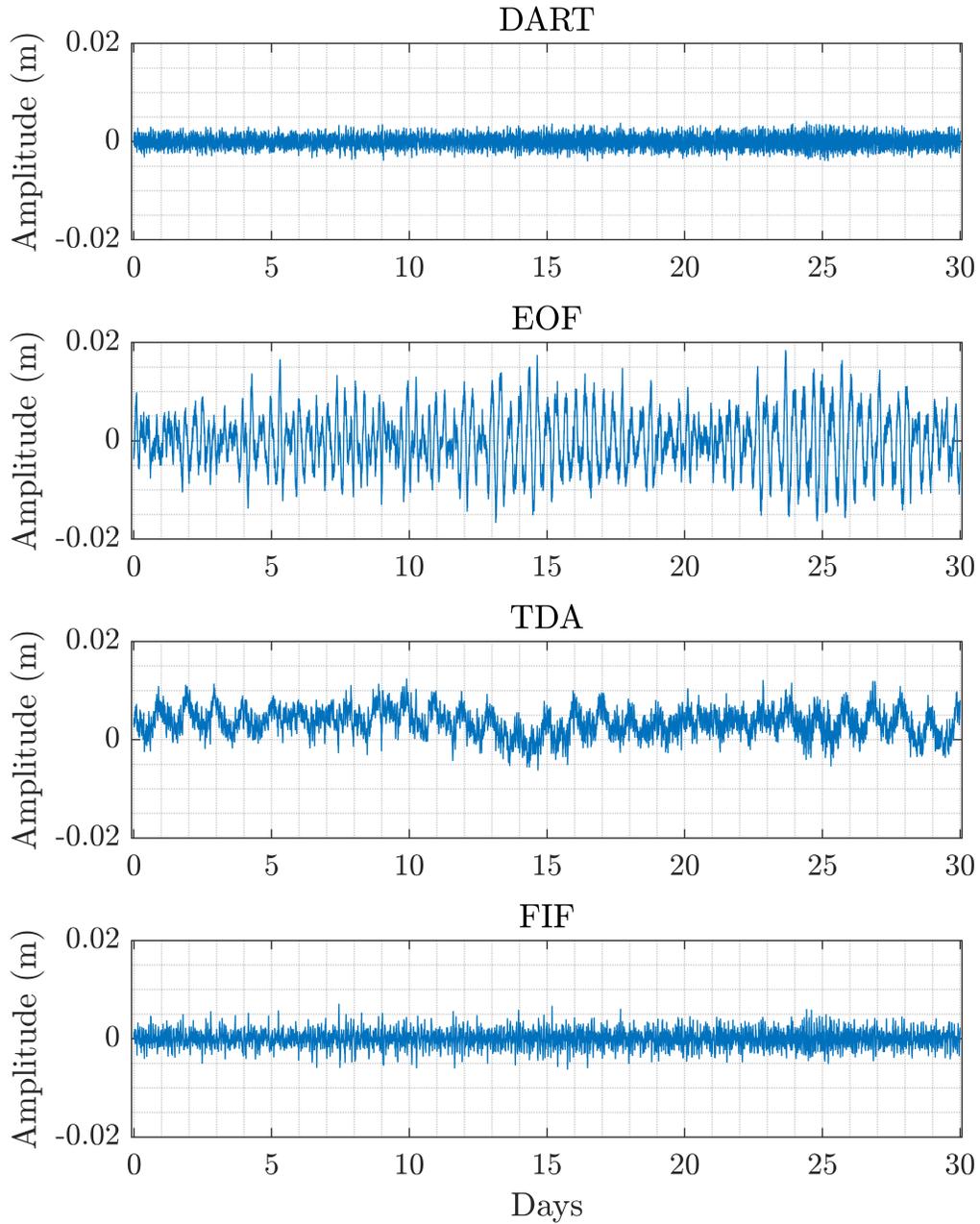


Figure 4.9: Detection curves for each technique for background case B. DART refers to Mofjeld's algorithm.

Histograms - case B

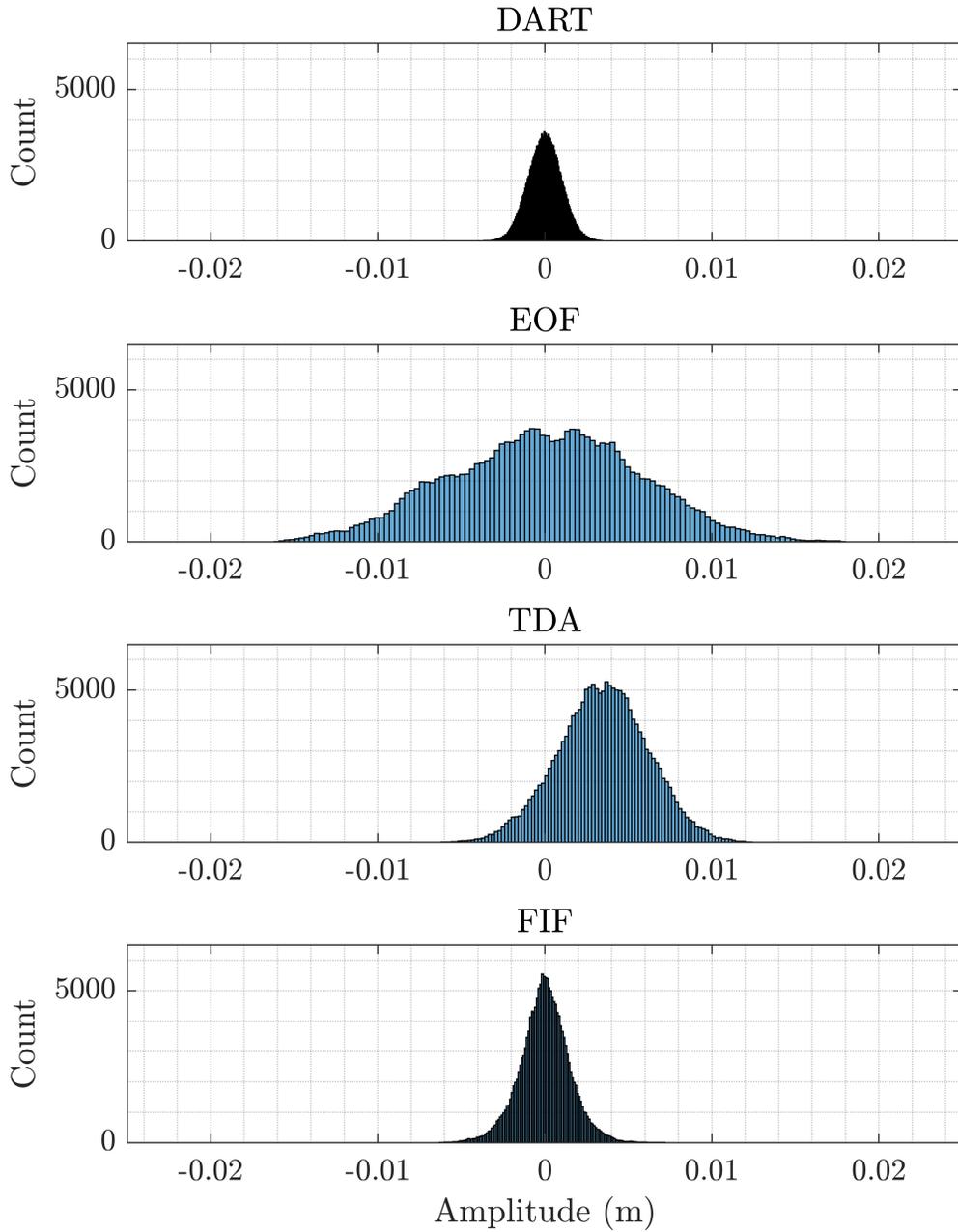


Figure 4.10: Histograms of the detection curves for each technique for background case B. DART refers to Mofjeld's algorithm.

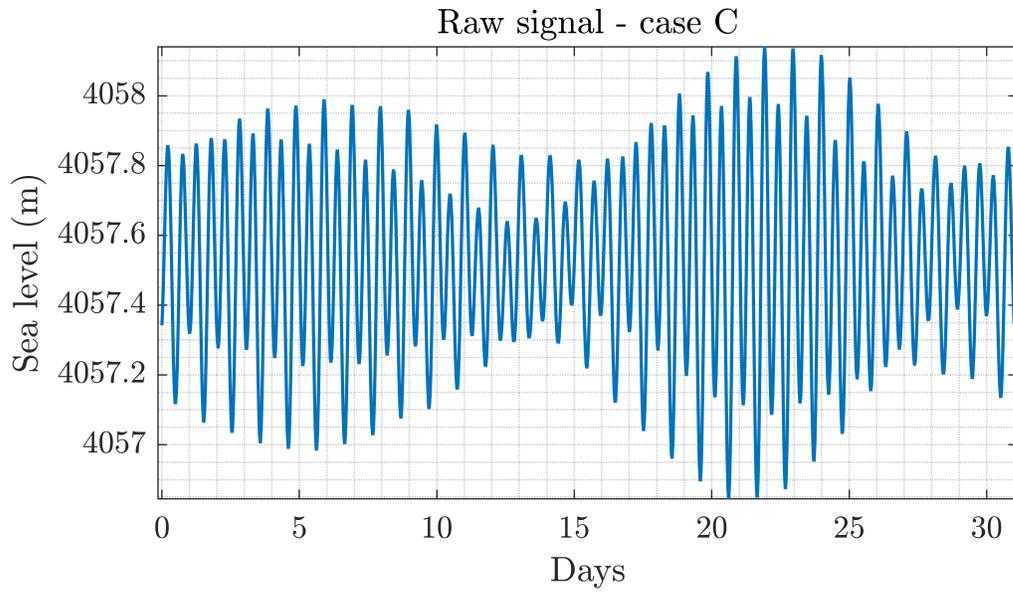


Figure 4.11: Raw data relative to case C in background analysis. Data from DART 32413, between 01/01/2019 and 31/01/2019.

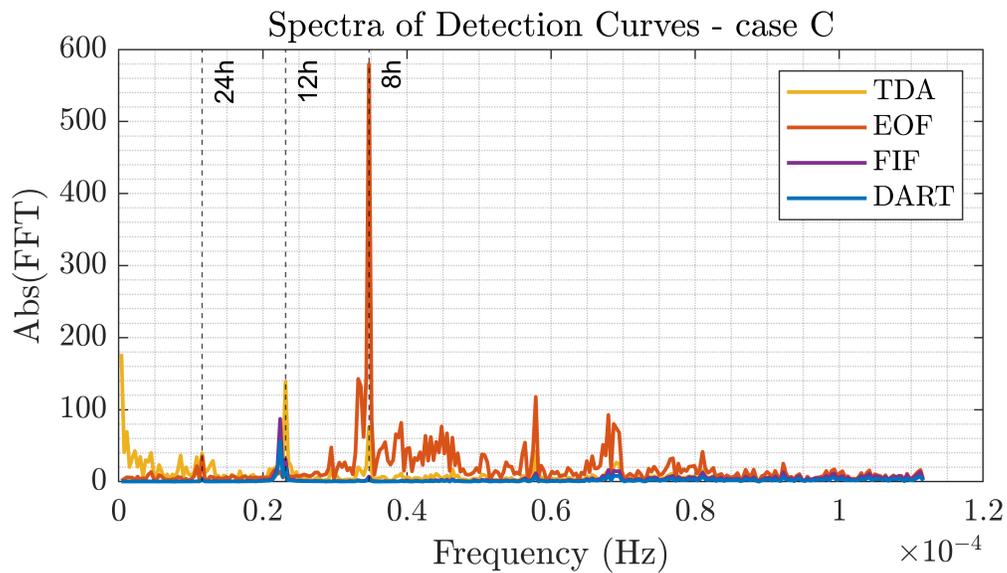


Figure 4.12: Absolute values of the FFT for detection curves obtained for background case C. DART refers to Mofjeld's algorithm.

Detection curves - case C

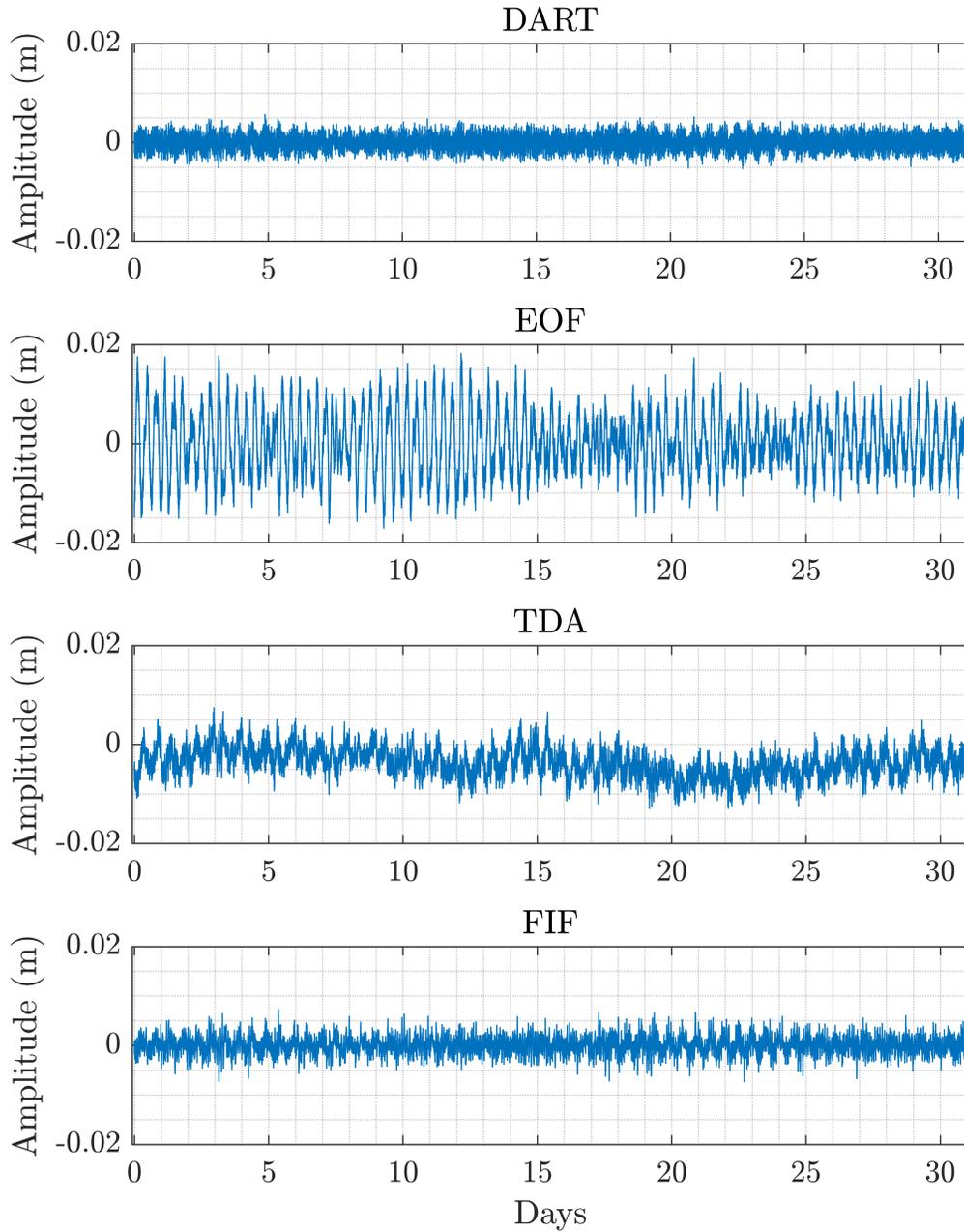


Figure 4.13: Detection curves for each technique for background case C. DART refers to Mofjeld's algorithm.

Histograms - case C

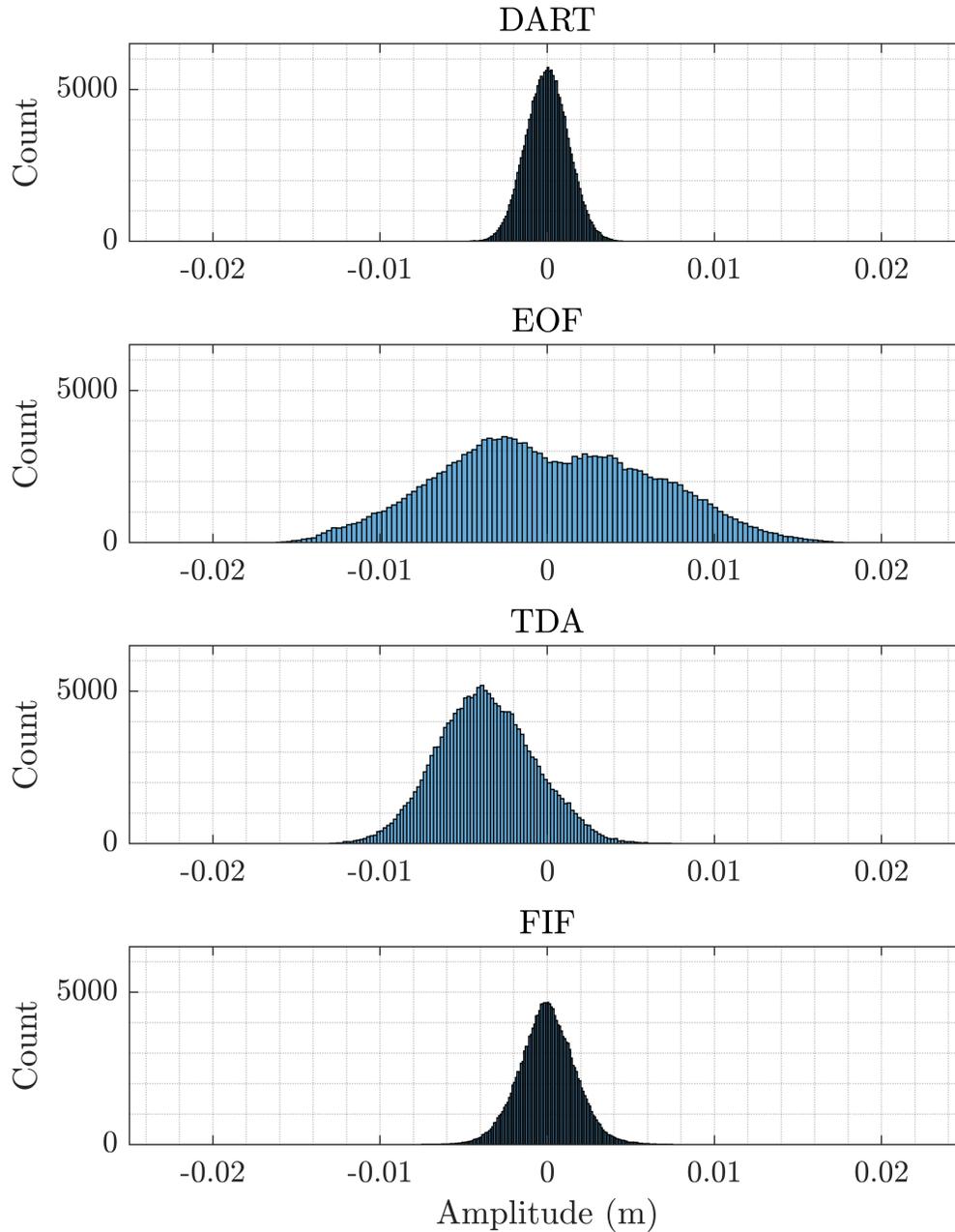


Figure 4.14: Histograms of the detection curves for each technique for background case C. DART refers to Mofjeld's algorithm.

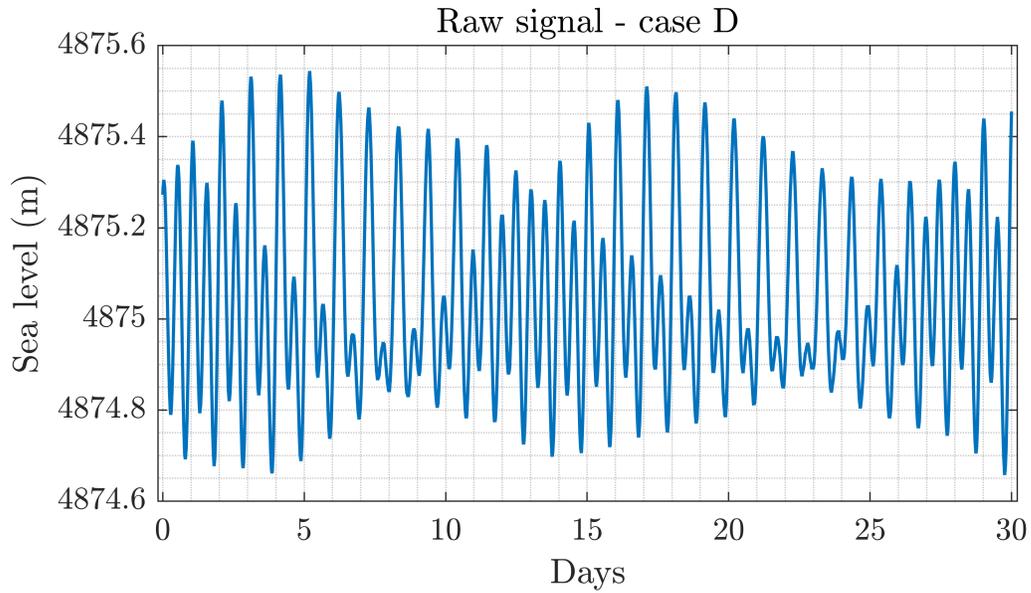


Figure 4.15: Raw data relative to case D in background analysis. Data from DART 51407, between 15/04/2022 and 14/05/2022.

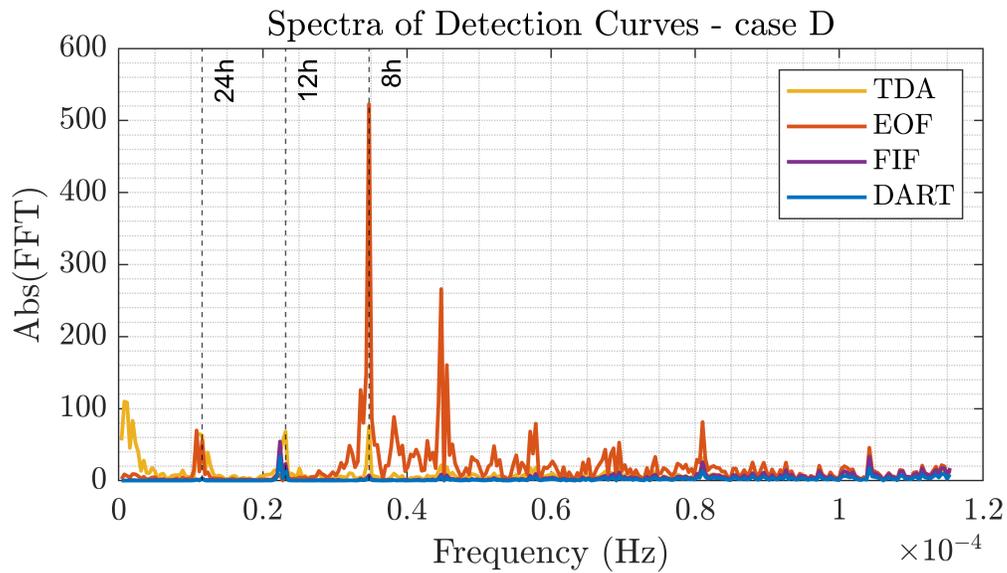


Figure 4.16: Absolute values of the FFT for detection curves obtained for background case D. DART refers to Mofjeld's algorithm.

Detection curves - case D

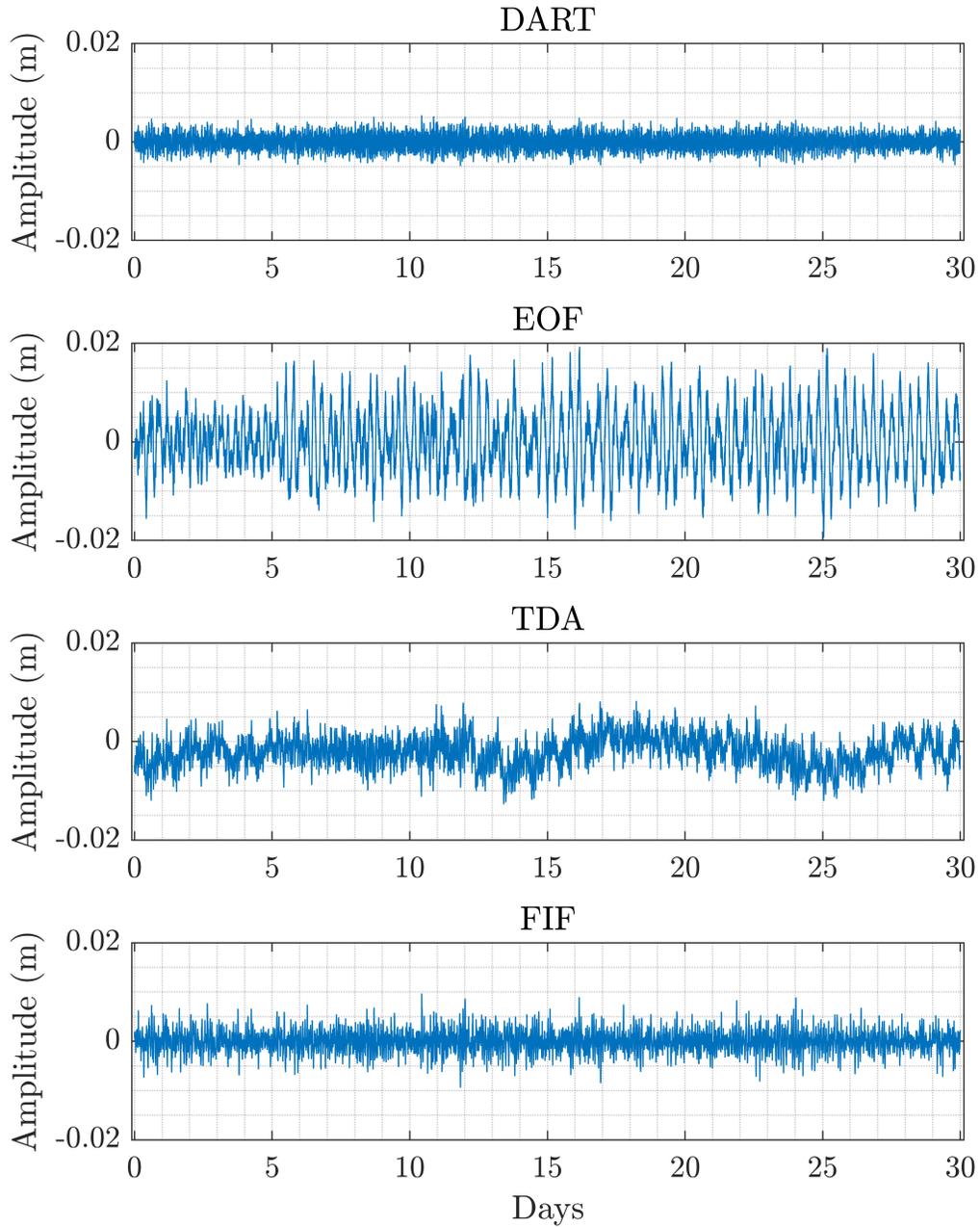


Figure 4.17: Detection curves for each technique for background case D. DART refers to Mofjeld's algorithm.

Histograms - case D

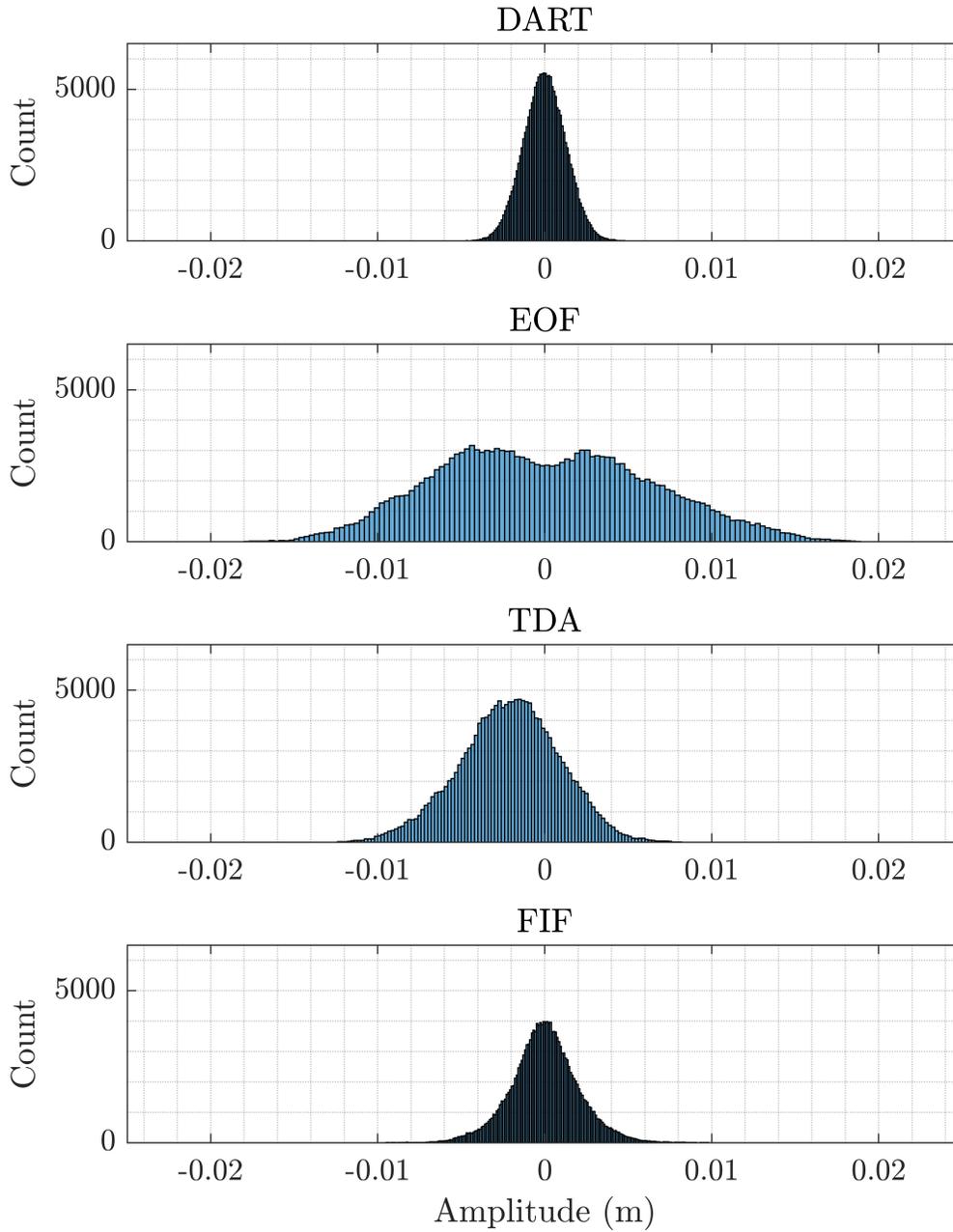


Figure 4.18: Histograms of the detection curves for each technique for background case D. DART refers to Mofjeld's algorithm.

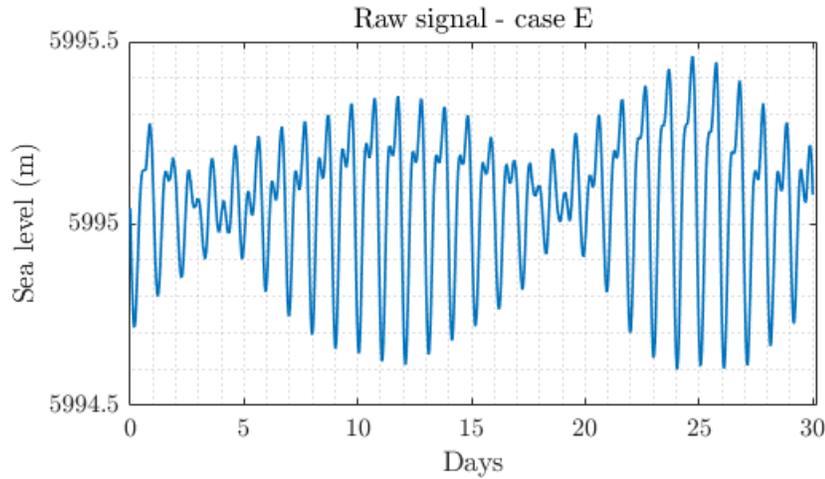


Figure 4.19: Raw data relative to case E in background analysis. Data from DART 21413, between 01/06/2021 and 30/06/2021.

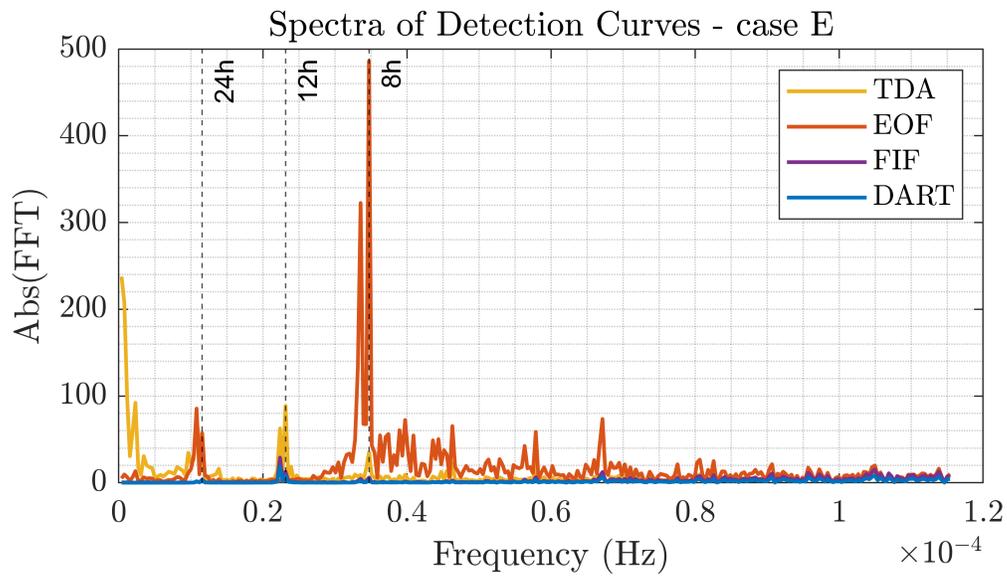


Figure 4.20: Absolute values of the FFT for detection curves obtained for background case E. DART refers to Mofjeld's algorithm.

Detection curves - case E

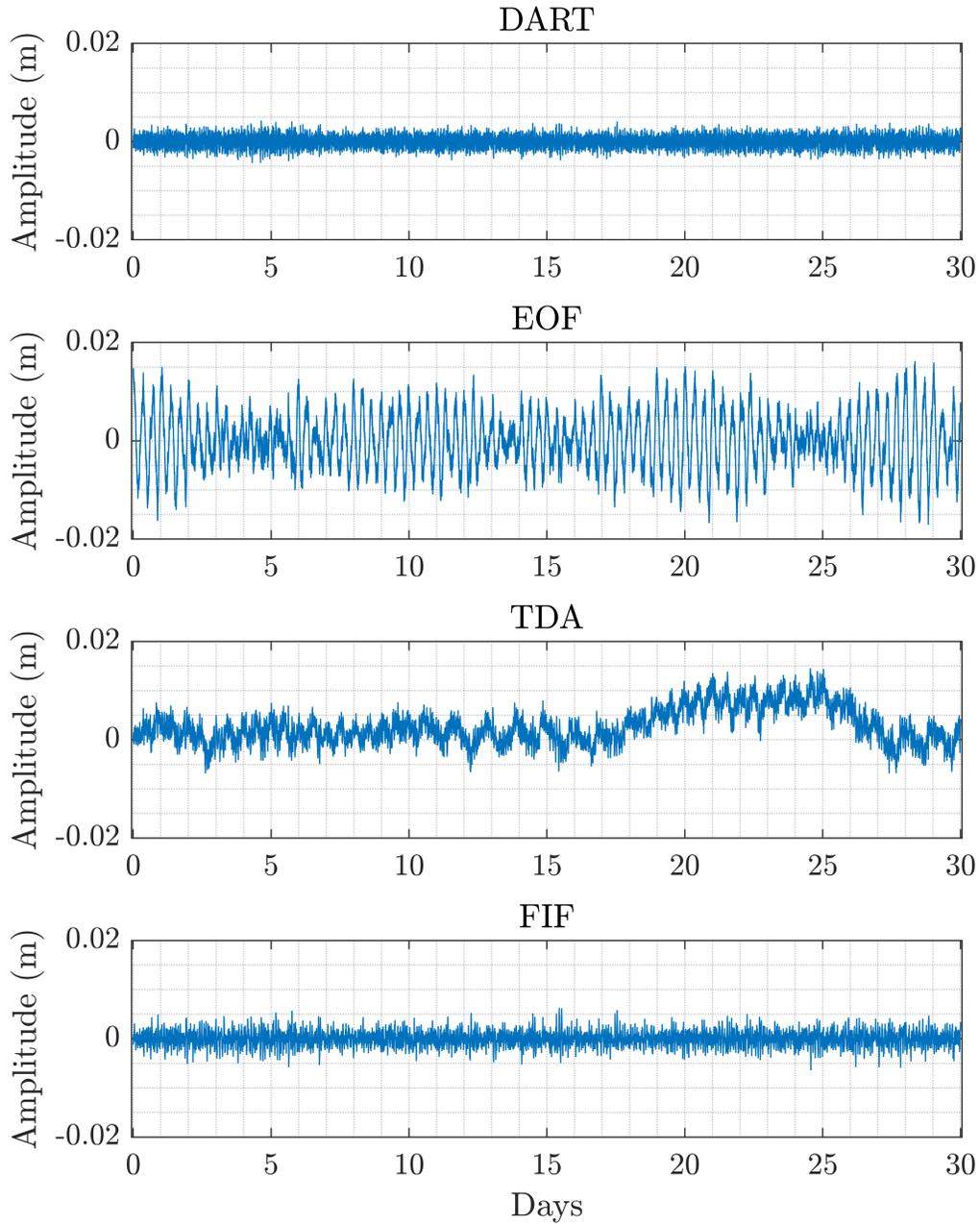


Figure 4.21: Detection curves for each technique for background case E. DART refers to Mofjeld's algorithm.

Histograms - case E

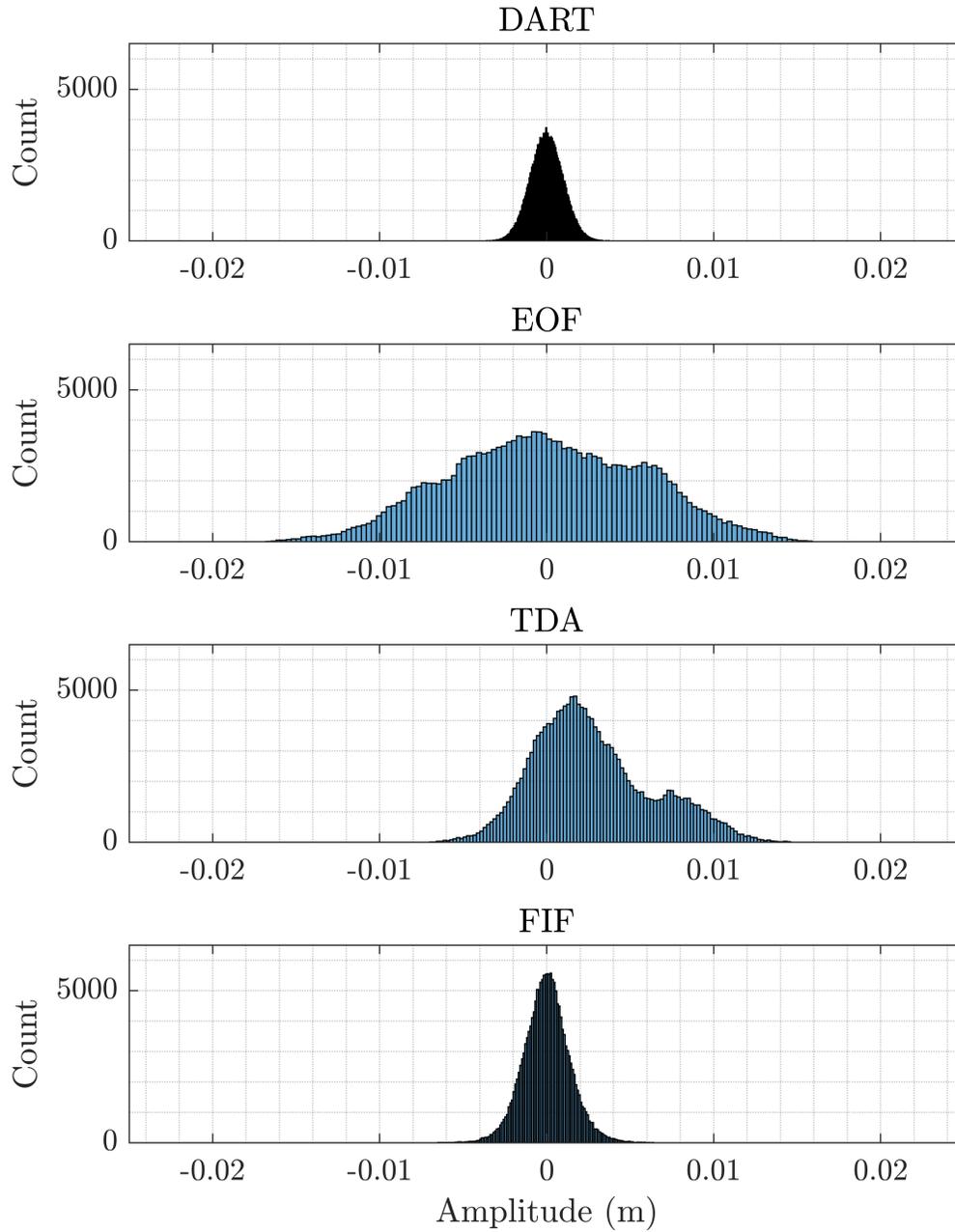


Figure 4.22: Histograms of the detection curves for each technique for background case E. DART refers to Mofjeld's algorithm.

Chapter 5

Detection tests on tsunami signals

In this chapter, we test the four detection techniques against signals recorded from DART® OBPGs during past tsunami events. The dataset is built in such a way to account for the variability that can be observed in real-time tsunami monitoring. In fact, records differ for the characteristics of the tsunami wave, due to factors like the polarity of the leading wave, dispersion effects and directivity in the propagation, presence of seismic shaking, which may overlap with the tsunami signal, different tidal and noise regimes, and even absence of any transient signal for far away instruments and/or weak tsunamis. The number of tsunami, earthquake and false detections are computed as functions of a variable detection threshold, whose impact on each technique's detection capability is shown and commented. Then, we check how detection curves relative to the FIF-based detection method reproduce the tsunami waveforms, in order to assess the possibility of characterizing period and amplitude in real time. Lastly, we want to note that the general procedure for detection testing presented here can be applied to any other tsunami detection method, provided enough observations are available.

Many of the detection tests presented in the literature have been limited in scope and have been applied to a very limited number of real tsunami records, such as in the works by Chierici et al. (2017), Di Risio and Beltrami (2014), Lee et al. (2016), Tolkova (2010), and Wang et al. (2020b), where no more than three experimental tsunami records have been used. Among more comprehensive tests, we can find the works by Bressan and Tinti (2011), where data

from seventeen events recorded by a single instrument (Adak Island coastal tide gauge) is used, Bressan and Tinti (2012), who analyzed 123 coastal tide gauge records relative to a single event (11 March 2011 Tohoku-Oki tsunami), and Bressan et al. (2013) and Lee et al. (2016), where simulated tsunami signals are used. The testing strategy with the highest statistical significance is probably the Montecarlo procedure presented by Chierici et al. (2017). However, the waves used are simple sinusoids superimposed over background data, since the number of waves needed is too high to use both simulated and real tsunami signals. The methodology proposed here takes into account variability in instrument location, background conditions and tsunami waveforms on a completely data-driven basis, without the need for simplified or simulated waveforms. While we recognize that some detection methods needs calibration for each instrument (e.g. Bressan and Tinti, 2011), and some may be appropriate for instruments where observations are limited (e.g. Di Risio and Beltrami, 2014), the detection methods proposed here are suitable for a completely data-driven instrument-independent calibration. For readability reasons, plots relative to section 5.3 are moved in a dedicated appendix at the end of this chapter.

5.1 Dataset creation

In order to test the tsunami detection technique exclusively on real data, our starting point is a catalogue of past tsunami events. In the present work, we start with the list of events used by Davies (2019). The catalogue is shown in Tab. 5.1 and is made of tsunami events triggered by thrust earthquakes from in the Global Centroid Moment Tensor catalogue (Ekström et al., 2012) that occurred between 2006 and 2016, with depth less than 71 km and $M_w > 7.7$. The reason behind this choice is that the distributions in magnitude and location result in a diverse list of tsunami records from OBPGs, in terms of amplitude, waveform and possible superposition with seismic waves. Another reason is that Davies (2019) used post-processed waveforms for this events, available online¹.

To build the signal catalogue, we proceeded as follows:

1. we selected all of NOAA's DART[®] gauges active at the time of each of the earthquakes in the Davies (2019) catalogue, whose data are available at the website of the NOAA Center for Tsunami Research²;

¹https://github.com/GeoscienceAustralia/ptha/blob/master/R/examples/austptha_template/SOURCE_ZONES/TEMPLATE/TSUNAMI_EVENTS/plots/README.md.

²<https://nctr.pmel.noaa.gov/Dart/>.

Event ID	Source zone	Date	M_w	Lon	Lat
KT1	Kermadec-Tonga	03/05/2006 15:26:40	8.0	-174.12	-20.19
KJ1	Kurils-Japan	15/11/2006 11:14:17	8.3	153.29	46.57
So1	Solomon Islands	01/04/2007 20:39:56	8.1	157.04	-8.46
SA1	South-America	15/08/2007 23:40:57	8.0	-76.60	-13.39
Su1	Sunda	12/09/2007 11:10:26	8.5	101.37	-4.44
SA2	South-America	14/11/2007 15:40:50	7.8	-69.89	-22.25
Pu1	Puysegur	15/07/2009 09:22:29	7.8	166.56	-45.76
KT2	Kermadec-Tonga	29/09/2009 17:48:11	8.1	-172.10	-15.49
NH1	New-Hebrides	07/10/2009 22:18:51	7.8	166.38	-12.52
SA3	South-America	27/02/2010 06:34:15	8.8	-72.71	-35.85
Su2	Sunda	06/04/2010 22:15:01	7.8	97.05	2.38
Su3	Sunda	25/10/2010 14:42:22	7.9	100.08	-3.49
KJ2	Kurils-Japan	11/03/2011 05:46:23	9.1	142.37	38.32
NH2	New-Hebrides	06/02/2013 01:12:25	7.9	165.11	-10.80
SA4	South-America	01/04/2014 23:46:47	8.2	-70.77	-19.61
SA5	South-America	16/09/2015 22:54:32	8.3	-71.67	-31.57
SA6	South-America	16/04/2016 23:58:36	7.8	-79.93	0.35
So2	Solomon Islands	08/12/2016 17:38:46	7.8	161.32	-10.68

Table 5.1: Catalogue of tsunamigenic earthquakes from Davies (2019), used to build a dataset of tsunami signals for detection testing.

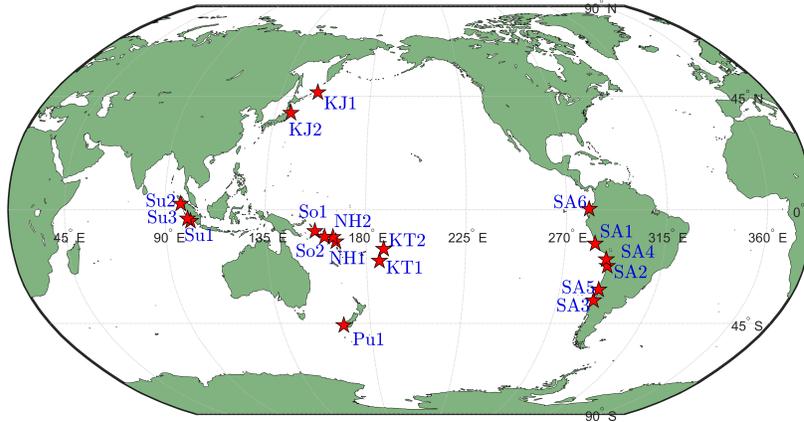


Figure 5.1: Map of the events in Davies' catalogue (Davies, 2019). The nomenclature is the same used by Davies (2019) and it is explained in Tab. 5.1.

2. from each DART we extracted data a variable time before the event (we will comment further on this later) up to 24 h after origin time;
3. we excluded signals with long segments of missing data and signals where only the transmitted data are available;
4. we removed instrumental spikes and filled the gaps by interpolation.

The reason behind point 3 is that we want to test the technique as if it was applied by the instrument in real time, i.e. on data sampled at 15s. On the other hand, point 4 is motivated by the fact that spikes in DARTs are usually caused by communication between the pressure gauge and the buoy and we assume here that they would be taken care of separately. The amount of needed data before the events depends on the technique. The data we are going to analyse are the detection curves starting from origin time lasting for 1 day, so we need the experimental signal to have enough data for each technique to work. We recall that Mofjeld's algorithm needs 3 h 5 min 15 s, or 741 samples at 15 s sampling intervals, of preceding data; EOF detiding needs one lunar day, or 5940 samples; TDA needs as many points as the order of the FIR filter, i.e. 4000, and FIF uses data segments that last 3 h, or 720 samples.

Since TDA assumes the tidal coefficients to be known during operation, the coefficients themselves have been computed in advance for each case using UTide as we did in the previous chapter for background signals. However, there are cases where this presents some difficulties. The first case is when we have instruments where not enough preceding data is available. This is the case for newly installed instruments, but also for instruments that have been recently resurfaced for maintenance and data recovery. In these cases, we did not apply TDA. A second case is when enough preceding data are available for UTide to give a convergent solution, but not enough to compute many tidal coefficients. In such cases TDA has been applied as described in Chapter 2, but we may find poor results, since large tidal residuals are passed to the detection curve. The problem is particularly evident in areas with large tidal oscillations, such as in the Gulf of Alaska.

In the end, we obtain a set of 437 signals. Due to the various possible features that each signal may present, determining if and what kind of detection occurs for each detection curve is carried out by visual inspection, based on the hypothesis that oscillations due to tsunamis and earthquakes are clearly distinguishable from one another and from oscillations in the detection curves not corresponding to an event. It is evident that a classification carried out in this way may be biased and contains a degree of arbitrariness. To alleviate the

situation, the analysis is carried out for two datasets. The first is the full set of 437 signals, while the second is composed by the 73 signals of our dataset for which a post-processed waveform is available in Davies (2019), so that a clear correspondence between detection curves and tsunami waves can be established. We will refer to the entire dataset simply as *full dataset* and to the second as *restricted dataset*.

5.2 Detection analysis

The first analysis we may carry out for the detection curves we just obtained is to look at the distribution of some basic statistical measures. Since we have seen in Chapter 4 that the amplitude of detection curves tend to follow bell-like distributions, we may give a global description of the full dataset by checking the distributions of signal averages and standard deviations. The signal average can be interpreted as a measure of symmetry in the amplitude distribution of the detection curves. As we can see from Fig. 5.2, the (absolute value of the) signal mean for a given technique may vary up to an order of magnitude. It is evident that the most asymmetric technique, i.e. the one that oscillates on average further from zero is TDA, where absolute means of up to a few centimeters are observed. As mentioned before, large residuals are to be expected whenever data to compute tidal coefficients are scarce. The relative ordering between techniques tend to be stable along the data set, with TDA performing the worst, followed by EOF detiding, followed by FIF-based detection and Mofjeld's algorithm. This is in line with the properties about asymmetry we observed on background signals in the previous chapter. As anticipated, TDA could not be applied to 55 of the signals, for which tides could not be computed and for this reason the curves for TDA in Fig. 5.2 are incomplete.

On the other hand, the standard deviation is an integral measure, so it is more correlated to the signal waveform. In fact, it shows a dependence on the event considered. If we look at weaker events, where most or all the signals are background, the standard deviations follow the same ordering as for the means, in accordance with what we observed in Chapter 4. If we look at bigger events, such as KJ1, SA3 and KJ2 (Tab. 5.1), we note an increase of up to two orders of magnitude and we can also observe that the techniques show better agreement in terms of standard deviation. Details about how the waveshapes are reproduced in the detection curves are omitted here and will be discussed in the next section.

To quantify detection rates of the techniques, we define

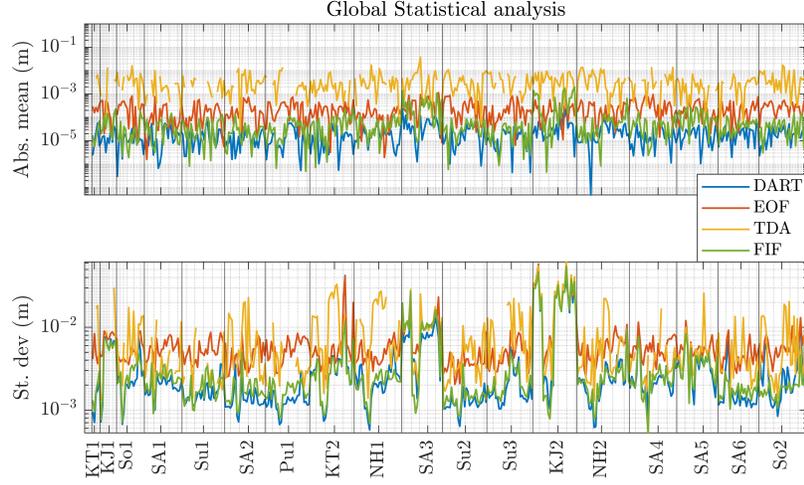


Figure 5.2: Statistical indicators for 24h long signals relative to the catalogue in Tab. 5.1. In the upper and lower plot, the absolute value of the signal mean value and the signal standard deviation are shown, respectively. Note that the TDA curve is incomplete due to the unavailability of data to compute tides in 55 signals.

- N : total number of signals in the dataset;
- n_F : number of signals with at least one false detection, i.e. at least one detection reducible neither to the tsunami wavetrain nor to earthquake shaking;
- n_E : number of signals with no false detections and at least one earthquake detection;
- n_T : number of signals with no false detections and at least one tsunami detection.

Furthermore, we define two further quantities, which we will call *detection scores* as

$$\theta_1 = \frac{n_T - n_F}{N} \quad (5.1)$$

and

$$\theta_2 = \frac{n_T - n_E - n_F}{N} \quad (5.2)$$

These detection scores are presented as metrics to measure the performance in some specific applications and we can choose which metric needs to be optimized based on the approach we want to apply to the detection network. One

possible approach can consist of a detection algorithm which gives as few false detections as possible, in order to reduce false alarms as much as possible. In this case, we would like a detection methods that minimizes n_F/N . However, this may lead to a too conservative algorithm. For this reason, we may decide to maximize θ_1 instead. In fact, this detection score measures the rate of tsunami detections, while also penalizing false detections. As already pointed out in previous chapters, the overlap of seismic shaking and tsunami oscillations may cause difficulties in the characterization of wave characteristics, especially if the instrument is placed close to seismogenic areas. So, we may want to also penalize the possibility of detecting earthquakes, in which case we may choose a detection algorithm that maximizes θ_2 .

To optimize the techniques, we computed n_F/N , n_E/N , n_T/N , θ_1 and θ_2 for each techniques for detection thresholds T varying from a minimum $T = 1$ cm to a maximum of $T = 4$ cm with a step of $T = 0.5$ cm. The choice of the minimum is due to the fact that, $T = 0.5$ cm would result in false detections on background signals using Mofjeld's algorithm, which gives the narrowest distribution in amplitude in detection curves. On the other hand, the choice of the maximum is motivated by the fact that choosing a too big T may lead to missing potentially dangerous events. The detection rates for earthquake and tsunami waves are shown in Fig. 5.3. The earthquake detection rates reflect very well the filtering capabilities of the techniques. In fact, TDA and the FIF-based method show much lower earthquake detection rates than EOF and Mofjeld's algorithm, though they still detect some earthquakes for every detection threshold. This is due to both the presence in the dataset of very large events, but also the boundary condition effects described in Chapter 2. The general trend is for the detection rate to decrease with T , i.e. the way to minimize earthquake detection is to use bigger thresholds. EOF detiding is an exception to the trend, showing reaching a maximum detection rate for $T = 2$ cm. Since we are counting signals with earthquake detection but no false detection, this shows that EOF is dominated by false detections for thresholds below 2 cm. The same happens for EOF detiding but also for TDA in the case of the tsunami detection rate, where false detections dominates over tsunami detections for thresholds below 1.5 cm for the full dataset and 2 cm for the restricted dataset. The FIF-based method has a mixed behaviour: on the restricted dataset, n_T/N reaches a maximum for 1.5 cm, while it decreases monotonically in the full dataset. The false detection rates in Fig. 5.4 confirms the strong role of false detections in the case of TDA and EOF detiding. These two techniques show much larger false detection rates for low T . However, EOF detiding reach an asymptotic behaviour

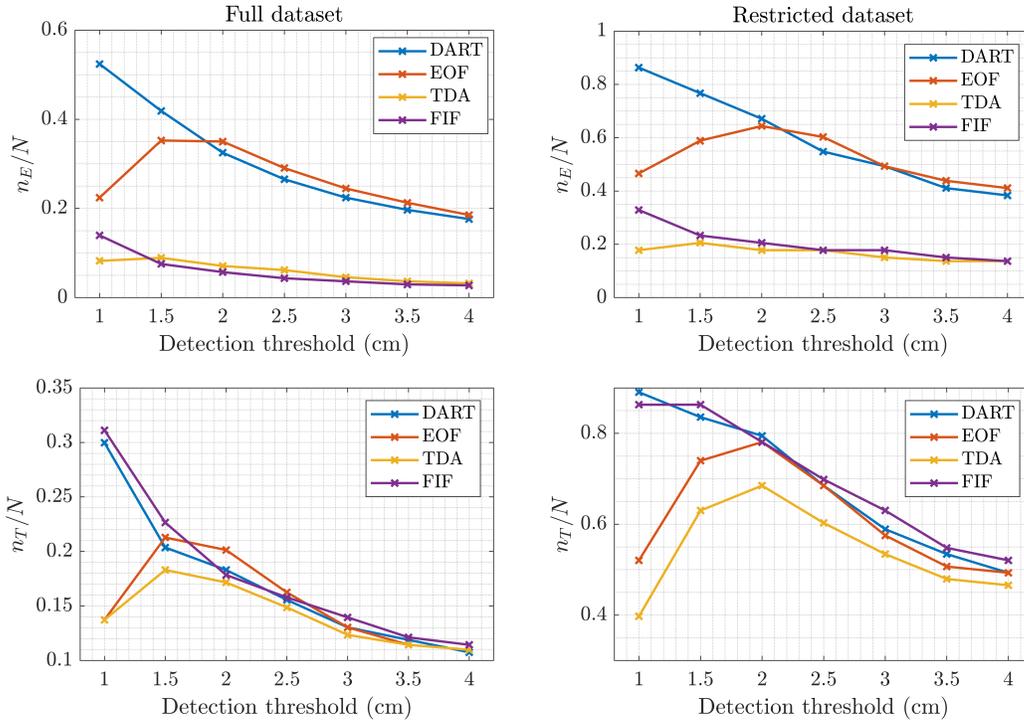


Figure 5.3: Rates of earthquake and tsunami detections as functions of a varying detection threshold in both the full and restricted datasets. Here, Mofjeld’s algorithm is indicated simply as DART.

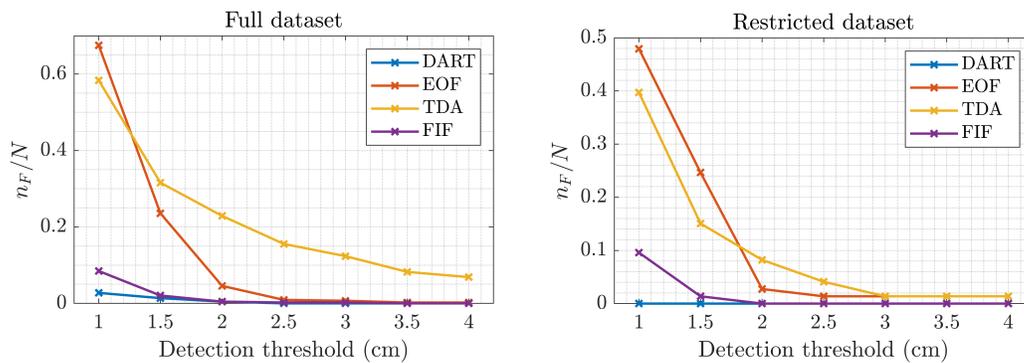


Figure 5.4: False detection rates as functions of a varying detection threshold for both the full and restricted datasets. Here, Mofjeld’s algorithm is indicated simply as DART.

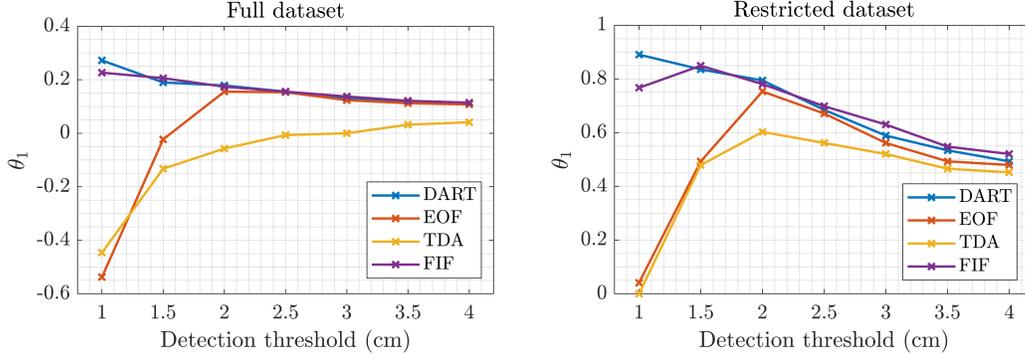


Figure 5.5: Detection score θ_1 , defined in eq. (5.1), as function of detection threshold for both the full and restricted dataset. Here, Mofjeld’s algorithm is indicated simply as DART.

for $T \geq 2.5$ cm, so we may interpret this as the optimal threshold to minimize false detections. TDA also reaches an asymptote in the restricted dataset for $T \geq 3.0$ cm, but it does not reach an asymptote in the full dataset, due to the higher number of signals with not enough data for precise tide prediction. Mofjeld’s algorithm and the FIF-based method show much better results. In the full dataset, Mofjeld’s algorithm show no false detections for $T \geq 2.5$ cm, while it shows none for any threshold in the restricted one. The FIF-based technique has a slightly worse performance for very low threshold and it converges to zero false detections for $T \geq 2.5$ cm in the full dataset and $T \geq 2.0$ cm in the restricted one. We also note that for $T = 2.0$ cm the method has the same number of false detections ($n_F = 2$) in the full dataset as Mofjeld’s algorithm. These are very promising results, due to the role as de facto standard tsunami detection method that Mofjeld’s algorithm has.

To find optimal thresholds, we may now look at the detection scores θ_1 and θ_2 . In the case of θ_1 , we have different behaviour between the two datasets. In the case of the restricted dataset, TDA, EOF detiding and the FIF-based technique all have an optimal threshold, equal to 2 cm for the first two and to 1.5 cm for the third, while Mofjeld’s algorithm has a monotonic behaviour. In the full dataset, only EOF detiding has a maximum value at 2.0 cm, while the other techniques all have a monotonic trend: Mofjeld’s algorithm and the FIF-based method decrease with the threshold, i.e. their performance gets better with lower T , while TDA increases with T , showing the strong influence of detection curves with large oscillations in the full dataset.

In the case of θ_2 , the FIF-based method shows a better performance than all

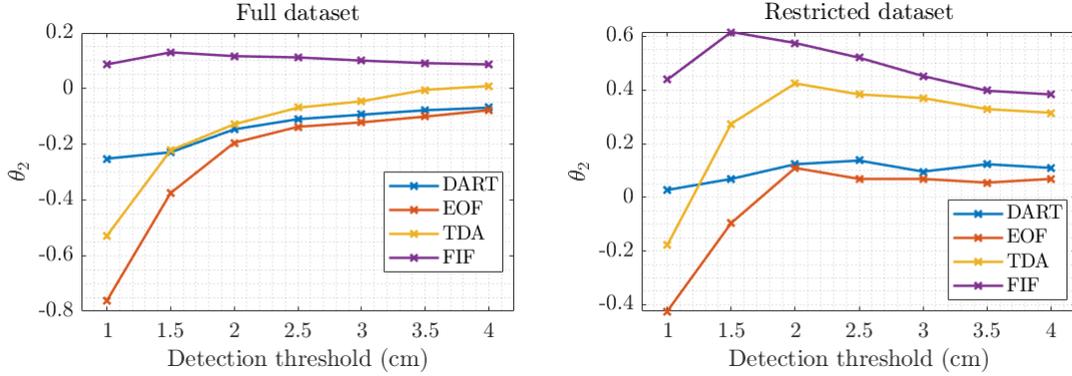


Figure 5.6: Detection score θ_2 , defined in eq. (5.2), as function of detection threshold for both the full and restricted dataset. Here, Mofjeld’s algorithm is indicated simply as DART.

three other techniques. In the case of EOF detiding and Mofjeld’s algorithm, the reason is found in the fact that θ_2 penalizes earthquake detections and they do not include any high frequency filtering modules. They increase monotonically in the full dataset, while they show a weakly concave trend in the restricted dataset. As for previous metrics, EOF gets significantly worse as soon as a threshold smaller than 2 cm is considered. On the other hand, TDA continues to suffer from the problem encountered previously: a number of signals with poor tidal model, which cause θ_2 to be monotonically increasing in the full dataset, and a set of signals for which tides cannot be computed at all, from which we get the offset in the restricted dataset between it and the FIF-based method. The FIF-based technique is the only one showing a consistent behaviour between the two datasets. In both cases, the optimal value of θ_2 is reached for $T = 1.5$ cm, which can be interpreted as the optimal threshold to detect tsunamis while also minimizing false or earthquake detections.

As we have done in chapter 4 for background signals, we can have try to visualize the effect of data availability on TDA detection curves. To do this, we can look at the relation between the maximum absolute amplitude of each detection curve with the amount of previous data available for computing a tidal model. For the maximum amplitude to be representative of the effects of tidal residuals and not tsunami waves or earthquake oscillations, curves need to be selected carefully. For this reason, we only keep TDA detection curves which do not trigger a detection or which trigger a false detection. However, in this work we label as false detection any detection curve where a detection not related to tsunami waves or earthquake oscillations occurs, so we also need to eliminate

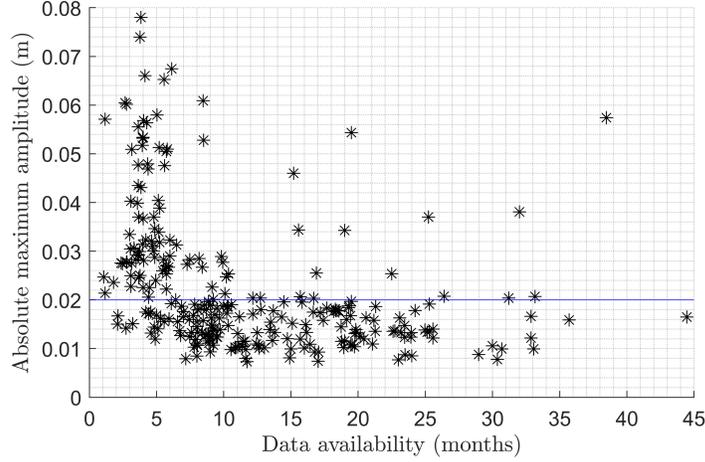


Figure 5.7: Scatterplot of maximum absolute amplitude as a function of previous data availability in months for TDA detection curves. The blue curve represent the amplitude threshold $T = 2$ cm.

curves where large tsunami waves occur. To this aim, we can compare with detection curves obtained with the FIF-based method. In fact, the FIF-based method presents no detection curves with false detections and visible tsunami waves for a detection threshold $T = 2$ cm. In the end, we select the detection curves which

- trigger either a false or no detection with TDA;
- trigger no detection with the FIF-based method.

In Fig. 5.7, we plot the maximum absolute amplitude for TDA detection curves in this subset as a function of the amount of data between the origin time and the beginning of the deployment of the instrument. We note that for relatively long data availability (approximately 6 or more months), not the full time series has been used for tidal modeling and for each case we proceeded on a case by case basis, to obtain the best model fitting tidal coefficients. Nonetheless, for shorter time series, where we used the full available data, the effects are more evident. Fig. 5.7 shows a correlation between amplitude of detection curves and data availability. In fact, we notice a general decrease of maximum absolute amplitude with increasing data availability, though some outliers are still present. The analysis shown here can thus be used to choose the optimal detection threshold for an instrumental network, following different criteria according

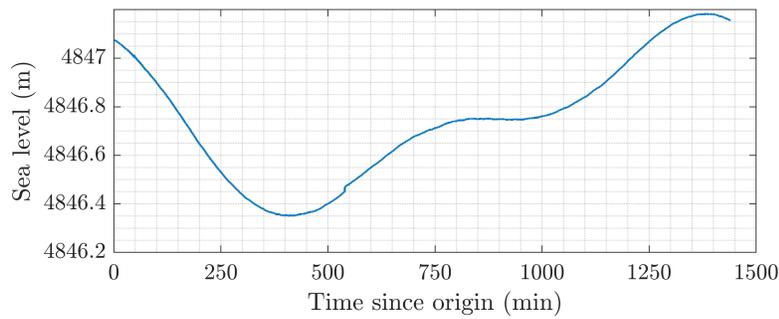


Figure 5.8: Raw data relative to event Pu1, DART 21415. Time is in minutes since origin time. Note the jump discontinuity between $t = 500$ min and $t = 550$ min.

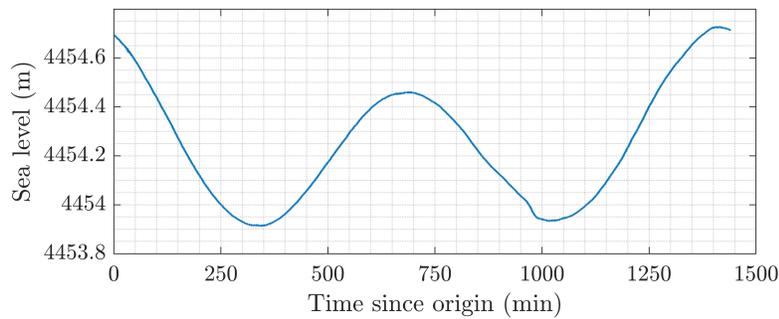


Figure 5.9: Raw data relative to event SA2, DART 44402. Time is in minutes since origin time. Note the sharp change of gradient between $t = 950$ min and $t = 1000$ min.

to the needs of the specific network. In particular, we may choose a threshold that minimize false detection at all cost, or one where we allow some false detections as long as we get improvements in detecting tsunamis. Mofjeld's algorithm works as expected: it has great performance in reducing non tsunami and non earthquake oscillations, resulting in very low false detection rates. However, if it is desirable to filter out seismic shaking, the technique performs worse than others. EOF detiding has somewhat of a hard threshold in efficiency around 2 cm, below which we get a sharp decrease. This is in line with the amplitude of residuals we expect after removing tides with EOF basis. TDA is strongly dependent on data quality before the segment of data under check and the absence, in some cases, or reduced length, in others, of data preceding the events strongly influence the overall performance. The technique may very well benefit from a different tide removal method.

Finally, we can also provide some considerations about optimal detection thresholds for the implementation and application of the FIF-based method. In particular, if we want to minimize the false detection rate, we propose $T = 2$ cm. For this value, the technique shows only 2 false detections. Upon further investigation, the two signals that give false detection are from event Pu1, DART 21415, and event SA2, DART 21415, which show unexpected sharp gradients, as shown in Fig. 5.8 and 5.9, respectively. In the detection curves in Fig. 5.10 and 5.11, we can see that such artifacts affect the detection curves of all techniques. For Mofjeld's algorithm, this happens because these gradients are too steep to be captured by the extrapolation. In other techniques, the problem is amplified by the effects that filtering operations have on discontinuities and jumps. The occurrence of such cases strongly suggests that, for operational purposes, any of the techniques should be accompanied by jump and spike detection algorithms. On the other hand, it also confirms that the FIF-based technique has no false detection on background signals for a threshold $T = 2$ cm, which we can consider appropriate for operational contexts. A threshold $T = 1.5$ cm presents false detection not caused by artifacts in the data, but it is shown to be the optimal parameter in terms of the detection score θ_1 .

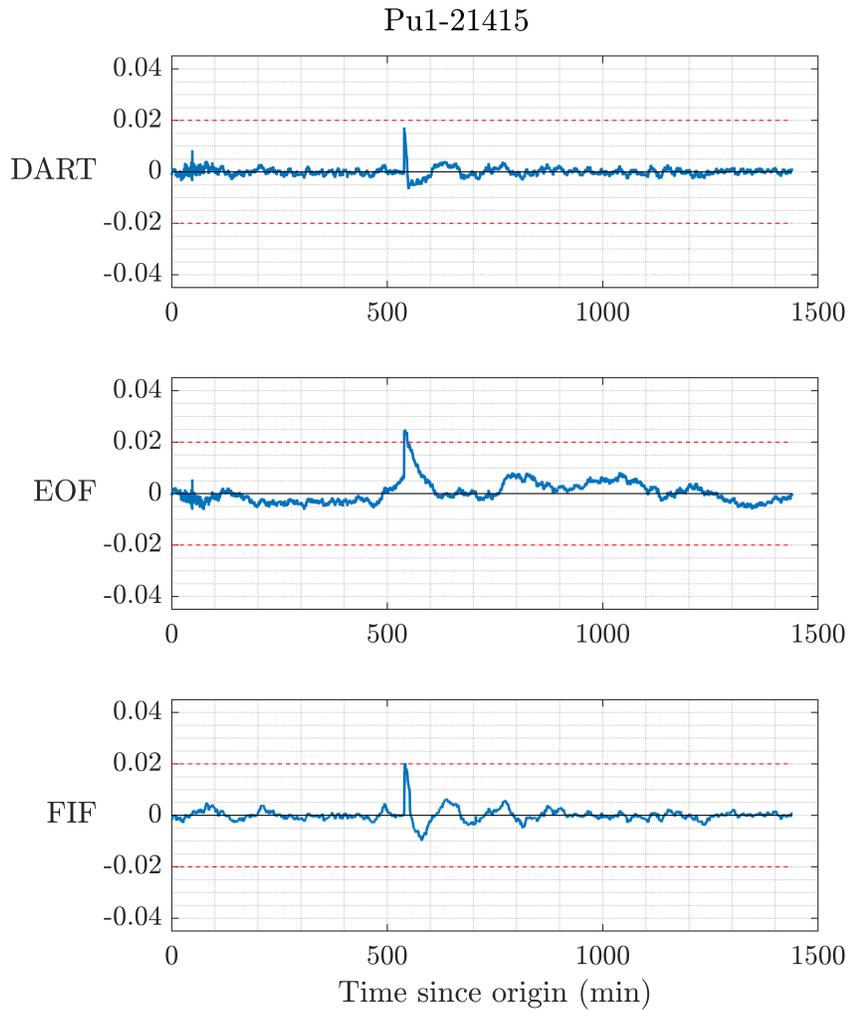


Figure 5.10: Example of detection curves for event Pu1, DART 21415. The amplitudes on the vertical axis are measured in centimeters. Here, Mofjeld's algorithm is indicated simply as DART. TDA was not applied due to the unavailability of preceding data to compute tidal coefficients.

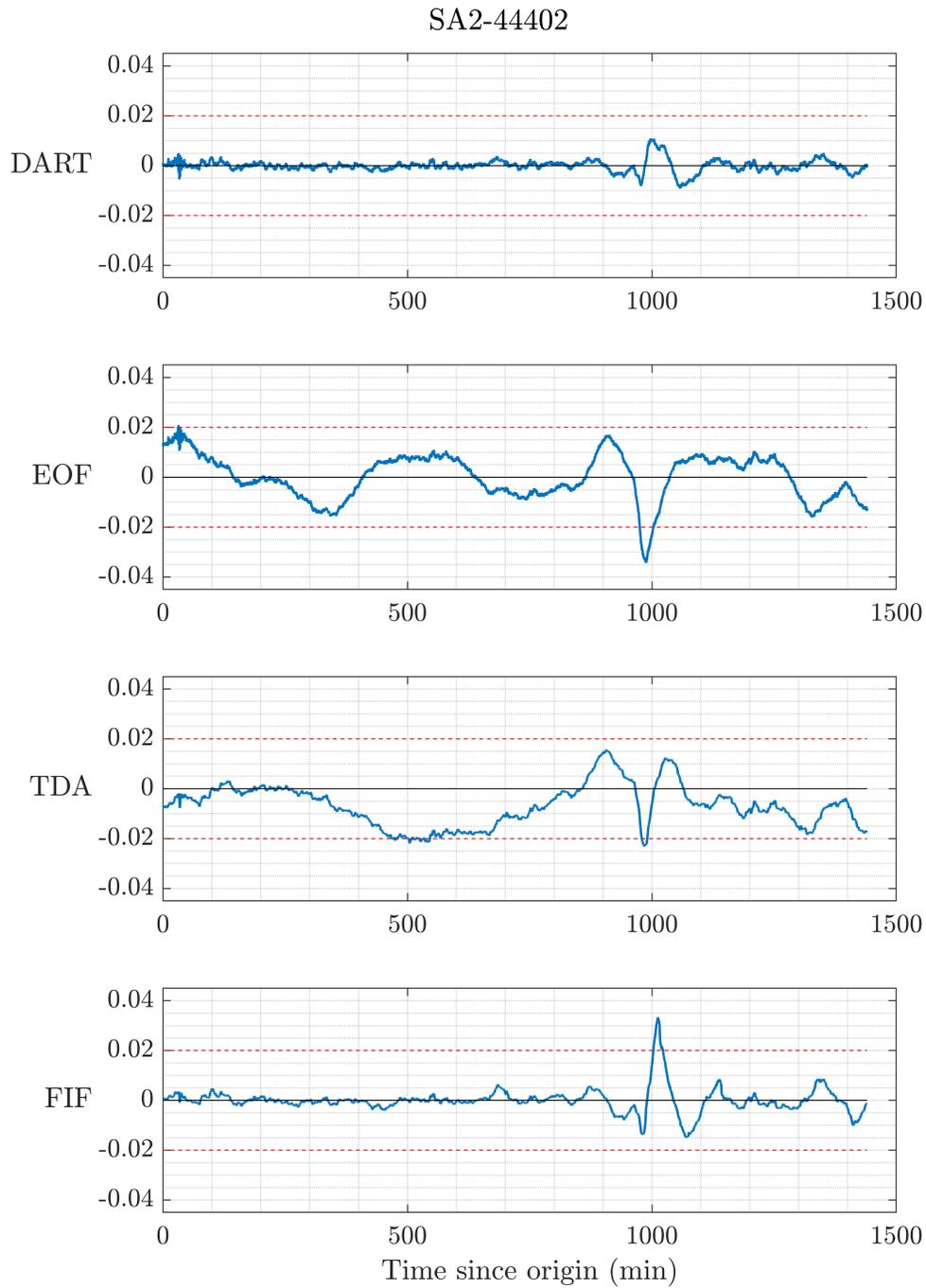


Figure 5.11: Example of detection curves for event SA2, DART 44402. The amplitudes on the vertical axis are measured in centimeters. Here, Mofjeld's algorithm is indicated simply as DART.

5.3 Tsunami characterization from detection curves

In the work by Beltrami (2011), the problem of real-time characterization of a detected tsunami wave was presented and it was shown that Mofjeld's algorithm, despite being very efficient in detecting, it is not able to characterize amplitude or period of the wave correctly. In fact, despite the detection curves showing recognizable oscillations in correspondence of the tsunami, the two waveshapes do not correspond for most of the tsunami period band. However, modern tsunami early warning systems greatly benefit from a correct estimation of wave amplitude and period, or even from the full waveform. The abilities of the techniques have already been addressed for a limited number of examples for Mofjeld's algorithm (Beltrami, 2011), EOF detiding (Tolkova, 2009; Tolkova, 2010) and TDA (Chierici et al., 2017). Obviously, no such analysis has been carried out for the FIF-based techniques. In general, the analysis has to be made on a case by case basis, but there are some generalizable characteristics that we are going to illustrate through selected examples from the analyzed dataset.

The characteristics that we want to investigate are

1. the behaviour of signals where no earthquake or tsunami is present;
2. if the oscillations in the detection curves reproduce correctly the tsunami waveform;
3. if the seismic wave is filtered and if it can be separated from the tsunami waves.

While the first point has been discussed in detail in the analysis of background signals (Chapter 4), an interesting behaviour is observed in the signals from DARTs 46404 and 46407 during event Su3. In both cases, the raw data appear to be much noisier than what is observed in background conditions. However, no recognizable signs of neither seismic nor tsunami waves is observed. In the detection curves in Fig. 5.12 and 5.13 obtained by Mofjeld's algorithm, this large amplitude noise is quite evident. Of course, such noise causes higher amplitude detection curves also for EOF detiding, which does not filter high frequencies, but also on TDA and the FIF-based technique, due to already discussed boundary effects. The situation, probably due to oceanographic effects, would lower the signal to noise ratio in presence of a tsunami, weakening the ability to detect and characterize it. These are the only cases where Mofjeld's algorithm triggers a false detection if a threshold $T = 2$ cm is adopted. These signals also exemplify the effect of long period residual oscillations that are observed in EOF and TDA detection curves, which would trigger a detection in both cases. The TDA

detection curve for DART 46404 shows the typical behaviour we have whenever a low amount of tidal coefficients is computed and the residual has an amplitude of a few centimeters. We also note that these cases are the only false detections produced by Mofjeld's algorithm for $T = 2$ cm.

The influence of these large tidal residuals on the detection of tsunamis is shown in the next examples. In Fig. 5.14, detection curves for event KJ1, DART 46411 are shown. All techniques reproduce the tsunami wavetrain quite accurately, except for Mofjeld's algorithm as expected, but show some unexpected behaviour after a few hours. EOF's detection curve increases about 2 h after the end of the tsunami wavetrain. For a threshold $T \leq 2$ cm, that would trigger a detection, this may lead to an overestimation of the tsunami duration and possibly too conservative measures for risk management in coastal settlements. A similar occurrence is observed for TDA a bit later.

Similarly, in the case of event NH2, DART 51425, the observed detections for $T = 2$ cm are quite diverse between techniques. While the postprocessed waveform reaches the threshold only at the peak of the first crest, Mofjeld's algorithm, EOF detiding and TDA all show multiple detections. Mofjeld's algorithm triggers a detection for the passage of Rayleigh waves and then, when the tsunami passes, a detection is triggered for the first trough, but not for the crest. In the case of EOF detiding, both the earthquake and first tsunami trough are detected, though, due to a residual increasing trend, later oscillations also exceed T . TDA and the FIF-based method both show signs of the earthquake oscillations, but it is greatly filtered out in both cases. While TDA triggers a detection for the first trough, FIF produces the results we expect from the postprocessed wave, i.e. a detection triggered for the first crest and nowhere else.

A more extreme example of this is observed in Fig. 5.16, relative to the same event (NH2), DART 52402. In this case, a detection is not expected for thresholds $T \geq 2$ cm. The FIF-based method is the only one giving the expected result. The passage of the earthquake triggers a detection for Mofjeld's algorithm and EOF detiding. Residual trends trigger detections during the passage of the tsunami wavetrain: a later detection for EOF and an immediate detection for TDA. A similar case is presented for event KT1, DART 51407, shown in Fig. 5.17, where TDA is the only technique triggering a detection, though the FIF-based method also over estimates the amplitude of the largest minimum. Lastly, the example in Fig. 5.18, relative to event KT2, DART 32412, shows good agreement between the tsunami waveform and detection curves in all cases, although all techniques overestimate the largest trough and trigger a

detection.

While all the aspects we just described can be easily dealt with by data processing, we remind that the techniques presented here are supposed to be automated and used in real time. Thus, a detection like the one in the TDA's detection curve in Fig. 5.16, while corresponding to the tsunami, gives an overestimation of the wave amplitude, due to the fact that the long terms trend is not clearly recognizable without successive data.

For larger events, the tsunami waveshape is better recognizable in the detection curves. In Fig. 5.19, the general properties of the techniques are well represented. Mofjeld's algorithm triggers at the correct time, but it deforms the tsunami quite significantly and it shows a way larger trough than what is observed. The other techniques are in better agreement with the postprocessed signal. EOF and TDA have a slight overestimation of the first trough, while FIF's detection curve is much closer to the expected signal. In Fig. 5.20, another characteristic is shown. While the tsunami waveform starts with a very slow decay before the first crest, a typical sign of dispersive effects that we observe a large distance from the source, all the detection techniques tend to underestimate this initial descent. Mofjeld's algorithm and the FIF-based method eliminate it completely.

The filtering abilities of TDA and FIF in case of seismic waves has already been highlighted by the much lower earthquake detection rates and by examples in Fig. 5.15, 5.16, 5.17, 5.19 and 5.20. We add two more examples, relative to event KT2, DARTs 51425 and 51426, shown in Fig. 5.21 and 5.22: they are recorded very close to the generating source, thus there is a significant overlap of Rayleigh waves and the tsunami waveshape. The technique may be differentiated based on two considerations. The first is that the amplitude of earthquake oscillations is reduced by an order of magnitude in TDA and FIF with respect to Mofjeld's algorithm and EOF detiding. The second is that in the formers, the tsunami waveshape is easily recognizable in the detection curve, while the seismic coda dominates the latter's curves. In a real time monitoring context, an easily recognizable tsunami waveshape is fundamental to issue precise warning and the detection curves produced by TDA and the FIF-based method are much more interpretable by an expert than the others, unless additional processing operations are performed. Thus, even in the case where a detection is triggered by an earthquake, the ability of a technique to filter high frequency oscillation is of great importance for the characterization of amplitude and period of the travelling tsunami.

Despite the numerous advantages offered by the FIF-based detection method,

there are some instances in which the detection curve does not represent tsunami wave accurately and they present sharp discontinuities. This is a manifestation of the non consistency of detection curves, discussed in previous chapters, caused by the fact that tidal fitting and the IMFs do not vary continuously from one time step to the next. However, an alternative way to estimate the tsunami waveform can be found in the decomposition steps. In particular, the detection method includes, as one of the steps, the extraction of the full tsunami component for the 3h segment of data, which can therefore be used as a representation of the tsunami wavefield in place of the detection curve. One possible issue with this approach would be that it involves the transmission of 3h of data, i.e. 720 measurements, instead of the single one used for detection. On the other hand, the estimation of the tsunami amplitude need not be carried out at each time step, so the extraction of the full tsunami components can be performed at fixed time after detection or on demand. In fact, this is what is needed in many forecasting methods, such as real time source inversion (Percival et al., 2011), probabilistic forecasting (Selva et al., 2021b) or data assimilation procedures (Wang et al., 2021).

In Fig. 5.23, 5.24 and 5.25, three examples are shown. In all cases, the detection curve is compared to the postprocessed waveform (Davies, 2019) and with the tsunami components extracted at time of detection t_0 , at $t_0 + 1$ h and $t_0 + 2$ h. We can see that the detection curves overestimate the maximum amplitude oscillations, i.e. the first crest in the case for event SA3, DART 46403, and the first trough for event KJ2, DART 52405 and event SA3, DART 21413, since the discontinuities appear to be related to strong gradient parts of the tsunami wave. On the other hand, the tsunami components seems to reproduce quite well the corresponding portion of the postprocessed waveform. Not only these components do not present any jump discontinuity as the detection curve does, but they also capture the initial descent of the wavefield that precedes the first crest in Fig. 5.23 and 5.24, which the detection curve was not able to reproduce (see previous examples). We point out at last, that using a tide removal method different from polynomial fitting in the FIF-based detection may alleviate the problem by producing tidal components that vary with more continuity between successive time steps. However, the computational efficiency of polynomial fitting makes it still a convenient choice.

5.4 Conclusions

In this chapter, the application of the four tsunami detection methods to pressure records from past tsunami events is presented and discussed. One first contribution consists in the methodology for calibrating the techniques, i.e. for choosing the detection threshold. This calibration methodology starts from building a catalog of events of different amplitude and origin location for which instrumental records are available. The records from every instrument active during the generation and propagation of the tsunami are used to build a testing dataset. Then, each detection technique is applied to the entire dataset with different detection thresholds. At last the optimal threshold is selected according to a chosen criterion. Since it uses events with different locations and recorded by different instruments, we are able to build large and comprehensive datasets made entirely of real tsunami records without the need of simulated tsunami waves. This is particularly useful for detection techniques which do not require an instrument specific calibration, such as the ones used in this work.

The event catalog used in the chapter consists of the tsunami records from NOAA's Unassessed Ocean Bottom Pressure data for the seismogenic tsunamis used by Davies (2019). The events cover the subduction zones around the Pacific Ocean and have been generated by earthquakes with magnitude between 7.7 and 9.1. The resulting dataset includes more than 400 records, which differ in background conditions, presence of seismic oscillations, tsunami amplitude and instrument location.

The criteria to choose the optimal detection threshold proposed in the chapter are based on optimizing simple metrics, that we called *detection scores*, that takes into account the number of tsunami, earthquake and false detections.

The techniques behaviour with varying detection thresholds differ quite substantially from each other. TDA is the technique that most critically depends on the quantity and quality of previous data available. In fact, it shows false detection rates consistently higher than other techniques, due to the presence of signals with large tidal residuals. In particular it is the only technique where the difference between tsunami and false detections increases monotonically with the detection threshold, i.e. there is no threshold that optimizes the detection score θ_1 . In the case of EOF detiding, we suggest that no threshold below 2.5 cm should be chosen, given the exponential increase in false detections below that. This is consistent with the properties of the technique presented in the original works by Tolкова (2010, 2009).

False detection rates are essentially related to the amplitude of residuals,

as we studied in the previous chapter. As expected, Mofjeld's algorithm and the FIF-based method have the lowest false detection rates, in line with the amplitude of residuals observed for background signals in Chapter 4.

A big difference between the technique is their ability to detect or filter out seismic shaking. As expected, the techniques with no high frequencies filtering capabilities, Mofjeld's algorithm and EOF detiding, show very large earthquake detection rates. If minimization of seismic shaking is a criterion of interest, this techniques should be avoided.

In general, the FIF-based detection methods performs quite well according to many different metrics. It filters out seismic shaking as efficiently as TDA, but it shows much smaller false detection rates. In fact, no false detection is triggered for thresholds equal or bigger than 2.5 cm and for a threshold of 2.0 cm the only false detections are related to discontinuities in the records. Thus, if we want to maximize tsunami detections while minimizing earthquake and false detections, FIF-based detection seems to be the best technique among the ones tested in this work.

At last, the technique also shows good abilities in accurately reproducing the tsunami waveshape, making the technique suitable for integration with tsunami data assimilation technologies. For a limited number of signals tested, the detection curves produced by FIF-based detection do not accurately represent the tsunami waveshape. In those cases, we propose that the entire tsunami component at fixed time after detection can be used for the determination of period and amplitude of the tsunami waves, or for applications in data assimilation systems.

Appendix - Plots

In the following, the plots relative to the previous sections are given. In Fig. 5.12 to 5.22, we show detection curves for exemplary signals from the analysed dataset. In each case, horizontal lines $y = \pm 2$ cm are given as reference for detection analysis. Wherever available, the postprocessed waveform from Davies (2019) is plotted, as well. Fig. 5.23 to 5.25 have examples of full tsunami components extracted during the propagation of the tsunami using the FIF-based techniques and are compared with both the detection curves and the postprocessed waveforms.

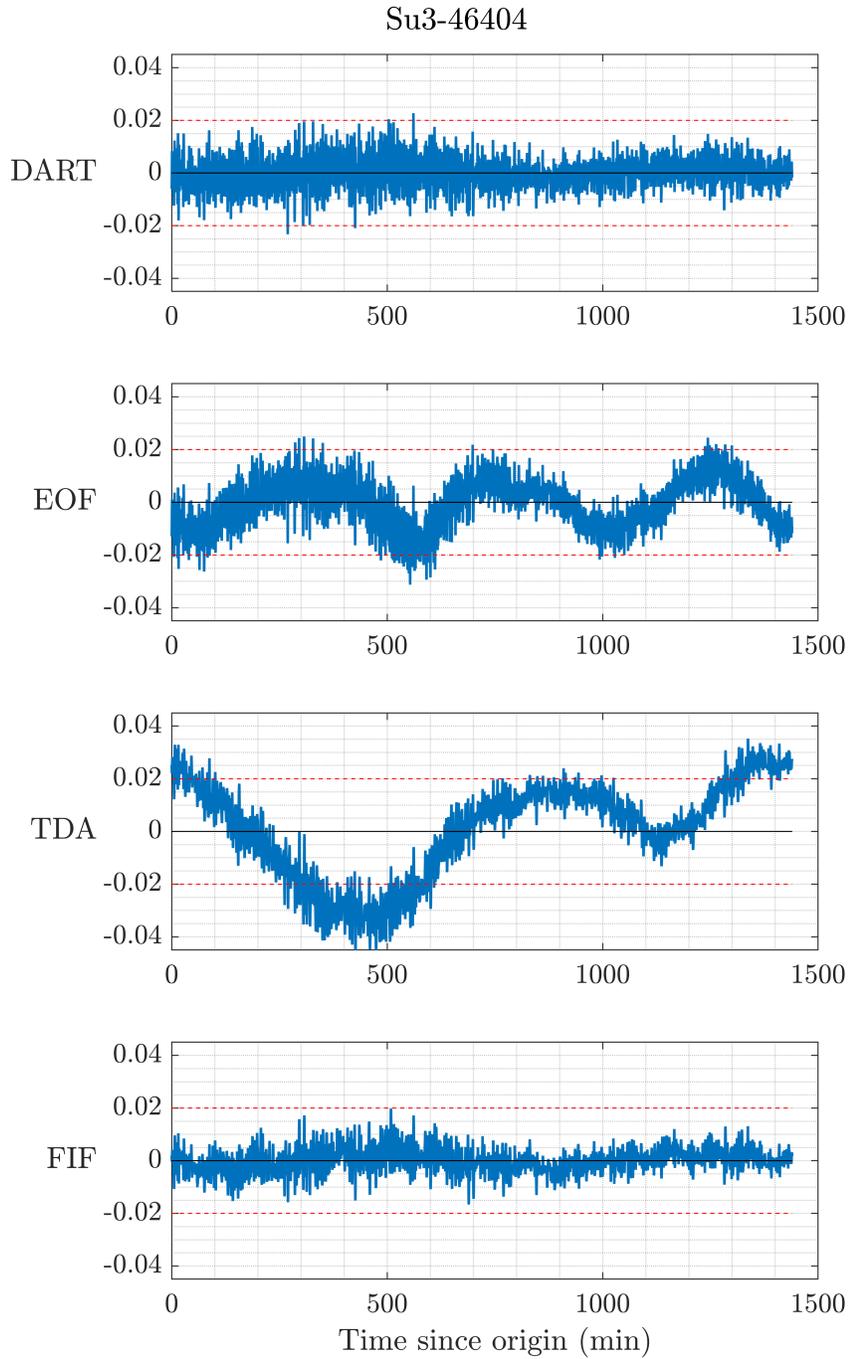


Figure 5.12: Example of detection curves for event Su3, DART 46404. The amplitudes on the vertical axis are measured in centimeters. Here, Mofjeld's algorithm is indicated simply as DART. Data availability for TDA tidal modelling 5.5 months.

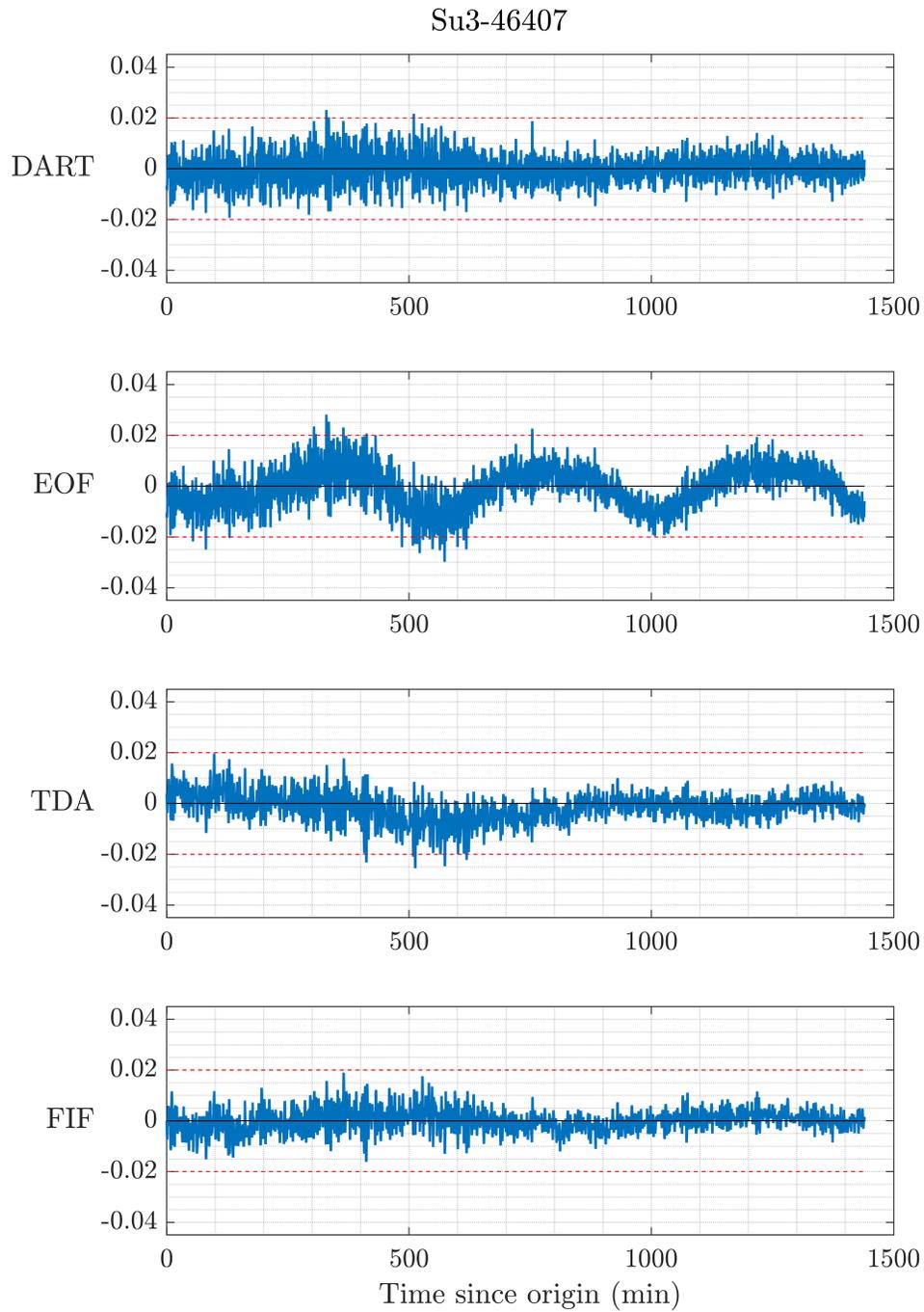


Figure 5.13: Example of detection curves for event Su3, DART 46404. The amplitudes on the vertical axis are measured in centimeters. Here, Mofjeld's algorithm is indicated simply as DART. Data availability for TDA tidal modelling 1 year 4 months.

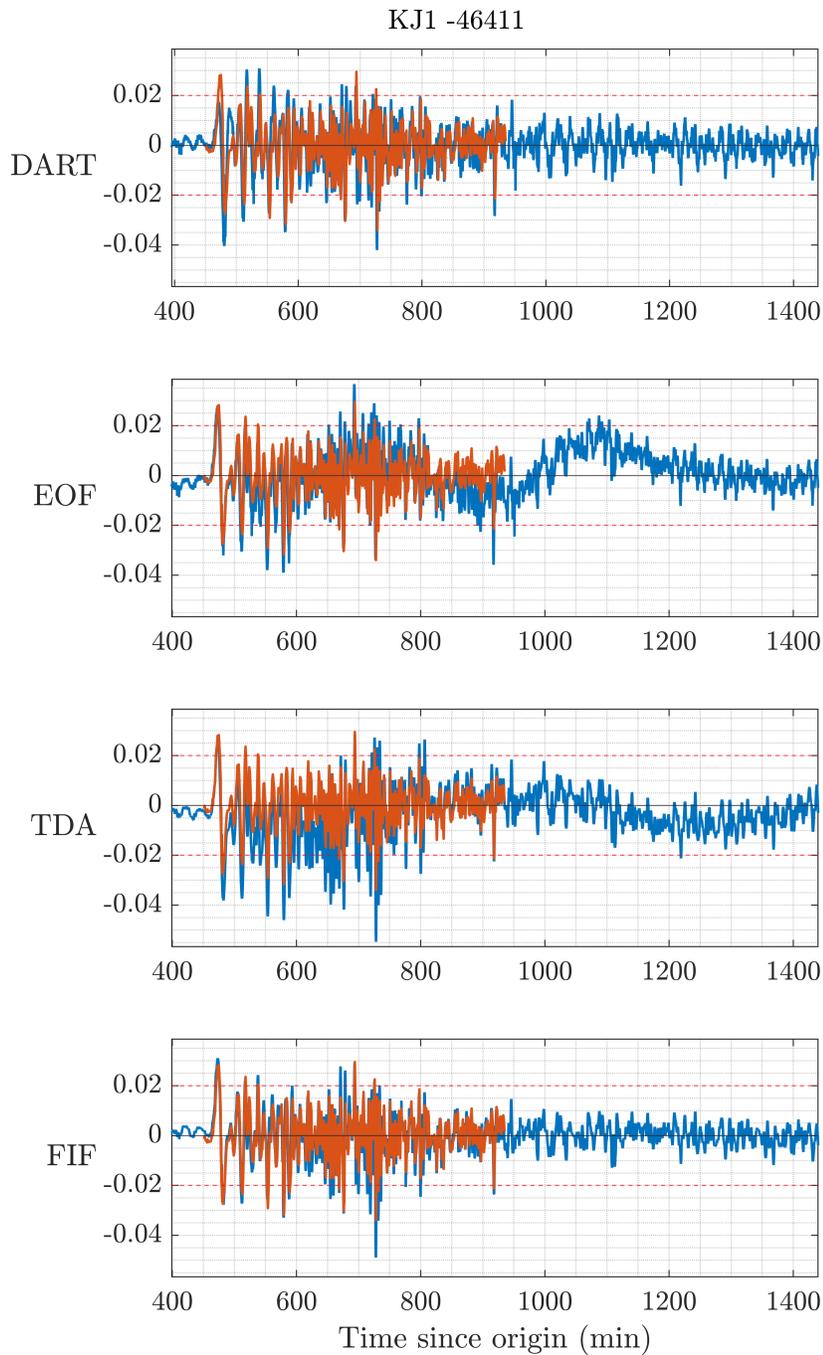


Figure 5.14: Example of detection curves for event KJ1, DART 46411. The amplitudes on the vertical axis are measured in centimeters. Here, Mofjeld's algorithm is indicated simply as DART. The orange curve is the postprocessed waveform from Davies (2019). Data availability for TDA tidal modelling 1 year 1 month.

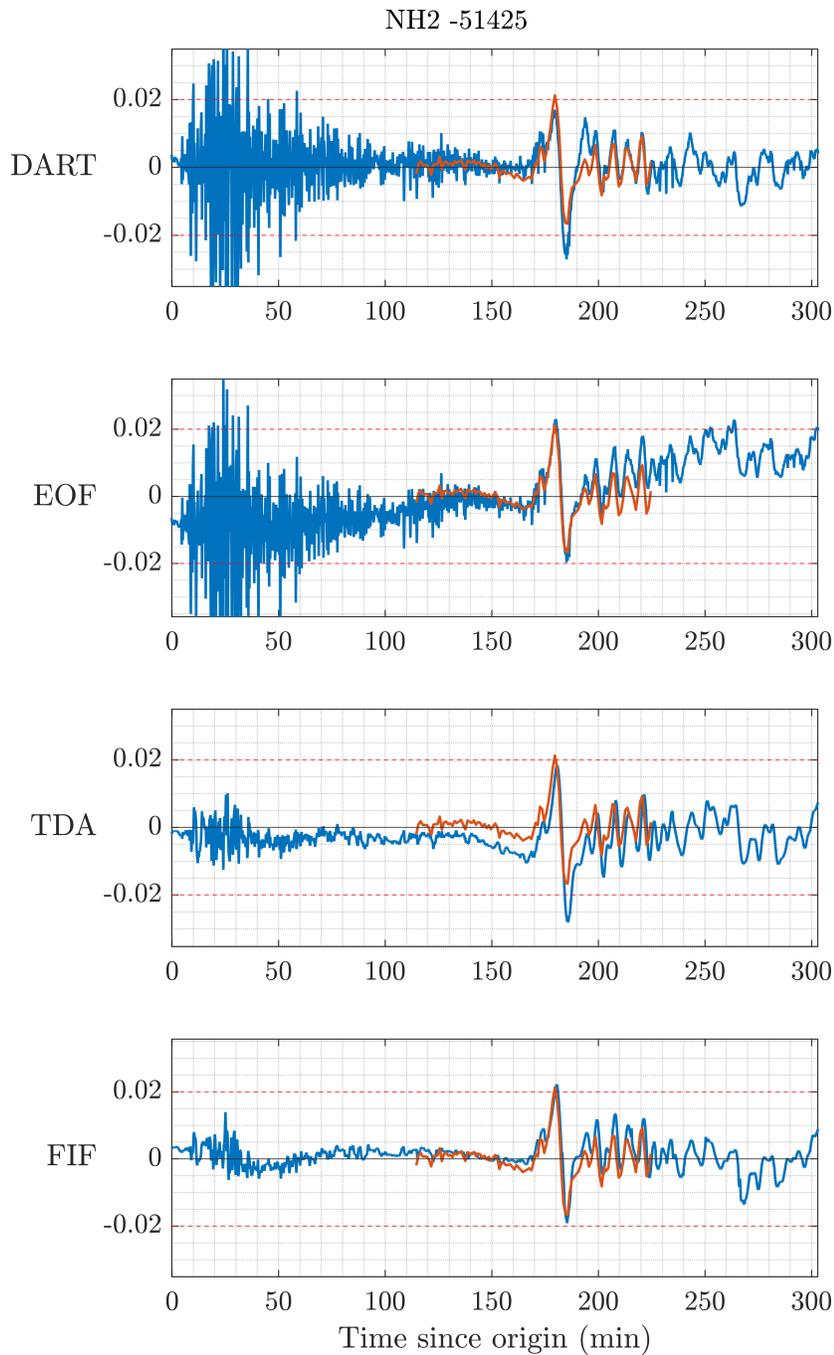


Figure 5.15: Example of detection curves for event NH2, DART 51425. The amplitudes on the vertical axis are measured in centimeters. Here, Mofjeld’s algorithm is indicated simply as DART. The orange curve is the postprocessed waveform from Davies (2019). Data availability for TDA tidal modelling 1 year 5 months.

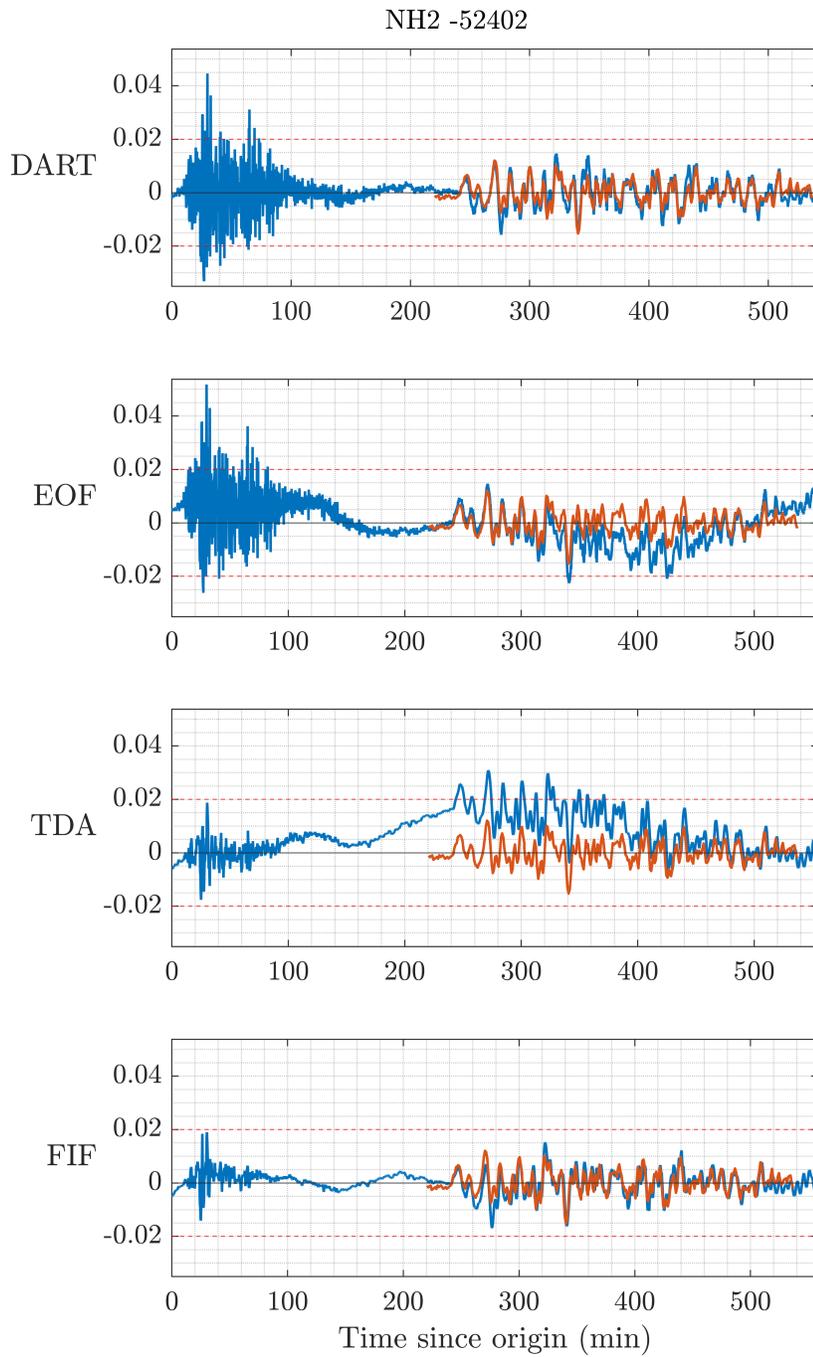


Figure 5.16: Example of detection curves for event NH2, DART 52402. The amplitudes on the vertical axis are measured in centimeters. Here, Mofjeld’s algorithm is indicated simply as DART. The orange curve is the postprocessed waveform from Davies (2019). Data availability for TDA tidal modelling 6 months.

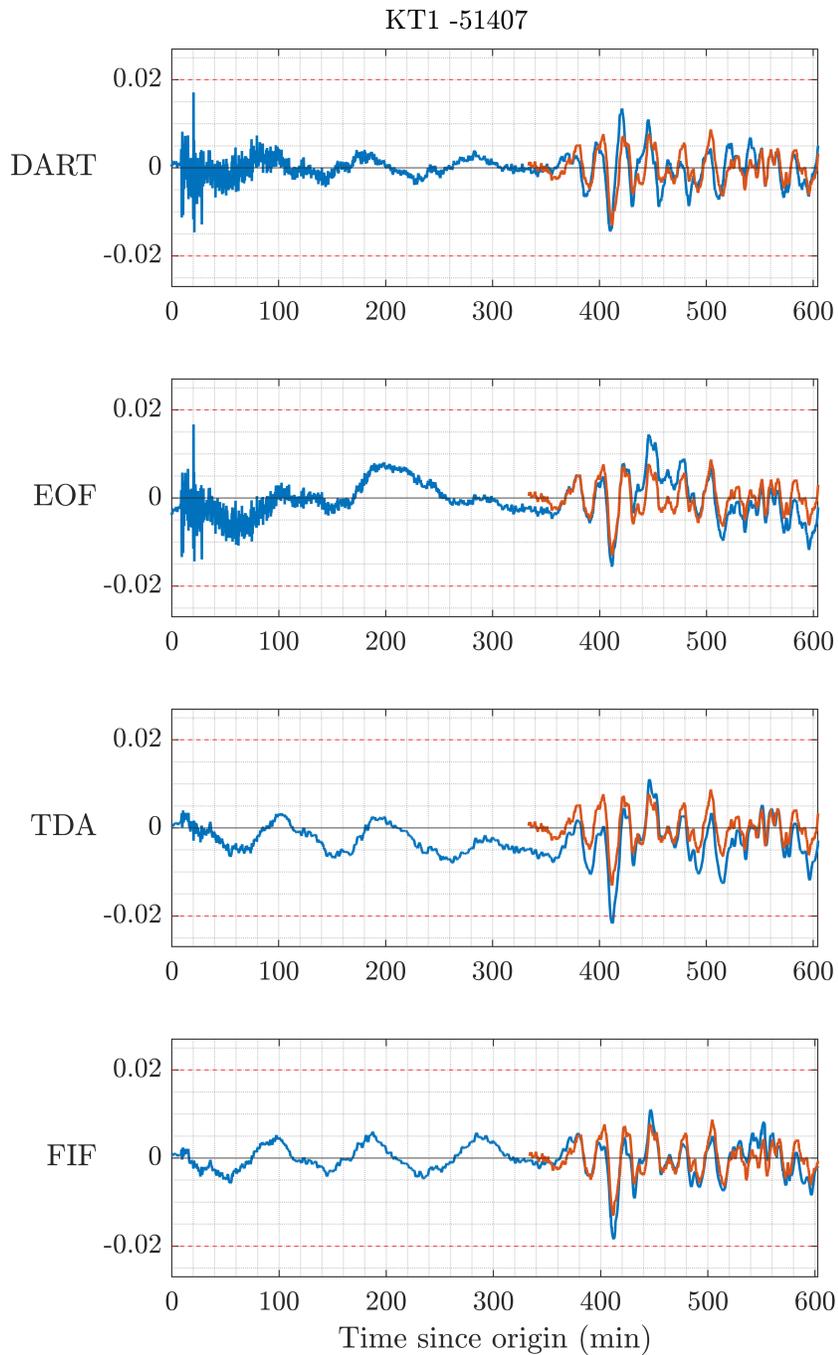


Figure 5.17: Example of detection curves for event KT1, DART 51407. The amplitudes on the vertical axis are measured in centimeters. Here, Mofjeld's algorithm is indicated simply as DART. The orange curve is the postprocessed waveform from Davies (2019). Data availability for TDA tidal modelling 1 year 1 month.

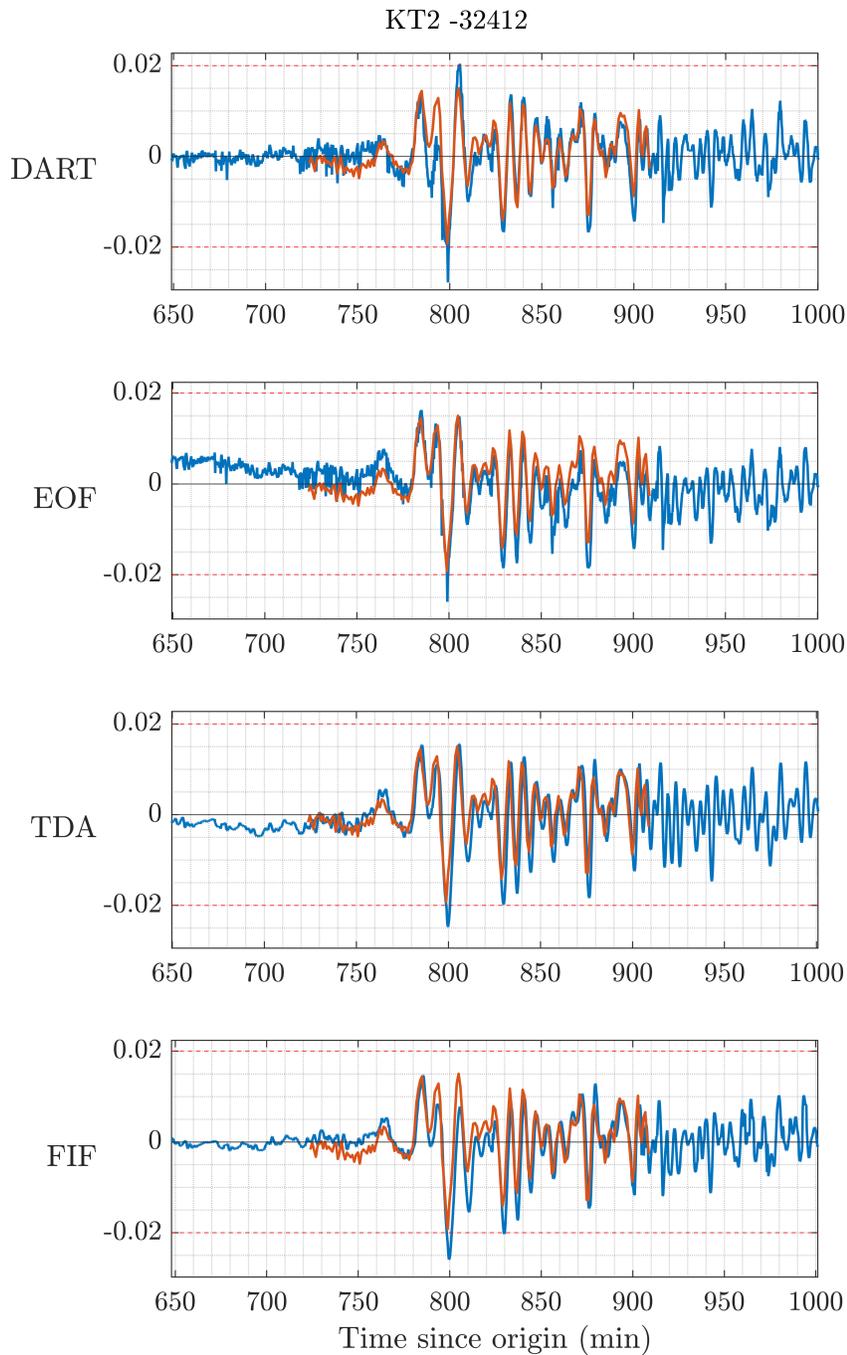


Figure 5.18: Example of detection curves for event KT2, DART 32412. The amplitudes on the vertical axis are measured in centimeters. Here, Mofjeld's algorithm is indicated simply as DART. The orange curve is the postprocessed waveform from Davies (2019). Data availability for TDA tidal modelling 1 year 10 months.

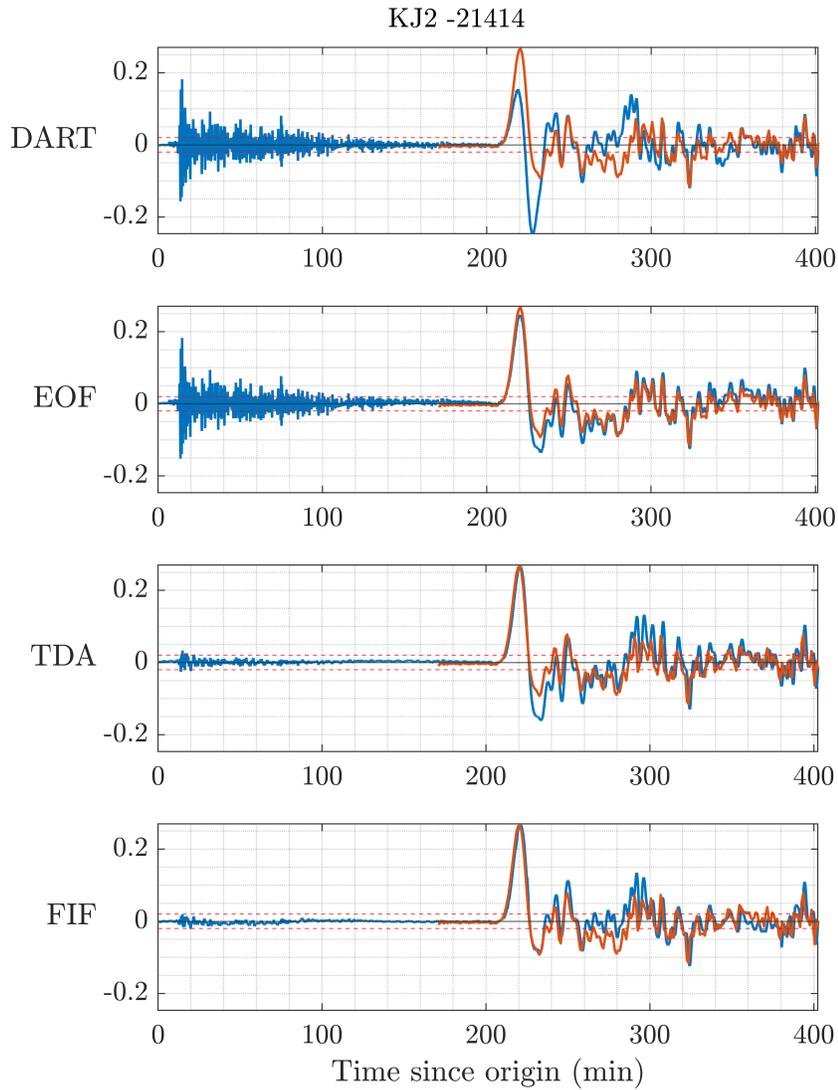


Figure 5.19: Example of detection curves for event KJ2, DART 21414. The amplitudes on the vertical axis are measured in centimeters. Here, Mofjeld's algorithm is indicated simply as DART. The orange curve is the postprocessed waveform from Davies (2019). Data availability for TDA tidal modelling 9 months.

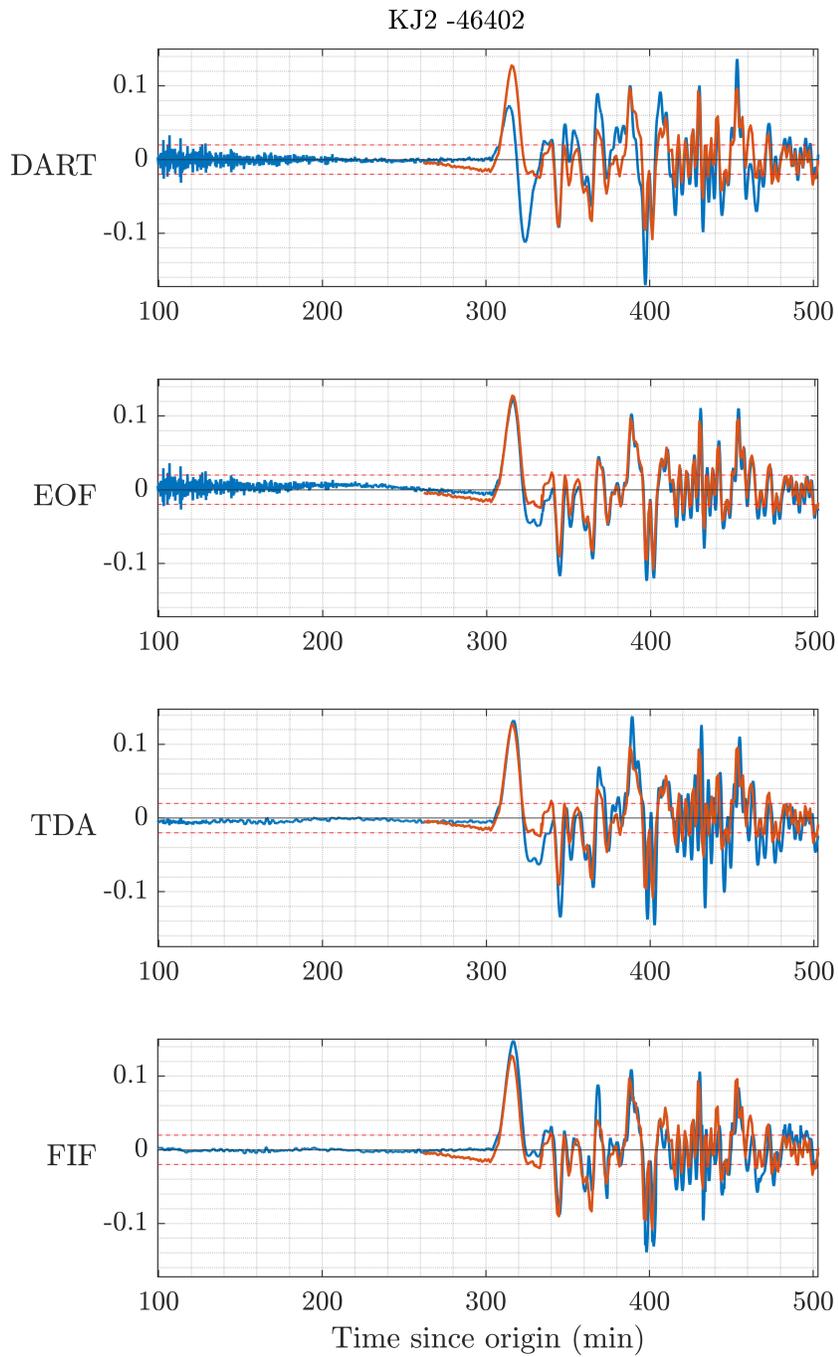


Figure 5.20: Example of detection curves for event KJ2, DART 46402. The amplitudes on the vertical axis are measured in centimeters. Here, Mofjeld's algorithm is indicated simply as DART. The orange curve is the postprocessed waveform from Davies (2019). Data availability for TDA tidal modelling 9.5 months.

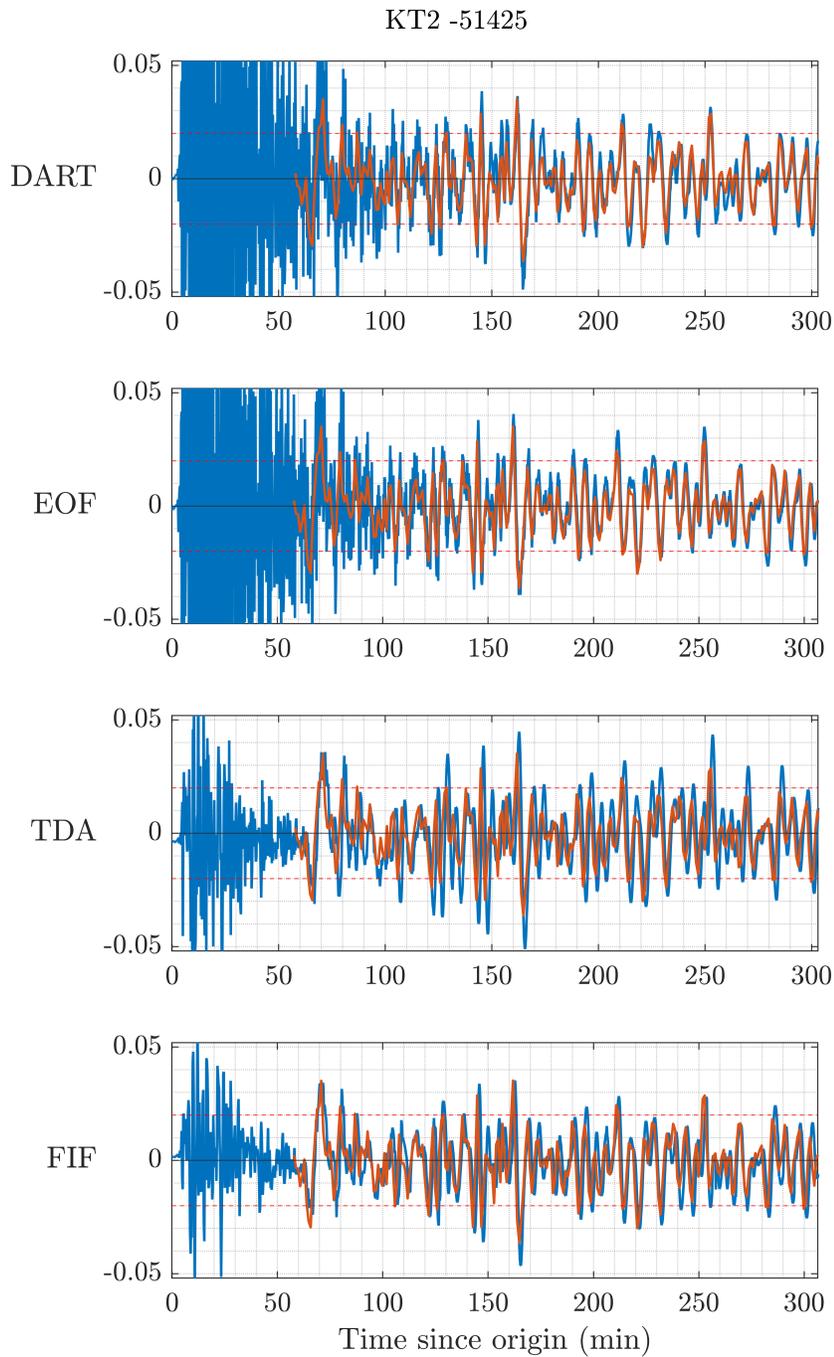


Figure 5.21: Example of detection curves for event KT2, DART 51425. The amplitudes on the vertical axis are measured in centimeters. Here, Mofjeld's algorithm is indicated simply as DART. The orange curve is the postprocessed waveform from Davies (2019). Data availability for TDA tidal modelling 1 year 7 months.

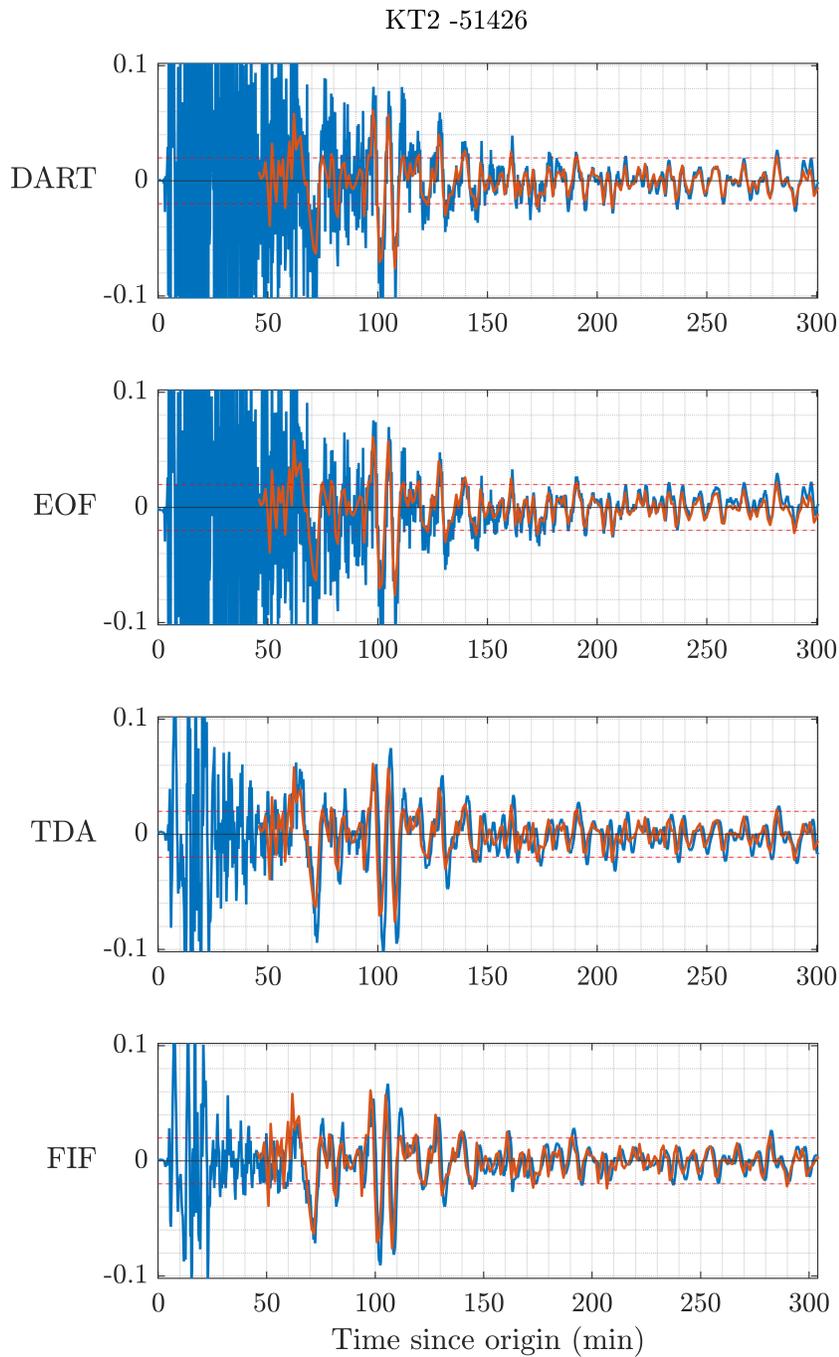


Figure 5.22: Example of detection curves for event KT2, DART 51426. The amplitudes on the vertical axis are measured in centimeters. Here, Mofjeld’s algorithm is indicated simply as DART. The orange curve is the postprocessed waveform from Davies (2019). Data availability for TDA tidal modelling 1 year 7 months.

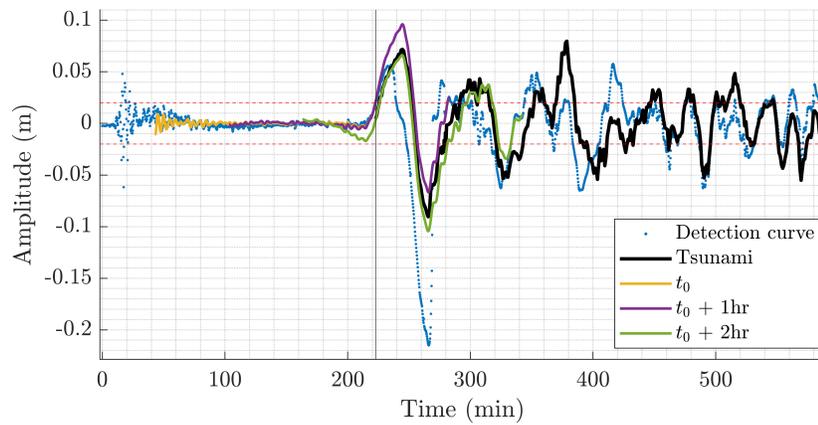


Figure 5.23: Comparison between tsunami components at different times, detection curve, obtained by the FIF-based method, and the tsunami waveform for event KJ2, DART 52405. The black curve indicated as “tsunami” is the post-processed waveform. The vertical line indicates the detection time from Davies (2019). $t_0 = 3\text{ h }43\text{ min }30\text{ s}$ after origin time. The red dashed horizontal lines represent the $T = 2\text{ cm}$ detection threshold. The coloured lines represent the tsunami components obtained at t_0 , $t_0 + 1\text{ h}$ and $t_0 + 2\text{ h}$.

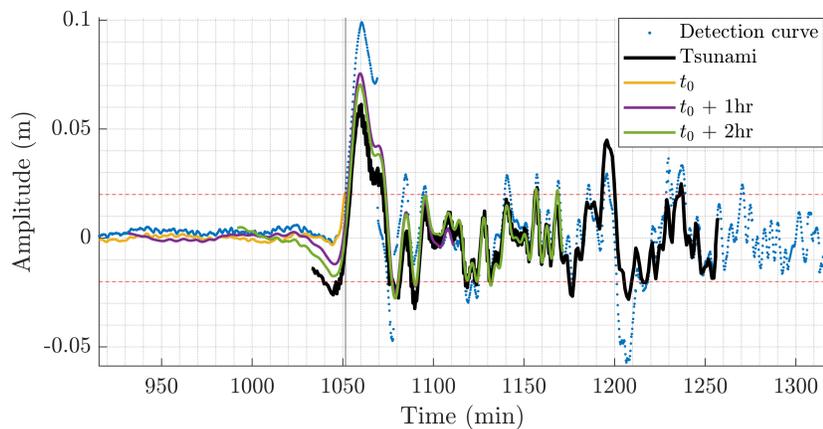


Figure 5.24: Comparison between tsunami components at different times, detection curve, obtained by the FIF-based method, and the tsunami waveform for event SA3, DART 46403. The black curve indicated as “tsunami” is the post-processed waveform. The vertical line indicates the detection time from Davies (2019). The vertical line indicates the detection time $t_0 = 17\text{ h }31\text{ min }45\text{ s}$ after origin time. The red dashed horizontal lines represent the $T = 2\text{ cm}$ detection threshold. The coloured lines represent the tsunami components obtained at t_0 , $t_0 + 1\text{ h}$ and $t_0 + 2\text{ h}$.

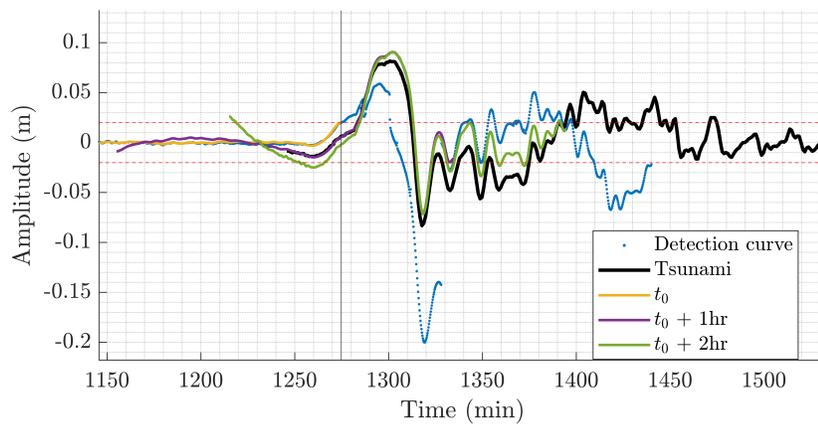


Figure 5.25: Comparison between tsunami components at different times, detection curve, obtained by the FIF-based method, and the tsunami waveform for event SA3, DART 21413. The vertical line indicates the detection time $t_0 = 21$ h 15 min after origin time. The red dashed horizontal lines represent the $T = 2$ cm detection threshold. The coloured lines represent the tsunami components obtained at t_0 , $t_0 + 1$ h and $t_0 + 2$ h

Conclusions

In this thesis, we presented four tsunami detection methods, namely Mofjeld’s algorithm, tide removal with Empirical Orthogonal Functions, the Tsunami Detection Algorithm (TDA) and a newly developed technique based on the Fast Iterative Filtering (FIF) and the IMFogram algorithms. The latter are data driven signal analysis techniques that have been applied to tsunami signals here for the first time. In particular, the FIF technique allows to decompose a signal into elementary oscillatory components, called Intrinsic Mode Functions (IMFs), from which the IMFogram algorithm can compute a time-frequency representation. By choosing IMFs based on their frequency content computed with the IMFogram, the two techniques combined can act as bandpass filters. We applied the techniques to selected examples, in particular a signal from the 02/05/2020 Crete tsunami and multiple records from the 06/02/2023 small tsunami off the coasts of Turkey. It is shown how the techniques give results consistent with classical techniques, such as Fourier analysis and wavelet transforms, but with some added benefits, among which:

- they are completely data driven;
- they can perform multiple operations, such as denoising, tide removal and bandpass filtering, all at once; the same operations are carried out separately with different ad hoc techniques in many studies;
- their use as bandpass filters is very robust with respect to the choice of corner frequencies, since each IMF is contained within a narrow frequency window.

We conclude that FIF and IMFogram can be very valuable and powerful tools in the context of tsunami data analysis.

The detection techniques have been tested against two datasets built from data from NOAA’s Deep-ocean Assessment and Reporting (DART). The first

dataset is made of several month-long time series that include only tides and noise, to analyse each technique's response in absence of transient oscillations. It is shown that Mofjeld's algorithm gives detection curves with the narrowest distribution around zero and the FIF-based detection gives close results. The other techniques need higher amplitude detection thresholds to have zero false detection on the background signals.

The second dataset includes the day-long signals acquired during past tsunamis from all instruments active at the time of their occurrence. The list of events is the tsunami catalogue used by Davies (2019) and includes events differing in magnitude and location. Using the signals from all the instruments active at the time of occurrence of each event allowed us to build a diverse dataset that includes background signals and records with seismic shaking and/or tsunami oscillations. The detection capabilities have been quantified through detection scores that takes into account tsunami detections, earthquake detections and false detections; the scores have been computed for different amplitude detection thresholds. Criteria about the choice of an optimal detection threshold for each technique have been discussed and presented.

Mofjeld's algorithm and the FIF-based method have the lowest optimal detection thresholds. In particular, for the latter we propose the use of a threshold $T = 1.5$ cm to optimize the detection of tsunami waves, while penalizing earthquake and false detections. EOF detiding also allows for an optimal threshold choice, though higher than the one for FIF, since its detection curves consistently show tide residual amplitude of up to 2 cm. TDA gives very good results on many signals, but shows difficulties in network wide applications. In fact, its performance is highly dependent on the quality of the precomputed set of tidal coefficients. With the methodology presented in this work, we are not able to find an optimal detection threshold for TDA and future applications of the technique should consider using a different tide removal method.

The ability of the FIF-based method to characterize correctly the tsunami waveshape has been highlighted and discussed. In particular, it has been shown that the detection curves reproduce the tsunami waveshape accurately most of the time and the influence of seismic shaking is greatly reduced. In particular, it has been shown that, even in cases where the earthquake triggers a detection, seismic waves are greatly reduced in amplitude with respect to the original signal and its effects on the tsunami waveshape are also greatly reduced. In the few cases where the tsunami is not accurately reproduced in the detection curves, it is shown that the tsunami component extracted as a byproduct of the detection method can be used to characterize amplitude and period of the wave.

The transmission of the full tsunami component can be used once a detection is triggered in the context of early warning.

The newly proposed FIF-based detection method is thus able to detect tsunamis in real time from continuously recording OBPGs, while reducing the influence of noise, tides and seismic shaking more than the other analyzed techniques. Applications of the techniques to different datasets is foreseen, especially in view of the planned installations of OBPGs in the Ionian Sea by the Italian Tsunami Alert Center (CAT-INGV, Amato et al., 2021). Furthermore, there are no theoretical aspects of the techniques that limit the applications to OBPG only. Thus, testing on other type of instruments, such as coastal tide gauges, is also planned. Other possible future applications regard site-specific tsunami alert system, such as the one installed at the island of Stromboli (Selva et al., 2021a) or possibly to the recent SMART cable installed in the Mediterranean (Marinero et al., 2024). At last, we should note that, despite the very good results obtained with the FIF-based method, we should be aware of its drawbacks, namely the computational costs and the possibility of deforming tsunami waveforms, discussed at the end of Chapter 5. For this second point, the solution could probably be found in equivalent algorithm with less errors introduced at the boundary during the decomposition, which is one of the lines of research we plan to pursue.

Acknowledgements

No man travels a road alone. No man crafts his achievements by himself. This has been particularly true in my academic path since the beginning.

Firstly, I would like to thank the members of the Tsunami Research Group of the University of Bologna, with which I had the privilege to work during my PhD. In particular, I would like to thank prof. Alberto Armigliato with which I will be forever indebted for being always supportive and for the immense amount of knowledge he was able to pass me regarding tsunamis.

I would like to thank my PhD colleagues, Anastasia, Martina, Riccardo and Letizia, for being a constant happy presence in our days at work. I would like to thank many professors and researchers with which I had the pleasure to interact, in particular Dr. Stefano Lorito and Dr. Fabrizio Romano, for the great help they gave me to carry out the work in this thesis, and prof. Ira Didenkulova for the warm hospitality during my months in Oslo.

I would like to thank the friends that helped and lived these years with me, who are too many to name.

Lastly, I want to thank my girlfriend, Francesca, my parents, Alessio and Emanuela, and my grandmother, Rosella. Without them, nothing would have been possible.

Bibliography

- Amato, A., A. Avallone, R. Basili, F. Bernardi, B. Brizuela, L. Graziani, A. Herrero, M. C. Lorenzino, S. Lorito, F. M. Mele, et al. (2021). “From seismic monitoring to tsunami warning in the Mediterranean Sea”. In: *Seismological Research Letters* 92.3, pp. 1796–1816.
- Aranguiz, R., P. Catalán, C. Cecioni, G. Bellotti, P. Henriquez, and J. González (2019). “Tsunami resonance and spatial pattern of natural oscillation modes with multiple resonators”. In: *Journal of Geophysical Research: Oceans* 124.11, pp. 7797–7816.
- Astafyeva, E. (2019). “Ionospheric detection of natural hazards”. In: *Reviews of Geophysics* 57.4, pp. 1265–1288.
- Baglione, E., S. Lorito, A. Piatanesi, F. Romano, R. Basili, B. Brizuela, R. Tonini, M. Volpe, H. B. Bayraktar, and A. Amato (2021). “Characterization of fault plane and coseismic slip for the 2 May 2020, M w 6.6 Cretan Passage earthquake from tide gauge tsunami data and moment tensor solutions”. In: *Natural Hazards and Earth System Sciences* 21.12, pp. 3713–3730.
- Barbe, P., A. Cicone, W. S. Li, and H. Zhou (2020). “Time-frequency representation of nonstationary signals: the IMFogram”. In: *arXiv preprint arXiv:2011.14209*.
- Basili, R., B. Brizuela, A. Herrero, S. Iqbal, S. Lorito, F. E. Maesano, S. Murphy, P. Perfetti, F. Romano, A. Scala, et al. (2021). “The making of the NEAM tsunami hazard model 2018 (NEAMTHM18)”. In: *Frontiers in Earth Science* 8, p. 753.
- Battista, B. M., C. Knapp, T. McGee, and V. Goebel (2007). “Application of the empirical mode decomposition and Hilbert-Huang transform to seismic reflection data”. In: *Geophysics* 72.2, H29–H37.
- Behrens, J., F. Løvholt, F. Jalayer, S. Lorito, M. A. Salgado-Gálvez, M. Sørensen, S. Abadie, I. Aguirre-Ayerbe, I. Aniel-Quiroga, A. Babeyko, et al. (2021). “Probabilistic tsunami hazard and risk analysis: A review of research gaps”. In: *Frontiers in Earth Science* 9, p. 628772.

- Bellotti, G. and A. Romano (2017). “Wavenumber-frequency analysis of landslide-generated tsunamis at a conical island. Part II: EOF and modal analysis”. In: *Coastal Engineering* 128, pp. 84–91.
- Beltrami, G. M. (2008). “An ANN algorithm for automatic, real-time tsunami detection in deep-sea level measurements”. In: *Ocean Engineering* 35.5-6, pp. 572–587.
- (2011). “Automatic, real-time detection and characterization of tsunamis in deep-sea level measurements”. In: *Ocean Engineering* 38.14-15, pp. 1677–1685.
- Beltrami, G. M. and M. Di Risio (2011). “Algorithms for automatic, real-time tsunami detection in wind-wave measurements Part I: Implementation strategies and basic tests”. In: *Coastal Engineering* 58.11, pp. 1062–1071. DOI: 10.1016/j.coastaleng.2011.06.004.
- Boyd, S. P. and L. Vandenberghe (2004). *Convex optimization*. Cambridge university press.
- Bressan, L. and S. Tinti (2012). “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)”. In: *Natural Hazards and Earth System Sciences* 12.5, pp. 1583–1606.
- (2011). “Structure and performance of a real-time algorithm to detect tsunami or tsunami-like alert conditions based on sea-level records analysis”. In: *Natural Hazards and Earth System Sciences* 11.5, pp. 1499–1521.
- Bressan, L., F. Zaniboni, and S. Tinti (2013). “Calibration of a real-time tsunami detection algorithm for sites with no instrumental tsunami records: application to coastal tide-gauge stations in eastern Sicily, Italy”. In: *Natural Hazards and Earth System Sciences* 13.12, pp. 3129–3144.
- Brigham, E. O. (1988). *The fast Fourier transform and its applications*. Prentice-Hall, Inc.
- Chen, J., B. Heincke, M. Jegen, and M. Moorkamp (2012). “Using empirical mode decomposition to process marine magnetotelluric data”. In: *Geophysical Journal International* 190.1, pp. 293–309.
- Chierici, F., D. Embriaco, and L. Pignagnoli (2017). “A new real-time tsunami detection algorithm”. In: *Journal of Geophysical Research: Oceans* 122.1, pp. 636–652.
- Cicone, A. (2020). “Iterative filtering as a direct method for the decomposition of nonstationary signals”. In: *Numerical Algorithms* 85.3, pp. 811–827.

- Cicone, A. (2019). “Nonstationary signal decomposition for dummies”. In: *Advances in Mathematical Methods and High Performance Computing*. Springer, pp. 69–82.
- Cicone, A. and P. Dell’Acqua (2020). “Study of boundary conditions in the iterative filtering method for the decomposition of nonstationary signals”. In: *Journal of Computational and Applied Mathematics* 373, p. 112248.
- Cicone, A., W. S. Li, and H. Zhou (2022). “New theoretical insights in the decomposition and time-frequency representation of nonstationary signals: the IMFogram algorithm”. In: *arXiv preprint arXiv:2205.15702*.
- Cicone, A., J. Liu, and H. Zhou (2016). “Adaptive local iterative filtering for signal decomposition and instantaneous frequency analysis”. In: *Applied and Computational Harmonic Analysis* 41.2, pp. 384–411.
- Cicone, A., S. Serra-Capizzano, and H. Zhou (2024). “One or two frequencies? the iterative filtering answers”. In: *Applied Mathematics and Computation* 462, p. 128322.
- Cicone, A. and H. Zhou (2021). “Numerical analysis for iterative filtering with new efficient implementations based on FFT”. In: *Numerische Mathematik* 147, pp. 1–28.
- Codiga, D. L. (2011). “Unified tidal analysis and prediction using the UTide Matlab functions”. In.
- Daubechies, I. (1992). *Ten lectures on wavelets*. SIAM.
- Davies, G. (2019). “Tsunami variability from uncalibrated stochastic earthquake models: tests against deep ocean observations 2006–2016”. In: *Geophysical Journal International* 218.3, pp. 1939–1960.
- Di Risio, M. and G. M. Beltrami (2014). “Algorithms for automatic, real-time tsunami detection in wind-wave measurements: using strategies and practical aspects”. In: *Procedia Engineering* 70, pp. 545–554.
- Duputel, Z., L. Rivera, H. Kanamori, G. P. Hayes, B. Hirshorn, and S. Weinstein (2011). “Real-time W phase inversion during the 2011 off the Pacific coast of Tohoku Earthquake”. In: *Earth, planets and space* 63.7, pp. 535–539.
- Ekström, G., M. Nettles, and A. Dziewoński (2012). “The global CMT project 2004–2010: Centroid-moment tensors for 13,017 earthquakes”. In: *Physics of the Earth and Planetary Interiors* 200, pp. 1–9.
- Eva, C. and A. B. Rabinovich (1997). “The February 23, 1887 tsunami recorded on the Ligurian coast, western Mediterranean”. In: *Geophysical Research Letters* 24.17, pp. 2211–2214.

- Foster, J. H., B. A. Brooks, D. Wang, G. S. Carter, and M. A. Merrifield (2012). “Improving tsunami warning using commercial ships”. In: *Geophysical Research Letters* 39.9. DOI: 10.1029/2012GL051367.
- Gómez, J. L. and D. R. Velis (2016). “A simple method inspired by empirical mode decomposition for denoising seismic data”. In: *Geophysics* 81.6, pp. V403–V413.
- Görüm, T., H. Tanyas, F. Karabacak, A. Yılmaz, S. Girgin, K. E. Allstadt, M. L. Süzen, and P. Burgi (2023). “Preliminary documentation of coseismic ground failure triggered by the February 6, 2023 Türkiye earthquake sequence”. In: *Engineering Geology* 327, p. 107315.
- Haralambous, H., M. Guerra, J. Chum, T. G. Verhulst, V. Barta, D. Altadill, C. Cesaroni, I. Galkin, K. Márta, J. Mielich, et al. (2023). “Multi-Instrument Observations of Various Ionospheric Disturbances Caused by the 6 February 2023 Turkey Earthquake”. In: *Journal of Geophysical Research: Space Physics* 128.12, e2023JA031691.
- Heidarzadeh, M. and A. R. Gusman (2021). “Source modeling and spectral analysis of the Crete tsunami of 2nd May 2020 along the Hellenic Subduction Zone, offshore Greece”. In: *Earth, Planets and Space* 73.1, pp. 1–16.
- Heidarzadeh, M., A. R. Gusman, and I. E. Mulia (2023). “The landslide source of the eastern Mediterranean tsunami on 6 February 2023 following the Mw 7.8 Kahramanmaraş (Türkiye) inland earthquake”. In: *Geoscience letters* 10.1, p. 50.
- Heidarzadeh, M., Y. Wang, K. Satake, and I. E. Mulia (2019). “Potential deployment of offshore bottom pressure gauges and adoption of data assimilation for tsunami warning system in the western Mediterranean Sea”. In: *Geoscience Letters* 6.1, pp. 1–12. DOI: 10.1186/s40562-019-0149-8.
- Howe, B. M., B. K. Arbic, J. Aucan, C. R. Barnes, N. Bayliff, N. Becker, R. Butler, L. Doyle, S. Elipot, G. C. Johnson, et al. (2019). “SMART cables for observing the global ocean: science and implementation”. In: *Frontiers in Marine Science* 6, p. 424. DOI: <https://doi.org/10.3389/fmars.2019.00424>.
- Hu, G., K. Satake, L. Li, and P. Du (2023). “Origins of the tsunami following the 2023 Turkey–Syria earthquake”. In: *Geophysical Research Letters* 50.18, e2023GL103997.
- Huang, N. E., Z. Shen, and S. R. Long (1999). “A new view of nonlinear water waves: the Hilbert spectrum”. In: *Annual review of fluid mechanics* 31, pp. 417–457.

- Huang, N. E., Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu (1998). “The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis”. In: *Proceedings of the Royal Society of London. Series A: mathematical, physical and engineering sciences* 454.1971, pp. 903–995.
- Imamura, F., S. P. Boret, A. Suppasri, and A. Muhari (2019). “Recent occurrences of serious tsunami damage and the future challenges of tsunami disaster risk reduction”. In: *Progress in Disaster Science* 1, p. 100009.
- Inbal, A. (2023). “Dynamic triggering along the Dead Sea transform uncovers tremor source physics”. In.
- Inbal, A., A. Ziv, I. Lior, R. N. Nof, and A. S. Eisermann (2023). “Non-Triggering and Then Triggering of a Repeating Aftershock Sequence in the Dead Sea by the 2023 Kahramanmaraş Earthquake Pair: Implications for the Physics of Remote Delayed Aftershocks”. In: *Geophysical Research Letters* 50.18, e2023GL104908.
- Kalligeris, N., V. Skanavis, N. Melis, E. Okal, A. Dimitroulia, M. Charalampakis, P. Lynett, and C. Synolakis (2022). “The Mw= 6.6 earthquake and tsunami of south Crete on May 2, 2020”. In.
- Kawaguchi, K., Y. Kaneda, and E. Araki (2008). “The DONET: A real-time seafloor research infrastructure for the precise earthquake and tsunami monitoring”. In: *OCEANS 2008-MTS/IEEE Kobe Techno-Ocean*. IEEE, pp. 1–4.
- Lang, X., N. ur Rehman, Y. Zhang, L. Xie, and H. Su (2020). “Median ensemble empirical mode decomposition”. In: *Signal Processing* 176, p. 107686.
- Lee, J.-W., S.-C. Park, D. K. Lee, and J. H. Lee (2016). “Tsunami arrival time detection system applicable to discontinuous time series data with outliers”. In: *Natural Hazards and Earth System Sciences* 16.12, pp. 2603–2622.
- Levin, B. W. and M. A. Nosov (2016). *Physics of tsunamis*. Springer.
- Li, H. (2022). “Deep-Ocean Assessment and Reporting of Tsunamis (DART) BUOY”. In: *Encyclopedia of Ocean Engineering*. Ed. by W. Cui, S. Fu, and Z. Hu. Singapore: Springer Nature Singapore, pp. 306–314. ISBN: 978-981-10-6946-8. DOI: 10.1007/978-981-10-6946-8_79. URL: https://doi.org/10.1007/978-981-10-6946-8_79.
- Lin, L., Y. Wang, and H. Zhou (2009). “Iterative filtering as an alternative algorithm for empirical mode decomposition”. In: *Advances in Adaptive Data Analysis* 1.04, pp. 543–560.
- Liu, C., T. Lay, R. Wang, T. Taymaz, Z. Xie, X. Xiong, T. S. Irmak, M. Kahraman, and C. Erman (2023). “Complex multi-fault rupture and triggering

- during the 2023 earthquake doublet in southeastern Türkiye”. In: *Nature Communications* 14.1, p. 5564.
- Lomax, A. and A. Michelini (2013). “Tsunami early warning within five minutes”. In: *Pure and Applied Geophysics* 170, pp. 1385–1395.
- Lorito, S., J. Selva, A. Amato, A. Babeyko, B. Bayraktar, F. Bernardi, M. Charalampakis, L. Cordrie, N. Kalligeris, A. Piatanesi, et al. (2023). *The Tsunami Warning triggered in the Mediterranean Sea by the 2023 February 6 Mw 7.8 Türkiye-Syria earthquake: from the present Decision Matrix (DM) to Probabilistic Tsunami Forecasting (PTF)*. Tech. rep. Copernicus Meetings.
- Løvholt, F., S. Glimsdal, C. B. Harbitz, N. Horspool, H. Smebye, A. De Bono, and F. Nadim (2014). “Global tsunami hazard and exposure due to large coseismic slip”. In: *International journal of disaster risk reduction* 10, pp. 406–418.
- Maeda, T., K. Obara, M. Shinohara, T. Kanazawa, and K. Uehira (2015). “Successive estimation of a tsunami wavefield without earthquake source data: A data assimilation approach toward real-time tsunami forecasting”. In: *Geophysical Research Letters* 42.19, pp. 7923–7932.
- Marinaro, G., S. D’Amico, D. Embriaco, A. Giuntini, F. Simeone, O’Neill, B. J. Nicholson, N. Watkiss, and F. Restelli (2024). *A 21 km SMART Cable for earthquakes and tsunami detection operating in the Ionian Sea*. Tech. rep. Copernicus Meetings. DOI: <https://doi.org/10.5194/egusphere-egu24-16261>.
- Matthäus, W. (1972). “On the history of recording tide gauges”. In: *Proceedings of the Royal Society of Edinburgh, Section B: Biological Sciences* 73, pp. 26–34.
- Medvedeva, A., D. Vydrin, A. Krylov, A. Shiryborova, D. Smirnova, E. Tsukanova, U. Kânoğlu, and I. Medvedev (2023). “The Turkish Tsunami of 6 February 2023 in the Northeastern Mediterranean”. In.
- Mochizuki, M., K. Uehira, T. Kanazawa, T. Kunugi, K. Shiomi, S. Aoi, T. Matsumoto, N. Takahashi, N. Chikasada, T. Nakamura, et al. (2018). “S-Net project: performance of a large-scale seafloor observation network for preventing and reducing seismic and tsunami disasters”. In: *2018 OCEANS-MTS/IEEE Kobe Techno-Oceans (OTO)*. IEEE, pp. 1–4.
- Mofjeld, H. O. (1997). *Tsunami Detection Algorithm*. URL: https://nctr.pmel.noaa.gov/tda_documentation.html.
- Moore, C. (2024). *Personal communication*. Private Communication.
- Mousa, W. A. (2020). *Advanced digital signal processing of seismic data*. Cambridge University Press.

- Mulia, I. E., A. R. Gusman, and K. Satake (2017). “Optimal design for placements of tsunami observing systems to accurately characterize the inducing earthquake”. In: *Geophysical Research Letters* 44.24, pp. 12–106.
- Mulia, I. E., T. Hirobe, D. Inazu, T. Endoh, Y. Niwa, A. R. Gusman, H. Tatehata, T. Waseda, and T. Hibiya (2020). “Advanced tsunami detection and forecasting by radar on unconventional airborne observing platforms”. In: *Scientific reports* 10.1, p. 2412. DOI: 10.1038/s41598-020-59239-1.
- Mungov, G., M. Eblé, and R. Bouchard (2013). “DART® tsunameter retrospective and real-time data: A reflection on 10 years of processing in support of tsunami research and operations”. In: *Pure and Applied Geophysics* 170, pp. 1369–1384.
- Occhipinti, G. (2011). “Tsunami detection by ionospheric sounding: New tools for oceanic monitoring”. In: *Tsunami: A Growing Disaster* 19.
- Ozkula, G., R. K. Dowell, T. Baser, J.-L. Lin, O. A. Numanoglu, O. Ilhan, C. G. Olgun, C.-W. Huang, and T. D. Uludag (2023). “Field reconnaissance and observations from the February 6, 2023, Turkey earthquake sequence”. In: *Natural Hazards* 119.1, pp. 663–700.
- Papadopoulos, G. A., E. Lekkas, K.-N. Katsetsiadou, E. Rovythakis, and A. Yahav (2020). “Tsunami alert efficiency in the Eastern Mediterranean Sea: The 2 May 2020 earthquake (Mw 6.6) and near-field tsunami south of Crete (Greece)”. In: *GeoHazards* 1.1, pp. 44–60.
- Pawlowicz, R., B. Beardsley, and S. Lentz (2002). “Classical tidal harmonic analysis including error estimates in MATLAB using T_TIDE”. In: *Computers & Geosciences* 28.8, pp. 929–937.
- Pelinovsky, E., B. Choi, A. Stromkov, I. Didenkulova, and H.-S. Kim (2005). “Analysis of tide-gauge records of the 1883 Krakatau tsunami”. In: *Tsunamis: Case studies and recent developments*, pp. 57–77.
- Percival, D. B., D. W. Denbo, M. C. Eblé, E. Gica, H. O. Mofjeld, M. C. Spillane, L. Tang, and V. V. Titov (2011). “Extraction of tsunami source coefficients via inversion of DART buoy data”. In: *Natural hazards* 58.1, pp. 567–590.
- Petersen, G., P. Büyükakpınar, F. Vera, M. Metz, J. Saul, S. Cesca, T. Dahm, and F. Tilmann (2023). *Rupture processes of the 2023 Türkiye earthquake sequence: Main-and aftershocks*. Tech. rep. Copernicus Meetings.
- Pires, C. and P. M. Miranda (2001). “Tsunami waveform inversion by adjoint methods”. In: *Journal of Geophysical Research: Oceans* 106.C9, pp. 19773–19796.

- Rabinovich, A. B. (1997). “Spectral analysis of tsunami waves: Separation of source and topography effects”. In: *Journal of Geophysical Research: Oceans* 102.C6, pp. 12663–12676.
- Rabinovich, A. B. and M. C. Eblé (2015). “Deep-ocean measurements of tsunami waves”. In: *Pure and Applied Geophysics* 172, pp. 3281–3312.
- Rilling, G. and P. Flandrin (2007). “One or two frequencies? The empirical mode decomposition answers”. In: *IEEE transactions on signal processing* 56.1, pp. 85–95.
- Romano, F., A. Piatanesi, S. Lorito, C. Tolomei, S. Atzori, and S. Murphy (2016). “Optimal time alignment of tide-gauge tsunami waveforms in non-linear inversions: Application to the 2015 Illapel (Chile) earthquake”. In: *Geophysical Research Letters* 43.21, pp. 11–226.
- Saito, T. (2019). *Tsunami generation and propagation*. Springer.
- Satake, K. (1987). “Inversion of tsunami waveforms for the estimation of a fault heterogeneity: Method and numerical experiments”. In: *Journal of Physics of the Earth* 35.3, pp. 241–254.
- Selva, J., A. Amato, A. Armigliato, R. Basili, F. Bernardi, B. Brizuela, M. Cerminara, M. de’Micheli Vitturi, D. Di Bucci, P. Di Manna, et al. (2021a). “Tsunami risk management for crustal earthquakes and non-seismic sources in Italy”. In: *La Rivista del Nuovo Cimento* 44.2, pp. 69–144.
- Selva, J., S. Lorito, M. Volpe, F. Romano, R. Tonini, P. Perfetti, F. Bernardi, M. Taroni, A. Scala, A. Babeyko, et al. (2021b). “Probabilistic tsunami forecasting for early warning”. In: *Nature communications* 12.1, p. 5677.
- Stallone, A., A. Cicone, and M. Materassi (2020). “New insights and best practices for the successful use of Empirical Mode Decomposition, Iterative Filtering and derived algorithms”. In: *Scientific reports* 10.1, pp. 1–15.
- Street, J. O., R. J. Carroll, and D. Ruppert (1988). “A note on computing robust regression estimates via iteratively reweighted least squares”. In: *The American Statistician* 42.2, pp. 152–154.
- Synolakis, C. E. and E. N. Bernard (2006). “Tsunami science before and beyond Boxing Day 2004”. In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 364.1845, pp. 2231–2265.
- Tang, L., s. Titov, and C. Chamberlin (2009). “Development, testing, and applications of site-specific tsunami inundation models for real-time forecasting”. In: *Journal of Geophysical Research: Oceans* 114.C12. DOI: 10.1029/2009JC005476.

- Titov, V. V., F. I. González, H. O. Mofjeld, and J. C. Newman (2003). “Short-term inundation forecasting for tsunamis”. In: *Submarine Landslides and Tsunamis*, pp. 277–284. DOI: 10.1007/978-94-010-0205-9_29.
- Tolkova, E. (2010). “EOF analysis of a time series with application to tsunami detection”. In: *Dynamics of Atmospheres and Oceans* 50.1, pp. 35–54.
- (2009). “Principal component analysis of tsunami buoy record: Tide prediction and removal”. In: *Dynamics of atmospheres and oceans* 46.1-4, pp. 62–82.
- Tolkova, E. and W. Power (2011). “Obtaining natural oscillatory modes of bays and harbors via Empirical Orthogonal Function analysis of tsunami wave fields”. In: *Ocean Dynamics* 61.6, pp. 731–751.
- Torres, M. E., M. A. Colominas, G. Schlotthauer, and P. Flandrin (2011). “A complete ensemble empirical mode decomposition with adaptive noise”. In: *2011 IEEE international conference on acoustics, speech and signal processing (ICASSP)*. IEEE, pp. 4144–4147.
- Volpe, M., S. Lorito, J. Selva, R. Tonini, F. Romano, and B. Brizuela (2019). “From regional to local SPTHA: efficient computation of probabilistic tsunami inundation maps addressing near-field sources”. In: *Natural Hazards and Earth System Sciences* 19.3, pp. 455–469.
- Wang, Y., M. Heidarzadeh, K. Satake, I. E. Mulia, and M. Yamada (2020a). “A tsunami warning system based on offshore bottom pressure gauges and data assimilation for Crete Island in the Eastern Mediterranean Basin”. In: *Journal of Geophysical Research: Solid Earth* 125.10, e2020JB020293.
- Wang, Y., T. Maeda, K. Satake, M. Heidarzadeh, H. Su, A. Sheehan, and A. Gusman (2019a). “Tsunami data assimilation without a dense observation network”. In: *Geophysical Research Letters* 46.4, pp. 2045–2053. DOI: 10.1029/2018GL080930.
- Wang, Y. and K. Satake (2021). “Real-time tsunami data assimilation of S-net pressure gauge records during the 2016 Fukushima earthquake”. In: *Seismological Society of America* 92.4, pp. 2145–2155.
- Wang, Y., K. Satake, R. Cienfuegos, M. Quiroz, and P. Navarrete (2019b). “Far-field tsunami data assimilation for the 2015 Illapel earthquake”. In: *Geophysical Journal International* 219.1, pp. 514–521.
- Wang, Y., K. Satake, T. Maeda, and A. R. Gusman (2017). “Green’s function-based tsunami data assimilation: A fast data assimilation approach toward tsunami early warning”. In: *Geophysical Research Letters* 44.20, pp. 10–282. DOI: 10.1002/2017GL075307.

- Wang, Y., K. Satake, T. Maeda, M. Shinohara, and S. Sakai (2020b). “A method of real-time tsunami detection using ensemble empirical mode decomposition”. In: *Seismological Research Letters* 91.5, pp. 2851–2861.
- Wang, Y., H. Tsushima, K. Satake, and P. Navarrete (2021). “Review on recent progress in near-field tsunami forecasting using offshore tsunami measurements: Source inversion and data assimilation”. In: *Pure and Applied Geophysics* 178, pp. 5109–5128.
- Wu, Z. and N. E. Huang (2009). “Ensemble empirical mode decomposition: a noise-assisted data analysis method”. In: *Advances in adaptive data analysis* 1.01, pp. 1–41.